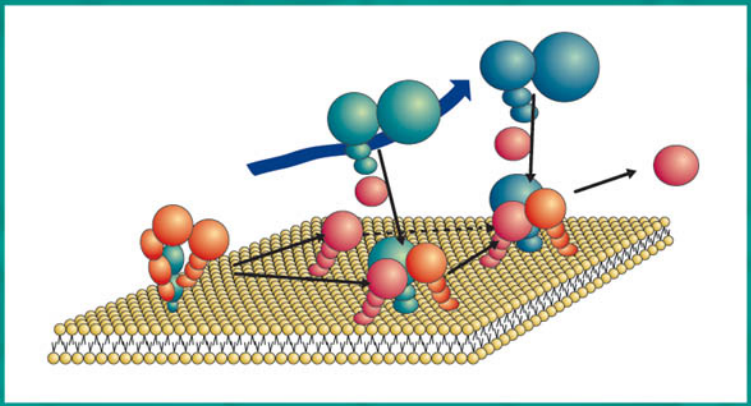
Molecular Hematology

SECOND EDITION

EDITED BY

D.Provan and J.Gribben



Blackwell Publishing

Molecular Hematology

Molecular Hematology

SECOND EDITION

Edited by

Drew Provan MD FRCP FRCPath
Senior Lecturer in Haematology
Department of Haematology
Barts and The Royal London School of Medicine and Dentistry
Whitechapel
London
UK

John G Gribben MD DSc FRCP FRCPath
Associate Professor
Division of Medical Oncology
Dana-Farber Cancer Institute
Department of Medicine
Brigham and Women's Hospital
Harvard Medical School
Boston
USA

Blackwell
Publishing

Blackwell Publishing, Inc., 350 Main Street, Malden, Massachusetts 02148-5020, USA
Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK
Blackwell Publishing Asia Pty Ltd, 550 Swanston Street, Carlton, Victoria 3053, Australia

First published 2000
Second edition 2005

Library of Congress Cataloging-in-Publication Data

Molecular haematology / edited by Drew Provan, John Gribben.
-- 2nd ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 1-4051-1255-7
1. Blood--Diseases--Molecular aspects. I. Provan, Andrew.
II. Gribben, John. III. Title: Molecular hematology,
[DNLM: 1. Hematologic Diseases. 2. Molecular Biology.
WH 120 M7187 2004]
RC636.M576 2004
616.1'5--dc22

2004000858

ISBN 1-4051-1255-7

A catalogue record for this title is available from the British Library

Set in 10/12 pt Minion by Sparks, Oxford – www.sparks.co.uk
Printed and bound in India by Gopsons Papers Ltd., NOIDA

Commissioning Editor: Maria Khan
Managing Editor: Rupal Malde
Production Editor: Fiona Pattison
Production Controller: Kate Charman

For further information on Blackwell Publishing, visit our website:
http://www.blackwellpublishing.com

Dedication

We would like to dedicate this book to our families, especially Val, Fraser and Peter, who provided constant encouragement and support throughout the project.

Contents

List of contributors

Nancy C Andrews MD PhD
Associate Investigator, Howard Hughes Medical Institute, Associate Professor, Harvard Medical School, and Associate in Medicine, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115, USA

Eyal C Attar MD
Instructor in Medicine, Harvard Medical School, Massachusetts General Hospital, Boston, MA 02129, USA

Luciano Baronciani PhD
Research Assistant, Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, IRCCS Maggiore Hospital, 20122 Milan, Italy

Anthony J Bench MA PhD
Senior Scientist, Department of Haematology, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK

P Leif Bergsagel MD FRCP(C)
Associate Professor of Medicine, Division of Hematology/Oncology, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA

Jaqueline Boultwood
University Research Lecturer and Co-Director of Leukaemia Research Fund Molecular Haematology Unit, Nuffield Department of Clinical Laboratory Sciences, John Radcliffe Hospital, Oxford OX3 9DU, UK

James B Bussel MD
Professor of Pediatrics, Division of Pediatrics, New York Presbyterian Hospital, Weill Cornell Medical Center, 525 E. 68th Street, Payson 609, New York, NY 10021, USA

Björn Dahlbäck MD PhD
Professor of Blood Coagulation Research, Department of Clinical Chemistry, University of Lund, University Hospital, Malmo S-20502, Sweden

Riccardo Dalla-Favera MD
Director, Institute for Cancer Genetics, Columbia University, 1150 St Nicholas Avenue, New York, NY 10032, USA

Silvana Debernardi PhD
Scientist, Cancer Research UK, Medical Oncology Unit, Barts and The London, Queen Mary's School of Medicine and Dentistry, Charterhouse Square, London EC1M 6BQ, UK

Katherine A Downes MD
Assistant Professor of Pathology, Department of Pathology, Case Western Reserve University School of Medicine/University Hospitals of Cleveland, 11000 Euclid Avenue Humphrey 5611, PTH 5077, Cleveland, OH 44106, USA, and Director of Coagulation, Associate Director of Blood Banking/Transfusion Medicine, University Hospitals of Cleveland, Cleveland, OH, USA

Brian J Druker MD
Investigator, Howard Hughes Medical Institute and JELD-WEN Chair of Leukemia Research, Department of Hematology/Medical Oncology, Oregon Health Sciences University Cancer Institute, L592, 3181 SW Sam Jackson Park Road, Portland, OR 97201-3011, USA

Paul L F Giangrande BSc MD FRCP FRCPath FRCPCH
Consultant Haematologist and Director, Oxford Haemophilia Centre and Thrombosis Unit, Churchill Hospital, Oxford OX3 7LJ, UK

D Gary Gilliland MD PhD
Professor of Medicine, Howard Hughes Medical Institute, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA

Anthony R Green PhD FRCP FRCPath
Wellcome Senior Fellow and Honorary Consultant Haematologist, Department of Haematology, University of Cambridge, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

John G Gribben MD DSc FRCP FRCPath
Associate Professor, Division of Medical Oncology, Dana-Farber Cancer Institute, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

Andreas Hillarp PhD
Associate Professor and Hospital Chemist, Department of Clinical Chemistry, University of Lund, University Hospital, Malmo S-20502, Sweden

Brian J P Huntley MB ChB MRCP DipRCPath
Clinical Research Fellow, Department of Haematology, University of Cambridge, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

Anastasios Karadimitris PhD MRCP MRCPath
Senior Lecturer and Honorary Consultant Haematologist, Leukaemia Research Fund Bennett Senior Fellow, Department of Haematology, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, UK

Ulf Klein PhD
Associate Research Scientist, Columbia University, Institute of Cancer Genetics, Russ Berrie Pavillion, 1150 St Nicholas Avenue, New York, NY 10032, USA

Anthony G Letai MD PhD
Assistant Professor in Medicine, Harvard Medical School, Smith 758, Dana-Farber Cancer Institute, 1 Jimmy Fund Way, Boston, MA 02115, USA

Debra M Lillington BSc
Head of Cytogenetics, ICRF Medical Oncology Laboratory, Barts and The London, Queen Mary's School of Medicine and Dentistry, Charterhouse Square, London EC1M 6BQ, UK

Lucio Luzzatto MD
Scientific Director, National Institute for Cancer Research, Istituto Scientifico Tumori, Largo, Rosanna Benzi, 10, 16132 Genova, Italy

Pier Mannuccio Mannucci MD
Professor and Chairman of Internal Medicine, Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, IRCCS Maggiore Hospital, 20122 Milan, Italy

Keith R McCrae MD
Associate Professor of Medicine Hematology-Oncology, BRB 3, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106-4937, USA

Jeffrey A Medin PhD
Associate Professor, Department of Medical Biophysics, University of Toronto, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Rm 7-411, Toronto, Ontario M5G 2M9, Canada

Graham Molineux PhD
Amgen Inc., Mailstop 15-2-A, One Amgen Center Drive, Thousand Oaks, CA 91320, USA

Ronald L Nagel MD
Irving D. Karpas Professor of Medicine, Head, Division of Hematology, Albert Einstein College of Medicine, The Bronx, New York, NY 10461, USA

Cristina Navarrete PhD
Head of Histocompatibility and Immunogenetics, National Blood Service, Colindale Avenue, London NW9 5BG, UK

Adrian C Newland MA FRCP FRCPath
Professor of Haematology, Department of Haematology, Barts and The London, Queen Mary's School of Medicine and Dentistry, Whitechapel, London E1 1BB, UK

Willem H Ouwehand MD PhD
Lecturer in Haematology, Division of Transfusion Medicine, Department of Haematology, University of Cambridge, Cambridge CB2 2PT, UK

John R Pawloski MD PhD
Assistant Professor of Medicine, Division of Hematology, Department of Medicine, Duke University, Durham, NC 27710, USA

Drew Provan MD FRCP FRCPath
Senior Lecturer in Haematology, Department of Haematology, Barts and The London, Queen Mary's School of Medicine and Dentistry, Whitechapel, London E1 1BB, UK

David T Scadden MD
Professor of Medicine, Harvard Medical School, AIDS Research Centre, Massachusetts General Hospital, 149, 13th Street, CNY 5212, Charleston, Boston, MA 02129-2020, USA

A Keith Stewart MD FRCP(C)
Associate Professor, Department of Medical Oncology/ Hematology, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada

Marilyn J Telen MD
Wellcome Professor of Hematology and Director, Duke-UNC Comprehensive Sickle Center, Box 2615 Duke University Medical Center, Durham, NC 27710, USA

George S Vassiliou MRCP(UK) DipRCPath BSc Hons (Lond)
Leukaemia Research Fund Clinical Research Fellow/Hon. Specialist Registrar, Department of Haematology, University of Cambridge, Cambridge Institute for Medical Research, Hills Road, Cambridge CB2 2XY, UK

James S Wainscoat FRCP FRCPath
Consultant Haematologist, Department of Haematology, John Radcliffe Hospital, Marston, Oxford OX3 9DU, UK

David Weatherall FRS
Regius Professor of Medicine, University of Oxford, Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK

Bryan D Young BSc PhD
Head and Professor of ICRF Medical Oncology Laboratory, Barts and The London, Queen Mary's School of Medicine and Dentistry, Charterhouse Square, London EC1M 6BQ, UK

Foreword

In 1968, after a quest lasting 30 years, X-ray analysis of crystalline horse hemoglobin at last reached the stage when I could build a model of its atomic structure. The amino acid sequences of human globin are largely homologous to those of horse globin, which made me confident that their structures are the same. By then, the amino acid substitutions responsible for many abnormal human hemoglobins had been determined. The world authority on them was the late Hermann Lehmann, Professor of Clinical Biochemistry at the University of Cambridge, who worked in the hospital just across the road from our Laboratory of Molecular Biology. I asked him to come over to see if there was any correlation between the symptoms caused by the different amino acids substituted in the abnormal hemoglobin and their positions in the atomic model. The day we spent going through them proved one of the most exciting in our scientific lives. We found hemoglobin to be insensitive to replacements of most amino acid residues on its surface, with the notable exception of sickle cell hemoglobin. On the other hand, we found the molecule to be extremely sensitive to even quite small alterations of internal non-polar contacts, especially those near the hemes. Replacements at the contact between the α and β subunits affected respiratory function.

In sickle cell hemoglobin an external glutamate was replaced by a valine. We wrote: '*A non-polar instead at a polar residue at a surface position would suffice to make each molecule adhere to a complementary site at a neighbouring one, that site being created by the conformational change from oxy to deoxy haemoglobin*'. This was soon proved to be correct. We published our findings under the title: 'The Molecular Pathology of Human Haemoglobin'. Our paper marked a turning point because it was the first time that the symptoms of diseases could be interpreted in terms of changes in the atomic structure of the affected protein. In the years that followed, the structure of the contact between the valine of one molecule of sickle cell hemoglobin and that of the complementary site of its neighbor became known in some detail. At a meeting at Arden House near Washington in 1980, several colleagues and I decided to use this knowledge for the design of anti-sickling drugs, but after an effort lasting 10 years, we realized that we were running up against a brick wall. Luckily, the work was not entirely wasted, because we found a series of compounds that lower the oxygen affinity of hemoglobin and we realized that this might be clinically useful. One of those compounds, designed by DJ Abraham at the University of Virginia in Richmond, is now entering phase 3 clinical trials. On the other hand, our failure to find a drug against sickle cell anemia, even when its cause was known in atomic detail, made me realize the extreme difficulty of finding drugs to correct a malfunction of a protein that is caused by a single amino acid substitution. Most thalassemias are due not to amino acid substitutions, but to either complete or partial failure to synthesize α- or β-globin chains. Weatherall's chapter shows that, at the genetic level, there may be literally hundreds of different genetic lesions responsible for that failure. Correction of such lesions is now the subject of intensive work in many laboratories.

Early in the next century, the human genome will be complete. It will reveal the amino acid sequences of all the 100000 or so different proteins of which we are made. Many of these proteins are still unknown. To discover their functions, the next project now under discussion is a billion dollar effort to determine the structures of all the thousands of unknown proteins within 10 years. By then we shall know the identity of the proteins responsible for most of the several thousand different genetic diseases. Will this lead to effective treatment or will medical geneticists be in the same position as doctors were early in this century when the famous physician Sir William Osler confined their task to the establishment of diagnoses? Shall we know the cause of every genetic disease without a cure?

Our only hope lies in somatic gene therapy. AK Stewart's chapter on Molecular Therapeutics describes the many ingenious methods now under development. So far, none of these has produced lasting effects, apparently because the transferred genes are not integrated into the mammalian genome, but a large literature already grown up bears testimony to the great efforts now underway to overcome this problem.

My much-loved teacher William Lawrence Bragg used to say '*If you go on hammering away at a problem, eventually it seems to get tired, lies down and lets you catch it*.' Let us hope that somatic gene therapy will soon get tired.

M.F. Perutz
Cambridge

Perutz MF, Lehmann H. (1968) Molecular pathology of human haemoglobin. *Nature*, **219**, 902–909.

Perutz MF, Muirhead H, Cox JM, Goaman LCG. (1968) Three-dimensional Fourier synthesis of horse oxyhaemoglobin at 2.8Å resolution: the atomic model. *Nature*, **219**, 131–139.

Preface to second edition

Hematology has seen many major developments since this book was first published, as molecular techniques become more powerful and the genetics of blood disorders are unraveled. The Human Genome Project has now been completed and a huge amount of new genetic information has become available.

Of course, most funding and resources continue to be spent on understanding the molecular basis of malignant disease and hence there has been a huge increase in our knowledge base for leukemias, lymphomas and other malignancies. Molecular biology has also begun to bring with it advances in treatment, with molecules devised to reduce the tumor burden through much more subtle and selective mechanisms than have been possible with conventional chemotherapy agents. Chronic myeloid leukemia (CML), one of the best studied human malignancies, is amenable to treatment with the molecule STI571, revolutionizing our therapy of this disease. Dr Brian Druker's seminal work on this molecule is explained in detail in his chapter devoted entirely to the biology and management of CML. We have updated all the chapters on malignant hematology and have included some new chapters, such as 'Stem cells', 'Secondary myelodysplasia/acute myelogenous leukemia—assessment of risk' and 'Gene expression profiling in the study of lymphoid malignancies'.

Non-malignant disease has also enjoyed the benefits of this new genetic information and the original chapters have been updated to reflect this new knowledge. New chapters dealing with the molecular basis of blood group antigens, the molecular basis of von Willebrand disease, and platelet disorders have been added by leading clinicians and researchers in these fields.

However, despite the growing complexity of the pathogenesis, diagnosis and management of patients with blood diseases, the ethos of the book remains the same: to provide a succinct account of the molecular biology of hematological disease written at a level at which it should be of benefit to the seasoned molecular biologist and the practicing clinician alike. We have retained the original structure for the chapters, with high-quality artwork and 'Further reading' sections, in order to make the book visually appealing and relevant to modern hematology practice.

As before, we welcome any comments or suggestions from readers, which we will attempt to incorporate into the next edition.

Drew Provan (a.b.provan@qmul.ac.uk)
John Gribben (John_Gribben@dfci.harvard.edu)

Suggested general reading

Anderson KC, Ness PM (eds). (2000) *Scientific Basis of Transfusion Medicine: Implications for Clinical Practice*, 2nd edn. Philadelphia: W.B. Saunders.

Beutler E, Lichtman MA (eds). (2001) *Williams Hematology* 6th edn. New York: McGraw-Hill.

Cooper GM. (1997) *The Cell: A Molecular Approach*. Washington, DC: ASM Press.

Cox TM, Sinclair J. (1997) *Molecular Biology in Medicine*. Oxford: Blackwell Science.

Jameson JL (ed.). (1998) *Principles of Molecular Medicine*. New York: Humana Press.

Mullis KB. (1990) The unusual origin of the polymerase chain reaction. *Scientific American*, **262**, 56–65.

Roitt I. (2001) *Roitt's Essential Immunology*, 10th edn. Oxford: Blackwell Science.

Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H (eds). (2000) *The Molecular Basis of Blood Diseases*, 2nd edn. Philadelphia: W.B. Saunders.

Watson JD, Gilman M, Witkowski J, Zoller M (eds). (1992) *Recombinant DNA*, 2nd edn. New York: Scientific American Books.

Acknowledgments

We would like to express thanks to Maria Khan (Commissioning Editor) and Rupal Malde (Managing Editor) for helping keep the project on track and for their continuing support throughout the project. We are very grateful to Peter Varney for his administrative assistance during this edition. As ever, David Gardner provided very high quality artwork, for which we are very grateful.

Abbreviations

AAV	adeno-associated virus
ADA	adenosine deaminase
AE1	anion exchanger protein 1
ALAS	δ-aminolaevulinate synthase
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia (Chapters 2, 4) acute myelogenous leukemia (Chapters 5, 6, 8)
APC	activated protein C; antigen-presenting cell
APL	acute promyelocytic leukemia
ASCT	autologous stem cell transplantation
AT	antithrombin
ATRA	all-*trans*-retinoic acid
B-CLL	B-cell chronic lymphocytic leukemia
BCR	breakpoint cluster region
BM	bone marrow
BMF	bone marrow failure
BPG	2,3-biphosphoglycerate
BSS	Bernard–Soulier syndrome
CBF	core binding factor
Cbl	cobalamin
CDA	congenital dyserythropoietic anemia
CDAE1	N-terminal cytoplasmic domain of anion exchanger protein 1
CDKI	cyclin-dependent kinase inhibitor
CDR	complementarity-determining region
CFC	colony-forming cell
CFU	colony-forming unit
CFU-S	colony-forming units–spleen
CGH	comparative genomic hybridization
CH	heavy-chain constant region
CHR	complete hematologic response
CLL	chronic lymphocytic leukemia
CLM	common lymphoid progenitor
cM	centimorgan
CML	chronic myeloid leukemia
CNS	central nervous system
CSF	colony-stimulating factor
CTLA	cytotoxic T-lymphocyte antigen
DC	dyskeratosis congenita
DLBL	diffuse large B-cell lymphoma
dm	double minute (chromosome)
EGF	epidermal growth factor
EPO	erythropoietin
ERT	enzyme replacement therapy
EST	expressed sequence tag
ET	essential thrombocythemia
FA	Fanconi's anemia
FAB	French–American–British (classification of myelodysplastic syndromes)
FACS	fluorescence-activated cell sorter
FcR	Fc receptor
FG	phenylalanine-glycine
FISH	fluorescence *in situ* hybridization
G6PD	glucose-6-phosphate dehydrogenase
GAP	glycine-alanine-proline
G-CSFR	granulocyte colony-stimulating factor receptor
GDP	guanosine diphosphate
Ge	Gerbich erythrocyte antigen
GEP	gene expression profiling
GM-CSF	granulocyte macrophage colony-stimulating factor
GP	glycoprotein
GPC, D	glycophorin C, glycophorin D
GPI	glucosylphosphatidylinositol
GT	Glanzmann thrombasthenia
GTP	guanosine triphosphate
GVHD	graft-versus-host disease
HDN	hemolytic disease of the newborn
HGF	hematopoietic growth factor
HHV-8	human herpesvirus
HLA	human leukocyte antigen
HMCL	human multiple myeloma cell line
HPA	human platelet antigen
HPFF	hereditary persistence of fetal hemoglobin 1
HPP-CFC	high proliferative potential colony-forming cell
HSC	hemopoietic stem cell
HSCT	hematopoietic stem cell transplantation
hsr	homogeneously staining region
IAA	idiopathic aplastic anemia
IDDm	insulin-dependent diabetes mellitus

IFN-γ	interferon γ
Ig	immunoglobulin
IL	interleukin
IMF	idiopathic myelofibrosis
IPI	International Prognostic Index
IPPS	International Prognostic Scoring System
ISC	irreversibly sickled cell
ITD	internal tandem duplications
ITP	idiopathic thrombocytopenic purpura
IVIg	intravenous immunoglobulin
IVS	intervening sequence
kb	kilobase pair (1000 base pairs)
KS	Kaposi's sarcoma
KTLS	c-kit^{pos}Thy-1.1^{low}Lin^{neg}Sca-1^{pos} cellular phenotype
LCS	locus control region
LOH	loss of heterozygosity
LTC-IC	long-term culture-initiating cell
LTR	long terminal repeat
mAb	monoclonal antibody
M-BCR	major breakpoint cluster region
m-BCR	minor breakpoint cluster region
MBR	major breakpoint region
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
MCL	mantle cell lymphoma
M-CSF	macrophage colony-stimulating receptor
MCV	mean cell volume
MDS	myelodysplastic syndrome
MGDF	megakaryocyte growth and development factor
MGUS	monoclonal gammopathy of undetermined significance
MHC	major histocompatibility complex
MM	multiple myeloma
MPD	myeloproliferative disorder
MRD	minimal residual disease
mtDNA	mitochondrial DNA
NAITP	neonatal/fetal alloimmune thrombocytopenia
NHL	non-Hodgkin's lymphoma
NO	nitric oxide
NOs	nitric oxide synthase
ORF	open reading frame
PAR	protease-activated receptor
PBMC	peripheral blood mononuclear cell
PBPC	peripheral blood progenitor cell
PCR	polymerase chain reaction
PEG-MGDF	pegylated megakaryocyte growth and development factor
PETS	paraffin-embedded tissue section
PMPS	Pearson's marrow–pancreas syndrome
PNH	paroxysmal nocturnal hemoglobinuria
PUBS	peri-umbilical blood sampling
PV	polycythemia vera
RBC	red blood cell
RSCA	reference strand conformational analysis
RT-PCR	reverse transcriptase–polymerase chain reaction
SA	sideroblastic anemia
SCF	stem cell factor
SDS-PAGE	sodium dodecylsulfate–polyacrylamide gel electrophoresis
SKY	spectral karyotyping
SLE	systemic lupus erythematosus
SNO-Hb	S-nitrosohemoglobin
SNP	single-nucleotide polymorphism
TBI	total body irradiation
TCR	T-cell receptor
TF	tissue factor
TGF-β	transforming growth factor-β
THF	tetrahydrofolate
V, D, J, C	variable, diversity, joining and constant regions
VCAM-1	vascular cell adhesion molecule-1
VH	heavy-chain variable region
vWD	von Willebrand disease
vWF	von Willebrand factor
vWF:RCo	von Willebrand factor ristocetin cofactor activity
WBC	white blood cell

Chapter 1 Beginnings: the molecular pathology of hemoglobin

David Weatherall

Historical background

Linus Pauling first used the term 'molecular disease' in 1949, after the discovery that the structure of sickle cell hemoglobin differed from that of normal hemoglobin. Indeed, it was this seminal observation that led to the concept of *molecular medicine*; that is, the description of disease mechanisms at the level of cells and molecules. However, until the development of recombinant DNA technology in the mid-1970s, knowledge of events inside the cell nucleus, notably how genes function, could only be the subject of guesswork based on the structure and function of their protein products. However, as soon as it became possible to isolate human genes and to study their properties, the picture changed dramatically.

Progress over the last 20 years has been driven by technological advances in molecular biology. At first it was possible only to obtain indirect information about the structure and function of genes by DNA/DNA and DNA/RNA hybridization; that is, by probing the quantity or structure of RNA or DNA by annealing reactions with molecular probes. The next major advance was the ability to fractionate DNA into pieces of predictable size with bacterial restriction enzymes. This led to the invention of a technique that played a central role in the early development of human molecular genetics, called 'Southern blotting' after the name of its developer, Edwin Southern. This method allowed the structure and organization of genes to be studied directly for the first time and led to the definition of a number of different forms of molecular pathology.

Once it was possible to fractionate DNA, it soon became feasible to insert the pieces into vectors that are able to divide within bacteria. The steady improvement in the properties of cloning vectors made it possible to generate libraries of human DNA growing in bacterial cultures. Ingenious approaches were developed to scan the libraries to detect genes of interest; once pinpointed, the appropriate bacterial colonies could be grown to generate larger quantities of DNA carrying a particular gene. Later it became possible to sequence these genes, persuade them to synthesize their products in microorganisms, cultured cells or even other species, and hence to define their key regulatory regions.

The early work in the field of human molecular genetics focussed on diseases in which there was some knowledge of the genetic defect at the protein or biochemical level. However, once linkage maps of the human genome became available, following the identification of highly polymorphic regions of DNA, it was possible to search for any gene for a disease, even where the cause was completely unknown. This approach, first called 'reverse genetics' and later rechristened 'positional cloning', led to the discovery of genes for many important diseases.

As even more DNA markers became available and as methods for sequencing were improved and automated, thoughts turned to the next major goal in this field, which was to determine the complete sequence of the bases that constitute our 30 000 or so genes and all that lies between them: the Human Genome Project. This remarkable endeavor was partially completed recently and should be finished within the next few years. The further understanding of the functions and regulation of our genes will require multidisciplinary research encompassing many different fields. The next stage in the Human Genome Project, called 'genome annotation', entails analyzing the raw DNA sequence in order to determine its biological significance. One of the main ventures in the era of functional genomics will be in what is termed 'proteomics', the large-scale analysis of the protein products of genes. The ultimate goal will be to try to define the protein complement, or proteome, of cells and how the many different proteins interact with one another. To this end, large-scale facilities are being established for isolating and purifying the protein

products of genes that have been expressed in bacteria. Their structure can then be studied by a variety of different techniques, notably X-ray crystallography and nuclear magnetic resonance spectroscopy. The crystallographic analysis of proteins is being greatly facilitated by the use of X-ray beams from a synchrotron radiation source.

During this remarkable period of technical advance, considerable progress has been made towards an understanding of the pathology of disease at the molecular level. This has had a particular impact on hematology, leading to advances in the understanding of gene function and disease mechanisms in almost every aspect of the field.

The inherited disorders of hemoglobin, the thalassemias and structural hemoglobin variants, the commonest human monogenic diseases, were the first to be studied systematically at the molecular level and a great deal is known about their genotype–phenotype relationships. This field led the way to molecular hematology and, indeed, to the development of molecular medicine. Thus, even though the genetics of hemoglobin is complicated by the fact that different varieties are produced at particular stages of human development, the molecular pathology of the hemoglobinopathies provides an excellent model system for understanding any monogenic disease and the complex interactions between genotype and environment that underlie many multigenic disorders.

In this chapter we will consider the structure, synthesis and genetic control of the human hemoglobins, describe the molecular pathology of the hemoglobin disorders individually, and discuss briefly how the complex interactions of their different genotypes produce a remarkably diverse family of clinical phenotypes. Readers who wish to learn more about the methods of molecular genetics, particularly as applied to the study of hemoglobin disorders, are referred to the reviews cited at the end of this chapter.

The structure, genetic control and synthesis of normal hemoglobin

Structure and function

The varying oxygen requirements during embryonic, fetal and adult life are reflected in the synthesis of different structural hemoglobins at each stage of human development. They all have the same general tetrameric structure, however, consisting of two different pairs of globin chains, each attached to one heme molecule. Adult and fetal hemoglobins have α chains combined with β chains (Hb A, $\alpha_2\beta_2$), δ chains (Hb A_2, $\alpha_2\delta_2$) and γ chains (Hb F, $\alpha_2\gamma_2$). In embryos, α-like chains called ζ chains combine with γ chains to produce Hb Portland ($\zeta_2\gamma_2$), or with ε chains to make Hb Gower 1 ($\zeta_2\varepsilon_2$), while α and ε chains form Hb Gower 2 ($\alpha_2\varepsilon_2$). Fetal hemoglobin is heterogeneous; there are two varieties of γ chain that differ only in their amino acid composition at position 136, which may be occupied by either glycine or alanine; γ chains containing glycine at this position are called G_γ chains, those with alanine, A_γ chains (Figure 1.1).

The synthesis of hemoglobin tetramers consisting of two unlike pairs of globin chains is absolutely essential for the effective function of hemoglobin as an oxygen carrier. The classical sigmoid shape of the oxygen dissociation curve, which reflects the allosteric properties of the hemoglobin molecule, ensures that, at high oxygen tensions in the lungs, oxygen is readily taken up and later released effectively at the lower tensions encountered in the tissues. The shape of the curve is quite different to that of myoglobin, a molecule which consists of a single globin chain with heme attached to it, which, like abnormal hemoglobins that consist of homotetramers of like-chains, has a hyperbolic oxygen dissociation curve.

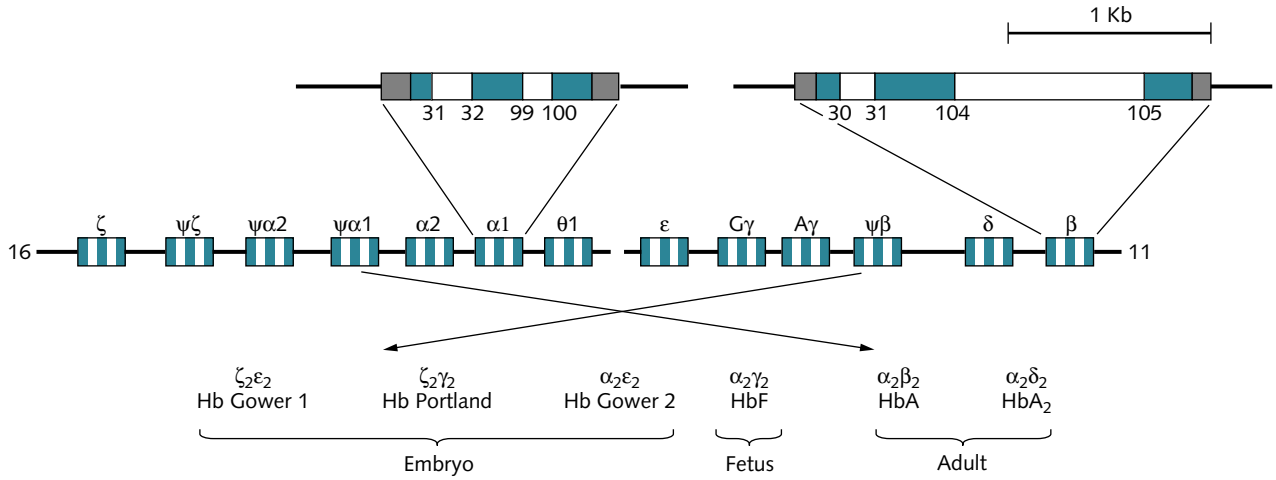


Fig. 1.1 The genetic control of human hemoglobin production in embryonic, fetal and adult life

The transition from a hyperbolic to a sigmoid oxygen dissociation curve, which is absolutely critical for normal oxygen delivery, reflects cooperativity between the four heme molecules and their globin subunits. When one of them takes on oxygen, the affinity of the remaining three increases markedly; this happens because hemoglobin can exist in two configurations, deoxy(T) and oxy(R), where T and R represent the tight and relaxed states, respectively. The T configuration has a lower affinity than the R for ligands such as oxygen. At some point during the addition of oxygen to the hemes, the transition from the T to the R configuration occurs and the oxygen affinity of the partially liganded molecule increases dramatically. These allosteric changes result from interactions between the iron of the heme groups and various bonds within the hemoglobin tetramer, which lead to subtle spatial changes as oxygen is taken on or given up.

The precise tetrameric structures of the different human hemoglobins, which reflect the primary amino acid sequences of their individual globin chains, are also vital for the various adaptive changes that are required to ensure adequate tissue oxygenation. The position of the oxygen dissociation curve can be modified in several ways. For example, oxygen affinity decreases with increasing CO_2 tension (the Bohr effect). This facilitates oxygen loading to the tissues, where a drop in pH due to CO_2 influx lowers oxygen affinity; the opposite effect occurs in the lungs. Oxygen affinity is also modified by the level of 2,3-biphosphoglycerate (2,3-BPG) in the red cell. Increasing concentrations shift the oxygen dissociation curve to the right, that is, they reduce oxygen affinity, while diminishing concentrations have the opposite effect. 2,3-BPG fits into the gap between the two β chains when it widens during deoxygenation, and interacts with several specific binding sites in the central cavity of the molecule. In the deoxy configuration the gap between the two β chains narrows and the molecule cannot be accommodated. With increasing concentrations of 2,3-BPG, which are found in various hypoxic and anemic states, more hemoglobin molecules tend to be held in the deoxy configuration and the oxygen dissociation curve is therefore shifted to the right, with more effective release of oxygen.

Fetal red cells have greater oxygen affinity than adult red cells, although, interestingly, purified fetal hemoglobin has an oxygen dissociation curve similar to that of adult hemoglobin. These differences, which are adapted to the oxygen requirements of fetal life, reflect the relative inability of Hb F to interact with 2,3-BPG compared with Hb A. This is because the γ chains of Hb F lack specific binding sites for 2,3-BPG.

In short, oxygen transport can be modified by a variety of adaptive features in the red cell that include interactions between the different heme molecules, the effects of CO_2 and differential affinities for 2,3-BPG. These changes, together with more general mechanisms involving the cardiorespiratory system, provide the main basis for physiological adaptation to anemia.

Genetic control of hemoglobin

The α- and β-like globin chains are the products of two different gene families which are found on different chromosomes (Figure 1.1). The β-like globin genes form a linked cluster on chromosome 11, spread over approximately 60 kb (kb = kilobase or 1000 nucleotide bases). The different genes that form this cluster are arranged in the order 5′–ε–^{G}γ–^{A}γ–ψβ–δ–β–3′. The α-like genes also form a linked cluster, in this case on chromosome 16, in the order 5′–ζ–ψζ–ψα1–α2–α1–3′. The ψβ, ψζ and ψα genes are pseudogenes; that is, they have strong sequence homology with the β, ζ and α genes but contain a number of differences that prevent them from directing the synthesis of any products. They may reflect remnants of genes that were functional at an earlier stage of human evolution.

The structure of the human globin genes is, in essence, similar to that of all mammalian genes. They consist of long strings of nucleotides that are divided into coding regions, or exons, and non-coding inserts called 'intervening sequences' (IVS), or introns. The β-like globin genes contain two introns, one of 122–130 between codons 30 and 31 and one of 850–900 base pairs between codons 104 and 105 (the exon codons are numbered sequentially from the 5′ to the 3′ end of the gene; that is, from left to right). Similar, though smaller, introns are found in the α and ζ globin genes. These introns and exons, together with short non-coding sequences at the 5′ and 3′ ends of the genes, represent the major functional regions of the particular genes. However, there are also extremely important regulatory sequences that subserve these functions, which lie outside the genes themselves.

At the 5′ non-coding (flanking) regions of the globin genes, as in all mammalian genes, there are blocks of nucleotide homology. The first, the ATA box, is about 30 bases upstream (to the left) of the initiation codon; that is, the start word for the beginning of protein synthesis (*see below*). The second, the CCAAT box, is about 70 base pairs upstream from the 5′ end of the genes. About 80–100 bases further upstream there is the sequence GGGGTG, or CACCC, which may be inverted or duplicated. These three highly conserved DNA sequences, called 'promoter elements', are involved in the initiation of transcription of the individual genes. Finally, in the 3′ non-coding region of all the globin genes there is the sequence AATAAA, which is the signal for cleavage and polyA addition to RNA transcripts (*see below: Gene action and globin synthesis*).

The globin gene clusters also contain several sequences that constitute regulatory elements, which interact to promote erythroid-specific gene expression and coordination of the changes in globin gene activity during development. These

include the globin genes themselves and their promoter elements—enhancers (regulatory sequences that increase gene expression despite being located at a considerable distance from the genes) and 'master' regulatory sequences called, in the case of the β globin gene cluster, the 'locus control region' (LCR); and, in the case of the α genes, HS40 (a nuclease-hypersensitive site in DNA 40 kb from the α globin genes). Each of these sequences has a modular structure made up of an array of short motifs that represent the binding sites for transcriptional activators or repressors.

Gene action and globin synthesis

The flow of information between DNA and protein is summarized in Figure 1.2. When a globin gene is transcribed, messenger RNA (mRNA) is synthesized from one of its strands, a process which begins with the formation of a transcription complex consisting of a variety of regulatory proteins together with an enzyme called RNA polymerase (*see below*). The primary transcript is a large mRNA precursor which contains both intron and exon sequences. While in the nucleus, this molecule undergoes a variety of modifications. First, the introns are removed and the exons are spliced together. The intron/exon junctions always have the same sequence: GT at their 5′ end, and AG at their 3′ end. This appears to be essential for accurate splicing; if there is a mutation at these sites this process does not occur. Splicing reflects a complex series of intermediary stages and the interaction of a number of different nuclear proteins. After the exons are joined, the mRNAs are modified and stabilized; at their 5′ end a complex CAP structure is formed, while at their 3′ end a string of adenylic acid residues (polyA) is added. The mRNA processed in this way moves into the cytoplasm, where it acts as a template for globin chain production. Because of the rules of base pairing—that is, cytosine always pairs with thymine, and guanine with adenine—the structure of the mRNA reflects a faithful copy of the DNA codons from which it is synthesized; the only difference is that, in RNA, uracil (U) replaces thymine (T).

Amino acids are transported to the mRNA template on carriers called transfer RNAs (tRNAs); there are specific tRNAs for each amino acid. Furthermore, because the genetic code is redundant (that is, more than one codon can encode a particular amino acid), for some of the amino acids there are several different individual tRNAs. Their order in the globin chain is determined by the order of codons in the mRNA. The tRNAs contain three bases, which together constitute an anticodon; these anticodons are complementary to mRNA codons for particular amino acids. They carry amino acids to the template, where they find the appropriate positioning by codon–anticodon base-pairing. When the first tRNA is in po-

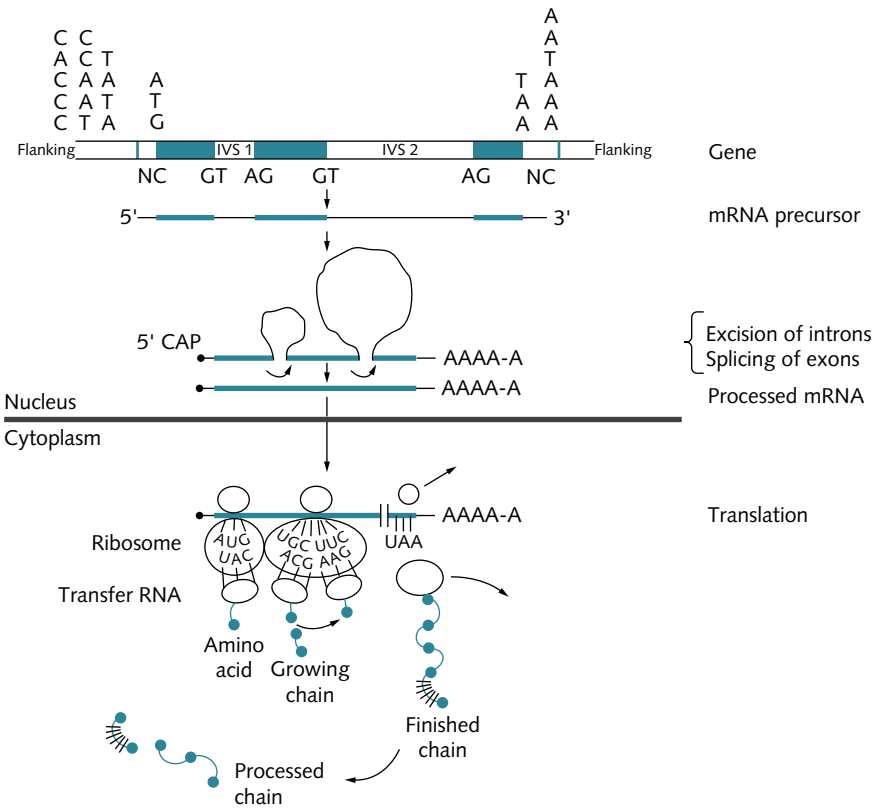


Fig. 1.2 The mechanisms of globin gene transcription and translation

sition, an initiation complex is formed between several protein initiation factors together with the two subunits which constitute the ribosomes. A second tRNA moves in alongside and the two amino acids that they are carrying form a peptide bond between them; the globin chain is now two amino acid residues long. This process is continued along the mRNA from left to right, and the growing peptide chain is transferred from one incoming tRNA to the next; that is, the mRNA is translated from 5′ to 3′. During this time the tRNAs are held in appropriate steric configuration with the mRNA by the two ribosomal subunits. There are specific initiation (AUG) and termination (UAA, UAG and UGA) codons. When the ribosomes reach the termination codon, translation ceases, the completed globin chains are released, and the ribosomal subunits are recycled. Individual globin chains combine with heme, which has been synthesized through a separate pathway, and then interact with one like chain and two unlike chains to form a complete hemoglobin tetramer.

Regulation of hemoglobin synthesis

The regulation of globin gene expression is mediated mainly at the transcriptional level, with some fine tuning during translation and post-translational modification of the gene products. DNA that is not involved in transcription is held tightly packaged in a compact, chemically modified form that is inaccessible to transcription factors and polymerases and which is heavily methylated. Activation of a particular gene is reflected by changes in the structure of the surrounding chromatin, which can be identified by enhanced sensitivity to nucleases. Erythroid lineage-specific nuclease-hypersensitive sites are found at several locations in the β globin gene cluster. Four are distributed over 20 kb upstream from the ε globin gene in the region of the β globin LCR (Figure 1.3). This vital regulatory region is able to establish a transcriptionally active domain spanning the entire β globin gene cluster. Several enhancer sequences have been identified in this cluster. A variety of regulatory proteins bind to the LCR, and to the promoter regions of the globin genes and to the enhancer sequences. It is thought that the LCR and other enhancer regions become opposed to the promoters to increase the rate of transcription of the genes to which they are related.

These regulatory regions contain sequence motifs for various ubiquitous and erythroid-restricted transcription factors. Binding sites for these factors have been identified in each of the globin gene promoters and at the hypersensitive-site regions of the various regulatory elements. A number of the factors which bind to these areas are found in all cell types. They include Sp1, Yy1 and Usf. In contrast, a number of transcription factors have been identified, including GATA-1, EKLF and NF-E2, which are restricted in their distribution to erythroid cells and, in some cases, megakaryocytes and mast cells. The overlapping of erythroid-specific and ubiquitous-factor binding sites in several cases suggests that competitive binding may

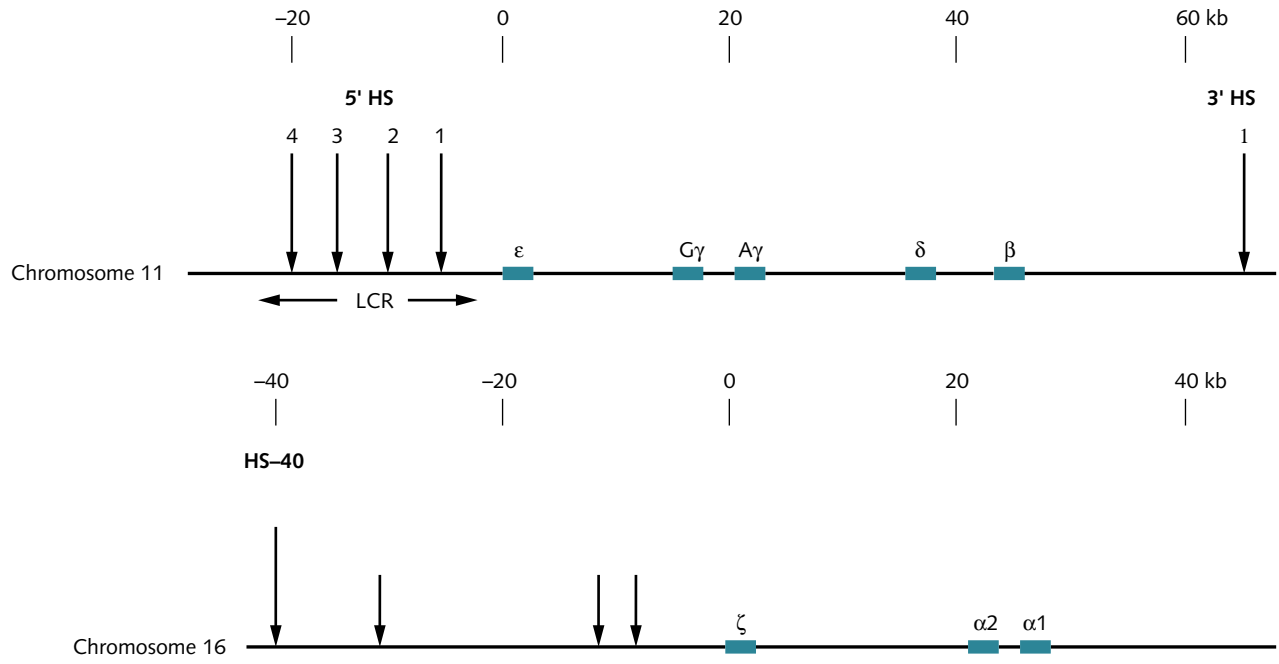


Fig. 1.3 The positions of the major regulatory regions in the β and α globin gene clusters
The arrows indicate the position of the erythroid lineage-specific nuclease-hypersensitive sites. HS = hypersensitive.

play an important part in the regulation of erythroid-specific genes. Another binding factor, SSP, the stage selector protein, appears to interact specifically with ε and γ genes.

The binding of hematopoietic-specific factors activates the LCR, which renders the entire β globin gene cluster transcriptionally active. These factors also bind to the enhancer and promoter sequences, which work in tandem to regulate the expression of the individual genes in the clusters. It is likely that some of the transcriptional factors are developmental-stage-specific, and hence may be responsible for the differential expression of the embryonic, fetal and adult globin genes. The α globin gene cluster also contains an element, HS40, which has some structural features in common with the β LCR, although it is different in aspects of its structure. A number of enhancer-like sequences have also been identified, although it is becoming clear that there are fundamental differences in the pattern of regulation of the two globin gene clusters.

In addition to the different regulatory sequences outlined above, there are also sequences which may be involved specifically with 'silencing' of genes, notably those for the embryonic hemoglobins, during development.

Some degree of regulation is also mediated by differences in the rates of initiation and translation of the different mRNAs, and at the post-transcriptional level by differential affinity for different protein subunits. However, this kind of post-transcriptional fine tuning probably plays a relatively small role in determining the overall output of the globin gene products.

Regulation of developmental changes in globin gene expression

During development, the site of red cell production moves from the yolk sac to the fetal liver and spleen, and thence to bone marrow in the adult. Embryonic, fetal and adult hemoglobin synthesis is approximately related in time to these changes in the site of erythropoiesis, although it is quite clear that the various switches, between embryonic and fetal and between fetal and adult hemoglobin synthesis, are beautifully synchronized throughout these different sites. Fetal hemoglobin synthesis declines during the later months of gestation and Hb F is replaced by Hbs A and A_2 by the end of the first year of life.

Despite a great deal of research, very little is known about the regulation of these different switches from one globin gene to another during development. Work from a variety of different sources suggests that there may be specific regions in the α and β globin gene clusters that are responsive to the action of transcription factors, some of which may be developmental-stage-specific. However, proteins of this type have not yet been isolated, and nothing is known about their regulation and how it is mediated during development.

The molecular pathology of hemoglobin

As is the case for most monogenic diseases, the inherited disorders of hemoglobin fall into two major classes. First, there are those that result from a reduced output of one or other globin genes, the *thalassemias*. Second, there is a wide range of conditions that result from the production of *structurally abnormal globin chains*; the type of disease depends on how the particular alteration in protein structure interferes with its stability or function. Of course, no biological classification is entirely satisfactory; those which attempt to define the hemoglobin disorders are no exception. There are some structural hemoglobin variants which happen to be synthesized at a reduced rate and hence are associated with a clinical picture similar to thalassemia. And there are other classes of mutations which simply interfere with the normal transition from fetal to adult hemoglobin synthesis, a family of conditions that is given the general title 'hereditary persistence of fetal hemoglobin'. Furthermore, because these diseases are all so common and occur together in particular populations, it is not uncommon for an individual to inherit a gene for one or other form of thalassemia and a structural hemoglobin variant. The rather heterogeneous group of conditions that results from all these different mutations and interactions is summarized in Table 1.1.

Table 1.1 The thalassemias and related disorders.

α **Thalassemia**
α^0
α^+
Deletion (–α)
Non-deletion (α^T)
β **Thalassemia**
β^0
β^+
Normal Hb A_2
'Silent'
δβ **Thalassemia**
$(\delta\beta)^+$
$(\delta\beta)^0$
$(^A\gamma\delta\beta)^0$
γ **Thalassemia**
δ **Thalassemia**
εγδβ **Thalassemia**
Hereditary persistence of fetal hemoglobin
Deletion
$(\delta\beta)^0$
Non-deletion
Linked to β globin genes
$^G\gamma\beta^+$
$^A\gamma\beta^+$
Unlinked to β globin genes

Over recent years, the determination of the molecular pathology of the two common forms of thalassemia, α and β, has provided a remarkable picture of the repertoire of mutations that can underlie human monogenic disease. Similarly, studies of the relationship between structure and function in the structurally abnormal hemoglobins have provided a great deal of information about normal human hemoglobin function.

In the sections that follow we will describe, in outline, the different forms of molecular pathology that underlie these conditions.

The β thalassemias

There are two main classes of β thalassemia, β^0 thalassemia, in which there is an absence of β globin chain production, and β^+ thalassemia, in which there is a variable reduction in the output of β globin chains. As shown in Figure 1.4, mutations of the β globin genes may cause a reduced output of gene product at the level of transcription or mRNA processing, translation, or through the stability of the globin gene product.

Defective β globin gene transcription

There are a variety of mechanisms that interfere with the normal transcription of the β globin genes. First, the genes may be either completely or partially deleted. Overall, deletions of the β globin genes are not commonly found in patients with β thalassemia, with one exception: a 619 bp deletion involving the 3′ end of the gene is found frequently in the Sind populations of India and Pakistan, where it constitutes about 30% of the β thalassemia alleles. Other deletions are extremely rare.

A much more common group of mutations, which results in a moderate decrease in the rate of transcription of the β globin genes, involves single nucleotide substitutions in or near the TATA box at about −30 nucleotides (nt) from the transcription start site, or in the proximal or distal promoter elements at −90 nt and −105 nt. These mutations result in decreased β globin mRNA production, ranging from 10 to 25% of the normal output. Thus, they are usually associated with the mild forms of β^+ thalassemia. They are particularly common in African populations, an observation which explains the unusual mildness of β thalassemia in this racial group. One particular mutation, C→T at position −101 nt to the β globin gene, causes an extremely mild deficit of β globin mRNA. Indeed, this allele is so mild that it is completely silent in carriers and can only be identified by its interaction with more severe β thalassemia alleles in compound heterozygotes.

Mutations that cause abnormal processing of mRNA

As mentioned earlier, the boundaries between exons and introns are marked by the invariant dinucleotides GT at the donor (5′) site and AG at the acceptor (3′) site. Mutations (base changes) that affect either of these sites completely abolish normal splicing and produce the phenotype of β^0 thalassemia. The transcription of genes carrying these mutations appears to be normal, but there is complete inactivation of splicing at the altered junction.

Another family of mutations involves what are called 'splice site consensus sequences'. Although only the GT dinucleotide is invariant at the donor splice site, there is conservation of adjacent nucleotides and a common, or consensus, sequence of these regions can be identified. Mutations within this sequence can reduce the efficiency of splicing to varying degrees because they lead to alternate splicing at the surrounding cryptic sites. For example, mutations of the nucleotide at

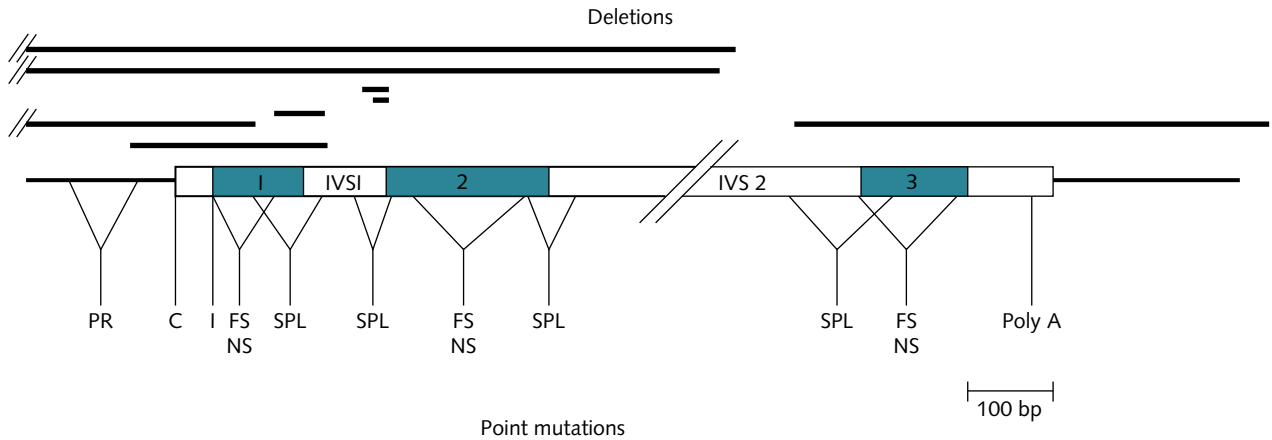


Fig. 1.4 The mutations of the β globin gene that underlie β thalassemia
The heavy black lines indicate the length of the deletions. The point mutations are designated as follows: PR, promoter; C, CAP site; I, initiation codon; FS, NS, frameshift and nonsense mutations; SPL, splice mutations; Poly A, poly A addition site mutations.

position 5 of IVS-1 (the first intervening sequence), G→C or T, result in a marked reduction of β chain production and in the phenotype of severe β^+ thalassemia. On the other hand, the substitution of C for T at position 6 in IVS-1 leads to only a mild reduction in the output of β chains.

Another mechanism that leads to abnormal splicing involves 'cryptic splice sites'. These are regions of DNA which, if mutated, assume the function of a splice site at an inappropriate region of the mRNA precursor. For example, a variety of mutations activate a cryptic site which spans codons 24–27 of exon 1 of the β globin gene. This site contains a GT dinucleotide, and adjacent substitutions that alter it so that it more closely resembles the consensus donor splice site result in its activation, even though the normal splice site is intact. A mutation at codon 24 GGT→GGA, though it does not alter the amino acid which is normally found in this position in the β globin chain (glycine), allows some splicing to occur at this site instead of the exon–intron boundary. This results in the production of both normal and abnormally spliced β globin mRNA and hence in the clinical phenotype of severe β thalassemia. Interestingly, mutations at codons 19, 26 and 27 result in both reduced production of normal mRNA (due to abnormal splicing) and an amino acid substitution when the mRNA which is spliced normally is translated into protein. The abnormal hemoglobins produced are hemoglobins Malay, E and Knossos, respectively. All these variants are associated with a mild β^+ thalassemia-like phenotype. These mutations illustrate how sequence changes in coding rather than intervening sequences influence RNA processing, and underline the importance of competition between potential splice site sequences in generating both normal and abnormal varieties of β globin mRNA.

Cryptic splice sites in introns may also carry mutations that activate them even though the normal splice sites remain intact. A common mutation of this kind in Mediterranean populations involves a base substitution at position 110 in IVS-1. This region contains a sequence similar to a 3′ acceptor site, though it lacks the invariant AG dinucleotide. The change of the G to A at position 110 creates this dinucleotide. The result is that about 90% of the RNA transcript splices to this particular site and only 10% to the normal site, again producing the phenotype of severe β^+ thalassemia (Figure 1.5). Several other β thalassemia mutations have been described which generate new donor sites within IVS-2 of the β globin gene.

Another family of mutations that interferes with β globin gene processing involves the sequence AAUAAA in the 3′ untranslated regions, which is the signal for cleavage and polyadenylation of the β globin gene transcript. Somehow, these mutations destabilize the transcript. For example, a T→C substitution in this sequence leads to only one-tenth of the normal amount of β globin mRNA transcript and hence to the phenotype of a moderately severe β^+ thalassemia. Another example of a mutation which probably leads to defective processing of function of β globin mRNA is the single base substitution, A→C, in the CAP site. It is not yet understood how this mutation causes a reduced rate of transcription of the β globin gene.

There is another small subset of rare mutations which involve the 3′ untranslated region of the β globin gene and are associated with relatively mild forms of β thalassemia. It is

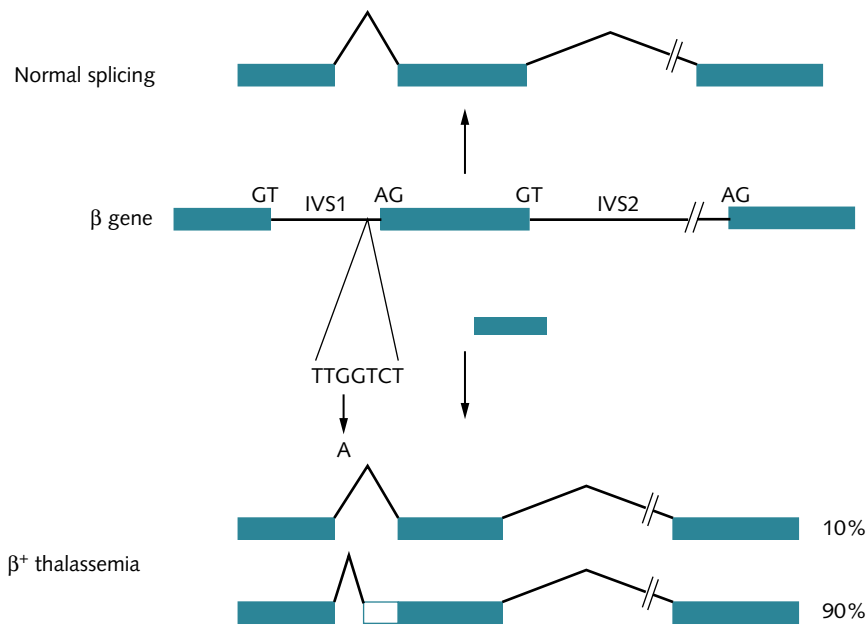


Fig. 1.5 The generation of a new splice site in an intron as the mechanism for a form of β^+ thalassemia
For details see text.

thought that these interfere in some way with transcription but the mechanism is unknown.

Mutations that result in abnormal translation of β globin mRNA

There are three main classes of mutations of this kind. Base substitutions that change an amino acid codon to a chain termination codon prevent the translation of β globin mRNA and result in the phenotype of β^0 thalassemia. Several mutations of this kind have been described; the commonest, involving codon 17, occurs widely throughout Southeast Asia. Similarly, a codon 39 mutation is encountered frequently in the Mediterranean region.

The second class involves the insertion or deletion of one, two or four nucleotides in the coding region of the β globin gene. These disrupt the normal reading frame, cause a frameshift, and hence interfere with the translation of β globin mRNA. The end result is the insertion of anomalous amino acids after the frameshift until a termination codon is reached in the new reading frame. This type of mutation always leads to the phenotype of β^0 thalassemia.

Finally, there are several mutations which involve the β globin gene initiation codon and which, presumably, reduce the efficiency of translation.

Unstable β globin chain variants

Some forms of β thalassemia result from the synthesis of highly unstable β globin chains which are incapable of forming hemoglobin tetramers, and which are rapidly degraded, leading to the phenotype of β^0 thalassemia. Indeed, in many of these conditions no abnormal globin chain product can be demonstrated by protein analysis and the molecular pathology has to be interpreted simply on the basis of a derived sequence of the variant β chain obtained by DNA analysis.

Recent studies have provided some interesting insights into how complex clinical phenotypes may result from the synthesis of unstable β globin products. For example, there is a spectrum of disorders that result from mutations in exon 3 which give rise to a moderately severe form of β thalassemia in heterozygotes. It has been found that nonsense or frameshift mutations in exons I and II are associated with the absence of messenger RNA from the cytoplasm of red cell precursors. This appears to be an adaptive mechanism, called 'nonsense-mediated decay', whereby abnormal messenger RNA of this type is not transported to the cytoplasm, where it would act as a template for the production of truncated gene products. However, in the case of exon III mutations, apparently because this process requires the presence of an intact upstream exon, the abnormal messenger RNA is transported into the cytoplasm and hence can act as a template for the production of unstable β globin chains. The latter precipitate in the red cell precursors together with excess α chains to form large inclusion bodies, and hence there is enough globin chain imbalance in heterozygotes to produce a moderately severe degree of anemia.

The molecular pathology of the α thalassemias

The molecular pathology of the α thalassemias is more complicated than that of the β thalassemias, simply because there are two α globin genes per haploid genome. Thus, the normal α globin genotype can be written αα/αα. As in the case of β thalassemia, there are two major varieties of α thalassemia, α^+ and α^0 thalassemia. In α^+ thalassemia one of the linked α globin genes is lost, either by deletion (–) or mutation (T); the heterozygous genotype can be written –α/αα or α^Tα/αα. In α^0 thalassemia the loss of both α globin genes nearly always results from a deletion; the heterozygous genotype is therefore written – –/αα. In populations where specific deletions are particularly common—Southeast Asia (SEA) or the Mediterranean region (MED)—it is useful to add the appropriate superscript, as follows: $--^{SEA}$/αα or $--^{MED}$/αα. It follows that when we speak of an 'α thalassemia gene' what we are really referring to is a haplotype; that is, the state and function of both of the linked α globin genes.

α^0 Thalassemia

Three main molecular pathologies, all involving deletions, have been found to underlie the α^0 thalassemia phenotype. The majority of cases result from deletions that remove both α globin genes and a varying length of the α globin gene cluster (Figure 1.6). Occasionally, however, the α globin gene cluster is intact but is inactivated by a deletion which involves the major regulatory region HS40, 40 kb upstream from the α globin genes. Finally, the α globin genes may be lost as part of a truncation of the tip of the short arm of chromosome 16.

As well as providing us with an understanding of the molecular basis for α^0 thalassemia, detailed studies of these deletions have yielded more general information about the mechanisms that underlie this form of molecular pathology. For example, it has been found that the 5′ breakpoints of a number of deletions of the α globin gene cluster are located approximately the same distance apart and in the same order along the chromosome as their respective 3′ breakpoints; similar findings have been observed in deletions of the β globin gene cluster. These deletions seem to have resulted from illegitimate recombination events which have led to the deletion of an integral number of chromatin loops as they pass through their nuclear attachment points during chromosomal replication. Another long deletion has been characterized

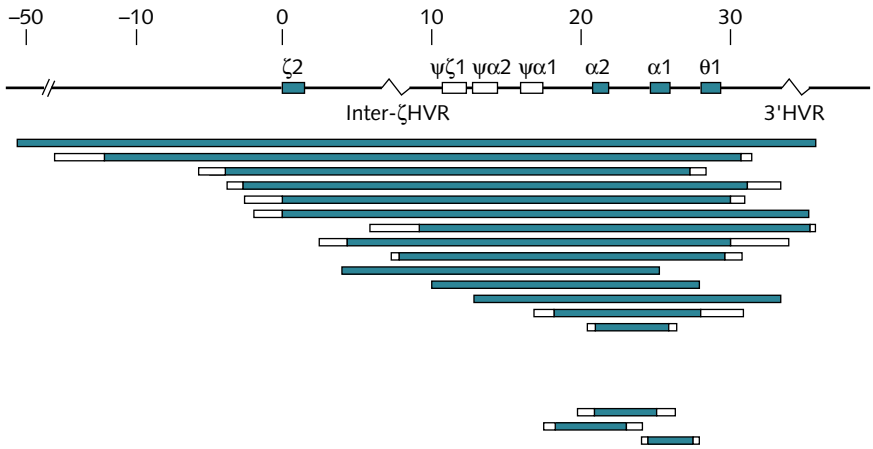


Fig. 1.6 Some of the deletions that underlie α^0 and α^+ thalassemia
The heavy red lines indicate the lengths of the deletions. The unshaded regions indicate uncertainty about the precise breakpoints. The three small deletions at the bottom of the figure represent the common α^+ thalassemia deletions.

in which a new piece of DNA bridges the two breakpoints in the α globin gene cluster. The inserted sequence originates upstream from the α globin gene cluster, where it normally is found in an inverted orientation with respect to that found between the breakpoints of the deletion. Thus it appears to have been incorporated into the junction in a way that reflects its close proximity to the deletion breakpoint region during replication. Other deletions seem to be related to the family of Alu-repeats, simple repeat sequences that are widely dispersed throughout the genome; one deletion appears to have resulted from a simple homologous recombination between two repeats of this kind that are usually 62 kb apart.

A number of forms of α^0 thalassemia result from terminal truncations of the short arm of chromosome 16 to a site about 50 kb distal to the α globin genes. The telomeric consensus sequence TTAGGGn has been added directly to the site of the break. Since these mutations are stably inherited, it appears that telomeric DNA alone is sufficient to stabilize the ends of broken chromosomes.

The molecular pathology of α^+ thalassemia

As mentioned earlier, the α^+ thalassemias result from the inactivation of one of the duplicated α globin genes, either by deletion or point mutation.

α^+ Thalassemia due to gene deletions

There are two common forms of α^+ thalassemia that are due to loss of one or other of the duplicated α globin genes, $-\alpha^{3.7}$ and $-\alpha^{4.2}$, where 3.7 and 4.2 indicate the sizes of the deletions. The way in which these deletions have been generated reflects the underlying structure of the α globin gene complex (Figure 1.7). Each α gene lies within a boundary of homology, approximately 4 kb long, probably generated by an ancient duplication event. The homologous regions, which are divided by small inserts, are designated X, Y and Z. The duplicated Z boxes are 3.7 kb apart and the X boxes are 4.2 kb apart. As the result of misalignment and reciprocal crossover between these segments at meiosis, a chromosome is produced with either a single ($-\alpha$) or triplicated ($\alpha\alpha\alpha$) α globin gene. As shown in Figure 1.7, if a crossover occurs between homologous Z boxes 3.7 kb of DNA are lost, an event which is described as a rightward deletion, $-\alpha^{3.7}$. A similar crossover between the two X boxes deletes 4.2 kb, the leftward deletion $-\alpha^{4.2}$. The corresponding triplicated α gene arrangements are called $\alpha\alpha\alpha^{anti\ 3.7}$ and $\alpha\alpha\alpha^{anti\ 4.2}$. A variety of different points of crossing over within the Z boxes give rise to different length deletions, still involving 3.7 kb.

Non-deletion types of α^+ thalassemia

These disorders result from single or oligonucleotide mutations of the particular α globin gene. Most of them involve the α2 gene but, since the output from this locus is two to three times greater than that from the α1 gene, this may simply reflect ascertainment bias due to the greater phenotypic effect and, possibly, a greater selective advantage.

Overall, these mutations interfere with α globin gene function in a similar way to those that affect the β globin genes. They affect the transcription, translation or post-translational stability of the gene product. Since the principles are the same as for β thalassemia, we do not need to describe them in detail with one exception, a mutation which has not been observed in the β globin gene cluster. It turns out that there is a family of mutations that involves the α2 globin gene termination codon, TAA. Each specifically changes this codon so that an amino acid is inserted instead of the chain terminating. This is followed by 'read-through' of α globin mRNA, which is not normally translated until another in-phase termination codon is reached. The result is an elongated α chain with 31 additional residues at the C terminal end. Five hemoglobin variants of this type have been identified. The commonest,

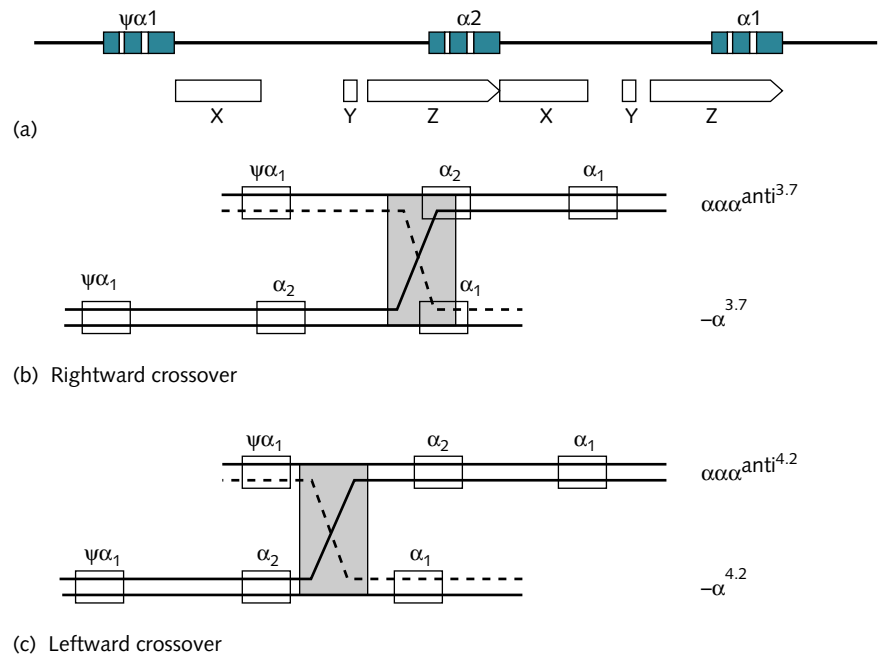


Fig. 1.7 Mechanisms of the generation of the common deletion forms of α^+ thalassemia
(a) The normal arrangement of the α globin genes, with the regions of homology X, Y and Z.
(b) The crossover that generates the $-\alpha^{3.7}$ deletion.
(c) The crossover that generates the $-\alpha^{4.2}$ deletion.

hemoglobin Constant Spring, occurs at a high frequency in many parts of Southeast Asia. It is not absolutely clear why the read-through of normally untranslated mRNAs leads to a reduced output from the α2 gene, although there is considerable evidence that it in some way destabilizes the mRNA.

α Thalassemia/mental retardation syndromes

There is a family of mild forms of α thalassemia which is quite different to that described in the previous section and which is associated with varying degrees of mental retardation. Recent studies indicate that there are two quite different varieties of this condition, one encoded on chromosome 16 (ATR-16) and the other on the X chromosome (ATR-X).

The ATR-16 syndrome is characterized by a relatively mild mental handicap with a variable constellation of facial and skeletal dysmorphisms. These individuals have long deletions involving the α globin gene cluster, but removing at least 1–2 Mb. This condition can arise in several ways, including unbalanced translocation involving chromosome 16, truncation of the tip of chromosome 16, and the loss of the α globin gene cluster and parts of its flanking regions by other mechanisms.

The ATR-X syndrome results from mutations in a gene on the X chromosome, Xq13.1-q21.1. The product of this gene is one of a family of proteins that are involved in chromatin-mediated transcriptional regulation. It is expressed ubiquitously during development and at interphase it is found entirely within the nucleus in association with pericentomeric heterochromatin. In metaphase, it is similarly found close to the centromeres of many chromosomes but, in addition, occurs at the stalks of acrocentric chromosomes, where the ribosomal (r) RNA is located. These locations provide important clues to the potential role of this protein in the establishment and/or maintenance of methylation of the genome. Although it is clear that ATR-X is involved in α globin transcription, it also must be an important player in early fetal development, particularly of the urogenital system and brain. Many different mutations of this gene have been discovered in association with the widespread morphological and developmental abnormalities which characterize the ATR-X syndrome.

α Thalassemia and the myelodysplastic syndrome

Since the first description of the finding of Hb H in the red cells of a patient with leukemia, many examples of this association have been reported. The condition usually is reflected in a mild form of Hb H disease, with typical Hb H inclusions in a proportion of the red cells and varying amounts of Hb H demonstrable by hemoglobin electrophoresis. The hematological findings are usually those of one or other form of the myelodysplastic syndrome. The condition occurs predominantly in males in older age groups. Very recently it has been found that some patients with this condition have mutations involving *ATR-X*. The relationship of these mutations to the associated myelodysplasia remains to be determined.

Rarer forms of thalassemia and related disorders

There are a variety of other conditions that involve the β globin gene cluster which, although less common than the β thalassemias, provide some important information about mechanisms of molecular pathology and therefore should be mentioned briefly.

The δβ thalassemias

Like the β thalassemias, the δβ thalassemias, which result from defective δ and β chain synthesis, are subdivided into the $(\delta\beta)^+$ and $(\delta\beta)^0$ forms.

The $(\delta\beta)^+$ thalassemias result from unequal crossing over between the δ and β globin gene loci at meiosis with the production of δβ fusion genes. The resulting δβ fusion chain products combine with α chains to form a family of hemoglobin variants called the hemoglobin Lepores, after the family name of the first patient of this kind to be discovered. Because the synthesis of these variants is directed by genes with the 5′ sequences of the δ globin genes, which have defective promoters, they are synthesized at a reduced rate and result in the phenotype of a moderately severe form of δβ thalassemia.

The $(\delta\beta)^0$ thalassemias nearly all result from long deletions involving the β globin gene complex. Sometimes they involve the $^{A}\gamma$ globin chains and hence the only active locus remaining is the $^{G}\gamma$ locus. In other cases the $^{G}\gamma$ and $^{A}\gamma$ loci are left intact and the deletion simply removes the δ and β globin genes; in these cases both the $^{G}\gamma$ and the $^{A}\gamma$ globin gene remains functional. For some reason, these long deletions allow persistent synthesis of the γ globin genes at a relatively high level during adult life, which helps to compensate for the absence of β and δ globin chain production. They are classified according to the kind of fetal hemoglobin that is produced, and hence into two varieties, $^{G}\gamma(^{A}\gamma\delta\beta)^0$ and $^{G}\gamma^{A}\gamma(\delta\beta)^0$ thalassemia; in line with other forms of thalassemia, they are best described by what is not produced: $(^{A}\gamma\delta\beta)^0$ and $(\delta\beta)^0$ thalassemia, respectively. Homozygotes produce only fetal hemoglobin, while heterozygotes have a thalassemic blood picture together with about 5–15% hemoglobin F.

Hereditary persistence of fetal hemoglobin (HPFH)

Genetically determined persistent fetal hemoglobin synthesis in adult life is of no clinical importance except that its genetic determinants can interact with the β thalassemias or structural hemoglobin variants; the resulting high level of Hb F production often ameliorates these conditions. The different forms of HPFH result from either long deletions involving the δβ globin gene cluster, similar to those that cause $(\delta\beta)^0$ thalassemia, or from point mutations that involve the promoters of the $^{G}\gamma$ or $^{A}\gamma$ globin gene. In the former case there is no β globin chain synthesis and therefore these conditions are classified as $(\delta\beta)^0$ HPFH. In cases in which there are promoter mutations involving the γ globin genes, there is increased γ globin chain production in adult life associated with some β and δ chain synthesis *in cis*, i.e. directed by the same chromosome, to the HPFH mutations. Thus, depending on whether the point mutations involve the promoter of the $^{G}\gamma$ or $^{A}\gamma$ globin gene, these conditions are called $^{G}\gamma\,\beta^+$ HPFH and $^{A}\gamma\,\beta^+$ HPFH, respectively.

There is another family of HPFH-like disorders in which the genetic determinant is not encoded in the β chain cluster. In one case the determinant encodes on chromosome 6, although its nature has not yet been determined.

It should be pointed out that all these conditions are very heterogeneous and that many different deletions or point mutations have been discovered that produce the rather similar phenotypes of $(\delta\beta)^0$ or $^{G}\gamma$ or $^{A}\gamma\,\beta^+$ HPFH.

Structural hemoglobin variants

Over 700 structural hemoglobin variants have been described, most of which are of no clinical significance. Only if the underlying mutation interferes with the stability or function of the hemoglobin molecule is there any important clinical accompaniment.

The majority of these variants result from missense mutations; that is, base substitutions which produce a codon change which encodes a different amino acid in the affected globin chain. Rarely, structural variants result from more subtle alterations in the structure of the α or β globin genes. Shortened chains may result from internal deletions of their particular genes, while elongated chains result either from duplications within genes or frameshift mutations which allow the chain termination codon to be read through and additional amino acids to be added to the C terminal end.

Genotype–phenotype relationships in the inherited disorders of hemoglobin

It is now necessary briefly to relate the remarkably diverse molecular pathology described in the previous sections to the phenotypes observed in patients with these diseases. It will not be possible to describe all these complex issues here. Rather we shall focus on those aspects that illustrate the more general principles of how abnormal gene action is reflected in a particular clinical picture. Perhaps the most important question that we will address is why patients with apparently identical genetic lesions have widely differing disorders, a problem that still bedevils the whole field of medical genetics, even in the molecular era.

The β thalassemias

As we have seen, the basic defect that results from the 200 or more different mutations that underlie these conditions is reduced β globin chain production. Synthesis of the α globin chain proceeds normally and hence there is imbalanced globin chain output with an excess of α chains (Figure 1.8). Unpaired α chains precipitate in both red cell precursors and their progeny with the production of inclusion bodies. These interfere with normal red cell maturation and survival in a variety of complex ways. Their attachment to the red cell membrane causes alterations in its structure, and their degradation products, notably heme, hemin (oxidized heme) and iron, result in oxidative damage to the red cell contents and membrane. These interactions result in intramedullary destruction of red cell precursors and in shortened survival of such cells as they reach the peripheral blood. The end result is an anemia of varying severity. This, in turn, causes tissue hypoxia and the production of relatively large amounts of erythropoietin; this leads to a massive expansion of the ineffective bone marrow, resulting in bone deformity, a hypermetabolic state with wasting and malaise, and bone fragility.

A large proportion of hemoglobin in the blood of β thalassemics is of the fetal variety. Normal individuals produce about 1% of Hb F, unevenly distributed among their red cells. In the bone marrow of β thalassemics, any red cell precursors that synthesize γ chains come under strong selection because they combine with α chains to produce fetal hemoglobin and

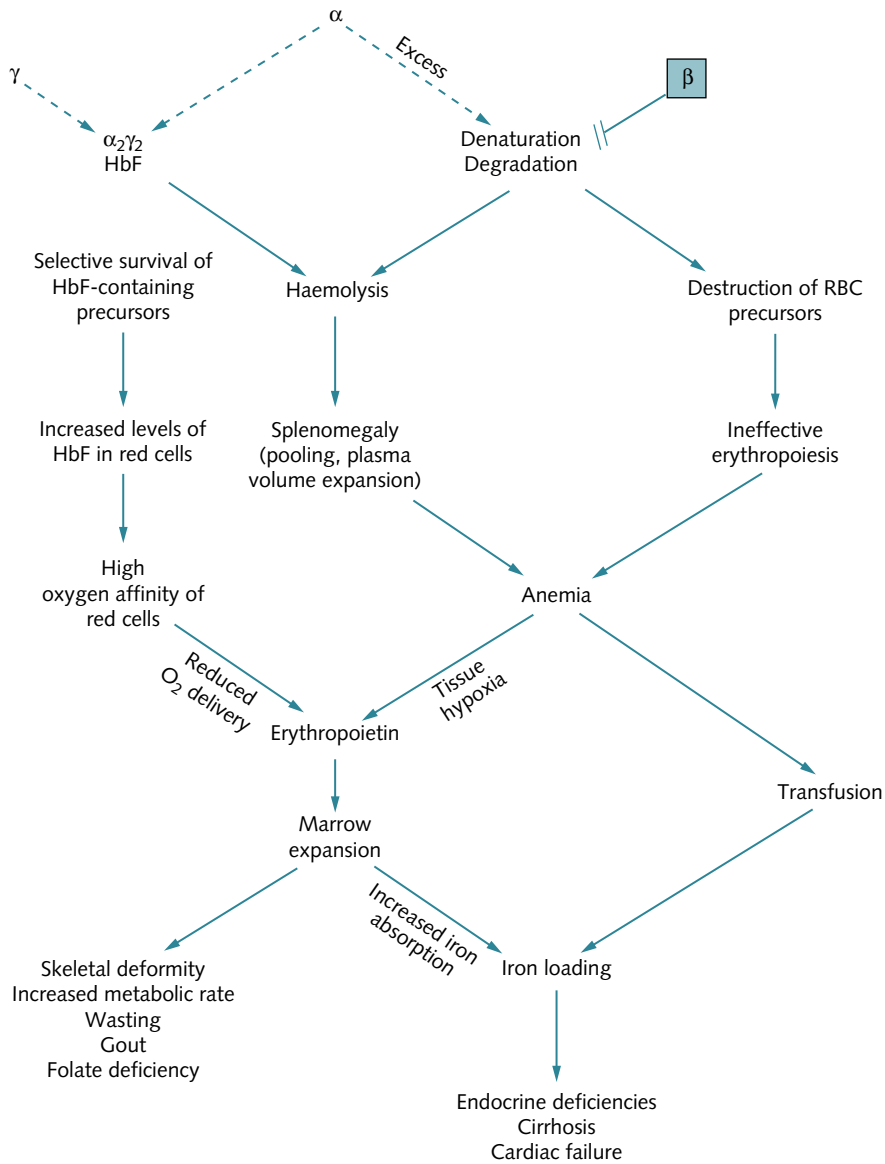


Fig. 1.8 The pathophysiology of β thalassemia

therefore the degree of globin chain imbalance is reduced. Furthermore, the likelihood of γ chain production seems to be increased in a highly stimulated erythroid bone marrow. It seems likely that these two factors combine to increase the relative output of hemoglobin F in this disorder. However, it has a higher oxygen affinity than hemoglobin A and hence patients with β thalassemia are not able to adapt to low hemoglobin levels as well as those who have adult hemoglobin.

The greatly expanded, ineffective erythron leads to an increased rate of iron absorption; this, combined with iron received by blood transfusion, leads to progressive iron loading of the tissues, with subsequent liver, cardiac and endocrine damage.

The constant bombardment of the spleen with abnormal red cells leads to its hypertrophy. Hence there is progressive splenomegaly with an increased plasma volume and trapping of part of the circulating red cell mass in the spleen. This leads to worsening of the anemia. All these pathophysiological mechanisms, except for iron loading, can be reversed by regular blood transfusion which, in effect, shuts off the ineffective bone marrow and its consequences.

Thus it is possible to relate nearly all the important features of the severe forms of β thalassemia to the primary defect in globin gene action. But can we also explain their remarkable clinical diversity? Part of it reflects the different mutations of the β globin genes. For example, some of the promoter or splice mutations cause an extremely mild form of β^+ thalassemia. Many β thalassemics are compound heterozygotes for either two severe β thalassemia alleles, a severe and mild allele, or different mild alleles, and this also accounts for a considerable amount of clinical diversity of the disease.

What of patients who have the same mutations at their β loci yet have completely different clinical phenotypes? The coinheritance of α thalassemia, which reduces the magnitude of the excess of α globin chains in β thalassemia, may ameliorate the clinical course. This remarkable experiment of nature provides unequivocal confirmation that the major pathophysiological mechanism that underlies β thalassemia is imbalanced globin chain synthesis. In other patients, especially those who are homozygous for β^0 thalassemia yet run a particularly mild course, it is apparent that unusually increased production of fetal hemoglobin is the main ameliorating factor. Although the precise mechanism is not yet understood, it is becoming apparent that genetic determinants both within the β globin gene cluster and on other chromosomes may be involved in this more effective production of Hb F. For example, a promoter polymorphism of the $^G\gamma$ globin gene may be associated with increased propensity to synthesize fetal hemoglobin, particularly in states of hemopoietic stress. Similarly, there is good evidence that a so-far unidentified gene on chromosome 6 may be involved in modifying the β thalassemia phenotype in this way. Undoubtedly other polymorphisms of this kind will be discovered.

There is increasing evidence that the complications of β thalassemia may be modified by polymorphisms at different loci. For example, variations at several loci that determine bone metabolism may be involved in modifying the severity of osteoporosis, a common complication of this disorder. Similarly, the occurrence of jaundice and iron loading from the intestine, both common complications of the intermediate forms of β thalassemia, are related to polymorphisms of genes involved in bilirubin and iron metabolism, respectively.

Recently it has been suggested that these complex layers of modifiers of β thalassemia should be placed in three classes: primary, including the different mutations that involve the β globin gene; secondary, including variation at the α globin locus or in fetal hemoglobin production; and tertiary, including modifiers that are involved in varying the severity of the different complications of the disease.

In short, while we have a reasonable idea of how the β thalassemia phenotype is modified, many questions remain and a considerable amount of the clinical variability of the disease remains unexplained.

The α thalassemias

The pathophysiology of the α thalassemias differs from that of the β thalassemias mainly because of the properties of the excess globin chains that are produced as a result of defective α chain synthesis. While the excess α chains produced in β thalassemia are unstable and precipitate, this is not the case in the α thalassemias, in which excess γ chains or β chains are able to form the soluble homotetramers γ_4 (Hb Bart's) and β_4 (Hb H) (Figure 1.9). Although these variants, particularly

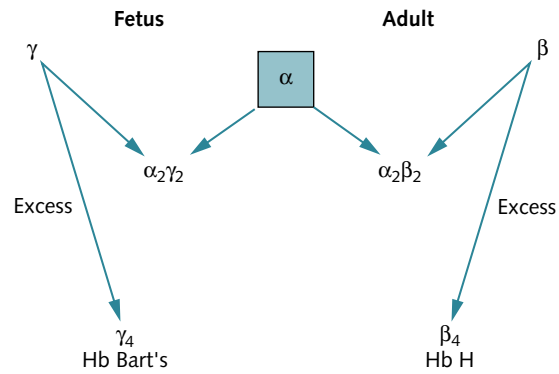


Fig. 1.9 The pathophysiology of α thalassemia

Hb H, are unstable and precipitate in older red cell populations, they remain soluble sufficiently long for the red cells to mature and develop relatively normally. Hence there is far less ineffective erythropoiesis in the α thalassemias and the main cause of the anemia is hemolysis associated with the precipitation of Hb H in older red cells. In addition, of course, there is a reduction in normal hemoglobin synthesis, which results in hypochromic, microcytic erythrocytes. Another important factor in the pathophysiology of the α thalassemias is the fact that Hb Bart's and Hb H are useless oxygen carriers, having an oxygen dissociation curve similar to that of myoglobin. Hence the circulating hemoglobin level may give a false impression of the oxygen-delivering capacity of the blood and patients may be symptomatic at relatively high hemoglobin levels.

The different clinical phenotypes of the α thalassemias are an elegant example of the effects of gene dosage (Figure 1.10). The heterozygous state for α^+ thalassemia is associated with minimal hematological changes. That for α^0 thalassemia (the loss of two α globin genes) is characterized by moderate hypochromia and microcytosis, similar to that of the β thalassemia trait. It does not matter whether the α genes are lost on the same chromosome or on opposite pairs of homologous chromosomes. Hence the homozygous state for α^+ thalassemia, $-\alpha/-\alpha$, has a similar phenotype to the heterozygous state for α^0 thalassemia ($--/\alpha\alpha$).

The loss of three α globin genes, which usually results from the compound heterozygous states for α^0 and α^+ thalassemia, is associated with a moderately severe anemia with the production of varying levels of hemoglobin H. This condition, hemoglobin H disease, is characterized by varying anemia and splenomegaly with a marked shortening of red cell survival.

Finally, the homozygous state for α^0 thalassemia ($--/--$) is characterized by death *in utero* or just after birth, with the clinical picture of hydrops fetalis. These babies produce no α chains and their hemoglobin consists mainly of Bart's with variable persistence of embryonic hemoglobin. This is reflected in gross intrauterine hypoxia; although these babies may have hemoglobin values as high as 8–9 g/dl, most of it is unable to release its oxygen. This is reflected in the hydropic changes, a massive outpouring of nucleated red cells, and hepatosplenomegaly with persistent hematopoiesis in the liver and spleen.

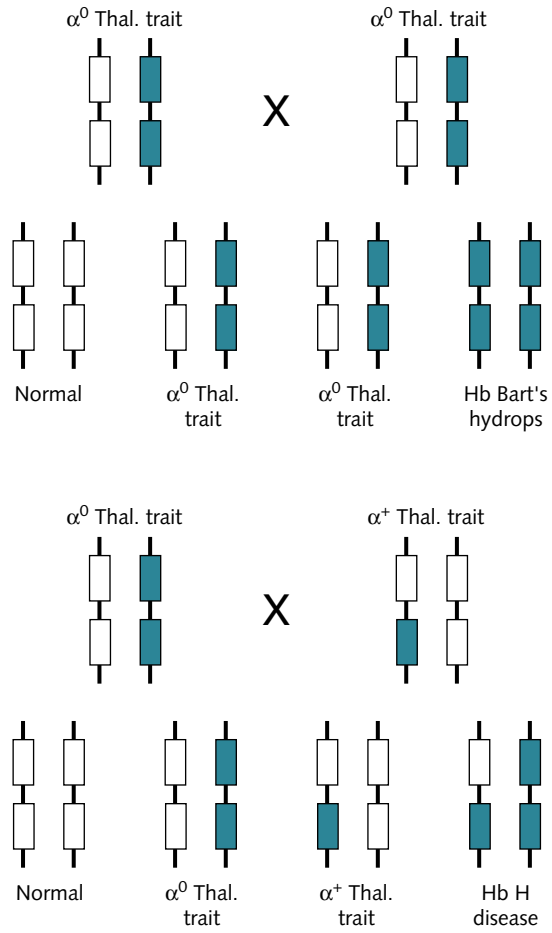


Fig. 1.10 The genetics of the common forms of α thalassemia
The light boxes represent normal α genes and the shaded boxes deleted α genes. The mating shown at the top shows how two α^0 thalassemia heterozygotes can produce a baby with the hemoglobin Bart's hydrops syndrome. In the mating at the bottom, between individuals with α^0 and α^+ thalassemia, one in four of the offspring will have hemoglobin H disease.

Structural hemoglobin variants

While most structural hemoglobin variants produce no clinical disability, a few, notably the sickling, and the rare variants are associated with instability or abnormal oxygen transport (*discussed in detail in Chapter 14*).

The sickling disorders

The sickling disorders represent the homozygous state for the sickle cell gene, sickle cell anemia, and the compound heterozygous state for the sickle cell gene and various structural hemoglobin variants, or β thalassemia. The chronic hemolysis and episodes of vascular occlusion and red cell sequestration that characterize sickle cell anemia can all be related to the replacement of the normal β6 glutamic acid by valine in hemoglobin S. This causes a hydrophobic interaction with another hemoglobin molecule, triggering aggregation into

large polymers. It is this change that causes the sickling distortion of the red blood cell and hence a marked decrease in its deformability. The resulting rigidity of the red cells is responsible for the vaso-occlusive changes that lead to many of the most serious aspects of all the sickling disorders.

The different conformations of sickle cells (banana-shaped or resembling a holly leaf) reflect different orientations of bundles of fibres along the long axis of the cell, the three-dimensional structure of which is constituted by a rope-like polymer composed of 14 strands. The rate and extent of polymer formation depend on the degree of oxygenation, the cellular hemoglobin concentration, and the presence or absence of Hb F. The latter inhibits polymerization and hence tends to ameliorate sickling. Polymerization of Hb S causes damage to the red cell membrane, the result of which is an irreversibly sickled cell. Probably the most important mechanism is cellular dehydration resulting from abnormalities of potassium/chloride cotransport and Ca^{2+}-activated potassium efflux. This is sufficient to trigger the Ca^{2+}-dependent (Gardos) potassium channel, providing a mechanism for the loss of potassium and water and leading to cellular dehydration.

The vascular pathology of the sickling disorders is not entirely related to the rigidity of sickled red cells, however. There is now a wealth of evidence that abnormal interactions between sickled cells and the vascular endothelium play a major role in the pathophysiology of the sickling disorders. Recently it has been demonstrated that nitric oxide may also play a role in some of the vascular complications of this disease. It has been found that nitric oxide reacts much more rapidly with free hemoglobin than with hemoglobin in erythrocytes and therefore it is possible that such decompartmentalization of hemoglobin into plasma, as occurs in sickle cell disease and other hemolytic anemias, diverts nitric oxide from its homeostatic vascular function.

These issues are discussed in greater detail in Chapter 14.

Unstable hemoglobin variants

There are a variety of different mechanisms underlying hemoglobin stability resulting from amino acid substitutions in different parts of the molecule. The first is typified by amino acid substitutions in the vicinity of the heme pocket, all of which lead to a decrease in the stability of the binding of heme to globin. A second group of unstable variants results from amino acids that simply disrupt the secondary structure of the globin chains. About 75% of globin is in the form of α helix, in which proline cannot participate except as part of one of the initial three residues. At least 11 unstable hemoglobin variants have been described that result from the substitution of proline for leucine, five that are caused by an alanine-to-proline change, and three in which proline is substituted for histidine. Another group of variants that causes disruption of the normal configuration of the hemoglobin molecule involves internal substitutions that somehow interfere with its stabilization by hydrophobic interactions. Finally, there are two groups of unstable hemoglobins that result from gross structural abnormalities of the globin subunits; many are due to deletions involving regions at or near interhelical corners. A few of the elongated globin chain variants are also unstable.

Abnormal oxygen transport

There is a family of hemoglobin variants that are associated with high oxygen affinity and hereditary polycythemia. Most result from amino acid substitutions that affect the equilibrium between the R and T states (*see earlier section—Structure and function*). Thus, many of them result from amino acid substitutions at the α_1/β_2 interface, the C terminal end of the β chain, and at the 2,3-DPG binding sites.

Congenital cyanosis due to hemoglobin variants

There is a family of structural hemoglobin variants that is designated hemoglobin M, to indicate congenital methemoglobinemia, and is further defined by their place of discovery. The iron atom of heme is normally linked to the imidazole group of the proximal histidine residue of the α and β chains. There is another histidine residue on the opposite side, near the sixth coordination position of the heme iron; this, the so-called distal histidine residue, is the normal site of binding of oxygen. Several M hemoglobins result from the substitution of a tyrosine for either the proximal or distal histidine residue in the α or β chain.

Postscript

In this short account of the molecular pathology of hemoglobin we have considered how mutations at or close to the α or β globin genes result in a diverse family of clinical disorders due to the defective synthesis of hemoglobin or its abnormal structure. Work in this field over the last 20 years has given us a fairly good idea of the repertoire of different mutations that underlie single-gene disorders and how these are expressed as discrete clinical phenotypes. Perhaps more importantly, however, the globin field has taught us how the interaction of a limited number of genes can produce a remarkably diverse series of clinical pictures, and something of the basis for how monogenic diseases due to the same mutation may vary widely in their clinical expression.

Further reading

Huisman THJ, Carver MFH, Erol Baysal E. (1997) *A Syllabus of Thalassemia Mutations*. Augusta, GA: The Sickle Cell Anemia Foundation.

Stamatoyannopoulos G, Grosveld F. (2001) Hemoglobin switching. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW *et al*., eds. *The Molecular Basis of Blood Diseases*, 3rd edn. Philadelphia: W.B. Saunders, pp. 136–182.

Steinberg MH, Forget BG, Higgs DR *et al*. (2001) *Disorders of Hemoglobin*. Cambridge: Cambridge University Press.

Weatherall DJ. (2001) Phenotype–genotype relations in monogenic disease: lessons from the thalassaemias. *Nature Genetics*, **2**, 245–255.

Weatherall DJ. (2001) Thalassemia. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW *et al*., eds. *The Molecular Basis of Blood Diseases*, 3rd edn. Philadelphia: W.B. Saunders, pp. 183–226.

Weatherall DJ, Clegg JB. (2001) *The Thalassaemia Syndromes*. Oxford: Blackwell.

Weatherall DJ, Clegg JB, Higgs DR *et al*. (2001) The hemoglobinopathies. In: Scriver CR, Beaudet AL, Sly WS *et al*., eds. *The Metabolic Basis of Inherited Disease*, 8th edn. New York: McGraw-Hill.

Chapter 2 Molecular cytogenetics

Debra M Lillington, Silvana Debernardi & Bryan D Young

Introduction

The use of cytogenetics and molecular cytogenetic analysis in hematology has both increased and improved over the last decade. Fluorescence *in situ* hybridization (FISH) has been incorporated into most diagnostic laboratories to complement chromosome analysis and further improve its accuracy. In the era of risk-adapted and mutation-directed therapy, accurate assessment of genetic status is of paramount importance. In many current studies, patients are stratified on the basis of their cytogenetic or molecular rearrangements, since numerous disease- or subtype-specific abnormalities have independent prognostic outcomes. Such is the specificity of certain chromosomal rearrangements that molecular cytogenetic information can provide an unequivocal diagnosis of the type of malignancy. This is true for both leukemia and lymphoma, although the number of recurrent chromosomal changes identified in leukemia exceeds that in lymphoma. Nevertheless, a number of highly specific lymphoma-associated changes have diagnostic and therapeutic value. In national treatment trials in the UK, such as those run by the Medical Research Council for acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL), cytogenetic information is vital to treatment stratification, and in other diseases, such as chronic lymphoblastic leukemia and myeloma, the impact of chromosomal abnormalities is beginning to be recognized.

Chromosomal analysis of metaphase cells provides a global assessment of karyotype and still plays a major role in modern tumor cytogenetics. FISH is used as a rapid, sensitive test to complement G-band analysis, allowing the detection of cryptic or subtle changes. In addition, FISH can be used to screen non-dividing cell populations, such as bone marrow smears, tumor imprints and paraffin-embedded tissue sections (PETS). A vast array of FISH probes is currently available, aimed at detecting fusion genes, numerical abnormalities, chromosomal imbalance, chromosomal rearrangement and complex events. FISH has been further developed to allow the global detection of tumor-associated gain and loss using tumor DNA as a FISH probe against normal metaphase chromosomes. This technique is known as 'comparative genomic hybridization' (CGH) and in turn has led to the very recent technique of array CGH. Array CGH promises to provide much higher resolution of genomic imbalance compared with chromosome-based CGH. Currently, custom-designed arrays containing oncogenes and tumor suppressor genes are available commercially and arrays with DNA clones spaced at 1 Mb intervals throughout the genome are becoming available. Gene expression profiling also offers exciting prospects in hematology and, coupled with the molecular cytogenetic and cytogenetic information, accurate diagnostic genetic analysis looks set to revolutionize patient management.

FISH on metaphase chromosomes

The production of metaphase chromosomes from malignant cells plays a fundamental role in genetic analysis, allowing both G-banded chromosome analysis and subsequent FISH analysis. The chromosome offers a more versatile target than interphase cells since many types of FISH probes can be applied. In leukemia and lymphoma, gene fusions are relatively frequent and well characterized at the molecular level. These novel disease-associated fusion events arise through chromosomal translocations, inversions or insertions and are usually visible by routine karyotype analysis, although subtle abnormalities do exist and some of the recurrent rearrangements can be cryptic. FISH probes mapping to the unique sequences involved in these fusions are readily available and detect their respective abnormalities by one of two methods. In the first strategy, probes mapping to the two genes involved are labeled in two distinct colors; as an example, the BCR–ABL (break cluster region–Abelson) fusion associated with the t(9;22)(q34;q11.2) is illustrated in Plate 2.1. BCR is represented by the green fluorescence and ABL by the red signal. The t(9;22) translocation results in both BCR–ABL and ABL–BCR fusions, and since the probe extends beyond the breakpoint for both genes, two fusion signals (red and green juxtaposed) are

generated (dual fusion probes), one on the der(9), the other on the der(22). A normal 9 and a normal 22 (single red and green signal) will also exist. To further complicate the analysis, however, deviations from this pattern may exist since some patients carry deletions around the breakpoint and some harbor cryptic insertions of part of one gene, thereby generating only one of the fusion sequences (Plate 2.2). The next generation of FISH probes look likely to use four fluorescent probes to enhance the sensitivity and specificity to simultaneously detect translocations and deletions around the breakpoint, which may confer independent prognostic value. Plate 2.3 shows a cryptic insertion of part of the RARA gene (chromosome 17) into the PML locus (chromosome 15) in a patient with acute promyelocytic leukemia. The t(15;17)(q21;q11) translocation is the hallmark of acute promyelocytic leukemia and is cytogenetically visible in 90% of patients.

The second common type of FISH strategy is the 'break apart' probe, specifically designed to detect abnormalities affecting one specific gene which rearranges with multiple partner loci, such as *MLL* (11q23). Over 60 different *MLL* gene translocations have been cytogenetically reported, and the FISH probe used most often for diagnosis consists of a probe mapping above the breakpoint labeled with one color and a second probe mapping below the breakpoint in another color. Translocations involving *MLL* therefore result in the separation of one set of probes (Plate 2.4) and the displaced *MLL* signal will map to the partner chromosome. Single-color probes extending across the breakpoints can also be used, resulting in a split signal.

Unique sequence probes can also be used to screen for copy number changes, particularly in cases with evidence of additional genetic material, by karyotyping such as double minute chromosomes (dm), homogeneously staining regions (hsr) or additional pieces of chromosomes. Dm and hsr are manifestations of gene amplification and in certain malignant diseases, particularly solid tumors, are well recognized mechanisms for oncogene activation. FISH probes mapping to the genes commonly associated with amplification can very quickly confirm the presence of multiple copies of genes; an example is N-MYC in neuroblastoma. Plate 2.5 shows a bone marrow aspirate infiltrated by neuroblastoma and multiple copies of N-MYC. Alpha satellite probes are often used to determine chromosome number. Hyperdiploidy is a frequent phenomenon in ALL and is associated with a common pattern of gain, namely chromosomes 4, 6, 10, 14, 17, 18, 21 and X. Using a selected cocktail of alpha satellite probes mapping to these chromosomes, hyperdiploidy can be detected in both metaphase and interphase cells (Plate 2.6). Metaphase cells derived from leukemic blasts of patients with ALL can often have poor morphology and be difficult to fully characterize. In such situations FISH can be of particular value since it may help elucidate chromosomal gains and losses.

Whole-chromosome painting probes (WCP), consisting of pools of DNA sequences mapping along the full length of a particular chromosome and labeled with a fluorochrome, can be used individually or in combination with other WCPs to characterize abnormalities whose origin is uncertain by G-banding. In simple karyotypes, requiring confirmation of a suspected rearrangement, two-color chromosome painting might be the most useful option (Plate 2.7). In more complex karyotypes, such as those associated with therapy-related leukemia, a mixture of paints mapping to all 24 human chromosomes (24-color karyotyping) is probably the most informative. M-FISH/SKY is not used routinely for diagnostic purposes but has revealed cryptic rearrangements in several studies. M-FISH/spectral karyotyping (SKY) uses a combinatorial labeling approach such that each individual chromosome paint is labeled with a unique combination of not more than five fluorochromes. The 24 differentially labeled paints are then applied in a single hybridization assay and visualization is achieved using one of two strategies. M-FISH uses a series of optical filters to collect the images from the different fluorochromes, which are then merged into a composite image; a pseudocolor is then assigned to each chromosome on the basis of its fluorochrome combination (Plate 2.8). SKY uses an interferometer with Fourier transformation to determine the spectral characteristics of each pixel in the image, and assigns a pseudocolor.

FISH on nuclei

The non-dividing cell population can be examined using FISH probes to yield both diagnostic and prognostic molecular cytogenetic information. Interphase cells from the sample sent for cytogenetic analysis (peripheral blood, bone marrow, lymph node, etc.), tumor touch imprints, PETS and bone marrow smears can be used as the target for FISH. In these instances screening for a specific chromosomal abnormality is performed, thereby allowing detection or exclusion of a single event per FISH assay. Techniques that involve the use of whole or chromosome-specific probes cannot readily be applied to interphase cells. PETS are sometimes the only tumor sample available for analysis since the paraffin treatment preserves the morphology of the tumor, thereby enabling histological diagnosis. However, if histology is equivocal, FISH for a tumor-associated chromosomal abnormality can be extremely valuable for diagnostic purposes. FISH on PETS does have inherent technical problems not found with FISH on other sample types; for example, probe accessibility is reduced, the thickness of the section means that the resulting FISH signals may not all be visible in the same focal plane, and the cells may be very tightly packed, making analysis more difficult. Interphase cells can be used to assess the copy number of unique

sequences/alpha satellites or to look for chimeric fusion genes. Plate 2.9 illustrates two PETS screened for the presence of the EWS/FLI1 rearrangement associated with Ewing's sarcoma/primitive neuroectodermal tumor.

CGH and array CGH

CGH provides a global assessment of copy number changes, revealing regions of the chromosome that are either gained or lost in the tumor sample. A key feature of this technique is that dividing tumor cells are not required. DNA is extracted from the tumor, labeled (usually) with a green fluorochrome and compared with DNA from a normal reference labeled with a red fluorochrome. Labeled test and reference DNA are combined and hybridized to normal chromosomes and the resulting ratio of the two signals along the length of the chromosomes reflects the differences in copy number between the tumor and reference DNA samples (Plate 2.10). Regions of gain in the tumor DNA are represented by an increased green/red ratio whereas deletions are indicated by a reduced ratio. CGH requires 50% abnormal cells to be present within the tumor sample for reliable detection of genomic imbalance and will not easily detect regions involving less than 10 Mb of DNA unless it involves high-level amplification. Nevertheless, CGH is particularly applicable to the analysis of solid tumors since DNA can be readily extracted from them and the karyotype frequently involves loss or gain of whole or partial chromosomes.

The use of genomic DNA arrays as the hybridization target essentially allows much higher resolution for the detection of copy number changes. Array CGH is based on the metaphase CGH method but uses mapped sequences as the hybridization target rather than metaphase chromosomes, and hence the resolution is limited only by the density of the sequences spotted on the slides. Array CGH was first introduced in 1997 (Solinas-Toldo *et al.*, 1997), and since then a number of groups have set up their own facilities for spotting sequences onto slides. In addition, some companies have produced commercial chips ranging from specialist arrays containing 287 targets (Vysis, Inc., Downers Grove, IL, USA) to 1 Mb arrays (e.g. Spectral Genomics, Inc., Houston, TX, USA) containing around 3000 target clones. Array CGH has the potential to provide a highly sensitive global assessment of gene copy number, simultaneously screening hundreds of individual gene sequences, although its reproducibility and robustness need to be thoroughly validated before it is used in the diagnostic setting. Figure 2.1 clearly demonstrates the sensitivity of array CGH. This figure shows regions of simultaneous gain and loss along chromosome 13 and demonstrates how the technique gives more precise definition than metaphase CGH.

Gene expression profiling

The introduction of microarrays for gene expression profiling now offers a new approach to molecular cytogenetics. Measurement of the expression of all the genes in a range of tissues or cell types allows the determination of the transcriptional status of the cell, identifying which genes are active and which are silent. Microarrays for expression profiling consist of systematic arrays of cDNA or oligonucleotides of known sequence that are spotted or synthesized at discrete loci on a glass or silicon surface. They allow the simultaneous analysis of a large number of genes at high resolution following the hybridization of labeled cDNA or cRNA derived from the samples to be examined. Microarray output is represented by a large number of individual data points that must be analyzed by a data-mining program in order to correlate the data, and to group them together in a meaningful manner. In recent years, the use of DNA microarrays has been largely devoted to the genetic profiling of tumor subtypes, with the aims of defining new classes with prognostic and diagnostic relevance and of increasing our knowledge of the mechanisms underlying the biology of these diseases. The pathological diagnosis and classification of human neoplasia is based on well-defined morphological, cytochemical, immunophenotypic and clinical criteria. For leukemia and lymphoma, the relevance of cytogenetics as one of the most valuable prognostic determinants at diagnosis, has come from the analysis of the leukemia karyotype. This has identified non-random, somatically acquired translocations, inversions and deletions, which are often associated with specific morphological subtypes. However, leukemias with apparently normal karyotypes do exist and constitute the largest single subgroup (up to 40% of cases). Thus, the application of microarray analysis may improve the classification of leukemias and offer clues to the underlying etiology. A molecular classification would have the potential to define new subgroups with more prognostic and therapeutic significance, linking the expression profile to the outcome. It could offer many advantages over conventional classification methods, including the possibility of deducing chromosomal data from non-dividing cells.

Three basic steps for efficient and effective data analysis are necessary: data normalization, data filtering, and pattern identification. To compare expression values directly, it is necessary to apply some sort of normalization strategy to the data, either between paired samples or across a set of experiments. To 'normalize' in the context of DNA microarrays means to standardize the data so as to be able to differentiate between real (biological) variations in gene expression levels and variations due to the measurement process. Gene expression data can then be subjected to a variation filter, which excludes uninformative genes, i.e. genes showing minimal variation across the samples, and genes expressed below or

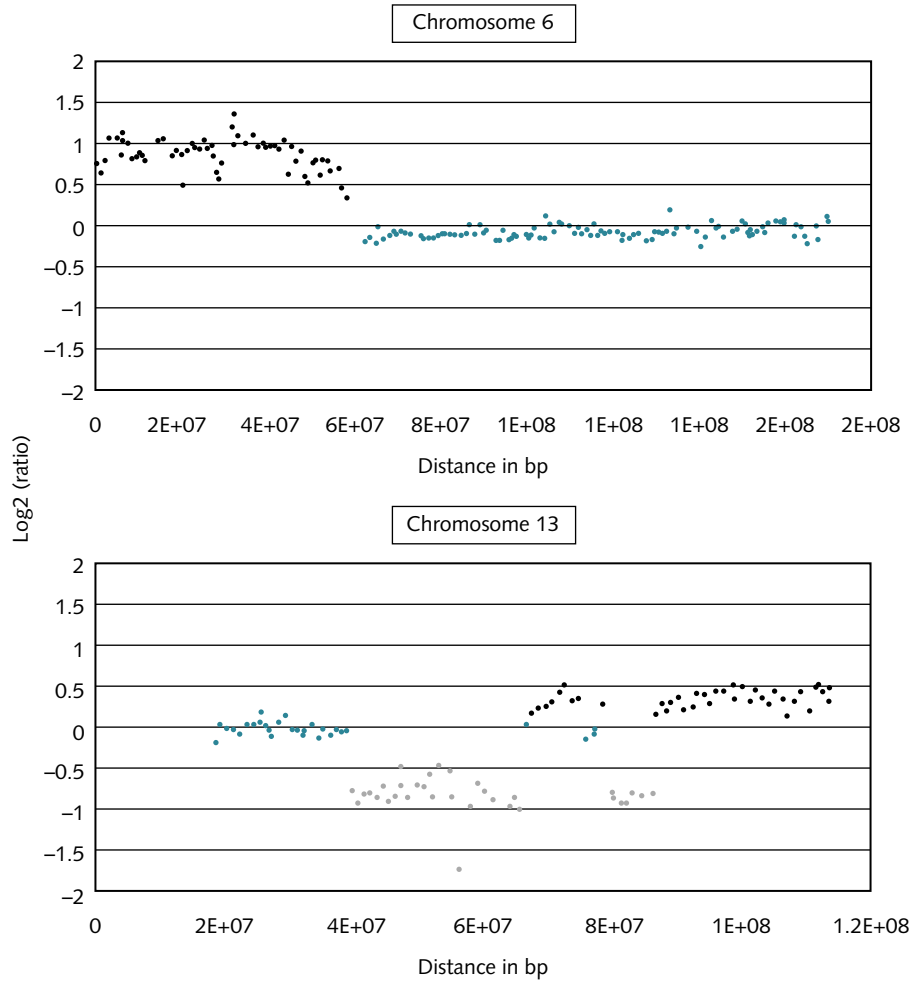


Fig. 2.1 A 1 Mb CGH-array (Sanger Centre) of a retinoblastoma tumor
(Top) Gain of BAC clones mapping to 6p (green). (Bottom) Two regions of gain on 13q (black) and two regions of deletion on 13q (gray).

above a user-defined threshold. This step facilitates the search for partners and groups in the data that can be used to assign biological meaning to the expression profiles, leading to the production of straightforward lists of increasing or decreasing genes or of more complex associations with the help of sophisticated clustering and visualization programs. Hierarchical clustering is used traditionally in phylogenetic analysis for the classification of organisms into trees; in the microarray context it is applied to genes and samples. Organisms sharing properties tend to be clustered together. The length of a branch containing two organisms can be considered a measure of how different the organisms are. It is possible to classify genes in a similar manner, gathering those whose expression patterns are similar into clusters in the tree. Such mock-phylogenetic trees are often referred to as 'dendrograms'. Genes can also be grouped on the basis of their expression patterns using *k-means clustering*. The goal is to produce groups of genes with a high degree of similarity within each group and a low degree of similarity between groups. The self-organizing map (SOM) is a clustering technique similar to k-means clustering, but in addition illustrates the relationship between groups by arranging them in a two-dimensional map. SOMs are useful for visualizing the number of distinct expression patterns in the data. A complex data set can also be reduced to a few specified dimensions by applying multidimensional scaling, so that the relationships between groups can be more effectively visualized.

The first classification of cancer on the basis of gene expression showed that it was possible to distinguish between myeloid and lymphoid acute leukemias by the use of arrays with approximately 6800 human genes. Since then, the approach has been applied successfully to the classification of hematological malignancies and a large variety of solid tumors. Recently, acute lymphoid leukemias with rearrangements of the *MLL* gene were shown to have expression patterns that could allow them to be distinguished from ALLs and AMLs without the *MLL* translocations. Further microarray analysis of AML cases with a favorable outcome—AML M2 with t(8;21), AML M3 or M3v with t(15;17) and AML M4eo with inv(16)—has shown a specific pattern of predictor genes associated with

the three subclasses In a subsequent microarray study, AML leukemia samples were specifically chosen to represent the spectrum of known karyotypes common in AML and included examples with AML-FAB phenotypes from M1 to M5. Hierarchical clustering sorted the profiles into separate groups, each representing one of the major cytogenetic classes in AML (i.e. t(8;21), t(15;17), inv(16), 11q23) and a normal karyotype, as shown in Plate 2.11. Statistical analysis identified genes whose expression was strongly correlated with these chromosomal classes. Importantly in this study, the AMLs with a normal karyotype were characterized by distinctive upregulation of certain members of the class I homeobox A and B gene families, implying a common underlying genetic lesion. These data reveal novel diagnostic and therapeutic targets and demonstrate the potential of microarray-based dissection of AML. The cluster analysis presented here illustrates the potential of expression profiling to distinguish the major subclasses. An important conclusion of expression profiling studies is that the major cytogenetic events in AML have associated expression signatures. This could form the basis of customized DNA arrays designed to classify leukemia.

Acknowledgement

The authors would like to thank Nigel Carter for his collaboration in the array CGH studies.

Further reading

FISH on metaphase chromosomes and nuclei

Bayani JM, Squire JA. (2002) Applications of SKY in cancer cytogenetics. *Cancer Investigation*, **20**, 373–386.

Dave BJ, Nelson M, Pickering DL *et al.* (2002) Cytogenetic characterization of diffuse large cell lymphoma using multi-color fluorescence in situ hybridization. *Cancer Genetics and Cytogenetics*, **132**,125–132.

Dierlamm J, Stul M, Vranckx H *et al.* (1998) FISH identifies inv(16)(p13q22) masked by translocations in three cases of acute myeloid leukemia. *Genes, Chromosomes and Cancer*, **22**, 87–94.

Fischer K, Scholl C, Salat J *et al.* (1996) Design and validation of DNA probe sets for a comprehensive interphase cytogenetic analysis of acute myeloid leukemia. *Blood*, **88**, 3962–3971.

Grimwade D, Gorman P, Duprez E *et al.* (1997) Characterization of cryptic rearrangements and variant translocations in acute promyelocytic leukemia. *Blood*, **90**, 4876–4885.

Hagemeijer A, de Klein A, Wijsman J *et al.* (1998) Development of an interphase fluorescent in situ hybridization (FISH) test to detect t(8;21) in AML patients. *Leukemia*, **12**, 96–101.

Jabber Al-Obaidi MS, Martineau M, Bennett CF *et al.* (2002) ETV6/AML1 fusion by FISH in adult acute lymphoblastic leukemia. *Leukemia*, **16**, 669–74.

Johnson PW, Leek J, Swinbank K *et al.* (1997) The use of fluorescent in situ hybridization for detection of the t(2;5)(p23;q35) translocation in anaplastic large-cell lymphoma. *Annals of Oncology*, **8** Suppl. 2, 65–69.

Kasprzyk A, Secker-Walker LM. (1997) Increased sensitivity of minimal residual disease detection by interphase FISH in acute lymphoblastic leukemia with hyperdiploidy. *Leukemia*, **11**, 429–435.

Kearney L, Bower M, Gibbons B *et al.* (1992) Chromosome 11q23 translocations in both infant and adult acute leukemias are detected by in situ hybridisation with a yeast artificial chromosome. *Blood*, **80**, 1659–1665.

Ketterling RP, Wyatt WA, VanWier SA *et al.* (2002) Primary myelodysplastic syndrome with normal cytogenetics: utility of 'FISH panel testing' and M-FISH. *Leukemia Research*, **26**, 235–240.

Lu XY, Harris CP, Cooley L *et al.* (2002) The utility of spectral karyotyping in the cytogenetic analysis of newly diagnosed pediatric acute lymphoblastic leukemia. *Leukemia*, **16**, 2222–2227.

Mancini M, Nanni M, Cedrone M *et al.* (1995) Combined cytogenetic, FISH and molecular analysis in acute promyelocytic leukaemia at diagnosis and in complete remission. *British Journal of Haematology*, **91**, 878–884.

Mathew P, Sanger WG, Weisenburger DD *et al.* (1997) Detection of the t(2;5)(p23;q35) and NPM–ALK fusion in non-Hodgkin's lymphoma by two-color fluorescence in situ hybridization. *Blood*, **89**, 1678–1685.

Monteil M, Callanan M, Dascalescu C *et al.* (1996) Molecular diagnosis of t(11;14) in mantle cell lymphoma using two-colour interphase fluorescence in situ hybridization. *British Journal of Haematology*, **93**, 656–660.

Mrozek K, Heinonen K, Theil KS *et al.* (2002) Spectral karyotyping in patients with acute myeloid leukemia and a complex karyotype shows hidden aberrations, including recurrent overrepresentation of 21q, 11q, and 22q. *Genes, Chromosomes & Cancer*, **34**, 137–153.

Nanjangud G, Rao PH, Hegde A *et al.* (2002) Spectral karyotyping identifies new rearrangements, translocations, and clinical associations in diffuse large B-cell lymphoma. *Blood*, **99**, 2554–2561.

Paternoster SF, Brockman SR, McClure RF *et al.* (2002) A new method to extract nuclei from paraffin-embedded tissue to study lymphomas using interphase fluorescence in situ hybridization. *American Journal of Pathology*, **160**, 1967–1972.

Schrock E, du Manoir S, Veldman T *et al.* (1996) Multicolor spectral karyotyping of human chromosomes. *Science*, **273**, 494–497.

Siebert R, Matthiesen P, Harder S *et al.* (1998) Application of interphase fluorescence in situ hybridization for the detection of the Burkitt translocation t(8;14)(q24;q32) in B-cell lymphomas. *Blood*, **91**, 984–990.

Sinclair PB, Green AR, Grace C *et al.* (1997) Improved sensitivity of BCR–ABL detection: a triple-probe three-color fluorescence in situ hybridization system. *Blood*, **90**, 1395–1402.

Speicher MR, Gwyn Ballard S, Ward DC. (1996) Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nature Genetics*, **12**, 368–375.

Suijkerbuijk RF, Matthopoulos D, Kearney L *et al.* (1992) Fluorescent in situ identification of human marker chromosomes using flow sorting and Alu element-mediated PCR. *Genomics*, **13**, 355–362.

Takashima T, Itoh M, Ueda Y *et al.* (1997) Detection of 14q32.33 translocation and t(11;14) in interphase nuclei of chronic B-cell leukemia/lymphomas by in situ hybridization. *International Journal of Cancer*, **72**, 31–38.

Telenius H, Pelmear AH, Tunnacliffe A *et al.* (1992) Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. *Genes, Chromosomes and Cancer*, **4**, 257–263.

Ueda Y, Nishida K, Miki T *et al.* (1997) Interphase detection of BCL6/IgH fusion gene in non-Hodgkin lymphoma by fluorescence in situ hybridization. *Cancer Genetics and Cytogenetics*, **99**, 102–107.

Van Limbergen H, Poppe B, Michaux L *et al.* (2002) Identification of cytogenetic subclasses and recurring chromosomal aberrations in AML and MDS with complex karyotypes using M-FISH. *Genes, Chromosomes & Cancer*, **33**, 60–72.

Veldman T, Vignon C, Schrock E *et al.* (1997) Hidden chromosome abnormalities in haematological malignancies detected by multicolour spectral karyotyping. *Nature Genetics*, **15**, 406–410.

von Bergh A, Emanuel B, van Zelderen-Bhola S *et al.* (2000) A DNA probe combination for improved detection of MLL/11q23 breakpoints by double-color interphase-FISH in acute leukemias. *Genes, Chromosomes & Cancer*, **28**, 14–22.

Comparative genomic hybridization and array CGH

Avet-Loiseau H, Andree-Ashley LE, Moore D *et al.* (1997) Molecular cytogenetic abnormalities in multiple myeloma and plasma cell leukemia measured using comparative genomic hybridization. *Genes, Chromosomes & Cancer*, **19**, 124–133.

Avet-Loiseau H, Vigier M, Moreau A *et al.* (1997) Comparative genomic hybridization detects genomic abnormalities in 80% of follicular lymphomas. *British Journal of Haematology*, **97**, 119–122.

Barth TFE, Dohner H, Werner CA *et al.* (1998) Characteristic pattern of chromosomal gains and losses in primary large B-cell lymphomas of the gastrointestinal tract. *Blood*, **91**, 4321–4330.

Bentz M, Plesch A, Stilgenbauer S *et al.* (1998) Minimal sizes of deletions detected by comparative genomic hybridization. *Genes, Chromosomes & Cancer*, **21**, 172–175.

Cai WW, Mao JH, Chow CW *et al.* (2002) Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays. *Nature Biotechnology*, **20**, 393–396.

Carter NP, Fiegler H, Piper J. (2002) Comparative analysis of comparative genomic hybridization microarray technologies: report of a workshop sponsored by the Wellcome Trust. *Cytometry*, **49**, 43–48.

Cigudosa JC, Rao PH, Calasanz MJ *et al.* (1998) Characterization of nonrandom chromosomal gains and losses in multiple myeloma by comparative genomic hybridization. *Blood*, **91**, 3007–3010.

El-Rifai W, Elonen E, Larramendy M *et al.* (1997) Chromosomal breakpoints and changes in DNA copy number in refractory acute myeloid leukemia. *Leukemia*, **11**, 958–963.

Forozan F, Karhu R, Kononen J *et al.* (1997) Genome screening by comparative genomic hybridization. *Trends in Genetics*, **13**, 405–409.

Haas O, Henn T, Romanakis K *et al.* (1998) Comparative genomic hybridization as part of a new diagnostic strategy in childhood hyperdiploid acute lymphoblastic leukemia. *Leukemia*, **12**, 474–481.

Houldsworth J, Mathew S, Rao PH *et al.* (1996) REL proto-oncogene is frequently amplified in extranodal diffuse large cell lymphoma. *Blood*, **87**, 25–29.

Joos S, Otano-Joos MI, Ziegler S *et al.* (1996) Primary mediastinal (thymic) B-cell lymphoma is characterized by gains of chromosomal material including 9p and amplification of the REL gene. *Blood*, **87**, 1571–1578.

Kallioniemi A, Kallioniemi OP, Sudar D *et al.* (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*, **258**, 818–821.

Karhu R, Knuutila S, Kallioniemi OP *et al.* (1997) Frequent loss of the 11q14–24 region in chronic lymphocytic leukemia: a study by comparative genomic hybridization. Tampere CLL Group. *Genes, Chromosomes and Cancer*, **19**, 286–290.

Karhu R, Siitonen S, Tanner M *et al.* (1997) Genetic aberrations in pediatric acute lymphoblastic leukemia by comparative genomic hybridization. *Cancer Genetics and Cytogenetics*, **95**, 123–129.

Monni O, Oinonen R, Elonen E *et al.* (1998) Gain of 3q and deletion of 11q22 are frequent aberrations in mantle cell lymphoma. *Genes, Chromosomes & Cancer*, **21**, 298–307.

Paszek-Vigier M, Talmant P, Mechinaud F *et al.* (1997) Comparative genomic hybridization is a powerful tool, complementary to cytogenetics, to identify chromosomal abnormalities in childhood acute lymphoblastic leukaemia. *British Journal of Haematology*, **99**, 589–596.

Rao PH, Houldsworth J, Dyomina K *et al.* (1998) Chromosomal and gene amplification in diffuse large B-cell lymphoma. *Blood*, **92**, 234–240.

Solinas-Toldo S, Lampel S, Stilgenbauer S *et al.* (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes, Chromosomes & Cancer*, **20**, 399–407.

Wessendorf S, Schwaenen C, Kohlhammer H *et al.* (2003) Hidden gene amplifications in aggressive B-cell non-Hodgkin lymphomas detected by microarray-based comparative genomic hybridization. *Oncogene*, **22**, 1425–1429.

Expression profiling

Armstrong SA, Staunton JE, Silverman LB *et al.* (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics*, **30**, 41–47.

Debernardi S, Lillington DM, Chaplin T *et al.* (2003) Genome-wide analysis of acute myeloid leukemia with normal karyotype reveal a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. *Genes, Chromosomes & Cancer*, **37**, 149–158.

Golub TR. (2001) Genomic approaches to the pathogenesis of hematologic malignancy. *Current Opinion in Hematology*, **8**, 252–261.

Golub TR. (2001) Genome-wide views of cancer. *New England Journal of Medicine*, **344**, 601–602.

Golub TR, Slonim DK, Tamayo P *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531–537.

Schoch C, Kohlmann A, Schnittger S *et al.* (2002) Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, **99**,10008–10013.

Shipp MA, Ross KN, Tamayo P *et al.* (2002) Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nature Medicine*, **8**, 68–74.

Tamayo P, Slonim D, Mesirov J *et al.* (1999) Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 2907–2912.

Yeang CH, Ramaswamy S, Tamayo P *et al.* (2001) Molecular classification of multiple tumor types. *Bioinformatics*, **17**, S316–S322.

Yeoh EJ, Ross ME, Shurtleff SA *et al.* (2002) Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*, **1**, 133–143.

Chapter 3 Stem cells

Eyal C Attar & David T Scadden

Introduction

The generation of sufficient numbers of blood cells to maintain homeostasis requires sustained production of mature cells. This process, called 'hematopoiesis', yields approximately 10^{10} blood cells daily, with capability for dramatic increases in the number and subsets of cells in response to physiological stress. Hematopoiesis is therefore a highly dynamic process dependent upon numerous modulating factors. Its prodigious production capability derives from the sustained presence of a cell type which is generally quiescent, but the descendants of which proliferate vigorously. This cell is the hematopoietic stem cell (HSC).

Stem cell definitions and distinctions

Stem cells derive their name from their ability to produce daughter cells of different types. Stem cells are defined by a combination of the traits of self-maintenance and the ability to produce multiple, varied offspring. Putting this in more biological terms, stem cells have the unique and defining characteristics of *self-renewal* and of *differentiation into multiple cell types*. Thus, with each cell division there is an inherent asymmetry in stem cells that is generally not found with other cell types.

While their name implies that stem cells have specific intrinsic characteristics, there are multiple different types of stem cells, each defined by their production ability. *Totipotent* stem cells are capable of generating any type of cell of the body, including those of the extra-embryonic membranes, such as the placenta (Figure 3.1). *Pluripotent* stem cells may give rise to any type of cell found in the body except those of the extra-embryonic membranes. They can produce ectoderm, mesoderm or endoderm cells. Pluripotent stem cells include embryonic stem cells, isolated from the inner cell mass of the blastocyst, embryonic germ cells, isolated from embryonic gonad precursors, and embryonic carcinoma cells, isolated from teratocarcinomas. Pluripotent stem cells may be maintained indefinitely in culture under specialized conditions that prevent differentiation. In particular, embryonic stem cells have been used to generate 'knockout' mice, animals harboring targeted gene disruptions via homologous recombination that permit the *in vivo* study of individual gene function. Lastly, *multipotent* stem cells, such as the HSCs of the bone marrow, are capable of giving rise to multiple mature cell types, but only those of a particular tissue, such as blood. Multipotent stem cells are found in adults, perhaps in all tissue, and function to replace dead or damaged tissue. Such stem cells are commonly referred to as 'adult' stem cells.

Hematopoietic stem cell concepts and their origin

The cellular compartment model

The short-lived nature of most blood cells was first deduced in the 1960s using thymidine labeling of reinfused blood. These studies demonstrated that the maintenance of normal numbers of blood cells in the adult required a process with the capacity to briskly generate large numbers of mature cells along multiple blood lineages. The early history of HSC research was largely shaped by cellular biology and animal transplantation experiments. It was advanced by experiments in the early 1960s demonstrating that injection of marrow cells could generate large hematopoietic colonies in the spleen of irradiated mice. Such colonies were the clonal progeny of single initiating cells, termed 'colony-forming units–spleen' (CFU-S), and contained hematopoietic populations of multiple lineages. CFU-S were further transplantable, demonstrating the self-renewing nature of CFU-S. HSCs are a minor component of marrow cells, able both to generate large numbers of progeny differentiated along multiple lines and to renew themselves.

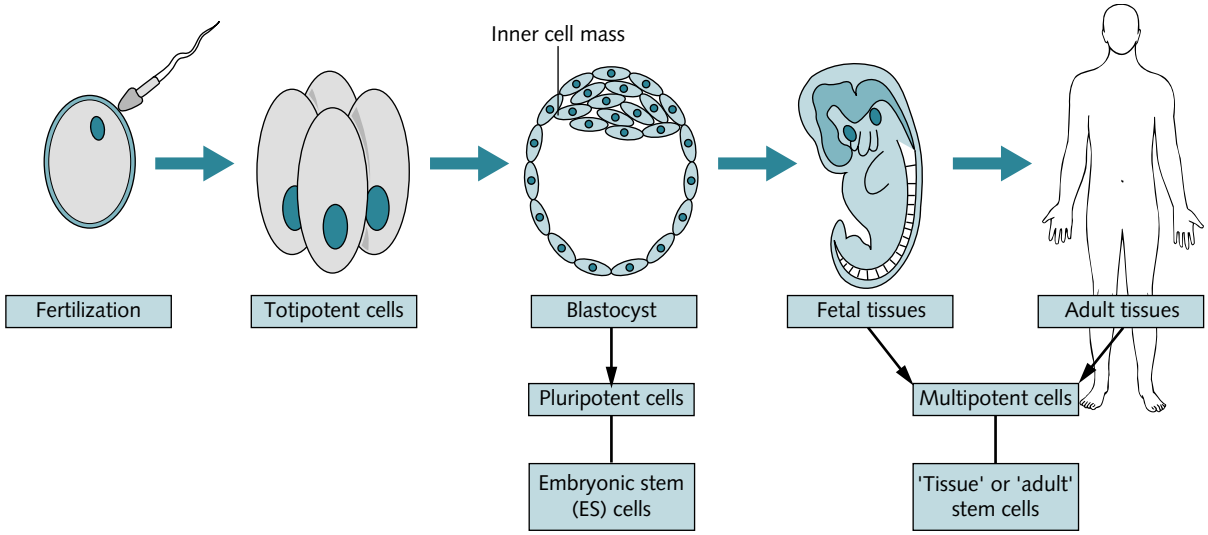


Fig. 3.1 Sources and types of stem cells. Adapted with gratitude from the National Institute of Health Stem Cell Infomation web site

The field was further advanced by the use of *in vitro* cell culture techniques; in particular, solid-state cultures of marrow and spleen cells furthered understanding of the colony-forming capacity of individual hematopoietic cells. The original technique demonstrated clonal colonies of granulocytes and/or macrophages, termed '*in vitro* colony-forming cells' (CFCs), which are now considered lineage-committed progenitor cells. These cells could be separated from whole marrow cells and from CFU-S, were more numerous than CFU-S and could be detected in splenic colonies as the progeny of CFU-S. These observations gave rise to the concept of the *three-compartment model of hematopoiesis*, the compartments being stem cells, progenitor cells, and dividing mature cells in increasing numbers; each compartment consists of the amplified progeny of cells in the preceding compartment.

Subsequent analyses have added further complexity to the compartment model of hematopoiesis. The term 'CFU-S' describes at least two groups of precursor cells. One group, arising from committed progenitors with little capacity for self-renewal, gives rise to colonies that peak in size by day 8, while a second, arising from a more primitive cell that is capable of self-renewal, yields colonies that peak in size at day 12. To further highlight the complexity of the hematopoietic hierarchy, a rarer population of hematopoietic cells provides longer-term repopulation of an irradiated host than CFU-S. These long-term repopulating cells have the capacity for sustained self-renewal and were considered the true adult stem cells. The presence of stromal cells in the cultures is important for the long-term culture of CFU-S and repopulating cells. Cells capable of long-term survival in culture on stroma were termed 'long-term culture-initiating cells' (LTC-ICs) and 'cobblestone area-forming cells'. These multipotential cell types were considered more primitive than lineage-committed progenitor cells but more mature than long-term repopulating cells.

Thus, a more complex version of the compartmental model has emerged. This provides a model with two populations of stem cells, the most immature group consisting of long-term repopulating cells and a more mature group of short-term repopulating cells. An intermediate group consisting of pre-progenitor cells (blast colony-forming cells) follows, leading to a larger population of lineage-committed progenitor cells. This large group of committed progenitors is stratified on the basis of the number of progeny they are able to generate. The immediate progeny of progenitor cells, cluster-forming cells, have less proliferative capacity. Subsequent progenitors (CFCs) have the capacity to give rise to colonies of clonal origin in semisolid media containing fully mature cells, permitting their analysis. A more mature set of precursor cells constitutes the bulk of bone marrow cells and has unique, identifiable features by light microscopy. Rapid division of precursor cells culminates in the production of mature cells. Although hematopoiesis proceeds according to this orderly scheme (Figure 3.2), special consideration must be given to the development of T and B lymphocytes. These cells are generated in the thymus and bone marrow, respectively, by a similar hierarchical process. Mature T and B lymphocytes enter peripheral lymphoid organs, where they encounter relevant antigens, leading to the production of new cells from reactivated mature cells. This process amplifies the *de novo* bone marrow formation of T and B lymphocytes. In addition, some members of this type of cell, memory T or B lymphocytes, are capable of sustained self-renewal. Their inability to produce multiple different types of daughter cells distinguishes them from stem cells.

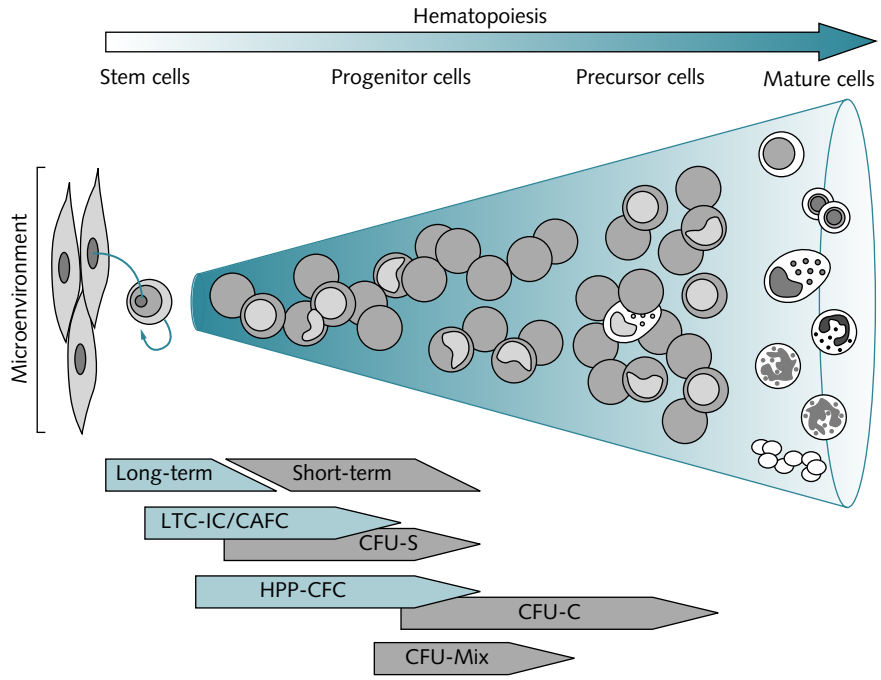


Fig. 3.2 Schematic view of hematopoiesis. Modified from Figure 12.1 in *Hematology: Basic Principles and Practice*, 3rd Edition, Ronald Hoffman, ed., 2000, with permission from Elsevier

In summary, the compartment model has given rise to terms that are generally applied to cells of hematopoietic origin. *Stem cells* are those that are multipotent and self-renewing. *Progenitor cells* have limited ability to self-renew and are likely to be unipotential or of very limited multipotential. *Precursor cells* are restricted to a single lineage, such as neutrophil precursors, and are the immediate precursors of the *mature cells* found in the blood. The mature cells are generally short-lived and preprogrammed to be highly responsive to cytokines, while the stem cells are long-lived, cytokine-resistant and generally quiescent.

Models of lineage commitment

Several theories have emerged to describe the manner by which HSCs undergo lineage commitment and differentiate. Some studies support a *deterministic theory* whereby the stem cell compartment encompasses a series of closely related cells maturing in a stepwise process. Other studies suggest that hematopoiesis is a random, *stochastic* process. The stochastic theory is based on *in vitro* observations that multilineage colonies develop variable combinations of lineages and that such lineage choices occur independently of external influences.

Similar controversy exists regarding the role of cytokines in cell lineage determination. An *instructive* model suggests that cytokine signaling forces the commitment of primitive cells along a particular lineage. Ectopic expression of the granulocyte macrophage colony-stimulating factor (GM-CSF) receptor in a common lymphoid progenitor (CLP) population was capable of converting the cells from a lymphoid to a myeloid lineage. The influence of the GM-CSF receptor was sufficiently dominant to change the entire differentiation program of cells, but only the CLP stage of development. A *permissive* model postulates that decisions about cell fate occur independently of extracellular signals. This model suggests that cytokines serve only to allow certain lineages to survive and proliferate. Evidence supporting this model is provided by the ectopic expression of growth receptors in progenitor cells. Expression of the erythropoietin receptor in a macrophage progenitor results in macrophage colony formation, whereas macrophage colony-stimulating receptor (M-CSF) in an erythroid progenitor results in erythroid rather than macrophage colony formation. Replacing the thrombopoietin receptor (c-mpl) with a chimeric receptor consisting of the extracellular domain of c-mpl with the cytoplasmic domain of the granulocyte colony-stimulating factor receptor (G-CSFR) results in normal platelet counts in homozygous knock-in mice. Therefore, the instructive and permissive models may both be correct, but at different stages of hematopoietic differentiation. Cells at earlier points in the differentiation cascade may be more plastic and susceptible to fate-altering stimuli, while more committed cells may be irreversibly determined, with only proliferation, cell death or the rate of differentiation susceptible to influence by external signals.

Stem cell plasticity and transdifferentiation

'Plasticity' refers to the concept that HSC development is not limited to hematopoietic cells but may also include cells of other tissue types. Studies have suggested that bone marrow-derived cells may develop into neural cells, skeletal muscle, cardiac muscle and hepatic cells in addition to epithelia of the gut, skin, lung and kidney. The possibility that HSCs have undergone 'transdifferentiation' serves as one explanation for these phenomena. However, it has been shown that hematopoietic cells may fuse with somatic cells and this may account for the finding ascribed to transdifferentiation and the issue of HSC plasticity remains controversial. It is possible that contaminating stem cells other than HSCs provide the other tissue types. Indeed, it is well established that mesenchymal stem cells reside in the bone marrow and have a broad range of capability, producing cartilage, bone, muscle or adipose cells upon proper stimulation.

Studies in which single cells were transferred have produced compelling evidence that cells of multiple tissue types do emerge. These events are very infrequent and are highly dependent on the method of cell selection before transplantation, and there are data indicating that a truly pluripotent stem cell may indeed exist in low abundance within the adult. Whether either fusion or transdifferentiation can ever be induced to occur at sufficient frequency to yield a therapeutic benefit is unclear, but is the subject of intense study.

Molecular regulation of hematopoiesis

The molecular nature of stem cell regulatory pathways has been determined using a variety of genetic approaches, including genetic loss-of-function and gain-of-function studies. These have provided several important concepts regarding the molecular control of hematopoiesis. First, some genes have binary functions and are either on or off in various biological states, while other genes function in a continuum and have different effects at different levels. Secondly, while perturbations in single genes may have dramatic cellular effects, cell cycle and lineage effects result from the combinatorial interplay of multiple genes and require coordinated expression of genes with both stimulatory and inhibitory functions. Finally, signal integration often depends on the assembly of large signaling complexes and the spatial proximity of molecules to facilitate interaction is therefore important.

Cell-intrinsic regulators of hematopoiesis

Cell cycle control

The quiescent nature of HSCs is supported by their low level of staining with DNA and RNA nucleic acid dyes, which is consistent with low metabolic activity. Various studies have sought to determine the cell-intrinsic regulators of hematopoiesis involved in HSC cycle control.

Single-cell reverse transcriptase–PCR has been used to profile pertinent transcription factors and other molecules in HSCs induced to differentiate along various lineages by the application of cytokines. This technique has demonstrated the presence of elevated levels of cyclin-dependent kinase inhibitors (CDKIs), suggesting that CDKIs present in HSCs function to exert a dominant inhibitory tone on HSC cell cycling. The bone marrow of mice deficient in CDKI $p21^{cip1/waf1}$ have increased numbers of HSCs and increased HSC cell cycling, suggesting that p21 functions as a dominant negative regulator of HSC proliferation, a function also noted for other CDKIs, such as $p27^{kip}$, a negative regulator of hematopoietic progenitor cells.

Self-renewal, commitment, and lineage determination

Many molecules suspected of having importance in stem cell biology have been identified in other systems, such as leukemia, and have encoded transcription factors. Thus, a major focus of molecular research in stem cell regulatory components has involved transcription factors.

Experimental results involving transcription factors have demonstrated cell-intrinsic roles in both global and lineage-specific hematopoietic development. Loss-of-function studies involving the transcription factors c-Myb, AML1 (CBF2), SCL (tal-1), LMO2 (Rbtn2), GATA-2 and TEL/ETV6 have demonstrated global effects on all hematopoietic lineages. Stem cells in animals deficient in these molecules fail to establish definitive hematopoiesis. To test the role of these genes in established hematopoiesis, a method of altering gene expression in the adult animal is required. A molecular technique to address this involves generating conditional knockouts. In these systems, transgenic animals are generated by swapping the wild-type gene of interest with a gene flanked at both ends with lox-p sites, target sites for Cre-recombinase. Such animals can then be mated with transgenic animals expressing the Cre-recombinase driven by different gene promotors. The Cre-recombinase can then be used to specifically excise the gene of interest in a global-, tissue- or developmental-specific manner, depending on the promoter driving the expression of the Cre gene. This approach is technically somewhat limited by the absence of stem cell-specific promoters thus far. However, this approach has aided the identification of critical roles for genes such as Notch-1, which are required for T-cell lineage induction, as Notch-1-deficient mice die during embryogenesis because of a requirement for the protein in other tissues. An interferon-inducible promotor (Mx-Cre) can also be used to turn on Cre-recombinase at specific times by injecting animals with nucleotides, a means of inducing endogenous interferon. This approach has been useful in defining a very different role for SCL in maintaining hematopoiesis in the

adult than in establishing it in the developing fetus. This gene product is absolutely required for establishing HSCs. Unexpectedly, there is not a requirement for SCL once the stem cell pool is present in the adult. Rather, SCL is required only for erythroid and megakaryocytic homeostasis. Therefore, transcription factor regulation of the stem cell compartment is highly dependent on the stage of development of the organism. Lineage-specific effects of transcription factors may also be stage-dependent.

Loss-of-function studies have also proved useful in identifying lineage-specific transcription factors. Mice genetically deficient in the transcription factor Ikaros lack T and B lymphocytes and natural killer cells, but maintain erythropoiesis and myelopoiesis. Mice lacking the ets-family transcription factor PU.1 demonstrate embryonic lethality. However, mutant embryos produce normal numbers of megakaryocytes and erythroid progenitors but have impaired erythroblast maturation and defective generation of progenitors for B and T lymphocytes, monocytes and granulocytes. While the outcome of such genetic lesions can be assessed, it remains unclear whether such lesions result in failure to establish a commitment program or the execution of an established program.

Gain-of-function studies have been used similarly to assess the roles of various global and lineage-specific transcription factors. Enforced expression of the HoxB-4 homeobox gene in HSCs confers heightened capacity for *in vivo* stem cell function. Similarly, ectopic expression of HoxB4 in embryonic stem cells combined with *in vitro* culture on stroma induces a switch to the definitive hematopoiesis phenotype that is transplantable into adult recipients. Mice deficient in the Pax-5 transcription factor suffer from severe impairment of the B-lymphoid lineage. This phenotype may be rescued by reintroduction of wild-type Pax-5.

In alternative model systems, lineage reprogramming may be achieved by ectopic expression of transcription factors. Introduction of the erythrocytic lineage transcription factor GATA-1 reprograms avian myeloblast cells down eosinophilic and thromboblastic lineages. Introduction of the dominant negative retinoic acid receptor-alpha (RARα) into murine stem cells permits the establishment of permanent cell lines that grow in response to stem cell factor and maintain the ability to differentiate along myeloid, erythroid and B-lineage lines. The points in the hematopoietic cascade at which specific transcription factors play a role are illustrated in Figure 3.3.

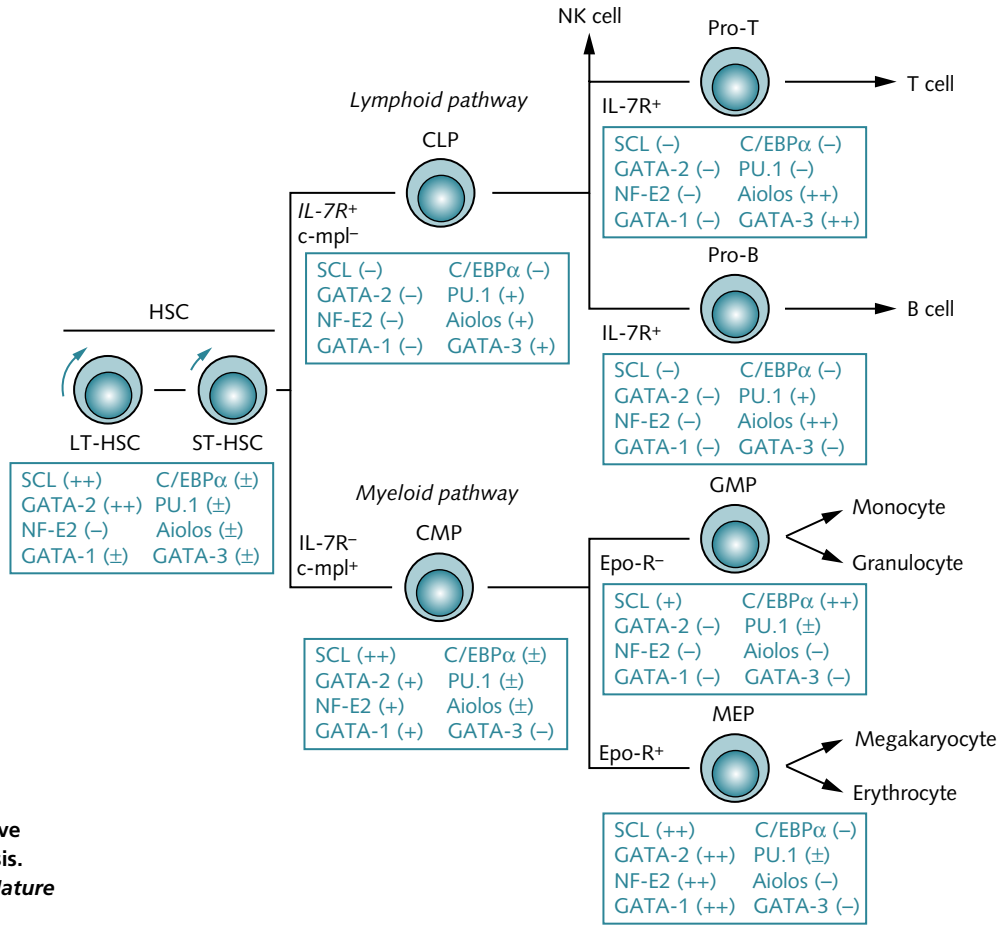


Fig. 3.3 Transcription factors active at various stages of hematopoiesis. Redrawn with permission from *Nature* (2000), 404, p.196

Cell-extrinsic regulators

Ultimately, hematopoietic stem and progenitor cell decisions are regulated by the coordinated action of transcription factors as modified by extracellular signals. Extracellular signals in the form of hematopoietic growth factors are mediated via cell surface hematopoietic growth factor receptors. Hematopoietic growth factors exert specific effects when acting alone and may have different effects when combined with other cytokines. There are at least six receptor superfamilies, and most growth factors are members of the type I cytokine receptor family. The effects of various cytokines during myelopoiesis are illustrated in Figure 3.4.

Type I cytokine receptors

Type I receptors do not possess intrinsic kinase activity but lead to phosphorylation of cellular substrates by serving as docking sites for adapter molecules with kinase activity. Examples of receptors in this family include leukemia inhibitory factor (LIF), interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-13, IL-18, GM-CSF, G-CSF, erythropoietin, prolactin, growth hormone, ciliary neurotrophic factor and c-mpl. These receptors share several features, including enhanced binding and/or signal transduction when expressed as hetero- or homodimers, four cysteine residues and fibronectin type III domains in the extracellular domain, WSXWS ligand-binding

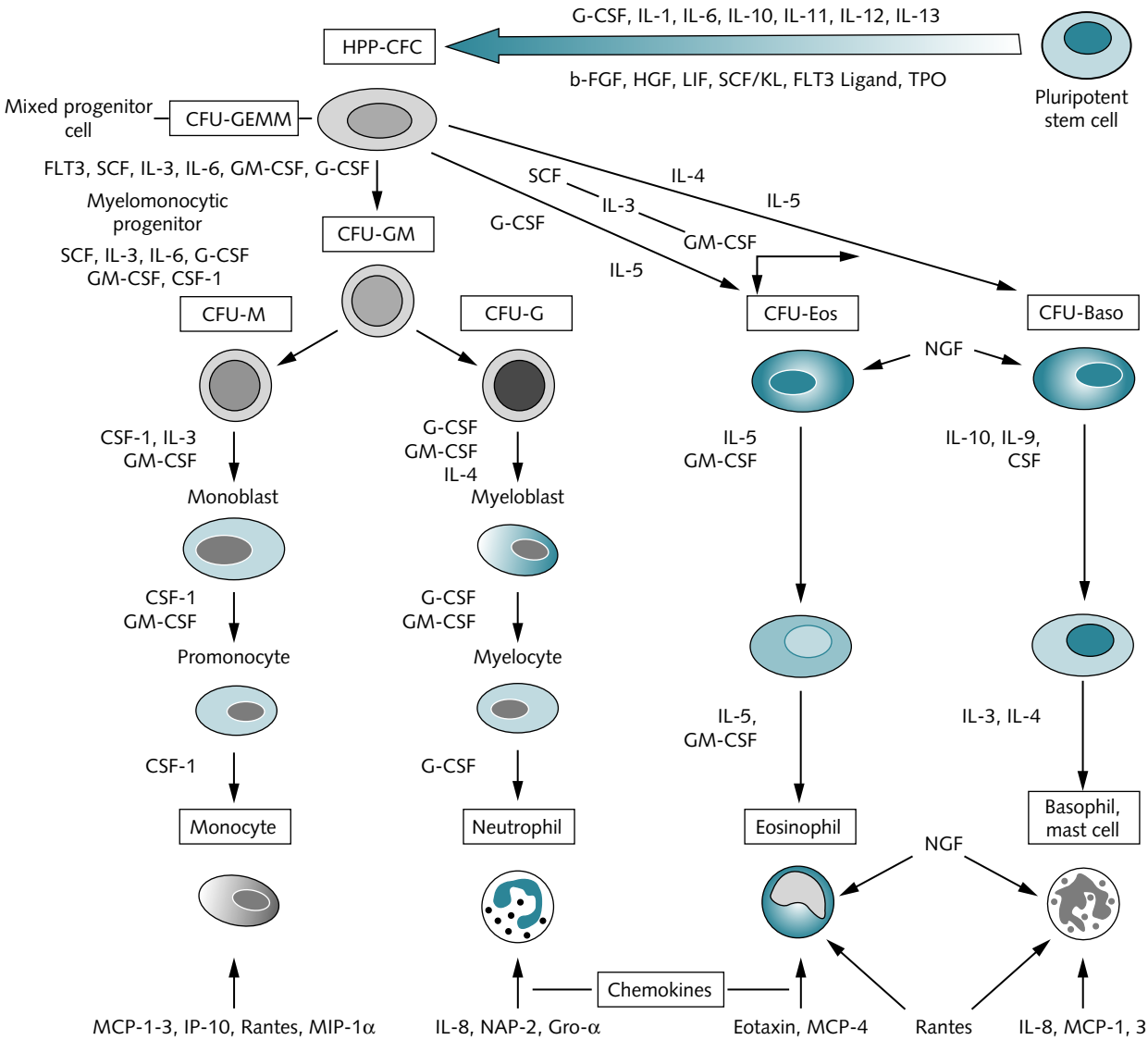


Fig. 3.4 Cytokines active at various stages of hematopoiesis. Modified from Figure 16.3 in *Hematology: Basic Principles and Practice*, 3rd Edition, Ronald Hoffman, ed., 2000, with permission from Elsevier

sequence in the extracellular cytokine receptor domains, and lack of a known catalytic domain in the cytoplasmic portion. Another shared feature of receptors in this family is the ability to transduce signals that prevent programmed cell death (apoptosis).

Several exceptions to this family with intrinsic kinase activity are the hematopoietic growth factor receptors platelet-derived growth factor receptor (PDGFR), flt-3 receptor and c-fms, which are the ligands for steel factor (SF), Flt-ligand (FL) and M-CSF, respectively (Table 3.1).

Type II cytokine receptors

This class includes the receptors for tissue factor, IL-10 and interferon γ (IFN-γ). This family contains a type III fibronectin domain in the extracellular domain, like the type I family.

Protein serine-threonine kinase receptors

This family includes the 30 members of the transforming growth factor-β (TGF-β) superfamily, which bind to their receptors as homodimers. Members of this family include the three TGF-β receptors: type I (TbRI, 53 kDa), type II (TbRII, 75 kDa) and type III (TbRIII, 200 kDa). Members of this family have a profound inhibitory effect on the growth and differentiation of hematopoietic cells and on auxiliary hematopoietic cells. Binding of TFG-β requires TbRII. After binding, signal transduction occurs via activation of serine–threonine kinase cytoplasmic domains of the receptor chains, which results in the phosphorylation of Smad molecules on serines. Phosphorylated Smad complexes translocate to the nucleus, where they induce or repress gene transcription. TGF-β is the best characterized negative regulator of hematopoiesis. It inhibits mitosis by inducing cell cycle inhibitors such as $p21^{cip1/waf1}$, $p27^{kip1}$ and $p16^{INK4a}$, inhibiting the cyclin-dependent kinases Cdk4 and Cdk6, and inducing phosphorylation of the retinoblastoma protein. The TGF-β receptor family and its downstream mediators act as braking factors for a number of cell types and are frequently inactivated by somatic mutation in a number of cancers.

Chemokine receptors

This family comprises seven transmembrane-spanning G-protein-coupled receptors that influence both cell cycle and cellular movement, or chemotaxis. These receptors are divided into three families, α or CXC, β or CC, and γ or C, on the basis of variability in cysteine residues. The best characterized is CXCR4, which mediates homing and engraftment of HSCs in bone marrow and is critical to hematopoietic development. IL-8 and Mip-1α act as inhibitors of progenitor cell proliferation. Members of this receptor family have also been implicated in cancer metastasis and the entry of HIV-1 into cells.

Tumor necrosis factor receptor family

Members of the tumor necrosis factor receptor (TNFR) family have varied effects, some having the ability to induce programmed cell death and others stimulating mesenchymal cells to secrete hematopoietic growth factors. These receptors contain Cys-rich extracellular domains and 80-amino acid cytoplasmic 'death domains', which are required for transducing the apoptotic signal and inducing NF-κB activation. Members of this family include TNFR1, TNFR2, fas, CD40, nerve growth factor (NGF) receptor, CD27, CD30 and OX40, each with at least one distinct biological effect.

Components of the hematopoietic microenvironmental niche

While soluble factors influence stem cell fate, these factors are seen by the cell in the context of cell–cell contact among heterologous cell types and cell–matrix contact, which make up the three-dimensional setting of the bone marrow. What actually constitutes the critical microenvironment for hematopoiesis is surprisingly poorly defined. The ability of primitive cells to mature *in vitro* in complex stromal cultures suggests that at least some elements of the regulatory milieu of the bone marrow can be recapitulated *ex vivo*. Studies based solely on *ex vivo* systems are suspect, however, as no fully satisfactory recreation of stem cell expansion or self-renewal has been defined. Recognizing this limitation, it has been determined that mesodermal cells of multiple types are needed to enable hematopoietic support. These include adipocytes, fibroblastic cells and endothelium. Recently, *in vivo* studies have indicated that the osteoblast may perform a key regulatory role in stem cell self-renewal, and the activation of this cell can affect the number of stem cells.

Trafficking of primitive hematopoietic cells

The migratory behavior characteristic of primitive hematopoietic cells is an area of intense research because of its relationship to bone marrow transplantation. Trafficking of HSCs can be divided into the components of homing, retention and engraftment. 'Homing' describes the tendency of cells to arrive at a particular environment, while 'retention' is their ability to remain in such an environment after arrival. Lastly, 'engraftment' reflects the ability of cells to divide and form functional progeny in a given microenvironment. Much has

Table 3.1 Factors affecting hematopoietic control.

Growth factor	Growth factor receptor	Produced by	Bioactivity	Deficient states
Erythropoiesis				
EPO (erythropoietin)	EPO-R	Adult kidney Liver during development	Stimulates clonal growth of CFU-E and BFU-E subsets Suppresses erythroid progenitor cell apoptosis Induces bone marrow release of reticulocytes Induces erythroid globin synthesis	Anemia
SF (steel factor), kit ligand, mast cell growth factor	c-kit (CD117)	Fibroblasts Endothelial cells Bone marrow stroma	Promotes proliferation and differentiation of pre-CFC cells Acts synergistically with IL-3, GM-CSF and TPO to support growth of CFU-GEMM, BFU-E, and CFU-Mk Expansion of committed progenitor cells *in vivo* Stimulates mast cell hyperplasia, degranulation, and IgE-dependent mediator release	Anemia Mast cell deficiency
IGF-1 (insulin-like growth factor, somatomedin C)	IGF-1R	Liver	Induces DNA synthesis and has anti-apoptotic effects in erythroid progenitors Simulates erythroid colony growth in the absence of EPO at high doses	Growth retardation, neurologic defects, homozygous deficient lethal
Granulopoiesis				
G-CSF (granulocyte colony-stimulating factor)	G-CSFR	Monocytes, macrophages, endothelial cells, fibroblasts	Stimulates growth of progenitors committed to neutrophil differentiation Activates neutrophil phagocytosis Stimulates quiescent HPCs to enter G_1-S Stimulates mobilization of HSCs and HPCs from bone marrow to periphery	Neutropenia, failure to develop neutrophilic leukocytosis in response to infection
GM-CSF (granulocyte-macrophage colony-stimulating factor)	GM-CSFR	Mast cells, T lymphocytes, endothelial cells, fibroblasts, thymic epithelial cells	Stimulates multilineage hematopoietic progenitor cells Stimulates BFU-E and granulocyte, macrophage, and eosinophil colony growth	Susceptibility to infections caused by obligate intracellular organisms
M-CSF (macrophage colony-stimulating factor)	c-fms	Monocytes, macrophages, fibroblasts, epithelial cells, vascular endothelium, osteoblasts	Induces monocyte/macrophage growth and differentiation and activation	Macrophage and osteoclast deficiency, hematopoietic failure
Thrombopoietin	c-mpl	Bone marrow stroma, spleen, renal tubule, liver, muscle, brain	Stimulates *in vitro* growth of CFU-Mk, megakaryocytes and platelets Stimulates clonal growth of individual $CD34^{+}38^{-}$ cells Synergizes with SF, IL-3 and FL Primes response to platelet activators ADP, epinephrine and thrombin but no effect on aggregation	Thrombocytopenia
IL-5	IL-5R	T lymphocytes	Stimulates eosinophil production and activation Activates cytotoxic T cells Induces immunoglobulin secretion	Inability to mount eosinophilic response
IL-11	IL-11R	Fibroblasts, bone marrow stroma	Acts synergistically with IL-3 or SF to stimulate the clonal growth of erythroid (BFU-E and CFU-E) and primitive megakaryocytic (BFU-Mk) progenitors Shortens duration of G_0 of HPCs Quickens hematopoietic recovery after chemotherapy and radiation	No hematological defect

Continued

Table 3.1 (Continued.)

Growth factor	Growth factor receptor	Produced by	Bioactivity	Deficient states
Lymphopoiesis				
IL-7	IL-7R	Bone marrow stroma, spleen, thymus	Induces clonal growth of pre-B cells Induces growth of pre-T cells	B- and T-cell lymphopenia
IL-2	IL-2R	T lymphocytes	Induces proliferation and activation of T cells, B cells and NK cells	Fatal immunoproliferative disorder, loss of self-tolerance
IL-15	IL-15R	Monocytes, macrophages, epithelial cells, skeletal muscle cells, bone marrow and thymic stroma	Induces proliferation and activation of T cells, B cells and NK cells	
IL-4	IL-4R	T lymphocytes	Induces proliferation of activated B cells Inhibits IL-2-stimulated proliferation of B cells Induces T cell proliferation	Defective T helper cell responses
IL-10			Inhibits monocyte/macrophage dependent synthesis of Th1- and Th2-derived cytokines	
Early-acting factors				
IL-3	IL-3R	T lymphocytes, mast cells	Stimulates multilineage colony growth and growth of primitive cell lines with multilineage potential Stimulates BFU-E proliferation	No hematopoietic defect in steady state, deficient delayed-type hypersensitivity
FLT3-ligand (FL)	FLT-3R, flk2	Most tissues, including spleen, lung, stromal cells, peripheral blood mononuclear cells	Weak colony-stimulating activity alone but synergizes with IL-3, GM-CSF, SF, IL-11, IL-6, G-CSF, IL-7, and others Augments retroviral transduction of HSCs when added to cytokine cocktails Mobilizes HSCs to periphery weakly alone but adds greatly to G-CSF	Reduction in pro-B cells, pre-B cells, B-cell colony-forming potential, reduced repopulating capacity of stem cells
IL-9 (T-cell growth factor)	IL-9R	T lymphocytes	Stimulates growth of BFU-E when combined with EPO Stimulates clonal growth of fetal CFU-Mix and CFU-GM	
IL-6	IL-6R	Macrophages, endothelial cells, fibroblasts, T lymphocytes	Synergistic with IL-3 for CFU-GEMM colony growth Synergistic with IL-4 in inducing T cell proliferation and colony growth Synergistic with M-CSF in macrophage colony growth Synergistic with GM-CSF in granulocyte colony growth Co-induces differentiation of B cells	Reduced HSC and progenitor cell survival, reduced T cell numbers, reduced proliferation and maturation of erythroid and myeloid cells

BFU-E, burst-forming unit—erythroid; CFU-mix, colony forming unit—mix; CFU-Mk, colony forming unit—megakaryocyte; CFU-GM, colony forming unit—granulocyte/macrophage; CFU-GEMM, colony forming unit granulocyte, erythroid, monocyte, megakaryocyte.

been learned about trafficking from the ontogeny of mouse and human HSCs.

Hematopoietic ontogeny

In both humans and mice, hematopoiesis occurs sequentially in distinct anatomical locations during development. These shifts in location are accompanied by changes in the functional status of the stem cells and reflect the changing needs of the developing organism. These are relevant for adult hematopoiesis since they offer insight into how the blood production process can be located in different places with distinct regulation.

There are essentially four sites of blood cell formation recognized in mammalian development, and these are best defined in the mouse. At about embryonic day 7.5 (E7.5), blood and endothelial progenitors emerge in the extra-embryonic yolk sac blood islands. The yolk sac supports the generation of primitive hematopoietic cells, which are primarily composed of nucleated erythrocytes. More sustained or definitive hematopoiesis may derive from the yolk sac, but this remains controversial. However, the aorta–gonadal–mesonephros (AGM) region has been clearly identified as the first site of definitive hematopoiesis in both the mouse (E8.5) and the human. It is not clear if the yolk sac seeds the AGM region or if the hematopoietic cells arise there *de novo*, but by E10 in the mouse the fetal liver assumes the primary role of cell production. By E14 in the mouse and the second trimester of human gestation, the bone marrow becomes populated with HSCs and it takes over blood cell production, along with the spleen and thymus. The spleen remains a more active hematopoietic organ in the mouse than in the human.

The transition in the location of hematopoiesis is roughly associated with changes in HSC function. Primitive hematopoiesis and definitive hematopoiesis in the AGM is dominated by the production of red blood cells and stem cells. Because the organism and its vascular supply become more complex when the fetal liver becomes a hematopoietic site, platelet production is added to the robust production of erythrocytes and stem cells. In keeping with the shifting needs of the organism, by late gestation a full spectrum of innate and adaptive immune system cells is added to the production repertoire. Stem cell proliferation begins to decrease and eventually reaches a state of relative quiescence shortly after gestation.

Homing and engraftment of HSCs following infusion

Despite the use of HSC transplantation for over three decades, the exact mechanisms whereby bone marrow cells home to the bone marrow are not fully understood. Other than lectins, no adhesion receptors have been identified that are exclusively present on HSCs. Furthermore, no adhesion ligands, other than hemonectin, have been identified that are exclusively present in the bone marrow microenvironment.

When first infused, HSCs lodge in the microvasculature of the lung and liver; they then colonize the bone marrow, first passing through marrow sinusoids, migrating through the extracellular space of the bone marrow, and ultimately land in the stem cell niches. Passage through endothelial barriers at first requires tethering, through endothelium-expressed addressins that bind hematopoietic cell selectins, and this is followed by firm attachment mediated by integrins.

Selectins are receptors expressed on hematopoietic cells (L- and P-selectins) and endothelium (E- and P-selectins). They have long extracellular domains containing an amino-terminal Ca^{2+}-binding domain, an epidermal growth factor domain, and a series of consensus repeats similar to those present in complement regulatory molecules. Ligands for selectins are sialylated fucosylglucoconjugates present on endothelium, termed 'addressins'. L-selectin is present on $CD34^+$ hematopoietic progenitors while L-selectin and P-selectin are present on more mature myeloid and lymphoid cells. Tethering by selectins allows integrin-mediated adhesion to the endothelium. Integrins, a family of glycoproteins composed of α and β chains responsible for cell–extracellular matrix and cell–cell adhesion, provide not only firm attachment but also allow migration of hematopoietic cells through the endothelium and bone marrow extracellular space. The functional state of integrins is only loosely tied to their expression level and depends on ligand affinity modulation regulated by the β subunit in response to cytokines and other stimuli.

The process of migration depends on the establishment of adhesion at the leading edge of the cell and simultaneous release at the trailing edge. The rate of migration depends on dynamic changes in the strength of the cell–ligand interactions, which is dictated by the number of receptors and their affinity state and the strength of the adhesion receptor–cytoskeleton interactions. Cell–ligand interaction strength may also be modulated by cytokines. Thus, successful engraftment relies not only on the presence of several different adhesion receptors but also on their functional state and ability to facilitate both migration and adhesion.

Egress of HSCs from bone marrow under physiological conditions

The majority of primitive HSCs are resident within the bone marrow space under steady-state physiological conditions. However, a population of $CD34^+$ cells capable of forming CFCs and LTC-ICs and capable of long-term repopulation may be found circulating in the peripheral blood and these may increase after physiological stressors such as exercise, stress and infection. Recent studies have suggested that a

relatively large number of bone marrow-derived stem cells circulates during the course of a day and that these cells periodically transit back into an engraftable niche to establish hematopoiesis. Defining the processes involved is important in guiding new approaches to peripheral blood stem cell mobilization for transplantation.

Examining mice in which specific adhesion molecules have been deleted has revealed several key molecular determinants of stem cell localization in the bone marrow. Among these, the chemokine receptor CXCR4 has perhaps the most striking phenotype. In the absence of this receptor, stem cells fail to traffic from the fetal liver to the bone marrow. Partly because of these studies, others have defined that CXCR4 is relevant for the engraftment of transplanted stem cells and that the modulation of CXCR4 signaling can affect adult stem cell localization in the bone marrow versus peripheral blood. As described, the integrin and selectin families are also important molecular participants in stem cell location. For example, HSCs from animals that are heterozygous-deficient for β_1 integrin cannot compete with wild-type cells for the colonization of hematopoietic organs. Pre-incubation of HSCs with α_4 integrin antibodies prior to transplantation results in decreased bone marrow and increased peripheral recovery of cells, while the continued presence of α_4 antibodies prevents engraftment. Evidence for selectin involvement has been demonstrated in animals deficient for single selectins or combinations of selectins. Endothelial P-selectin mediates leukocyte rolling in the absence of inflammation, while L-, P- and E-selectins contribute to leukocyte rolling in the setting of inflammation. L-selectin is important in lymphocyte homing. Transplantation studies performed in animals deficient in P- and E-selectins demonstrate severely decreased engraftment due to impaired homing, an effect that is further compromised by blocking vascular cell adhesion molecule-1 (VCAM-1).

Mature hematopoietic cells are thought to migrate from the marrow to the blood by similar mechanisms, though these are not well defined. One purported mechanism is a shift in expression from molecules thought to interact with stromal proteins to those that interact with endothelium. For example, myeloid progenitors express functional $\alpha_4\beta_1$ and $\alpha_5\beta_1$ integrins that act to ensure that these progenitors are retained in the bone marrow through interactions with VCAM and fibronectin. Mature neutrophils, in contrast, express β_2 integrins that permit interaction with ligands, such as intercellular adhesion molecule (ICAM), expressed by endothelial cells. Mature neutrophils also express β_1 integrins that permit interaction with collagen and laminin present in basal membranes, perhaps regulating a progressive shift in cell affinities for specific microenvironmental determinants that ultimately results in cell egress into the blood. Mobilization of murine HSCs induced by cyclophosphamide or granulocyte colony-stimulating factor (Cy/G-CSF) is accompanied by changes in integrin expression levels and functional changes in homing, thus linking cellular localization with adhesion molecule receptor expression.

Manipulating hematopoietic stem cells for clinical use

Mobilization of HSCs

Mobilization of HSCs in response to chemotherapy or cytokines was first documented in the 1970s and 1980s. This process may be induced by a variety of molecules, including cytokines such as G-CSF, GM-CSF, IL-7, IL-3, IL-12, stem cell factor and flt-3 ligand; and chemokines such as IL-8, Mip-1α, Gro-β and SDF-1. The one that is most often used clinically is G-CSF, which may be combined with chemotherapeutic agents for added benefit. This mobilizing capability has resulted in a dramatic change in the manner by which HSCs are harvested for transplantation. Up to 25% of candidates for autologous transplantation are unable to mobilize sufficient cells to enable the procedure to be safely performed. The study of mobilization and its counterpart, engraftment, has implications of great significance for patient care. The ability of G-CSF to mobilize bone marrow HSCs has several apparent mechanisms. The first is reported to be the activation of neutrophils, causing the release of neutrophil elastases capable of cleaving CXCR4 on HSCs, thus reducing HSC–bone marrow interaction. Other receptors that undergo cleavage are VCAM-1 and c-kit. A second mechanism of G-CSF-induced mobilization is via CD26, an extracellular dipeptidase present on primitive HSCs that is able to cleave SDF-1 to an inactive form. Other proposed options for improving mobilization include co-administration of G-CSF and kit ligand, antibodies directed against VLA-4, and infusion of IL-8.

Isolating stem cells for manipulation

Characteristics of HSCs used for isolation

Physical Early attempts to isolate HSCs were based on cell size and density. In order to clarify whether the heterogeneity of CFU-S was due to differences in the input cells used, velocity sedimentation was performed to separate cells by size, demonstrating that smaller cells were more likely to produce secondary CFU-S than larger cells. HSCs are similar in size to mature lymphocytes and, when flow cytometry is performed, overlap the lymphocyte region on plots of forward and side scatter.

Using cell-cycle-active drugs Because HSCs are largely in a quiescent portion of the cell cycle (G_0 or G_1), investigators have used cell-cycle-active drugs to deplete bone marrow populations of cycling cells and thereby enrich for primitive HSCs. Treatment of mice with nitrogen mustard resulted in a 30-fold enrichment in CFU-S. HSCs may be isolated by *in vitro* treatment with 5-fluorouracil, and this remains the most commonly used agent. In addition, HSC populations may be further enriched by first stimulating cells to enter the cell cycle with the early-acting cytokines c-kit ligand and IL-3 before forcing them to metabolic death. This strategy is useful for human cells but not murine cells, probably because of different cycling characteristics. It should be noted that these techniques might result in a decrease in the quality of HSCs obtained.

Markers of primitive HSCs A variety of strategies have been used to identify potential HSC markers. CFU-S in rat bone marrow, fetal liver and neonatal spleen express Thy-1 antigen at high levels, and this was the first important HSC marker discovered. Pluripotent stem cells could be enriched from the bone marrow 150-fold on the basis of combination of size and Thy-1 expression. Negative selection of cells using soybean agglutination resulted in enrichment in the colony-forming unit culture assay (CFU-C). In addition, a FACS-based negative selection strategy involving labeling hematopoietic cells with a cocktail of antibodies directed against mature hematopoietic cell antigens has been developed. These lineage-directed antibodies included B220, against mature B lymphocytes, CD8, against T cells, Mac-1, against macrophages, and Gr-1, against granulocytes. When this negative selection protocol (Lin^{neg}) was combined with a positive selection protocol to enrich for cells that expressed low levels of Thy-1 ($Thy\text{-}1^{low}$), 200-fold enrichment of day-10 CFU-S could be achieved.

Using a magnetic bead selection strategy to enrich for Thy-1-expressing cells followed by a FACS-based strategy to deplete cells expressing lineage markers, murine cells were isolated that were found to express a newly defined stem cell antigen, Sca-1. These $Lin^{neg}Thy\text{-}1^{low}Sca\text{-}1^{pos}$ cells represented 1 in 1000 bone marrow cells and had heightened stem cell activity compared with whole bone marrow in the CFU-S assay. This highly selected cell population produced day-13 CFU-S at 1 colony per 10 cells and day-8 CFU-S at 1 per 100 cells. Also, these cells were 1000- to 2000-fold enriched in their ability to rescue irradiated animals and could give rise to all blood cell lineages. Self-renewal capability was demonstrated by the ability of these cells to rescue lethally irradiated animals upon secondary transplantation. Interestingly, the $Sca\text{-}1^{neg}$ population had similar CFU-S activity but could not produce T cells or confer radioprotection.

More recently, the receptor for stem cell factor, c-kit, was demonstrated to be present on HSCs. Populations expressing this phenotype ($c\text{-}kit^{pos}Thy\text{-}1.1^{low}Lin^{neg}Sca\text{-}1^{pos}$), also known as KTLS, are 2000-fold enriched in HSC activity compared with unfractionated bone marrow. Thus, KTLS has come to be regarded by many as a profile that represents, but is not specific for, HSCs in the mouse. Equivalent markers are not as well defined in the human, though it is apparent that cells expressing kit ligand without lineage markers (including the CD38 antigen) are enriched in stem cells. The antigen CD34 has long been regarded as a marker for a stem cell population, but it is now clear that the vast majority of $CD34^+$ cells are progenitors, and stem cells may or may not express CD34. CD133 is a marker more recently shown to be expressed on primitive human hematopoietic cells. A summary of proposed HSC markers for the mouse and human is presented in Table 3.2.

Supravital stains Since HSCs are inherently quiescent, spend most of their time in inactive portions of the cell cycle and are resistant to toxins, exclusion of dyes has been used as a method of isolation. The DNA dye Hoechst 33342 was first used to separate quiescent cells from the bone marrow. Cells with low-intensity staining were enriched for HPP-CFC and day-12 CFU-S. The red and blue emissions from this dye have been recently used to define a small subset of bone marrow cells known as the side population (SP). SP cells have extremely low fluorescence emission in these channels, resulting from efflux of Hoechst 33342 by multidrug resistance pumps that are highly expressed on HSCs. SP cells constitute approximately 0.1% of the bone marrow and are highly enriched in reconstitution potential.

The mitochondrial dye rhodamine-123 (Rh-123) has also been used to subdivide primitive stem cells. Mitochondria in quiescent cells bind low levels of Rh-123 and FACS can be used to separate $Rh\text{-}123^{low}$ cells. These cells were enriched for day-13 CFU-S and multilineage reconstituting potential.

The combination of supravital stains with fluorescent antibodies against cell surface markers provides the ability to enrich for highly primitive HSCs such that fewer than ten cells are required to reconstitute hematopoiesis.

Table 3.2 Proposed markers of primitive HSCs.

Hematopoietic stem cell surface markers	
Mouse	**Human**
$CD34^{low/-}$	$CD34^+$
$Sca\text{-}1^+$	$CD59^+$
$Thy1^{+/low}$	$Thy1^+$
$CD38^+$	$CD38^{low/-}$
$C\text{-}kit^+$	$C\text{-}kit^{-/low}$
lin^-	lin^-

Methods of isolation of HSCs

FACS While the flow cytometer may be used for analysis of cells, the apparatus may also physically sort cells of desired fluorescence or fluorescence pattern, size and granularity characteristics. Using a magnetic field, these cells may be diverted to a collection tube during analysis and later analyzed using techniques of molecular and cellular biology. Sorting is both expensive and labor-intensive as it requires costly machines, a high degree of expertise, and time to sort samples consisting of single-cell suspensions. Many FACS machines are now available with high-speed sorting. This was once a technique available to only a few laboratories, but many centers are developing 'core' laboratories to provide cell analysis and sorting services for investigators. FACS may be used to isolate HSCs using both positive and negative selection strategies with fluorescence-labeled antibodies directed against primitive hematopoietic cell antigens, as described above.

Magnetic bead columns Large-volume isolation of HSC subsets has been facilitated by the use of magnetic bead columns. Using this system, cells are incubated with antibodies directed against primitive hematopoietic cells. These antibodies are typically coupled to a hapten. A second-step incubation is then performed using a magnetic microbead conjugated to a hapten that is able to bind the first-step hapten. The effect is to label HSCs with a magnetic bead. Cells are then passed through a column mounted adjacent to a magnet. Labeled cells are retained within the column and unbound cells can be washed through. Then, the column is removed from the magnet and the desired cells may be eluted.

Alternatively, negative selection may be performed by capturing only the cells that pass through the column. For example, a sample may be depleted of mature cells by labeling with antibodies directed against mature blood cell antigens (Lin^{pos}). Cells can then be passed over a column in which the mature cells adhere and immature cells pass through and may be isolated.

Systems of these types permit rapid isolation of large numbers of primitive cells of relatively high purity.

Ex vivo expansion

Given the possible clinical applications of HSCs for such uses as bone marrow transplantation, there is increasing interest in strategies that both result in an increase in the quantity of HSCs and the ability to manipulate HSCs *ex vivo*. Thus, *ex vivo* expansion of HSCs represents a highly prioritized goal of clinically oriented HSC research.

The first benefit of expanding HSCs is to provide sufficient cells for transplantation when insufficient numbers exist. For example, cord blood represents a rich source of primitive $CD34^+$ cells that are less immunocompetent and are therefore transplantable across partial HLA disparity barriers. However, the absolute quantity of HSCs within a single cord blood is low and transplantation is followed by periods of aplasia. *Ex vivo* expansion would thereby facilitate cord blood transplantation. Similarly, selective expansion of HSC subsets would permit the extension of tumor-free cells from patients with limited quantities of normal bone marrow due to bone marrow-infiltrating diseases, such as leukemia, for the purpose of autologous transplantation.

The second benefit of *ex vivo* manipulation is that HSCs have a relative growth advantage over other cell types, such as tumor cells. Therefore, *ex vivo* growth provides a purging effect. Furthermore, specific tumor cell purging may be achieved via the application of certain cytokines (IL-2, IFN-γ), antitumor agents such as 5-fluorouracil or cyclophosphamide, tumor-specific antibodies combined with complement-mediated lysis, and oncogene-specific tyrosine kinase inhibitors, in addition to other targeted therapies, such as antisense oligonucleotides, prior to use of the graft.

The third benefit is the support of gene transfer into HSCs for the purpose of gene therapy. A variety of gene-transfer mechanisms, including retroviral infection, are conveyed during mitosis. Thus, the *ex vivo* stimulation of cells using cytokines results in heightened transfer of exogenous genes to HSCs.

Strategies to expand HSCs *ex vivo* have used cytokine cocktails such as IL-11, flt3-ligand and steel factor, stimulation with the purified WNT-3a glycoprotein, neutralizing antibodies of TGF-β alone or in combination with inhibition of CDKI p27, inhibition of the CDKI p21, and stimulation with Notch ligands. While these efforts have resulted in encouraging laboratory results, none to date has translated into accepted clinical practice. Testing regarding these methods continues with intensity and relies heavily on specific functional analyses.

Functional analysis of HSCs

Functional assays for HSCs do not actually measure the activity of HSCs but instead assess more differentiated progeny, such as progenitor and precursor cells. Whereas *in vitro* assays measure mature populations, *in vivo* assays detect the activity of primitive cells capable of homing and engrafting in the proper microenvironment to produce functional hematopoietic progeny.

In vitro *assays* The CFU-C measures hematopoietic progenitor function and is performed by plating cells in semisolid media containing methylcellulose and one or more cytokines. After 5–14 days, colonies comprising mature cell populations committed to either myeloid or lymphoid lineages may be observed. While most colonies obtained using this

assay are composed of cells of a single lineage, less frequently multipotent progenitors can yield colonies containing multiple lineages. Another type of primitive cell, known as the 'high proliferative potential colony-forming cell' (HPP-CFC), which possesses a high degree of proliferative and multilineage potential, may be detected in this culture system. Formation of HPP-CFC colonies, characterized by size greater than 0.5 mm and multilineage composition, requires the use of multiple cytokines in order to proliferate.

The LTC-IC assay correlates more closely to HSCs. Here, hematopoietic cells are plated on top of stromal cell lines or irradiated primary bone marrow stroma. Primitive HSCs are able to initiate growth and to generate progeny *in vitro* for up to 12 weeks. Progenitor cells and mature myeloid cells are removed weekly to prevent overgrowth. Ultimately, HSCs, characterized by high proliferative and self-renewal capabilities, are able to sustain long-term culture and may be enumerated at the conclusion of the assay.

The cobblestone area-forming cell assay represents a type of LTC-IC that similarly measures the ability of cells to initiate growth and generate progeny *in vitro* for up to 12 weeks. However, the readout is slightly different. Hematopoietic cells are plated at limiting dilution on top of a monolayer consisting of irradiated bone marrow stroma or a stromal cell line. The growth of colonies consisting of at least five small, nonrefractile cells reminiscent of cobblestones, found underneath the stromal layer, are counted. Such cultures are maintained using weekly half-media changes until up to 5 weeks after seeding. In this assay, more primitive cells appear later, and day-35 cobblestone area-forming cells (CAFCs) represent a close correlate of a cell with *in vivo* long-term multilineage repopulating potential. LTC-ICs may be enumerated after day 35 by completely removing the CAFC medium, overlaying methylcellulose and counting the number of colonies produced after 8–10 days.

In vivo *assays* The CFU-S assay, first developed by Till and McCulloch in 1961, is described earlier in this chapter (*see Hematopoietic stem cell concepts and their origin*). Bone marrow or spleen cells are transplanted to irradiated recipients and animals are killed after 8 or 12 days for analysis of spleen colonies, termed 'CFU-S_8' and 'CFU-S_{12}', respectively. Cells that give rise to CFU-S_8 are predominantly unipotential and produce erythroid colonies. CFU-S_{12} colonies consist of several types of myeloid cells, including erythrocytes, megakaryocytes, macrophages and granulocytes. Cells giving rise to CFU-S_{12} represent a more primitive population of multipotent cells than those that result in CFU-S_8.

The long-term repopulation assay (LTRA) is a more accurate measure of HSC activity. Whole collections of hematopoietic cells or fractionate subpopulations are transplanted to lethally irradiated syngeneic mice, typically by tail vein injection. Recipients are screened for ongoing hematopoiesis 8–10 weeks after transplantation. By this time, hematopoiesis is firmly established and donor-derived blood is produced by transplanted HSCs. This assay requires that cells fulfill the two central features of HSCs: multilineage reconstitution, consistent with multipotentiality, and indefinite hematopoiesis, indicative of self-renewal.

Tracking of transplanted cells was originally conducted using radiation-induced chromosomal abnormalities or by retrovirally marking donor cells. However, a major advance in the ability to track transplanted cells has been the development of congenic mice with minor allelic differences in the leukocyte common antigen (Ly5), which is expressed on all nucleated blood cells. The C57/BL6 ('black-6') strain contains the Ly5.2 antigen while the BL6/SJL strain contains a separate allele, Ly5.1. However, these syngeneic strains may be transplanted interchangeably. Both antibodies are available with distinct fluorescent labels. FACS analysis using these antibodies permits measurement of donor-derived reconstitution of the nucleated blood lineages. However, erythrocytes and platelets do not express the Ly5 antigen and cannot be tracked using this technique. Instead, investigators use congenic strains with allelic variants of hemoglobin and glucose phosphate isomerase to track erythroid and platelet engraftment, respectively.

A modification of this assay permits quantitation of HSCs within the graft. Here, HSCs are quantified by transplanting limiting-dilution numbers of bone marrow into lethally irradiated recipients. Each recipient also receives 1×10^5 cells of the host's marrow to ensure survival during the period of pancytopenia immediately after irradiation. At 10–12 weeks, host peripheral blood is assessed to determine whether donor-derived reconstitution has occurred. Donor cells must constitute at least 1% of the peripheral blood to contend that at least one HSC was present in the donor population. Also, both lymphoid and myeloid lineages must demonstrate at least 1% donor derivations. The percentage of reconstituted animals in each group may be plotted against the number of input cells to determine a limiting-dilution estimate of the frequency of HSCs within the donor population. This assay is termed a 'competitive repopulation assay', as transplanted HSCs compete with the host's HSCs that survive irradiation-induced death, in addition to host cells transplanted with the graft. The HSCs detected are termed 'competitive repopulation units' (CRU). The competitive repopulation assay using congenic mouse strains is depicted in Figure 3.5.

Summary

Hematopoietic stem cell investigation has been facilitated by the development of *in vitro* and *in vivo* assays of hematopoiet-

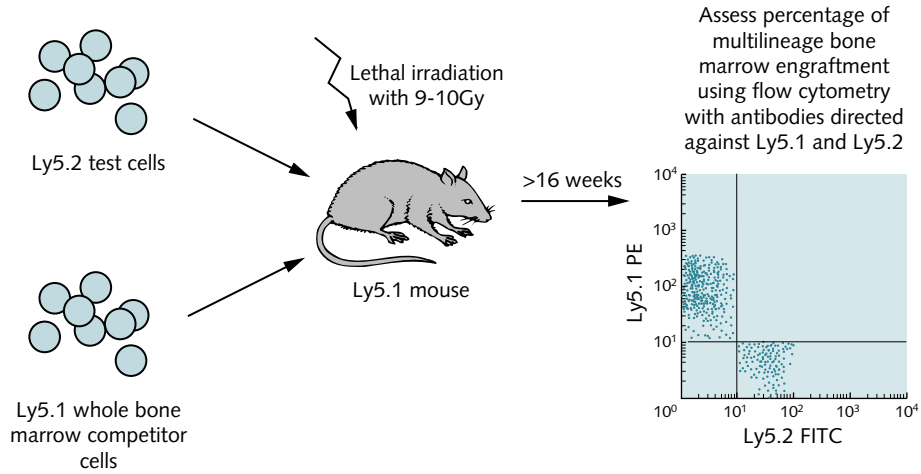


Fig. 3.5 Competitive repopulation assay

ic cell function followed by the identification of molecular cell surface markers that permit the isolation of purified subsets of cells with defined characteristics. Studies in this field have contributed greatly to the understanding of both general stem cell biology and hematopoiesis. Further investigation of cell-intrinsic and cell-extrinsic regulators of hematopoiesis will enable rational manipulation of HSCs and thereby extend the current uses of stem cells in clinical practice.

Further reading

Identification of stem cells and colony assays

Bradley TR, Metcalf D. (1966) The growth of mouse bone marrow cells in vitro. *Australian Journal of Experimental Biology and Medical Science*, **44**, 287–299.

Curry JL, Trentin JJ. (1967) Hemopoietic spleen colony studies. I. Growth and differentiation. *Developmental Biology*, **15**, 395–413.

Ichikawa Y, Pluznik DH, Sachs L. (1966) In vitro control of the development of macrophage and granulocyte colonies. *Proceedings of the National Academy of Sciences of the United States of America*, **56**, 488–495.

Metcalf D. (1984) *The Hematopoietic Colony Stimulating Factors*. Amsterdam: Elsevier.

Till J, McCulloch E. (1961) A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, **14**, 213–222.

Lineage commitment

Morrison SJ, Hemmati HD, Wandycz AM *et al.* (1995) The purification and characterization of fetal liver hematopoietic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 10302–10306.

Nakahata T, Ogawa M. (1982) Clonal origin of murine hemopoietic colonies with apparent restriction to granulocyte-macrophage-megakaryocyte (GMM) differentiation. *Journal of Cellular Physiology*, **111**, 239–246.

Stem cell plasticity

Brazelton TR, Rossi FM, Keshet GI *et al.* (2000) From marrow to brain: expression of neuronal phenotypes in adult mice. *Science*, **290**, 1775–1779.

Camargo FD, Green R, Capetenaki Y *et al.* (2003) Single hematopoietic stem cells generate skeletal muscle through myeloid intermediates. *Nature Medicine*, **9**(12), 1520–1527.

Ferrari G, Cusella-De Angelis G, Coletta M *et al.* (1998) Muscle regeneration by bone marrow-derived myogenic progenitors. *Science*, **279**, 1528–1530.

Gussoni E, Soneoka Y, Strickland CD *et al.* (1999) Dystrophin expression in the mdx mouse restored by stem cell transplantation. *Nature*, **401**, 390–394.

Lagasse E, Connors H, Al-Dhalimy M *et al.* (2000) Purified hematopoietic stem cells can differentiate into hepatocytes in vivo. *Nature Medicine*, **6**, 1229–1234.

Mezey E, Chandross KJ, Harta G *et al.* (2000) Turning blood into brain: cells bearing neuronal antigens generated in vivo from bone marrow. *Science*, **290**, 1779–1782.

Orlic D, Kajstura J, Chimenti S *et al.* (2001) Bone marrow cells regenerate infarcted myocardium. *Nature*, **410**, 701–705.

Wagers AJ, Sherwood RI, Christensen JL *et al.* (2002) Little evidence for developmental plasticity of adult hematopoietic stem cells. *Science*, **297**, 2256–2259.

Wang X, Willenbring H, Akkari Y *et al.* (2003) Cell fusion is the principle source of bone-marrow-derived hepatocytes. *Nature*, **422**, 897–901.

Molecular regulators of hematopoiesis

Cheng T, Rodrigues N, Dombkowski D *et al.* (2000) Stem cell repopulation efficiency but not pool size is governed by p27(kip1). *Nature Medicine*, **6**, 1235–1240.

Cheng T, Rodrigues N, Shen H *et al.* (2000) Hematopoietic stem cell quiescence maintained by p21cip1/waf1. *Science*, **287**, 1804–1808.

Kuhn R, Schwenk F, Aguet M *et al.* (1995) Inducible gene targeting in mice. *Science*, **269**, 1427–1429.

Orkin SH. (1996) Development of the hematopoietic system. *Current Opinion in Genetics and Development*, **6**, 597–602.

Reya T, Duncan AW, Ailles L *et al.* (2003) A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature*, **423**, 409–414.

Shivdasani RA, Orkin SH. (1996) The transcriptional control of hematopoiesis. *Blood*, **87**, 4025–4039.

Shivdasani RA, Mayer EL, Orkin SH. (1995) Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature*, **373**, 432–434.

Homing and engraftment

Cashman JD, Lapidot T, Wang JC *et al.* (1997) Kinetic evidence of the regeneration of multilineage hematopoiesis from primitive cells in normal human bone marrow transplanted into immunodeficient mice. *Blood*, **89**, 4307–4316.

Peled A, Petit I, Kollet O *et al.* (1999) Dependence of human stem cell engraftment and repopulation of NOD/SCID mice on CXCR4. *Science*, **283**, 845–848.

Wright DE, Wagers AJ, Gulati AP *et al.* (2001) Physiological migration of hematopoietic stem and progenitor cells. *Science*, **294**, 1933–1936.

Isolation of stem cells

Baum CM, Weissman IL, Tsukamoto AS *et al.* (1992) Isolation of a candidate human hematopoietic stem-cell population. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 2804–2808.

Goodell MA, Brose K, Paradis G *et al.* (1996) Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *Journal of Experimental Medicine*, **183**(4), 1797–1806.

Spangrude GJ, Heimfeld S, Weissman IL. (1988) Purification and characterization of mouse hematopoietic stem cells. *Science*, **241**, 58–62.

Chapter 4 The genetics of acute myeloid leukemias

D. Gary Gilliland

Introduction

Acute myeloid leukemia (AML) is a complex and heterogeneous disease. However, major strides have been made in our understanding of the molecular pathogenesis of AML based on the cloning of more than 100 disease alleles. These insights in turn have generated strategies for improving treatment outcome and minimizing the toxicity of therapies. Although the genotypic diversity of AML might suggest that it would be challenging to develop molecular targeted therapies for each genotypic variant, there are common themes in the pathways of transformation. For example, many translocations and point mutations in myeloid leukemias target shared signal transduction pathways that confer proliferative and survival advantages on hematopoietic progenitors. Other mutations target hematopoietic transcription factors, and phenotypically result in impaired hematopoietic differentiation. It may therefore be possible to develop therapeutic strategies that target these shared pathways of transformation.

Genetics of AML

There are considerable data showing that AML, like other human cancers, is the consequence of more than one mutation. These data include epidemiological and genotypic data demonstrating that many AMLs have more than one recurring mutation, as either point mutations or chromosomal translocations. In addition, emerging data from animal models of leukemia strongly support the multistep pathogenesis of disease. Rare inherited leukemia syndromes provide strong evidence in support of more than one mutation, as do childhood leukemias in which the expression of a leukemia oncogene can be detected at birth but the leukemia does not develop until later in life.

Analysis of the catalog of mutations that have been cloned in human acute leukemias has suggested that disease alleles can be broadly divided into two categories: those that confer a proliferative and/or survival advantage on hematopoietic progenitors and those that impair hematopoietic differentiation.

Mutations that confer a proliferative and/or survival advantage

Gain-of-function mutations in certain genes confer proliferative and/or survival advantages on hematopoietic progenitors, usually as a consequence of aberrantly activating signal transduction pathways. Examples in myeloid leukemias include activating mutations in RAS family members, in the receptor tyrosine kinases KIT and FLT3 (discussed in more detail below), loss of function of NF-1, and, more recently, gain-of-function mutations in the hematopoietic phosphatase SHP-2. It is of note that although these mutations collectively account for as many as 50% of cases of AML, with rare exceptions only one of these is mutant in any given patient. This epidemiological observation suggests that these mutations can be viewed as a complementation group, and that any one of these is sufficient to confer proliferative and survival advantages on a leukemic cell (Figure 4.1).

The most common of these is the hematopoietic receptor tyrosine kinase FLT3, which is constitutively activated by acquired mutation in approximately 30–35% of AMLs. In 20–25% of cases of AML, there are internal tandem duplications (ITDs) in the juxtamembrane domain of FLT3, ranging in size from several to more than 50 amino acids. In each case, the consequence of the ITDs is constitutive activation of FLT3 tyrosine kinase activity. Although the mechanism of FLT3 activation by ITD mutations is not fully understood, a working hypothesis is that the ITDs result in loss of function of negative autoregulatory modules in the juxtamembrane

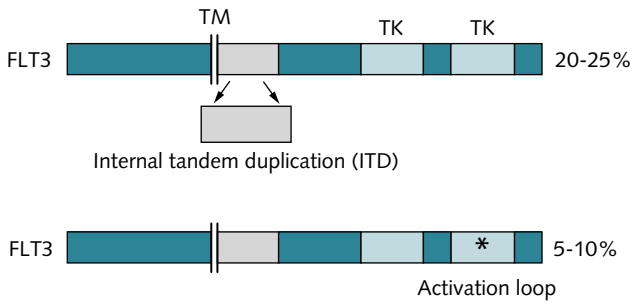


Fig. 4.1 FLT3-ITD and activation loop mutations in AML
Twenty to twenty-five percent of cases of AML have internal tandem duplications (ITD) in the juxtamembrane domain of FLT3. These can range from several to more than 50 amino acids in length, and result in ligand-independent activation of FLT3 tyrosine kinase activity. The mechanism of activation is not clear, but may be the disruption of a negative autoregulatory module in the juxtamembrane domain, with subsequent ligand-independent activation. Point mutations in the 'activation loop' occur in about 5–10% of cases of AML and result in constitutive activation of the FLT3 tyrosine kinase.

region, with subsequent kinase activation. In addition, mutations may also occur in the so-called activation loop of FLT3 in about 5–10% of AMLs. These also result in constitutive kinase activation. Although the structure of FLT3 is not yet available, these mutations, in the context of other tyrosine kinases, result in folding out of the activation loop, providing access of the catalytic site to ATP and substrate. There are rare examples of both ITD and activation loop mutations in the same allele of FLT3, suggesting that the combination of mutations may hyperactivate the kinase and provide added proliferative advantage to cells that harbor both mutations.

FLT3 mutant AML constitutes an important subset of AMLs that has a poor prognosis in most studies of children and adults as an independent prognostic indicator. FLT3 has not been reported as a poor prognostic indicator in adults with AML over the age of 65, which may reflect the overall worse prognosis of this group compared with younger individuals.

As one surrogate of transformation, FLT3-ITD confers interleukin (IL)-3-independent growth on the murine hematopoietic cell line Ba/F3, which is normally dependent on IL-3 for growth and survival. FLT3-ITDs activate several signal transduction pathways in Ba/F3 cells that are known to confer proliferative and/or survival advantage, including the RAS/MAPK, STAT and PI3K/AKT pathways. FLT3-ITD also induces a myeloproliferative disease in primary hematopoietic progenitors in a murine bone marrow transplantation assay. In this assay, mice are treated with 5-fluorouracil to induce transient pancytopenia. Cytopenias induce hematopoietic progenitors into the cell cycle and render them susceptible to transduction with a murine ecotropic retrovirus. Bone marrow is harvested from donor animals, and transduced with retrovirus containing a FLT3-ITD, followed by transplantation into lethally irradiated syngeneic recipient mice by tail vein injection. When this experiment is performed with various FLT3-ITDs, mice develop a myeloproliferative disease and die with a median latency of about 45 days. However, these animals never develop AML, but rather a leukocytosis with normal maturation and differentiation of myeloid lineage cells, and splenomegaly due to extramedullary hematopoiesis. The disease is not transplantable into secondary recipient mice. These data indicate that FLT3-ITDs alone are not sufficient to induce an AML phenotype in primary murine hematopoietic progenitors.

The FLT3-ITD phenotype is similar to that reported in the murine bone marrow transplantation assay for other constitutively activated tyrosine kinases associated with myeloproliferative phenotypes in humans, including BCR-ABL, TEL-PDGFβR, TEL-ABL and TEL-JAK2. Taken together, these data indicate that constitutive activation of tyrosine kinases is sufficient to induce a myeloproliferative phenotype, but not AML.

Mutations associated with AML that affect hematopoietic differentiation

Core binding factor in acute leukemias

In contrast with chromosomal translocations in chronic myeloid leukemias, which almost invariably involve constitutively activated tyrosine kinases, the cloning of recurring chromosomal translocations associated with AML usually identifies fusion genes involving transcription factors or transcriptional co-activators that are important for normal hematopoietic development. Multiple translocations target the core binding factor (CBF) in acute leukemias. Of these, the most extensively studied are the AML1-ETO, CBFβ-SMMHC and TEL-AML1 fusions. CBF is a heterodimeric transcription factor comprising AML1 (also known as RUNX1) and CBFβ subunits. Homozygous loss of function of either AML1 or CBFβ in genetically engineered mice results in a complete lack of definitive hematopoiesis, indicating that both components of CBF are necessary for normal hematopoietic development.

One might therefore predict that an acquired gene rearrangement or mutation that resulted in CBF loss of function might impair hematopoietic differentiation. Indeed, several lines of evidence indicate that the leukemia-associated fusion genes result in loss of CBF function through dominant negative activity mediated by aberrant recruitment of the nuclear co-repressor complex. For example, the use of homologous recombination strategies to express AML1-ETO or CBFβ-SMMHC from their endogenous promoters in mice results in a phenotype nearly identical to that of the AML1 or CBFβ knock-outs, namely a loss of definitive hematopoiesis.

Additional evidence that implicates loss of function of CBF in the pathogenesis of leukemias has come from the analysis of pedigrees with an inherited predisposition to develop leukemia. For example, the familial platelet disorder with propensity to develop acute myeloid leukemia (FPD/AML syndrome) is an autosomal dominant disorder that is caused by haploinsufficiency of the AML1 gene. Furthermore, in sporadic AML there are loss-of-function point mutations involving AML1 in about 3–5% of cases, most of which impair AML1 DNA binding activity. In many AML patients there is loss of function of both alleles of AML1, suggesting that homozygous loss may contribute to disease progression or severity of disease.

Although mutations and gene rearrangements affecting CBF function are clearly important in the pathogenesis of AML, in part through the disruption of normal hematopoietic differentiation programs, it is equally clear that they are not sufficient to cause AML. For example, conditional alleles of AML1-ETO expressed in adult hematopoietic progenitors are not sufficient to cause AML. It is necessary to treat animals with chemical mutagens such as ethyl-nitrosourea to induce AML. However, AML1-ETO expression does confer an immortalization phenotype, in that AML1-ETO-expressing progenitors can be propagated in serial transfer assays *in vitro*. It is not clear whether AML1-ETO expression itself induces a transcriptional program that confers the immortalization phenotype, or whether the phenotype simply reflects a block in differentiation at the level of the hematopoietic stem cell that has self-renewal capacity.

Mutations involving retinoic acid receptor-alpha (RARα)

As with CBF, there are also multiple chromosomal translocations that involve the RARα locus. Each of these is associated with an acute promyelocytic leukemia (APL) characterized by a block in differentiation at the promyelocyte stage of hematopoietic development. The most extensively studied, and the most common, is the PML–RARα fusion associated with t(15;17). PML–RARα expression is associated with a block in differentiation due to aberrant recruitment of the nuclear co-repressor complex, similar to observations in the context of the AML1-ETO, CBFβ-MYH11 and TEL-AML1 fusions. All-*trans*-retinoic acid (ATRA), a ligand for RARα, has proved to be an effective therapy for APL, especially when given in combination with conventional induction chemotherapy with anthracyclines and cytosine arabinoside. The efficacy of ATRA in the treatment of APL appears to be related to the ability of ATRA to bind to the fusion protein, with resultant dissociation of the nuclear co-repressor complex. Promyelocytes are then able to engage normal hematopoietic differentiation programs that ultimately result in apoptotic cell death. The efficacy of agents that induce normal differentiation, such as ATRA, has suggested that inhibitors of histone deacetylase, a key component of the nuclear co-repressor complex, might have therapeutic efficacy not only in APL, but also in other leukemias characterized by aberrant recruitment of the nuclear co-repressor complex, such as AML1-ETO and CBFβ-MHY11.

Expression of PML–RARα and/or its reciprocal is not sufficient to induce AML. There are several lines of evidence that support this assertion. These include transgenic murine models of PML–RARα-induced AML. Expression of PML–RARα has been directed to the promyelocyte compartment using promyelocyte-specific promoters, including the cathepsin G promoter, and the MRP8 promoter. However, although the fusion gene is present in the germline and is expressed during embryonic and adult development, these animals do not develop AML until 3–6 months after birth, and even then with a modest penetrance of only 15–30%, and often with acquisition of secondary cytogenetic abnormalities. Although co-expression of the reciprocal RARα–PML and PML–RARα under the control of the cathepsin G promoter in double transgenic mice increases penetrance to about 60%, double transgenic mice do not have shortened latency of disease. These data indicate that second mutations are necessary in the pathogenesis of APL in this murine model system.

A second line of reasoning that supports a need for more than one mutation in the pathogenesis of APL is derived from genotyping. At least 30% of APL patients harbor activating mutations in FLT3-ITD mutations in addition to t(15;17), which gives rise to the PML–RARα fusion. These mutations are not observed in normal individuals; thus, their concordance in this context indicates that both are required for pathogenesis of APL in at least a subset of patients.

Multistep pathogenesis of AML

These data suggest that at least two mutations are required for the development of AML. Genotypic analysis of known leukemia oncogenes indicates that there are at least two broad complementation groups of mutations. One class of mutations, exemplified by FLT3-ITD or oncogenic RAS mutations, confer a proliferative and/or survival advantage on hematopoietic progenitors, but have minimal effects on differentiation programs in hematopoietic progenitors. These mutations are each relatively frequent in AML, but only very rarely are observed together in the same patient. In contrast, mutations resulting in loss of function of hematopoietic transcription factors result in a block in differentiation at a specific stage in hematopoietic development, as exemplified by the PML–RARα fusion, which is associated with a block in differentiation at the promyelocyte stage. Again, although CBF mutations and PML–RARα mutations occur in a significant proportion of AML patients, they are never observed

together in the same patient, suggesting that they also constitute a complementation group. These mutations may also confer an immortalization phenotype, but are not sufficient to cause acute myeloid leukemia. On the basis of these observations, and genotype data, a model emerges for the pathogenesis of AML in which there are at least two broad classes or complementation groups of mutations. When mutations that confer proliferative and/or survival advantage are expressed alone, they result in a myeloproliferative disease with leukocytosis and normal differentiation. When mutations such as AML1-ETO are expressed alone, they impair differentiation and confer an immortalization phenotype, reminiscent of the behavior of myelodysplastic cells. In this model, co-expression of a mutant that confers a proliferative and/or survival advantage, such as an FLT3-ITD, and a mutation than impairs hematopoietic differentiation, such as PML–RARα, results in AML (Figure 4.2).

To test this hypothesis, bone marrow from transgenic C3H/C57BL6 mice expressing the PML–RARα fusion under the control of the cathepsin G promoter was harvested and transduced with retrovirus containing an FLT3-ITD mutant. In control experiments, PML–RARα transgenic bone marrow transduced with an empty vector control resulted in an APL-like disease in secondary recipient mice, with a latency of approximately 6 months and a penetrance of about 15–30%, in agreement with previous reports. In an additional control experiment, FLT3-ITD retrovirus transduced into the C3H/C57BL6 wild-type background resulted in T-cell lymphomas with a latency of 3–6 months (not shown).

PML–RARα bone marrow transduced with FLT3-ITD resulted in a shortened latency, with 100% penetrance of disease (Figure 4.3). This murine model of cooperativity not only provides experimental evidence for cooperation between these two types of mutations, but also provides a model in which FLT3 inhibitors can be tested alone and in combination with ATRA in the treatment of APL. Indeed, initial data in a similar system using MRP8-PML–RARα transgenic mice and FLT3-ITD indicate that small-molecule inhibitors of FLT3 have at least additive effects with ATRA in treating APL in this model system. Additional experimentation will be

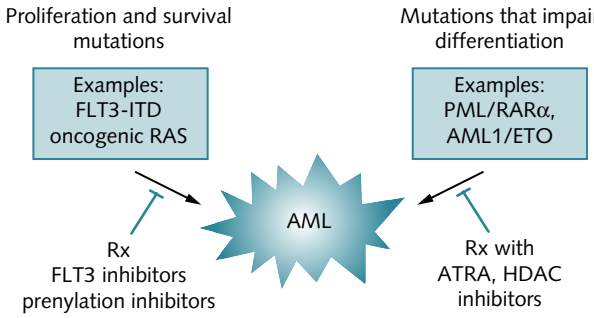


Fig. 4.2 Therapeutic insights from a multistep model of pathogenesis of AML

AML appears to be due to cooperation between at least two broad classes of mutations. The first class, exemplified by activating mutations in FLT3 or oncogenic RAS, confers a proliferative and/or survival advantage on hematopoietic progenitors but does not affect differentiation. A corollary of the hypothesis, supported by murine models of disease, is that expression of FLT3-ITD alone would result in a myeloproliferative phenotype. A second class of mutations, exemplified by loss-of-function mutations in hematopoietic transcription factors such as core binding factor (CBF) and the PML–RARα fusion, results in impaired hematopoietic differentiation and may confer an immortalization phenotype due to the inability to undergo terminal differentiation and apoptosis. These mutations alone are not sufficient to cause AML, but appear to confer a phenotype most similar to myelodysplastic syndrome. Together, these two mutations would result in the acute myeloid leukemia phenotype, characterized by a proliferative and/or survival advantage of hematopoietic progenitors and by impaired hematopoietic differentiation. The hypothesis has important clinical therapeutic implications. For example, targeting the proliferative and survival pathways with FLT3 inhibitors or with inhibitors of RAS, such as farnesyltransferase inhibitors (FTI), might have therapeutic benefit. Alternatively, agents that relieve the block in differentiation in AML, as ATRA does in APL, may have therapeutic benefit.

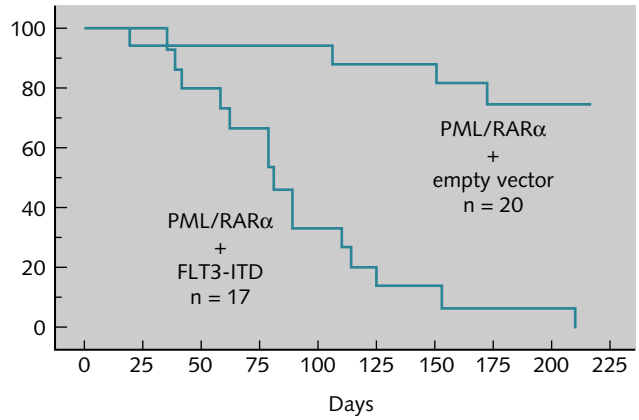


Fig. 4.3 Cooperativity between FLT3-ITD and PML–RARα in the induction of an APL-like disease in a murine model

About 30–40% of human APL patients have both the t(15;17) mutation, which gives rise to the PML–RARα fusion, and the FLT3-ITD mutation. Expression of PML–RARα in the promyelocyte compartment under the control of the cathepsin G promoter results in the development of leukemia with a long latency (3–6 months) and incomplete penetrance (15–30%), strongly indicating the need for a second mutation in disease pathogenesis. Retroviral transduction of FLT3-ITD into cathepsin G-PML–RARα-transgenic bone marrow results in shortened latency and 100% penetrance, indicating cooperativity between FLT3-ITD and PML–RARα in the induction of APL. These data indicate that FLT3-ITD and PML–RARα may both be targets for molecularly targeted therapy with FLT3-specific small-molecule inhibitors and ATRA, respectively. APL patients with FLT3-ITD appear to have a worse prognosis than FLT3-negative patients, suggesting that introduction of FLT3-ITD inhibitors could have therapeutic benefit in this subgroup of patients.

required to assess the transforming properties of oncogenic RAS expressed from its endogenous reporter, and co-expression of oncogenic RAS with potential cooperating mutations. Overall, these data support the multistep model of disease described in Figure 4.2.

Inhibition of FLT3 as a strategy for improving outcome in AML

FLT3 mutations are an independent indicator of a poor prognosis in AML patients under the age of 65 in most studies, and are thus an attractive target for therapeutic intervention. This approach of targeting constitutively activated kinases with selective small-molecule inhibitors has been validated by Druker, Sawyers, Kantarjian and colleagues in demonstrating the efficacy of the ABL kinase inhibitor imatinib (Gleevec) in BCR-ABL-positive chronic myelogenous leukemia.

Similar strategies have been used to identify selective inhibitors of FLT3. We have developed cell-based screens for specific inhibitors of FLT3. Using this approach, in collaboration with Novartis Pharma and Millenium respectively, we have identified several FLT3-selective inhibitors that have suitable properties for use in clinical trials in humans. These include PKC412 from Novartis Pharma and MLN518 (CT53518) from Millenium. Other agents with similar activity include SU11248, SU5614 and SU5416 from SuGen, and CEP-701 developed by Small and colleagues in collaboration with Cephalon. Each of these inhibitors is selective rather than specific. For example, MLN518 is also a potent inhibitor of KIT and platelet-derived growth factor receptor (PDGFR); CEP-701 also inhibits transforming tyrosine kinase protein (TRKA); PKC412 inhibits KIT, PDGFR and protein kinase C; SU11248 also inhibits KIT and PDGFR.

Preclinical activity of FLT3 inhibitors

FLT3 inhibitors induce apoptosis in cell lines harboring activating mutations in FLT3. Each of the inhibitors listed above inhibits Ba/F3 cells transformed with FLT3-ITD, and this effect can be rescued by addition of IL-3. Several of these inhibitors also induce apoptotic cell death in human AML cell lines containing the FLT3-ITD mutation, and in some cases even in AML cell lines overexpressing wild-type FLT3. It is not clear why inhibitors appear to be generally more effective for cells expressing mutant receptors in some studies. These observations indicate that it may be appropriate to test FLT3 inhibitors in AML patients with overexpression of wild-type FLT3 as well as mutant FLT3.

Murine models have also been developed to test FLT3 inhibitors in preclinical analysis, including injection of FLT3-ITD transformed Ba/F3 cells into syngeneic recipient mice and murine bone marrow transplant models of FLT3-ITD-induced disease. In each of the model systems FLT3 inhibitors demonstrate statistically significantly prolonged survival, indicating that these agents are effective *in vivo*, and have appropriate pharmacokinetic properties for inhibition of FLT3 *in vivo*.

Based on these data, Phase I/II trials of FLT3 inhibitors have been initiated by several groups in AML. Most trials have focussed on the treatment of relapsed AML patients who have a mutant FLT3. In some cases, FLT3 inhibitors such as PKC412 and CEP-701 have already been tested in Phase I trials for other disease indications, and are currently in Phase II trials for AML. Other agents, such as MLN518, are currently in Phase I trials. Although it is still early in the evaluation process, preliminary indications are that several of these agents have reasonable safety profiles and have activity in this clinical context. Extensive additional testing will be necessary to determine whether these agents will have a place in the armamentarium used to treat AML, whether they can be used in combination or sequentially with available therapies for AML, and whether they will be efficacious in both FLT3 mutant and wild-type AML.

Summary

We have new insights into potential therapeutic approaches to AML based on a thorough and comprehensive effort to identify disease alleles that are causally implicated in disease pathogenesis. These include strategies that target the block in differentiation, with ATRA treatment in APL as the paradigm for this approach. In addition, inhibitors of proliferative and/or survival mutations such as FLT3 may also prove therapeutically useful. Future directions include the identification of additional proliferative and survival mutations that may be targets for small-molecule inhibitions, and screens for compounds that override the block in differentiation. Eventually, it may be possible to use combinations of molecularly targeted therapies, such as FLT3 inhibitors plus ATRA, in selected clinical contexts to improve outcome and reduce toxicity.

To the extent that these therapies are successful, we can anticipate the development of resistance to single agents, such as FLT3 inhibitors and ATRA. ATRA resistance develops in most APL patients, which is one reason why current therapy always includes a combination of ATRA with conventional induction chemotherapy. Imatinib resistance is well described, particularly in chronic myelogenous leukemia blast crisis patients, and we should anticipate this problem and begin to develop strategies to circumvent or prevent resistance to FLT3 inhibitors. It may be possible to address the problem of resistance to small-molecule kinase inhibitors using alternative inhibitors with different chemical structures.

Further reading

Molecular genetics of leukemia

Dash A, Gilliland DG. (2001) Molecular genetics of acute myeloid leukaemia. *Best Practice in Research in Clinical Haematology*, **14**, 49–64.

Gilliland DG. (2001) Hematologic malignancies. *Current Opinion in Hematology*, **8**, 189–191.

Gilliland DG. (2002) Molecular genetics of human leukemias: new insights into therapy. *Seminars in Hematology*, **39** (4 Supplement 3), 6–11.

Gilliland DG, Tallman MS. (2002) Focus on acute leukemias. *Cancer Cell*, **1**, 417–420.

Speck NA, Gilliland DG. (2002) Core-binding factors in haematopoiesis and leukemia. *Nature Reviews. Cancer*, **2**, 502–513.

AML

Brown D, Kogan S, Lagasse E *et al.* (1997) A PMLRARalpha transgene initiates murine acute promyelocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, **18**, 2551–2556.

Grisolano JL, Wesselschmidt RL, Pelicci PG *et al.* (1997) Altered myeloid development and acute leukemia in transgenic mice expressing PML–RAR alpha under control of cathepsin G regulatory sequences. *Blood*, **89**, 376–387.

He LZ, Tribioli C, Rivi R *et al.* (1997) Acute leukemia with promyelocytic features in PML/RARalpha transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 5302–5307.

Tallman MS, Nabhan C, Feusner JH *et al.* (2002) Acute promyelocytic leukemia: evolving therapeutic strategies. *Blood*, **99**, 759–767.

FLT3 in leukemia

Gilliland DG, Griffin JD. (2002) The roles of FLT3 in hematopoiesis and leukemia. *Blood*, **100**, 1532–1542.

Kelly LM, Kutok JL, Williams IR *et al.* (2002) PML/RARalpha and FLT3-ITD induce an APL-like disease in a mouse model. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 8283–8288.

Kelly LM, Liu Q, Kutok JL *et al.* (2002) FLT3 internal tandem duplication mutations associated with human acute myeloid leukemias induce myeloproliferative disease in a murine bone marrow transplant model. *Blood*, **99**, 310–318.

Sohal J, Phan VT, Chan PV *et al.* (2003) A model of APL with FLT3 mutation is responsive to retinoic acid and a receptor tyrosine kinase inhibitor, SU11657. *Blood*, **101**, 3188–3197.

Stirewalt DL, Radich JP. (2003) The role of FLT3 in haematopoietic malignancies. *Nature Reviews. Cancer*, **3**, 650–665.

Translocation and leukemiagenesis

Armstrong SA, Staunton JE, Silverman LB *et al.* (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics*, **30**, 41–47.

Higuchi M, O'Brien D, Kumaravelu P *et al.* (2002) Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell*, **1**, 63–74.

Michaud J, Wu F, Osato M *et al.* (2002) In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. *Blood*, **99**, 1364–1372.

Osato M, Asou N, Abdalla E *et al.* (1999) Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2alphaB gene associated with myeloblastic leukemias. *Blood*, **93**, 1817–1824.

Schwaller J, Frantsve J, Aster J *et al.* (1998) Transformation of hematopoietic cell lines to growth-factor independence and induction of a fatal myeloid and lymphoproliferative disease in mice by retrovirally transduced TEL/JAK2 fusion gene. *EMBO Journal*, **17**, 5321–5333.

Song W-J, Sullivan MG, Legare RD *et al.* (1999) Haploinsufficiency of CBFA2 (AML1) causes familial thrombocytopenia with propensity to develop acute myelogenous leukemia (FPD/AML). *Nature Genetics*, **23**, 166–175.

Tomasson MH, Sternberg DW, Williams IR *et al.* (2000) Fatal myeloproliferation, induced in mice by TEL/PDGFβR expression, depends on PDGFβR tyrosines 579/581. *Journal of Clinical Investigation*, **105**, 423–432.

Targeted therapy

Kelly LM, Yu JC, Boulton CL *et al.* (2002) CT53518, a novel selective FLT3 antagonist for the treatment of acute myelogenous leukemia. *Cancer Cell*, **1**, 421–432.

Levis M, Tse KF, Smith BD *et al.* (2001) A FLT3 tyrosine kinase inhibitor is selectively cytotoxic to acute myeloid leukemia blasts harboring FLT3 internal tandem duplication mutations. *Blood* **98**, 885–887.

Levis M, Allebach J, Tse KF *et al.* (2002) A FLT3-targeted tyrosine kinase inhibitor is cytotoxic to leukemia cells in vitro and in vivo. *Blood*, **99**, 3885–3891.

Pandolfi PP. (2001) Histone deacetylases and transcriptional therapy with their inhibitors. *Cancer Chemotherapy and Pharmacology*, **48** (Suppl. 1), S17–S19.

Weisberg E, Boulton C, Kelly LM *et al.* (2002) Inhibition of mutant FLT3 receptors in leukemia cells by the small molecule tyrosine kinase inhibitor PKC412. *Cancer Cell*, **1**, 433–443.

Chapter 5 Secondary myelodysplasia/acute myelogenous leukemia—assessment of risk

D Gary Gilliland & John G Gribben

Introduction

A major complication of chemotherapy and radiotherapy for the treatment of cancer is the subsequent development of therapy-related myelodysplastic syndromes and secondary acute myelogenous leukemia (t-MDS/AML). Although devastating in their impact, t-MDS/AML allow the opportunity to study the development of malignancy since many such patients have serial blood and bone marrow samples available from the time of initial therapy to their subsequent diagnosis with leukemia. There is abundant evidence that t-MDS/AML are clonal disorders that are the consequence of acquired somatic mutations and confer a proliferative and/or survival advantage on hematopoietic progenitors. No single mutation or gene rearrangement appears to be sufficient for the development of tMDS/AML. Indeed, the identification of a single gene rearrangement or point mutation may not necessarily be predictive of its subsequent development. Methods for assessing risk are based on the presence of clonal abnormalities in hematopoietic cells, including standard cytogenetics, interphase fluorescence *in situ* hybridization (FISH), analysis for loss of heterozygosity (LOH), PCR for point mutations, and X-inactivation-based clonality assays. Each of these approaches has strengths and weaknesses, and they are discussed in more detail below.

The actuarial risk of developing therapy-related leukemia (t-MDS/AML) varies with the therapy used to treat the cancer. Although some agents are associated with particularly increased risk, in general the more intensive the therapy the higher the risk. The risk of the development of t-MDS/AML after high-dose chemoradiotherapy and autologous stem cell transplantation (ASCT) for lymphoma is substantial, ranging from 3% to as many as 24% of patients. In our own series of patients with non-Hodgkin's lymphoma who have undergone high-dose therapy and ASCT at the Dana-Farber Cancer Institute, development of t-MDS/AML has emerged as the second most common cause of death, after relapse of disease, in these patients. On review of these cases, it became clear that strict criteria were required to make the diagnosis of secondary MDS. A number of patients have relative pancytopenia after ASCT and many patients have dysplastic features and cytogenetic abnormalities, but only 30% of these patients develop secondary MDS. The criteria that we use to define t-MDS are shown in Table 5.1. On the basis of these criteria, some patients initially reported to have developed t-MDS in the original report from this center have now been excluded, and it is of note that none of these patients have progressed to t-MDS or AML. The median time from high-dose therapy to the development of t-MDS was 47 months, with a range from 12 to 129 months after ASCT. The actuarial risk of development of t-MDS in these patients is shown in Figure 5.1. The prognosis of these patients remains dismal and these patients have a median survival after diagnosis of less than 1 year.

Table 5.1 Criteria used to define t-MDS after ASCT at Dana-Farber Cancer Institute.

- Significant marrow dysplasia in at least 2 cell lineages
- Peripheral cytopenia without alternative explanation
- Blast counts in marrow defined by FAB classification

Risk factors for t-MDS/AML after autologous stem cell transplantation

Because t-MDS/AML is frequently a fatal complication, there is a need to better understand the risk factors and to identify individuals at risk prior to ASCT. There are three contributors to the risk of t-MDS/AML in this context that have been

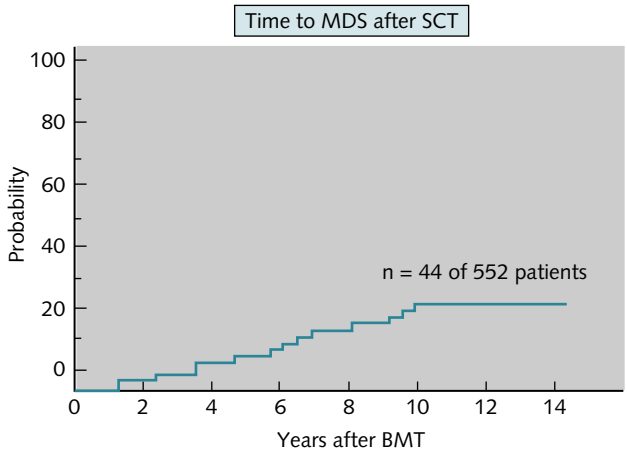


Fig. 5.1 Actuarial probability of development of t-MDS/AML after ASCT

identified: pretransplant therapy, the method of stem cell mobilization, and the transplant conditioning regimens.

The data implicating pretransplant therapy as a risk factor include the demonstration that patients who do not undergo ASCT have a risk of development of t-MDS/AML, although there is a relatively longer latency of development of t-MDS/AML in patients who have undergone ASCT. In patients who have undergone ASCT, there is increased risk with increased cumulative exposure to alkylating agents and with prior external beam irradiation. Specific cytogenetic abnormalities and clonal hematopoiesis have been identified at the time of stem cell harvest in patients who subsequently develop t-MDS/AML.

The method of stem cell collection may influence the risk of developing t-MDS/AML. For example, patients undergoing ASCT using peripheral blood stem cells (PBSC) have a higher risk of developing t-MDS/AML than those receiving bone marrow stem cells. There may be several explanations for this observation, including the possibility that previously only those patients who had inadequate marrow harvests had PBSC collected. It has also been reported that patients primed by receiving etoposide as part of their mobilization regimen had a higher risk of developing t-MDS/AML; this included patients with 11q23 and 21q22 chromosomal abnormalities.

The use of total body irradiation (TBI) in the conditioning regimen is associated with an increased risk of t-MDS/AML after ASCT, although a randomized trial to determine the contribution of TBI to the risk of t-MDS/AML has not been done. Only one study has compared patients who did and did not receive TBI at a single institution, and this did not demonstrate increased incidence in patients who received a TBI containing regimen.

Therapy-related AML is a clonal disorder

There is abundant evidence that t-MDS and AML are clonal disorders. Multiple cytogenetic abnormalities, including deletions (5q, 7q, 20q), numerical abnormalities (trisomy 8, deletion 7) and translocations [11q23, t(3;21), t(15;17) and clonal point mutations of RAS, FLT3 and AML1], have been identified in t-MDS/AML. Population-based analysis of clonality using X-inactivation assays in females has convincingly demonstrated that t-MDS/AML is a clonal disease. Thus, t-MDS/AML is a clonal disease that is the consequence of an acquired somatic mutation that confers a proliferative and/or survival advantage on hematopoietic progenitors.

More than one mutation is necessary to cause AML

It is plausible to determine risk through the analysis of molecular markers of disease. However, no single mutation or gene rearrangement appears to be sufficient for the development of therapy-related AML.

Several lines of evidence support the requirement for second mutations in leukemias associated with mutations of core binding factor (CBF), including analysis of the heritable FPD/AML syndrome (familial platelet disorder with propensity to develop AML), the TEL/AML1 leukemias in syngeneic twins, and murine models of AML1/ETO and CBFβ/MYH11 leukemias. In addition, point mutations that cause loss of function of AML1 have been identified in both inherited and sporadic leukemias. CBF is a heterodimeric transcription factor comprising AML1 (also known as RUNX1) and CBFβ subunits. It is a common target of gene rearrangements as a consequence of chromosomal translocations, giving rise to the AML1/ETO, CBFβMYH11 and TEL/AML1 fusions. FPD/AML syndrome is an autosomal dominant trait characterized by a qualitative and quantitative platelet defect, progressive pancytopenia and dysplasia with age, and progression to AML associated with acquisition of secondary mutations. FPD/AML is caused by loss-of-function mutations in the *AML1* gene, demonstrating that mutations in the AML1 component of CBF are not sufficient to cause leukemia, but require second mutations during the lifetime of affected individuals to cause leukemia. TEL/AML1 leukemias have been studied in syngeneic twins, each of whom harbored the same clone of cells containing the TEL/AML1 gene rearrangement at the time of birth, presumably as a result of intrauterine transmission of a TEL/AML1-positive clone. However, despite the syngeneic host background and the carriage of an identical TEL/AML1 clone, the twins devel-

oped leukemia at widely different ages, indicating the need for additional mutations to cause leukemia.

Murine models of leukemia also provide convincing evidence for 'multiple-hit' pathogenesis of disease. Expression of either AML1/ETO or CBFβ/MYH11 fusion proteins alone in hematopoietic cells is not sufficient to cause leukemia, and chemical-induced mutagenesis must be added to generate a leukemia phenotype. Similar data emerge for PML/RARα-mediated leukemias in transgenic murine models. PML/RARα is expressed in promyelocytes in the germline of transgenic animals under the control of the cathepsin G promoter. However, despite germline expression, animals require 4–6 months to develop leukemia and have karyotypically evident second mutations. Similarly, in MLL/AF9 knock-in mice there is a long latency required for the development of leukemia, and MLL/CBP leukemias in a murine bone marrow transplant model require long latencies, indicative of the need for second mutations.

Furthermore, leukemogenic fusions have been detected using sensitive PCR-based assays in normal individuals. Examples include IgH/BCL2, BCR/ABL, MLL tandem duplication and the TEL/AML1 fusion. The frequency of these rearrangements is much higher in the general population than the risk of developing the respective leukemias. These data indicate that carriage of even a known leukemogenic fusion gene does not provide useful information about the likelihood of progression to leukemia. Indeed, there are currently no data demonstrating that PCR-detectable fusions are a risk factor for the eventual development of leukemia. Collectively, these data indicate that the identification of a single gene rearrangement or point mutation may not necessarily be predictive of the development of therapy-related AML in the post-ASCT setting.

Methods for assessing the risk of therapy-related leukemia

Methods for assessing risk are based on the presence of clonal abnormalities in hematopoietic cells. These methods are shown in Table 5.2 and include standard cytogenetics, interphase FISH, analysis for LOH, PCR for point mutations, and X-inactivation-based clonality assays. Each of these approaches has strengths and weaknesses in this context.

Standard cytogenetics analyzes a limited number of cells that must be capable of mitosis and therefore lacks sensitivity and specificity. Most patients who develop t-MDS/AML after ASCT may have normal cytogenetics at the time of stem cell harvest, whereas some patients who have characteristic cytogenetic abnormalities will not develop t-MDS/AML.

Interphase FISH (Figure 5.2) circumvents some of the frailties of conventional cytogenetics. For example, abnormal clones (5q–, –7, +8, –11) were detectable in pre-ASCT specimens from nine out of 12 patients who developed t-MDS/AML. An advantage of interphase FISH is that hundreds of non-mitotic cells can be analyzed. However, the technique is locus-specific and requires prior selection of markers for analysis, such as 5q–, 7q– and +8. In addition, interphase FISH is

Table 5.2 Methods for assessing the risk of t-MDS/AML before and after ASCT.

1. Standard cytogenetics
2. Interphase FISH
3. Loss of heterozygosity (LOH)
4. PCR for point mutations
5. X-inactivation-based clonality

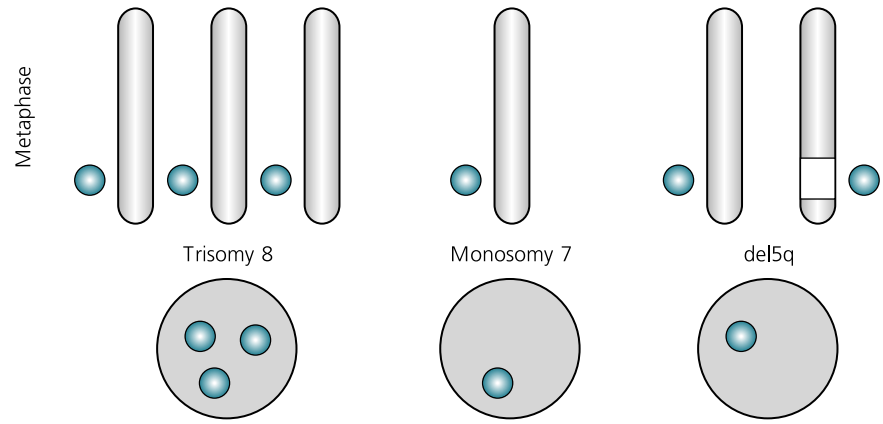


Fig. 5.2 Applications of interphase FISH to detect trisomy, monosomy or chromosomal deletions

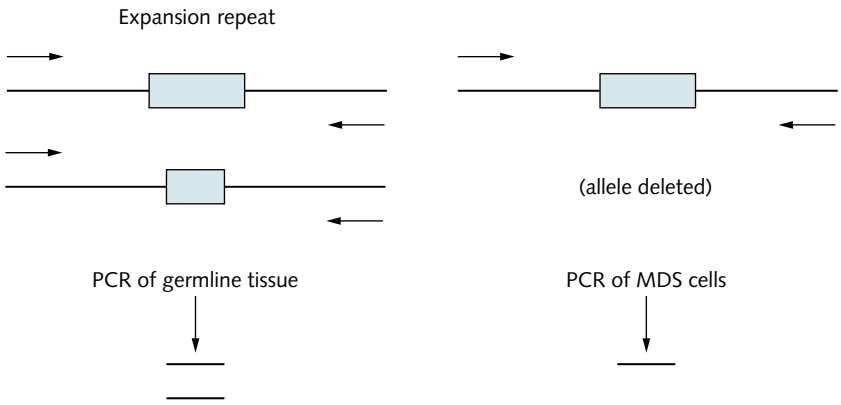


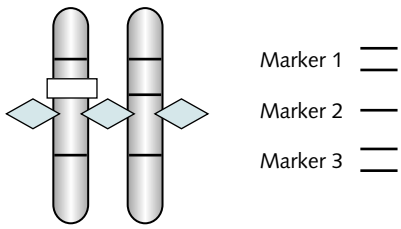
Fig. 5.3 Schematic representation of loss of heterozygosity (LOH) in MDS
In the presence of two alleles, two bands will be apparent after PCR amplification, whereas a single band will be seen if there is allelic deletion.

not sensitive below the level of approximately 5–10% of cells. However, the identification of clonal abnormalities in a high percentage of cells may indicate a proliferative advantage for these cells, and may be more predictive of the development of t-MDS/AML. The specificity of interphase FISH is also unknown, since we do not know how many patients who do not develop t-MDS/AML have interphase FISH abnormalities at the time of stem cell harvest. The test has been validated only in retrospective studies, and it is time- and labor-intensive as a screening test.

LOH analysis is based on the loss of one allele at a particular locus, usually by PCR analysis. This strategy can be used to identify LOH and to define the excursion of large deletions (Figures 5.3 and 5.4). It is a population-based assay and requires prior selection of loci to be analyzed. It lacks sensitivity and is probably unable to detect fewer than 20% of cells with LOH at a given locus. However, it is more likely to be specific, in that a positive test indicates clonal expansion of cells with LOH. It is amenable to high-throughput strategies, but has not yet been validated as a predictor of post-ASCT t-MDS/AML in prospective studies, although such studies are under way.

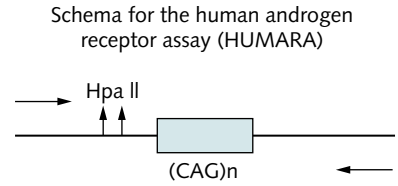
PCR for point mutations or chromosomal translocations is emerging as a potentially useful predictor of t-MDS/AML as we learn more about the molecular genetics of the disease. Markers that may be useful include mutations in RAS, FLT3, AML1 and MLL. In addition, PCR can be used to identify fusion transcripts, including AML1/EVI1, PML/RARa and 11q23 gene rearrangements. The PCR approach is also by definition locus-specific, and there are relatively few markers known to date. It is highly sensitive and capable of detecting only a few cells. But, as noted above, since some normal individuals harbor PCR-detectable rearrangements, the specificity of the assay remains to be determined in this context. The test is probably most informative when performed using quantitative techniques, such as the Taqman PCR, and is amenable to high-throughput analysis, but has not yet been validated as a predictor of t-MDS/AML.

X-inactivation-based clonality assays require no locus-specific information, or indeed any information about the nature of the mutation that causes t-MDS/AML (Figure 5.5). It detects only clonal populations of cells that have a proliferative advantage over normal polyclonal cells. It uses DNA, is PCR-based and is readily amenable to high-throughput analysis, but is only applicable to female patients. There are several potential pitfalls of this test, including false-positive



- The deletion is bordered by markers 1 and 3

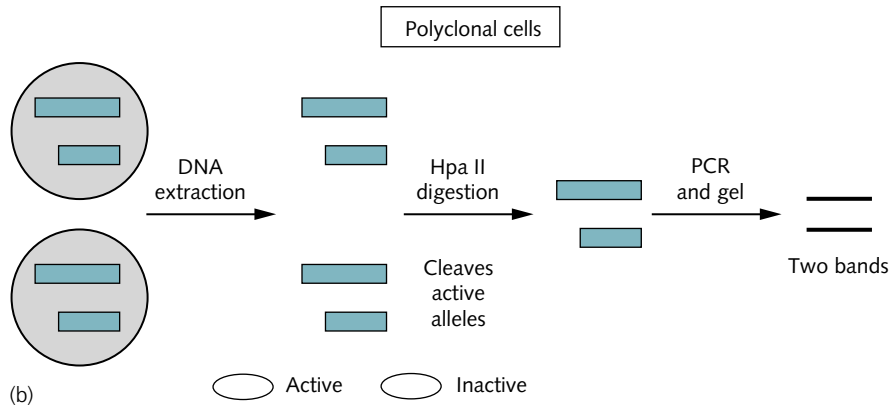
Fig. 5.4 Schematic representation of the use of LOH to map deletions

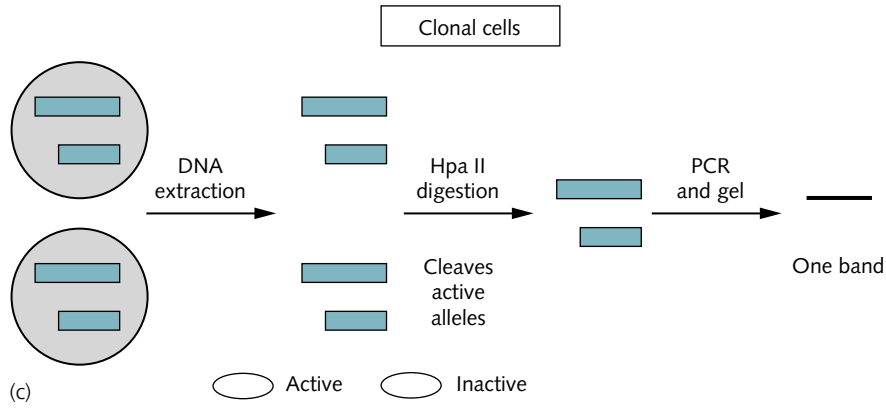


- Variable length CAG expansion repeat distinguishes the maternal from paternal X in >90% of females
- Variably methylated Hpa II sites distinguish active from inactive X chromosomes

(a)

Fig. 5.5 Human androgen receptor assay (HUMARA)
(a) Schema of the assay, which uses the variable-length CAG repeat pattern to distinguish the maternal and paternal X chromosomes.
(*Continued*.)





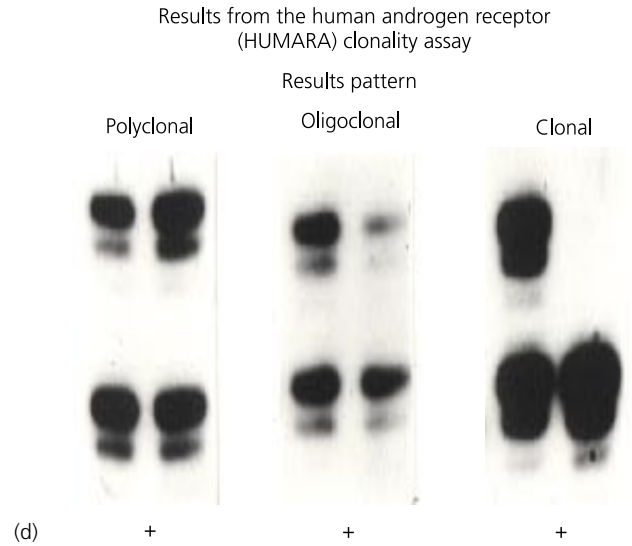


Fig. 5.5 Human androgen receptor assay (HUMARA) (*continued*)
(b) Two bands will be seen after PCR amplification in polyclonal cells where there is random inactivation.
(c) A single band will be seen in a clonal population.
(d) Results from patients studied, showing polyclonal, oligoclonal and clonal populations.

tests due to germline or acquired skewing of the pattern of X-inactivation. This problem can be overcome in part by the appropriate use of related tissue controls. However, the test may be difficult to interpret in cases with severe skewing of the X-inactivation pattern. This technique has been validated in retrospective studies and prospective studies are ongoing.

Approaches to minimizing the risk of t-MDS/AML

It may be appropriate to minimize, where possible, agents that are particularly associated with the greatest risk, including alkylating agents, external beam irradiation and topoisomerase inhibitors. This can be accomplished in part by the identification of high-risk individuals who are likely to require ASCT as part of their therapy. Recent innovations in the application both of standard prognostic indicators and of global expression arrays may help in the identification of such patients, and efforts to assess risk using molecular markers should be further explored and validated. It seems advisable to avoid TBI as part of the conditioning regimen, although it may be best to directly determine the risk/benefit ratio of using TBI in a randomized trial. If standard cytogenetics are abnormal, allogeneic rather than autologous stem cell transplantation may be indicated. Selected FISH loci, such as 5q, 7q, +8, 20q and –11 should be explored prospectively as predictors of outcome, as should X-inactivation-based clonality assays. Effort should be devoted to pilot retrospective studies to evaluate the role and validity of genome-wide LOH screens, quantitative PCR for specific mutations and gene rearrangements, and the assessment of global expression patterns to identify signatures predictive of t-MDS/AML.

Further reading

Castilla LH, Garrett L, Adya N *et al.* (1999) The fusion gene Cbfβ blocks myeloid differentiation and predisposes mice to acute myelomonocytic leukemia. *Nature Genetics*, **23**, 144–146.

Ford AM, Bennett CA, Price CM *et al.* (1998) Fetal origins of the TEL-AML1 fusion gene in identical twins with leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 4584–4588.

Friedberg JW, Neuberg D, Stone RM *et al.* (1999) Outcome in patients with myelodysplastic syndrome after autologous bone marrow transplantation for non-Hodgkin's lymphoma. *Journal of Clinical Oncology*, **17**, 3128–3135.

Mach-Pascual S, Legare RD, Lu D *et al.* (1998) Predictive value of clonality assays in patients with non-Hodgkin's lymphoma undergoing autologous bone marrow transplant: a single institution study. *Blood*, **91**, 4496–4503.

Pedersen-Bjergaard J, Andersen MK, Christiansen DH. (2000) Therapy-related acute myeloid leukemia and myelodysplasia after high-dose chemotherapy and autologous stem cell transplantation. *Blood*, **95**, 3273–3279.

Zimonjic DB, Pollock JL, Westervelt P *et al.* (2000) Acquired, nonrandom chromosomal abnormalities associated with the development of acute promyelocytic leukemia in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 13306–13311.

Chapter 6 Detection of minimal residual disease in hematological malignancies

Drew Provan & John G Gribben

Introduction

Despite advances in the treatment of human hematological malignancies, a significant proportion of patients relapse, usually with the same malignant clone found at diagnosis. Until recently, detection of residual leukemia or lymphoma cells in marrow, blood or lymph nodes relied on light microscopy and flow cytometry. However, these techniques are not sensitive for the detection of small numbers of malignant cells. Other, more sensitive, methods are now available to assess whether early detection of residual tumor might allow intervention and prevent relapse of disease. Molecular techniques, such as the polymerase chain reaction (PCR), seem to offer highly sensitive detection of malignant DNA sequences, and this technique has been applied to a wide variety of diseases.

Many studies have now been carried out in a variety of disorders, and whilst it is true that for many hematological cancers the persistence of PCR-detectable disease predicts which patients will do less well, this does not hold true for all diseases studied. It appears that patients with some malignancies may harbor residual tumor cells for many years without ever showing any evidence of clinical relapse. This will be discussed in detail later in this chapter.

This chapter outlines the methods available, with particular emphasis on PCR amplification, and their clinical application to a variety of hematological malignancies, including lymphomas and leukemias. The molecular basis of leukemia and lymphoma is discussed in detail in Chapters 4 and 10.

What is minimal residual disease?

Minimal residual disease (MRD) describes the *lowest level of disease detectable using available methods*. Previously, light microscopy, cytogenetic analysis and flow cytometry were standard techniques used for the detection of residual malignant cells in the blood and marrow of patients after treatment. However, the sensitivities of these methods do not allow identification of low levels of disease, nor do they allow accurate quantitation of malignant cell numbers. Since these residual malignant cells may be the source of ultimate relapse, there has been great interest in developing molecular techniques for the detection of residual tumor. For many years Southern blot hybridization has been the gold standard for the detection of DNA sequence alterations at specific genetic loci, but it has been largely superseded by PCR amplification of DNA sequences. Because of the power of PCR technology, we are now able to detect one residual malignant cell in a background of one million normal cells. Molecular targets for PCR-based approaches include chromosomal translocations and antigen receptor (immunoglobulin and T-cell receptor) gene rearrangements.

Methods available for the detection of residual disease

To date, several methods have been used to determine the presence of residual neoplastic cells in blood, bone marrow or other tissue following therapy (Figure 6.1). The ideal assay

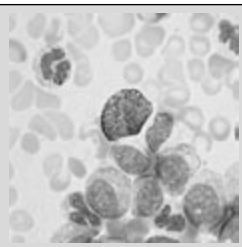
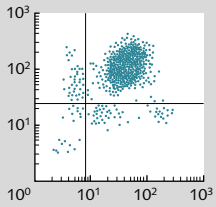
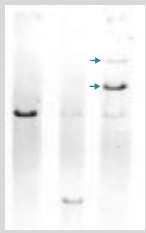
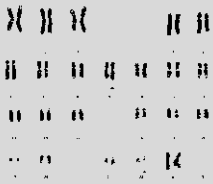
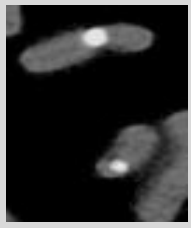
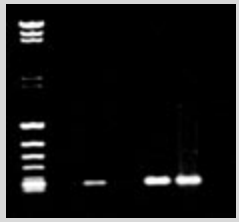
Technique		Features	Sensitivity
Morphology		Low sensitivity	5%
Immunophenotyping		Lacks specificity	1–5%
Southern blotting		Labour intensive Slow	1%
Cytogenetics		Labour intensive Slow Requires metaphase chromosome preparations	1%
FISH		Labour intensive Interphase FISH obviates need for high quality metaphases (cf. standard cytogenetics)	0.3–5%
PCR amplification		DNA sequence information required False +ve results	0.001%

Fig. 6.1 Methods of detection of marrow infiltration in non-Hodgkin's lymphoma, showing the sensitivity of each

system for the detection of small numbers of malignant cells in a marrow or blood sample should fulfill the following criteria: the method should be applicable in most cases of the disease under investigation; the method should be specific for the neoplastic cell type; the method should be sensitive; and the method should allow quantitation of tumor burden for prognostic purposes.

Methods include:

- morphology
- flow cytometry and immunophenotypic analyses
- cell culture assays
- karyotypic analysis
- fluorescence *in situ* hybridization techniques
- molecular analyses, including Southern blotting and PCR.

Morphology

In acute leukemia, remission is the term used to describe a bone marrow containing fewer than 5% blast (i.e. leukemic) cells using conventional light microscopy, but this may still represent a considerable tumor burden since, at diagnosis, the leukemic cell number may be 10^{12} and, following therapy, the neoplastic cell number may drop only by 2 logs to 10^{10} even in the presence of fewer than 5% marrow blasts. Standard morphology alone is not a sensitive method for determining low levels of disease and is a poor indicator to attempt to predict impending relapse (Table 6.1).

Flow cytometry and immunophenotyping

Immunophenotypic analysis using single monoclonal antibodies to cell membrane or cytoplasmic proteins lacks absolute specificity for leukemia or lymphoma cells and is therefore of limited value. Combining monoclonal antibodies allows the more specific detection of residual disease and quantitation is possible, although the tumor cell burden may be underestimated. The technique is further hampered by the lack of true 'specific–specific' surface determinants and tumor-associated antigens are normal differentiation antigens present on developing hematopoietic progenitor cells. Using combinations of monoclonal antibodies and multicolor flow cytometric analysis, the sensitivity of this technique can be greatly enhanced. Except in the most expert hands, this technique is limited to a sensitivity of around 10^{-4} (i.e. 1 malignant cell in 10 000 normal cells).

Table 6.1 Sensitivity of methods for MRD detection.

Standard morphology	1–5%
Cytogenetics	5%
Fluorescence *in situ*	0.3–5%
Immunophenotyping	10^{-4}
Translocations	
PCR	10^{-6}
Gene rearrangements	
Southern blotting	1–5%
PCR	10^{-4} to 10^{-6}

Cell culture assays

These involve growing T-cell-depleted marrow in culture after the patient has undergone treatment, followed by subsequent morphological, immunophenotypic and karyotypic analyses on the colonies produced. Due to the variability of culture techniques between and within laboratories, this method has proved unreliable and insensitive for detecting persisting blasts. In addition, culture techniques do not provide any estimate of cell number and hence provide little information about tumor cell burden.

Karyotypic analysis

Detection of non-random chromosomal translocations is of great value in the diagnosis of leukemias and lymphomas. Chromosomal abnormalities are present in at least 70% of patients with acute lymphoblastic leukemia (ALL) and 50% of patients with chronic lymphocytic leukemia (CLL). However, karyotypic analysis is of limited value following therapy, with a sensitivity level of around 5%, making it little better than standard morphological analysis. In addition, cytogenetics relies on obtaining adequate numbers of suitable metaphases for analysis, which is difficult in some malignancies.

Fluorescence in situ hybridization (FISH)

FISH can detect smaller chromosomal abnormalities than standard karyotyping and allows analysis of interphase nuclei (cf. metaphase preparations in standard karyotyping). The method involves the binding of a nucleic acid probe to a specific chromosomal region. Preparations are counterstained with fluorescent dye, allowing the chromosomal region of interest to be detected. The technique is useful in the diagnosis of trisomies and monosomies and has been particularly useful in identifying deletions that have prognostic significance in CLL. The sensitivity of the technique is around 1%, making it considerably more useful than standard karyotyping for follow-up marrows in patients with leukemias or lymphomas, but is still of limited value for MRD detection.

Molecular techniques — Southern blot hybridization

Initially described by its inventor, Professor Ed Southern, in the 1970s, Southern blotting involves the digestion of

chromosomal DNA using bacterial restriction enzymes, with size separation of the DNA fragments using electric current and gel electrophoresis before transferring these to a nylon support membrane. A labeled probe for the gene of interest is applied, which binds to its complementary sequence on the membrane and visualization of the gene is by autoradiography (Figure 6.2).

Southern blotting is useful for the initial diagnosis of leukemia and lymphoma using probes specific for translocations or gene rearrangements. With Southern blotting, a non-germline or rearranged gene pattern may be seen in DNA from a population of cells where more than 1% of the total population is made up by a clone of malignant lymphoid cells. In other words, Southern blotting will detect a rearranged gene provided the cells containing the rearranged gene exceed 1 in 100 normal cells. The disadvantage of Southern blotting is that the technique is not sufficiently sensitive for the detection of small numbers of malignant cells persisting after therapy and giving rise to disease relapse. For this reason, Southern blotting has largely been replaced by PCR for the detection of MRD.

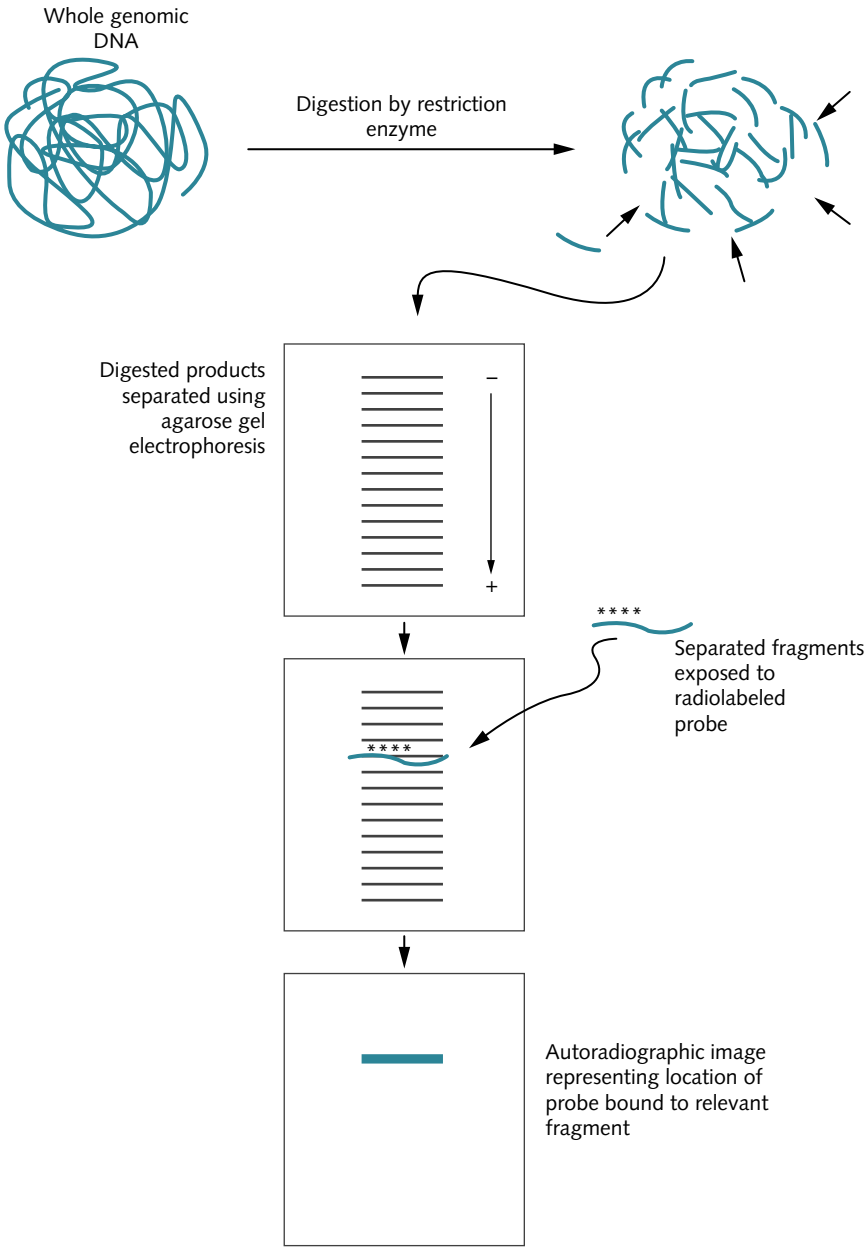


Fig. 6.2 Principle of Southern blotting Genomic DNA is digested using a restriction enzyme, after which the fragments are separated on the basis of size using agarose gel electrophoresis, and are finally transferred to a nylon membrane. Radiolabeled probe for the gene of interest is hybridized to the DNA on the membrane and, after removal of the non-specifically hybridized probe, the location and size of the fragment are determined using autoradiography.

PCR amplification of DNA

As described above, Southern blotting is a useful technique for assessing whether there is a clone of abnormal cells in blood, marrow or other tissue but is not useful if these cells are present in only very small amounts. In this case, techniques that involve amplification of specific DNA sequences are required. PCR has filled the void in this respect and has found a place in diagnostic laboratories investigating oncogenes, hematological malignancies, single-gene disorders and infectious diseases. Part of the attraction of a PCR-based approach is its extreme simplicity and the speed with which results are obtained.

What is PCR amplification?

In the PCR reaction, two short oligonucleotide DNA primers are synthesized that are complementary to the DNA sequence on either side of the translocation or gene of interest. The region between the primers is filled in using a heat-stable bacterial DNA polymerase (*Taq*) from the hot-spring bacterium *Thermus aquaticus*. After a single round of amplification has been performed, the whole process is repeated (Figure 6.3). This takes place 30 times (i.e. through 30 cycles of amplification) and leads to a million-fold increase in the amount of specific sequence. When the 30 cycles are complete, a sample of the PCR is electrophoresed on agarose or polyacrylamide gel. Information about the presence or absence of the region or mutation of interest is obtained by assessing the sizes and numbers of different PCR products obtained after 30 cycles of amplification.

The specificity of PCR can be further increased by the use of nested PCR, which involves re-amplification of a small amount of the amplified product (obtained using outside, external, primers) using internal oligonucleotide primers.

PCR has the advantage that very little tissue sample is required for analysis and the technique can be applied to a variety

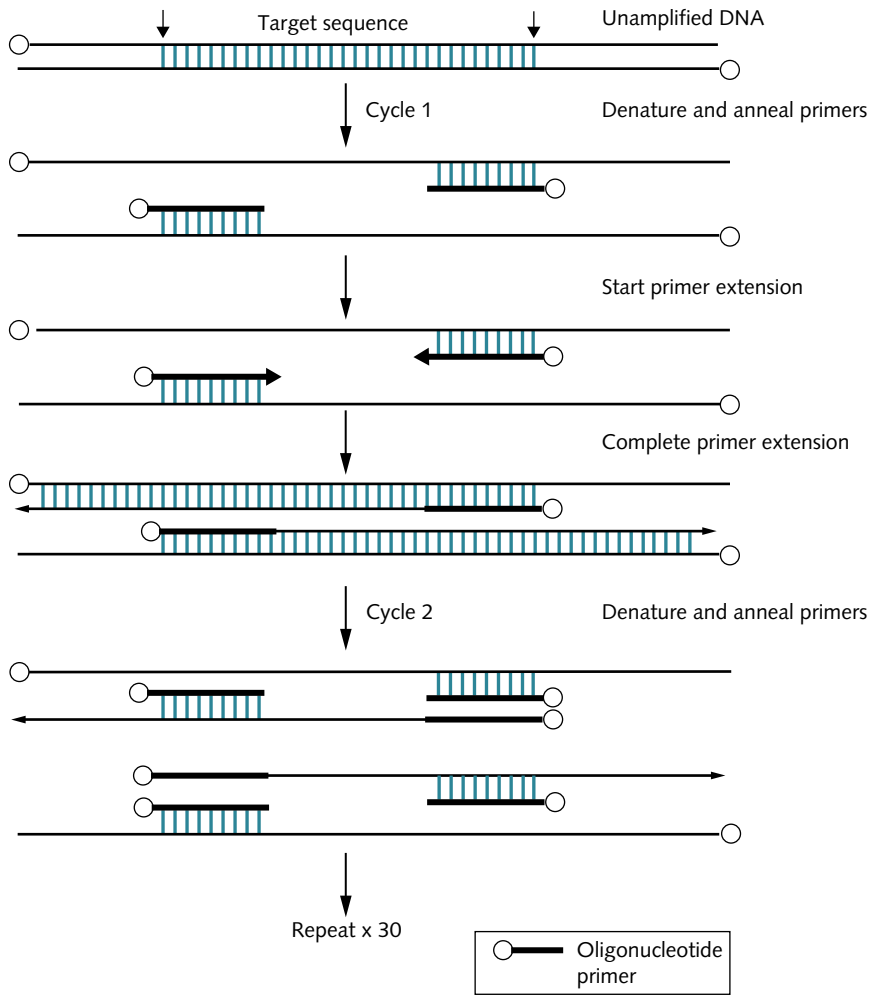


Fig. 6.3 Simplified PCR schema
Double-stranded DNA is denatured to allow binding of specific oligonucleotides on either side of the region of interest. *Taq* DNA polymerase extends the oligonucleotides before the double-stranded molecules are denatured and the process is repeated.

of different sample types, for example fresh, unfixed, cryopreserved and formalin-fixed paraffin-embedded tissue as well as hematoxylin and eosin-stained and formalin-fixed tissue.

PCR may be used to detect the presence of chromosomal translocations. The most commonly investigated rearrangements include the t(9;22) chromosomal translocation in chronic myeloid leukemia (CML), t(1;19), found in a subset of pre-B-cell ALL, the t(14;18) found in 85% of follicular and 15% of diffuse large cell lymphomas, and several others. Alternatively, in the lymphoid malignancies, if the tumor being investigated does not carry a translocation marker, PCR may be used to amplify rearranged antigen receptor [Ig or T-cell receptor (TCR)] genes.

Molecular targets

Chromosomal translocations

Translocations, which involve the transfer of DNA between chromosomes, are found in many of the hematological malignancies. Other chromosomal abnormalities include chromosomal deletions and inversions. Table 6.2 shows some of the translocations described in myeloid and lymphoid malignancies. As a result of chromosomal translocation, a gene from one chromosome ends up adjacent to a gene on the chromosome to which the DNA has been translocated, and this may have important consequences for the cell (and the patient). If a potentially cancerous gene (proto-oncogene), which is generally not transcriptionally active, abuts onto a gene that is being actively transcribed, this may result in upregulation of expression of that proto-oncogene. This is exactly the situation in many translocations described to date. In some cases, such as the translocation between chromosomes 14 and 18 found in many cases of follicular lymphoma, the *BCL-2* gene is moved to chromosome 14 and comes under the transcriptional control of the immunoglobulin heavy chain (IgH) gene, which is transcribed actively. The increase in BCL-2 protein prevents apoptosis (programmed cell death) and this may explain, in part, the underlying pathogenesis of some lymphomas.

The first non-random chromosome translocation described was the Philadelphia chromosome, in which recip-

Table 6.2 PCR-amplifiable chromosomal translocations and gene rearrangements in human hematological disorders.

Disease	Translocation	Genes involved	
Acute myeloid leukemia			
M2	t(8;21)	ETO–AML1	
M2 or M4	t(6;9)	DEK–CAN	
M3	t(15;17)	PML–RARα	
M4	inv(16)	CBFβ–MYH11	
Acute lymphoblastic leukemia			
	B-lineage	t(9;22)	BCR–ABL
		t(1;19)	E2A–PBX1
		t(17;19)	HLF–E2A
		t(12;21)	TEL–AML-1
		t(4;11)	AF4–MLL
		t(8;14)	MYC–IgH
	T-lineage	TAL interstitial deletion	TAL
		t(1;14)	TAL-1–TCRδ
		t(10;14)	HOX11–TCRα
		t(11;14)	11p13–TCRδ
Lymphomas			
Follicular and diffuse NHL	t(14;18)	BCL-2–IgH	
Mantle cell lymphoma	t(11;14)	BCL-1–IgH	
Burkitt's lymphoma	t(8;14)	MYC–IgH	
Anaplastic lymphoma	t(2;5)	ALK–NPM	
Gene rearrangements			
Immunoglobulin heavy chain	B-cell lymphoma/leukemia		
T-cell receptors	T-cell lymphoma/leukemia		

rocal translocation of DNA between chromosomes 9 and 22 takes place. In the t(9;22), the distal ends of chromosomes 9 and 22 are exchanged in a so-called reciprocal translocation; that is, there is no overall net loss or gain of genetic material. The C-*ABL* proto-oncogene from chromosome 9 becomes joined to BCR (breakpoint cluster region) on chromosome 22, resulting in a chimeric fusion protein which has tyrosine kinase properties, and through some unknown mechanism leads to the typical CML phenotype (*discussed in detail in Chapter 7*).

Detecting the presence of translocations

(Table 6.3)

Some translocations are disease-specific

Follicular lymphoma is characterized by the t(14;18), which is found in almost 90% of cases. However, this translocation is found in other types of non-Hodgkin's lymphoma (NHL), so that the t(14;18) is not, in itself, diagnostic of one particular malignancy. Acute promyelocytic leukemia (AML M3) is characterized by a reciprocal translocation between chromosomes 15 and 17. This is found in the majority of cases but, unlike t(14;18), the t(15;17) is not found in any other neoplasm or in health and so serves as a diagnostic marker for this disease (although its absence does not exclude the diagnosis). Although t(9;22) is characteristic of CML, it is important to detect this in cases in which blastic transformation has occurred. In addition, t(9;22) occurs in a subset of patients with acute lymphoblastic leukemia (ALL) and in these cases is associated with a particularly poor prognosis. It is therefore important to identify these patients at diagnosis since their prognosis and treatment differ from those for other cases of ALL.

More recently a number of chromosomal translocations that were thought to be leukemia- or lymphoma-specific have been found in the blood of normal individuals when assessed by PCR amplification, including t(14;18), t(8;14), t(2;5), t(9;22), t(4;11), t(15;17) and t(12;21). The implication of this finding is that these rearrangements are not themselves sufficient for malignant transformation of cells, in keeping with a multi-hit hypothesis for tumor development.

Table 6.3 Detecting the presence of translocations.

Standard cytogenetics
If the translocation alters the appearance of banded chromosomes using standard cytogenetic analysis.
Fluorescence *in situ* hybridization (FISH)
Using metaphase or interphase techniques.
Polymerase chain reaction
Requires the DNA on either side of the breakpoint to be sequenced to allow oligonucleotide primers to be constructed.

Translocations may be used for detecting residual disease

Translocations serve as useful diagnostic disease markers at presentation for a variety of leukemias and lymphomas. For the detection of MRD, standard cytogenetic analysis for the detection of translocations is not sufficiently sensitive for follow-up but other techniques can be applied, including FISH and PCR. FISH techniques are constantly being improved (*see Chapter 2*) and may be of value for MRD detection. However, more sensitive MRD detection is possible using PCR in cases where the translocations are well characterized and DNA on either side of the breakpoints has been sequenced. MRD using the chromosomal translocations t(14;18) and t(9;22) and other translocations are described later.

Antigen receptor gene rearrangements: immunoglobulin and TCR genes as molecular markers

Many hematopoietic malignancies have no detectable translocation suitable for PCR amplification, and in these cases an alternative strategy is required. In the lymphoid malignancies there is rearrangement of the antigen receptor at the immunoglobulin H (IgH) or TCR genes. The Ig and TCR molecules belong to a group of related proteins termed the immunoglobulin superfamily. Other members include CD8, N-CAM and MHC molecules. The Ig and TCR molecules have many similarities and have been shown to share common amino acid motifs. It is estimated that the immune system requires in excess of 10^{10} specific antibodies to respond to antigenic determinants encountered in the environment. If each Ig molecule were encoded separately in the germline, most of our genome would consist simply of Ig genes. Elegant work by Tonegawa has shown that Ig and TCR genes exist in the germline state as non-contiguous DNA segments that are rearranged during lymphocyte development (Table 6.4). Gene rearrangement involves recombination of germline gene segments that results in a permanently altered non-germline configuration (Figure 6.4). The process of Ig and TCR gene assembly ensures almost limitless variation of Ig and TCR molecules using only a limited amount of chromosomal DNA. Other features that ensure Ig and TCR variability include imprecise joining of individual V, D and J segments, duplication and inversion of segments, and somatic mutation (in Ig genes) of V, D and J.

Table 6.4 Immunoglobulin and T-cell receptor diversity is achieved through rearrangement of separate germline segments.

	Diversity of immunoglobulin and TCR genes						
	Immunoglobulin			T-cell receptor			
	H	κ	λ	α	β	γ	δ
V segments	250	100	100	60	80	8	6
D segments	15	0	0	0	2	0	3
J segments	6	5	4	50	13	5	3
VDJ recombination	104	500	400	3000	2000	40	18
N regions	2	0	0	1	2	1	4
N region additions	V–D, D–J	None	None	V–J	V–D, D–J	V–J	V–D1, D1–D2,D1–J
V domains	10^{10}	10^4	10^4	10^6	10^9	10^4	10^{13}
V domain pairs		10^{14}			10^{15}		10^{17}

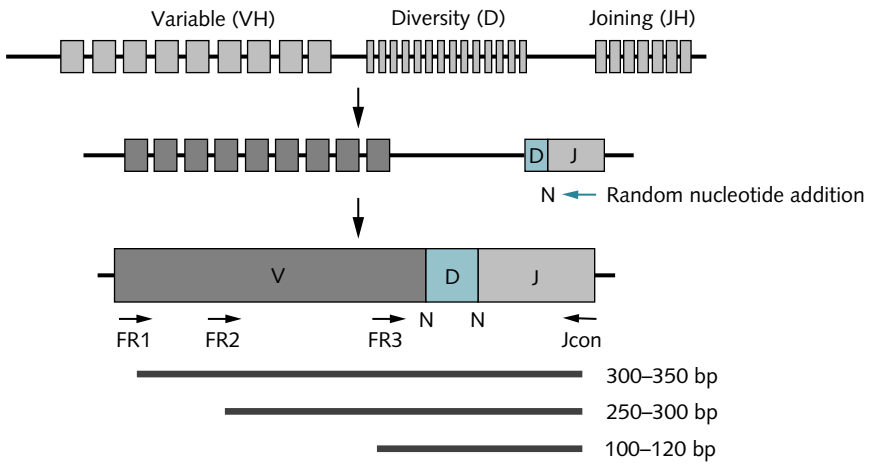


Fig. 6.4 VDJ rearrangement Rearrangement of non-contiguous germline V-, D- and J-region segments generates a complete V–D–J complex, which serves as a useful marker of malignancy. FR1, 2 and 3 refer to framework regions 1, 2 and 3, respectively; N, random N nucleotides; Jcon, JH consensus primer. The sizes of the various PCR products are shown (FR1 + Jcon generates a fragment of 300–350 bp, and so on).

The immunoglobulin heavy chain locus

During normal lymphoid development, both B and T lymphocytes undergo rearrangement of their antigen receptor genes, i.e. Ig genes in B-cells and TCR genes in T-cells, and their clonal progeny bear this identical antigen receptor rearrangement. B-cell neoplasms, including NHL, ALL, myeloma and CLL, undergo irreversible somatic rearrangement of the IgH locus, providing a useful marker of clonality and the stage of differentiation in these tumors. Until recently, the lineage of Hodgkin lymphoma cells was unclear. PCR amplification of Ig genes has demonstrated that the vast majority of cases of Hodgkin disease are of B-cell lineage.

The IgH locus is located on chromosome 14q32.3. Unlike the light chain (IgL) locus, IgH contains diversity segments in addition to V, J and C segments. In humans there are around 250 heavy-chain variable region (VH) segments, of which two-thirds are probably pseudogenes, representing ancestral gene remnants (denoted by ψ). The VH elements fall into seven families (VH1, VH2, VH3, VH4a, VH4b, VH5 and VH6). Unlike the TCR and IgL loci, the IgH locus contains multiple heavy-chain constant region (CH) segments (Figure 6.5), some 11 in total, including two pseudogenes (Cμ, Cδ, Cγ3, Cγ1, Cψε, Cα1, Cψγ, Cγ2, Cγ4, Cε and Cα2). Each C segment contains multiple exons corresponding to the functional domains in the heavy-chain protein (CH1, CH2, CH3, etc.). The multiple C elements correspond to the different classes of heavy chain encountered during class switching. Cμ generates IgM, Cα generates IgA, and so on. This mechanism ensures that, although the heavy chains are of varying class, they will all bear identical V–D–J sequences.

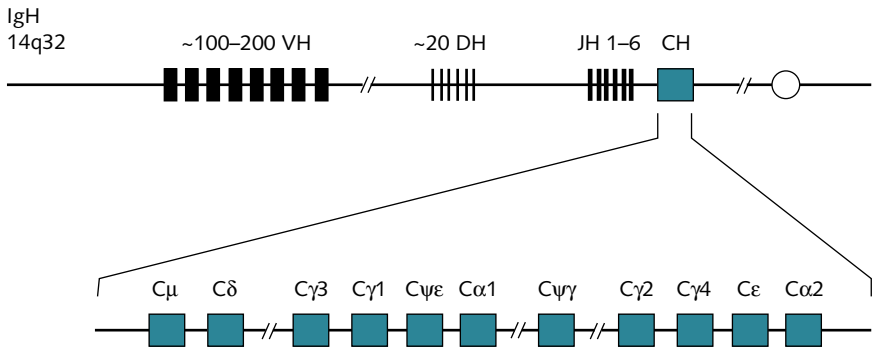


Fig. 6.5 Genetic map of region 14q32
The CH segments are shown towards the 3′ end of the region.

Third complementarity-determining region (CDR3)

The CDR3 region of the IgH gene is generated early in B-cell development and is the result of rearrangement of germline sequences on chromosome 14. One diversity segment is joined to a joining region (D→J). The resulting D–J segment then joins one variable-region sequence (V→DJ), producing a V–D–J complex (Figure 6.4). The enzyme terminal deoxynucleotide transferase (TdT) inserts random nucleotides at two sites: the V–D and D–J junctions. At the same time random deoxynucleotides are removed by exonucleases. Antibody diversity is further increased by somatic mutation, a process that is not found in TCR genes. The final V–N–D–N–J sequence (CDR3) is unique to that cell, and if the cell multiplies to form a clone this region will act as a unique marker for that malignant clone. The V(D)J product corresponds to part of the variable region of the antibody molecule.

TCR genes undergo a similar process of rearranging their germline segments to produce complete TCR genes

Junctional region diversity

Imprecise recombination involving V(D)J region DNA enhances the number of possible different antibody molecule polypeptides due to loss or gain of additional nucleotides during the recombination event. The resulting V(D)J product may be functional, i.e. generate antibody molecules, or, if the reading frame is lost, non-functional. Whether functional or not, the CDR3 remains a unique marker for the malignant clone.

TdT inserts N region nucleotides into the CDR3

N region nucleotide insertion is seen at the boundary of V, D or J coding segments and is template-independent. These N regions contain between one and 12 nucleotides and are more often guanine or cytosine rather than adenine or thymidine, reflecting the role played by the enzyme TdT in this process.

Combinatorial association

The TCR molecules are dimeric proteins, usually $\alpha + \beta$ (TCR α:β), although 5% of circulating T-cells bear the γ:δ TCR. The random combination of subunits in the TCR dimers further enhances the generation of diversity. The recombination events on one chromosome leading to the production of a functional molecule, such as TCR α:β, result in the inhibition of recombination at that locus on the other chromosome. This so-called allelic exclusion ensures that any given lymphocyte will express only one type of receptor molecule.

Somatic hypermutation

This describes the random introduction of mutations within the V, D and J segments and is well documented in Ig genes but does not contribute to diversity in the TCR genes. Rearranged V region sequences in B cells have been analyzed and found to differ from those of the germline V sequences from which they were generated. Most of these mutated V regions are found in the secondary immune response on rechallenge of B cells with antigen. During this process the antibody of the primary response (IgM) is switched to IgG or IgA. The somatic mutation rate has been estimated to be as high as 10^{-3} per base pair per cell generation, and the process occurs predominantly in variable regions of the molecule. The presence of somatic mutation can be useful in determining the stage of lineage in B-cell malignancies. In CLL it has been shown that cells either do or do not have mutated Ig genes. This has important prognostic significance since those cases that have undergone somatic hypermutation have a better prognosis than those cases that have no mutations.

The clinical utility of the CDR3 DNA sequence

The description of V–D–J recombination may appear arcane, with no obvious relevance in clinical terms, but it is the formation of this unique recombination product that generates a powerful specific–specific marker that we can use for the detection of malignant clones and MRD. The DNA sequence within the V–D–J is determined by sequencing, following which the individual V, D and J segments are delineated. This allows accurate identification of the N region nucleotides (which are generated randomly by the enzyme TdT) that form the basis of the unique clone-specific (patient-specific) probe (Figure 6.6).

There are two sites available for design of the customized probe—the DNA of the V–N–D sequence and that of the D–N–J sequence. Does it matter which one we use to make the probes? The V–N–D sequence generally has a larger N region with more random nucleotides inserted, but the D–N–J site appears preferable for use as a clone-specific probe since there is less base deletion of the 3′ end of the framework region 3 (FR3) than of the 5′ end of the J region. In addition, the D–J segments appear to be inherently more stable than V–D segments. Finally, where there is V→V switching, as happens in some diseases, such as ALL, the D–J segment remains unchanged and the probe will still detect the clone even if the V regions alter. The consensus view at present is that the D–N–J is probably the best DNA sequence to use to make probes for MRD detection.

Quantitation of the neoplastic cells using PCR

Until fairly recently, PCR amplification simply confirmed the presence (+) or absence (–) of tumor DNA sequences with little scope for quantifying the tumor bulk, particularly when using DNA as the PCR template. A band on agarose gel may represent the DNA from one cell—or many millions of cells. Clearly, this is of clinical importance if the information obtained is to be of value in determining the need for further chemotherapy, which is the main rationale for attempting to detect MRD in the first place.

In the early years of PCR detection of MRD the starting template was usually DNA, but more recently PCR amplification of reverse-transcribed mRNA (termed 'complementary DNA' or 'cDNA') has been used. This refinement in PCR amplification has evolved where analysis of translocations such as t(9;22) or t(15;17) is impossible using a DNA template, simply because of the enormous size of the target being amplified. In these translocations the primer binding sites are so far apart on the DNA template that amplification is virtually impossible. However, the mRNA transcribed from these translocations undergoes considerable modification, with excision of introns making the mRNA counterpart of the translocation much smaller than the DNA.

Quantitation using competitive PCR templates has been possible for RNA-based PCR, and so we are able to quantitate the tumor cell burden in those diseases where RNA is the nucleic acid used for the PCR assays. Diseases in which reverse transcriptase PCR (RT-PCR) is possible, with quantitation of the tumor burden, include CML [with t(9;22)], AML M3 [t(15;17)] and AML M2 [t(8;21)].

DNA templates are more difficult to quantitate, although competitive PCR templates may be of value here also. Recent technologies such as the TaqMan® real-time PCR machines may allow true quantitation using DNA as starting material. This system uses an internal oligonucleotide probe with added reporter and quenching activities (Figure 6.7). After primer and probe annealing, the reporter dye is cleaved off by the 5′–3′ nuclease activity of Taq DNA polymerase during primer extension (Figure 6.8). This cleavage of the probe separates the reporter from quencher dye, greatly increasing the reporter dye signal. The sequence detector is able to detect the fluorescent signal during thermal cycling. The advantages of this system are the elimination of post-PCR processing and the ability to examine the entire PCR process—not simply the endpoint of amplification. Moreover, since the probe is designed to be sequence-specific, non-specific amplification products are not detected.

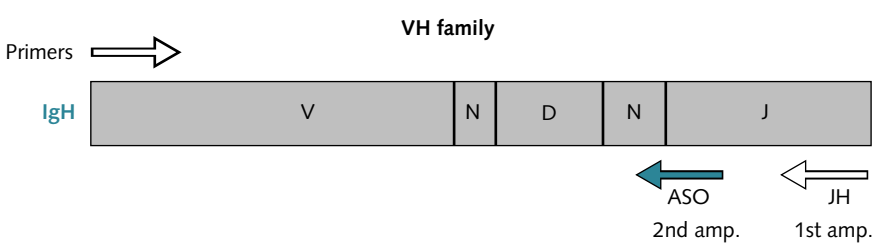


Fig. 6.6 Semi-nested PCR of IgH region in patient with B-cell tumor
V- and J-region primers are used to generate the initial PCR product. Using DNA sequence information, an allele-specific oligonucleotide (ASO) primer unique to that patient is constructed and used with the V region primer to amplify an aliquot of the first-round PCR product.

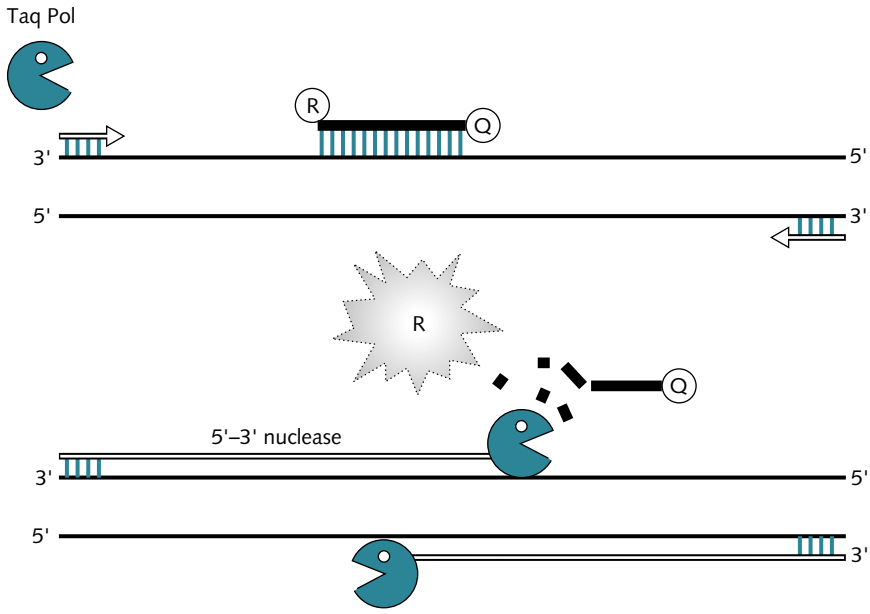


Fig. 6.7 Real-time PCR amplification
See text for details.

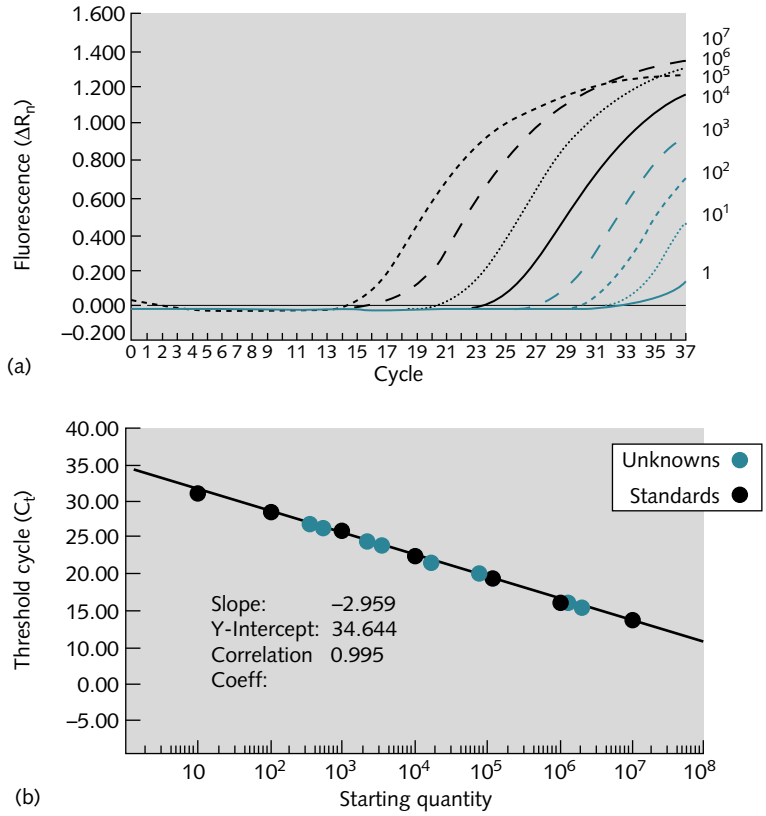


Fig. 6.8 Standard curves generated for accurate quantitation of leukemic cell burden
Quantitation by real-time PCR requires generation of a standard curve. A known amount of template DNA is diluted into genomic DNA and amplified by PCR. The threshold cycle is the cycle number at which reported fluorescence is first detected above background and is proportional to the amount of starting template DNA. The threshold cycle number is then plotted against the known amounts and a standard curve can be generated. The threshold cycle of the unknown samples can then be quantified by reading off the standard curve.

Application of PCR for residual disease detection in non-Hodgkin's lymphoma and leukemia

The standard technique used for the diagnosis of NHL is light microscopy of stained sections of lymph node or other tissue. This allows accurate classification of lymphoma subtype. In terms of detecting MRD, this technique has the limitation of detecting lymphoma cells only when they constitute approximately 5% or more of all cells (i.e. one malignant cell in 20 normal cells). Application of flow cytometric analysis for the detection of NHL has been hampered by the lack of lymphoma-specific monoclonal antibodies since all the cell surface antigens identified to date on the surface of lymphoma cells are also present on normal B cells or B-cell precursor cells (Table 6.5).

Table 6.5 Clinical utility of PCR-based studies in patients with leukemia and lymphoma.

- Detection of bone marrow infiltration as part of staging procedure.
- Detection of circulating lymphoma cells in peripheral blood.
- Detection of minimal residual disease following therapy.
- Assessing contribution of reinfused lymphoma cells to relapse in patients undergoing autologous bone marrow transplantation.
- Assessing ability of purging techniques to eradicate residual malignant cells in marrow.

Chromosomal translocations

As shown in Table 6.2, a number of chromosomal translocations and gene rearrangements associated with NHL have been identified; the breakpoints have been sequenced and are applicable for PCR amplification.

t(14;18) translocation

One of the most widely studied non-random chromosomal translocations in NHL is the t(14;18), occurring in 85% of patients with follicular lymphoma and 30% of patients with diffuse large cell lymphoma. In the t(14;18) the *BCL-2* protooncogene on chromosome 18 is juxtaposed with the IgH locus on chromosome 14 (Figure 6.9). The breakpoints have been cloned and sequenced, and have been shown to cluster at two main regions 3′ to the *BCL-2* coding region: the major breakpoint region (MBR) within the 3′ untranslated region of the *BCL-2* gene, and the minor breakpoint cluster region (m-BCR) located 20 kb downstream. Juxtaposition of the transcriptionally active IgH with the *BCL-2* gene results in upregulation of the *BCL-2* gene product and subsequent resistance to programmed cell death by apoptosis.

The clustering of the breakpoints at these two main regions at the *BCL-2* gene and the availability of consensus regions of the IgH joining (J) regions make this an ideal candidate for PCR amplification to detect lymphoma cells containing the t(14;18) translocation. A major advantage in the detection of lymphoma cells bearing the *BCL-2*/IgH translocation is that DNA rather than RNA can be used to detect the translocation. In addition, since there is variation at the site of the breakpoint at the *BCL-2* gene, the PCR products for individual patients differ in size and have unique sequences. The size of the PCR product can be assessed by gel electrophoresis and used as confirmation that the expected size fragment is amplified from a specific patient.

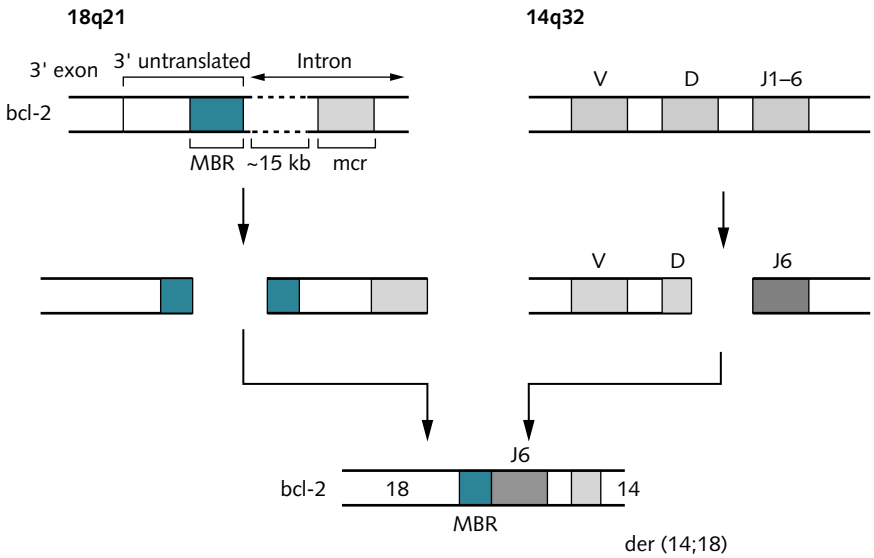


Fig. 6.9 t(14;18) translocation
In the t(14;18) the *BCL-2* locus on chromosome 18 is juxtaposed to the IgH locus on chromosome 14. The breakpoints on chromosome 18 cluster at two main regions: the major breakpoint region (MBR) in the 5′ untranslated region of the *BCL-2* gene, and the minor cluster region (mcr) downstream in the intron. The chimeric gene product provides a unique tumor marker that can be PCR-amplified using primers upstream of the MBR or mcr region with consensus primers within the J region of the IgH gene.

Other translocations in non-Hodgkin's lymphoma

The t(11;14)(q13;q32) is associated with a number of B-cell malignancies, particularly mantle cell lymphomas (MCL). In this translocation the proto-oncogene *BCL-1* (also called *PRAD-1*) on chromosome 11 is juxtaposed to the IgH chain locus on chromosome 14.

One-third of anaplastic lymphomas express the chromosomal translocation t(2;5)(p23;q35), which involves a novel protein tyrosine kinase and nucleophosmin, resulting in a p80 fusion protein. This translocation is detected by RT-PCR where the mRNA sequence is converted into cDNA before PCR amplification.

Utility of molecular techniques for detection of minimal residual disease

Lymphoma

Over the past decade a number of methods capable of detecting MRD have been developed. These techniques have clearly illustrated that patients in clinical complete remission often harbor malignant cells in low numbers. The clinical significance of the detection of such MRD is still being evaluated and remains unclear. The results of these studies will likely have great impact on the clinical management of patients as we understand more about the contribution of minimal disease to subsequent relapse. The prognostic significance of the achievement of molecular complete remission remains elusive, and few studies to date have demonstrated the importance of eradicating MRD in the patient to achieve cure. The majority of studies have been performed using as a target the t(14;18) in follicular lymphoma, but more recently studies have examined other translocations as well as Ig or TCR rearrangements and have been reporting similar results. These studies have suggested that the goal of therapy should be to eradicate the malignant clone and achieve molecular complete remission.

PCR detection of bone marrow infiltration as a staging procedure

Lymphomas generally originate in lymphoid tissue, but as the disease progresses there may be spread to other sites, such as bone marrow and blood. At initial presentation all patients undergo staging investigations to determine the extent of disease as a means of planning treatment. A number of studies have examined the use of PCR detection of t(14;18) as a staging procedure to detect lymphoma cells in the bone marrow and peripheral blood at the time of initial presentation. However, PCR analysis cannot replace morphological assessment of bone marrow (BM) since not all patients have translocations detectable by PCR, and these techniques are essentially complementary. These PCR studies have all detected lymphoma cells in the BM in a number of patients who had no overt evidence of marrow infiltration by morphology. Of great interest are those studies that have evaluated the clinical utility of MRD detection in those patients presenting with localized disease. Although the patient numbers studied are small, a significant number of patients can be found who would be upstaged from early stage to advanced stage disease by the results of PCR analysis. Whether PCR detection of minimal marrow infiltration will eventually lead to modifications in therapy in those patients currently treated with localized radiotherapy remains to be determined.

PCR detection of MRD following chemotherapy

In follicular lymphoma, long-term analysis of patients after completion of conventional chemotherapy has shown that conventional-dose chemotherapy does not eradicate PCR-detectable disease, but this may not be associated with poor outcome. One study has shown no association between the presence or absence of PCR-detectable lymphoma cells and clinical outcome. Moreover, this confirms the previous observation that some patients can indeed remain in long-term continuous complete remission despite the presence of PCR-detectable lymphoma cells, strongly suggesting that the detection of residual lymphoma cells has no prognostic significance. Cells containing t(14;18) might not always represent residual lymphoma cells, but may be cells with this translocation but without the additional necessary cellular changes required for malignant transformation. However, an alternative explanation is that conventional chemotherapy might not cure any patients with advanced stage follicular lymphoma and that all patients with persistent lymphoma cells are destined to relapse. The long-term remission status of these small numbers of patients might therefore represent merely the very long duration of their disease course.

These studies suggest that conventional-dose chemotherapy did not result in molecular remission. More novel treatment approaches, including more aggressive induction therapy and combinations of monoclonal antibody therapy with chemotherapy and the use of stem cell transplantation, have all been reported to be capable of eradicating PCR-detectable disease, achieving so-called molecular complete remission. In all of these circumstances, eradication of PCR-detectable disease has been shown to be associated with improved outcome in follicular lymphoma, strongly suggesting that eradication of MRD may be required for cure. With longer follow-up this question should be answered.

Detection of circulating lymphoma cells in peripheral blood

Blood is less frequently involved than marrow at presentation, but becomes more frequent as disease progresses. Studies at the time of initial presentation have suggested a high level of concordance between the detection of lymphoma cells in the peripheral blood (PB) and BM when assessed by PCR. However, other studies have found that the BM is more likely than PB to contain infiltrating lymphoma cells in previously untreated patients. The presence of residual lymphoma in the BM but not in the PB argues strongly that the marrow is indeed infiltrated with lymphoma in these patients and does not simply represent contamination from the PB. The findings of PB contamination with NHL when assessed by PCR are likely to have profound implications since there is now increasing interest in the use of PB stem cells, rather than BM, as a source of hematopoietic progenitors. A number of studies have demonstrated that PB stem cell collections may also be contaminated with lymphoma cells when assessed by PCR techniques. In addition, much work is being performed to monitor the effects of chemotherapy and the growth factors that are used to mobilize hematopoietic progenitor cells, to ensure that these agents do not also mobilize lymphoma cells.

Contribution of re-infused lymphoma cells to relapse after autologous stem cell transplantation

In low-grade NHL there has been increasing interest in the use of high-dose therapy as salvage therapy for patients who have failed conventional-dose chemotherapy regimens. The resulting ablation of a patient's marrow after high-dose therapy can be rescued by infusion of allogeneic or autologous stem cells. Autologous stem cell transplantation (ASCT) has several potential advantages over allogeneic stem cell transplantation for marrow rescue: there is no need for a histocompatible donor and there is no risk of graft-versus-host disease. ASCT can therefore be performed more safely, and in older patients, and has become a major treatment option for an increasing number of patients with hematological malignancies.

The major obstacle to the use of ASCT is that the infusion of occult tumor cells harbored within the stem cell collection may result in more rapid relapse of disease. To minimize the effects of the infusion of significant numbers of malignant cells, stem cells are collected when the patient either is in complete remission or has no evidence of lymphoma in the blood. In addition, a variety of methods have been developed to purge malignant cells from the stem cell collection to attempt to eliminate any contaminating malignant cells and leave intact the hematopoietic stem cells that are necessary for engraftment. The development of purging techniques has led to a number of studies of ASCT in patients with either a previous history of BM infiltration or even overt marrow infiltration at the time of BM harvest. Because of their specificity, monoclonal antibodies are ideal agents for the selective elimination of malignant cells. Clinical studies have demonstrated that immunological purging can deplete malignant cells *in vitro* without significantly impairing hematological engraftment.

Assessing purging efficacy by PCR

PCR has been used to assess the efficacy of immunological purging in models using lymphoma cell lines, demonstrating that PCR is a highly sensitive and efficient method to determine the efficacy of purging residual lymphoma cells. The efficacy of purging varies between the cell lines studied, making it likely that there would also be variability between patient samples.

PCR amplifications of the t(14;18), t(11;14) and IgH rearrangements have all been used to detect residual lymphoma cells in the BM before and after purging in patients undergoing autologous BM transplantation to assess whether the efficiency of purging had any impact on disease-free survival. In one study, 114 patients with B-cell NHL and the *BCL-2* translocation were studied. Residual lymphoma cells were detected by PCR analysis in the harvested autologous BM of all patients. Following three cycles of immunological purging using anti-B-cell monoclonal antibodies and complement-mediated lysis, PCR amplification detected residual lymphoma cells in 50% of these patients. The incidence of relapse was significantly increased in the patients who had residual detectable lymphoma cells compared with those in whom no lymphoma cells were detectable after purging.

Detection of residual lymphoma cells in the marrow after transplantation is associated with increased incidence of subsequent relapse

Since PCR analysis detected residual lymphoma cells after conventional-dose chemotherapy in the majority of patients studied, it is not surprising that it has not been possible to determine any prognostic significance for the persistence of PCR-detectable lymphoma cells. At the Dana-Farber Cancer Institute, PCR analysis was performed on serial BM samples obtained after ASCT to assess whether high-dose therapy might be capable of depleting PCR-detectable lymphoma cells. The persistence or reappearance of residual detectable lymphoma cells had a great adverse influence on the disease-free survival of patients in this study after high-dose therapy. In contrast to previous findings that all patients had BM infiltration following conventional-dose therapy, no PCR-detectable lymphoma cells could be detected in the most recent BM sample obtained from more than 50% of patients following high-dose chemoradiotherapy and ASCT. A number of stud-

ies have now demonstrated that persistent detection of MRD by PCR following ASCT in patients with lymphoma identifies those patients who require additional treatment for cure, and also suggest that our therapeutic goal should be to eradicate all PCR-detectable lymphoma cells. In addition, quantitative PCR analysis has further shown that a rising tumor burden is a particularly poor prognosis feature.

Acute leukemias

Acute lymphoblastic leukemia

The treatment of childhood ALL has been one of the great success stories of modern chemotherapy and cure rates approaching 80% have been achieved in recently reported series. ALL cells usually rearrange either the IgH or TCR genes or both, and these provide markers that can be used to assess the clinical significance of MRD detection in a disease with such a high likelihood of cure. Despite near uniformity in approaches to the management of newly diagnosed ALL, the issue of whether eradication of PCR-detectable MRD is necessary for cure remains highly controversial in this disease. Most studies have suggested that modern aggressive induction regimens are often associated with rapid elimination of PCR-detectable disease. Additional studies have suggested an association between the rate of decrease of detectable disease and subsequent prognosis. However, a significant number of patients have persistence of PCR-detectable disease and yet have not relapsed. One study suggested that, using the most sensitive PCR analysis, it is possible to detect persistent cells bearing the associated antigen receptor rearrangement in the majority of patients, indicating that it is neither possible nor necessary to eradicate PCR-detectable disease for cure. For this reason, many current studies are addressing whether the quantitative assessment of MRD may be required to predict which patients are ultimately fated to relapse. Most studies addressing this issue have suggested that a quantitative increase in tumor burden is almost invariably associated with impending relapse.

t(12;21)

The *TEL/AML-1* gene rearrangement results from the cryptic reciprocal translocation t(12;21). This is the most common gene rearrangement found in childhood ALL and accounts for 25% of pre-B-cell ALL in children, but is rarely found in adult ALL. Some data are suggestive that the presence of this rearrangement is associated with a good prognosis. However, qualitative and quantitative PCR analysis studies have suggested that the persistence of residual leukemia cells or a slower rate of eradication of the leukemic cells is associated with a poorer prognosis.

APML: acute promyelocytic leukemia (AML M3)

AML M3 is associated with a balanced translocation between chromosomes 15 and 17, resulting in t(15;17)(q22;q21) and leading to rearrangement of the *RARα* gene (also termed *RARA*) on chromosome 17 and *PML* on chromosome 15 (Figure 6.10). With rearrangement of DNA, the chromosomal translocation produces two novel fusion genes involving *PML* and *RARα*, namely *PML/RARα* and *RARα/PML*. It is believed that *PML/RARα* is responsible for the development of aberrant hematopoiesis. There are two isoforms of the *PML/RARα* fusion gene: long and short. Patients who possess the short isoform have a poorer clinical outcome than those in whom the long isoform is found, but the exact mechanism involved is unclear at present.

The resultant fusion protein (PML/RARα) contains functional domains in both PML and RARα, and binds all-*trans* retinoic acid (ATRA), to which the leukemic cells in AML M3 are exquisitely sensitive. In fact ATRA, which induces differentiation of the leukemic cells, may alone achieve remission in 80% of *de novo* cases of AML M3. Two classes of retinoic acid receptor mediate the effects of retinoids: RAR and RXR, both of which are members of a superfamily of related ligand-inducible transcriptional regulatory factors. RAR (α, β and γ) are activated by ATRA and 9-*cis* retinoic acid. RXR (α, β and γ) is activated by 9-*cis* retinoic acid only.

Patients with APML and t(15;17) who achieve remission are now regarded as good-risk patients, with a 60% chance of achieving long-term remission. The presence of the fusion gene may be inferred from cytogenetic analysis (i.e. the presence of typical translocation) or, more recently, by an RT-PCR method. In this, the *PML/RARα* mRNA is reverse-transcribed into cDNA, which is then used for PCR detection of the abnormal transcript. The RT-PCR assay has been used to quantitate residual leukemic cells in patients with M3 undergoing chemotherapy.

Trial data suggest that persistence of t(15;17) determined by the PCR approach predicts outcome: those patients who fail to become PCR-negative or who become PCR-positive following a period of PCR negativity subsequently suffer overt clinical relapse. More recent data have suggested that quantitative PCR monitoring of *PML/RARα* can identify patients at high risk of relapse and suggest that clinically practical monitoring at more frequent intervals may improve predictive accuracy for relapse or continuing complete remission in many patients with persistent, fluctuating MRD levels.

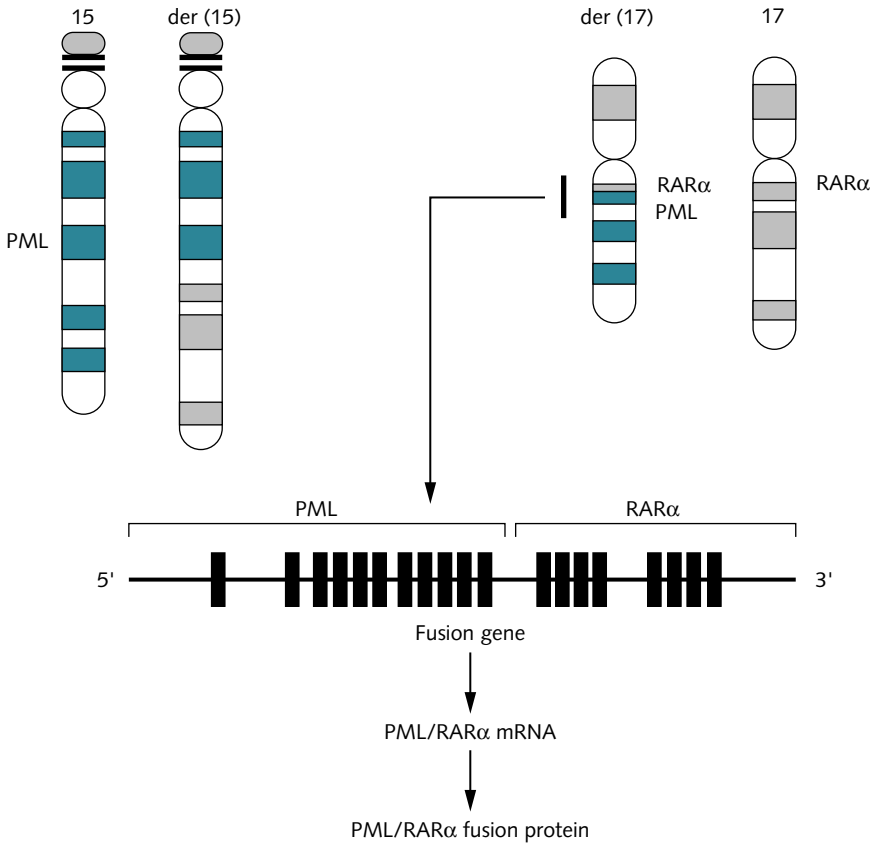


Fig. 6.10 t(15;17)(q22;q21) translocation
A balanced translocation involving the *RARα* gene (at 17q21) and the *PML* gene (15q22), found in AML M3 (APML) in >90% of cases. The chimeric PML/RARα protein plays a role in the differentiation block characteristic of APML.

Acute myelogenous leukemia

t(8;21)

The non-random chromosomal translocation t(8;21) occurs in up to 10% of *de novo* AML cases. It is more common in AML with features of maturation. This gene fuses the *AML* gene on 21q22 with the *ETO* gene on 8q22. The breakpoints in this translocation invariably occur within defined regions in the *AML* and *ETO* genes, resulting in a fairly uniform fusion product. Early studies of this translocation suggested that there was persistence of this transcript in almost all cases studied, even in patients in long-term remission. This suggests that the *AML1/ETO* translocation may be necessary, but in itself insufficient, for leukemic transformation. Recent studies using quantitative PCR analysis have suggested, however, that a quantitative increase in the fusion transcript is predictive of subsequent relapse.

Chronic leukemias

Chronic myeloid leukemia

Detection of t(9;22) by PCR amplification

The t(9;22), termed the Philadelphia chromosome, was described in 1960 by Nowell and Hungerford, and represented the first non-random chromosomal abnormality shown to be associated with a specific neoplasm, namely CML (although it is found in other disorders). The t(9;22) is formed by the fusion of the *BCR* gene on chromosome 22 with the *ABL* proto-oncogene on chromosome 9 and occurs in the vast majority of patients with CML and in up to 20% of adult patients with ALL. The chronic myeloid leukemic cells transcribe an 8.5-kb chimeric mRNA that is translated into a 210-kDa protein (p210) with tyrosine kinase activity. The breakpoints at the *ABL* gene can occur at any point up to 200

kb upstream in the intron and therefore cannot easily be amplified by PCR using genomic DNA as described earlier in this chapter. In contrast, the chimeric mRNA will usually be of two possible types. It is therefore possible to amplify the chimeric mRNA by first reverse-transcribing to cDNA. Using this technique, it is possible to detect one leukemic cell in up to 10^6 normal cells (*see Chapter 7*).

Detection of MRD after bone marrow transplantation in chronic myeloid leukemia

CML is incurable using standard chemotherapy, and allogeneic bone marrow transplantation (BMT) remains the treatment of choice for suitable patients. However, 20% of patients transplanted in the chronic phase and more than 50% of patients transplanted in the accelerated phase or blast crisis will relapse. Considerable effort has been made to establish whether persistence of MRD after allogeneic BMT is predictive of relapse. Early studies yielded conflicting results about the clinical implications of persistence of PCR-detectable disease. However, a recent large study from Seattle, including analysis of data from 346 patients, showed a clear association between the relapse and persistence of PCR-detectable disease. Detection of MRD early after BMT does not necessarily suggest a poor prognosis, and a PCR-positive sample 3 months after BMT was not informative for the clinical outcome. In contrast, a PCR-positive BM or peripheral blood sample at or after 6 months post-BMT was closely associated with subsequent relapse. Statistical analysis of the data revealed that the PCR assay for the *BCR-ABL* fusion transcript 6–12 months after BMT is an independent predictor of subsequent relapse. In contrast, no clear prediction of clinical outcome could be made in patients who tested PCR-positive more than 3 years after BMT. This study and others have clearly demonstrated that most patients are PCR-positive 3 months after BMT, indicating that BMT preparative regimens alone do not eradicate CML cells effectively. Nevertheless, since this treatment leads to cure in more than 50% of patients, other mechanisms, for example immunological mechanisms, must be responsible for tumor eradication.

Chronic lymphocytic leukemia

This is the commonest leukemia in adults and predominantly affects the elderly. A full description of CLL and its molecular abnormalities is provided in Chapter 10. Most are B-cell neoplasms (95%), which demonstrate a variety of cytogenetic abnormalities that are of value for molecular diagnosis and residual disease detection. Karyotypic abnormalities include trisomy 12 and deletions or translocations of chromosomes 11 and 13. Since these tumors are of B-cell origin, rearranged IgH genes may be used to confirm clonality and to detect residual tumor following chemotherapy and, in younger poor-risk patients, BMT.

Bone marrow transplantation is generally precluded in most patients with CLL due to the advanced age of the patients affected. However, recent studies of younger patients with aggressive disease who have undergone either autologous or allogeneic BMT have shown that PCR-detectable disease is often present at, or shortly after, transplantation, but this does not predict relapse. In the largest single-center study, the methods used for the analysis involved PCR amplification of the IgH locus with sequencing of the CDR3 products, before constructing patient-specific oligonucleotide probes, which were then used to probe the PCR products from marrow or blood samples taken after transplantation. Data suggest that patients who remain PCR-positive in the months following transplantation or become PCR-positive, having been PCR-negative initially, tend to relapse. Those that remain PCR-negative or become negative remain in clinical and morphological remission (Figure 6.11). Obviously, with an indolent, slow-growing disease like CLL, we must wait some years before the data can be interpreted fully, since it may be that ultimately all patients will relapse.

Problems with PCR analysis for detection of minimal residual disease

The major concern with PCR-based disease detection will always be the fear of false-positive results because of the ability of the technique to amplify even minute amounts of contaminating DNA. Unlike cell culture assays, it is not possible to determine whether cells detected by PCR are clonogenic (i.e. capable of division and causing relapse). Cells bearing a translocation may be committed progenitors incapable of further proliferation, or might have been sufficiently damaged by previous exposure to chemotherapy or radiotherapy

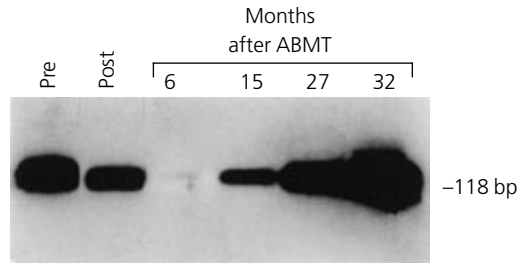


Fig 6.11 Detection of relapse of CLL in a patient undergoing autologous bone marrow transplantation using PCR
In the samples before (pre) and after (post) purging of the patient's marrow, PCR positivity is clearly seen. Six months after ABMT no PCR-detectable signal is seen. However, 15, 27 and 32 months after the transplant PCR positivity is easily detected. These findings were confirmed clinically and using standard morphological examination of the patient's bone marrow.

to be already dead but will still be detectable by PCR analysis. A potential problem with the use of PCR of the *BCL-2*/IgH translocation is that this translocation may not be specific for lymphoma cells. Cells bearing the translocation have been detected in hyperplastic lymphoid tissue in healthy individuals with no evidence of lymphoma, and more recently have been shown to occur rarely in normal B cells.

Conclusions

Methodologies have been developed for the sensitive detection of MRD in lymphoma and leukemia that are applicable to many patients. The question that now remains to be answered is whether these techniques will have any clinical utility and will predict which patients will relapse. In NHL these studies are most advanced in patients with t(14;18). In these patients, conventional-dose chemotherapy does not appear to be capable of depleting PCR-detectable lymphoma cells, although lymphoma cells were detectable in PB in only half of the patients studied. Following ASCT, the persistence or reappearance of PCR-detectable lymphoma cells in the BM was associated with an increased likelihood of relapse. In lymphomas that do not express the t(14;18), it is not yet clear whether failure to detect MRD in PB and BM will predict which patients will relapse since other subtypes of lymphoma may relapse in nodal sites without detectable lymphoma cells in the circulation.

From the available data, there are clearly diseases in which the persistence of PCR-detectable disease following treatment predicts relapse and others in which it does not. The full relevance of these findings will become clearer as we understand more about the biology of the diseases and additional data are generated as part of ongoing major clinical trials.

Further reading

Polymerase chain reaction

Kwok S, Higuchi R. (1989) Avoiding false positives with PCR. *Nature*, **339**, 237–238.

Saiki RK, Gelfand DH, Stoffel S *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491.

Antigen receptor genes

Aisenberg AC. (1993) Utility of gene rearrangements in lymphoid malignancies. *Annual Review of Medicine*, **44**, 75–84.

Griesser H, Tkachuk D, Reis MD *et al.* (1989) Gene rearrangements and translocations in lymphoproliferative diseases. *Blood*, **73**, 1402–1415.

Tonegawa S. (1983) Somatic generation of antibody diversity. *Nature*, **302**, 575–581.

Toyonaga B, Mak TW. (1987) Genes of the T-cell antigen receptor in normal and malignant T cells. *Annual Review of Immunology*, **5**, 585–620.

Acute leukemias

Cave H, van der Werff ten Bosch J, Suciu S *et al.* (1998) Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. European Organization for Research and Treatment of Cancer—Childhood Leukemia Cooperative Group. *New England Journal of Medicine*, **339**, 591–598.

Diverio D, Rossi V, Avvisati G *et al.* (1998) Early detection of relapse by prospective reverse transcriptase-polymerase chain reaction analysis of the PML/RARα fusion gene in patients with acute promyelocytic leukemia enrolled in the GIMEMA-AIEOP multicenter 'AIDA' trial. *Blood*, **92**, 784–789.

Gameiro P, Moreira I, Yetgin S *et al.* (2002) Polymerase chain reaction (PCR)- and reverse transcription PCR-based minimal residual disease detection in long-term follow-up of childhood acute lymphoblastic leukaemia. *British Journal of Haematology*, **119**, 685–696.

Mancini M, Nanni M, Cedrone M *et al.* (1995) Combined cytogenetic, FISH and molecular analysis in acute promyelocytic leukemia at diagnosis and in complete remission. *British Journal of Haematology*, **91**, 878–884.

Marcucci G, Livak KJ, Bi W *et al.* (1998) Detection of minimal residual disease in patients with AML1/ETO-associated acute myeloid leukemia using a novel quantitative reverse transcription polymerase chain reaction assay. *Leukemia*, **12**, 1482–1489.

Sykes PJ, Brisco MJ, Hughes E *et al.* (1998) Minimal residual disease in childhood acute lymphoblastic leukemia quantified by aspirate and trephine: is the disease multifocal? *British Journal of Haematology*, **103**, 60–65.

van Dongen JJ, Seriu T, Panzer-Grumayer ER *et al.* (1998) Prognostic value of minimal residual disease in acute lymphoblastic leukemia in childhood. *Lancet*, **352**, 1731–1738.

Yamada M, Wasserman R, Lange B *et al.* (1990) Minimal residual disease in childhood B-lineage lymphoblastic leukemia. *New England Journal of Medicine*, **323**, 448–455.

Chronic leukemias

Bose S, Deininger M, Gora-Tybor J *et al.* (1998) The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood*, **92**, 3362–3367.

Khouri IF, Keating MJ, Vriesendorp HM *et al.* (1994) Autologous and allogeneic bone marrow transplantation for chronic lymphocytic leukemia: preliminary results. *Journal of Clinical Oncology*, **12**, 748–758.

Nowell PC, Hungerford DA. (1960) A minute chromosome in human granulocytic leukemia. *Science*, **132**, 125–132.

Provan D, Bartlett-Pandite L, Zwicky C *et al.* (1996) Eradication of polymerase chain reaction-detectable chronic leukemia cells is associated with improved outcome after bone marrow transplantation. *Blood*, **88**, 2228–2235.

Rabinowe SN, Soiffer RJ, Gribben JG *et al.* (1993) Autologous and allogeneic bone marrow transplantation for poor prognosis patients with B-cell chronic lymphocytic leukemia. *Blood*, **82**, 1366–1376.

Radich JP, Gehly G, Gooley T *et al.* (1995) Polymerase chain reaction detection of the BCR-ABL fusion transcript after allogeneic marrow transplantation for chronic myeloid leukemia: results and implications in 346 patients. *Blood*, **85**, 2632–2638.

Non-Hodgkin's lymphoma

Corradini P, Tarella C, Olivieri A *et al.* (2002) Reduced-intensity conditioning followed by allografting of hematopoietic cells can produce clinical and molecular remissions in patients with poor-risk hematologic malignancies. *Blood*, **99**, 75–82.

Gribben JG, Neuberg D, Freedman AS *et al.* (1993) Detection by polymerase chain reaction of residual cells with the bcl-2 translocation is associated with increased risk of relapse after autologous bone marrow transplantation for B-cell lymphoma. *Blood*, **81**, 3449–3457.

Gribben JG, Neuberg DN, Barber M *et al.* (1994) Detection of residual lymphoma cells by polymerase chain reaction in peripheral blood is significantly less predictive for relapse than detection in bone marrow. *Blood*, **83**, 3800–3807.

Lopez-Guillermo A, Cabanillas F, McLaughlin P *et al.* (1998) The clinical significance of molecular response in indolent follicular lymphomas. *Blood*, **91**, 2955–2960.

Sharp JG, Joshi SS, Armitage JO *et al.* (1992) Significance of detection of occult non-Hodgkin's lymphoma in histologically uninvolved bone marrow by culture technique. *Blood*, **79**, 1074–1080.

Taniwaki M, Nishida K, Ueda Y *et al.* (1995) Interphase and metaphase detection of the breakpoint of 14q32 translocations in B-cell malignancies by double-color fluorescence *in situ* hybridization. *Blood*, **85**, 3223–3228.

Weiss LM, Warnke RA, Sklar J *et al.* (1987) Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas. *New England Journal of Medicine*, **317**, 1185–1189.

Yuan R, Dowling P, Zucca E *et al.* (1993) Detection of bcl-2/JH rearrangement in follicular and diffuse lymphoma: concordant results of peripheral blood and bone marrow analysis at diagnosis. *British Journal of Cancer*, **67**, 922–925.

Zwicky CS, Maddocks AB, Andersen N *et al.* (1996) Eradication of polymerase chain reaction detectable immunoglobulin gene rearrangement in non-Hodgkin's lymphoma is associated with decreased relapse after autologous bone marrow transplantation. *Blood*, **88**, 3314–3322.

Quantitation

Cross NC, Feng L, Chase A *et al.* (1993) Competitive polymerase chain reaction to estimate the number of BCR-ABL transcripts in chronic myeloid leukemia patients after bone marrow transplantation. *Blood*, **82**, 1929–1936.

Donovan JW, Ladetto M, Poor C *et al.* (2000) Immunoglobulin heavy chain consensus probes for real-time PCR quantification of residual disease in acute lymphoblastic leukemia. *Blood*, **95**, 2651–2658.

Mensink E, van de Locht A, Schattenberg A *et al.* (1998) Quantitation of minimal residual disease in Philadelphia chromosome positive chronic myeloid leukemia patients using real-time quantitative RT-PCR. *British Journal of Haematology*, **102**, 768–774.

Pongers-Willemse MJ, Verhagen OJ, Tibbe GJ *et al.* (1998) Real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia using junctional region specific TaqMan probes. *Leukemia*, **12**, 2006–2014.

Chapter 7 Chronic myeloid leukemia

Brian J Druker

Introduction

Chronic myeloid leukemia (CML) was the first hematological malignancy to be associated with a specific chromosome abnormality. This abnormality, known as the Philadelphia chromosome, results from a reciprocal translocation between chromosomes 9 and 22. This finding led to the identification of *BCR-ABL* as the causative molecular event of this disease and ultimately to the development of a therapy that targets this pathogenetic event. In this chapter, the clinical and molecular aspects of CML will be reviewed along with molecular methods of detection and monitoring of the disease. Treatment options are described, with emphasis on molecularly targeted therapy for this disease. Lastly, the signal transduction pathways affected by the BCR-ABL fusion protein will be discussed.

Clinical features

CML is a malignant, clonal hematopoietic stem cell disorder. It accounts for 15–20% of all cases of leukemia, with an annual incidence of 1–1.5 cases per 100 000. Although CML affects all age groups, the median age at diagnosis for patients not selected by referral is close to 60 years. Clinically, the course of CML is divided into three phases. Most patients will present in the chronic or stable phase of the disease, which, after a variable length of time, progresses through an accelerated phase to an invariably fatal acute leukemia, also known as 'blast crisis'.

Patients with CML in the chronic phase may be asymptomatic and up to 50% of patients are identified through routine blood tests. Common presenting symptoms include fatigue, night sweats, and splenomegaly with abdominal discomfort and early satiety. Occasionally, patients may present with a hyperviscosity syndrome, with manifestations such as stroke, priapism, stupor, or visual changes caused by retinal hemorrhage.

The white blood cell count in stable-phase CML is usually more than 50×10^9/L at the time of diagnosis, with a range of 20×10^9 to 800×10^9/L. Anemia, proportional to the degree of leukocytosis, is present at diagnosis in 50% of cases of CML. The platelet count is elevated in up to half of patients. During the chronic phase, leukemic cells retain the capacity to differentiate normally with the peripheral blood smear, showing a full spectrum of myeloid cells from blasts to neutrophils, with blasts comprising less than 5% of the white blood cell differential. Basophilia is invariable present and eosinophilia is common. The bone marrow of untreated chronic-phase patients shows hypercellularity with granulocytic and megakaryocytic hyperplasia, basophilia and fewer than 5% blasts.

After a median of 4–5 years in untreated patients, there is a transition from chronic-phase CML to an accelerated or blastic phase. The accelerated phase is a transitional period that is poorly defined. Patients may have progressive splenomegaly, bone pain and/or constitutional complaints such as fever, night sweats, or weight loss. Frequently, the accelerated phase is marked by difficulty in controlling the white blood cell count with standard doses of chemotherapy. Laboratory hallmarks of the accelerated phase that have been correlated with survival of less than 18 months include an increase in the percentage of blasts in the peripheral blood to at least 15%, an increase in basophilia to over 20%, the presence of more than 30% blasts plus promyelocytes in the peripheral blood, or a platelet count less than 100×10^9/L. Approximately 20–40% of patients with CML progress to the blastic phase without an intervening accelerated phase.

The transition of CML to the blastic phase is accompanied by the loss of the capacity of the malignant clone for terminal differentiation. Morphologically, the bone marrow of blastic-phase CML resembles acute leukemia, with more than 30% blasts. Approximately 65% of patients evolve to blastic crisis with myeloid blasts, 30% have lymphoid blast crisis, and 5% of cases have biphenotypic, undifferentiated or T-cell blasts. Blastic-phase CML responds poorly to cytotoxic chemotherapy. Median survival of patients in blastic phase is 3–6 months,

although patients with lymphoid blast crisis have a somewhat better outcome.

Staging and prognosis

Several pretreatment clinical characteristics have been found to have prognostic significance in CML. The prognostic scoring system proposed by Sokal reproducibly segregates chemotherapy-treated patients into high- and low-risk groups for disease progression. Sokal's model identified four independent prognostic factors: (1) older age; (2) splenomegaly; (3) higher platelet count; and (4) higher peripheral blast percentage. Sokal's index is less efficient in discriminating outcome in interferon-α-treated patients. However, a revised score, the Euro score, which also incorporates peripheral blood eosinophils and basophils, can identify risk groups in interferon-α-treated CML patients. Whether either of these scores will discriminate among outcomes for patients treated with imatinib is unknown.

More recently, there has been increasing interest in using microarray studies to assist in defining risk groups as it is clear that there is significant heterogeneity among patients who present with the same stage of disease. This heterogeneity could be based on the individual genetics of the particular leukemia or on the genes involved in drug metabolism and drug resistance. Radich and colleagues compared a pool of blast crisis samples with a pool of chronic-phase samples and found approximately 500 genes that are significantly different between the two disease states. Blast crisis and chronic phase patients demonstrate clear differences in gene expression, and occasional cases occur in which the clinical and pathological diagnosis is quite discordant from the gene expression pattern.

Molecular pathogenesis of CML

Molecular anatomy of BCR-ABL translocations

In 1960, Nowell and Hungerford described a consistent chromosomal abnormality in CML patients, an acrocentric chromosome that was thought to represent a chromosomal deletion. This was the first example of a chromosomal abnormality linked to a specific malignancy. As chromosomal banding techniques improved, it became apparent that the abnormality was a shortened chromosome 22. Rowley later clarified that the shortened chromosome, the so-called Philadelphia chromosome, was the product of a reciprocal translocation between the long arms of chromosomes 9 and 22, t(9:22)(q34;q11) (Figure 7.1). The molecular consequence of this translocation is the fusion of the *ABL* gene from chromosome 9 to sequences on chromosome 22, the breakpoint cluster region (*BCR*), giving rise to a chimeric *BCR-ABL* gene.

The breakpoints within the *ABL* gene at 9q34 can occur anywhere over a large area (greater than 300 kb) at its 5′ end—upstream of the first alternative exon Ia, downstream of the second alternative exon Ib, or, more frequently, between the two (Figure 7.2). Regardless of the exact location of the breakpoint, splicing of the primary hybrid transcript yields an mRNA molecule in which *BCR* sequences are fused to the second exon of *ABLI*, exon a2. In the majority of patients with CML and in approximately one-third of patients

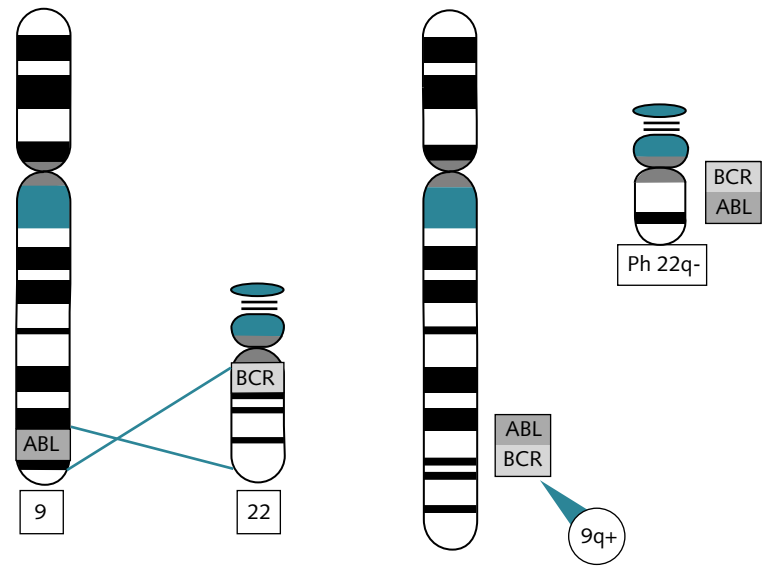


Fig. 7.1 Schematic diagram of the translocation that creates the Philadelphia chromosome

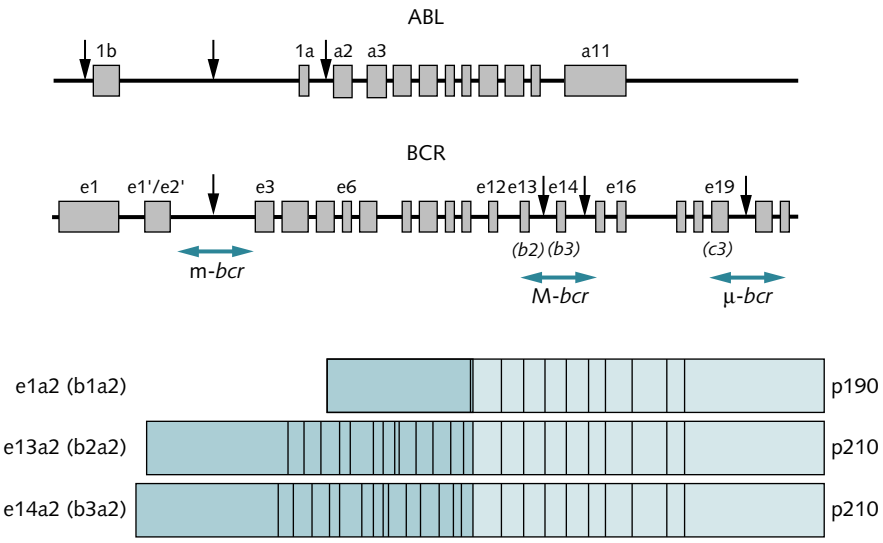


Fig. 7.2 Structure of the *BCR* and *ABL* genes, showing locations of the breakpoints and various mRNAs created

with Philadelphia chromosome-positive acute lymphoblastic leukemia (ALL), the break occurs within a 5.8-kb area spanning *BCR* exons 13 and 14, originally referred to as exons b2 and b3, and also known as the major breakpoint cluster region (M-*BCR*). Depending of the site of the translocation, fusion transcripts with either *BCR* exon b2 or b3 fused to *ABL* exon a2 can be formed. Each of these fusion mRNAs is transcribed into a 210-kDa chimeric protein ($p210^{BCR-ABL}$). In the remaining patients with ALL and rarely in patients with CML, characterized clinically by prominent monocytosis, the breakpoints are further upstream in the region between *BCR* exons 1 and 2, also known as the minor breakpoint cluster region (m-*BCR*). The resulting e1a2 mRNA is translated into a 190-kDa protein ($p190^{BCR-ABL}$). Recently, a third breakpoint cluster region (μ-*BCR*) has been identified downstream of exon 19, giving rise to a 230-kDa fusion protein ($p230^{BCR-ABL}$) associated with some, but not all, cases of the rare Philadelphia chromosome-positive chronic neutrophilic leukemia.

Animal models of CML

Several experimental approaches have demonstrated the ability of *BCR-ABL* to cause leukemia. In one set of experiments, transgenic mice that express *BCR-ABL* were shown to develop a rapidly fatal acute leukemia. Using a different approach, a *BCR-ABL*-expressing retrovirus was used to infect murine bone marrow. These *BCR-ABL*-expressing marrow cells were used to repopulate irradiated mice. The transplanted mice developed a variety of myeloproliferative disorders, including a CML-like syndrome. Although these approaches demonstrate the leukemogenic potential of *BCR-ABL*, it is possible that secondary changes are required for leukemia to develop. Recently, C. Huettner and colleagues placed *BCR-ABL* under the control of a tetracycline-repressible promoter. Mice expressing this transgene develop a reversible leukemia dependent on the presence or absence of tetracycline, thus demonstrating the leukemic potential of *BCR-ABL* as a sole oncogenic abnormality.

BCR-ABL signaling and CML pathogenesis

BCR-ABL functions as a constitutively activated tyrosine kinase and significant advances have been made in determining the signaling pathways that are activated by BCR-ABL kinase activity (Figure 7.3). Numerous substrates and binding partners have been identified and current efforts are directed at linking these pathways to the specific pathological defects that characterize CML. The pathological defects identified in CML cells include increased proliferation or decreased apoptosis of a hematopoietic stem or progenitor cell, leading to a massive increase in myeloid cell numbers. Since patients have circulating immature myeloid progenitors, it has been postulated that there is a defect in the adherence of myeloid progenitors to marrow stroma. An example of a cellular pathway that links to an increased proliferative rate is activation of the RAS pathway. STAT-5-mediated upregulation of the anti-apoptotic molecule BCL_{XL} and the phosphorylation and inactivation of the pro-apoptotic molecule BAD by AKT are postulated to lead to a protection from programmed cell death. CML cells also exhibit reduced adhesion to fibronectin, possibly as a downstream effect of CRKL phosphorylation. Despite the seemingly endless expansion of the list of pathways activated by BCR-ABL and the increasing complexity that is being revealed in these pathways, all of the transforming functions of BCR-ABL are dependent on its tyrosine kinase activity.

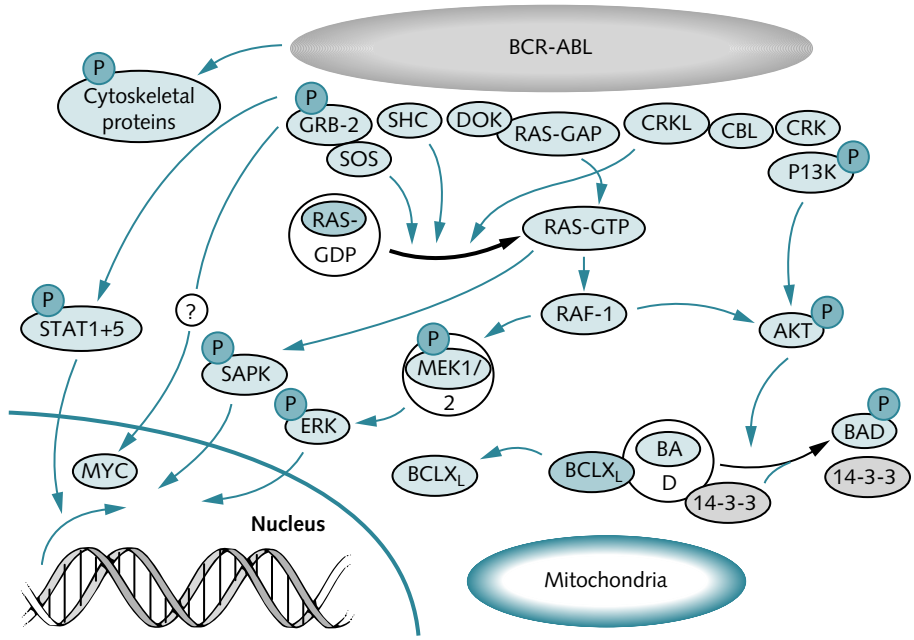


Fig. 7.3 Signaling pathways affected by BCR-ABL expression

Treatment options for CML

The treatment of CML has rapidly evolved. Currently, the only curative therapy for CML is allogeneic stem cell transplantation. Non-transplant options include busulfan, hydroxyurea, interferon-α, and imatinib. Each of these therapies will be reviewed briefly.

Conventional chemotherapy

Either hydroxyurea or busulfan as single agents can provide hematological control in the majority of patients (>75%) with chronic-phase CML. These hematological 'remissions' are characterized by persistence of Philadelphia chromosome-positive cells in the marrow, with inevitable progression to the blastic phase. Suppression of Philadelphia chromosome positivity has been observed occasionally during treatment with busulfan or hydroxyurea, usually in the setting of therapy-induced myelosuppression, and is transient. A prospective, randomized trial comparing these two agents in chronic-phase CML found a significantly prolonged median duration of the chronic phase in hydroxyurea-treated patients: 47 months versus 37 months with busulfan ($P = 0.04$). Overall survival was also superior in hydroxyurea-treated patients, with a median survival of 58 months versus 45 months ($P = 0.008$) in the busulfan group. Additionally, hydroxyurea has a lower toxicity profile than busulfan.

Interferon-α

Interferon-α is a member of a family of glycoproteins that have antiviral and antiproliferative properties. It was first shown to be an active agent in CML in the early 1980s. Although interferon-α has significant activity in early chronic-phase patients, it has only modest activity in late chronic-phase CML and minimal activity in the accelerated or blastic phase. In several large, prospective, randomized trials, interferon-α increased survival when administered in the chronic phase, and a meta-analysis of several major trials showed a significant survival advantage for interferon-α over hydroxyurea and busulfan. In this meta-analysis, 5-year survival rates were 57 and 42% for interferon-α and chemotherapy, respectively. A complete hematological response (CHR) is seen in most (70%) patients treated with interferon-α. A cytogenetic response, defined as a decrease in the percentage of Philadelphia chromosome-positive metaphases in the marrow to less than 35%, is seen in 30–50% of patients, with complete cytogenetic responses in only 10–20% of interferon-treated patients. This group of patients with complete cytogenetic responses gains the greatest survival advantage and has a median survival of more than 8 years. Up to 20% of patients tolerate interferon-α poorly, necessitating discontinuation of treatment. These side effects include fevers, chills, arthralgias, myalgias, fatigue, depression, weight loss, peripheral neuropathy, myelosuppression and hepatotoxicity.

Investigators have sought to improve on the success of treatment with interferon-α by adding cytosine arabinoside (also called cytarabine or Ara-C), another agent with antileukemic activity. One randomized trial of the combination showed significantly improved response rates over interferon-α alone, which translated into an overall survival advantage. A subsequent study found improved response rates but no survival advantage for the combination. In addition, the combination of interferon-α and cytarabine is associated with increased gastrointestinal and marrow toxicity and, not surprisingly, many patients tolerate it poorly. Another potential problem with the use of interferon-α is that it may compromise the outcome of subsequent transplantation owing to an increase in graft rejection or graft-versus-host disease (GVHD). However, there are conflicting data on this issue and many recommend discontinuing interferon-α at least 3 months before transplantation.

Imatinib: molecular biology to therapy

From the above discussion, it should be clear that the BCR-ABL possesses many characteristics of an ideal therapeutic target. It is expressed in the majority of patients with CML and it has been shown to be the cause of CML. BCR-ABL functions as a constitutively activated tyrosine kinase and mutagenic analysis has shown that this activity is essential for the transforming function of the protein. Thus, an inhibitor of the BCR-ABL kinase would be predicted to be an effective and selective therapeutic agent for CML (Figure 7.4).

Imatinib mesylate (Gleevec, Glivec, formerly STI571) is a relatively specific inhibitor of the BCR-ABL tyrosine kinase. Other tyrosine kinases inhibited by imatinib include the platelet-derived growth factor receptor KIT, and ARG (*ABL*-related gene). Preclinical data showed significant specific activity against *BCR-ABL*-expressing cell lines *in vitro* and *in vivo*. In addition, imatinib could select for the growth of *BCR-ABL*-negative hematopoietic cells from CML patient samples in colony-forming assays and long-term marrow cultures.

Based on the favorable preclinical data, imatinib was tested in Phase I and II clinical trials in CML patients. These studies

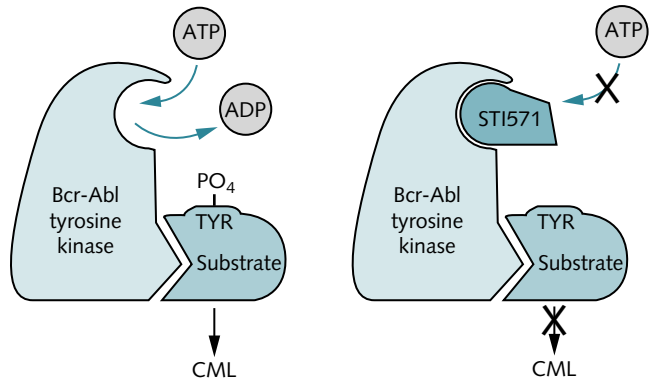


Fig. 7.4 Schematic representation of the mechanism of action of the BCR-ABL tyrosine kinase and its inhibition by imatinib

showed that imatinib has significant activity in all phases of CML, and the results of the Phase II trials are summarized in Table 7.1. In chronic-phase patients who had failed interferon, 96% achieved a CHR with an imatinib dose of 400 mg per day. Imatinib induced major cytogenetic responses (≤35% Philadelphia chromosome-positive metaphases) in 64% of patients, with a complete cytogenetic response rate of 48%. The estimated progression-free survival at 24 months was 87%. Cytogenetic responses have been durable and correlate with improved progression-free and overall survival. Thus, once a patient achieves a major cytogenetic response, it is estimated that, 24 months later, 91% of these patients will not have progressed. Achievement of a major cytogenetic response was associated with a statistically significant improvement in overall survival. For example, if patients achieve a major cytogenetic response within 12 months, the estimated survival at 24 months was 99%, compared with 86% for patients with less than a major cytogenetic response ($P < 0.001$). Baseline features that independently predicted a high rate of major cytogenetic responses were the absence of blasts in the peripheral blood, a hemoglobin concentration above 12 g/dl, fewer than 5% blasts in the marrow, CML disease duration of less than 1 year, and a prior cytogenetic response to interferon-α.

Table 7.1 Phase II results with imatinib.

	Chronic phase (IFN failure)	Accelerated phase	Blast crisis
CHR	96%	40%	9%
MCR	64%	28%	16%
CCR	48%	20%	7%
Disease progression*	13%	50%	90%

IFN, interferon; CHR, complete hematological response; MCR, major cytogenetic response (Philadelphia chromosome-positive metaphases ≤35%); CCR, complete cytogenetic response; *at 24 months.

In the Phase II trial in the accelerated phase, patients were required to have 15–30% blasts or more than 30% blasts plus promyelocytes in the peripheral blood or marrow, greater than 20% peripheral basophils, or a platelet count less than 100×10^9/L, unrelated to therapy. With follow-up of up to 3 years, 83% of patients showed some form of hematological response, 40% of patients achieving a CHR. Twenty-eight percent of patients achieved a major cytogenetic response, with 20% complete responses. In this study, significantly improved outcomes for responses and survival were observed for patients treated with 600 mg per day of imatinib compared with patients treated with 400 mg per day.

In patients with myeloid blast crisis, the overall response rate to imatinib was 52%, with sustained hematological responses lasting at least 4 weeks in 31% of patients. Nine percent of patients achieved complete remission (fewer than 5% blasts) with peripheral blood recovery and another 4% of patients cleared their marrows to less than 5% blasts but did not meet the criteria for complete remission because of persistent cytopenia. Finally, 18% of patients either returned to the chronic phase or had partial responses. Major cytogenetic responses were seen in 16% of patients, 7% having complete responses. Median survival was 6.9 months, with an estimated survival of 17% at 24 months. The baseline features predictive of prolonged survival were a platelet count at least 100×109/L, peripheral blood blasts fewer than 50% and a hemoglobin concentration of at least 10 g/dl. For patients with all three of these features, the median survival was 21 months. However, if patients had none of these features their median survival was only 4 months. Patients with a CHR or marrow blasts less than 5% at 2 months had a median survival of more than 24 months and their survival was significantly longer than that of patients who either returned to the chronic phase or had no response. These results with single-agent imatinib compare favorably with historical controls treated with chemotherapy for myeloid blast crisis, in which the median survival is approximately 3 months. However, the high relapse rates suggest either that imatinib should be viewed as a bridge to allogeneic stem cell transplantation or that patients should be enrolled in clinical trials combining imatinib with other agents.

In patients with Philadelphia chromosome-positive ALL, 29/48 (60%) responded to single-agent imatinib. Unfortunately, the duration of response was relatively short, with a median estimated time to disease progression of only 2.2 months.

A Phase III randomized study, comparing imatinib at 400 mg per day with interferon-α plus Ara-C in newly diagnosed patients with chronic-phase CML, enrolled 1106 patients between June 2000 and January 2001. Five hundred fifty-three patients were randomized to each treatment. Baseline characteristics were well balanced for all features evaluated, including age, white blood count, Sokal and Euro score, and time from diagnosis. With a median follow-up of 19 months, patients randomized to imatinib had statistically significant better results than patients treated with interferon-α plus Ara-C in all parameters measured (Table 7.2) including rates of CHR, major and complete cytogenetic responses, tolerance of therapy, and freedom from disease progression. Despite the fact that 76% of patients randomized to imatinib achieved a complete cytogenetic response, the majority of these patients had detectable leukemia, as analyzed by reverse transcriptase–polymerase chain reaction (RT-PCR) for *BCR-ABL* transcripts. When analyzed by log reduction in *BCR-ABL* transcript levels, 39% of patients achieved at least a three-log reduction in *BCR-ABL* levels, but only 13 and 3% achieved a four- and five-log reduction, respectively.

Table 7.2 Phase III results of imatinib versus interferon-α plus cytarabine (Ara-C) for newly diagnosed chronic phase CML patients.

	Imatinib 400 mg	Interferon + Ara-C
CHR	97%	69%
MCR	87%	35%
CCR	76%	14%
Intolerance	3%	31%
Progressive disease	3%	8.5%

Ara-C, cytosine arabinoside; CHR, complete hematological response; MCR, major cytogenetic response (Philadelphia chromosome-positive metaphases <35%); CCR, complete cytogenetic response. 'Intolerance' means intolerance leading to discontinuation of first-line therapy. 'Progressive disease' indicates progression to accelerated phase or blast crisis. All differences are highly statistically significant ($P < 0.001$). Median follow up is 18 months.

Resistance to imatinib therapy

In evaluating relapse mechanisms, patients can be separated into two categories: those with persistent inhibition of the BCR-ABL kinase (BCR-ABL-independent) and those with reactivation of the BCR-ABL kinase (BCR-ABL-dependent) at relapse (Figure 7.5). In the largest studies of resistance or relapse, several consistent themes emerge. In patients with primary resistance, i.e. patients who do not respond to imatinib therapy, BCR-ABL-independent mechanisms are most common. In contrast, most of the patients who relapse on therapy with imatinib reactivate the BCR-ABL kinase. In these studies, more than 50% and perhaps as many as 90% of patients with hematological relapse have *BCR-ABL* point mutations in at least 13 different amino acids scattered throughout the ABL kinase domain (Figure 7.6). Other patients have amplification of *BCR-ABL* at the genomic or transcript level.

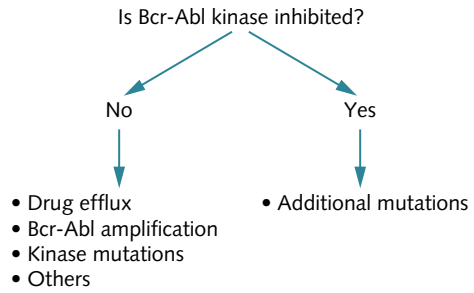


Fig. 7.5 Distinguishing between potential mechanisms of resistance to imatinib

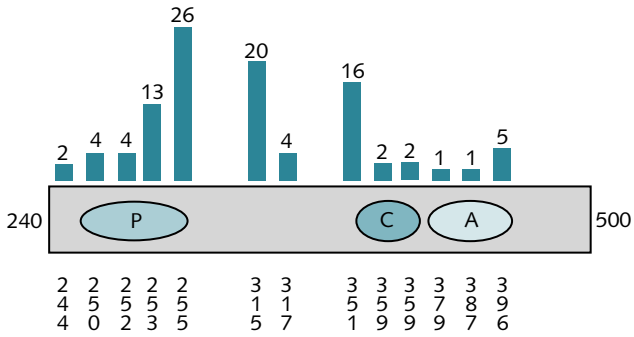


Fig. 7.6 ABL kinase domain mutations that have been identified that are associated with resistance to imatinib

The ABL kinase domain, from amino acid 240 to amino acid 500, is shown with the ATP binding domain (P), the catalytic domain (C) and the activation loop (A). The numbers below the kinase domain are amino acids that are mutated in patients who relapsed on therapy with imatinib. The vertical lines above the kinase domain indicate the number of times each amino acid has been found to be mutated (compiled from various series).

It should be noted that the studies described above represent a minority of CML patients. Most of the patients described in the studies of resistance had advanced phase disease. In contrast, the majority of patients diagnosed with CML will be in the chronic phase, most will obtain a complete cytogenetic response with imatinib, and very few have relapsed. However, only a minority attains molecular remission. Current hypotheses for the mechanism of molecular persistence can also be divided into BCR-ABL-dependent and -independent mechanisms. For example, stem cell quiescence has been postulated as a potential BCR-ABL-independent mechanism of resistance to imatinib. BCR-ABL-dependent mechanisms of molecular persistence include the possibility that low levels of BCR-ABL kinase activity prevent cells from proliferating, but are sufficient to protect cells from apoptosis. This could be due to imatinib not being capable of completely inhibiting the BCR-ABL kinase or due to the fact that stem cells express high levels of P-glycoprotein that result in efflux of imatinib. Each of these scenarios suggests specific therapeutic interventions to overcome molecular persistence.

Allogeneic stem cell transplantation

Allogeneic stem cell transplantation, using myeloablative doses of chemotherapy and/or radiation followed by infusion of allogeneic stem cells, remains the only proven curative therapy for CML. Stem cells used for allogeneic transplantation may be obtained from either the bone marrow or the peripheral blood of the donor. Allogeneic stem cell transplantation in CML has a long-term survival rate of 55–80%, with median survival of more than 10 years for patients transplanted during chronic phase with non-T-cell-depleted marrow, using modern GVHD and cytomegalovirus prophylaxis. Results with allogeneic stem cell transplantation are much better for patients in chronic-phase CML than for those in the accelerated or blastic phase of the disease. Long-term overall survival rates are 15–20% for patients transplanted during the accelerated phase and below 10% for patients transplanted during the blastic phase.

The major causes of treatment failure are relapse and transplant-related mortality. Relapse rates are lowest for chronic-phase patients (10–20%), intermediate for patients transplanted in the accelerated phase (approximately 40%), and highest for patients allografted during the blastic phase, for whom relapse rates in excess of 60% are reported. The majority of relapses occur within the first 3 years after transplantation, though late relapses can occur.

Allogeneic stem cell transplantation is associated with high early rates of morbidity and mortality due to regimen-related toxicity, GVHD, and infectious complications. This early mortality is highly dependent on patient age and the degree of graft mismatch. Transplant-related mortality also increases with the interval from diagnosis of CML and with the stage of disease. Patients younger than 20 years with an HLA-matched sibling donor have a transplant-related mortality of 5–10%, with a long-term disease survival of 60–80% in most studies. Patients over the age of 40 have a risk of transplant-related mortality double that of patients under age 20. However, data from the Fred Hutchinson Cancer Research Center show that selected patients in the 50- to 60-year age group may have a disease-free survival at 4 years of approximately 80%.

The widespread application of stem cell transplantation has been limited by donor availability and the high toxicity of the procedure in older patients. Ongoing advances in alternative donor sources (including unrelated donors and, to a lesser degree, cord blood), more accurate molecular HLA typing, and less toxic regimens, including reduced intensity or non-myeloablative transplants, are broadening the potential use of this treatment modality.

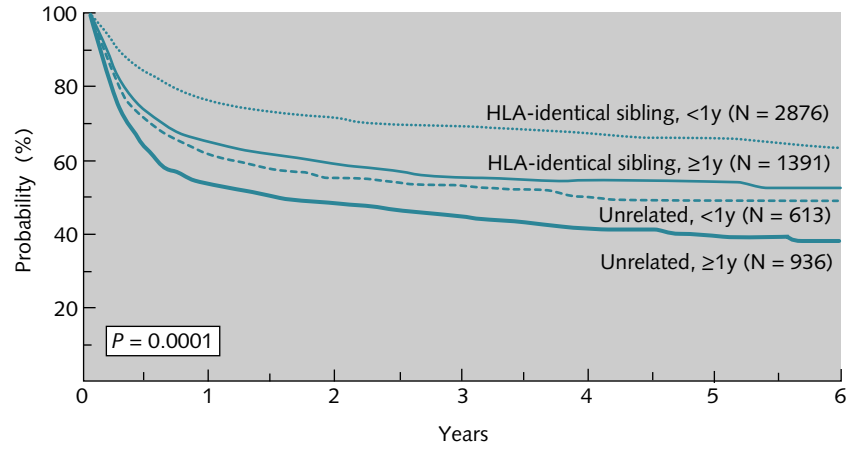


Fig. 7.7 The IBMTR (International Bone Marrow Transplant Registry) experience of CML chronic-phase patients
The curves demonstrate both the effect of a related versus unrelated donor and the benefit of early transplantation for survival.

CML is the most frequent indication for an unrelated donor transplant because of the lethality of the disease and the fact that disease progression is typically slow enough to allow adequate time for identification of a suitable donor. The use of volunteer unrelated donor transplants is associated with a higher incidence of viral infections, extensive chronic GVHD (up to 50%) and engraftment failure (16%) compared with related donor transplants. Recent reports show disease-free survival rates of 57–64% at 4 or 5 years after unrelated donor transplantation. In highly selected patients under age 50 who are transplanted within 1 year of diagnosis, a 74% survival rate at 5 years has been reported.

T-cell depletion of the unrelated allograft has been used in an effort to decrease the incidence of GVHD. Although this strategy does effectively reduce the incidence and severity of GVHD, it results in increased graft failure and relapse.

Non-myeloablative approaches use less intensive conditioning therapy than the standard allogeneic stem cell transplantation. The primary aim of non-myeloablative approaches is the induction of sufficient immunosuppression in the recipient to allow durable engraftment of both donor stem cells and donor lymphocytes, which can then exert a potent graft-versus-leukemia effect. However, GVHD remains a major limiting factor in this approach. It is not clear which patients should undergo a standard allogeneic transplant as opposed to a non-myeloablative transplant.

Post-transplant relapses

In the event of relapse after allogeneic bone marrow transplantation, donor lymphocyte infusions have resulted in obliteration of the malignant clone with long-lasting hematological and cytogenetic remissions. In the largest series, 72% of 75 evaluable patients were reported to achieve a complete cytogenetic response following post-transplant relapse, and the 3-year survival after donor lymphocyte infusions was 67%. Disease stage at relapse is fundamental to the prediction of outcome. When relapse is detected by cytogenetic analysis or molecular analysis, donor lymphocyte infusion is typically effective, while patients with hematological relapse are less likely to respond. Donor lymphocyte infusion is believed to work via immune modulation, in which the donor T cells induce a graft-versus-leukemia effect that allows the normal cells to re-expand as the malignant clone is destroyed. Acute GVHD and myelosuppression complicate this procedure and are associated with a 1-year mortality of 18%. These complications may be reduced by adjusting the dose and the subsets of lymphocytes infused.

Imatinib has also shown significant activity in patients who relapse after transplantation. However, there are insufficient data to determine whether imatinib or donor lymphocyte infusions should be the preferred therapy for this patient population.

Molecular diagnosis and monitoring of CML

The traditional method of diagnosing CML and monitoring disease status is the cytogenetic analysis of bone marrow-derived metaphases for the presence of the Philadelphia chromosome. However, cytogenetics is negative in approximately 10% of patients at diagnosis and at least half of these patients will have a *BCR-ABL* fusion detectable by either fluorescence *in situ* hybridization (FISH) or RT-PCR. The level of sensitivity of cytogenetics is about 1%, which means that in a patient in apparent cytogenetic remission, more than 10^{10} leukemic cells may still be present. The most sensitive method for the detection of residual leukemic cells is RT-PCR, which can detect one cell in a background of 10^5 to 10^6 cells. This means that, even with a negative result, a large number of malignant

cells may still be present and could contribute to disease relapse.

FISH analysis relies on the co-localization of large genomic probes specific to the *BCR* and *ABL* genes. FISH has several advantages over conventional cytogenetics. It can be performed on metaphase or interphase cells and on peripheral blood. Comparison of marrow and blood samples by FISH analysis shows high concordance. One potential problem with FISH is the random co-localization of the signals from the *BCR* and *ABL* probes, such that 8–10% of normal cells will appear positive. This has been circumvented, in part, by D-FISH (double FISH), which uses probes that span the breakpoint region. However, at diagnosis, when most cells are *BCR-ABL*-positive, FISH is a highly accurate diagnostic test as false-negative results are a rarity and false positives are not a concern when more than 90% of cells are positive.

RT-PCR can be used to amplify the region around the splice junction between *BCR* and *ABL*. The high sensitivity of this technique makes it ideal for the detection of minimal residual disease. RT-PCR and, more recently, quantitative RT-PCR have been used to detect residual CML cells following allogeneic stem cell transplantation. These studies have shown that *BCR-ABL* transcripts can be detected for several months after transplantation. Whether persistence of *BCR-ABL* transcripts beyond 6 months is predictive of relapse has not been resolved, but a rising level of *BCR-ABL* transcripts by quantitative RT-PCR after transplantation is highly predictive of relapse. As with FISH in peripheral blood and marrow, RT-PCR the values show a high level of concordance. Both false-positive and false-negative results are possible with RT-PCR and rigorous controls are required to detect these events. False negatives can be due to poor-quality RNA or failure of the reaction, while false positives can be due to contamination. As the majority of patients treated with imatinib will be cytogenetically and FISH-negative, PCR monitoring has been increasingly incorporated into monitoring strategies. Studies of the levels of molecular response to imatinib are being conducted to determine if this correlates with outcome.

Conclusions

From the discovery of the Philadelphia chromosome to the molecularly targeted therapy of CML with imatinib, there is perhaps no other disease which has been so well characterized at the molecular level and in which the knowledge of the molecular biology has been exploited in diagnosis, disease monitoring and therapy. The current task for patients and their treating physicians is to make increasingly difficult decisions between treatment options. For the future, the goals will be to continue to improve and apply the knowledge of the molecular and immunological features of this disease to further improve treatment outcomes.

Further reading

Molecular biology of CML

Ben-Neriah Y, Daley GQ, Mes-Masson A-M Witte ON *et al.* (1986) The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. *Science*, **233**, 212–214.

Daley GQ, Van Etten RA, Baltimore D. (1990) Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science*, **247**, 824–830.

Deininger MW, Goldman JM, Melo JV. (2000) The molecular biology of chronic myeloid leukemia. *Blood*, **96**, 3343–3356.

de Klein A, Geurts van Kessel A, Grosveld G *et al.* (1982) A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. *Nature*, **200**, 765–767.

Heisterkamp N, Jenster G, ten Hoeve J *et al.* (1990) Acute leukaemia in bcr/abl transgenic mice. *Nature*, **344**, 251–253.

Huettner CS, Zhang P, Van Etten RA *et al.* (2000) Reversibility of acute B-cell leukaemia induced by BCR-ABL1. *Nature Genetics*, **24**, 57–60.

Kelliher MA, McLaughlin J, Witte ON *et al.* (1990) Induction of a chronic myelogenous leukemia-like syndrome in mice with v-abl and BCR/ABL. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 6649–6653.

Nowell PC, Hungerford DA. (1960) A minute chromosome in human chronic granulocytic leukemia. *Science*, **132**, 1497(Abstr).

Rowley JD. (1973) A new consistent abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*, **243**, 290–293.

Shtivelman E, Lifshitz B, Gale RP *et al.* (1985) Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature*, **315**, 550–554.

Prognostic scores

Druker BJ, O'Brien SG, Cortes J *et al.* (2002) Chronic myelogenous leukemia. *Hematology (American Society Hematological Education Program)*, pp. 111–135.

Gratwohl A, Hermans J, Goldman JM *et al.* (1998) Risk assessment for patients with chronic myeloid leukaemia before allogeneic blood or marrow transplantation. Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation. *Lancet*, **352**, 1087–1092.

Hasford J, Pfirrmann M, Hehlmann R *et al.* (1998) A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon alfa. Writing Committee for the Collaborative CML

Prognostic Factors Project Group. *Journal of the National Cancer Institute*, **90**, 850–858.

Sokal JE, Cox EB, Baccarani M *et al.* (1984) Prognostic discrimination in 'good-risk' chronic granulocytic leukemia. *Blood*, **63**, 789–799.

Hydroxyurea, busulfan and interferon-α

Chronic Myeloid Leukemia Trialists' Collaborative Group. (1997) Interferon alfa versus chemotherapy for chronic myeloid leukemia: a meta-analysis of seven randomized trials. *Journal of the National Cancer Institute*, **89**, 1616–1620.

Guilhot F, Chastang C, Michallet M *et al.* (1997) Interferon alfa-2B combined with cytarabine versus interferon alone in chronic myelogenous leukemia. *New England Journal of Medicine*, **337**, 223–229.

Silver RT, Woolf SH, Hehlmann R *et al.* (1999) An evidence-based analysis of the effect of busulfan, hydroxyurea, interferon, and allogeneic bone marrow transplantation in treating the chronic phase of chronic myeloid leukemia: developed for the American Society of Hematology. *Blood*, **94**, 1517–1536.

Allogeneic stem cell transplantation

Carella AM, Champlin R, Slavin S *et al.* (2000) Mini-allografts: ongoing trials in humans. *Bone Marrow Transplantation*, **25**, 345–350.

Hansen JA, Gooley TA, Martin PJ *et al.* (1998) Bone marrow transplants from unrelated donors for patients with chronic myeloid leukemia. *New England Journal of Medicine*, **338**, 962–968.

Horowitz MM, Rowlings PA, Passweg JR. (1996) Allogeneic bone marrow transplantation for CML: a report from the International Bone Marrow Transplant Registry. *Bone Marrow Transplantation*, **17** (Supplement 3), S5–S6.

Kolb HJ, Schattenberg A, Goldman JM *et al.* (1995) Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients. European Group for Blood and Marrow Transplantation Working Party Chronic Leukemia. *Blood*, **86**, 2041–2050.

van Rhee F, Szydlo RM, Hermans J *et al.* (1997) Long-term results after allogeneic bone marrow transplantation for chronic myelogenous leukemia in chronic phase: a report from the Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation. *Bone Marrow Transplantation*, **20**, 553–560.

Imatinib

Druker BJ, Lydon NB. (2000) Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *Journal of Clinical Investigation*, **105**, 3–7.

Druker BJ, Sawyers CL, Kantarjian H *et al.* (2001) Activity of a specific inhibitor of the Bcr-Abl tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *New England Journal of Medicine*, **344**, 1038–1042.

Druker BJ, Talpaz M, Resta D *et al.* (2001) Efficacy and safety of a specific inhibitor of the Bcr-Abl tyrosine kinase in chronic myeloid leukemia. *New England Journal of Medicine*, **344**, 1031–1037.

Goldman JM, Druker BJ. (2001) Chronic myeloid leukemia: current treatment options. *Blood*, **98**, 2039–2042.

Hochhaus A, Kreil S, Corbin AS *et al.* (2002) Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy. *Leukemia*, **16**, 2190–2196.

Kantarjian H, Sawyers C, Hochhaus A *et al.* (2002) Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. *New England Journal of Medicine*, **346**, 645–652.

O'Brien SG, Guilhot F, Larson RA *et al.* (2003) Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *New England Journal of Medicine*, **348**, 994–1004.

Sawyers CL, Hochhaus A, Feldman E *et al.* (2002) Imatinib induces hematologic and cytogenetic responses in patients with chronic myeloid leukemia in myeloid blast crisis: results of a phase II study. *Blood*, **99**, 3530–3539.

Shah NP, Nicoll JM, Nagar B *et al.* (2002) Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell*, **2**, 117–125.

Talpaz M, Silver RT, Druker BJ *et al.* (2002) Imatinib induces durable hematologic and cytogenetic responses in patients with accelerated phase chronic myeloid leukemia: results of a phase 2 study. *Blood*, **99**, 1928–1937.

Molecular diagnosis and monitoring

Dewald GW, Wyatt WA, Juneau AL *et al.* (1998) Highly sensitive fluorescence in situ hybridization method to detect double BCR/ABL fusion and monitor response to therapy in chronic myeloid leukemia. *Blood*, **91**, 3357–3365.

Hochhaus A, Reiter A, Saussele S *et al.* (2000) Molecular heterogeneity in complete cytogenetic responders after interferon-alpha therapy for chronic myelogenous leukemia: low levels of minimal residual disease are associated with continuing remission. German CML Study Group and the UK MRC CML Study Group. *Blood*, **95**, 62–66.

Schoch C, Schnittger S, Bursch S *et al.* (2002) Comparison of chromosome banding analysis, interphase- and hypermetaphase-FISH, qualitative and quantitative PCR for diagnosis and for follow-up in chronic myeloid leukemia: a study on 350 cases. *Leukemia*, **16**, 53–59.

Chapter 8 Myelodysplastic syndromes

Jaqueline Boultwood & James S Wainscoat

Pathogenesis

The myelodysplastic syndromes (MDS) are clonal disorders of hematopoietic stem cells. Typically there is peripheral pancytopenia despite a hypercellular bone marrow. It is generally considered that this phenomenon is due to apoptosis of hematopoietic bone marrow cells, resulting in ineffective hematopoiesis. Typically, morphological examination of the bone marrow shows trilineage dysplasia and the molecular lesions occur in pluripotent hematopoietic precursors; the myeloid, monocytic, erythroid, and megakaryocytic lineages are all affected. There is susceptibility to progression to acute myelogenous leukemia (AML), and the patients are often refractory to current therapeutic modalities. There is accumulating evidence suggesting that there is a multistep process with abnormalities in genes for cell cycle control, growth factor receptors, RAS signaling molecules, and transcription factors. A good proportion of all chromosome abnormalities in MDS involve deletions of chromosomes 5, 7, 11, 12, 13 or 20. This provides circumstantial evidence that tumor suppressor genes play an important role in MDS.

It is helpful to consider the molecular genetic model emerging for AML, which is essentially based on two classes of mutations acting together to give rise to AML. The class I mutations provide a proliferative and/or survival signal to hematopoietic progenitors. These include activating point mutations in receptor tyrosine kinases such as FLT3 and c-KIT. Class II mutations are those targeting hematopoietic transcription factors and serve primarily to impair differentiation and subsequent apoptosis. This model may be directly applicable to a small proportion of MDS cases such as those with balanced translocations, although it seems unlikely to be more generally applicable to MDS. Figure 8.1 shows a proposed model for genetic progression of MDS to AML which will be refined over the next few years as new insights are gained from microarray gene expression and other studies.

Etiology

The vast majority of MDS cases are primary; that is, they have no known predisposing event. Epidemiological studies have reported weak associations between MDS and smoking, petroleum exposure, organic solvents, fertilizers, and arsenic and thallium. Inherited gene mutations (e.g. *AML1* and *NF1*; *see below*) probably account for a very small number of cases. Studies of polymorphic gene variants are in progress but as yet no definitive associations have been reported. Some cases of MDS relate to previous treatment with chemotherapy, particularly alkylating agents or topoisomerase inhibitors; these are the secondary MDS. In addition, MDS is seen following bone marrow transplantation; for example, autologous bone marrow transplantation for non-Hodgkin's lymphoma.

Classification and diagnosis

The classification and diagnosis of MDS are achieved chiefly through the morphological examination of blood and bone marrow. The karyotype can also provide important prognostic information. At present, molecular diagnostics are not routinely employed in the diagnosis of MDS, although the application of cDNA microarray technology may come to play an important role in the future.

The French–American–British (FAB) classification of MDS was published in 1982. This classification, criticized in recent years, has in fact been enormously useful to the field. At the very least it has allowed physicians to communicate about

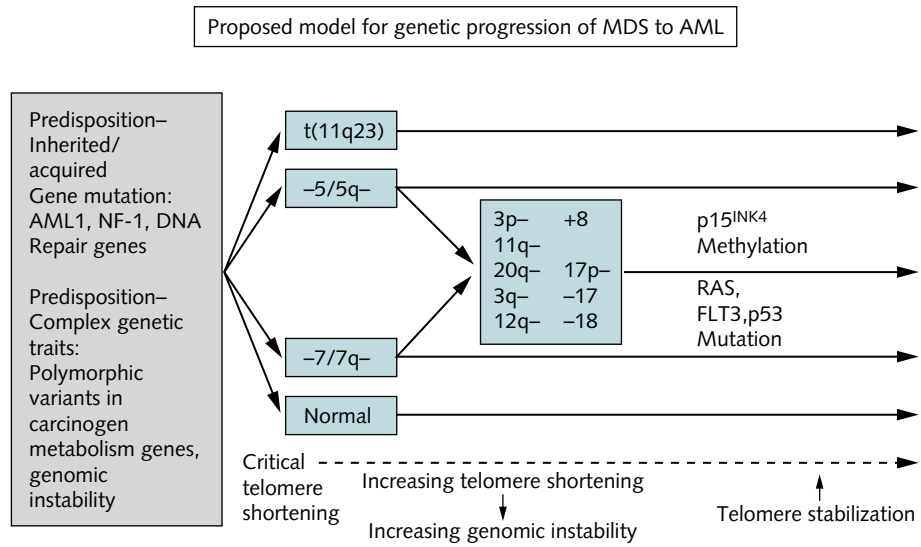


Fig. 8.1 Scheme illustrating some of the genetic changes favoring progression of MDS to AML

a very heterogeneous group of disorders and allow comparisons to be made of clinical trials. Consequently it is still widely used by hematologists. In the FAB classification there are five subtypes, as determined by peripheral blood and bone marrow morphology:

- refractory anemia (RA)
- refractory anemia with ringed sideroblasts (RARS)
- refractory anemia with excess blasts (RAEB)
- refractory anemia with excess blasts in transformation (RAEB-T)
- chronic myelomonocytic leukemia (CMML).

In 1999 the World Health Organization (WHO) published a revised classification of MDS. RA and RARS are defined as conditions with dysplastic features in the erythroid lineages only. RAEB may be split into RAEBI (5–10% blasts) and RAEBII (10–20% bone marrow blasts). The new groups are refractory cytopenia with multilineage dysplasia, del(5q) syndrome, and MDS unclassifiable. The RAEB-T of the FAB classification has disappeared because of similarities to AML, and CMML is now part of a new group of myelodysplastic/myeloproliferative diseases. In addition there is the International Prognostic Scoring System (IPSS), a risk-based classification system in which scores are given for marrow blast percentage, cytogenetic features and cytopenias (Table 8.1). Not surprisingly, the overall score correlates with median survival. All these classification systems are in use but the need now is for more basic information about the fundamental nature of MDS rather than debates about classification.

A caution

At this early point in the chapter a cautionary point may be in order. Some patients will have a disorder which is in fact masquerading as an MDS. Patients diagnosed with refractory anemia (as part of the MDS) without a cytogenetic abnormality and with no excess blasts in the bone marrow are those in whom this problem may occasionally arise. There may well be a variety of causes of such anemias. Perhaps most obviously, there are many causes of sideroblastic anemia in addition to MDS.

Clinical and laboratory features

Most MDS patients are elderly and present with symptoms of anemia. The family doctor may suspect B_{12} deficiency because of the combination of anemia and macrocytosis. However, the anemia does not respond to B_{12} or other vitamins and is by definition refractory. Occasionally patients may present more dramatically with a severe infection or, even more uncommonly, with hemorrhage. The typical laboratory finding is of peripheral blood cytopenia and a hypercellular bone marrow. Of course, secondary MDS may present at any age consequent on chemotherapy for a primary tumor. Overall, the sex ratio of patients with MDS is equal, but the 5q– syndrome has a particular female preponderance and CMML has a male preponderance.

Cytogenetic abnormalities

Clearly, since MDS can progress to AML, all of the cytogenetic abnormalities found in MDS are also found in AML, although their incidence will differ between MDS and AML. However, certain balanced translocations found in AML are never found in MDS, including the t(15;17), inv(16), t(8;21)

Table 8.1 International Prognostic Scoring System.

	Score	
Blast % (marrow)		
<5%	0.5	
5–10%	1	
11–20%	1.5	
21–30%	2.0	
Cytopenia*		
0–1	0.5	
2 or 3	1.5	
Karyotype		
Good (–Y, 5q–, 20q–)	0	
Intermediate (other)	0.5	
Poor (chromosome 7) or complex (≥3 abnormalities)	1.0	
Total score		**Median survival (years)**
0	Low risk	5.7
0.5–1.0	Low to intermediate risk	3.2
1.5–2.0	High to intermediate risk	1.2
≥2.5	High risk	0.4

*Cytopenia is defined as platelet count $<100 \times 10^9$/L, hemoglobin <10 g/dl, or neutrophil count $<1.5 \times 10^9$/L.

and t(9;11). In MDS the typical abnormalities are partial and complete chromosome loss, most often involving del(5q), –7, –Y and del(20q), and chromosome gain, most often of +8. Such abnormalities are present in approximately 50% of primary MDS and 80% of secondary MDS. These abnormalities provide important prognostic information (as reflected in the IPSS). They also give clues to location of tumor suppressor genes whose deletion is postulated to play an important role in the etiology of MDS.

The del(5q) is the most commonly reported deletion in *de novo* MDS and is found in 10–15% of all patients (Figure 8.2). Those MDS patients with RA and the del(5q) as the sole abnormality have the 5q– syndrome. This syndrome was described by Van den Berghe and Cassiman in 1974 and has the following hematological features in association with an interstitial deletion of the long arm of chromosome 5: refractory anemia, female preponderance, macrocytosis, normal/high platelet count, hypolobulated megakaryocytes, and a low transformation rate to AML. There is a clear genotype–phenotype association and the prognosis of the 5q– syndrome is good (Figure 8.3).

The –7/del(7q) is found in 5–10% of patients with *de novo* MDS and in approximately 50% of all therapy-related cases. The del(20q) is found in 3–4% of patients with MDS. The commonly deleted regions (CDRs) of the del(5q), del(7q) and del(20q) in MDS have been defined using fluorescence *in situ* hybridization and molecular mapping. Recently, many of these regions have been significantly narrowed. Gene prediction programs have been used for the complete genomic annotation of the CDRs and this has greatly facilitated the identification of candidate genes. The identification of more than one CDR of the del(5q) and del(7q) suggests the existence of more than one pathogenetically relevant gene.

It remains possible that the deletions involving 5q, 7q and 20q contribute to myeloid malignancy by haploinsufficiency; that is, a dosage affect resulting from the loss of a single allele of a gene. Using Cre-*loxP* based genome engineering, we and others are currently producing chromosomal deletions in mice that mimic those seen in MDS both to define the effect of gene deletion on the development of MDS and for the production of viable models of this disease.

Apoptosis

It is generally agreed that ineffective hemopoiesis in MDS is caused by excessive apoptosis of myeloid precursors. An illustrative study is that of Huh and colleagues (2002), which showed more cell death in refractory anemia with excess blasts in transformation than in acute myeloid leukemia. It is postulated that apoptosis is a reactive phenomenon caused by cytokines. The progression of MDS to leukemia is associated

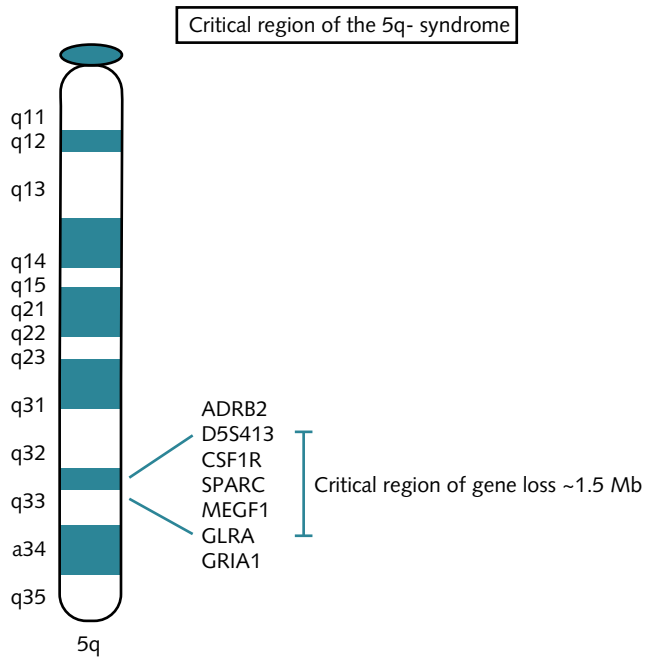


Fig. 8.2 Diagrammatic representation of the long arm of chromosome 5 showing the position of the critical region of gene loss and some of the genes within this region

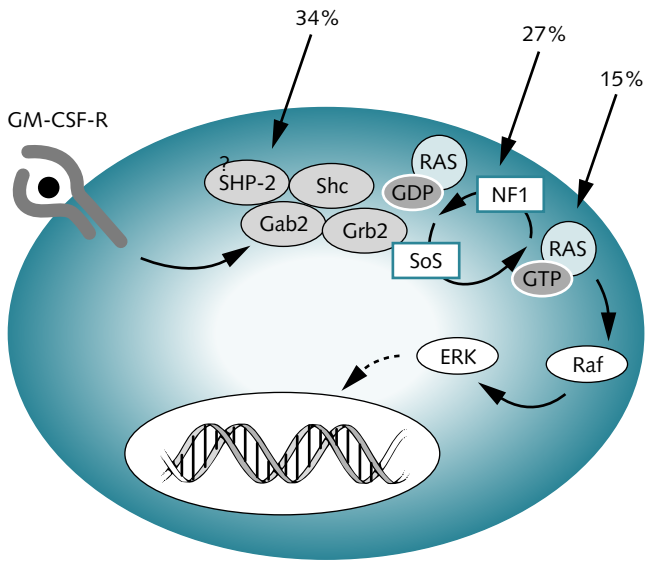


Fig. 8.3 Diagram illustrating the RAS/MAPK signaling in JMML and the percentage of cases with mutations in the *PTPN11* (encoding SHP), *NF1* and *RAS* genes

with a diminution in the level of apoptosis. The molecular basis for these observations is unknown. Hopefully, studies of apoptosis alongside gene expression studies will help elucidate this phenomenon.

Interestingly the expression of the anti-apoptotic gene *survivin* has been shown to be common in RA (11/12 patients) and undetectable in CMML (0/4 patients).

Angiogenesis

It is accepted that increased angiogenesis is important in the pathophysiology of solid tumors. However, some recent studies show that angiogenesis and angiogenic factors may be important in MDS and AML. Both disorders are associated with an increase in the vascularity in the bone marrow as well as increased levels of various angiogenic factors, such as vascular endothelial growth factor, basic fibroblast growth factor, angiogenin, platelet-derived growth factor, hepatocyte growth factor, epidermal growth factor, tumor necrosis factor-α, and transforming growth factor-β. The angiogenic factors may be secreted by the neoplastic cells and promote the growth proliferation of leukemic cells in an autocrine fashion. Blood levels of angiogenin and vascular endothelial factors are elevated in myelodysplastic syndromes and in acute myeloid leukemia. Again, as for apoptosis, the molecular basis of these observations is unknown.

Molecular abnormalities

Mutations of *RAS*, *NF1* and *PTPN11* genes

The *RAS* gene encodes proteins that regulate signal transduction by cycling between an active guanosine triphosphate (GTP) bound state and an inactive guanosine diphosphate (GDP) bound state. These proteins regulate cellular proliferation and differentiation. *RAS* gene mutations (mostly N-*RAS*) are the most common molecular abnormalities in MDS. The reported incidence of *RAS* mutations in MDS has varied but is perhaps 10% of MDS cases overall, with a relatively high incidence in CMML. Often N-*RAS* mutations are found at diagnosis but have also been found to occur during disease progression. Their real significance remains unknown, although most studies report them as unfavorable prognostic markers. It seems unlikely that they are initiating events.

Neurofibromin is the protein encoded by the gene neurofibromatosis type-1 (*NF1*). It contains a domain with sequence homology to GTPase-activating proteins. The binding of neurofibromin to RAS protein accelerates the conversion of RAS-GTP to RAS-GDP. In the majority of children with neurofibromatosis type-1 and MDS, both alleles of *NF1* have been inactivated. This is further evidence for the importance of a deregulated RAS pathway in the molecular pathogenesis

of MDS, and also evidence for the relevance of Knudson's model, at least in some cases of MDS.

Juvenile myelomonocytic leukemia (JMML) accounts for 30% of childhood cases of MDS. It is known that the RAS/MAPK pathway is deregulated in JMML due to mutations in N-*RAS*, K-*RAS* or *NF1* in approximately 40% of cases. Recently, germline mutations of the *PTPN11* gene, which encodes the protein tyrosine phosphatase SHP-2, were discovered to cause Noonan syndrome, a developmental disorder rarely associated with JMML. Subsequently it has been found that somatic mutations in *PTPN11* account for 34% of non-syndromic (i.e. non-Noonan) JMML. The special interest here is that defects in RAS, neurofibromin and SHP-2 are all involved in the regulation of the MAPK cascade and are mutually exclusive in JMML.

Mutations of the *p53* gene

The *p53* gene is the most commonly mutated gene in cancer but *p53* abnormalities are much less common in hematological malignancies. As in solid tumors, mutations tend to occur in exons 5–8 of the *p53* gene. They are found principally in RAEB, RAEB-T and in CMML. Interestingly, they generally occur in association with a deletion of the other allele. It is possible that this may represent an example in MDS of Knudson's hypothesis, whereby the two hits are: (1) a deletion of one chromosome 17, and (2) a *p53* mutation on the partner chromosome. The cases harboring *p53* mutations generally have complex karyotypic abnormalities, making it difficult to assess the contribution of the *p53* abnormalities to the leukemia.

These *p53* mutations are associated with the pseudo-Pelger–Hüet anomaly and vacuoles in neutrophils giving rise to the 17p– syndrome, one of the rare phenotype–genotype associations reported in MDS (the principal one being the 5q– syndrome; *see below*).

Other gene mutations

The *FLT3* duplication is a rare mutation in MDS, occurring in up to 5% of patients and almost always in patients with advanced MDS. Similarly, mutations of the *AML1* gene are rare in primary MDS, with an incidence below 3%. However, a recent study has shown that mutation of *AML1* is a frequent event in radiation-associated and therapy-associated MDS (t-MDS/t-AML), occurring in 40% of patients.

p15 and p16

p15 Hypermethylation

p15 is an inhibitor of cyclin-dependent kinase-4 and a negative regulator of the cell cycle. Transcriptional silencing of *p15* by hypermethylation commonly occurs in MDS (in contrast to inactivation of *p16*, which is rarely inactivated in MDS). Hypermethylation of *p15* is associated with disease progression such that most MDS patients with 10% or more bone marrow blasts show this phenomenon. The importance of gene inactivation by hypermethylation in MDS generally is unknown, but the whole subject has gained in importance from clinical trials of drugs such as 5-aza-2-deoxycytidine, which can reverse hypermethylation.

Telomere shortening in MDS

Telomere shortening is associated with disease evolution and with the presence of complex karyotypic abnormalities in MDS. Telomere shortening in MDS is frequent in patients in the high-risk subgroup and correlates with poor prognosis, suggesting that telomere dynamics may be linked to clinical outcome.

Balanced translocations

As already discussed, the presence of balanced translocations is much more typical of AML than of MDS. Nevertheless they do occur occasionally in MDS and have led to the identification of several novel genes.

The t(5;12) is found in patients with CMML. The translocation creates a fusion gene containing the 5′ portion of the *TEL* gene and the 3′ portion of the platelet-derived growth factor gene (*PDGFRβ*), a tyrosine kinase gene. Treatment of such cases with inatimib mesylate has given durable responses.

A small number of cases of MDS with balanced translocations involving chromosomal band 11p15 and another chromosome, for example t(11;17)(p15;q21) and t(11;12)(p15;q13), have been described, mostly arising after cytotoxic therapy. In all cases analyzed to date, the *NUP98* gene at 11p15 is fused to another gene on the partner chromosome, resulting in a novel fusion protein. The t(11;16)(p23;p13.3) has been reported in patients with t-MDS. The *MLL* gene on chromosome 11 is fused to the *CBP* gene on chromosome 16; MLL is a multifunction protein that regulates *Hox* gene expression during development.

A number of rearrangements involving chromosome band 3q26 have been reported in *de novo* MDS and t-MDS. The two related genes *MDS1/EVI1* and *EVI1* map to 3q26 and encode nuclear transcription factors containing DNA-binding zinc finger domains. They have opposite functions as transcription factors. *EVI1* is frequently activated inappropriately by chromosomal rearrangements at 3q26, leading to the development of myeloid leukemia.

The t(3;3)(q21;26) has been reported in *de novo* MDS and t-MDS. The t(3;21)(q26;q22) has been described in t-MDS/t-AML. In this translocation the *AML1* gene at 21q22 is fused to either *EVI1* or *MDS1* at 3q26 or to both (*AML1–MDS1–EVI1* complex fusion) or *EAP*. *EVI1*, in particular, is activated as a result of the translocation. *AML1* fuses with *MDS1–EVI1* in frame, resulting in the loss of the first 12 amino acids and producing a novel EVI1 protein. The novel EVI1 protein causes arrested differentiation, which leads to apoptosis *in vitro*. It has been shown recently that MDS–EVI1 enhances transforming growth factor-β1 (TGF-β1) signaling and strengthens its growth-inhibitory effect, but the leukemia-associated fusion protein AML1–MDS1–EVI1 product of the t(3;21) abrogates growth inhibition in response to TGF-β1. Forced expression of *EVI1* in embryonic stem cells has been demonstrated to increase cell growth—of megakaryocytic colonies in particular.

α-Thalassemia myelodysplastic syndromes

Rarely, MDS is found in association with acquired α-thalassemia (ATMDS). This disorder has a strong male preponderance and has the features of MDS in addition to α-thalassemia, such as microcytic red cells and HbH on Hb electrophoresis. Previous investigations had failed to find any structural abnormality in the α-globin genes themselves. The mutation responsible for ATMDS has now been found to occur in the *ATRX* gene. The initial clue as to the cause of ATMDS was the finding of markedly diminished expression levels of ATRX in neutrophils by microarray experiments. Subsequently, mutations of *ATRX* were detected in neutrophil DNA of ATMDS patients.

Microarray studies

It is anticipated that a substantial improvement in our understanding of MDS will come from the application of microarray analyses for gene expression profiling. By analogy with published results in other disorders, such as breast cancer, it is feasible that this technology will allow better classification, give a better prognostic guide and, importantly, give new insights into the biology of MDS. The study by Schoch and colleagues (2002) clearly shows that AML with reciprocal rearrangements can be distinguished by its specific gene expression profile. A minimum of 13 genes were able to separate AML subtypes with t(15;17), t(8;21) and inv(16).

The early studies of MDS illustrate the promise of this development. Miyazato *et al.* (2001) investigated gene expression profile in MDS and AML patients. This study identified 20 MDS-specific genes and 21 AML-specific genes and reported preferential expression of the *DLK* gene in MDS. Hofmann *et al.* (2002) studied the gene expression profile of 18 MDS patients compared with four controls. The study identified many genes expressed differentially between the low-risk and high-risk groups and between the high-risk or low-risk groups and the healthy control group. Class membership prediction was used to identify a subset of 11 predictor genes that could differentiate between patients with low-risk MDS, patients with high-risk MDS, and healthy controls.

Treatment

In the treatment of leukemia there is a long history of drug discovery by chance and serendipity, and hopefully this will continue. Nevertheless, better understanding of the molecular basis of MDS will lead to more targeted drug approaches. The best example of this approach in hematology is the treatment of chronic myeloid leukemia (CML) with imatinib. As noted above, there is spin-off for MDS in the treatment of the t(5;12) with the same drug (imatinib), although targeting a different tyrosine kinase. The commonest abnormalities in MDS are the *RAS* point mutations and hypermethylation of *p15*, and both are under scrutiny from the pharmaceutical companies. Farnesyltransferase inhibitors target multiple pathways, including the RAS pathways. In phase I trials of the farnesyltransferase inhibitor Zarnestra, clinical responses were shown in approximately 30% of patients with high-risk leukemia or MDS.

Low-dose 5-aza-2′deoxycytidine (DAC), a DNA hypomethylating agent, has been used for the treatment of MDS. A phase II study shows that DAC therapy was effective in half of the studied patients with high-risk MDS and was especially active in the patients with the worst prognosis. Subsequently it was demonstrated that frequent, selective *p15* hypermethylation was reversed in responding MDS patients after treatment with a methylation inhibitor. Of course, the methylation status of many genes will be affected by DAC treatment and much more work is needed to resolve the complicated biological issues surrounding this treatment.

Conclusions

One strategy for scientific progress in MDS would be to break the disease down into different disorders or syndromes—much as the REAL (Revised European American Lymphoma Classification) classification does for lymphoma. However, although most patients are not readily classified in this way, some particular syndromes do exist amongst the MDS. These include the 5q– syndrome, the 17p– syndrome, CMML with

t(5;12), and ATMDS. The central problem of MDS is its genetic heterogeneity. Fortunately, we now have the technology to achieve a better understanding of the genetic and biochemical basis for this heterogeneity, and this in turn will lead to improved therapies.

Further reading

Pathogenesis

Chen H, Sandler DP, Taylor JA *et al.* (1996) Increased risk for myelodysplastic syndromes in individuals with glutathione transferase theta 1 (GSTT1) gene defect. *Lancet*, **347**, 295–297.

Daskalakis M, Nguyen TT, Nguyen C *et al.* (2002) Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2′-deoxycytidine (decitabine) treatment. *Blood*, **100**, 2957–2964.

Paquette RL, Landaw EM, Pierre RV *et al.* (1993) N-ras mutations are associated with poor prognosis and increased risk of leukemia in myelodysplastic syndrome. *Blood*, **82**, 590–599.

Smith MT, Wang Y, Kane E *et al.* (2001) Low NAD(P)H:quinone oxidoreductase 1 activity is associated with increased risk of acute leukemia in adults. *Blood*, **97**, 1422–1426.

Classification

Kita-Sasai Y, Horiike S, Misawa S *et al.* (2001) International prognostic scoring system and TP53 mutations are independent prognostic indicators for patients with myelodysplastic syndrome. *British Journal of Haematology*, **115**, 309–312.

Cytogenetic and molecular abnormalities

Ahuja HG, Felix CA, Aplan PD. (1999) The t(11;20)(p15;q11) chromosomal translocation associated with therapy-related myelodysplastic syndrome results in an NUP98-TOP1 fusion. *Blood*, **94**, 3258–3261.

Atoyebi W, Kusec R, Fidler C *et al.* (1997) Glutathione S-transferase gene deletions in myelodysplasia. *Lancet*, **349**, 1450–1451.

Baxter EJ, Kulkarni S, Vizmanos JL *et al.* (2003) Novel translocations that disrupt the platelet-derived growth factor receptor beta (PDGFRB) gene in BCR-ABL-negative chronic myeloproliferative disorders. *British Journal of Haematology*, **120**, 251–256.

Bench AJ, Nacheva EP, Hood TL *et al.* (2000) Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. UK Cancer Cytogenetics Group (UKCCG). *Oncogene*, **19**, 3902–3913.

Boultwood J, Fidler C, Kusec R *et al.* (1997) Telomere length in myelodysplastic syndromes. *American Journal of Hematology*, **56**, 266–271.

Boultwood J, Fidler C, Strickson AJ *et al.* (2002) Narrowing and genomic annotation of the commonly deleted region of the 5q- syndrome. *Blood*, **99**, 4638–4641.

Gibbons RJ, Pellagatti A, Garrick D *et al.* (2003) Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the alpha-thalassemia myelodysplasia syndrome (ATMDS). *Nature Genetics*, **34**, 446–449.

Golub TR, Barker GF, Lovett M *et al.* (1994) Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell*, **77**, 307–316.

Harada H, Harada Y, Tanaka H *et al.* (2003) Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood*, **101**, 673–680.

Hofmann WK, de Vos S, Komor M *et al.* (2002) Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow. *Blood*, **100**, 3553–3560.

Horiike S, Yokota S, Nakao M *et al.* (1997) Tandem duplications of the FLT3 receptor gene are associated with leukemic transformation of myelodysplasia. *Leukemia*, **11**, 1442–1446.

Le Beau MM, Espinosa R 3rd, Davis EM *et al.* (1996) Cytogenetic and molecular delineation of a region of chromosome 7 commonly deleted in malignant myeloid diseases. *Blood*, **88**, 1930–1935.

Miyazato A, Ueno S, Ohmine K *et al.* (2001) Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. *Blood*, **98**, 422–427.

Padua RA, Guinn BA, Al-Sabah AI *et al.* (1998) RAS, FMS and p53 mutations and poor clinical outcome in myelodysplasias: a 10-year follow-up. *Leukemia*, **12**, 887–892.

Pellagatti A, Esoof N, Watkins F *et al.* (2004) Gene expression profiling in the myelodysplastic syndromes using cDNA microarray technology. *British Journal of Haematology*, **125**, 576–583.

Ross TS, Bernard OA, Berger R *et al.* (1998) Fusion of huntingtin interacting protein 1 to platelet-derived growth factor beta receptor (PDGFbetaR) in chronic myelomonocytic leukemia with t(5;7)(q33;q11.2). *Blood*, **91**, 4419–4426.

Rowley JD, Reshmi S, Sobulo O *et al.* (1997) All patients with the t(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders. *Blood*, **90**, 535–541.

Rubin CM, Larson RA, Anastasi J *et al.* (1990) t(3;21)(q26;q22): a recurring chromosomal abnormality in therapy-related myelodysplastic syndrome and acute myeloid leukemia. *Blood*, **76**, 2594–2598.

Quesnel B, Guillerm G, Vereecque R *et al.* (1998) Methylation of the p15(INK4b) gene in myelodysplastic syndromes is frequent and acquired during disease progression. *Blood*, **91**, 2985–2990.

Schoch C, Kohlmann A, Schnittger S *et al.* (2002) Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 10008–10013.

Side L, Taylor B, Cayouette M *et al.* (1997) Homozygous inactivation of the NF1 gene in bone marrow cells from children with neurofibromatosis type 1 and malignant myeloid disorders. *New England Journal of Medicine*, **336**, 1713–1720.

Song WJ, Sullivan MG, Legare RD *et al.* (1999) Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nature Genetics*, **23**, 166–175.

Sood R, Talwar-Trikha A, Chakrabarti SR *et al.* (1999) MDS1/EVI1 enhances TGF-beta1 signaling and strengthens its growth-inhibitory effect but the leukemia-associated fusion protein AML1/MDS1/EVI1, product of the t(3;21), abrogates growth-inhibition in response to TGF-beta1. *Leukemia*, **13**, 348–357.

Tartaglia M, Niemeyer CM, Fragale A *et al.* (2003) Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nature Genetics*, **34**, 148–150.

Uchida T, Kinoshita T, Nagai H *et al.* (1997) Hypermethylation of the p15INK4B gene in myelodysplastic syndromes. *Blood*, **90**, 1403–1409.

Van den Berghe H, Cassiman JJ *et al.* (1974) Distinct haematological disorder with deletion of long arm of no. 5 chromosome. *Nature*, **251**, 437–438.

Wattel E, Preudhomme C, Hecquet B *et al.* (1994) p53 mutations are associated with resistance to chemotherapy and short survival in hematologic malignancies. *Blood*, **84**, 3148–3157.

Yoneda-Kato N, Look AT, Kirstein MN *et al.* (1996) The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid leukemia produces a novel fusion gene, NPM-MLF1. *Oncogene*, **12**, 265–275.

Zhao N, Stoffel A, Wang PW *et al.* (1997) Molecular delineation of the smallest commonly deleted region of chromosome 5 in malignant myeloid diseases to 1–1.5 Mb and preparation of a PAC-based physical map. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 6948–6953.

Apoptosis

Badran A, Yoshida A, Wano Y *et al.* (2003) Expression of the anti-apoptotic gene survivin in myelodysplastic syndrome. *International Journal of Oncology*, **22**(1), 59–64.

Huh YO, Jilani I, Estey E *et al.* (2002) More cell death in refractory anemia with excess blasts in transformation than in acute myeloid leukemia. *Leukemia*, **16**, 2249–2252.

Parker JE, Mufti GJ. (2001) The role of apoptosis in the pathogenesis of the myelodysplastic syndromes. *International Journal of Hematology*, **73**, 416–428.

Therapy

Cortes JE, Kurzrock R, Kantarjian HM. (2002) Farnesyltransferase inhibitors: novel compounds in development for the treatment of myeloid malignancies. *Seminars in Hematology*, **39**, 26–30.

Gilliland DG. (2002) Molecular genetics of human leukemias: new insights into therapy. *Seminars in Hematology*, **39**, 6–11.

Chapter 9 Myeloproliferative disorders

Anthony J Bench, George S Vassiliou, Brian J P Huntly & Anthony R Green

Introduction

The myeloproliferative disorders (MPDs) comprise polycythemia vera (PV), essential thrombocythemia (ET) and idiopathic myelofibrosis (IMF). They were first classed together as an overlapping spectrum of preleukemic disorders along with chronic myeloid leukemia (CML) by Dameshek in 1951. CML, with its pathognomonic chromosomal translocation, greater clinical homogeneity and increased acute leukemic potential, is now considered a separate entity. The MPDs are believed to result from acquired genetic changes in hematopoietic stem cells that perturb stem cell behavior and result in the overproduction of one or more myeloid lineages. These disorders are not mutually exclusive and considerable overlap is seen between distinct categories, with progression from one disease to another a frequent occurrence (Figure 9.1). The MPDs are also preleukemic and a small proportion of patients transform to acute myeloid leukemia (AML).

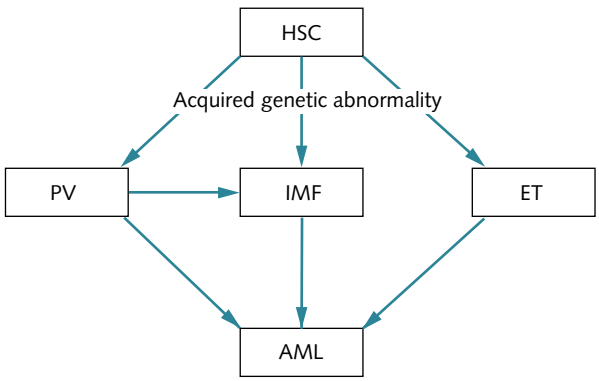


Fig. 9.1 Relationships of the myeloproliferative disorders and transformation to acute myeloid leukemia
The percentage of patients who transform varies according to the initial disease, duration of disease and treatment. HSC, hematopoietic stem cell; PV, polycythemia vera; IMF, idiopathic myelofibrosis; ET, essential thrombocythemia; AML, acute myeloid leukemia.

The MPDs are relatively uncommon disorders. PV has an annual incidence of 5–10 cases per million population. Less is known about the incidence of ET. However, it is likely to be similar to, if not greater than, that for PV. Although little is known of the epidemiology of IMF, it is thought to have an incidence of 5 cases per million population annually but, since many patients are diagnosed coincidentally and up to 25% are asymptomatic at the time of diagnosis, this figure is likely to be an underestimate.

Classification

The diagnosis of the MPDs is made according to a defined set of criteria originally put forward by the Polycythemia Vera Study Group (PVSG) and subsequently modified by Pearson and colleagues to reflect advances in diagnostic practice (Table 9.1). These have led to greater diagnostic uniformity and have helped clinicians make diagnostic decisions in difficult and borderline cases.

A more recent attempt to revise these criteria under the auspices of the World Health Organization has introduced a number of controversial features (Table 9.2). In PV, for example, many feel it is inappropriate to use a raised hemoglobin level (greater than 18.5 g/dl in men; greater than 16.5 g/dl in women) as a specific diagnostic criterion in the absence of a raised red cell mass. In ET and PV, there is concern about the increased reliance on bone marrow histology, the interpretation of which is frequently subjective. This is especially true for ET, in which an abnormal bone marrow biopsy is one of only two positive criteria. In addition, the degree of fibrosis allowable in a diagnosis of ET is vague; the distinction between 'prefibrotic IMF' and ET relies on subtle differences in megakaryocyte morphology and there are no data on interobserver variation when attempting to apply these criteria.

We therefore feel it is most appropriate to use the Pearson modification of the PVSG criteria (Table 9.1). Even with these strict criteria, it can be difficult to differentiate the MPDs from

Table 9.1 Modified PVSG diagnostic criteria for myeloproliferative disorders.

(a) Modified criteria for the diagnosis of polycythemia vera

A1 Raised red cell mass (>25% above normal predicted value)
A2 Absence of cause of secondary polycythemia
A3 Palpable splenomegaly
A4 Acquired cytogenetic abnormality

B1 Thrombocytosis (platelet count >400 × 10^9/L)
B2 Neutrophil leukocytosis (neutrophil count >10 × 10^9/L)
B3 Splenomegaly demonstrated on isotope/ultrasound scanning
B4 Characteristic BFU-E growth or reduced serum erythropoietin

A1 + A2 + A3 or A4 establishes PV. A1 + A2 + two of B establishes PV.

(b) Diagnostic criteria for essential thrombocythemia

1 Platelet count >600 × 10^9/L
2 Packed cell volume (PCV) <0.51 for males or <0.48 for females or normal red cell mass in those with a high normal PCV and splenomegaly
3 Stainable iron in marrow or normal serum ferritin or normal red cell mean corpuscular volume (MCV). If measurement suggests iron deficiency then PV cannot be excluded unless a trial of iron therapy fails to increase the red cell mass into the erythrocytotic range.
4 No Philadelphia chromosome or *BCR-ABL* gene rearrangement
5 Collagen fibrosis of marrow either absent or less than one-third of biopsy area without marked splenomegaly and leukoerythroblastic reaction
6 No cytogenetic or morphological evidence for a myelodysplastic syndrome
7 No cause for reactive thrombocytosis

To make a diagnosis of ET, all criteria need to be met.

(c) Characteristic features of idiopathic myelofibrosis

1 Bone marrow fibrosis
2 Extramedullary hemopoiesis
3 Splenomegaly
4 Leukoerythroblastic blood picture
5 Absence of another chronic myeloproliferative disorder
6 Absence of a condition associated with secondary bone marrow fibrosis

All patients demonstrate points 1, 4, 5 and 6. Most patients would also have points 2 and 3.

Information from Pearson (1998) and Reilly (1998).

reactive causes. Hence, the identification of target genes in these disorders would produce much-needed diagnostic tools for clinicians as well as providing important insights into the regulation of normal hematopoiesis.

X-linked clonality assays

The original evidence that MPDs arise as a result of transformation of a hematopoietic stem cell has come from two main sources: analysis of karyotypic abnormalities and analysis of X-chromosome inactivation patterns to determine the clonality of different cell lineages. X-inactivation assays offer a means of assessing the clonality of a population of cells without any acquired cytogenetic or molecular markers. X inactivation occurs in females as a means of dosage compensation so that equal levels of expression of X-linked genes occur in males and females. Early in embryogenesis one X chromosome is inactivated at random in each cell—a process termed 'Lyonization'. The progeny of each cell inherits this inactivation pattern. As a consequence, a normal adult female is a 'mosaic'—some cells carry an active maternal X chromosome, others an active paternal X chromosome.

A population of cells such as a tumor, being clonally derived from a single cell, will therefore all carry either an active maternal X chromosome or an active paternal X chromosome. By contrast, a polyclonal population contains some cells with an active maternal X chromosome and others with an active paternal X chromosome. Assessment of X-inactivation patterns requires the ability to distinguish paternal and maternal X chromosomes and so the various assays are all based on polymorphic X-linked genes. It is also necessary to determine which X chromosome is active. This can be achieved by monitoring expression of an X-linked gene at the protein or RNA level. Alternatively, DNA methylation can be used as a surrogate marker for gene inactivity since, for many genes, methylation status correlates well with transcriptional activity. An example of a DNA-based technique, the human androgen receptor assay (HUMARA), is shown in Figure 9.2.

Initial studies on clonality used a rare polymorphism in the *G6PD* gene which results in two distinct protein products. In patients with PV, a single isoform was present in erythrocytes, granulocytes, platelets and bone marrow buffy coat, implying that these cells were clonally derived. Both isoforms were expressed in lymphocytes and skin fibroblasts. The majority of erythroid and granulocyte progenitors were

Table 9.2 Summary of WHO diagnostic criteria for myeloproliferative disorders.

(a) Diagnostic criteria for polycythemia vera

A1 Raised red cell mass (>25% above normal predicted value) or Hb >18.5 g/dl (men) or >16.5 g/dl (women)
A2 Absence of cause of secondary polycythemia
A3 Splenomegaly
A4 Acquired clonal genetic abnormality other than Philadelphia chromosome or *BCR-ABL*
A5 Endogenous erythroid colony formation *in vitro*

B1 Thrombocytosis (platelet count >400 × 10^9/L)
B2 Neutrophil leukocytosis (neutrophil count >12 × 10^9/L)
B3 Bone marrow biopsy showing panmyelosis with prominent erythroid and megakaryocytic proliferation
B4 Low serum erythropoietin

A1 + A2 + one of A3, A4 or A5 establishes PV. A1 + A2 + two of B establishes PV.

(b) Diagnostic criteria for essential thrombocythemia

1 Platelet count >600 × 10^9/L
2 Bone marrow biopsy showing megakaryocytic lineage proliferation with an increased number of enlarged mature megakaryocytes
3 No evidence for polycythemia vera
Normal red cell mass or Hb <18.5 g/dl (men), 16.5 g/dl (women)
Stainable iron in marrow, normal serum ferritin or normal MCV. If this condition is not met (iron deficiency) PV cannot be excluded unless a trial of iron fails to increase red cell mass or Hb levels to the PV range.
4 No evidence for CML
No Philadelphia chromosome or *BCR-ABL* gene rearrangement
5 No evidence for IMF
Collagen fibrosis is absent
Reticulin fibrosis is minimal or absent
6 No evidence for a myelodysplastic syndrome
No del(5q), t(3;3)(q21;q26), inv (3)(q21;q26)
No significant granulocytic dysplasia; few if any micromegakaryocytes
7 No cause for reactive thrombocytosis due to
Underlying inflammation or infection
Underlying neoplasia
Prior splenectomy

All seven criteria need to be met to confirm a diagnosis of ET.

(c) Characteristic features of idiopathic myelofibrosis

	Histology	**Clinical/laboratory findings**
Prefibrotic IMF	Minimal or absent reticulin fibrosis Hypercellular BM with increased number of neutrophils Increased numbers and clustering of atypical megakaryocytes No or mild leukoerythroblastosis No or minimal red cell poikilocytosis	Mild anemia Mild to moderate leukocytosis Mild to marked thrombocytosis No or mild splenomegaly or hepatomegaly
Fibrotic IMF	Reticulin and/or collagen fibrosis Decreased BM cellularity Diluted marrow sinuses containing pockets of hematopoietic precursors Increased numbers and clustering of atypical megakaryocytes New bone formation Leukoerythroblastosis Prominent red cell poikilocytosis (dacrocytes)	Splenomegaly and hepatomegaly Moderate to marked anemia WBC and platelet count variable

also clonally derived. Similar patterns were also observed for ET and IMF.

Subsequent studies using methylation of X-linked genes as a marker of gene inactivity have demonstrated an unbalanced pattern of methylation consistent with clonality in white blood cells or bone marrow cells from a high proportion of patients with ET, PV or IMF. However, in most of the early studies no control was performed to exclude the possibility of

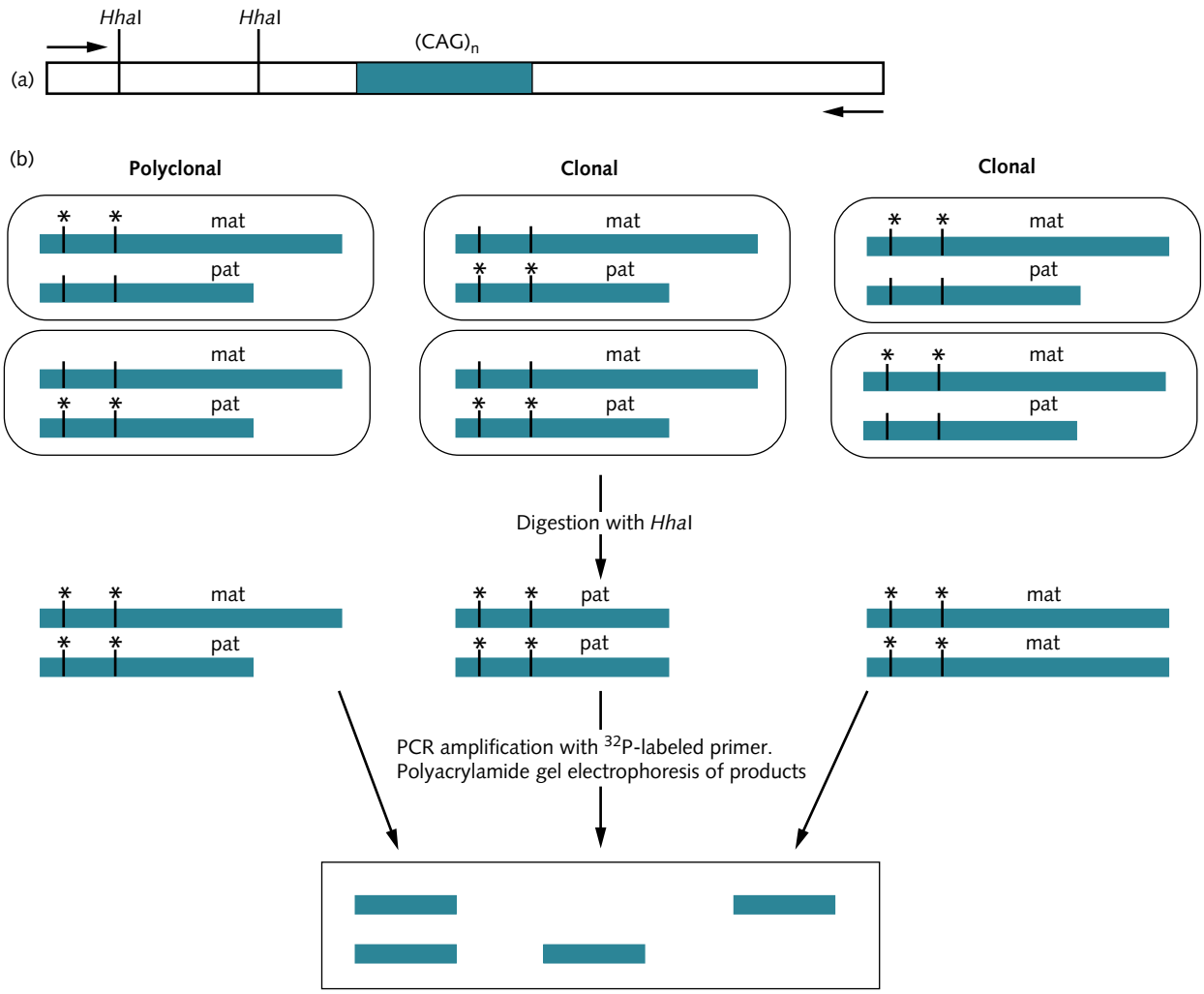


Fig. 9.2
(a) First exon of the human androgen receptor locus (HUMARA). The first exon contains two recognition sites for the methylation-sensitive restriction enzyme *Hha*I. Methylation of these sites correlates with X-chromosome inactivation. The sites lie very close to a highly polymorphic $(CAG)_n$ repeat. PCR primers can therefore be designed which flank both the *Hha*I site and the polymorphic $(CAG)_n$ repeat.
(b) Determination of clonality of a population of cells using the HUMARA. *Methylated restriction site. *Hha*I sites on the active X chromosome are unmethylated and are therefore digested with the enzyme, whereas methylated sites will not be cleaved. If the maternally and paternally derived alleles contain different numbers of the (CAG) repeat, their size can be distinguished following amplification by PCR and separation by polyacrylamide gel electrophoresis. Adapted from Allen *et al*. (1992).

skewed Lyonization, a situation which can mimic true clonality. The most appropriate tissue to use for such a control is T cells, and the demonstration of a polyclonal pattern in T cells together with a skewed pattern in granulocytes (or other appropriate lineage) in female patients (Figure 9.3) was generally considered as evidence for the presence of a clonal myeloid malignancy.

Interestingly, a small number of well-characterized patients with PV and ET possess polyclonal T cells and polyclonal granulocytes (Figure 9.3). These results suggest that, in some patients, only a small proportion of granulocytes are part of the neoplastic clone or that the granulocytic lineage is not involved at all. Alternatively, polyclonal hematopoiesis may still be present in some patients fulfilling current criteria for an MPD. These results have been confirmed in patients with ET by analysis of polymorphisms at the level of mRNA and hence do not merely reflect alterations in the methylation status of these patients' granulocytes. The use of RNA-based methods has also shown that a small number of ET patients possess clonal platelets but polyclonal granulocytes and T

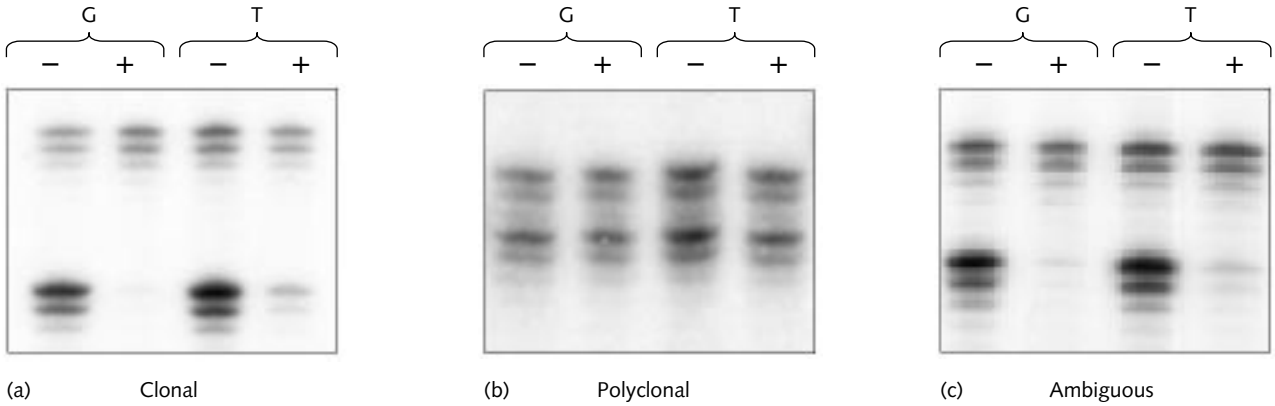


Fig. 9.3 HUMARA assay results

Examples of results obtained with three MPD patient samples are shown.
G, granulocytes; T, T cells; + and – refer to the presence and absence of predigestion with the methylation-sensitive restriction enzyme *Hha*I. (a) Clonal pattern observed in most MPD patients. T cells exhibit balanced X-inactivation whereas granulocytes display a skewed pattern. (b) A small number of patients have polyclonal T cells and polyclonal granulocytes. (c) This result is ambiguous since both the granulocytes and T cells give a skewed pattern. This is consistent with either excessive Lyonization or with T cells arising from the malignant clone.

cells. These various observations have contributed to a growing realization that heterogeneity exists within each diagnostic category and especially for ET. It has also been suggested that ET patients demonstrating polyclonal hemopoiesis may be at a lower risk of suffering a thrombotic event.

Although a minority of MPD patients possess polyclonal granulocytes, the finding of a skewed pattern in granulocytes with polyclonal T cells was thought to be potentially useful as a positive diagnostic test. Initial studies suggested that this pattern was rare in normal women. However, these control women were usually young and the MPDs are mainly observed in the elderly. When a large number of normal elderly women were studied, a significant number showed a clonal pattern in unfractionated blood cells. It was subsequently found that a skewed pattern in granulocytes with polyclonal T cells, identical to that found in MPD patients, is present in 25–50% of normal elderly women.

A number of mechanisms, not necessarily mutually exclusive, could account for this age-related skewed X inactivation. Firstly, an acquired mutation could lead to a proliferative advantage. Secondly, stem cell depletion could result in stochastic predominance of one cell type. However, several lines of evidence suggest that the predominant mechanism contributing to this phenomenon is selection for polymorphic X-linked differences. Twin studies in humans have suggested the presence of X-linked genes which regulate stem cell kinetics; elderly cats develop skewing towards one parental *G6PD* allele; and strain-dependent differences in the response of mouse stem cells to cytokines have been linked to a number of genetic loci, including some on the X chromosome.

Chromosome abnormalities

Unlike the Philadelphia chromosome in CML, there is no pathognomonic chromosomal abnormality associated with the MPDs. However, a number of recurrent chromosomal abnormalities have been documented in the MPDs. Chromosomal abnormalities are seen in approximately one-third of patients with PV and IMF. The most frequent chromosomal abnormalities in these disorders, as detected by G-banding, are shown in Table 9.3. Detection of cytogenetic abnormalities is infrequent in ET patients.

In PV, although the numbers of patients were small or not well matched, there is some evidence that the survival of patients with a chromosomal abnormality is less than that of patients with a normal karyotype. For IMF, studies with large numbers of patients have demonstrated that an abnormal karyotype at diagnosis tends to be associated with a poorer prognosis.

Some chromosomal changes, such as deletions or monosomies of chromosomes 5 and 7, are almost invariably seen after exposure to myelosuppressive therapy and frequently as part of a complex karyotype. It is therefore unlikely that these chromosomes contain genes involved in the etiology of the MPDs. By contrast, deletions of part of the long arm of chromosomes 20 (del 20q) and 13 (del 13q), trisomies of chromosomes 8 and 9 and duplication of part of the long arm of chromosome 1 have all been seen in untreated patients and are frequently present as sole abnormalities. They are likely, therefore, to mark the sites of genes which play an early role in the pathogenesis of the MPDs.

Table 9.3 Summary of major karyotypic abnormalities identified by G-banding in PV and IMF.

	Polycythemia vera		Idiopathic myelofibrosis	
Abnormality	**Number of occurrences of abnormality**	**% of all patients**	**Number of occurrences of abnormality**	**% of all patients**
Deletion of 20q	36	8.6	46	10.6
Deletion of 13q	15	3.6	39	9.0
Trisomy 8	27	6.4	23	5.3
Trisomy 9	25	6.0	13	3.0
Duplication of 1q	14	3.3	24	5.5
Deletion of 7q or monosomy 7	5	1.2	14	3.2
Deletion of 5q or monosomy 5	17	4.1	7	1.6
Total number of patients with one or more abnormality	145	34.6	185	42.5
Total patients	419		435	

Data taken from Berger *et al*., 1984; Demory *et al*., 1988; Rege-Cambrin *et al*., 1987; Swolin *et al*., 1988; Diez-Martin *et al*., 1991; Mertens *et al*., 1991; Dupriez *et al*., 1996; Reilly *et al*., 1997; Tefferi *et al*., 2001.

Deletions

20q deletion

The importance of del(20q) is exemplified in a study of 3996 consecutive abnormal bone marrow samples performed by Dewald *et al*. (1993). Almost 3000 of these samples possessed a sole chromosomal abnormality and, of these, del(20q) was the second most common structural abnormality after t(9;22). In addition to the MPDs, 20q deletions are also seen in approximately 4% of patients with myelodysplastic syndrome (MDS) and in 1–2% of patients with AML. However, 20q deletions are rarely seen in lymphoid malignancies. This pattern of disease association suggests that the deleted region of chromosome 20 marks the site of one or more genes, loss or inactivation of which perturbs the regulation of hematopoietic progenitors. The finding of 20q deletions at diagnosis and as a sole abnormality suggests that, in at least some cases, it plays an early role in disease pathogenesis.

There have been a number of studies concerning the prognostic significance of the 20q deletion in both MPD and MDS. As far as MPD is concerned, there is no significant difference in the survival rate of patients with and without a 20q deletion, although only a small number of patients have been studied. For MDS, a 20q deletion is associated with a relatively good prognosis, if it is observed without any other karyotypic abnormalities. Deletions of chromosome 20q may be particularly associated with a subset of myeloid disorders characterized by megakaryocytic and erythroid dysplasia with only infrequent granulocytic dysplasia. However, it is not clear whether such a pattern is also seen in patients lacking a 20q deletion and so the significance of these findings is unclear.

Since both MPD and MDS are believed to result from transformation of a hematopoietic stem cell, it was of interest to determine whether 20q deletions arise in a pluripotent progenitor or a later, committed progenitor. White *et al.* (1994) described a patient with MDS whose granulocytes and monocytes were clonal (as assessed by X-inactivation patterns) and clearly contained the deletion, whereas B cells and T cells were polyclonal and did not contain the deletion. However, EBV-transformed B-cell lines carrying the 20q deletion were derived from this patient. Similarly, in another patient, a 20q deletion was reported in EBV-transformed B-cell lines as well as in CFU-GM, CFU-GEMM and BFU-E colonies. Clearly, the 20q deletion can arise in a very early progenitor with both lymphoid and myeloid potential.

In most patients with a 20q deletion, the deletion can readily be detected in peripheral blood neutrophils using microsatellite PCR to demonstrate loss of heterozygosity (LOH) (Figure 9.4). However, this is not always the case. Asimakopoulos *et al.* (1996) described an interesting subset of patients with a 20q deletion in the majority of bone marrow metaphases but with no deletion detectable in peripheral blood granulocytes by microsatellite PCR. This observation suggests that, in some patients, granulocytes carrying the deletion may be preferentially destroyed or retained within the bone marrow. Granulocytes from the female patients displayed a clonal pattern of X inactivation, implying that either the 20q deletion was not the initiating event or the clonal X-inactivation pattern in granulocytes was age-related (*see above*) and not part of the pathogenesis of these disorders.

Molecular analysis of the 20q deletion has been undertaken to identify the gene or genes involved. Using fluorescence *in situ* hybridization (FISH), microsatellite PCR (Figure 9.4)

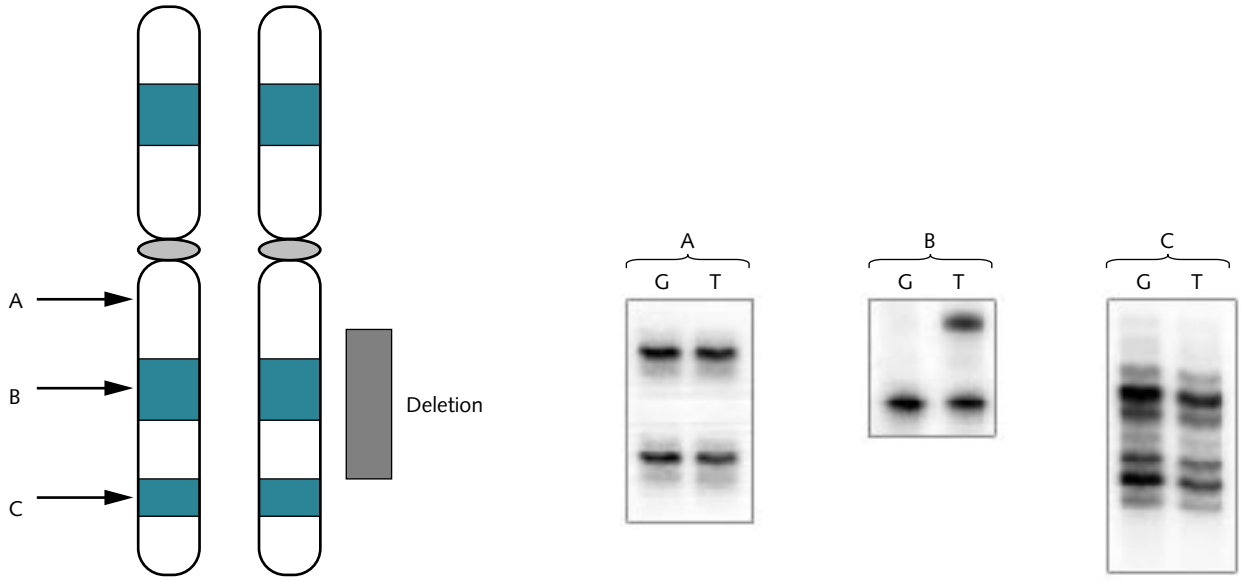


Fig. 9.4 Use of microsatellite PCR to map deletions
Granulocytes (G) contain the deletion whereas T cells (T) do not. Using markers A and C, two alleles are present in granulocytes and therefore these markers lie outside the deletion. Using marker B, two alleles are present in T cells but only one in granulocytes. Hence, marker B lies within the deletion. Reproduced from Bench *et al*. (1998) with permission from Elsevier Science.

and quantitative Southern blotting, a common deleted region (CDR) spanning 20q11–20q13 has been defined and is likely to contain one or more tumor suppressor genes. Deletion of part of 20q has also been demonstrated by LOH and FISH in CML patients.

Given that MPD and MDS are overlapping but clinically different diseases, it remains possible that different or additional genes are involved in the two disorders. Therefore, two overlapping common deleted regions have been defined. The MPD CDR spans 3 Mb. Two different MDS/AML CDRs have been constructed (Figure 9.5). A 3-Mb CDR based entirely on patients with a simple interstitial deletion overlaps the MPD CDR by 2 Mb. A smaller 700-kb CDR has been constructed using patients with complex rearrangements of chromosome 20. Within the CDRs identified by Bench *et al*. (Figure 9.5), 40 genes lie within the MPD CDR and 18 within the MDS CDR. Of the 16 that lie within both CDRs, 6 are expressed within normal $CD34^+$ progenitor cells, making them good positional and expression candidates (Figure 9.5). These include the *SFRS6*, *L3MBTL* and *MYBL2* genes.

13q deletion

In contrast to 20q deletions, molecular analysis of chromosome 13q deletions in myeloid malignancies has only recently been initiated. CDRs of 4 centimorgans (cM) and of 14 cM have been constructed in MDS and MPD patients respectively. LOH at the *RB1* locus has been observed in the bone marrow or peripheral blood from 13 of 30 MPD patients, suggesting that abnormalities of chromosome 13 may be relatively more common than conventional karyotyping suggests. No mutations of the *RB1* gene have been found in MPD patients.

Deletion of chromosome 13q is the commonest structural chromosomal abnormality seen in chronic lymphocytic leukemia (CLL). Deletions and unbalanced translocations involving 13q14 are seen in approximately 18% of all CLL cases. Several small common deleted regions have been constructed and a number of candidate genes identified. These include *RFP2*, *BCMSUN* and *BCMS*, a large gene spanning all the CDRs and likely to represent non-coding RNA. Quantitative RT-PCR has demonstrated that many genes in this region are underexpressed compared with normal B cells. However, only *RFP2* showed significant loss of expression in B-CLL patients without a 13q deletion. No mutations have been found in any of these genes, suggesting that downregulation of one or more of these genes by an unknown mechanism may contribute to the pathogenesis of CLL. Whether the same mechanism contributes to the pathogenesis of myeloid malignancies with 13q deletions is not known.

Models of deletion syndromes

Loss and/or inactivation of candidate genes on 20q may be responsible for the pathogenesis of MPD by a number of possible mechanisms (Figure 9.6). A simple 'two-hit' model with a single target gene, reminiscent of Knudson's two-hit

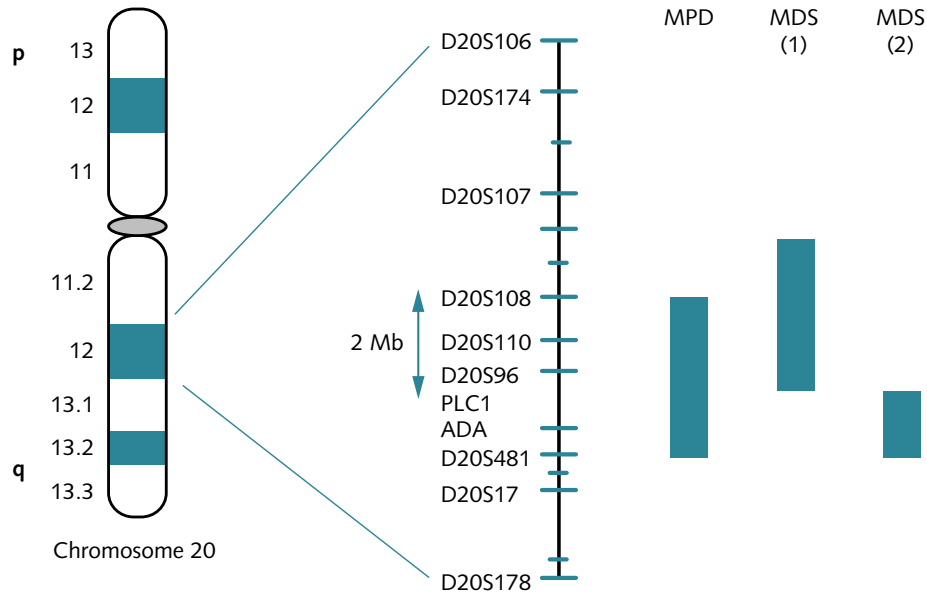


Fig. 9.5 Summary of common deleted regions on 20q for MPD and MDS
The MPD CDR is bordered by D20S108 and D20S481. Two distinct CDRs for MDS/AML have been published. MDS CDR (1) is flanked by PACs 620E11 and 196H17. MDS CDR (2) is flanked by PACs 29M7 and 179M20. Data taken from Bench *et al*. (2000) and Wang *et al*. (2000). Reproduced from Bench *et al*. (2001) with permission from Elsevier Science.

hypothesis, might involve inactivation of one copy of the gene by a subtle genetic alteration, such as a point mutation followed by loss of the second copy by deletion. Alternatively, the intact copy may be transcriptionally silenced; for example, by methylation, as has been demonstrated for the *VHL*, *p16* and *p15* genes. In a 'one-hit' model, loss of only a single copy of the gene may result in haploinsufficiency and be sufficient to contribute to disease pathogenesis. A number of tumor suppressor genes, such as *p27* and *p53*, show retention of one active allele in human tumors. Furthermore, heterozygous knockout mice develop tumors without any apparent inactivation of the wild-type allele. In addition, loss of two or more genes may be required (Figure 9.6). Again, inactivation of one or both copies of critical genes may be necessary to perturb progenitor cell behavior. Perhaps the inactivation of different combinations of genes is responsible for distinct myeloid disorders. The generation of mouse models of acquired deletions may offer ways of investigating such multigene hypotheses.

Gain of chromosomal material

Duplication of segments of 1q

Duplication of part of 1q has been demonstrated in a number of patients with MPD as well as other myeloid malignancies. It has been found at all stages of disease progression, including at diagnosis. Dupl(1q) was found in both erythroid and myeloid precursors of an MPD patient, confirming its origin within a multipotent progenitor. A common duplicated region spanning 1q23–q32 has been identified but, as yet, no molecular analysis has been undertaken. Duplication of the long arm of chromosome 1 can also result from a non-reciprocal translocation. For example, der(1)t(1;9) leads to trisomy 1q and 9p, both of which may be pathogenetically important. Similarly, monosomy 7 and der(1)t(1;7) results in monosomy 7q and trisomy 1q.

Trisomy 8

Trisomy 8 has been detected in 10–15% of patients with MDS, 5% of AML patients and in 35% of patients with CML blast crisis. The detection of trisomy 8 is particularly amenable to FISH-based techniques using centromeric probes and chromosome painting, which are particularly useful for the analysis of samples in which metaphases are absent or of poor quality.

The use of such techniques has provided a number of interesting observations. Firstly, in some patients it has been suggested that cells with trisomy 8 have a proliferative advantage over normal cells, at least in culture, since the frequency of trisomy 8 cells is greater in bone marrow metaphases than in interphase nuclei. Secondly, trisomy 8 has been detected in minor subclones of a small number of patients for whom conventional karyotyping showed no such aberration, indicating that trisomy 8 may be present in a greater proportion of patients than originally thought. As far as MPD is concerned,

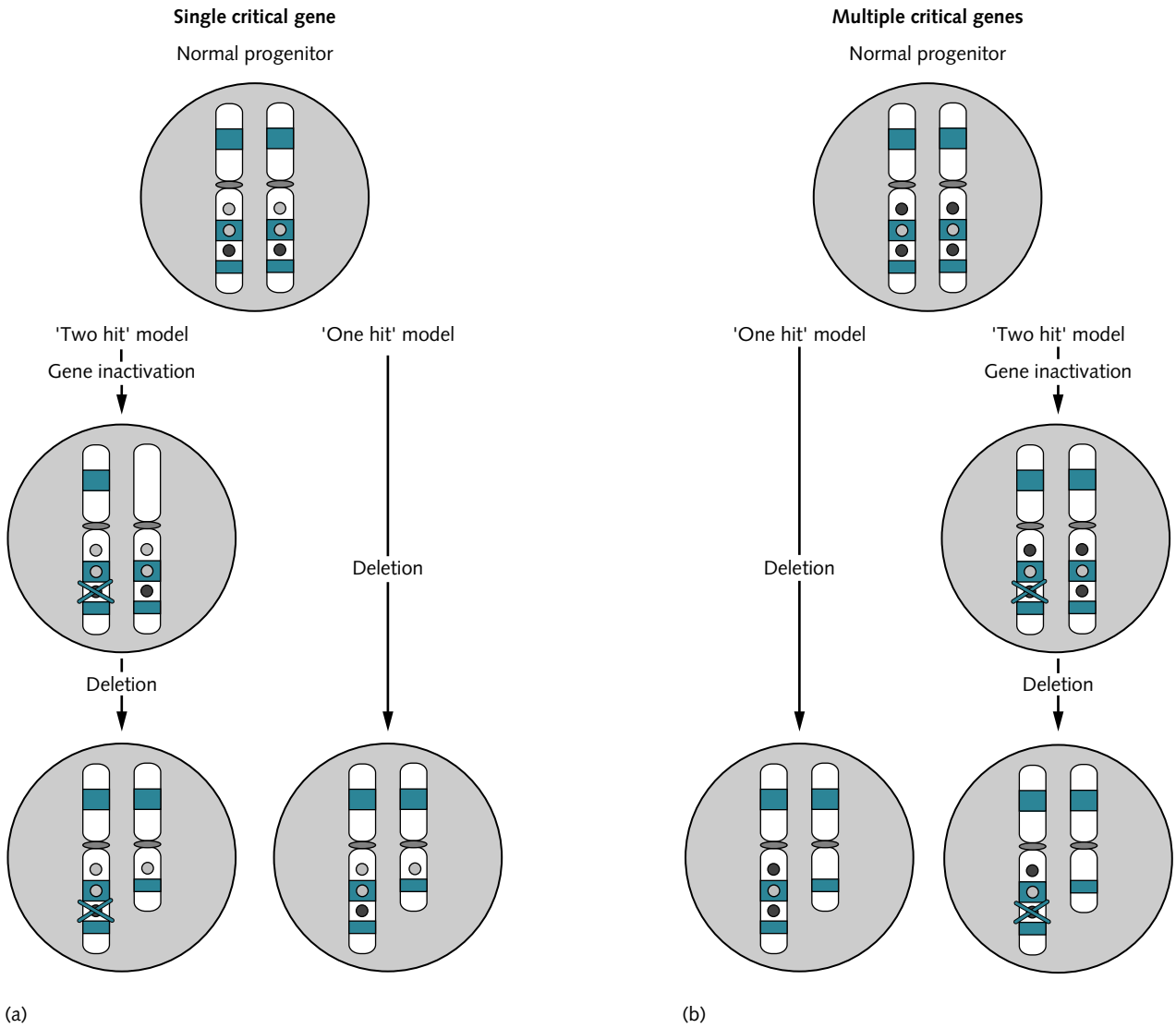


Fig. 9.6 Potential mechanism of the pathogenesis of deletions
(a) If there is a single target gene, for example, on chromosome 20q (●) it may be necessary for both copies to be lost/inactivated (two-hit model). Alternatively, loss of a single copy (haploinsufficiency) may be adequate to produce a phenotypic effect (one-hit model). (b) If there is more than one target gene, the one-hit model would entail haploinsufficiency for two or more genes, perhaps scattered over a large distance of the chromosome. In contrast, the two-hit model would involve biallelic inactivation of at least one of the target genes. Adapted from Asimakopoulos and Green (1996) with permission from Blackwell Science.

initial observations of subclones containing trisomy 8 in one-quarter of patients have not been confirmed by larger studies.

Combining FISH with immunophenotyping, it has been possible to follow the lineage involvement of trisomy 8. Price *et al.* (1992) demonstrated trisomy 8 in the majority of BFU-E and CFU-GM colonies from two patients with PV and trisomy 8. Furthermore, trisomy 8 was present in $CD34^+$ cells and mature myeloid cells but not in lymphoid cells. In AML patients, trisomy 8 has been detected in multipotent progenitor cells as well as in a subpopulation of flow-sorted lymphoid and erythroid cells. Therefore, it seems likely that, in MPD and MDS, trisomy 8, like del(20q), arises in primitive progenitor cells with myeloid and lymphoid potential. The genetic consequence of chromosome 8 amplification is not clear as no molecular mapping has yet been undertaken.

Abnormality of chromosome 9

As was the case for trisomy 8, suggestions that trisomy 9 may

be present in a significant number of patients with MPD have not been substantiated. However, most FISH studies use a chromosome 9 centromere probe and so amplification involving either the short or long arm only may be missed. As described above, trisomy 9p resulting from an unbalanced translocation with chromosome 1 is a recognized abnormality in MPD and gain of 9p due to other mechanisms has also been reported, suggesting that this region may be pathogenetically important. Interestingly, LOH of a large part of 9p has been demonstrated in six out of 20 patients with PV. No loss of material was observed in these patients, indicating that mitotic recombination rather than deletion was responsible for the 9p LOH. One gene from this region, *NFI-B*, was found to be overexpressed in patients with 9p LOH, but its pathogenetic significance remains unclear.

Translocations

The BCR-ABL fusion

The *BCR-ABL* fusion oncogene is a hallmark of CML. Using sensitive RT-PCR techniques, a number of reports have suggested that a proportion of patients with ET, PV or IMF express the *BCR-ABL* fusion transcript. However, this observation has not been universally confirmed and is likely to represent identification of very low levels of the transcript. Detection of such low levels of *BCR-ABL* transcripts is unlikely to be relevant to the pathogenesis of the MPDs since similar results have been obtained using blood from normal individuals.

The 8p11 myeloproliferative syndrome

The 8p11 syndrome is a rare, atypical MPD with a balanced translocation involving chromosome 8p11. Patients commonly present with lymphadenopathy and general malaise. Bone marrow and peripheral blood show myeloid proliferation, frequently with eosinophilia, and there is a high incidence of T-cell non-Hodgkin's lymphoma. Almost invariably, the disease progresses to AML and has a poor prognosis. Cytogenetic examination has revealed a t(8;13)(p11;q12) translocation or, more rarely, t(8;9)(p11;q34) or t(6;8)(q27;p11) translocations. This syndrome appears to result from transformation of a progenitor with both myeloid and lymphoid potential since cytogenetic analysis of lymph node and bone marrow from the same patient revealed the same chromosomal abnormality.

Molecular characterization of these translocations revealed that they all disrupt the *FGFR1* gene on 8p11. *FGFR1* encodes fibroblast growth factor receptor 1 and is a member of the receptor tyrosine kinase family. As a consequence, *FGFR1* becomes fused with three different partner genes: *ZNF198* on chromosome 13q12, *CEP110* at 9q34 and *FOP* at 6q27. Subsequently, a translocation t(8;22)(p11;q11) in patients with a CML-like disease was shown to result in a fusion between *BCR* and *FGFR1* and a rare t(8;19)(p12;q13) translocation fused a retrovirus gene, *HERV-K*, to *FGFR1* in a patient with atypical MPD. In addition, a further three translocations involving *FGFR1* have been reported, although partner genes have not been cloned.

Fusion proteins involving tyrosine kinases are a common theme amongst myeloid malignancies (Figure 9.7). The first to be described was the *BCR-ABL* fusion gene generated by the t(9;22) translocation in CML. The BCR-ABL fusion protein contains the coiled-coil oligomerization motif from BCR fused to the majority of the ABL protein containing a tyrosine kinase domain. The t(5;12) translocation in patients with chronic myelomonocytic leukemia (CMML) results in the fusion of the helix–loop–helix domain of the ETS-like transcription factor, TEL, with the tyrosine kinase domain of the PDGFRβ protein. In a similar fashion, ZNF198-FGFR1 contains a proline-rich oligomerization domain from ZNF198 fused to the tyrosine kinase domain of FGFR1. Thus, a common activation mechanism links these fusion products. Ligand-independent dimerization leads to constitutive tyrosine kinase activation and phosphorylation of downstream targets. These oncogenic products are also potentially linked

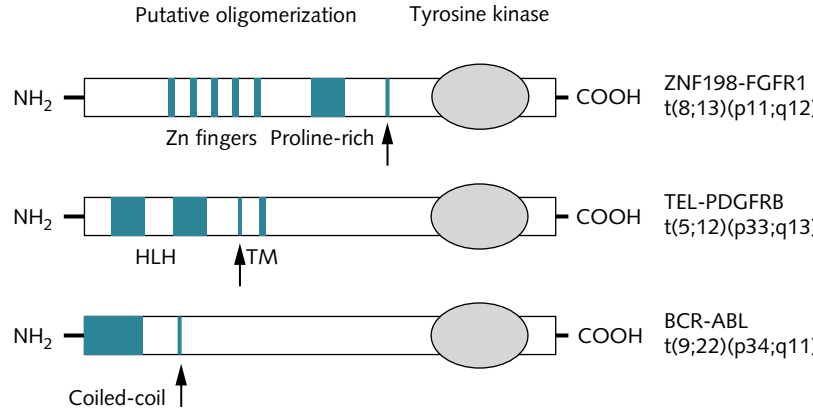


Fig. 9.7 Comparison of fusion proteins involving tyrosine kinase proteins in myeloid malignancies
Each fusion protein includes an oligomerization domain plus a tyrosine kinase domain. Breakpoints are indicated by an arrow. HLH, helix–loop–helix domain; TM, transmembrane domain. Reproduced from Bench *et al*. (2001) with permission from Elsevier Science.

to the RAS signaling pathway. In common with BCR-ABL and TEL-PDGFRB, ZNF198-FGFR1 possesses transforming ability.

Rare translocations in MPD

Balanced translocations are rare in 'true' MPD (PV, ET, IMF), although a small number have been reported. Such rare events may mark the site of pathogenetically important genes, the identification of which could offer short cuts to the understanding of the molecular pathogenesis of MPD.

Identification of novel cytogenetic changes

New molecular cytogenetic techniques now available offer the researcher the ability to analyze cytogenetic changes in much greater detail. In particular, color karyotyping (M-FISH) permits the identification of cryptic translocations not previously detectable by G-banding and also reveals the chromosomal origin of genetic material present in marker chromosomes. Comparative genomic hybridization is a complementary technique which detects quantitative changes (gain or loss of sequences) throughout the genome without the need to obtain metaphase chromosomes. These techniques have been applied to cases of PV with a normal karyotype. However, no cryptic cytogenetic changes have been detected.

Molecular genetic alterations

Clues about the molecular pathogenesis of the MPDs have also been obtained from the biology of the diseases. Prchal and Axelrad (1974) first demonstrated that erythropoietin (EPO)-independent BFU-E colonies could be grown from PV patients. Spontaneous BFU-E colonies have been observed in the majority of PV patients in addition to a large number of ET patients. Similarly, spontaneous CFU-MEG colony formation occurs, without the addition of exogenous growth factors, in most ET patients. Although it is now believed that progenitors are hypersensitive to a number of growth factors, these results have focussed attention on the EPO and TPO signaling pathway (Figure 9.8).

Receptors and ligands

Mutations in the erythropoietin receptor gene, *EPOR*, have been found in some families with primary familial and congenital polycythemia (PFCP). Most of these mutations result in truncation of the polypeptide and loss of part of the cytoplasmic negative regulatory domain, although missense mutations in the same region have also been reported. The finding of *EPOR* mutations in some cases of PFCP stimulated

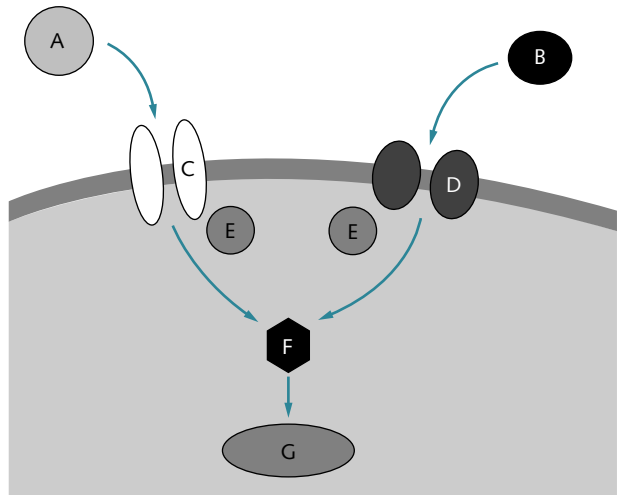


Fig. 9.8 Possible defects in signal transduction in MPD
Progenitor cells from patients are hypersensitive to a number of cytokines (A and B), such as EPO, which act through specific receptor molecules (C and D). There may be a defect in a common receptor component (E), a common signaling intermediate (F) or a common effector molecule (G) such as a transcription factor. Reproduced from Hinshelwood *et al*. (1997) with permission from Elsevier Science.

an intensive search for acquired mutations of the *EPOR* gene in PV. However, no acquired mutations in the *EPOR* gene were identified in patients with PV or other MPDs. Interestingly, not all families with PFCP have *EPOR* mutations, suggesting the existence of other target genes, in which mutations can give rise to polycythemia. An autosomal recessive inherited polycythemia endemic in the Russian region of Chuvashia results from mutation of the *VHL* gene (von Hippel–Lindau) on chromosome 3p25. The normal role of the VHL protein is to degrade HIF-1α (hypoxia inducible factor 1 α subunit) by ubiquitination. The decreased association of VHL with HIF-1α in Chuvash polycythemia leads to an increased level of HIF-1α, the upregulation of downstream target genes, including *EPO*, and increased erythropoiesis.

Rare examples of familial thrombocythemia exist. In a number of such families, a mutation in the 5′ UTR of the *TPO* (thrombopoietin) gene has been detected. The consequence of this mutation is to increase the expression level of TPO, leading to increased megakaryopoiesis. Other families show no linkage to either the *TPO* or the *C-MPL* (TPO receptor) locus, implying the existence of other genetic loci which contribute to familial thrombocythemia. No acquired mutations in the *TPO* gene have been found in patients with acquired ET. In addition, no acquired mutations in the *C-MPL* gene have been detected in ET or PV.

Despite the lack of *C-MPL* mutations in MPD, there is some evidence that C-MPL is defective in a proportion of MPD patients. Platelets derived from some PV patients show reduced

amounts of C-MPL protein when assessed with an antibody to the C-terminal (cytoplasmic) end of the protein. Using an antibody to the N-terminal end, Moliterno and Spivak (1999) demonstrated the presence, at a reduced level, of two isoforms of C-MPL in PV platelets. One isoform represented normal C-MPL and the other was the result of defective C-MPL glycosylation. Platelets from patients with reduced C-MPL expression failed to show the appropriate protein tyrosine phosphorylation in response to TPO, indicating a defect in C-MPL signaling. In ET, reduced amounts of C-MPL protein and mRNA have been demonstrated in the platelets of some but not all patients. However, results are inconsistent between groups, possibly because of different sources of antibodies or pathogenetic heterogeneity. Despite this, decreased C-MPL expression within bone marrow megakaryocytes may help distinguish some PV and ET patients from patients with reactive conditions.

Given the lack of mutations in acquired MPDs in EPO and TPO and their receptors, attention turned to other growth factors that may play a role in MPD. TGF-β1 is a strong negative regulator of early hemopoietic stem cells and, in particular, plays an important role in the regulation of megakaryopoiesis. TGF-β1 is itself largely produced by megakaryocytes and platelets, and negatively regulates megakaryopoiesis. Within the bone marrow, TGF-β1 stimulates the production of TPO, which in turn induces the expression of TGF-βRI and TGF-βRII by megakaryocyte progenitors. This renders them susceptible to the inhibitory effects of TGF-β1, completing the negative feedback loop. Megakaryocyte progenitors from ET patients show reduced sensitivity to the inhibitory effects of TGF-β1, which may in part be responsible for the increased rate of platelet production. One possible mechanism for this effect is reduced activity of TGF-βRII. Expression of TGF-βRII is decreased in $CD34^+$ cells from IMF patients and in neutrophils from PV, ET and CML patients. Acquired mutations are present in many solid tumors, leading to a truncated protein that lacks the kinase domain and is insensitive to the effect of TGF-β1. No acquired mutations of TGF-βRII have been detected in PV or ET patients and methylation of the promoter and CpG island has been ruled out as a mechanism for reduced expression, at least for PV. Further studies of the TGF-β1 pathway in MPD are warranted.

Transcript levels of a novel cell surface receptor termed *PRV-1* (also termed *NB1* and *CD177*) are increased in granulocytes from most PV patients and in a significant number of ET patients. In normal individuals, *PRV-1* mRNA is expressed in bone marrow cells but its level is reduced in differentiated neutrophils. The PRV-1 protein is a member of the uPAR/CD59/Ly6 family, a group of cell membrane proteins attached to the extracellular membrane by a lipid anchor. These proteins play a role in signal transduction including, but not exclusively, activation of the JAK/STAT pathway, although the specific function of PRV-1 is not known. Despite the significant increase in *PRV-1* mRNA levels in PV granulocytes, levels of the cell surface protein are not significantly different between patients and normal individuals. Detection of increased *PRV-1* transcript levels in granulocytes by quantitative RT-PCR appears to discriminate patients with PV from normal individuals and those with reactive erythrocytosis. No structural alterations of the *PRV-1* gene have been observed in MPD patients, suggesting that *PRV-1* overexpression is likely to be a secondary consequence of the acquired genetic defect that causes PV.

Signaling pathways

In retrospect, lack of receptor and ligand mutations may not be surprising. It is now recognized that PV progenitor cells are abnormally responsive to multiple growth factors. This suggests that the defect may be in a common signaling pathway (Figure 9.8). Several pathways are involved in processing the response to a cytokine. Each pathway incorporates a complex array of both activating and inhibitory components to ensure an appropriate response is achieved. For example, the EPO/EPOR pathway involves activation of the JAK2/STAT5 and other signal transduction pathways. Negative regulators, such as SHP-1 and the SOCS proteins, ensure that the cellular response to EPO is controlled and appropriate.

Only a small number of signal transduction components have been assessed for mutations in MPD. The SHP-1 phosphatase was considered a good candidate since it interacts with a number of cytokine receptors, including EPOR, and negatively regulates EPO-induced signal transduction. Hence, such a phosphatase could be a candidate tumor suppressor protein. No mutations were detected in PV or ET patients, and transcript, protein and promoter methylation status were normal. Signaling via TGF-β1 is achieved through a series of SMAD proteins. Mutations of *SMAD2* and *SMAD4* are common in solid tumors, such as colorectal cancer. However, mutations of the *SMAD* genes have not been detected in hematological malignancies.

Other changes in MPD

Mutations in the *p53* and *RAS* genes have been observed in a small proportion of patients with an MPD. However, most such mutations were found in patients with a later or acute phase of the disease. Therefore, mutations of *p53* and *RAS* genes are more frequently associated with disease progression and transformation to acute leukemia rather than initiation of the MPD itself.

In addition to positional cloning strategies, several other approaches are being pursued to identify genetic defects that contribute to the pathogenesis of MPD. One approach is to

search for differentially expressed genes between cells from normal individuals and PV patients. This approach has a number of inherent problems. Progenitor cell populations are heterogeneous, so differences in gene expression may be secondary to shifts in the composition of the progenitor cell compartment. Comparison of neutrophils uses a more homogeneous population but is unlikely to identify primary pathogenetic changes. Secondly, genes that are affected by point mutations but whose expression level is unchanged will not be identified. However, even if the primary target is not identified, this approach may uncover consistent secondary changes which may be helpful diagnostically and shed light on the altered behavior of progenitor cells in PV. Such an approach was used to identify the overexpression of *PRV-1* in PV granulocytes.

Acknowledgments

Work in the authors' laboratory is supported by the Leukaemia Research Fund, the Medical Research Council and the Wellcome Trust.

Further reading

Classification

Dameshek W. (1951) Some speculations on the myeloproliferative syndromes. *Blood*, **6**, 372–375.

Jaffe ES, Harris NL, Stein H, Vardiman JW. (2001) *World Health Organization Classification of Tumors. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon: IARC Press.

Pearson TC. (1998) Diagnosis and classification of erythrocytosis and thrombocytosis. *Baillière's Clinical Haematology*, **11**, 695–720.

Reilly JT. (1998) Pathogenesis and management of idiopathic myelofibrosis. *Baillière's Clinical Haematology*, **11**, 751–767.

X-linked clonality assays

Adamson JW, Fialkow PJ, Murphy S *et al.* (1976) Polycythemia vera: stem cell and probable clonal origin of the disease. *New England Journal of Medicine*, **295**, 913–916.

Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. (1992) Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *American Journal of Human Genetics*, **51**, 1229–1239.

Busque L, Mio R, Mattioli J *et al.* (1996) Nonrandom X-inactivation patterns in normal females: Lyonization ratios vary with age. *Blood*, **88**, 59–65.

Champion KM, Gilbert JGR, Asimakopoulos FA, Hinshelwood S, Green AR. (1997) Clonal haemopoiesis in normal elderly women: implications for the myeloproliferative disorders and myelodysplastic syndromes. *British Journal of Haematology*, **97**, 920–926.

El-Kassar N, Hetet G, Briere J *et al.* (1997) Clonality analysis of hematopoiesis in essential thrombocythemia: advantages of studying T lymphocytes and platelets. *Blood*, **89**, 128–134.

Harrison CN, Gale RE, Machin SJ *et al.* (1999) A large proportion of patients with a diagnosis of essential thrombocythemia do not have a clonal disorder and may be at lower risk of thrombotic complications. *Blood*, **93**, 417–424.

Vickers MA, McLeod E, Spector TD *et al.* (2001) Assessment of mechanism of acquired skewed X inactivation by analysis of twins. *Blood*, **97**, 1274–1281.

Chromosome abnormalities

Asimakopoulos FA, Green AR. (1996) Deletions of chromosome 20q and the pathogenesis of myeloproliferative disorders. *British Journal of Haematology*, **95**, 219–226.

Bench AJ, Nacheva EP, Champion KM, Green AR. (1998) Molecular genetics and cytogenetics of myeloproliferative disorders. *Baillière's Clinical Haematology*, **11**, 819–848.

Bench AJ, Cross NC, Huntly BJ *et al.* (2001) Myeloproliferative disorders. *Best Practice & Research. Clinical Haematology*, **14**, 531–551.

Berger R, Bernheim A, LeConiat M *et al.* (1984) Chromosome studies in polycythemia vera patients. *Cancer Genetics & Cytogenetics*, **12**, 217–223.

Demory JL, Dupriez B, Fenaux P *et al.* (1988) Cytogenetic studies and their prognostic significance in agnogenic myeloid metaplasia: a report on 47 cases. *Blood*, **72**, 855–859.

Diez-Martin JL, Graham DL, Petitt RM *et al.* (1991) Chromosome studies in 104 patients with polycythemia vera. *Mayo Clinic Proceedings*, **66**, 287–299.

Dupriez B, Morel P, Demory JL *et al.* (1996) Prognostic factors in agnogenic myeloid metaplasia: a report on 195 cases with a new scoring system. *Blood*, **88**, 1013–1018.

Mertens F, Johansson B, Heim S *et al.* (1991) Karyotypic patterns in chronic myeloproliferative disorders: report on 74 cases and review of the literature. *Leukemia*, **5**, 214–220.

Rege-Cambrin G, Mecucci C, Tricot G *et al.* (1987) A chromosomal profile of polycythemia vera. *Cancer Genetics & Cytogenetics*, **25**, 233–245.

Reilly JT, Snowden JA, Spearing RL *et al.* (1997) Cytogenetic abnormalities and their prognostic significance in idiopathic myelofibrosis: a study of 106 cases. *British Journal of Haematology*, **98**, 96–102.

Swolin B, Weinfeld A, Westin J. (1988) A prospective long-term cytogenetic study in polycythemia vera in relation to treatment and clinical course. *Blood*, **72**, 386–395.

Tefferi A, Mesa RA, Schroeder G *et al.* (2001) Cytogenetic findings and their clinical relevance in myelofibrosis with myeloid metaplasia. *British Journal of Haematology*, **113,** 763–771.

Chromosome deletions

Asimakopoulos FA, Holloway TL, Nacheva EP *et al.* (1996) Detection of chromosome 20q deletions in bone marrow metaphases but not peripheral blood granulocytes in patients with myeloproliferative disorders or myelodysplastic syndromes. *Blood*, **87**, 1561–1570.

Bench AJ, Nacheva EP, Hood TL *et al.* (2000) Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. *Oncogene*, **19**, 3902–3913.

Cook WD, McCaw BJ. (2000) Accommodating haploinsufficient tumor suppressor genes in Knudson's model. *Oncogene*, **19**, 3434–3438.

Dewald GW, Schad CR, Lilla VC *et al.* (1993) Frequency and photographs of HGM11 chromosome anomalies in bone marrow samples from 3996 patients with malignant hematologic neoplasms. *Cancer Genetics & Cytogenetics*, **68**, 60–69.

Greenberg P, Cox C, LeBeau MM *et al.* (1997) International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*, **89**, 2079–2088.

Hollings PE, Rosman I, Beard MEJ. (1994) A 20q deletion originating in a pluripotent stem cell. *Blood*, **83**, 305–306.

Kurtin PJ, Dewald GW, Shields DJ *et al.* (1996) Hematologic disorders associated with deletions of chromosome 20q. *American Journal of Clinical Pathology*, **106**, 680–688.

La Starza R, Wlodarska I, Aventin A *et al.* (1998) Molecular delineation of 13q deletion boundaries in 20 patients with myeloid malignancies. *Blood*, **91**, 231–237.

Mertens D, Wolf S, Schroeter P *et al.* (2002) Down-regulation of candidate tumor suppressor genes within chromosome band 13q14.3 is independent of the DNA methylation pattern in B-cell chronic lymphocytic leukemia. *Blood*, **99**, 4116–4121.

Mori N, Morosetti R, Lee S *et al.* (1997) Allelotype analysis in the evolution of chronic myelocytic leukemia. *Blood*, **90**, 2010–2014.

Pastore C, Nomdedeu J, Volpe G *et al.* (1995) Genetic analysis of chromosome 13 deletions in Bcr/Abl negative chronic myeloproliferative disorders. *Genes, Chromosomes and Cancer*, **14**, 106–111.

Wang PW, Eisenbart JD, Espinosa III R, *et al.* (2000) Refinement of the smallest commonly deleted segment of chromosome 20 in malignant myeloid diseases and development of a PAC-based physical and transcription map. *Genomics*, **67**, 28–39.

White NJ, Nacheva E, Asimakopoulos FA *et al.* (1994) Deletion of chromosome 20q in myelodysplasia occurring in multipotent precursor of both myeloid cells and B cells. *Blood*, **83**, 2809–2816.

Gain of chromosomal material

Jenkins RB, Le Beau MM, Kraker WJ *et al.* (1992) Fluorescence *in situ* hybridization: a sensitive method for trisomy 8 detection in bone marrow specimens. *Blood*, **79**, 3307–3315.

Knuutila S, Ruutu T, Partanen S *et al.* (1983) Chromosome 1q+ in erythroid and granulocyte-monocyte precursors in a patient with essential thrombocythemia. *Cancer Genetics & Cytogenetics*, **9**, 245–249.

Kralovics R, Guan Y, Prchal JT. (2002) Acquired uniparental disomy of chromosome 9p is a frequent stem cell defect in polycythemia vera. *Experimental Hematology*, **30**, 229–236.

Mehrotra B, George TI, Kavanau K *et al.* (1995) Cytogenetically aberrant cells in the stem cell compartment (CD34+lin−) in acute myeloid leukemia. *Blood*, **86**, 1139–1147.

Price CM, Kanfer EJ, Colman SM *et al.* (1992) Simultaneous genotypic and immunophenotypic analysis of interphase cells using dual-colour fluorescence: a demonstration of lineage involvement in polycythemia vera. *Blood*, **80**, 1033–1038.

Swolin B, Weinfeld A, Westin J. (1986) Trisomy 1q in polycythemia vera and its relation to disease transition. *American Journal of Hematology*, **22**, 155–167.

Translocations

Bench AJ, Cross NC, Huntly BJ *et al.* (2001) Myeloproliferative disorders. *Best Practice & Research. Clinical Haematology*, **14,** 531–551.

Cross NC, Reiter A. (2002) Tyrosine kinase fusion genes in chronic myeloproliferative diseases. *Leukemia*, **16**, 1207–1212.

Guasch G, Mack GJ, Popovici C *et al.* (2000) FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33). *Blood*, **95**, 1788–1796.

Guasch G, Popovici C, Mugneret F *et al.* (2003) Endogenous retroviral sequence is fused to FGFR1 kinase in the 8p12 stem cell myeloproliferative disorder with t(8;19)(p12;q13.3). *Blood*, **101**, 286–288.

Popovici C, Zhang B, Gregoire MJ *et al.* (1999) The t(6;8)(q27;p11) translocation in a stem cell myeloproliferative disorder fuses a novel gene, FOP, to fibroblast growth factor receptor 1. *Blood*, **93**, 1381–1389.

Smedley D, Hamoudi R, Clark J *et al.* (1998) The t(8;13)(p11;q11–12) rearrangement associated with an atypical myeloproliferative disorder fuses the fibroblast growth factor receptor 1 gene to a novel gene RAMP. *Human Molecular Genetics*, **7**, 637–642.

Xiao S, McCarthy JG, Aster JC *et al.* (2000) ZNF198-FGFR1 transforming activity depends on a novel proline-rich ZNF198 oligomerization domain. *Blood*, **96**, 699–704.

Xiao S, Nalabolu SR, Aster JC *et al.* (1998) FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. *Nature Genetics*, **18**, 84–87.

Molecular genetic alterations

Ang SO, Chen H, Gordeuk VR *et al.* (2002) Endemic polycythemia in Russia: mutation in the VHL gene. *Blood Cells Molecules & Diseases*, **28**, 57–62.

Asimakopoulos FA, Hinshelwood S, Gilbert JGR *et al.* (1997) The gene encoding hematopoietic cell phosphatase (SHP-1) is structurally and transcriptionally intact in polycythemia vera. *Oncogene*, **14**, 1215–1222.

Axelrad AA, Eskinazi D, Correa PN *et al.* (2000) Hypersensitivity of circulating progenitor cells to megakaryocyte growth and development factor (PEG-rHu MGDF) in essential thrombocythemia. *Blood*, **96**, 3310–3321.

Correa, PN, Eskinazi D, Axelrad AA. (1994) Circulating erythroid progenitors in polycythemia vera are hypersensitive to insulin-like growth factor-1 *in vitro*: studies in an improved serum-free medium. *Blood*, **83**, 99–112.

Gregg XT, Prchal JT. (1997) Erythropoietin receptor mutations and human disease. *Seminars in Hematology*, **34**, 70–76.

Hinshelwood S, Bench AJ, Green AR. (1997) Pathogenesis of polycythaemia vera. *Blood Reviews*, **11**, 224–232.

Horikawa Y, Matsumura I, Hashimoto K *et al.* (1997) Markedly reduced expression of platelet c-mpl receptor in essential thrombocythemia. *Blood*, **90**, 4031–4038.

Mesa RA, Hanson CA, Li C *et al.* (2002) Diagnostic and prognostic value of bone marrow angiogenesis and megakaryocyte c-Mpl expression in essential thrombocythaemia. *Blood*, **99**, 4131–4137.

Moliterno AR, Hankins WD, Spivak JL. (1998) Impaired expression of the thrombopoietin receptor by platelets from patients with polycythemia vera. *New England Journal of Medicine*, **338**, 572–580.

Moliterno AR, Spivak JL. (1999) Posttranslational processing of the thrombopoietin receptor is impaired in polycythemia vera. *Blood*, **94**, 2555–2561.

Prchal JF, Axelrad AA. (1974) Bone marrow responses in polycythemia vera. *New England Journal of Medicine*, **290**, 1382.

Temerinac S, Klippel S, Strunck E *et al.* (2000) Cloning of PRV-1, a novel member of the uPAR receptor superfamily, which is overexpressed in polycythemia rubra vera. *Blood*, **95**, 2569–2576.

Wiestner A, Schlemper RJ, van der Maas APC *et al.* (1998) An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. *Nature Genetics*, **18**, 49–52.

Chapter 10 Lymphoid neoplasms

Anthony G Letai & John G Gribben

Introduction

As with all cancers, lymphomas were originally categorized primarily on morphology and clinical behavior. The use of antibodies against cell surface markers allowed the study of lymphoma specimens with antibody panels that could, along with morphological criteria, usually place a given lymphoma into a diagnostic category. Even within a given lymphoma category, however, there is considerable heterogeneity of clinical behavior. A prominent example is the category of diffuse large B-cell lymphomas, of which approximately 40% are cured with chemotherapy but 60% of patients die of disease, usually within a few years of diagnosis. In this case, as with many of the lymphomas, a prognostic index (the International Prognostic Index or IPI) based on a few pretreatment criteria is able to subdivide the category and provide very useful prognostic information. However, even within IPI classes, significant clinical heterogeneity persists. Furthermore, it is likely that the IPI defines subclasses of lymphomas according to biological differences among these lymphomas. Studies of genetic abnormalities are proving important tools for the improved classification and prognostication of diseases. In addition, a better understanding of the molecular pathophysiology of the disease will likely lead to improvements in treatment of lymphoma.

Techniques

Techniques for studying genetic abnormalities in tumor specimens have undergone a revolution in the past 5–10 years (Table 10.1). Initial genetic analyses were based on the technique of chromosomal study by Giemsa–trypsin banding. In these studies, cells are grown in short-term culture, usually in the presence of mitogens. Colcemid treatment results in the accumulation of cells in metaphase, at which point the cells are fixed and dropped onto glass slides. The slides are treated with trypsin followed by Giemsa to give a banding pattern. An experienced cytogenetic technician can then identify normal chromosomes, translocations, numerical abnormalities, and sometimes more subtle deletions. The technique can identify only genetic changes large enough to disrupt a Giemsa-stained band.

Table 10.1 Techniques to study lymphoma genetics.

Cytogenetic analysis
Southern blot analysis
Polymerase chain reaction (PCR) analysis
Fluorescent *in situ* hybridization (FISH)
Comparative genomic hybridization (CGH)
CGH microarray
Gene expression profiling
Proteomic profiling

More modern techniques are able to detect abnormalities with greater sensitivity. Southern hybridization starts with the electrophoretic separation of tumor DNA on a gel, followed by transfer to a membrane. This membrane is then probed with radioactively labeled polynucleotide probes specific for certain genes of interest. Changes in the expected size or intensity of the band of interest can indicate mutation, translocation, amplification, or deletion of the gene of interest.

Polymerase chain reaction (PCR) technology has allowed the detection of genetic abnormalities using only a small amount of tumor DNA. PCR for the detection of lymphoma cells is discussed in more detail in Chapter 6. Using primers designed to flank the genomic region of interest, repetitive cycles of annealing, DNA polymerization and thermal melting eventually yield a PCR product. The presence and size of this product may be analyzed by gel electrophoresis to determine

the presence of a translocation. Furthermore, a PCR product may be sequenced to look for point mutations.

Fluorescent *in situ* hybridization (FISH) uses fluorescently labeled DNA probes to bind to specific regions of genomic DNA. Images are then analyzed under a fluorescence microscope. Numerical chromosomal abnormalities may be detected by simply counting the number of signals per cell: more than two indicates the addition of a chromosome, whereas fewer than two indicates a deletion (Plate 10.1). To investigate a potential translocation, two probes are used, one to detect the genomic DNA on each side of the known translocation. If the two probes are consistently approximated, this indicates the presence of a translocation (Plate 10.2). FISH may be performed on interphase cells, so that growth in culture is not a requirement for this type of analysis as it is for conventional cytogenetics. Tests for small deletions or other more subtle abnormalities may be better performed on metaphase cells. FISH requires knowledge of the area to be labeled.

Since they rely on the annealing of a labeled specific DNA probe or primer, Southern hybridization, PCR and FISH are techniques to determine the presence or absence of a *known* genetic abnormality. Two modern techniques that provide a genome-wide scan for abnormalities and require no prior suspicion of a particular abnormality are comparative genomic hybridization (CGH) and gene expression profiling (GEP). These are beginning to have a clinical impact.

In the original CGH techniques (Figure 10.1), DNA is isolated from the tumor sample and a normal control sample. The DNA in each is labeled with a different fluorescent dye; for example, green for the tumor DNA and red for the normal DNA. These samples are then mixed and hybridized onto slides of metaphase spreads of normal cells. Images of metaphase spreads are then analyzed for the green:red color ratio. Regions of chromosomes that have a high green:red ratio contain a putative area of amplification. Regions that have a low green:red ratio contain a putative deletion. In this way, the entire genome may be examined for abnormalities. Other techniques, generally beyond what is performed in clinical laboratories, are required to determine the critical genes involved in areas of amplification and deletion. Small abnormalities and balanced translocations cannot be observed using this technique. In a variation of the CGH technique, DNA is hybridized to defined arrays of genomic DNA fragments (Figure 10.2). These arrays can cover the genome, and this improves resolution to the size of the DNA fragments within the microarray, currently down to the level of 1 Mb.

GEP is described in more detail in Chapter 24. This technique allows the comprehensive, quantitative examination of the mRNA transcripts of a tumor sample, a group of molecules that has been termed the 'transcriptome'. In this technique, mRNA is purified from a fresh or frozen tumor sample. Formalin-fixed tissue cannot be used. It is important that, when a group of samples is being compared, tissue acquisition, mRNA preparation and all subsequent steps are performed as identically as possible. When possible, steps should be performed on all samples in parallel, with identical reagents, and simultaneously. Techniques exist to amplify very small amounts of mRNA to obtain usable quantities; while

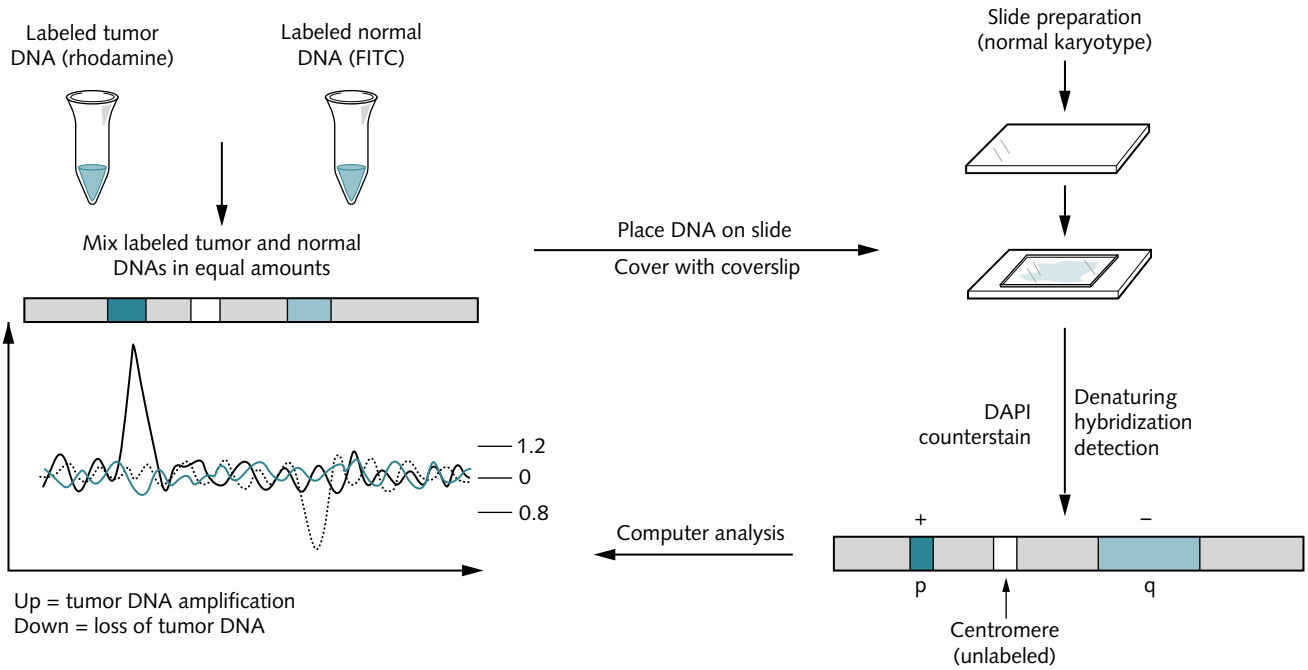


Fig. 10.1 Comparative genomic hybridization

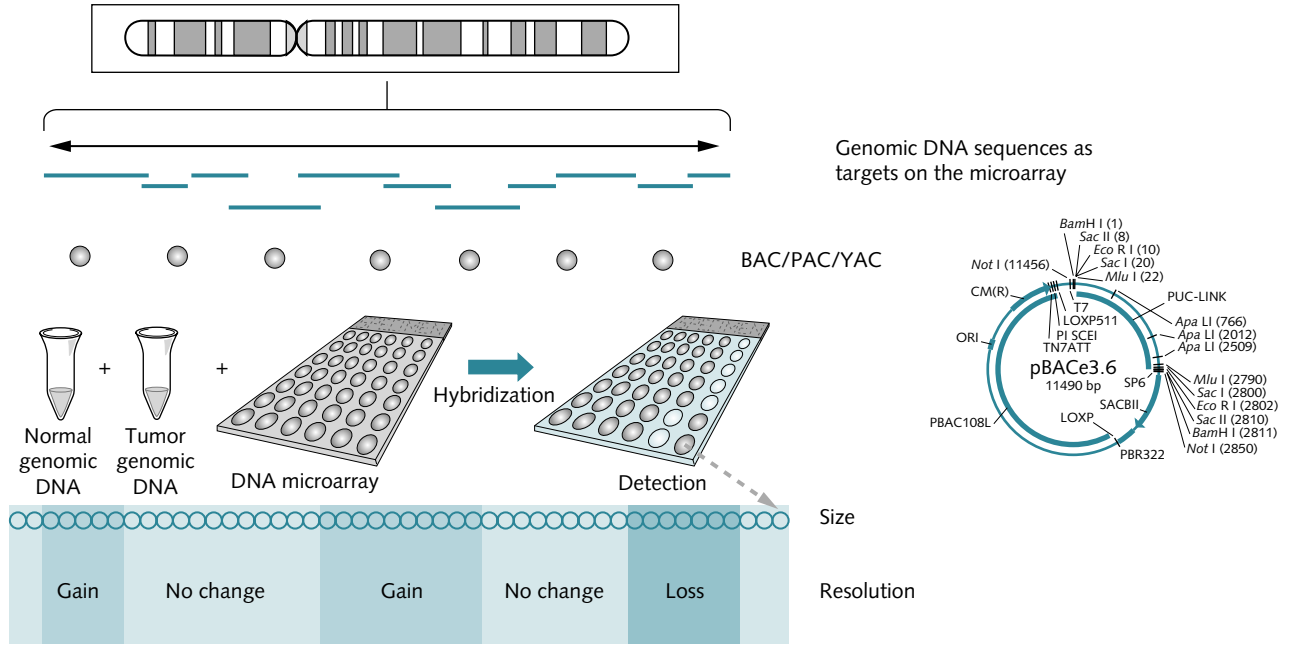


Fig. 10.2 Array comparative genomic hybridization (array CGH)

some have reported that this may be done without bias to the relative quantities of transcripts, this may not be universally true. The mRNA is reverse-transcribed, then transcribed with fluorescently labeled nucleotides to develop a fluorescently labeled complementary RNA (cRNA) representation of the original mRNA mixture. The labeled cRNA is then used for hybridization to immobilized, indexed oligonucleotide or cDNA probes. The signals at each of the loci on the slide or 'gene chip' may then be quantitated by microscopy and image analysis software. The strength of fluorescent signal may then be related to the abundance of a particular mRNA in the original sample. Comparison of different tumor samples and comparison with wild-type can then allow the determination of transcripts that are over- or under-represented in certain conditions or tumors. Tumors may be categorized using these profiles, and subgroups of messages may be used to create predictors of clinical behavior. This powerful technique has the potential to analyze an entire transcriptome of tens of thousands of genes simultaneously. It cannot determine the genomic abnormalities that lead to the differences in expression pattern, however. Given the amount of data this technique can encompass, it is possible that it will some day be a part of the routine pathological analysis of cancers, providing, like conventional pathology today, categorical and prognostic information, and possibly even directing therapeutic decision-making. The technique is currently limited in its clinical application by its relatively high cost and the relatively limited testing on tumor samples for which ample clinical details are available.

Types of genetic abnormality

The types of genetic abnormality found in lymphoma may be crudely divided into two main classes: those that foster increased proliferation, and those that inhibit programmed cell death, or apoptosis. The classical gene in lymphomagenesis that induces proliferation is c-*MYC*. Burkitt's lymphoma, one of the most rapidly dividing lymphomas, is the archetype of a lymphoma that overexpresses c-MYC by the t(8;14). c-*MYC* is a helix–loop–helix leucine zipper transcription factor which requires heterodimerization with the protein MAX to activate transcription and induce proliferation. Targets of this dimer include genes controlling cell cycle progression, cell growth, metabolism, differentiation, and apoptosis. The net effect of c-MYC expression is generally an increase in proliferation; however, this effect is context-specific. In some cells, c-MYC overexpression can induce cell cycle arrest or apoptosis via p53. Therefore, it may require an apoptotic defect to permit c-MYC overexpression.

BCL-2 is an oncogene that does not directly foster increased proliferation, but rather opposes apoptosis. It does this at least in part by binding and sequestering pro-apoptotic BCL-2 family members, preventing them from communicating or executing death signals, especially at the mitochondrion. It is classically overexpressed in the indolent follicular lymphoma due to the t(14;18). Other apoptotic defects often found in lymphoma include those allowing the activation or stabilization of NF-κB transcription factors.

A frequent hallmark of translocations found in B-cell lymphomas is their exploitation of immunoglobulin (Ig) gene regulatory elements to drive expression of an oncogene in a malignant B cell or B-cell precursor. Burkitt's lymphoma is an example of a lymphoma characterized by the overexpression of c-MYC. While the most common translocation is t(8;14), which puts c-*MYC* under the control of the Ig heavy chain (IgH) transcription elements, the less common t(2;8) and t(8;22) are also found, putting c-*MYC* transcription under the control of the light-chain κ and λ transcription elements, respectively. The t(14;18) found in follicular lymphoma drives BCL-2 expression using IgH transcription elements. The BCL-6 expression found in many DLBL cases is often driven by IgH, Igκ, and Igλ elements in the t(3;14), t(2;3) and t(3;22) respectively. PAX 5 expression in lymphoplasmacytoid lymphoma and cyclin D1 expression in mantle cell lymphoma are likewise driven by the t(9;14) and t(11;14) which exploit the IgH locus.

Improved techniques of genetic study have allowed the identification of a large number of chromosomal translocations, the most common of which are shown in Table 10.2. Those abnormalities which are the most common, or which have been demonstrated to have the greatest impact on prognosis or treatment, are described. Figure 10.3 shows the molecular pathogenesis, putative cell of origin within B-cell development within the lymph node and germinal center, and the immunophenotype of the most common types of lymphomas. The molecular pathogenesis of chronic lymphocytic leukemia/small lymphocytic lymphoma remains unknown.

Burkitt's lymphoma

Burkitt's lymphoma is a very high-grade B-cell malignancy. Pathologically it is characterized by small, non-cleaved cells. The presence of many apoptotic malignant cells gives rise to tangible-body macrophages and the 'starry sky' appearance characteristic of this and other very rapidly dividing tumors. Frequent mitotic figures demonstrate the rapid cell division characteristic of this tumor. Though rapidly dividing, it is one of the most curable lymphomas, and more than 90% of adults enjoy long-term survival when treated with a regimen similar to that proposed by MacGrath. As the MacGrath regimen is quite different and yields much improved results when compared with the CHOP regimen that is used for other aggressive B-cell lymphomas, it is important to make the diagnostic distinction between Burkitt's and large B-cell lymphoma.

Genetic testing plays a key role in making the diagnosis of Burkitt's lymphoma. The genetic hallmark of Burkitt's lymphoma is overexpression of the c-*MYC* oncogene due to a translocation which places c-*MYC* transcription under the control of elements at an immunoglobulin locus. The most common translocation, t(8;14), is a chromosomal rearrangement involving c-*MYC* and the immunoglobulin heavy chain locus. Other translocations involve c-*MYC* with the κ [t(2;8)]

Table 10.2 Chromosomal translocations in non-Hodgkin's lymphomas.

NHL histological type	Translocation	% of cases involved	Proto-oncogene	Function	Mechanism of activation of oncogene
Burkitt's lymphoma	t(8;14) t(2;8) (t8;22)	80% 15% 5%	c-*MYC*	Cell proliferation and growth	Transcriptional deregulation
Diffuse large cell lymphoma	der(3)	35%	BCL-6	Transcriptional repressor, required for GC formation	Transcriptional deregulation
Mantle cell lymphoma	t(11;14)	>70%	BCL-1	Cell cycle regulator	Transcriptional deregulation
Follicular lymphoma	t(14;18)	90%	BCL-2	Anti-apoptotic	Transcriptional deregulation
Lymphoplasmacytic lymphoma	t(9;14)	50%	PAX-5	Transcription factor regulation B-cell proliferation	Transcriptional deregulation
MALT lymphoma	t(11;18) t(1;14)	50% Rare	API-2-MLT BCL-10	API-2 is anti-apoptotic ? Anti-apoptotic	Fusion protein Transcriptional deregulation
Anaplastic large T-cell lymphoma	t(2;5)	60% in adults 85% in children	NPM–ALK	ALK is a tyrosine kinase	Fusion protein

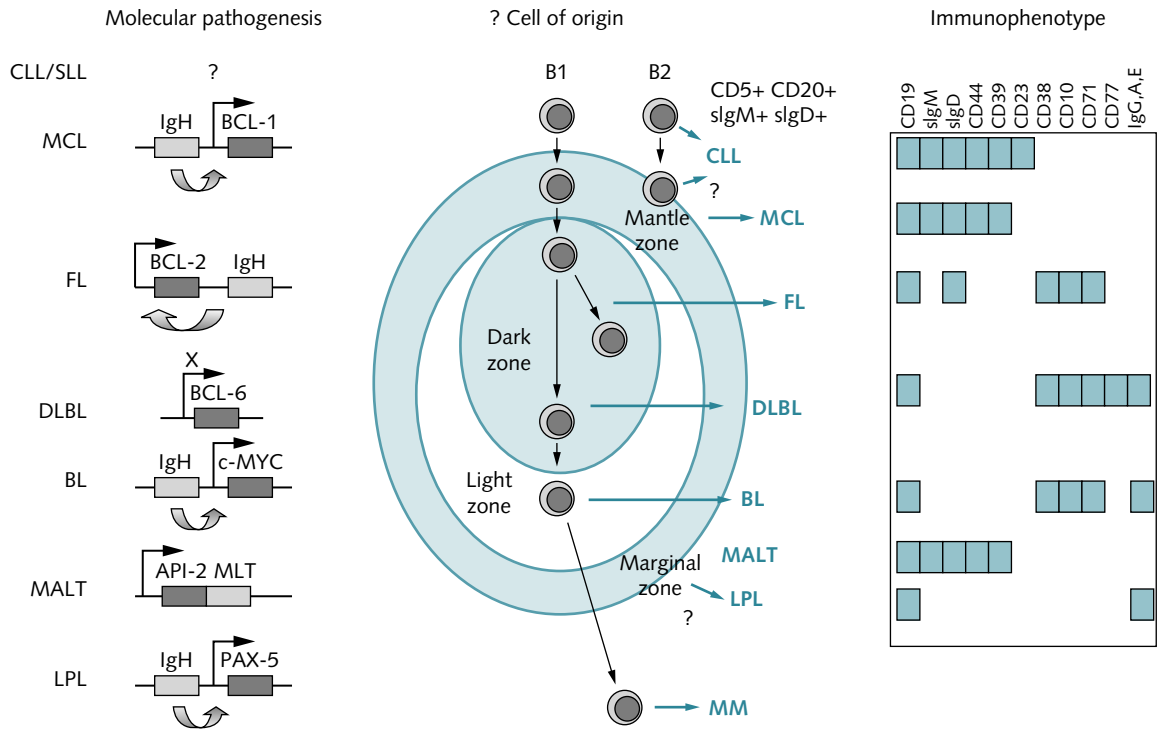


Fig. 10.3 Molecular and cytological pathogenesis of the most common types of lymphomas

or λ [t(8;22)] light chain loci. It is difficult to make the diagnosis of Burkitt's lymphoma in the absence of evidence for a c-*MYC* translocation by cytogenetics, FISH, or PCR. The c-*MYC* (myelocytomatosis) oncogene is a helix–loop–helix, zinc finger-containing transcription factor. The expression of the transcriptional targets of c-*MYC* is associated with a proliferative phenotype.

There is a type of lymphoma which lies histologically and clinically between the Burkitt's and the diffuse large B-cell lymphomas. These Burkitt's-like lymphomas lack c-*MYC* translocations. Thirty percent possess rearrangements involving the *BCL-2* gene. The prognosis of these tumors is generally inferior to that of the true Burkitt's lymphomas.

Evidence for latent Epstein–Barr virus infection is found in nearly all of the African endemic Burkitt's lymphoma but in only 20% of the sporadic form found outside Africa. It has been suggested that Epstein–Barr virus plays a causative role by opposing apoptosis.

Diffuse large B-cell lymphoma

Diffuse large B-cell lymphomas (DLBL) are a heterogeneous group of lymphomas of aggressive clinical behavior. The majority likely derive from follicular center cells, and roughly one-fifth of large B-cell lymphomas derive from transformation of a pre-existing follicular lymphoma. As the name suggests, DLBL has a diffuse histological pattern of large lymphoid cells. Approximately 40% of patients with this disease will be cured. The mainstay of therapy is combination chemotherapy including adriamycin. Some relapsing patients may be rescued by autologous stem cell transplantation following high-dose therapy.

Numerous heterogeneous genetic abnormalities have been reported for DLBL. These lymphomas are not characterized by a single, archetypical translocation, as with the t(14;18) in follicular lymphoma or the t(8;14) of Burkitt's lymphoma. Of the abnormalities that have been identified, those involving the *BCL-6* gene at 3q27 are the most common.

BCL-6 was initially described as the gene involved in translocations involving the 3q27 locus in a group of follicular and large B-cell lymphomas. Its expression is often deregulated via translocation with heterologous promoters, including immunoglobulin promoters. While only 10% of large B-cell lymphomas demonstrate the 3q27 translocation by cytogenetics, gene rearrangements involving 3q27 can be found by Southern hybridization analysis in 40% of large B-cell lymphomas. Additionally, somatic mutation of 5′ non-coding sequences has been shown. Overall, BCL-6 expression is found in >80% of DLBL. BCL-6 is required for germinal center formation. Expression of BCL-6 is now used in clinical pathology laboratories as a marker for germinal center origin. Containing six

zinc fingers, BCL-6 functions as a transcriptional repressor, at least in part by recruiting histone deacetylases. Likely gene targets of BCL-6 repression include chemokines, cell cycle proteins, and other transcriptional effectors. How repression of the heterogeneous BCL-6 targets leads to oncogenesis is unclear.

Approximately 20% of DLBL have the t(14;18) resulting in BCL-2 expression, which confers a worse prognosis. A significant proportion of these tumors likely arise via transformation of a follicular cell lymphoma. Overexpression of BCL-2 by amplification of the *BCL-2* allele has been observed by quantitative Southern hybridization and by comparative genomic hybridization in 11–31% of DLBL cases tested. Other genes which have demonstrated amplification by these techniques include *REL*, *MYC*, *CDK4* and *MDM2*.

Expression profiling

Pathological diagnosis is perhaps most important to the oncologist to the extent that it can inform about prognosis and treatment choice. Current diagnostic categorization of a lymphoma as DLBL relies on a fairly small number of data, including cell surface markers, nuclear and cytoplasmic appearance, and tissue morphology. When these data lead to the diagnosis of DLBL, the oncologist is left with a diagnostic grouping that includes those who will die of unresponsive disease in the first 6 months after diagnosis despite the most aggressive treatment approaches, and those who will rapidly obtain and maintain a durable complete remission after administration of anthracycline-based combination chemotherapy. It seems odd to call two diseases that behave so differently by the same name.

In an attempt to better divide the heterogeneous group of diseases encompassed by the label DLBL, Shipp and colleagues developed the IPI. The IPI uses just four pieces of clinical and laboratory data to further subclassify DLBL into four groups. While this formulation does provide a useful refinement of prognosis, it still falls short of the ideal predictor: a predictor that would definitively determine, prior to a particular therapy, whether that therapy will work.

While the ideal predictor may be unattainable in practice, attempts are being made to improve prognostic prediction using the many bits of molecular data provided by GEP. Two groups, one based at the National Cancer Institute and one based at the Dana-Farber Cancer Institute, have published results of applying GEP to lymphoma samples for which clinical data were available. In both cases, predictors generated by GEP were able to identify new subclasses of lymphomas and also to further refine prognosis even within IPI subgroups. Furthermore, when prognosis is predicted by a molecular signature, the molecules involved in that signature can be immediately identified as potential targets of anti-cancer therapy, a feat not possible when prognosis is determined by purely clinical criteria. The Dana-Farber group identified protein kinase C-β as such a target, and clinical trials incorporating a PKC-β inhibitor in DLBL are under way.

GEP potentially places tens of thousands of bits of data at the disposal of the pathologist and oncologist. As experience with this fascinating technology grows, its use in prognosis and therapeutic development will only improve.

Mantle cell lymphoma

Mantle cell lymphoma is a B-cell lymphoma thought to be the malignant counterpart of the memory B cells found in the mantle zone of lymphoid follicles. It has characteristic cell surface markings of $CD5^+$, $CD10^-$, $CD23^-$. Clinically, it is characterized by a moderate rate of growth. While it often responds to cytotoxic chemotherapy, it has frustrated attempts at cure with chemotherapy, though there are reports of long-term survivors following allogeneic bone marrow transplantation. The median survival is generally 3–5 years.

Mantle cell lymphoma is almost uniformly characterized, using classical cytogenetics or PCR, by a t(11;14) which puts the cyclin D1 (also known as *BCL-1*, or B-cell leukemia/lymphoma 1) gene under control of the Ig heavy chain transcription control elements. Cyclin D1 binds to and activates cyclin-dependent kinases. An important target of this activated cyclin-dependent kinase complex is the retinoblastoma (RB) gene product. In its hypophosphorylated state, RB inhibits entry into S-phase of the cell cycle by binding the transcription factor E2F. When RB is phosphorylated, E2F is freed to activate the transcription of genes which propel the cell into S-phase. Therefore, overexpression of cyclin D1 acts to overcome this late G1-phase checkpoint and maintain continuous proliferation.

Follicular lymphoma

Follicular lymphoma is an indolent lymphoma. The cell of origin is thought to be the follicular center B cell. While it can be cured by local therapy in the early stages, it is more usually diagnosed in an advanced stage, when cure is exceedingly rare. It is generally quite responsive to chemotherapy but almost always relapses. The clinical course is commonly marked by a series of chemotherapy-induced remissions followed by relapses, the interval between these decreasing over time. The end-stage of the disease may be characterized by insuperable resistance to chemotherapy or by transformation to an aggressive large B-cell phenotype. Despite the very low cure rate, many patients nonetheless survive more than a decade due to the indolent nature of the disease.

Histologically, it is characterized by a follicular pattern in the lymph node. The appearance can be similar to that of the non-malignant follicular hyperplasia. Light-chain restriction can be useful in suggesting the clonality of the tumor, which distinguishes it from benign hyperplasia.

Genetics

A t(14;18) translocation (Plate 10.2; Figure 10.4) is found in >85% of follicular lymphomas. This rearrangement puts the *BCL-2* gene under the transcriptional control of elements from the immunoglobulin heavy chain locus. The BCL-2 protein functions to oppose programmed cell death. It is presumed that BCL-2 expression in malignancies such as follicular lymphoma permits survival of the cancer cells under conditions (cell cycle checkpoint violation, metastatic location, genomic instability) that would otherwise trigger programmed cell death. The cloning of *BCL-2* led to the identification of a family of related proteins. While some are anti-apoptotic, like BCL-2, many are pro-apoptotic, but all function in the control of apoptosis.

Follicular lymphoma can transform into a higher-grade lymphoma with DLBL morphology. Numerous genetic changes have been associated with this transformation, including trisomy 7, loss of *p53*, and c-*MYC* rearrangements.

Lymphoplasmacytoid lymphoma

Lymphoplasmacytoid lymphoma is an indolent lymphoma. The cells of this lymphoma have a phenotype that lies midway between those of mature lymphocytes and plasma cells, for which reason they are often nicknamed 'plymphocytes'. This lymphoma commonly expresses IgM, which can lead to the syndrome of Waldenstrom's macroglobulinemia. Waldenstrom's macroglobulinemia is characterized by IgM expression, hyperviscosity, bleeding, Raynaud's phenomenon, visual disturbances, and other neurological symptoms.

Roughly half of all lymphoplasmacytoid lymphoma cases will demonstrate the t(9;14), which juxtaposes the *PAX-5* gene and the *IgH* locus. *PAX-5* encodes the BSAP (B-cell specific activator protein), which is a transcription factor. Its expression is associated with increased expression of genes important in early B-cell development and decreased expression of the *p53* tumor suppressor.

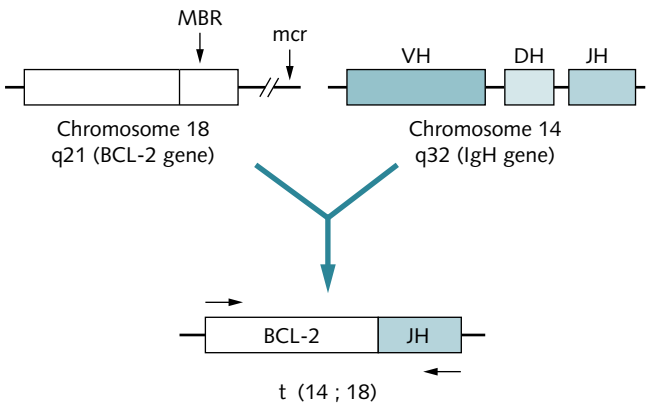


Fig. 10.4 Detection of t(14;18) by PCR amplification

MALT lymphoma

The mucosa-associated lymphoid tissue (MALT) lymphomas are thought to arise from the extranodal counterpart to post-follicular memory B cells found in the marginal zone of lymph node follicles. These tumors are often localized and their behavior is generally indolent. At least some depend on continued antigen stimulation for survival, as demonstrated by the prolonged complete responses that are seen when early-stage gastric MALT lymphoma is treated with an antibiotic regimen to eradicate chronic *Helicobacter pylori* infection.

t(11;18)(q21;q21) is found in more than half of all low-grade MALT lymphomas, with a preference for gastric lymphomas. The translocation is not typically found in high-grade MALT lymphomas. API-2-MALT 1 fusion protein is expressed from the mutant locus. API-2 (also known as IAP-2) belongs to a family of inhibitors of apoptosis that prevent death, likely due to their direct interaction with caspases, the proteases activated by programmed cell death. The physiological function of the MALT 1 protein is less well understood, though it possesses a caspase-like domain at its C-terminus. The function of the fusion protein is unclear, though there is some evidence that it activates NF-κB, perhaps leading to inhibition of apoptosis.

BCL-10 is overexpressed in a minority of MALT lymphoma cases via the t(1;14)(p22;q32), putting the coding region of *BCL-10* under the influence of the immunoglobulin heavy chain enhancer. The function of this protein is unclear, but some have suggested an interaction between BCL-10 and MALT-1, leading to NF-κB activation. Others have shown that API2-MALT 1 correlates with the nuclear location of BCL-10. These findings suggest that these two translocations may be involved in activating the same pathway.

Trisomy 3 is observed in 20–60% of all MALT lymphomas. The oncogenic properties of this numerical chromosomal abnormality are not understood.

Anaplastic large cell lymphoma

Anaplastic large cell lymphoma (ALCL) is characterized by

strong surface expression of the CD30 (Ki-1) antigen, a cytokine receptor in the tumor necrosis factor receptor family. The majority of ALCL demonstrate T-cell surface markers and/or clonal rearrangements of the T-cell receptor locus. There are two main clinical forms: systemic and cutaneous. The cutaneous form is particularly indolent. While it is clinically aggressive, systemic ALCL is generally sensitive to chemotherapy. Approximately 30% of those diagnosed die of the disease.

Approximately 50% of the systemic ALCL carry the t(2;5), which confers a good prognosis. Long-term survival of t(2;5)-positive patients is 80% while that of t(2;5)-negative patients is 25%. The t(2;5)(p23;q35) results in a chimeric gene encoding a fusion of the nucleophosmin (NPM) and anaplastic lymphoma kinase (ALK) proteins. NPM is a multifunctional protein which has been implicated in ribosome assembly, control of centrosome duplication, and nuclear transport as a shuttle protein; it also possesses chaperonin and ribonuclease activities. ALK is a member of the insulin family of receptor tyrosine kinases. Its natural ligand is unknown. The NPM–ALK fusion contains the oligomerization domain of NPM and the tyrosine kinase domain of ALK. It results in a self-oligomerizing, constitutively active tyrosine kinase with transforming properties. NPM–ALK can activate numerous downstream effectors, including phospholipase C-γ, phosphoinositol 3′-kinase and RAS.

Chronic lymphocytic leukemias/small cell lymphoma

Chronic lymphocytic leukemia is a low-grade lymphoma marked by a peripheral lymphocytosis of $CD5^+$, $CD20^+$, $CD23^+$ small lymphocytes that are similar in morphology to normal lymphocytes. BCL-2, which is expressed at low levels in normal lymphocytes, is expressed at high levels in more than 70% of CLL cases but this is rarely, if ever, due to a t(14;18). Staging based on presence of lymphadenopathy, organomegaly anemia or thrombocytopenia can provide prognostic information, those in the best prognostic groups enjoying normal mean survival times.

Prognosis can also be estimated by purely molecular criteria. In about half of CLL cases, lymphocytes are $CD38^+$, IgD^-, and contain V_H genes which exhibit somatic hypermutation. In the other half of CLL cases, the malignant lymphocytes resemble naive B cells, with surface marking $CD38^-$, IgD^+, and they lack V_H mutations. CLL with immunoglobulin V_H genes that exhibit more than 2% somatic hypermutation has significantly better survival than the latter.

Table 10.3 Abnormal genes in CLL.

Abnormality	%	Median survival (months)
13q deletion	50	133
11q deletion	18	79
12 trisomy	16	114
17p deletion	7	7
Normal karyotype	18	111

Other B-cell malignancies are characterized by chromosomal translocations, but there are no chromosomal translocations that characterize a significant subset of CLL. There are, however, several important genetic abnormalities in the absence of translocations. Whereas conventional Giemsa–trypsin banding analysis of chromosomes from CLL cells detected cytogenetic abnormalities in about half of CLL cases, the higher sensitivity of FISH has allowed the detection of genomic aberrations in 82% of cases. As FISH is a directed rather than a screening technique, these abnormalities had previously been demonstrated by conventional banding techniques. The abnormalities include del 13q (50%), del 11q (18%), +12q (16%), del 17p (7%) and del 6q (7%) (Table 10.3). Regression analysis allowed the assignment of 90% of these cases to one of five prognostic classes based on genetic abnormalities. The best prognostic group included those that had 13q deletion as their sole abnormality, with a median survival of 133 months. The worst prognostic group contained those with a 17p deletion, with a mean survival of 32 months. CLL with 17p and 11q deletions were more likely to have extensive lymphadenopathy, splenomegaly, cytopenias, and B symptoms. Examples of chromosome 13 and trisomy 12 detected by FISH and by CGH microarray are shown in Plates 10.3 and 10.4 respectively. These data raise the question of whether the classical clinical staging, which can be used to predict survival, is partly just a surrogate for particular genetic abnormalities, and it is the genetic abnormalities and resulting expression patterns that are more important in determining prognosis.

The specific genes affected by these abnormalities that are important for CLL oncogenesis are not known. The critical tumor suppressor lost in the 13q deletion is probably not *RB*, but rather a gene that lies telomeric and has so far defied definitive identification. Tumor suppressor *p53* is involved in 17p deletions. Overall, *p53* abnormalities have been found in at least 15% of patients, and are associated with an increased percentage of prolymphocytes and a poorer outcome.

Conclusions

Identification of the genes involved in lymphoma pathogenesis has allowed better characterization of the disease. A fuller understanding of the mechanisms causing specific subgroups of lymphomas should provide us with the means to develop specific therapies that will provide rational targets for improved therapies in these diseases.

Further reading

Burkitt's lymphoma

Battey J, Moulding C, Taub R *et al.* (1983) The human c-myc oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. *Cell*, **34**, 779–787.

Diffuse large B-cell lymphoma

Alizadeh AA, Eisen MB, Davis RE *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, **403**, 503–511.

Houldsworth J, Mathew S, Rao PH *et al.* (1996) REL proto-oncogene is frequently amplified in extranodal diffuse large cell lymphoma. *Blood*, **87**, 25–29.

Monni O, Joensuu H, Franssila K *et al.* (1997) BCL2 overexpression associated with chromosomal amplification in diffuse large B-cell lymphoma. *Blood*, **90**, 1168–1174.

Pasqualucci L, Migliazza A, Basso K *et al.* (2003) Mutations of the BCL-6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood*, **101**, 2914–2923.

Shipp MA, Ross KN, Tamayo P *et al.* (2002) Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nature Medicine*, **8**, 68–74.

Ye BH, Lista F, Lo Coco F *et al.* (1993) Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. *Science*, **262**, 747–750.

Follicular lymphoma

Bernell P, Jacobsson B, Liliemark J *et al.* (1998) Gain of chromosome 7 marks the progression from indolent to aggressive follicle centre lymphoma and is a common finding in patients with diffuse large B-cell lymphoma: a study by FISH. *British Journal of Haematology*, **101**, 487–491.

Cheng EH, Wei MC, Weiler S *et al.* (2001) BCL-2, BCL-X(L) sequester BH3 domain-only molecules preventing BAX- and BAK-mediated mitochondrial apoptosis. *Molecular Cell*, **8**, 705–711.

Lo Coco F, Gaidano G, Louie DC *et al.* (1993) p53 mutations are associated with histologic transformation of follicular lymphoma. *Blood*, **82**, 2289–2295.

McDonnell TJ, Korsmeyer SJ. (1991) Progression from lymphoid hyperplasia to high-grade malignant lymphoma in mice transgenic for the t(14;18). *Nature*, **349**, 254–256.

McDonnell TJ, Deane N, Platt FM *et al.* (1989) bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. *Cell*, **57**, 79–88.

Sander CA, Yano T, Clark HM *et al.* (1993) p53 mutation is associated with progression in follicular lymphomas. *Blood*, **82**, 1994–2004.

Tsjimoto Y, Yunis J, Onorato-Showe L *et al.* (1984) Molecular cloning of the chromosomal breakpoint of B-cell lymphomas and leukemias with the t(11;14) chromosome translocation. *Science*, **224**, 1403–1406.

Vaux DL, Cory S, Adams JM. (1988) Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature*, **335**, 440–442.

Lymphoplasmacytoid lymphoma

Iida S, Rao PH, Nallasivam P *et al.* (1996) The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. *Blood*, **88**, 4110–4117.

MALT lymphoma

Lucas PC, Yonezumi M, Inohara N *et al.* (2001) Bcl10 and MALT1, independent targets of chromosomal translocation in malt lymphoma, cooperate in a novel NF-kappa B signaling pathway. *Journal of Biological Chemistry*, **276**, 19012–19019.

Maes B, Demunter A, Peeters B, De Wolf-Peeters C. (2002) BCL10 mutation does not represent an important pathogenic mechanism in gastric MALT-type lymphoma, and the presence of the API2-MLT fusion is associated with aberrant nuclear BCL10 expression. *Blood*, **99**, 1398–1404.

Uren AG, O'Rourke K, Aravind LA *et al.* (2000) Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. *Molecular Cell*, **6**, 961–967.

Chronic lymphocytic leukemia

Damle RN, Wasil T, Fais F *et al.* (1999) Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*, **94**, 1840–1847.

Dohner H, Stilgenbauer S, Benner A *et al.* (2000) Genomic aberrations and survival in chronic lymphocytic leukemia. *New England Journal of Medicine*, **343**, 1910–1916.

Anaplastic large cell lymphoma

Kutok JL, Aster JC. (2002) Molecular biology of anaplastic lymphoma kinase-positive anaplastic large-cell lymphoma. *Journal of Clinical Oncology*, **20**, 3691–3702.

Morris SW, Kirstein MN, Valentine MB *et al.* (1994) Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science*, **263**, 1281–1284.

Chapter 11 The molecular biology of multiple myeloma

P. Leif Bergsagel

Introduction

On January 1 2000 there were estimated to be 47 000 patients with multiple myeloma (MM) in the USA, reflecting a yearly incidence of nearly 14 000 and a median survival of about 3 years. It remains an incurable malignancy that is often preceded by an exceptionally common (3.4% of the population over the age of 50) premalignant tumor—monoclonal gammopathy of undetermined significance (MGUS). Monoclonal gammopathies are usually asymptomatic, but they may sometimes cause primary amyloidosis as a result of pathological, and ultimately lethal, deposits of monoclonal immunoglobulin (Ig) in critical tissues. Although MGUS is stable, it progresses stochastically to frankly malignant MM at a rate of 0.6–3% per year depending on the level of monoclonal Ig. For both MGUS and MM, the incidence is markedly age-dependent, about twofold higher in American blacks than in Caucasians, and significantly higher in males. Although for many years the incidence of MM has appeared to be increasing, since 1992 the incidence appears to have become stable. The roles of genetic background and environment are poorly defined, although there may be clustering within families.

Stages of multiple myeloma

Multiple myeloma is usually—but perhaps not always—preceded by an age-dependent premalignant tumor called monoclonal gammopathy of undetermined significance (MGUS), which is present in 1% of adults over the age of 25 (rarely before the age of 40 but 10% of individuals in their tenth decade). MGUS cells secrete monoclonal immunoglobulin (Ig) and progress to malignant MM, expressing the same Ig, at the rate of 1% per year. Amyloidosis, which accounts for about 4000 deaths per year in the USA, usually has the same pathology as MGUS except that the monoclonal Ig forms pathological deposits in various tissues (generally the intact or fragmented Ig light chain). These tumors occur at multiple intramedullary (within the bone marrow) sites. Multiple myeloma is distinguished from MGUS by having a greater intramedullary tumor cell content (>10%), osteolytic bone lesions and/or an increasing tumor mass. Smoldering myeloma has a stable intramedullary tumor cell content of >10% but no osteolytic lesions or other complications of malignant myeloma. Progression of intramedullary myeloma is associated with increasingly severe secondary features (lytic bone lesions, anemia, immunodeficiency, renal impairment), and in a fraction of patients the occurrence of tumor in extramedullary locations. Extramedullary MM is a more aggressive tumor that is often called secondary or primary plasma cell leukemia, depending on whether preceding intramedullary myeloma has been recognized. Immortalized, Epstein–Barr virus-negative human multiple myeloma cell lines (HMCL) can sometimes be generated but, with very rare exceptions, only from extramedullary myeloma. The highly proliferative HMCL can be viewed as the ultimate stage of tumor progression, providing

a repository of genetic changes that have accumulated during *in vivo* tumor progression, but also some additional genetic changes that have occurred during *in vitro* culture. As is true for cell lines from all malignancies, the incidence and type of genetic abnormalities in HMCL may not precisely reflect the *in vivo* tumor cells. Multiple myeloma is a low-proliferative tumor. The plasma cell labeling index, typically detecting fewer than 1% of tumor cells actively synthesizing DNA until late in the disease, is a better prognostic indicator than the tumor cell content in the bone marrow.

Multiple myeloma is a plasma cell tumor with frequent Ig translocations

Germinal center B cells uniquely modify the DNA of Ig genes through sequential rounds of somatic hypermutation and antigen selection, and also by IgH switch recombination. Post-germinal center B cells can generate plasmablasts that have successfully completed somatic hypermutation and IgH switching before migrating to the bone marrow, where stromal cells enable terminal differentiation into long-lived plasma cells. MGUS and MM are characterized by the accumulation of transformed plasmablasts/plasma cells at multiple sites in the bone marrow. Importantly, although MM is more proliferative than MGUS, both tumors have an extremely low rate of proliferation. The combination of karyotypic complexity, inability to efficiently perform conventional cytogenetics on low proliferative tumors, and the telomeric location of some translocation partners delayed the identification of Ig translocations in MGUS and MM. An important initial step in solving this problem was the identification and cloning of IgH switch-region translocation breakpoints in HMCL. Interphase fluorescence *in situ* hybridization (FISH) using probes flanking the cloned breakpoints identifies karyotypic abnormalities even in non-dividing cells and has enabled the analysis of primary MGUS and MM tumors. Several studies have shown that most MM tumors have an IgH translocation that non-randomly involves one of many potential chromosomal partners. The prevalence of IgH translocations varies with the stage of disease: 46–48% in MGUS or smoldering MM, 55–73% in intramedullary MM, 85% in primary plasma cell leukemia, and >90% in HMCL.

Marked karyotypic instability

The karyotypes of MM are more similar to those of epithelial tumors and the blast phase of chronic myelogenous leukemia than are other hematopoietic tumors. However, the ratio of balanced translocations versus unbalanced translocations is substantially higher in MM than in epithelial tumors. Numerical chromosomal abnormalities are present in virtually all MM tumors and most, if not all, MGUS tumors. There is nonrandom involvement of different chromosomes in different myeloma tumors, and often heterogeneity among cells within a tumor. Comparative genomic hybridization (CGH) studies show that unbalanced chromosome structural changes are present in all plasma cell leukemias and most, if not all, MM tumors. Chromosomal gains that recur in more than 30% of MM tumors include 1q, 3q, 9q, 11q and 15q, the consequences of which remain to be determined, and the most frequent chromosome loss is 13q (*see below*). As CGH does not detect balanced translocations, a more comprehensive view is provided by spectral karyotype (SKY) analyses, although these are complicated by the fact that metaphase spreads can be obtained in only about 20% of cases. It is thought that karyotypic complexity increases during tumor progression, although karyotypic progression has not been well documented. Understanding how the karyotype correlates with disease severity is important because the detection of an abnormal karyotype correlates with an increase in the plasma cell labeling index and poor prognosis, and hypodiploidy is associated with a poorer prognosis than hyperdiploidy.

Recurrent chromosomal partners for Ig translocations: 11q13, 6p21, 4p16, 16q23 and 20q11

There are five well-defined recurrent chromosomal partners (oncogenes) that are involved in IgH translocations in MGUS and MM: 11q13 (cyclin D1), 6p21 (cyclin D3), 4p16 (*FGFR3* and *MMSET*), 16q23 (*c-maf*) and 20q11 (*mafB*). The combined prevalence of these five IgH translocation partners is about 40%, with approximately 15% 11q13, 3% 6p21, 15% 4p16, 5% 16q23 and 2% 20q11. The t(4;14) translocation is unusual in that it appears to dysregulate two potential oncogenes, *MMSET* on der(4), and *FGFR3* on der(14), although *FGFR3* on der(14) is lost or not expressed in about 20% of MM tumors that have a t(4;14) translocation. The apparently lower incidence of 4p16 and/or 16q23 in MGUS/smoldering MM compared with MM may be due to these translocations resulting in *de novo* MM without preceding MGUS, or a more rapid progression of MGUS to MM, an hypothesis supported by the fact that patients with translocations involving 4p16 or 16q23 have an extremely poor prognosis.

Secondary translocations dysregulate MYC

Primary translocations occur as early and perhaps initiating events during tumor pathogenesis, whereas secondary trans-

locations occur as progression events. Most translocations involving the five recurrent translocation partners described above appear to be primary translocations that occurred as a result of errors in IgH switch recombination during B-cell development in germinal centers. In contrast, translocations of *c-myc* appear to be very late secondary events that do not involve B-cell specific recombination mechanisms, are often complex, and sometimes do not involve Ig loci. By FISH analysis, rearrangements of *c-myc* are reported in only 15% of MM tumors (with frequent heterogeneity within a tumor), but in nearly 40% of advanced MM tumors and 90% of HMCL. Regarding the approximately 20% of IgH translocations not involving the six recurrent partners above, little is known about the multitude of partners and oncogenes, the mechanisms that mediate these translocations, and the time(s) at which these translocations occur.

Dysregulation of cyclin D1, 2 or 3: a unifying oncogenic event in multiple myeloma

MGUS and MM appear closer to normal, non-proliferating plasma cells than to normal but highly proliferating plasmablasts, for which 30% or more of the cells can be in S phase. It is surprising, therefore, that an analysis by Bergsagel and Kuehl of combined gene expression profiling data published from two laboratories shows that the expression level of cyclin D1, cyclin D2 or cyclin D3 mRNA in MM and MGUS is distinctly higher than in normal plasma cells—comparable to the levels of cyclin D2 mRNA expressed in normal proliferating plasmablasts (Figure 11.1). Normal hematopoietic cells, including normal B lymphocytes, plasma cells and plasmablasts, express cyclin D2 and/or D3, but little or no cyclin D1. Given the lack of cyclin D1 expression in normal lymphocytes, the occurrence of Ig translocations that dysregulate cyclin D1 or cyclin D3 in about 20% of MM tumors, the expression of cyclin D1 in nearly 40% of tumors lacking a t(11;14) translocation, and the increased expression levels of cyclin D2 in most remaining tumors, it seems apparent that almost all MM tumors dysregulate at least one of the cyclin D genes.

A model for the molecular pathogenesis of multiple myeloma

Based on the results summarized above, a model for the molecular pathogenesis of MM has been proposed. Chromosome

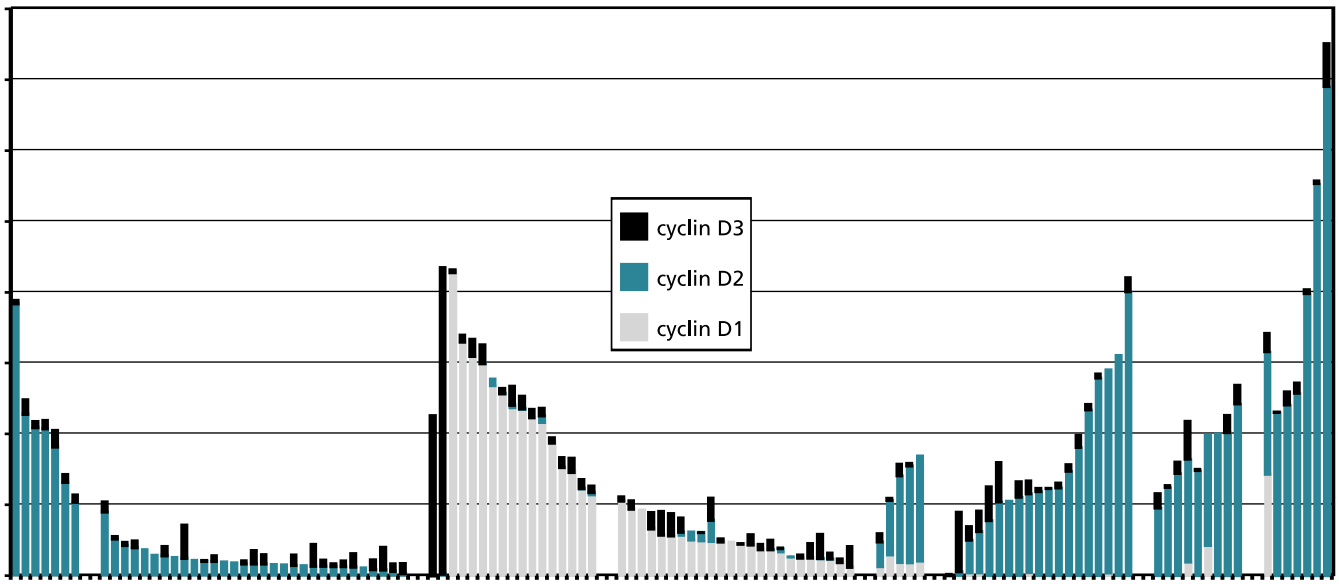


Fig. 11.1 Cyclin D expression in normal and malignant plasma cells

The raw scores for each of the three D-cyclins (D1, D2, D3) from the Affymetrix HuFL data set published by Tarte *et al*. and Zhan *et al*. are plotted one above the other. The samples are divided into nine groups, and arranged by the level of expression of the predominant cyclin D within each group. The samples are $CD138^{+}$-selected cells from six peripheral blood-generated plasmablasts and one reactive plasmacytosis (PPC), 31 bone marrow plasma cell (BMPC) from normal volunteers, and 78 samples from patients with newly diagnosed MM and three with plasma cell leukemia. Among these, there are two with high CCND3 (6p21) and 15 with high CCND1 (11q13); 25 with lower levels of CCND1 without t(11;14) (D1); four with lower levels of D1 and elevated CCND2 (D1+D2), 17 remaining patients with elevated CCND2 (D2), and two patients without an elevated cyclin D (D2); nine with elevated FGFR3 (4p16) (4p); seven with elevated CX3CR1 and β_7-integrin, a marker of Maf dysregulation (c-*maf*, 16q23; *mafb*, 20q11) (maf).

content appears to identify two different, but perhaps overlapping, pathways of pathogenesis: non-hyperdiploid tumors with a very high incidence of IgH translocations involving the five recurrent partners (above) and a relatively high incidence of chromosome 13/13q14 loss; and hyperdiploid tumors associated with multiple trisomies involving chromosomes 3, 5, 7, 9, 11, 15, 19 and 21, but a low incidence of both chromosome 13/13q14 loss and IgH translocations involving the five recurrent partners (Plate 11.1). In about half of the tumors, a primary chromosome translocation results in the dysregulated expression of an oncogene. This may lead directly to cyclin D dysregulation, either directly for cyclin D1 for translocations involving 11q13 or for cyclin D3 for translocations involving 6p21, or indirectly for cyclin D2 translocations involving 4p16, 16q23 and other translocations. Alternatively, the remaining tumors are mostly hyperdiploid, and cyclin D1 (or, less often, cyclin D2) is usually dysregulated by an undefined mechanism. The dysregulation of one of three cyclin D genes may render the cells more susceptible to proliferative stimuli, resulting in selective expansion as a result of interaction with bone marrow stromal cells that produce interleukin 6 (IL-6) and other cytokines. Karyotypic abnormalities, most notably IgH translocations, trisomies of chromosomes 3, 5, 7, 9, 11, 15, 19 and 21 and monosomy of chromosome 13 or 13q14 deletion, are often present in premalignant MGUS, the earliest identified stage of tumorigenesis. Even though dysregulation of a cyclin D gene appears to be a nearly universal event in early pathogenesis, there is evidence that the retinoblastoma (Rb) pathway is further disrupted by p16INK4a methylation and inactivation in a substantial fraction of MGUS and MM tumors. Tumor progression is associated with secondary chromosome translocations, of which *c-myc* provides a paradigm. Mutually exclusive activating mutations of K- or N-*Ras* [or *FGFR3* when there is a t(4;14) translocation] are rare or absent in MGUS, whereas *Ras* mutations are present in 30–40% of early MM and *FGFR3* mutations occur more frequently in advanced MM. Mutations and/or monoallelic deletion of *p53* occur frequently but only late in the course of the disease. Further disruption of the Rb pathway by inactivation of Rb or p18INK4c can also occur at a low frequency, most likely as a late progression event. The frequency and timing of other events, such as inactivation of *PTEN*, remain to be determined.

Proposal for a six-group translocation/cyclin D classification of multiple myeloma tumors

In addition to determining the expression level of cyclin D1, 2 and 3, gene expression profiling can effectively identify MM tumors that overexpress the oncogenes dysregulated by the five recurrent IgH translocations: 11q13 (cyclin D1); 6p21 (cyclin D3); 4p16 (*MMSET* and usually *FGFR3*); 16q23 (*c-maf*); and 20q11 (*mafB*). We propose six translocation/cyclin D (TC) groups (Table 11.1) that can be distinguished on the basis of five recurrent Ig translocations and cyclin D expression. (1) The 11+6 group (19%) expresses high levels of either cyclin D1 or cyclin D3 as a result of an Ig translocation. (2) The D1 group (31%) ectopically expresses low to moderate levels of cyclin D1 despite the absence of a t(11;14) translocation. (3) The D1+D2 group (8%) expresses low levels of cyclin D1 and moderate levels of cyclin D2. (4) The D2 group (19%) is a mixture of tumors that do not fall into one of the other groups; most of them express cyclin D2 but a few also express very low levels of each D cyclin. (5) The 4p group (16%) expresses high levels of cyclin D2 and also multiple myeloma SET domain (MMSET) [and in most cases

Table 11.1 Translocation and cyclin D (TC) molecular classification of multiple myeloma of 231 newly diagnosed and 30 relapsed multiple myelomas.

TC group	Primary translocation	Gene(s) at breakpoint	Cyclin	Multiple trisomies[1] (%)	Proliferation index >0.2[2] (%)	Frequency (%)
11+6	11q13	CCND1	D1	0	18	16
	6p21	CCND3	D3	29	14	3
D1	None	None	D1	93	5	31
D1+D2	None	None	D1+D2	86	43	8
D2	?	?	D2	40	16	19
4p	4p16	*FGFR3/MMSET*	D2	29	21	16
MAF	16q23	c-maf	D2	16	32	7
	20q11	mafB				

[1]Percentage of samples with gene expression profiling evidence for three or more trisomies of chromosomes 3, 5, 7, 9, 11, 15, 19 or 21; [2]percentage of samples with a gene expression profiling proliferation index greater than 0.2. The proliferation index was greater than 0.2 in 67% of relapsed samples and 10% of newly diagnosed multiple myelomas.

fibroblast growth factor receptor 3 (FGFR3)] as a result of a t(4;14) translocation. (6) The Maf group (7%) expresses the highest levels of cyclin D2 and also high levels of either c-Maf or MafB, consistent with evidence that both Maf transcription factors upregulate the expression of cyclin D2. With the exception of the 4p group, the focus of this classification is the increased expression of cyclin D1 (11q, D1, D1+D2 groups), cyclin D2 (Maf, D1+D2, D2 groups) or cyclin D3 (6p group). In fact, even though we do not understand the mechanism(s) responsible, the 4p group invariably shows increased expression of cyclin D2. Although this classification is not unequivocally established, we think that the basis for it is focussed on very early, if not initiating, oncogenic events, but the D1+D2 group might represent an exception.

The D1 and D1+D2 groups share several key features, including biallelic expression of cyclin D1, a very high incidence of hyperdiploidy with associated multiple trisomies of odd chromosomes, and a similar gene expression signature. The D1+D2 group is strikingly different from the D1 group, not only in the increased expression of cyclin D2 but also in the lack of increased expression of some genes that cluster with the D1 group and the increased expression of some genes that cluster with the D2 group. In addition, more than 40% of tumors in the D1+D2 group have a high proliferation index, whereas only 5% of tumors in the D1 group have a high proliferation index. Although more evidence is needed, we suspect that some tumors in the D1 group progress to the D1+D2 group.

The clear correlation of the various groups with ploidy provides significant support for this six-group classification. The supervised clustering also provides evidence that unique patterns of gene expression can be identified for each of the major groups, with the exception of the D2 group, which may well be a heterogeneous default group. Most importantly, the unsupervised cluster results demonstrate that the Maf, 4p and D1 groups (and to a lesser extent the 11+6 group) mostly associate as tight clusters. This implies to us that the ultimate phenotype of the tumor is determined mainly by early events, and is consistent with the possibility that MM represents several disease entities with distinct pathways of pathogenesis.

The translocation/cyclin D molecular classification predicts prognosis and response to therapies

In addition to tumor mass and secondary features, which represent a host response to MM (anemia, thrombocytopenia, bone disease, immunodeficiency, etc.), the intrinsic properties of the tumor cell are also informative in predicting the prognosis and the response to existing therapies. For example, it has been well documented that an unfavorable outcome is associated with each of the following: increased plasma cell labeling index; the generation of tumor cells with an abnormal karyotype (perhaps a surrogate for increased proliferation); hypodiploidy compared with hyperdiploidy; monosomy of chromosome 13/13q; monosomy of chromosome 17/deletion of *p53*; and lack of cyclin D1 expression. It also has been reported and independently confirmed that activating mutations of K-*Ras* (but not N-*Ras*) represent an adverse prognostic factor. More recently, it has become clear that specific IgH translocations also have a profound prognostic significance (Table 11.1). In particular, patients with tumors that have a t(4;14) translocation have substantially shortened survival, either with standard or high-dose therapy (median overall survival 26 and 33 months respectively), and patients with a t(14;16) have a similarly poor if not worse prognosis (median overall survival 16 months with conventional therapy). By contrast, patients with tumors that have a t(11;14) translocation appear to have a marginally better survival after conventional chemotherapy (median overall survival 50 months) but apparently a remarkably better survival after intense therapy (predicted overall survival at 80 months, 88%). These results suggest that the TC classification, which appears to be based on the earliest events in pathogenesis, may be a clinically useful way to classify patients into groups that have distinct subtypes of MM (and MGUS) tumors. The TC classification identifies clinically important molecular subtypes of MM with different prognoses and with unique responses to different treatments (e.g. high-dose therapy and 11+6, microenvironment-directed therapy and D1, FGFR3 inhibitor and 4p, *maf* dominant-negative and Maf).

A tumor suppressor gene on chromosome 13?

Monoallelic loss of 13q sequences is one of the most frequent abnormalities in MM (~50% of untreated cases by interphase FISH analyses) and is an independent predictor of a poor prognosis. Most often there is 13 monosomy, with selective loss of 13q sequences by interstitial deletion or translocation occurring much less frequently. The minimum region of deletion appears to be at 13q14, but biallelic deletion is rare. Notably, trisomy of chromosome 13 is also rare. The frequency of chromosome 13q loss increases with disease stage, from 20% in MGUS to nearly 70% in plasma cell leukemia or HMCL. In most patients, only a subset of MGUS tumor cells have the 13q abnormality, whereas for MM patients with an abnormality of 13q virtually all tumor cells have this abnormality. These results indicate that chromosome 13 losses begin in MGUS and increase in myeloma, but the nearly uniform presence of this abnormality in MGUS and myeloma tumors with t(4;14) or t(14;16) raises the possibility that this abnormality might

occur as a very early event in tumors with these translocations. Studies on other kinds of tumors show that both copies of the retinoblastoma (*RB*) gene at 13q14 must be inactivated to eliminate its tumor suppressor function. However, biallelic deletion, inactivating mutations and lack of *RB* expression appear to occur only rarely, even in advanced myeloma tumors and cell lines. Notably, loss of 13q14 sequences is frequent in chronic lymphocytic leukemia, but is not associated with a poor prognosis. So, there might be a tumor suppressor gene on 13q that is unique to MM. In addition, there is evidence that a number of well-known tumor suppressor genes are sometimes involved in myeloma.

Activating *Ras* and *FGFR3* mutations

Activating mutations of N- or K-*Ras* oncogenes distinguish MM from MGUS. In one large study, activating mutations of *Ras* at codon 12, 13 or 61 were identified in approximately 40% of MM tumors at the time of diagnosis, and a limited analysis indicated mutations in 49% of tumors at the time of relapse; 60% of the mutations were in N-*Ras* and 40% in K-*Ras*. A second large study reports a slightly higher incidence of *Ras* mutations, with K-*Ras* affected more often than N-*Ras*. The frequency of activating *Ras* mutations is relatively independent of the plasma cell labeling index and stage of myeloma. The same mutations are found in 17/38 (45%) of HMCL. Strikingly, fewer than 5% of MGUS tumors have *Ras* mutations. Although H-RAS is expressed at high levels in some MM tumors, no activating mutations of *H-Ras* have been reported (possibly because most studies have focussed on K- and N-*ras*). Tumors that overexpress FGFR3 as a result of a t(4;14) translocation can have activating mutations of *Ras* or *FGFR3* but not both, consistent with constitutive activation of the mitogen-activated protein kinase (MAPK) pathway in each case. Transfection studies of an IL-6-dependent HMCL show that activating mutations of either N- or K-*Ras*, or *FGFR3*, enhance growth and decrease the amount of IL-6 that is required for survival and growth. K-*Ras* mutation is associated with shortened survival, whereas patients whose myeloma tumors have N-*Ras* mutations have a similar prognosis to those who do not have *Ras* mutations.

Identification of novel therapeutic strategies targeting genetic abnormalities

The critical role of cyclin D dysregulation in the pathogenesis of MM highlights the importance of the cyclin D/RB pathway, and suggests that there may be a therapeutic window in targeting this pathway for all molecular subtypes of MM (Figure 11.2). For example, epigenetic silencing of cyclin-dependent protein kinase (CDK) inhibitor mRNA expression might be reversed by histone deacetylase inhibitors (suberoylanilide hydroxamic acid, depsipeptide) or inhibitors of DNA methyl transferase (5 aza-2′deoxy-cytidine). To target cyclin D *per se*, there are a number of possible strategies, including modulation of mRNA translation (e.g. desferroxamine, eicosapentaenoic acid), post-translational modifications (ubiquitination and proteasomal degradation), enzyme function (selective CDK inhibitors), and perhaps even inhibition of expression

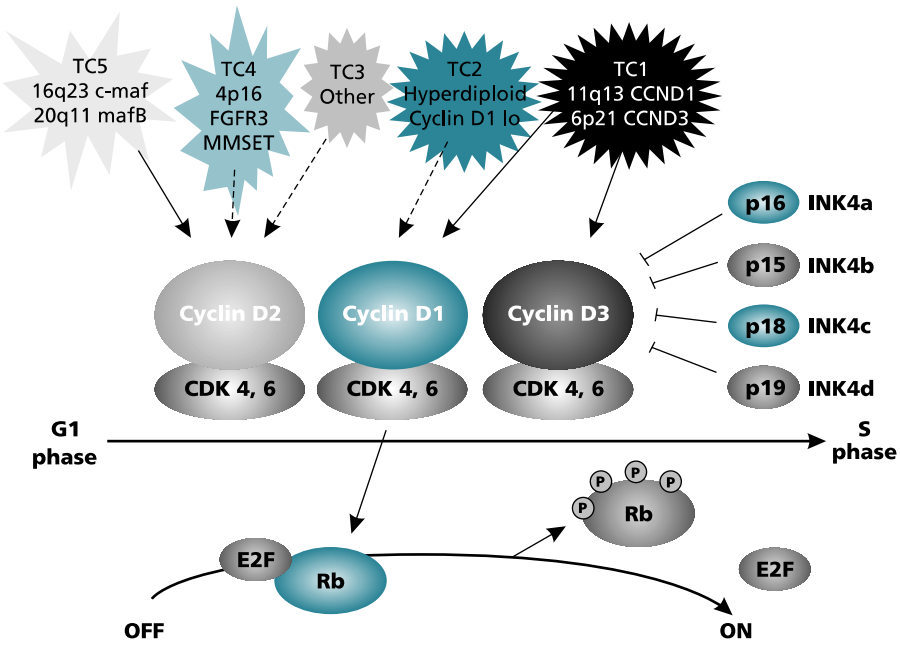


Fig. 11.2 Cyclin dysregulations as protential therapeutic targets in MM
With greater understanding of the specific dysregulations induced in the different groups in myeloma, it should be possible to develop strategies to correct the defects in these pathways.

of cyclin D mRNA (the TC2 group may be particularly dependent on interaction with bone marrow stromal cells for the ectopic expression of cyclin D1). Additional specificity may be achieved by targeting the genes that are directly dysregulated by translocations. This seems to be especially true in the case of the t(4;14), where two enzymes are overexpressed: FGFR3, a tyrosine kinase receptor, and MMSET, which has homology to histone methyltransferases. As a surface receptor, FGFR3 may be targeted by monoclonal antibodies, and as a tyrosine kinase by selective tyrosine kinase inhibitors. Preclinical studies have validated FGFR3 as a therapeutic target in t(4;14) MM, inhibitors of histone methyltransferases are being developed, and studies are under way to validate MMSET as a target in t(4;14) MM.

Critical but variable role for the bone marrow microenvironment

Like their normal bone marrow plasma cell counterpart, MGUS and MM tumors are dependent on mutual interactions with cells and extracellular components of the bone marrow for survival and growth. Exceptions to this include primary plasma cell leukemia (PCL) and terminal phases of MM, which sometimes extends to extramedullary sites. Significantly, virtually all HMCL are derived from PCL or extramedullary tumor. Although not yet well understood, there is increasing evidence that some of the earliest oncogenic events differentially affect the interaction of tumor cells with bone marrow components. First, tumors in the 11+6 and D1 groups are more strongly associated with lytic bone lesions than tumors in the 4p and Maf groups (P. L. Bergsagel, unpublished results). Secondly, the Maf transcription factor-stimulated expression of β7 integrin and other surface receptors or cytokines seems likely to influence the interactions of the Maf tumor group in the bone marrow. Thirdly, in contrast to tumors in the other TC groups, D1 tumors [hyperdiploid with multiple trisomies and cyclin D1 expression without a t(11;14)] are greatly under-represented or absent in primary PCL and HMCL. Thus, D1 tumors may be uniquely dependent on the bone marrow environment, with the possibility that the ectopic/increased expression of cyclin D1 is dependent on the bone marrow microenvironment. For example, IL-6 secreted by bone marrow stromal cells (BMSCs) triggers phosphorylation of Akt, and downstream glycogen synthasekinase 3 (GSK-3α) in turn induces phosphorylation of cyclin D1 followed by degradation through the ubiquitin-proteasome pathway, thereby promoting the transition from G1 to S phase. Tumor necrosis factor α (TNF-α) in the bone marrow milieu activates NFκB, thereby modulating expression of adhesion molecules on both MM cells and BMSCs, and inducing IL-6 transcription and secretion in BMSCs. Activated NFκB also binds to the promoter of cyclin D1, thereby regulating its expression.

Therapies that target the multiple myeloma cell and the microenvironment

Conventional therapy of MM with oral melphalan and prednisone was developed in the 1960s, results in median overall survival of about 36 months, and is still the standard for patients who are not treated with high-dose (200 mg) melphalan and stem cell support. Two randomized controlled studies in patients under the age of 65 have shown that this latter therapy extends median overall survival by about 12 months. Recently two new agents, thalidomide (1999) and bortezomib (2003) have shown remarkable activity in MM and are rapidly changing the management and outcome for patients. Both of these agents have unique and novel mechanisms of action, appear to be more effective in MM than in other malignancies, and are postulated to target both the MM cell, and its supporting microenvironment. Thalidomide was originally marketed as a sedative, but was withdrawn in the 1960s because of teratogenicity. It has many activities, including inhibition of angiogenesis, inhibition of TNF-α signaling, and T-cell stimulatory activity. Several derivative compounds have been developed to overcome the devastating side effects, and two of these are showing promise in early clinical trials: CC5013 (Revimid) and CC4047 (Actimid). Bortezomib is the first of a new class of drug that targets the proteasome. Cancer cells in general, and MM cells in particular, appear to be particularly dependent on proteasome function, and reliably undergo apoptosis when its function is inhibited by 60% or more. Although the exact nature of the critical proteins that accumulate is still unclear, some of the implicated candidates include cell cycle regulated proteins, p53-mdm2, NfκB-IkB and caspases. An important challenge for the future will be determining the critical pathways by which these novel therapies exert their actions in MM and developing a rational approach to their clinical development. The long-term promise is that a detailed knowledge of the molecular pathogenesis, together with the availability of specific targeted therapies, will allow the individualized tailoring of a non-toxic cocktail that will effectively control the disease and its devastating complications.

Conclusions

There appear to be two pathways involved in the early pathogenesis of premalignant MGUS and malignant MM tumors. Nearly half of these tumors are non-hyperdiploid and most have IgH translocations that involve five recurrent

chromosomal loci, including 11q13 (cyclin D1), 6p21 (cyclin D3), 4p16 (*FGFR3* and *MMSET*), 16q23 (*c-maf*) and 20q11 (*mafB*). The remaining tumors are hyperdiploid and contain multiple trisomies involving chromosomes 3, 5, 7, 9, 11, 15, 19 and 21, but infrequently have IgH translocations involving the five recurrent loci. Dysregulated expression of cyclin D1, D2 or D3 appears to occur as an early event in virtually all of these tumors, providing a unifying pathogenic event. This may render the cells more susceptible to proliferative stimuli, resulting in selective expansion as a result of interaction with bone marrow stromal cells that produce IL-6 and other cytokines. Delineation of the mechanisms mediating MM cell proliferation, survival and migration in the bone marrow microenvironment may both enhance understanding of pathogenesis and provide the framework for identification and validation of novel molecular targets.

Further reading

Cohen HJ, Crawford J, Rao MK *et al.* (1998) Racial differences in the prevalence of monoclonal gammopathy in a community-based sample of the elderly. *American Journal of Medicine*, **104**, 439–444.

Kuehl WM, Bergsagel PL. (2002) Multiple myeloma: evolving genetic events and host interactions. *Nature Reviews. Cancer*, **2**, 175–187.

Kyle RA, Therneau TM, Rajkumar SV *et al.* (2002) A long-term study of prognosis in monoclonal gammopathy of undetermined significance. *New England Journal of Medicine*, **346**, 564–569.

Lynch HT, Sanger WG, Pirruccello S *et al.* (2001) Familial multiple myeloma: a family study and review of the literature. *Journal of the National Cancer Institute*, **93**, 1479–1483.

Cytogenetics

Avet-Loiseau H, Brigaudeau C, Morineau N *et al.* (1999) High incidence of cryptic translocations involving the Ig heavy chain gene in multiple myeloma as shown by fluorescence in situ hybridization. *Genes, Chromosomes and Cancer*, **24**, 9–15.

Avet-Loiseau H, Facon T, Grosbois B *et al.* (2002) Oncogenesis of multiple myeloma: 14q32 and 13q chromosomal abnormalities are not randomly distributed, but correlate with natural history, immunological features, and clinical presentation. *Blood*, **99**, 2185–2191.

Bergsagel PL, Kuehl WM. (2001) Chromosomal translocations in multiple myeloma. *Oncogene*, **20**, 5611–5622.

Fonseca R, Bailey RJ, Ahmann GJ *et al.* (2002) Genomic abnormalities in monoclonal gammopathy of undetermined significance. *Blood*, **100**, 1417–1424.

Fonseca R, Blood E, Rue M *et al.* (2003) Clinical and biologic implications of recurrent genomic aberrations in myeloma. *Blood*, **101**, 4569–4575.

Rajkumar SV, Fonseca R, Dewald GW *et al.* (1999) Cytogenetic abnormalities correlate with the plasma cell labeling index and extent of bone marrow involvement in myeloma. *Cancer Genetics and Cytogenetics*, **113**, 73–77.

IgH translocations

Bergsagel PL, Chesi MC, Nardini E *et al.* (1996) Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 13931–13936.

Chesi M, Bergsagel PL, Brents LA *et al.* (1996) Dysregulation of cyclin D1 by translocation into an IgH gamma switch region in two multiple myeloma cell lines. *Blood*, **88**, 674–681.

Chesi M, Nardini E, Brents LA *et al.* (1997) Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. *Nature Genetics*, **16**, 260–264.

Chesi M, Nardini E, Lim RSC *et al.* (1998) The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood*, **92**, 3025–3034.

Chesi M, Bergsagel PL, Shonukan OO *et al.* (1998) Frequent dysregulation of the c-maf proto-oncogene at 16q23 by translocation to an Ig locus in multiple myeloma. *Blood*, **91**, 4457–4463.

Garand R, Avet-Loiseau H, Accard F *et al.* (2003) t(11;14) and t(4;14) translocations correlated with mature lymphoplasmacytoid and immature morphology, respectively, in multiple myeloma. *Leukemia*, **17**, 2032–2035.

Shaughnessy J, Gabrea A, Qi Y *et al.* (2001) Cyclin D3 at 6p21 is dysregulated by recurrent Ig translocations in multiple myeloma. *Blood*, **98**, 217–223.

Molecular genetic classification

Barille-Nion S, Barlogie B, Bataille R *et al.* (2003) Advances in biology and therapy of multiple myeloma. *Hematology (American Society of Hematology Education Program)*, 248–278.

Bergsagel PL, Kuehl WM. (2003) Critical roles for immunoglobulin translocations and cyclin D dysregulation in multiple myeloma. *Immunological Reviews*, **194**, 96–104.

Fonseca R, Debes-Marun CS, Picken EB *et al.* (2003) The recurrent IgH translocations are highly associated with non-hyperdiploid variant multiple myeloma. *Blood*, **102**, 2562–2567.

Fonseca R, Barlogie B, Bataille R *et al.* (2004) Genetics and cytogenetics of multiple myeloma: a workshop report. *Cancer Research*, **64**, 1546–58.

Smadja NV, Bastard C, Brigaudeau C *et al.* (2001) Hypodiploidy is a major prognostic factor in multiple myeloma. *Blood*, **98**, 2229–2238.

Tarte K, De Vos J, Thykjaer T *et al.* (2002) Generation of polyclonal plasmablasts from peripheral blood B cells: a normal counterpart of malignant plasmablasts. *Blood*, **100**, 1113–1122.

Tian E, Zhan F, Walker R *et al.* (2003) The role of the Wnt-signaling antagonist DKK1 in the development of osteolytic lesions in multiple myeloma. *New England Journal of Medicine*, **349**, 2483–2494.

Zhan F, Hardin J, Kordsmeier B *et al.* (2002) Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood*, **99**, 1745–1757.

Genetic prognostic factors

Avet-Loiseau H, Li JY, Morineau N *et al.* (1999) Monosomy 13 is associated with the transition of monoclonal gammopathy of undetermined significance to multiple myeloma. Intergroupe Francophone du Myelome. *Blood*, **94**, 2583–2589.

Facon T, Avet-Loiseau H, Guillerm G *et al.* (2001) Chromosome 13 abnormalities identified by FISH analysis and serum beta2-microglobulin produce a powerful myeloma staging system for patients receiving high-dose therapy. *Blood*, **97**, 1566–1571.

Fonseca R, Harrington D, Oken MM *et al.* (2002) Biological and prognostic significance of interphase fluorescence in situ hybridization detection of chromosome 13 abnormalities (delta13) in multiple myeloma: an Eastern Cooperative Oncology Group study. *Cancer Research*, **62**, 715–720.

Moreau P, Facon T, Leleu X *et al.* (2002) Recurrent 14q32 translocations determine the prognosis of multiple myeloma, especially in patients receiving intensive chemotherapy. *Blood*, **100**, 1579–1583.

Shaughnessy J, Tian E, Sawyer J *et al.* (2000) High incidence of chromosome 13 deletion in multiple myeloma detected by multiprobe interphase FISH. *Blood*, **96**, 1505–1511.

Soverini S, Cavo M, Cellini C *et al.* (2003) Cyclin D1 overexpression is a favorable prognostic variable for newly diagnosed multiple myeloma patients treated with high-dose chemotherapy and single or double autologous transplantation. *Blood*, **102**, 1588–1594.

Secondary genetic events

Avet-Loiseau H, Gerson F, Margrangeas F *et al.* (2001) Rearrangements of the c-myc oncogene are present in 15% of primary human multiple myeloma tumors. *Blood*, **98**, 3082–3086.

Billadeau D, Liu P, Jelinek D *et al.* (1997) Activating mutations in the N- and K-ras oncogenes differentially affect the growth properties of the IL-6-dependent myeloma cell line ANBL6. *Cancer Research*, **57**, 2268–2275.

Chesi M, Brents LA, Ely SA *et al.* (2001) Activated fibroblast growth factor receptor 3 is an oncogene that contributes to tumor progression in multiple myeloma. *Blood*, **97**, 729–736.

Corradini P, Ladetto M, Voena C *et al.* (1993) Mutational activation of N- and K-ras oncogenes in plasma cell dyscrasias. *Blood*, **81**, 2708–2713.

Drach J, Ackermann J, Fritz E *et al.* (1998) Presence of a p53 gene deletion in patients with multiple myeloma predicts for short survival after conventional-dose chemotherapy. *Blood*, **92**, 802–809.

Fonseca R, Price-Troska T, Blood E *et al.* (2003) Implication of N-ras and K-ras mutation in clinical outcome and biology of multiple myeloma. *Blood*, **102**, 113a.

Ge NL, Rudikoff S. (2000) Expression of PTEN in PTEN-deficient multiple myeloma cells abolishes tumor growth in vivo. *Oncogene*, **19**, 4091–4095.

Guillerm G, Gyan E, Wolowiec D *et al.* (2001) p16(INK4a) and p15(INK4b) gene methylations in plasma cells from monoclonal gammopathy of undetermined significance. *Blood*, **98**, 244–246.

Hyun T, Yam A, Pece S *et al.* (2000) Loss of PTEN expression leading to high Akt activation in human multiple myelomas. *Blood*, **96**, 3560–3568.

Kulkarni MS, Daggett JL, Bender TP *et al.* (2002) Frequent inactivation of the cyclin-dependent kinase inhibitor p18 by homozygous deletion in multiple myeloma cell lines: ectopic p18 expression inhibits growth and induces apoptosis. *Leukemia*, **16**, 127–134.

Liu P, Leong T, Quam L *et al.* (1996) Activating mutations of N- and K-ras in multiple myeloma show different clinical associations: analysis of the Eastern Cooperative Oncology Group Phase III Trial. *Blood*, **88**, 2699–2706.

Neri A, Baldini L, Trecca D *et al.* (1993) p53 gene mutations in multiple myeloma are associated with advanced forms of malignancy. *Blood*, **81**, 128–135.

Shou Y, Martelli ML, Gabrea A *et al.* (2000) Diverse karyotypic abnormalities of the c-myc locus associated with c-myc dysregulation and tumor progression in multiple myeloma. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 228–233.

Uchida T, Kinoshita T, Ohno T *et al.* (2001) Hypermethylation of p16INK4A gene promoter during the progression of plasma cell dyscrasia. *Leukemia*, **15**, 157–165.

Urashima M, Teoh G, Ogata A *et al.* (1997) Characterization of p16(INK4A) expression in multiple myeloma and plasma cell leukemia. *Clinical Cancer Research*, **3**, 2173–2179.

Targeting genetic abnormalities

Caraglia M, Tagliaferri P, Budillon A *et al.* (1999) Post-translational modifications of eukaryotic initiation factor-5A (eIF-5A) as a new target for anti-cancer therapy. *Advances in Experimental Medicine and Biology*, **472**, 187–198.

Dragnev KH, Freemantle SJ, Spinella MJ *et al.* (2001) Cyclin proteolysis as a retinoid cancer prevention mechanism. *Annals of the New York Academy of Sciences*, **952**, 13–22.

Hurt EM, Wiestner A, Rosenwald A *et al.* (2004) Overexpression of c-maf is a frequent oncogenic event in multiple myeloma that promotes proliferation and pathological interactions with bone marrow stroma. *Cancer Cell*, **5**, 191–199.

Malumbres M, Barbacid M. (2001) To cycle or not to cycle: a critical decision in cancer. *Nature Reviews. Cancer*, **1**, 222–231.

Palakurthi SS, Fluckiger R, Aktas H *et al.* (2000) Inhibition of translation initiation mediates the anticancer effect of the n-3 polyunsaturated fatty acid eicosapentaenoic acid. *Cancer Research*, **60**, 2919–2925.

Trudel S, Ely SA, Farooqi Y *et al.* (2004) Inhibition of fibroblast growth factor receptor 3 induces differentiation and apoptosis in t(4;14) myeloma. *Blood*, **103**, 3521–3528.

Tsutsumi S, Yanagawa T, Shimura T *et al.* (2003) Regulation of cell proliferation by autocrine motility factor/phosphoglucose isomerase signaling. *Journal of Biological Chemistry*, **278**, 32165–32172.

Zhu WG, Otterson GA. (2003) The interaction of histone deacetylase inhibitors and DNA methyltransferase inhibitors in the treatment of human cancer cells. *Current Medicinal Chemistry. Anti-cancer Agents*, **3**, 187–199.

Current therapy

Adams J. (2004) The proteasome: a suitable antineoplastic target. *Nature Reviews. Cancer*, **4**, 349–360.

Attal M, Harousseau JL, Stoppa AM *et al.* (1996) A prospective, randomized trial of autologous bone marrow transplantation and chemotherapy in multiple myeloma. Intergroupe Francais du Myelome. *New England Journal of Medicine*, **335**, 91–97.

Attal M, Harousseau JL, Facon T *et al.* (2003) Single versus double autologous stem-cell transplantation for multiple myeloma. *New England Journal of Medicine*, **349**, 2495–2502.

Bartlett JB, Dredge K, Dalgleish AG. (2004) The evolution of thalidomide and its IMiD derivatives as anticancer agents. *Nature Reviews. Cancer*, **4**, 314–322.

Bergsagel DE, Griffith KM, Haut A *et al.* (1967) The treatment of plasma cell myeloma. *Advances in Cancer Research*, **10**, 311–359.

Child JA, Morgan GJ, Davies FE, *et al.* (2003) High-dose chemotherapy with hematopoietic stem-cell rescue for multiple myeloma. *New England Journal of Medicine*, **348**, 1875–1883.

Gregory WM, Richards MA, Malpas JS. (1992) Combination chemotherapy versus melphalan and prednisolone in the treatment of multiple myeloma: an overview of published trials. *Journal of Clinical Oncology*, **10**, 334–342.

Hideshima T, Anderson KC. (2002) Molecular mechanisms of novel therapeutic approaches for multiple myeloma. *Nature Reviews. Cancer*, **2**, 927–937.

Richardson PG, Schlossman RL, Weller E *et al.* (2002) Immunomodulatory drug CC-5013 overcomes drug resistance and is well tolerated in patients with relapsed multiple myeloma. *Blood*, **100**, 3063–3067.

Richardson PG, Barlogie B, Berenson J *et al.* (2003) A phase 2 study of bortezomib in relapsed, refractory myeloma. *New England Journal of Medicine*, **348**, 2609–2617.

Singhal S, Mehta J, Desikan R *et al.* (1999) Antitumor activity of thalidomide in refractory multiple myeloma. *New England Journal of Medicine*, **341**, 1565–1571.

Chapter 12 The molecular basis of anemia

Lucio Luzzatto & Anastasios Karadimitris

Introduction

The title of this chapter is quite ambitious, and in this respect we beg the reader's indulgence. As hematologists, we must have the ambition to explain anemia at the molecular level; this has been done successfully in some cases but not yet in others. One major reason is that anemia is not a disease, but a vast collection of diseases, extremely heterogeneous in terms of etiology, pathophysiology and clinico-hematological manifestations, as well as in our ability to treat them effectively. Since this book focusses on molecular pathophysiology, it is particularly pertinent to blood diseases that have a genetic basis. There are three main groups of anemias that qualify in this respect: (1) the hemoglobinopathies, covered in Chapters 1 and 14; (2) red cell membrane cytoskeleton disorders, which are not covered in this book; and (3) inherited hemolytic anemia due to enzyme abnormalities, covered in this chapter. In addition, considering the space allocated to this chapter, we have included several other types of inherited and acquired anemias: our main inclusion criterion was that, based on current knowledge, we could offer at least some meaningful discussion of their molecular basis. In this area there have been at least three breakthrough additions since the first edition of this book: (1) autosomal dominant dyskeratosis congenita has been found to be due to mutations of telomerase; (2) the gene mutated in congenital dyserythropoietic anemia type I has been identified; and (3) the gene encoding red cell 5′ nucleotidase has been identified, enabling at long last the molecular diagnosis of one of the commonest red cell enzymopathies.

Megaloblastic anemia

Megaloblastic anemia is defined by a highly characteristic set of morphological changes which affect cells of the erythroid, myeloid and megakaryocytic lineages in the peripheral blood and bone marrow. These changes include macrocytosis, Howell–Jolly bodies, hypersegmented neutrophils, giant metamyelocytes and giant platelets. Despite the multitude of these signs, the one pathognomonic feature which we regard as a *sine qua non* for the diagnosis of megaloblastic anemia is the peculiarly finely stippled chromatin of erythroid cells, combined with derangement of the normally precisely ordered parallel pattern of maturation of the nucleus and of the cytoplasm. This asynchrony, whereby the maturation of the nucleus lags behind that of the cytoplasm, is the *morphological hallmark* of what we call megaloblastic erythropoiesis.

Etiology

Megaloblastic anemia has multiple etiologies (Table 12.1). Indeed, it is the main hematological manifestation of, on the one hand, classic inherited disorders (e.g. Lesch–Nyhan syndrome, orotic aciduria, transcobalamin deficiency) and, on the other hand, classic acquired disorders such as pernicious anemia and nutritional deficiency of either vitamin B_{12} (= cobalamin, Cbl) or folate. In this respect, the acquired (and far more common) conditions can be regarded as *phenocopies* of the much more rare inherited conditions.

Pathophysiology

While megaloblastosis is directly defined by its morphology, the mechanism of the anemia is complex. In general, anemia due to excessive destruction of red cells is characteristically associated with a cellular marrow and a high output of reticulocytes, whereas anemia due to decreased production of red cells is characteristically associated with hypocellular marrow and a low output of reticulocytes. In megaloblastic anemia, the marrow is hypercellular, often to an extreme degree, but the output of reticulocytes is low. This contrast is the *pathophysiological hallmark* of megaloblastic anemia, and it signifies that a large proportion of megaloblasts fail to mature into viable red cells; in other words, there is a vast component of *ineffective erythropoiesis*.

Table 12.1 Classification of the megaloblastic anemias.

FOLATE DEFICIENCY	
Inherited	**Acquired**
Inborn errors of folate metabolism	*Decreased intake*
Congenital folate malabsorption	Old age, alcoholism
Dihydrofolate reductase deficiency	Hemodialysis
Methylene tetrahydrofolate reductase deficiency	*Impaired absorption*
	Coeliac disease
	Tropical sprue
	Increased requirements
	Pregnancy
	Other increased cell turnover
	Chronic hemolytic anemia
	Drugs
	Inhibitors of dihydrofolate reductase deficiency
	Anticonvulsants
COBALAMIN DEFICIENCY	
Inherited	**Acquired**
Inborn errors of cobalamin transport	*Impaired absorption*
Imerslund–Grasbeck disease	Gastric causes
Congenital deficiency of intrinsic factor	Pernicious anemia
Gastrectomy	Intestinal causes
Inborn errors of cobalamin metabolism	Blind loop syndrome
Methionine synthase deficiency (cblE and cblG)	Fish tapeworm
cblC and cblD disease	Pancreatic insufficiency
	Decreased intake
	Vegetarians
	Other
	Nitrous oxide exposure
OTHER	
Hereditary orotic aciduria	
Lesch–Nyhan syndrome	
Thiamine-responsive megaloblastic anemia	

Modified from Babior BM. (1995) The megaloblastic anemias. In: Beutler E, Lichtman MA, Coller BS, Kipps TJ. (eds). *Williams Hematology*, 5th edn. New York: McGraw-Hill, p. 471.

Molecular pathogenesis

In view of the above, it is clear that any explanation of the pathogenesis of megaloblastic anemia at the molecular level must, on the one hand, rationalize it as the final common pathway of a variety of underlying lesions and, on the other hand, account for the phenomenon of ineffective erythropoiesis. Since the inherited causes of megaloblastic anemia are defects in the purine or pyrimidine biosynthetic pathways, and folate is the coenzyme of these pathways, it is natural to focus on this area of metabolism. It has been held for a long time that in megaloblastic anemia a decreased concentration of nucleotide precursors becomes rate-limiting for DNA synthesis, and as a result cell proliferation is curtailed and therefore few cells are produced. However, since cell proliferation is most active, this model can be rejected out of hand: the marrow is hypercellular rather than hypocellular, suggesting that the underlying defect may be qualitative and not merely quantitative.

Since the hematological consequences of the deficiency of either folate or Cbl are indistinguishable, it seems reasonable to surmise that there must be at least one point in common in their action, or that one depends on the other (Figure 12.1). Indeed, Cbl is required for the conversion of methyltetrahydrofolate (methylTHF), the main form of folate in the serum, to THF, the active form in bone marrow cells. Various derivatives of THF intervene in several steps of the biosynthesis of purine and pyrimidine nitrogen bases, but when folate is in

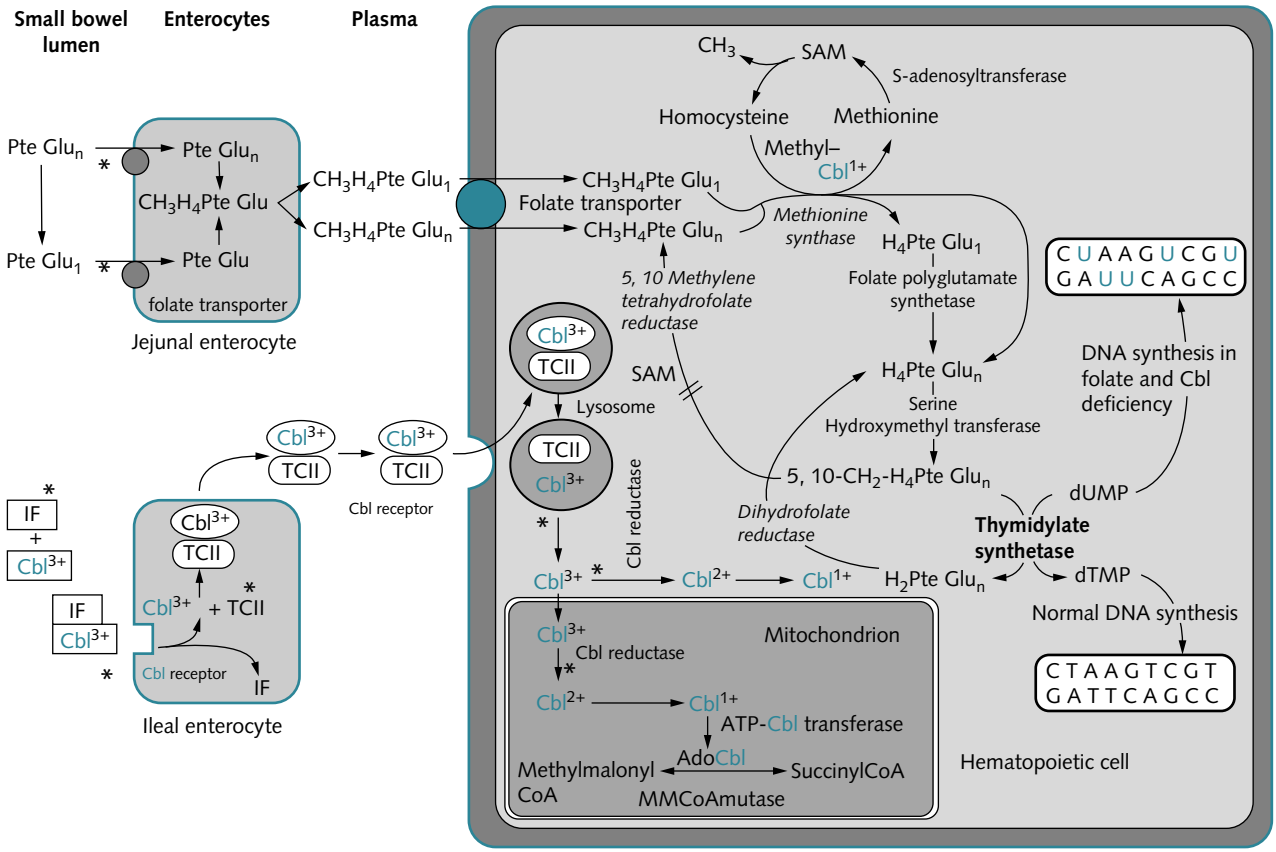


Fig. 12.1 The metabolic basis of megaloblastosis in folate and vitamin B_{12} (cobalamin, Cbl) deficiency

The absorption of folates takes place in the proximal small bowel while Cbl bound to intrinsic factor (IF) is absorbed in the ileum. Folate enters the cells in the form of methyltetrahydrofolate (methylTHF). Cbl is transferred to and enters the cells bound to transcobalamin II (TCII). In the cytoplasm, Cbl is necessary for the reaction catalysed by methionine synthase, whereby the CH_3 group of methylTHF is transferred to homocysteine; as a result, THF and methionine respectively are produced. The polyglutamated form of THF is converted to 5,10-methyleneTHF, which donates the single carbon group CH_2^- to the reaction catalysed by thymidylate synthetase, whereby dUMP is converted to dTMP, which is used in DNA synthesis. Under conditions of folate and/or Cbl deficiency, there is a shortage of 5,10-methyleneTHF. The result of this is, on the one hand, that dTMP is drastically reduced and not available for DNA synthesis and, on the other hand, that dUMP is in excess. Strong evidence exists suggesting that, under such circumstances, dUMP is misincorporated in the DNA, leading eventually to changes characteristic of megaloblastosis (*see text*). Note also that under Cbl-replete conditions the conversion of 5,10-methyleneTHF to methylTHF is inhibited by S-adenosylmethionine (SAM); by contrast, in Cbl deficiency SAM is in short supply and consequently this inhibition is relaxed, diverting the formation of 5,10-methyleneTHF to methylTHF, thus exacerbating the shortage of dTMP and the accumulation of dUMP. Cbl also plays a significant role in the mitochondrial metabolic pathways necessary for the conversion of the products of propionate metabolism (i.e. methylmalonyl-CoA) into easily metabolized products. As is evident from the metabolic inter-relationships of folate and Cbl, in folate deficiency homocysteine levels will increase; in Cbl deficiency, not only homocysteine but also methylmalonyl-CoA (and methylmalonic acid) will be increased: indeed, measurement of the serum and urine levels of homocysteine and methylmalonic acid is used in clinical practice for the diagnosis of folate and Cbl deficiency, especially at early stages.

Explanatory notes

In the above figure, the various forms of Cbl are shown in red. Pte$Glu_{1/n}$, mono- or polyglutamated forms of folate; CH_3H_4PteGlu, methylTHF; 5,10-CH_2H_4PteGlu_n, 5,10-methyleneTHF; H_2PteGlu_n, dihydrofolate. Enzymes shown in italics are those the hereditary deficiency of which causes megaloblastic anemia. Asterisks indicate steps in folate and Cbl metabolism whose defects can also cause hereditary megaloblastic anemia.

short supply these steps can be bypassed by using preformed bases (the so-called salvage pathway). The one reaction for which folate (in the form of 5,10-methyleneTHF) is irreplaceable is the conversion of dUMP to dTMP, for which it is the methyl group donor; as a result, this conversion will be impaired when either folate or Cbl is deficient. This reaction is crucially important because thymidine is, of course, the one base in which DNA normally differs from RNA. In principle, one might have expected that a block of this reaction would prevent DNA synthesis, but we have seen that this

is not the case. On the other hand, it is well established that *in vitro* DNA polymerase is able to incorporate dUTP into DNA, especially if the dUTP concentration is much higher than that of dTTP, which will be the case when the conversion of the former to the latter is impeded. Several studies indicate that this also takes place *in vivo*; indeed, the major molecular lesion in megaloblastic anemia may be this misincorporation of dUridine (dU) instead of T into DNA. The cell's DNA replicating machinery includes an enzyme, uracil glucosidase, which has the specific function of removing dU, should it be occasionally and illegitimately incorporated into DNA. Therefore dU will be retained in newly synthesized DNA only when the capacity of uracil glucosidase to remove it has been exceeded. For this reason, very little dU is actually found in megaloblastic DNA, but even that may have a disruptive effect on chromatin structure. Perhaps more importantly, if dU incorporation has been rampant rather than occasional, the numerous strand breaks produced by uracil glucosidase may exceed the capacity of other repair enzymes, thus causing the accumulation of damaged DNA and eventually cell death. Surprisingly, a careful study has not detected in megaloblastic bone marrow features regarded as characteristic of apoptosis. We must therefore presume that cell death takes place by a different pathway or that phagocytosis of dead cells is rapid and highly efficient.

If shortage of dTTP is the fundamental metabolic defect underlying megaloblastic anemia, since its production is 5,10-methyleneTHFA-dependent, the formation of which is in turn Cbl-dependent, it is clear that this same mechanism explains why megaloblastic anemia is the common manifestation not only of nutritional folate and Cbl deficiency but also of all genetically determined lesions of the transport and metabolism of folate or Cbl (Table 12.1). It is not clear why megaloblastic anemia should occur in other inherited conditions, such as HPRT deficiency and orotic aciduria, in which it is the salvage pathway rather than the *de novo* pathway of the nitrogen bases that is compromised. At the moment we can only speculate by analogy. Perhaps the metabolic blocks in these conditions also entail serious alterations in the absolute and/or relative pool sizes of the various deoxynucleoside triphosphates. Once again, this could cause misincorporation followed by repair attempts that are not always successful. In summary, although the rate of uracil incorporation into DNA has been controversial for years, a recent review validates its role in the pathogenesis of megaloblastic anemia.

Finally, despite the significant advances in the understanding of megaloblastosis, the molecular basis of demyelination that underlies the neurological complications of advanced Cbl deficiency remains elusive.

Congenital dyserythropoietic anemias

Congenital dyserythropoietic anemia (CDA) is the current designation for a group of rare inherited disorders that have a common feature: abnormalities in the maturation of the erythroid lineage. It is evident from genetics and from morphology that they are heterogeneous, and it is likely that they may be even more heterogeneous at the molecular level. Of the three classical forms of CDA (Table 12.2), CDA II (or HEMPAS) is the best defined, on account of a pathognomonic serological test; CDA I is defined by characteristic ultrastructural

Table 12.2 Defining features of CDA types I–III.

	Type I	Type II	Type III
Inheritance	Autosomal recessive	Autosomal recessive	(a) Autosomal dominant (b) Autosomal recessive
Location of gene	15q15.1–15.3	20q11.2	15q21–25
Identity of gene	CDAN1	Not known	Not known
Red cells	Macrocytes	Normocytes	
Erythroblasts			
(a) Light microscopy	Megaloblastic; internuclear chromatin bridges	Normoblastic; binuclearity predominates	Megaloblastic; up to 12 nuclei per cell
(b) Electron microscopy	Swiss cheese appearance of heterochromatin	Peripheral double membranes	
Serology			
Ham's test	Negative	Positive	Negative
Anti-i agglutinability	Normal/strong	Strong	Normal/strong
SDS-PAGE	Normal	Band 3 thinner and faster	Band 3 slightly faster

Modified from Wickramasinghe SN. (1997) Dyserythropoiesis and congenital dyserythropoietic anaemias. *British Journal of Haematology*, **98**, 785–797.

changes in the chromatin of erythroblasts and by autosomal recessive inheritance; CDA III is defined by large, sometimes multinucleated, erythroblasts and by autosomal dominant inheritance. A variety of terms have been used to classify patients who have features of CDA but do not fit neatly in any of these three categories.

As a result of a deranged developmental program, the mature red cells that are produced in CDA are macrocytic and abnormal in their membrane; this often entails a hemolytic component in their anemia. In addition, and most characteristically, a significant proportion of erythroid cells fail to achieve full maturity, and as a result they are destroyed in the bone marrow. Thus, the pathophysiological hallmark in CDA, just as in acquired megaloblastic anemias (*see above*), is ineffective erythropoiesis.

The biochemical basis for abnormal maturation has been well characterized in the case of CDA II. Sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of red cell membrane proteins reveals an increased sharpness of band 3, the size heterogeneity of which is normally produced by the variable size of its carbohydrate moiety. This finding has focussed attention on the enzymes required for glycosylation of membrane proteins: decreased activity of α-mannosidase and of fucosyl transferase has been reported in individual cases. Targeted inactivation of the gene encoding the latter enzyme has produced mice with features of CDA II. On the other hand, very recently it has been reported that inactivation of *AE1*, the gene encoding band 3, produces in zebrafish some features of human CDA II. However, by linkage analysis CDA II maps to 20q11.2 in most families, whereas fucosyl transferase maps to 11q21 and *AE1* maps to 17q21-q22. Thus, the gene that is mutated in human CDA II still needs to be identified. Variability of clinical expression of CDA II could be due to different underlying genetic lesions (and of course also to different mutant alleles at the same locus). In fact, there is some indirect evidence that alleles causing mild CDA may be relatively common, because two cases have been reported as causing chronic hemolytic anemia in association with glucose-6-phosphate dehydrogenase (G6PD)-deficient variants which do not, on their own, cause this condition. We do not know whether the CDA mutations present in these patients would have caused clinical manifestations in the absence of G6PD deficiency.

The gene for CDA I has been mapped to chromosome 15 by linkage analysis in a single Swedish family (one of the first from which the concept of CDA developed). Very recently this linkage has been confirmed in Bedouin families, and this has led to the identification of a gene that is mutated in all of these families, which has been called codanin-1. Codanin-1 has a 150-residue amino-terminal domain with sequence similarity to collagens, and two shorter segments that show weak similarities to the microtubule associated proteins, MAP1B (neuraxin) and synapsin.

In view of the fact that in the various forms of CDA the abnormal phenotype is almost exclusively restricted to the erythroid lineage, it is likely that the genes that are mutated in any patient with CDA serve some important role in the program of erythroid differentiation. For instance, it seems likely that in erythroid cells codanin-1 is involved in nuclear envelope integrity, conceivably related to microtubule attachments. Thus, each one of the CDA genes will be of great interest, quite out of proportion to the rarity of CDAs as clinical entities.

It has been reported that three patients with CDA I have responded to treatment with interferon-α with near normalization of the hemoglobin values. Although the clinical data seem convincing, at the moment it is not clear by what mechanism an intrinsic erythroid molecular abnormality can benefit from this treatment.

The sideroblastic anemias

The term 'sideroblastic anemia' (SA) encompasses a diverse collection of diseases in which different causes and different mechanisms converge to produce the same rather spectacular feature: the accumulation of inorganic iron in the cytoplasm of erythroid cells in sufficient quantities to be easily demonstrated in the form of granules by using Perl's Prussian Blue staining. Characteristically, the iron is found in mitochondria positioned around the nucleus — hence the term 'ring sideroblast'. SAs are broadly divided into *acquired* and *inherited* forms (Table 12.3). Acquired SAs are the most common and are discussed in Chapter 8. Here, we discuss briefly the three inherited forms of SA for which the genetic and biochemical bases have been elucidated.

ALAS2 deficiency

Normally, about 90% of the iron obtained daily from the diet reaches the erythroblasts, where it is required for the final step of heme biosynthesis, namely the incorporation of iron into the tetrapyrrolic ring of protoporphyrin IX. The first and the last three of the eight steps of the heme biosynthetic pathway take place in the mitochondrion. The first and rate-limiting step consists of the condensation of glycine and succinyl-coenzyme A to δ-aminolevulinate. This reaction is catalyzed by δ-aminolaevulinate synthase (ALAS) and it requires pyridoxal 5′-phosphate (PLP) as a cofactor. Two isoforms of ALAS are known; both have a homodimeric structure but they are encoded by different genes. *ALAS1* is an ubiquitously expressed housekeeping gene, whereas *ALAS2* is erythroid-specific.

Table 12.3 Classification of the sideroblastic anemias.

INHERITED				
Mode of inheritance	**Chromosomal locus**	**Gene**	**Type of mutation**	**Clinical manifestations other than SA**
X-linked	Xp11.21	ALAS2	Missense	None
X-linked	Xq1.3-q13.3	*hABC7* (Fe-Su cluster transporter protein)	Missense	Cerebellar ataxia
Autosomal recessive	1q23.2–23.3	*SLC19A2* (thiamine transporter protein)	Missense, nonsense, frameshift	Thiamine-responsive megaloblastic anemia, diabetes mellitus, sensorineural deafness
Autosomal dominant	Not known	Not known		None
Mitochondrial, e.g. Pearson's MPS	Usually from nt 8469 to nt 13447 (see Fig. 12.3)	Deletions		Pancreatic exocrine dysfunction, cytopenia, metabolic acidosis
ACQUIRED				
Refractory anemia with ring sideroblasts (RARS) Drug-induced (e.g. isoniazid, chloramphenicol, ethanol) Secondary to systemic, metabolic, malignant disorders				

ALAS2, δ-aminolevulinate synthase 2; PGK1, phosphoglycerate kinase 1; SA, sideroblastic anemia; MPS, marrow-pancreas syndrome.

Genetics of ALAS2 deficiency

The first family with what is now called X-linked SA was reported by T. Cooley in 1945. Subsequently, more families with apparent X-linked sideroblastic anemia were described. The observation that some affected males responded to pharmacological doses of pyridoxine focussed attention on the erythroid-specific ALAS: indeed, the *in vitro* activity of this enzyme was invariably reduced. The erythroid-specific *ALAS* (*ALAS2*) gene was subsequently cloned (Figure 12.2) and mapped to Xp11.12. This locus is subject to X-chromosome inactivation and as a result female relatives of affected males may demonstrate red cell mosaicism; that is, two populations of red cells, one hypochromic and one normochromic. In general, these women do not have anemia and therefore the disease is regarded as recessive. However, full-blown SA may develop in women, especially late in life, as a result of a skewed X-inactivation pattern: in such cases we would have to regard the same disease as dominant.

Transcriptional and translational control of ALAS2 expression

The expression of *ALAS2* is regulated at both the transcriptional and the translational level. Transcriptional control is effected through the programmed differentiation of the erythroid lineage, mainly through the erythropoietin-induced transcription factors, which include GATA-1, EKLF and NF-E2. In contrast, as one of the few examples of control of gene expression at the translational level, *ALAS2* is involved in the regulation of heme synthesis in relation to iron availability. Central to this regulatory mechanism is a protein called IRE-binding protein (IRE-BP), which requires four atoms of iron to function and is also able to bind to an mRNA motif, called iron-responsive element (IRE). IRE has a characteristic hairpin secondary structure, present within the 5′ untranslated region of the *ALAS2* mRNA. IRE-BP does not bind to the IRE when it is iron-replete, thus allowing the translation of the *ALAS2* mRNA to proceed unimpeded. When iron is in scarce supply, the iron-depleted IRE-BP binds to IRE and translation of the *ALAS2* mRNA is repressed, thus decreasing the synthesis of heme and consequently of mature hemoglobin and leading to the production of hypochromic red cells. Interestingly, IRE-BP can also bind to IREs present in the mRNAs of ferritin and transferrin receptor. In conditions of iron depletion, real or functional, the translational regulation is such that translation of ferritin mRNA is repressed whereas translation of the transferrin receptor is favored. Reverse effects are seen in iron-replete conditions.

Molecular pathology of ALAS2 deficiency

Over twenty different point mutations have been described (Figure 12.3). They all map to exons 5–11, which are highly

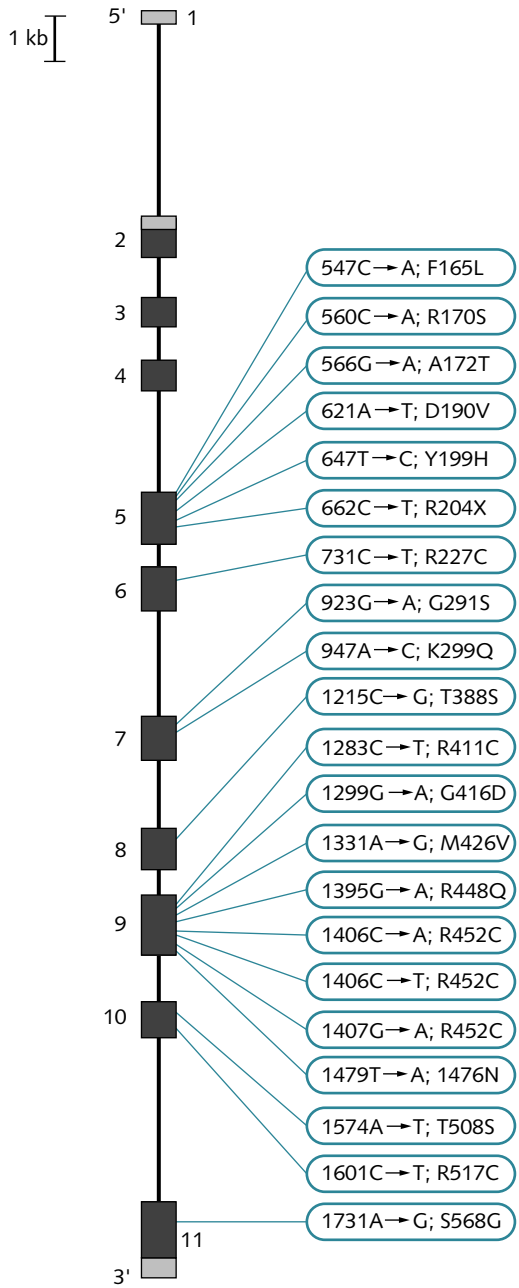


Fig. 12.2 The erythroid *ALAS2* gene and its pathogenic mutations

ALAS2 spans about 22 kb and consists of 11 exons. Two mature forms of mRNA exist: one full-length and one shorter as a result of alternative splicing of exon 4. The functional significance of these splicing variants is not known. Exon 1 and part of exon 2 (shown in gray) form the 5′-UTR and contain the iron-responsive element (*see text*). The full-length mRNA (1.95 kb) encodes a protein of 522 amino acids (64.4 kDa). The mature protein derives from the cleavage of the first 49 amino-terminal amino acids (mitochondrial signal sequence) upon entry in the mitochondrion.

conserved across species. Parts of these exons are thought to contribute to the formation of the catalytic site, the pyridoxal 5′-phosphate-binding site and the substrate-binding site of ALAS2. The presence of only missense mutations implies that hemizygotes with *ALAS2* null mutations would not be viable. Very recently a nucleotide replacement at position –206 from the transcription start of *ALAS2* has been found to cause SA; this is a rare example of a human disease resulting from a 'promoter-down' mutation in a gene encoding an enzyme.

Clinical aspects and treatment of ALAS2 deficiency

The clinical picture of ALAS2 deficiency is that of a hypochromic microcytic anemia with bone marrow erythroid hyperplasia as a result of ineffective erythropoiesis. The characteristic ring sideroblasts are found mainly in the late erythroid precursors. There is considerable heterogeneity in the severity of the disease, not only between individuals bearing different *ALAS2* mutations but also between related individuals with the same mutation. Patients at one extreme may present a few months after birth with severe anemia, severe microcytosis and no response to pyridoxine; at the other extreme they may present in the ninth decade of life with anemia fully responsive to pyridoxine.

Although the anemia of pyridoxine-responsive SA is treatable, the main complication of the disease is iron overload; if left untreated, it has the same deleterious results as hereditary hemochromatosis. Iron overload can be biochemically evident as early as in adolescence, does not correlate with the degree of anemia, and can affect mildly anemic females. The importance of effectively treating iron overload cannot be overemphasized for one further reason: excess iron interferes with the function of ALAS2 and patients previously unresponsive to pyridoxine, after effective iron chelation, occasionally become responsive.

The advances in the molecular aspects of X-linked SA make prenatal diagnosis and counseling feasible, especially for families with the severe, pyridoxine-resistant forms of the disease.

X-linked sideroblastic anemia with ataxia

This pyridoxine-resistant form of SA with congenital cerebellar ataxia has been described in two families. It is due to missense mutations of the human adenosine triphosphate binding cassette 7 (hABC7) gene, which maps to Xq1.3-q13.3. ABC7 belongs to a wider family of proteins involved in the transportation of substrates across cell and organelle membranes. As studies of its yeast ortholog indicate, hABC7 is involved in the transportation of clusters of Fe-Su (iron–sulfur) across the mitochondrial membrane to the cytosol. These clusters are required for the functional maturation of a number of

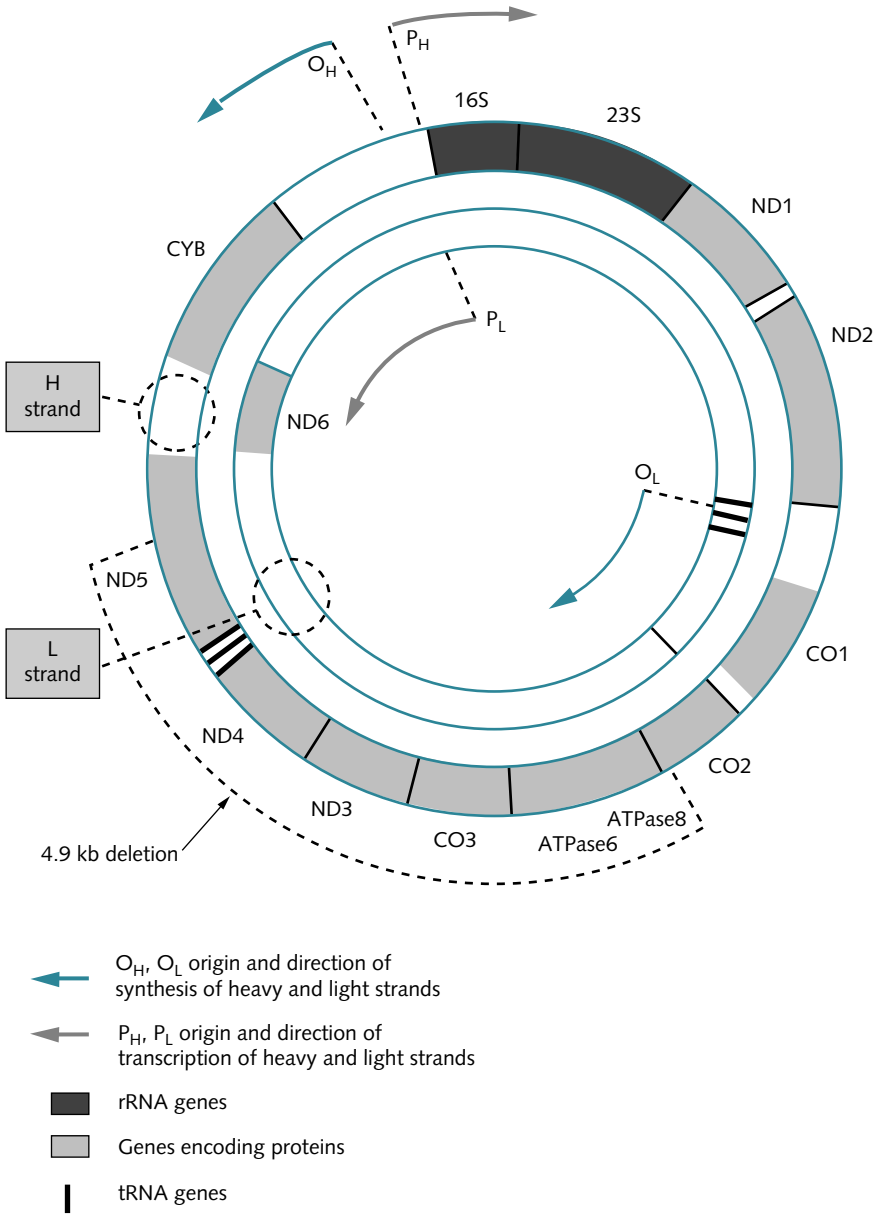


Fig. 12.3 The human mitochondrial genome
Both replication and transcription of the heavy (H) and light (L) strands of mtDNA run in opposite directions, as indicated in the figure. The products of transcription are large, and are cleaved to generate RNAs for individual genes. Note the absence of introns and close apposition of genes. *ATPase 8* and *ATPase 6* overlap. The approximate location of the common 4.9 kb deletion (*see text*) is indicated by the broken line. π*ND1–ND6* encode NADH dehydrogenase subunits; *CO1–CO3* encode cytochrome *c* oxidase subunits and *CYB* encodes cytochrome *b*. Modified from Strachan T, Read AP. (1996) *Human Molecular Genetics*, 1st edn. BIOS.

proteins. Mutations of hABC7 interrupt this process and lead to iron accumulation in the mitochondrion.

Pearson's marrow–pancreas syndrome

Pearson's marrow–pancreas syndrome (PMPS) is unique among the SAs because the underlying genetic lesion is not in the nuclear DNA but in the mitochondrial DNA (mtDNA).

Properties of mtDNA

Each somatic cell has hundreds to several thousand mitochondria, and each mitochondrion contains two to ten copies of mtDNA. mtDNA consists of a single double-stranded circular molecule, and its 16 569 nucleotide-long genome encodes 24 structural RNAs (22 tRNAs and 2 rRNAs), which are required for mitochondrial protein synthesis; in addition it encodes 13 protein subunits belonging to four enzyme complexes, all involved, directly or indirectly, in the generation of adenosine triphosphate (ATP) through oxidative phosphorylation (Figure 12.3). Replication, transcription and translation of mtDNA are all quite distinct from their nuclear DNA equivalents.

Since mitochondria are abundant in the cytoplasm of the mature oocyte but absent from that part of the sperm cell that

enters it at the time of fertilization, the hereditary transmission of mtDNA is exclusively through the maternal germline. Consequently, a mitochondrial disease can be transmitted from a mother to all of her children, whether male or female — a non-Mendelian form of inheritance.

mtDNA is estimated to be at least 10 times more vulnerable than nuclear DNA to mutations and their consequences for various reasons. Firstly, it is constantly exposed to oxygen free radicals generated by oxidative phosphorylation; secondly, mitochondria largely lack effective DNA repair mechanisms; thirdly, because most of the mtDNA sequence is coding, many more mutations will be reflected in the structure of its protein products. On the other hand, since there are so many copies of mtDNA in each cell, mutant mtDNA may coexist with normal mtDNA, a situation called 'heteroplasmy' (the opposite of homoplasmy). The phenotype of a particular cell or organ will depend on the relative proportions of normal and mutated mtDNA. Since, unlike nuclear chromosomes, mtDNA molecules have no rigorous mechanism for segregation at either mitosis or meiosis, the inheritance of mtDNA mutations is rather unpredictable; and since heteroplasmy may exist even in oocytes, this further complicates the non-Mendelian inheritance of diseases caused by mtDNA mutations.

Clinical aspects of PMPS

PMPS usually presents within the first few months of life with hypoproliferative SA, variable cytopenias and pancreatic exocrine dysfunction. Bone marrow examination reveals, in addition to ring sideroblasts, striking vacuolation of the erythroid and myeloid precursors (Plate 12.1). Metabolic acidosis is another frequent manifestation, and renal disease, liver failure, hypoparathyroidism and diabetes mellitus may also occur. Despite treatment with blood products, pancreatic enzymes and various vitamins (e.g. coenzyme Q), about half of the patients do not survive beyond the third year of life. Of the patients who do survive, most develop complications in other organs, particularly ophthalmoplegia, pigmentary degeneration of the retina and cardiomyopathy; that is, the features originally described as the Kearns–Sayre syndrome (KSS). The overlap between PMPS and KSS is not surprising, since similar genetic lesions are found in both conditions (*see below*).

Molecular pathology

In all but one of the cases that have been adequately investigated, PMPS was due to deletions within mtDNA. The size of the deletion is variable; in one study, 43% of the patients had an identical 4.9-kb deletion which has also been found in patients with other mitochondrial diseases (such as KSS and progressive external ophthalmoplegia; Figure 12.3). In some cases, deletion dimers and/or deletion multimers were observed. There is no obvious correlation between the size or location of the deletion and the clinical severity of the disease; rather, what probably largely determines the clinical phenotype is the proportion of mutated mtDNA in a particular tissue. For example, patients with PMPS have mutated mtDNA in all tissues examined, whereas in patients with classical KSS the mutated mtDNA is restricted to muscle and is not found in blood cells. Since the mothers of PMPS patients are invariably unaffected, it can be expected that the deletion has taken place *de novo*, and this has been documented in a number of cases.

A difficult practical problem is that of genetic counseling for couples who have had the misfortune of having a child with PMPS. The risk of recurrence depends on whether the deletion is present in most, or only in some, of the mother's gonadal cells. At the moment there is no established method for determining this. At an experimental level, it could be achieved by inducing multiple ovulation and using the polymerase chain reaction (PCR) to test individual oocytes for the deletion previously detected in an affected sib. The same could be done in very early embryos after *in vitro* fertilization (pre-implantation prenatal diagnosis). Because of the high copy number of mtDNA, testing for mtDNA mutations is much easier than testing for mutations in nuclear genes.

Conditions associated with bone marrow failure

The disease entities falling under this heading are quite diverse, but they are grouped on account of a common pathogenesis—the loss of hemopoietic stem cells (HSCs). The pace of HSC depletion varies widely, from weeks (e.g. in idiopathic aplastic anemia, IAA) to years (e.g. in Fanconi's anemia, FA). In fact, IAA and FA exemplify two broad categories of bone marrow failure (BMF) syndromes; that is, *acquired* and *inherited*, respectively (Table 12.4).

Acquired bone marrow failure syndromes

Idiopathic aplastic anemia

Idiopathic aplastic anemia (IAA) accounts for the majority (~80–90%) of cases of acquired BMF syndromes, with an incidence estimated at two cases per million (somewhat higher in the Orient).

Immunopathogenesis of idiopathic aplastic anemia

The most direct evidence that IAA may be an autoimmune disorder has come from the clinical observation that patients

Table 12.4 Classification of the bone marrow failure (BMF) syndromes.

INHERITED

Disease	Mode of inheritance	Chromosomal locus	Gene	Clinical manifestations
Fanconi's anemia	Autosomal recessive	See Table 12.5	See Table 12.5	See text
Dyskeratosis congenital	X-linked autosomal recessive	Xq28 3q26	DKC1, hTR	See text
Diamond–Blackfan anemia	Autosomal dominant Autosomal recessive	19q13.2 8p22-p23.31	RPS19 Not known	See text
Shwachman–Diamond syndrome	Autosomal recessive	Not known	Not known	Neutropenia, exocrine pancreatic insufficiency, metaphyseal dysostosis
Amegakaryocytic thrombocytopenia	Autosomal recessive	1p34	c-mpl (thrombopoeitin receptor)	Absent megakaryocytes in bone marrow, late BMF
Thrombocytopenia with absent radii syndrome	Autosomal recessive?	Not known	Not known	Bilateral radial aplasia, lower limb anomalies, cow's milk intolerance, renal and cardiac anomalies
Congenital thrombocytopenia and radio-ulnar synostosis	Autosomal dominant	7p15-p14	HOXA11? (needs confirmation)	Aplastic anemia, proximal radial ulnar synostosis, clinodactyly, syndactyly, hip dysplasia and sensorineural hearing loss
Pearson's marrow–pancreas syndrome	Mitochondrial	Usually from nt 8469 to nt 13447	Contiguous genes deleted	Pancreatic exocrine dysfunction, sideroblastic anemia (*see also text*)
Familial aplastic anemias	There are numerous reports of family clustering of BMF other than the above. They have shown different patterns of inheritance and a variety of associated clinical manifestations (e.g. malformations), suggesting that this group may be genetically heterogeneous.			

ACQUIRED

Idiopathic
Radiation
Drugs and chemicals
  Regular: cytotoxic, benzene
  Idiosyncratic: chloramphenicol, NSAIDs, antiepileptics, gold
Viruses
  Epstein–Barr virus
  Hepatitis
  Parvovirus
  Human immunodeficiency virus
Immune diseases
  Thymoma
  Pregnancy
  Paroxysmal nocturnal hemoglobinuria

NSAID, non-steroidal anti-inflammatory drug.

with IAA have complete or partial reversion of their pancytopenia when they are treated with antilymphocyte globulin (ALG). Subsequently, it was shown that patients with IAA often have increased numbers of 'activated' $CD8^+$ $CD25^+$ T cells in their blood and bone marrow. In addition, T cells from IAA patients can inhibit the growth of autologous *in vitro* hemopoietic colonies, and the growth of colonies from HLA-identical siblings. Based on these observations, a cur-

rent model of the pathogenesis of IAA predicts that autoreactive T cells attack HSCs, causing their depletion—hence the reduction of HSCs in severe IAA to about 1% of normal. The primary event that triggers this aberrant immune response remains elusive: a possible viral cause has long been sought but never proved. The identity of the putative autoantigen on HSCs also remains unknown. There is evidence, however, that the inhibitory effect of autoreactive T cells is mediated, at least in part, through interferon-γ (IFN-γ). In addition, IFN-γ upregulates Fas receptor on the surface of HSCs, thus facilitating activation of the Fas-dependent apoptotic pathways.

As in other autoimmune diseases there is over-representation of specific HLA alleles in IAA patients compared with population controls: the HLA-DR2 allele is over-represented in patients with IAA of European ancestry, whereas another HLA class II haplotype is over-represented in Japanese patients with IAA.

Recently, a pathogenetic link between IAA and myelodysplastic syndrome (MDS) has surfaced. Clinically, IAA overlaps with the hypoplastic form of MDS; in fact, the differential diagnosis between the two is often difficult. It is now recognized that 20–30% of patients with MDS, especially those who would otherwise be classified as having refractory anemia with hypocellular marrow, respond to immunosuppressive therapy with alleviation of their cytopenias, indicating that an immune process, similar to the one operating in IAA, is also involved in the pathogenesis of some forms of MDS. In addition, as in IAA, in 20% of patients with refractory anemia, paroxysmal nocturnal hemoglobinuria (PNH) populations of small size are detected (*see below*). What is more, their presence is predictive of response to immunosuppressive therapy, providing further evidence of an immune process in the pathogenesis of MDS in this select group of patients, and linking them to those with IAA.

Clinical aspects and treatment of idiopathic aplastic anemia

The clinical picture of IAA generally reflects the extent of HSC loss and the subsequent cytopenias. Typically, a patient with severe IAA presents with bruising and mucosal bleeding, anemia and septic episodes (bacterial or fungal), but without hepatosplenomegaly. The differential diagnosis of IAA, as well as hypoplastic MDS, includes inherited BMF syndromes, the aplastic form of childhood ALL, infectious and malignant processes that may infiltrate the bone marrow. Thus, the diagnosis of IAA is made eventually by exclusion. The contemporary treatment of IAA is dictated by the severity of IAA (as determined by the degree of pancytopenia, reticulocytopenia and bone marrow cellularity) and by the patient's age. HSC transplantation (HSCT) from an HLA-identical sibling or from an alternative donor is the treatment of choice for younger patients with severe IAA, and offers better than 65% long-term survival. In the absence of an appropriate donor, or when the patient is older or the disease milder, immunosuppressive treatment (in particular the combination of ALG/ATG and cyclosporin A) results in a complete or partial response in most cases. As well as PNH (*see below*), IAA bears a significant risk of late clonal disorders, especially after immunosuppressive therapy: the risk of MDS and acute myeloid leukemia (AML) is 7% 10 years after immunosuppression, and the risk of tumors of other organs brings the total risk to 18% (compared with 3.1% after bone marrow transplantation).

Paroxysmal nocturnal hemoglobinuria

This is a rare acquired hematological disorder with three main clinical features: intravascular hemolysis, a tendency to thrombosis and BMF of variable severity. As in IAA, the precise cause of the BMF remains unclear; by contrast, the molecular and cellular events responsible for hemolysis are now explained. For this reason, the space devoted here to this condition is out of proportion to its prevalence.

Molecular pathogenesis. The initiating event in the pathogenesis of PNH consists in somatic mutation(s) in the X-linked gene *PIG-A* (Figure 12.4) in multipotent HSCs. The protein product of *PIG-A*, although not yet physically isolated, is thought to be the enzymatically active subunit of a N-acetylglycosamine transferase. This enzyme catalyzes an early step in the formation of a complex glycolipid molecule called glucosylphosphatidylinositol (GPI; Figure 12.4). The synthesis of GPI takes place initially on the cytoplasmic surface of the endoplasmic reticulum, and is completed on its luminal surface. Once formed, the GPI molecules (anchors) are attached through a transpeptidation reaction to the carboxy-terminus of a variety of proteins. The GPI-linked proteins, after post-translational modifications in the Golgi apparatus, emerge eventually on the cell surface, to which they remain attached through the GPI anchor. As a result of the impaired synthesis of GPI, blood cells are either completely (PNH III) or partially (PNH II) deficient in GPI-linked proteins. Although for most of these proteins the functional consequences of this deficiency are not known, some are directly implicated in the pathogenesis of PNH: CD55 and especially CD59, which normally protect red cells from the lytic effect of activated complement. Intravascular hemolysis in PNH (manifesting clinically as hemoglobinuria) is the direct consequence of the fact that CD55- and CD59-deficient PNH red cells are susceptible to complement-mediated lysis. The *in vitro* counterpart of this phenomenon is the Ham test, whereby the patient's red cells are lysed by either autologous or ABO-compatible donor acidified serum. As for thrombosis, this may result in part from complement-mediated activation of platelets deficient in CD55 and CD59. Other, as yet unidentified, genetic or

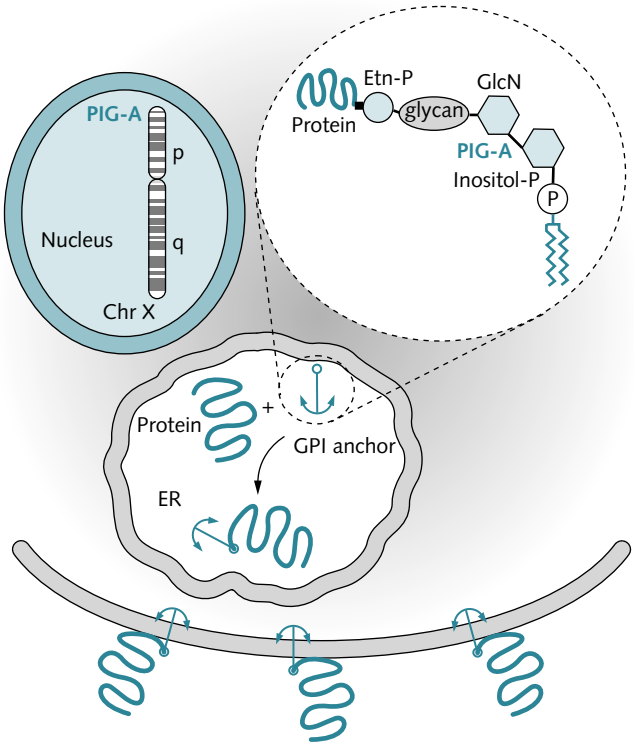


Fig. 12.4 *PIG-A*, GPI anchors, GPI-anchored proteins and PNH
PIG-A is a protein encoded by the X-linked gene *PIG-A*. PIG-A is a member of a multi-subunit enzymatic complex which catalyzes, in the endoplasmic reticulum (ER), the first step in the biosynthesis of GPI: the addition of acetylglucosamine (GlcN) to phosphatidylinositol (inositol-P; inset). The synthesis of the GPI anchor is completed by the serial addition of a glycan moiety consisting of three mannose molecules and a molecule of phosphoethanolamine (Etn-P), to which, through a transpeptidation reaction, proteins with the appropriate carboxy-terminal amino acid motif are attached covalently. The GPI–protein complex subsequently travels to the cell surface, where the protein becomes anchored to the lipid bilayer through GPI. In PNH, *PIG-A* has suffered a somatic mutation in one or few HSCs. As a result, very little GPI is synthesized, or none at all, with consequent severe deficiency of GPI-anchored proteins on the surface of the mutated HSCs and their progeny. From Karadimitris A, Luzzatto L. (2001) The cellular pathogenesis of paroxysmal nocturnal hemoglobinuria. *Leukemia*, **15**, 1148–1152.

acquired factors affecting coagulation and/or fibrinolysis may have an additive or synergistic effect in producing thrombosis, which may be devastating, especially as it tends to take place in the abdominal veins.

Cellular pathogenesis. Since PNH-HSCs lack GPI-linked proteins, in principle they would be expected to be poorly competitive in growth with respect to normal HSCs. Instead, PNH-HSCs can expand until they largely supplant normal hematopoiesis. As a first approximation, this paradox may be explained by invoking an intrinsic proliferative advantage of PNH-HSCs over normal HSCs, as is the case with leukemic cells: however, the fact that patients with PNH can live for decades with normal and PNH hemopoiesis coexisting in their bone marrow militates against this notion. Lack of a competitive growth advantage by the PNH hemopoiesis has also been demonstrated experimentally in *pig-a* null mouse models.

An alternative pathogenetic model—the escape model—was suggested by the long-known association between PNH and IAA. In this model, the link between PNA and IAA is the effect of HSC-specific T cells. It is predicted that in PNH, such cells would selectively target normal HSCs but not PNH-HSCs. Under these circumstances, PNH-HSCs would expand and contribute to hemopoiesis (in some cases as much as 90% of it). Clinical evidence in support of the escape model is provided by the appearance of small PNH clones in as many as 50% of patients with bona fide IAA and by the presence of minute PNH clones (~1 in 10^5 of granulocytes) even in normal individuals. Further evidence supporting the immune model is provided by (1) the presence of expanded T-cell clones in the blood of PNH patients at a frequency three-fold higher than in appropriate controls; and (2) the increased representation of the HLA-DR2 allele in PNH patients compared with population controls (as in IAA).

The molecular target of the postulated autoreactive T cells is not known. Potential targets are a surface GPI-linked protein, failure of PNH-HSCs to present in the context of HLA an immunogenic peptide derived from GPI-linked protein or the GPI molecule itself. Experimental evidence has ruled out the first mechanism; the last two are still to be tested. Recently, it was reported that NK cells are more effective in lysing GPI^+ targets than GPI^- targets. However, two previous studies had found no difference; clearly, the role of NK cells in the pathogenesis of PNH needs further exploration.

Molecular pathology. All types of mutations have been observed in the *PIG-A* gene in patients with PNH: a few are large deletions, the majority (~75%) are small insertions or deletions causing frameshifts, and the rest are nonsense and missense point mutations. Interestingly, the nonsense and frameshift mutations are spread throughout the coding sequence (exons 2–6), presumably because they cause complete inactivation of the gene product wherever they fall, whereas missense mutations are clustered mainly within exon 2, where it is presumed that amino acid residues critical for catalytic activity must be located.

Clinical aspects and treatment of PNH. Although hemoglobinuria is paroxysmal by definition in the prototypical PNH patient, the brisk intravascular hemolysis is in fact continuous. Additionally, patients with florid PNH (Table 12.5) often experience acute exacerbations of their hemolysis during intercurrent illnesses, such as infections (presumably because this is associated with activation of complement through either the classical or the alternative pathway), other stressful events, or for no obvious reason.

Table 12.5 PNH: clinical heterogeneity and proposed terminology.

Predominant clinical features	Blood findings	Size of PNH clone	Designation
Hemolysis ± thrombosis	Anemia; little or no other cytopenia	Large	Florid PNH
Hemolysis ± thrombosis	Anemia; mild to moderate other cytopenia(s)	Large	PNH, hypoplastic
Purpura and/or infection	Moderate to severe pancytopenia	Large	AA/PNH
Purpura and/or infection	Severe pancytopenia	Small	AA with PNH clone
Thrombosis	Normal or moderate cytopenia(s)	Small	Mini-PNH

From Tremml G, Karadimitris A, Luzzatto L. (1998) Paroxysmal nocturnal hemoglobinuria: learning about PNH cells from patients and from mice. *Haema*, **1**, 12–20.

Hemolytic anemia with macrocytosis (partly due to reticulocytosis and partly due to BMF), different degrees of thrombocytopenia and leukopenia, iron deficiency and hemosiderinuria should raise the suspicion of PNH. The diagnosis can be confirmed with the Ham test (which only detects the PNH abnormality in erythrocytes) and/or flow cytometric analysis of erythrocytes and leukocytes; for this purpose, anti-CD59 is the most reliable antibody. In about 40% of PNH patients, venous thrombosis of small or large vessels (particularly at unusual sites, such as the abdomen or the brain) is a serious and potentially life-threatening complication. If promptly diagnosed, venous thrombosis may respond to thrombolytic therapy, which must be followed by short- and long-term anticoagulants. Supportive treatment consists of blood transfusions, folic acid and iron supplements when needed. The advent of the humanized monoclonal antibody Eculizumab, an inhibitor of the C5 component of complement, has been a major therapeutic advance. In a pilot study of 12 patients hemolysis was almost completely halted and at least half of the patients became transfusion dependent; the rest had a dramatic reduction of their transfusion requirements. Side effects were mild. It seems very likely, that if these results are confirmed in a larger study, Eculizumab will replace blood transfusion as the treatment of choice for anemia.

Long-term therapeutic options for PNH include immunosuppressive agents (e.g. combination of ALG/ATG and cyclosporin A, especially for patients with moderate to severe cytopenias) and, for selected patients, HSC transplantation from an HLA-identical sibling. In principle, PNH should also be amenable to reduced intensity allografting protocols.

PNH and other clonal disorders. Patients with PNH have a small risk (<4%) of developing MDS (the reverse, i.e. patients with MDS developing PNH, is discussed above) and AML. Because a similar risk exists in IAA, it is probably the perturbed marrow environment that allows the emergence of premalignant (MDS) or malignant (AML) clones, rather than the *PIG-A* mutations *per se* predisposing to MDS and AML.

Inherited bone marrow failure syndromes

Fanconi's anemia

Fanconi's anemia (FA) is the most common cause of hereditary BMF. Early studies indicated that FA was a genetically heterogeneous disease, a notion confirmed by the identification, through the use of somatic cell hybridization, of as many as eight complementation groups (FA-A, B, C, D1, D2, E, F, and G). Seven FA genes have now been cloned and characterized. The overall frequency of heterozygotes for FA mutant genes in the general population is estimated at 1 in 300. In Ashkenazi Jews and in the Afrikaans population of the Republic of South Africa it is much higher (1 in 100 and 1 in 89, respectively), most likely as a result of founder effects.

Clinical aspects. Gradual onset of BMF (median age, 7 years; range, birth to 31 years), skeletal abnormalities (most commonly of the radius and thumb), skin lesions (hyperpigmentation, café-au-lait spots), renal and urinary tract malformations, and gonadal dysfunction are the most common clinical manifestations. However, the clinical spectrum is even wider, as it includes congenital defects of the gastrointestinal system, heart and central nervous system. BMF is often heralded by thrombocytopenia, macrocytosis and increased Hb F levels. Patients with FA have an unusually high risk of developing treatment-resistant MDS and AML, estimated at 52% in total by the age of 40 years. Furthermore, the risk of a variety of solid tumors, especially squamous cell carcinomas of the skin, is several times higher than in the general population.

Molecular genetics. FA is an autosomal recessive disorder. The most characteristic cellular feature of FA cells is the formation of DNA double-strand breaks upon exposure to DNA inter- and intrastrand adducting agents (clastogens), such as mitomycin and diepoxybutane. The *in vitro* response to clastogens has made it possible to test cells from different patients for their ability to cross-correct each other's defect by somatic cell fusion. As discussed earlier, this has led to the classification of patients in eight complementation groups (FA A–H).

Table 12.6 The identity of the FA genes and proteins.

Subtype	Chromosome	Relative frequency	Exon	Protein (kDa)
A	16q24.3	66	43	163
B	13q12-13	4	27	380 (BRCA2)
C	9q22.3	12	14	63
D1	13q12-13	Rare	27	380 (BRCA2)
D2	3p25.3	Rare	44	155,162
E	6p21-22	Rare	10	60
F	11p15	Rare	1	42
G	9p13	10	14	68

Using the technique of complementation cloning, *FANCC* was identified first. In this approach, a cDNA library, inserted into an EBV-based episomal expression vector, is transfected into immortalized B-cell lines derived from patients with FA. Cells complemented with the correct cDNA survive the treatment with a clastogenic agent; the episomal DNA recovered from the surviving cells contains inserts that are candidates for the disease gene. The disease gene is identified by its capacity to correct the phenotype of deficient cells and by screening for pathogenic mutations in affected individuals and family members. Using the same approach, another four genes were cloned: *FANCA*, *FANCG*, *FANCF* and *FANCE* (Table 12.6). The recent identification of *FANCD2* was achieved through a combination of complementation and positional cloning. Complementation was achieved using a panel of microcells lacking one specific chromosome. By fusing such cells with fibroblasts derived from FA-D patients, it was found that the gene of interest lay on chromosome 3. Further complementation experiments using cells with partially deleted chromosome 3 led to the identification of a critical area of 200 kb which contained the *FANCD2* gene.

A major advance in the clarification of the genetics of FA and also in our understanding of the molecular pathogenesis of the disease was the finding that the breast cancer susceptibility gene *BRCA2* is mutated in FA-B and FA-D1 patients (*see below*).

Molecular pathology and population genetics. (Figure 12.5) Mutations of *FANCA* account for about 60% of FA cases and

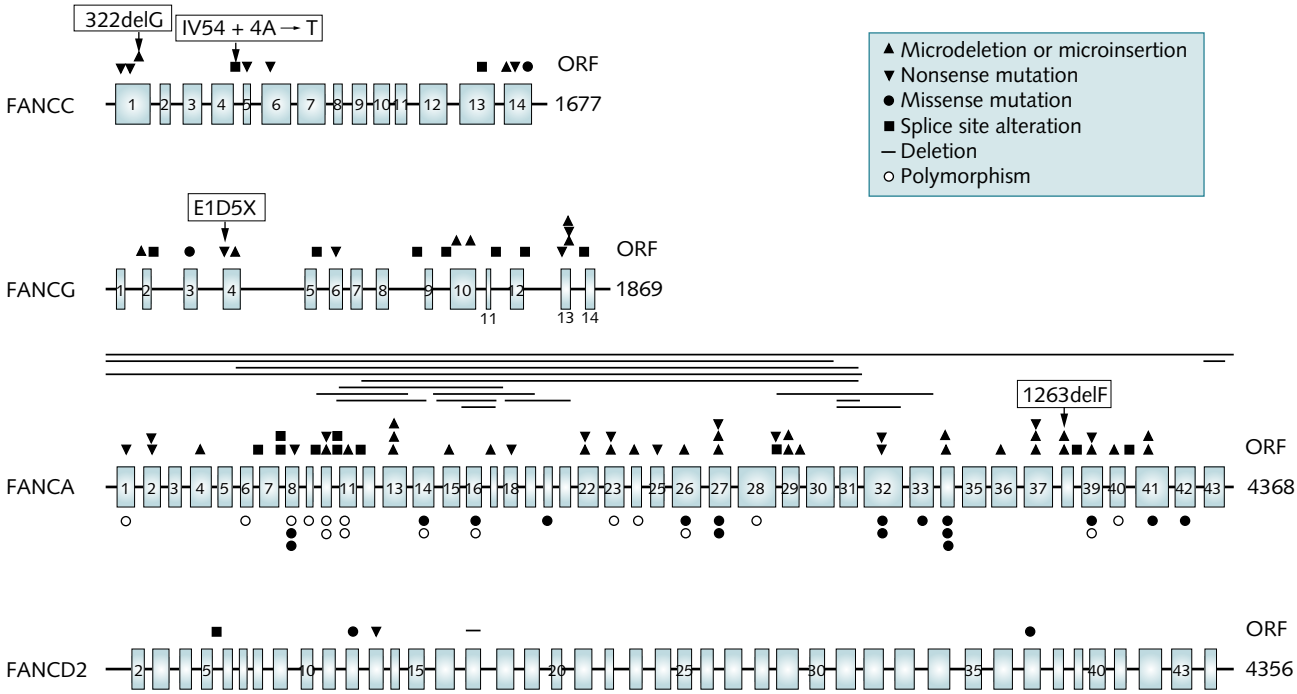


Fig 12.5 Molecular pathogenesis of FA: a proposed model

are spread throughout the gene. None of the mutant alleles is common and few have been encountered more than once. Thus, identifying mutations in newly diagnosed cases of FA is laborious, as one needs to scan the entire coding sequence of several genes, unless the complementation group is known.

Mutations in the *FAC* gene account for about 10–15% of FA cases. The IVS4+A→T and del322G mutations comprise >75% of *FANCC* mutations. The IVS4+A→T allele is found in Ashkenazi Jews at a polymorphic frequency (1 in 80), and it is responsible for 85% of the FA cases in this population. Patients with IVS4 or exon 14 mutations tend to have earlier onset of hematological complications (BMF and MDS/AML), and to have a shorter survival time compared with patients with exon 1 mutations or patients with non-*FANCC*-related FA. Mutations in the *FANCG* gene are found in 10% of FA cases. The stop codon mutation E105X accounts for 44% of mutations in German FA-G patients. In a handful of FA patients, mutations in the *FANCD2* gene and in *BRCA2* genes have been identified. *BRCA2* mutations affecting the carboxy-terminus of the protein have been found in patients of FA-B and FA-D1 groups; therefore *BRCA2* is the mutant gene for these complementation groups.

Cellular phenotype and function of the FA proteins. Chromosomal instability, the most striking cellular phenotypic feature in FA, is due to spontaneous development of chromosomal double-strand breaks during replication, a phenomenon accentuated by clastogenic agents at low concentrations. Chromosomal instability is a feature shared by many single-gene (ataxia–telengiectasia, Bloom syndrome, Werner syndrome) or multiple-gene (xeroderma pigmentosum, hereditary non-polyposis colorectal cancer, hereditary breast/ovary cancer) diseases. They all demonstrate a defect in pathways responsible for the maintenance of the integrity of the genome. As a result, a number of mutations with cell-transforming potential accumulate, leading eventually to neoplasia. The crucial biochemical evidence linking FA to the DNA repair pathways, in particular to those involving repair through homologous recombination, was unveiled recently. Mono-ubiquinated FANCD2 interacts in the so-called nuclear foci (these nuclear structures appear during DNA replication in response to DNA damage) with BRCA1 and BRCA2, two proteins with a central role in DNA damage repair pathways; as discussed above, *BRCA2* is an FA gene, mutated in FA-B and -D1 patients. In the currently evolving model of FA molecular pathophysiology, it is predicted that the FANCA, B, C, F, G and E proteins form a nuclear complex, the assembly of which commences in the cytoplasm (Figure 12.5). This core complex could function as the sensor of DNA damage that relays the signal for repair to the FANCD2/BRCA1 complex, possibly by activating FANCD2 through mono-ubiquitination. An important observation supporting the above is the observation that FANCD2 is not ubiquitinated in any of the FA A, B, C, F, G and E complementation groups. However, currently it is unclear whether the core complex itself provides the enzymatic activity for the ubiquitination of FANCD2. None of the proteins involved has a domain with recognized ubiquitin ligase activity; instead, BRCA1 contains a domain which is thought to provide this enzymatic function. BRCA2 is the downstream effector that leads to homologous recombination repair of the double-strand break through activation of the RAD51 complex; however it is also likely that BRCA2 acts upstream at the level of the FA protein core complex. Other biochemical defects reported in FA cells include increased sensitivity to reactive oxygen species, defects of the cell cycle and increased,

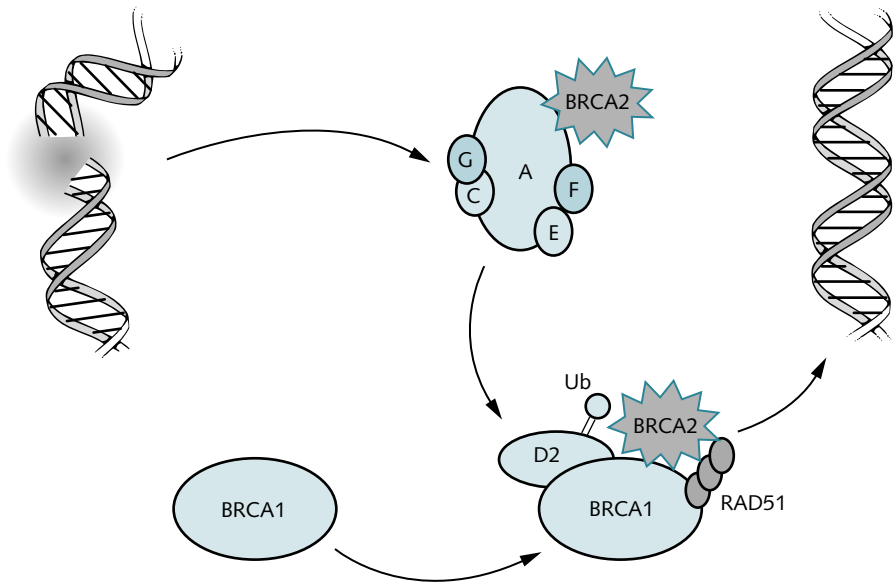


Fig. 12.6 The exon structure of the FA genes and their pathogenic mutations (see also text)
From Joenje H, Patel KJ. (2001) The emerging genetic and molecular basis of Fanconi anemia. *Nature Reviews in Genetics*, **2**, 446–459.

IFN-γ-mediated, susceptibility to apoptosis. How such diverse functional defects relate to each other and fit with the subcellular localization of the FA proteins remains to be elucidated.

Diagnosis. The clastogen test remains the gold standard clinical diagnostic test; however, complementation studies using retroviral vectors containing all seven genes have now entered clinical practice and this now allows rapid assignment of patients to a specific complementation group. This is followed by screening for pathogenic mutations in the corresponding gene. With this approach, genetic counseling and prenatal and preimplantation diagnosis are now feasible for most families with affected children. In areas with a large Ashkenazi Jewish population (e.g. New York City), screening for polymorphic FA-C alleles is feasible on a wider basis and can be offered to all couples at risk.

Treatment. The conventional management of FA focusses on the consequences of BMF and includes hematopoietic growth factors, blood product support and androgens. About half of the patients respond to androgens initially, but often suffer from significant side-effects, including androgen-induced hepatic adenomas. Eventually all patients become refractory.

HSCT from an HLA-identical, unaffected sibling or from alternative sources is currently the only therapeutic approach that can successfully achieve long-term correction of BMF and possible prevention of MDS and AML. Currently, clinical research focusses on the use of reduced-intensity, non-myeloablative conditioning HSCT regimens in order to reduce short- and long-term mortality. Unfortunately, even after HSCT the patient remains at increased risk of developing solid tumors.

The cloning of *FA* genes has opened the way to gene therapy for FA patients. There is clinical and experimental evidence that HSCs with corrected phenotype have a survival and growth advantage over uncorrected cells and can support long-term hemopoiesis. Successful transfer of FA genes to a small number of autologous HSCs should therefore be adequate to reduce the severity of BMF. Four FA-C patients underwent gene therapy in a pilot study using a retroviral vector. Although a transient improvement in the clonogenic capacity of the bone marrow progenitors was observed, this was not reflected in the clinical parameters. The use of lentivirus-based vectors holds greater promise and preclinical results are encouraging.

Dyskeratosis congenita

Dyskeratosis congenita (DC) is another rare, genetically heterogeneous, inherited disorder with BMF as a major feature. Lacy, reticulated skin, nail dystrophy, abnormal pigmentation, and mucosal leukoplakia in the first years of life and the later development of BMF classically define DC. Leukoplakia, particularly of the oral mucosa and also of the gastrointestinal and genital tracts, is common. Other clinical manifestations include developmental delay, short stature, ocular, dental and skeletal abnormalities, hyperhidrosis, hyperkeratinization of the palms and soles, bullae on minimal trauma, hair loss, sometimes gonadal failure, and features of premature ageing.

Molecular genetics and molecular pathogenesis. In the majority of cases (~85%) the mode of inheritance is X-linked recessive, and many cases are sporadic; in the remaining cases the mode of inheritance is either autosomal dominant or autosomal recessive.

The X-linked form of DC results mostly from missense mutations in the housekeeping gene *DKC1*. *DKC1* maps to Xq28, consists of 15 exons, its cDNA is 2465 bp long and it encodes a protein (dyskerin) of 514 amino acids with a predicted molecular weight of 57.6 kDa. Mutations in the *DKC1* gene account for ~40% of all cases. Hoyerall–Hreidarsson syndrome, a rare syndrome characterized by severe growth failure, immunodeficiency, cerebellar abnormalities and BMF, has been shown to be allelic to DC; indeed, in all cases tested mutations in *DKC1* were identified.

Dyskerin is homologous to Cbf5p, a yeast protein that, in conjunction with the H/ACA class of small nucleolar RNAs, is involved in the pseudo-uridylation of preribosomal RNA (rRNA). However, one study addressing this issue failed to demonstrate any qualitative or quantitative defect of rRNA in DC patients. Instead, it was shown that dyskerin binds to a H/ACA domain of the human telomerase RNA (hTR); that is, the template used by the reverse transcriptase component of telomerase (TERT) for telomere lengthening. Telomerase is a ribonucleoprotein complex and dyskerin is among a number of proteins interacting, directly or indirectly, with hTR, ensuring its stability and ability to function as a template for TERT. Telomerase activity is highest in rapidly dividing cells (e.g. dividing HSCs), during organogenesis, in early life and also in tumor cells. The critical evidence that DC is a disease of telomere dysfunction was provided by the finding of monoallelic mutations in the 451-nucelotide, intronless hTR gene in some families with the autosomal dominant form of DC (Figure 12.7). Thus, the current evidence links the X-linked and autosomal dominant types of DC in a common molecular pathogenetic pathway: namely, the maintenance of telomere length. In keeping with this notion, all types of blood cells in DC patients have significantly shorter telomeres than age-matched controls. Shortening of telomeres to a critical length can result either in growth arrest and apoptosis, hence the development of BMF, or, on some occasions, to neoplastic transformation; hence the increased propensity to epithelial tumors.

As for FA, there is no definitive treatment for DC and most patients die before the age of 30. The main causes of death are complications of BMF, opportunistic infections, pulmonary complications and malignancy. Few patients have had successful correction of BMF by HSCT.

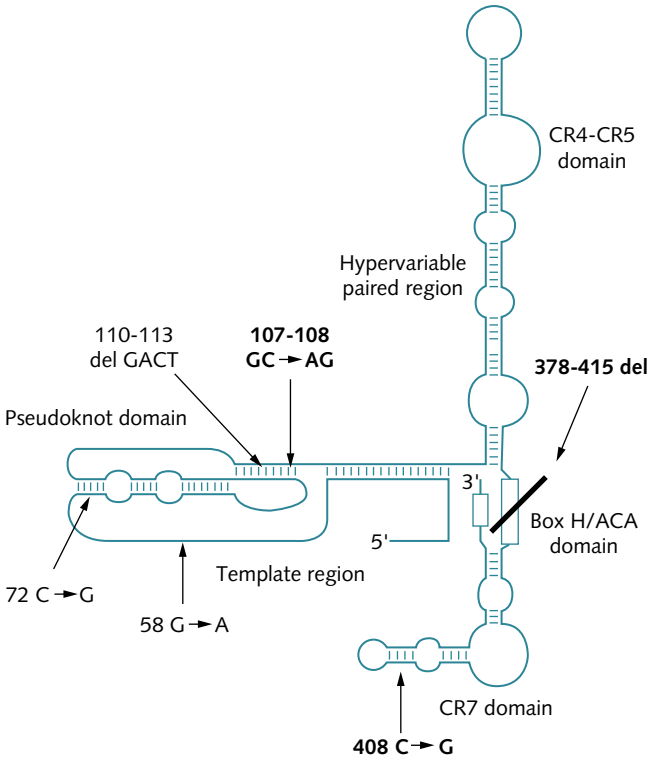


Fig 12.7 Model of secondary structure of the human hTR
Mutations identified in families with autosomal dominant dyskeratosis congenita are shown (courtesy of Dr T. Vulliamy).

Blackfan–Diamond anemia

Blackfan–Diamond anemia (DBA) is a rare inherited syndrome with isolated anemia and erythroblastopenia as the main features. Clinically it manifests itself in the first months of life with macrocytic anemia, reticulocytopenia, increased Hb F and elevated erythrocyte adenosine deaminase levels; occasional patients have thrombocythemia. The bone marrow is normocellular with characteristic absence of erythroblasts (<5% of nucleated cells). Growth retardation and other developmental abnormalities (including skeletal, cardiac and urogenital defects) are seen in about a third of patients. There is also a small risk (~2%) of transformation to MDS/AML.

Molecular genetics and pathogenesis. Most cases of DBA are sporadic but in multiplex families both autosomal dominant and autosomal recessive patterns of inheritance have been observed. There is a considerable degree of penetrance and phenotypic variability even in the same families.

In about 25% of cases, mutations in the ribosomal protein S19 gene (*RPS19*) have been identified in familial as well as in sporadic cases. The patients are heterozygous for these mutations, which would be expected to reduce the level of protein by about 50%, in keeping with haploinsufficiency. It is unclear at the moment how haploinsufficiency of *RPS19* is implicated in the molecular pathogenesis of DBA, and particularly why it should have a selective effect on erythropoiesis. It is noteworthy that mutations of ribosomal proteins in *Drosophila melanogaster* lead to the so-called Minute phenotype, comprising delayed larval development, small size, and reduced viability and fertility. Interestingly, haploinsufficiency of some ribosomal proteins may be associated with additional phenotypic features; for example, mutated *RPS6* leads to hypertrophy of lymphohemopoietic organs. It is possible, therefore, that haploinsufficiency of the ribosomal function of RPS19 accounts for the somatic developmental defects, whereas an as-yet unidentified extraribosomal function accounts for the defect in erythropoiesis. Recent evidence suggests that RPS19 expression decreases during normal erythroid maturation, implying a role of RPS19 in early erythropoiesis. In another study, overexpression of RPS19 in $CD34^+$ cells from DBA patients improved the erythroid clonogenic activity but did not restore it to normal.

Blood transfusion for moderate to severe anemia is the mainstay of treatment for DBA. Over 50% of patients respond to steroids and become independent of blood transfusion. Some sustain remission after steroid withdrawal but most either become steroid-resistant or require an unacceptably high maintenance dose. Such patients may benefit from a sibling allogeneic HSCT. The median long-term survival in a small series from the North American DBA registry was 87%; by contrast, the survival of patients receiving HSCT from an alternative donor was only 14%.

Red cell enzyme deficiencies

Inherited abnormalities of red cell enzymes (red cell enzymopathies) are a distinct set of genetic disorders with one important clinical manifestation in common: chronic hemolytic anemia. Most of the enzymes involved are housekeeping enzymes present by definition in all cells. Therefore, one might expect that a severe reduction in activity of any of these might have generalized clinical manifestations. However, we can identify at least two reasons why red cells are more severely affected: firstly, red cells have a much more limited metabolic machinery than most other somatic cells; if a particular enzyme is deficient, other cells may cope by the use of alternate or surrogate metabolic pathways. Secondly, mature red cells are not competent for protein synthesis: therefore, if a particular enzyme is made highly unstable by a mutation, other cells can compensate by increased enzyme synthesis, but red cells cannot. Nevertheless, the fact that enzymopathies are not purely red cell disorders is highlighted by the coexistence, in some cases, of clinical manifestations in other systems, particularly the muscles and the nervous system; indeed, in some enzymopathies neurological damage may dominate the clinical picture.

This section deals with those enzymopathies affecting red cell metabolism for which the molecular basis has been elucidated. We will not discuss conditions in which an enzyme abnormality is expressed also in red cells but the main clinical manifestations are elsewhere (such as the porphyrias, galactosemia, and Lesch–Nyhan syndrome). For the sake of brevity, most of the enzymopathies will be discussed in groups.

Enzymopathies of glycolysis

Clinical features

All of these defects are rare to very rare (Table 12.6) and all cause hemolytic anemia with varying degrees of severity. It is not unusual for the presentation to be in the guise of severe neonatal jaundice that may require exchange transfusion; if the anemia is less severe it may present later in life, or it may even remain asymptomatic and be detected incidentally when a blood count is done for unrelated reasons. The spleen is often enlarged. When other systemic manifestations occur, they involve the central nervous system, sometimes entailing severe mental retardation, or the neuromuscular system, or both.

Diagnosis

The diagnosis of hemolytic anemia is usually not difficult, thanks to the triad of normomacrocytic anemia, reticulocytosis and hyperbilirubinemia. Enzymopathies should be considered in the differential diagnosis of any chronic Coombs-negative hemolytic anemia. In most cases of glycolytic enzymopathies, the morphological abnormalities of red cells characteristically seen in membrane disorders are conspicuous by their absence. A definitive diagnosis can be made only by demonstrating the deficiency of an individual enzyme by a quantitative assay. For the sake of economy, it is sensible to carry out these rather laborious tests in order of frequency of occurrence of the various enzymopathies [e.g. first pyruvate kinase (PK), then GPI, etc.; Table 12.7]. If a particular molecular abnormality is already known in the family, of course one could test directly for that at the DNA level, bypassing the need for enzyme assays.

Molecular pathophysiology

Since the main physiological significance of the glycolytic pathway (*see Figure 12.8 for an overview of glycolysis*) in the red cell is to produce chemical energy in the form of ATP, the main consequence of any glycolytic enzymopathy is a shortage of energy supply. Since glycolytic enzymes are apparently present in cells in considerable excess, the 50% residual enzyme activity seen in heterozygotes does not become rate-limiting; thus, heterozygotes do not have hemolytic anemia, and that is why these enzymopathies show a recessive pattern of inheritance. As seen in Table 12.7, most of the mutations so far identified in the genes encoding glycolytic enzymes are of the missense type, causing single amino acid replacements. This is important, because the low level of residual enzyme activity can still support some metabolic flow through the glycolytic pathway, and helps explain how red cells survive in circulation, even though their lifespan is reduced. With respect to the precise reason why enzyme activity is reduced, we must consider at the protein level two basic mechanisms. (1) In the majority of cases loss of activity is probably due to a decreased stability of the protein. In such cases we would predict that other cells might be much less affected than red cells, because the former can compensate for decreased stability through increased synthesis of the enzyme. (2) In some cases the amino acid replacement may affect the active center of the enzyme, which in turn may affect either substrate binding (K_m) or the catalytic rate of the enzyme (K_{cat}), or both: in this case other cells in which the rate of glycolysis is critical will be affected, as well as red cells.

Management

There is no specific treatment for these conditions. Patients with moderate anemia may require occasional blood transfusion when they experience exacerbations of the anemia due to an increased rate of hemolysis, or to decreased red cell production secondary to infection (the most extreme example being aplastic crisis from parvovirus infection). Patients with chronic severe anemia may require regular blood transfusion therapy with associated iron chelation. In some patients splenectomy has been beneficial. In severe cases bone marrow transplantation would be a rational form of treatment (for patients who have a suitable donor), provided there are no systemic manifestations other than hemolytic anemia, and provided it is carried out before there is organ damage (e.g. from iron overload).

Glucose-6-phosphate dehydrogenase deficiency

Epidemiology

Glucose-6-phosphate dehydrogenase (G6PD) deficiency is distributed worldwide with a high prevalence in populations of Africa, Southern Europe, the Middle East, Southeast Asia and parts of Oceania, as well as in areas to which migrations from these areas have taken place. The overall geographic distribution of G6PD deficiency and its heterogeneity, together with clinical field studies and *in vitro* culture experiments, strongly support the view that this common genetic trait has been selected by *Plasmodium falciparum* malaria, by virtue of

Table 12.7 Synopsis of red cell enzymopathies*.

Enzyme (abbreviation)	Isoenzyme[a] characteristic of red cells	Prevalence of enzyme deficiency	Main clinical features associated with enzyme deficiency[b]	Benefit from splenectomy[c]	Chromosomal localization
Hexokinase (HK)	R (I)	Very rare	HA	Partial	10q22
Glucose 6-phosphate isomerase (GPI)		Rare	HA, NM, CNS	Partial	19q13.1
Phosphofructokinase (PFK)[g]	M	Very rare	HA, myopathy		12q13
	L				21q22.3
Aldolase	A	Very rare	HA, myopathy		16q22-24
Triosephosphate isomerase (TPI)		Very rare	HA, CNS, NM,	None	12p13
Glyceraldehyde 3-phosphate dehydrogenase (GAPD)[h]		Very rare	HA		12p13.31-p13.1
Diphosphoglycerate mutase (DPGM)		Very rare	Polycythemia		7q31-q34
Phosphoglycerate kinase (PGK)	1	Very rare	HA, CNS, NM	Partial	Xq13
Monophosphoglycerate mutase (PGAM-B)	B				10q25.3
Enolase[h]	1 (α)	Very rare	HA		1pter-p36.13
Pyruvate kinase (PK)	R[l]	Rare	HA	Partial	1q21
Glucose 6-phosphate dehydrogenase (G6PD)	B	Common	HA	None	Xq28
Cytochrome B5 reductase		Rare	pseudo-cyanosis, CNS		22q13.31-qter
Adenylate kinase (AK)	1	Very rare	HA, CNS	Partial	9q34.1
γ-Glutamylcysteine synthetase (GLCLC)[l]		Very rare	HA, CNS(?)		6p12
Glutathione synthetase (GSS)		Very rare	HA, CNS		20q11.2
Glutathione peroxidase (GSH-Px)		Very rare[m]	?[m]		3q11-q12
Pyrimidine 5′ nucleotidase (P5′N1)		Rare	HA	Partial	7p15-p14

			Number of known mutations[e]					
					Deletion–insertion			
Number of exons	Number of amino acids[d]	5′-UTR	Missense	Nonsense	In frame	With frameshift	Affecting splicing	Total
19	916 (917)[f]		1		3–0			4
18	558		16	2			2	20
24	780		7	1		1–0	6	15
22	784							
12	364		2					2
7	249	1	9	2		1–0		13
9	335							

Continued

Table 12.7 *Continued*

3	259		1		1–0			2
11	417		8		1–0		2	11
	254							
	434							
12	574[i]	2	90	9	3–3	7–6	12	132
13	515		122[j]	1	6–0		1	130
9	276[k]		18	5	3–0		5	31
7	194		2	1				3
16	637		3					3
12	474		14		1–0	1–0	1	17
	201							
10	286[n]		3	1		0–1	2	7

*We have listed in the table all enzymes in the intermediary metabolism of red cells for which, to the best of our knowledge, the corresponding cDNA/gene has been cloned. Modified from Luzzatto L. and Notaro R. (1998) Red cell enzymopathies. In: Jameson JL (ed.) *Principles of Molecular Medicine.* Clifton, NJ: Humana Press, pp. 197–207.

[a]No entry in this column means that there are no known isoenzymes; therefore, it is assumed that the same enzyme type is present in all tissues.

[b]CNS, central nervous system involvement; HA, hemolytic anemia; NM, neuromuscular manifestations.

[c]Data available only for some patients.

[d]Including N-terminal methionine, which is cleaved off in most or all cases.

[e]Each individual molecular change, if observed in more than one patient, has been counted only once.

[f]The *HK-I* gene encodes both the erythroid-specific (HK-R) and ubiquitous (HK-I) isoforms. HK-R transcription starts from the erythroid-specific promoter upstream exon I (exon II is missing in the erythroid HK-R mRNA). HK-I transcription starts from the ubiquitously expressed promoter upstream exon II.

[g]PFK in normal red cells consists of a mixture of the five tetrameric species that can be formed from random association of the M (muscle) and L (liver) highly homologous subunits (M4, M3L, M2L2, ML3 and L4).

[h]Since no mutations have yet been reported, there is no formal proof that HA associated with this enzyme deficiency is due to mutation of the corresponding gene.

[i]The red cell form of PK, called R, is produced by the gene encoding the L (liver) subunit. Because a different promoter is used, the size of liver PK is 543 amino acids.

[j]The 122 missense mutations include two variants with normal activity, A and São Borja. Six variants have two missense mutations each and one variant has three different missense mutations.

[k]The cytoplasmic form of this enzyme, present in red cells, differs from the microsomal form present in other cells because, as a result of an alternative splicing pathway, it lacks the first 25 N-terminal amino acids. Therefore, in other cells the size of the enzyme is 301 amino acids.

[l]γ-Glutamylcysteine synthetase consists of two subunits: a catalytic subunit and a regulatory subunit. Data for the catalytic subunit are shown here.

[m]There is no clear evidence that inherited deficiency of glutathione peroxidase exists.

[n]Two alternatively spliced forms are present in reticulocytes. The main protein is the long 286 amino acid protein (the third exon is spliced out); a longer protein contains 11 extra amino acids at its C-terminus.

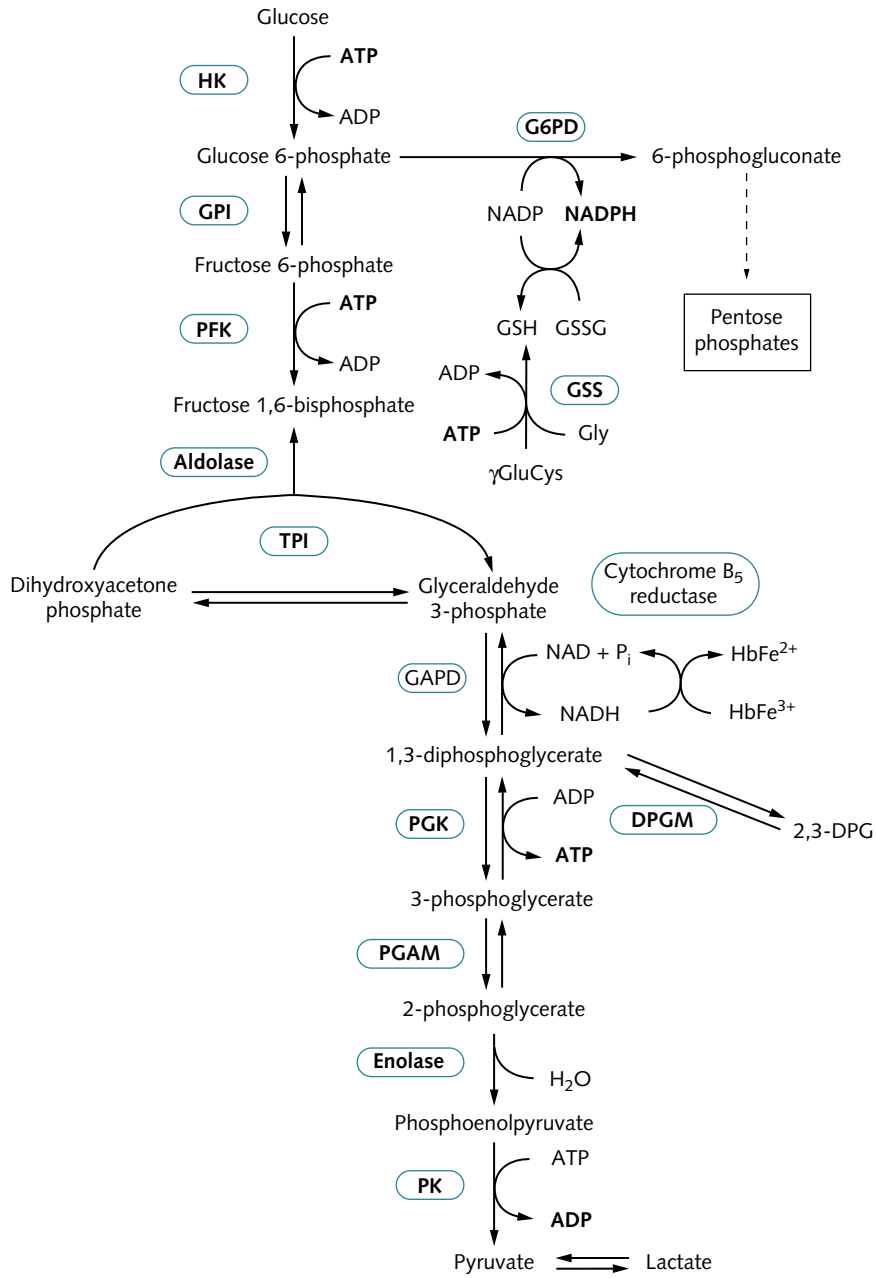


Fig. 12.8 Intermediary metabolism in red cells
The diagram shows the glycolytic pathway and related reactions (not the complete metabolic machinery of the red cells). Enzymes are enclosed in rounded boxes. Abbreviations as in Table 12.7. Additional abbreviations: DPG, diphosphoglycerate; GSH, reduced glutathione; GSSG, glutathione; $HbFe^{2+}$, hemoglobin; $HbFe^{3+}$, methemoglobin; γGluCys, γ-glutamylcysteine. From Luzzatto L, Notaro R. (1998) Red cell enzymopathies. In: Jameson JL (ed.). *Principles of Molecular Medicine*. Clifton, NJ: Humana Press.

the fact that it confers relative resistance against this highly lethal infection.

Clinical features

Three types of clinical presentations are well characterized. (1) The vast majority of G6PD-deficient people are asymptomatic most of the time, but they are at risk of developing acute hemolytic anemia, which may be triggered by drugs, infections or fava beans. (2) The risk of developing neonatal jaundice is much greater in G6PD-deficient than in G6PD-normal newborns. This is of great public health importance, because untreated severe neonatal jaundice can lead to permanent neurological damage. (3) In contrast to the first two presentations, chronic non-spherocytic hemolytic anemia (CNSHA) is very rare. The clinical picture is rather similar to CNSHA associated with glycolytic enzymopathies (*see above*), and again it is of variable severity. However, the hemolysis

is characteristically exacerbated by the same agents that can cause acute hemolytic anemia in people with the ordinary type of G6PD deficiency.

Diagnosis

The anemia is usually normocytic and normochromic, and its severity ranges from moderate to extreme. Acute hemolytic anemia is due largely to intravascular hemolysis, and hence is associated with hemoglobinemia and hemoglobinuria. The blood film may show spectacular evidence of hemolysis in the guise of anisocytosis, polychromasia, spherocytes, bite cells, blister cells and hemighosts. Supravital staining reveals the presence of Heinz bodies, consisting of precipitates of denatured hemoglobin. In CNSHA the morphology is less characteristic. The final diagnosis must rely on the direct demonstration of decreased activity of G6PD in red cells by an appropriate enzyme assay.

Genetic basis

G6PD is a homodimeric molecule, and its single subunit is encoded by an X-linked gene. As a result of the phenomenon of X-chromosome inactivation in somatic cells, female heterozygotes are genetic mosaics in whom approximately one-half of the red cells are normal and approximately one-half are G6PD-deficient. However, in some cases the ratio is imbalanced. Therefore, clinical manifestations, such as favism, can occur in both hemizygous males and heterozygous females, but they tend to be milder in the latter, roughly in proportion to the fraction of red cells that are G6PD-deficient.

Function of G6PD

In intermediary metabolism G6PD is aptly depicted as the first step in the pentose phosphate pathway. However, several lines of evidence indicate that its most essential role is not to produce pentose, but rather to produce reductive potential in the form of NADPH. Recently, G6PD-null mouse embryos have been obtained by targeted inactivation of G6PD in embryonic stem cells. From a detailed analysis of hemizygous mutant embryos, which die by day 10.5, it was inferred that the cause of death is precisely the onset of aerobic metabolism. Interestingly, heterozygous embryos also die, somewhat later, only if their G6PD-null gene is of maternal origin: in this case the cause of death is a defective placenta, as a consequence of the selective inactivation of the paternal X chromosome in extra-embryonic tissues.

Molecular pathophysiology

Acute hemolytic anemia is seen with variants of G6PD where-by red cells retain some 10% of the normal G6PD activity, resulting in a limited capacity of these cells to withstand the oxidative action of an exogenous factor (oxidative hemolysis). By contrast, with other variants, the steady-state level of G6PD is so low that it becomes limiting for red cell survival, even in the absence of any oxidant challenge; the result is CNSHA. Numerous point mutations in the *G6PD* gene causing CNSHA have been identified (Figure 12.9). Although we cannot explain the reason for a severe clinical phenotype in every case, a cluster of mutations causing CNSHA in exons 10 and 11 corresponds closely to the region of the molecule where the two subunits interface. It is not surprising that amino acid replacements in this region will interfere with dimer formation or will cause marked instability of the dimer.

Management

The commonest manifestations of G6PD deficiency, neonatal jaundice and acute hemolytic anemia, are largely preventable or controllable by screening, surveillance and avoidance of triggering factors, particularly fava beans, by G6PD-deficient subjects. When a patient presents with acute hemolytic anemia, and once the cause is diagnosed, no specific treatment may be needed if the episode is mild. At the other end of the spectrum, and especially in children, acute hemolytic anemia may be a medical emergency requiring immediate blood transfusion. The management of neonatal jaundice does not differ from that of neonatal jaundice due to causes other than G6PD deficiency and, in order to prevent neurological damage, treatment with phototherapy and/or exchange blood transfusion may be required. The management of CNSHA is similar to that of CNSHA due to glycolytic enzymopathies, but in addition it is important to avoid exposure to potentially hemolytic drugs. Again, although there is no evidence of selective red cell destruction in the spleen (as seen in hereditary spherocytosis), splenectomy has proven beneficial in severe cases.

Recently, life-long expression of human G6PD at therapeutic levels has been obtained in red blood cells and in white blood cells of mice through retroviral-mediated transfer into hematopoietic stem cells.

5′ Nucleotidase deficiency

This enzyme deficiency, known for some 30 years, is of interest for several reasons. Firstly, it is probably the third most common red cell enzymopathy (trailing G6PD deficiency and PK deficiency). Secondly, the diagnosis can be suspected from red cell morphology because it is associated with basophilic stippling (accounted for by persistence of RNA in mature red cells). Thirdly, although we do not really understand the precise mechanism, 5′ nucleotidase (5′N) deficiency is a

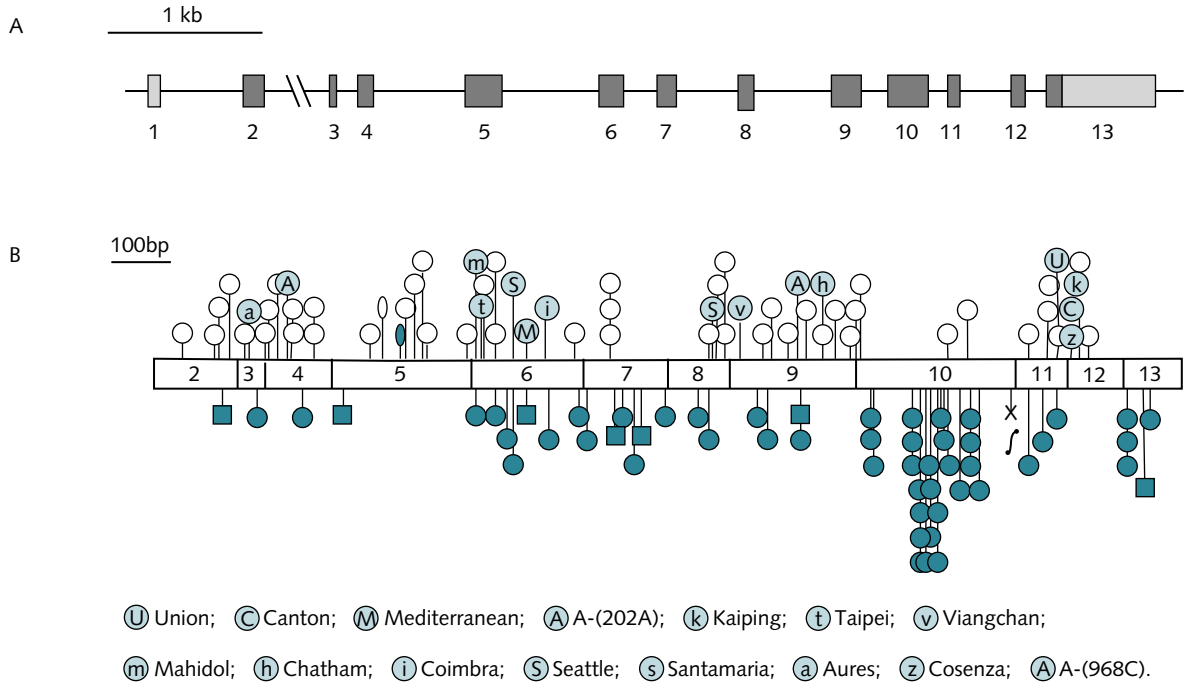


Fig. 12.9 Distribution of mutations along the human G6PD gene
(a) Genomic structure of the human G6PD gene. Exons are shown as numbered rectangles (black rectangles represent coding sequences; shaded rectangles represent non-coding sequences). (b) The locations of amino acid substitutions are shown along the coding sequence of the gene, in which the exons are shown as open boxes. Substitutions giving rise to the more severe (class I) variants are shown as filled circles below. Small deletions, a nonsense mutation, and a splice site mutation are shown as filled rectangle, a cross and a ∫, respectively. The milder class II and class III variants are shown as open circles above: polymorphic variants are shown as a letter in a colored circle . Class IV variants are shown as open ellipses. From Luzzatto L, Mehta A, Vulliamy T. (2001) Glucose-6-phosphate dehydrogenase deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds). *The Metabolic and Molecular Basis of Inherited Disease*, 8th edn. New York: McGraw-Hill, pp. 4517–4553.

good example of how the red cell has virtually only one way to manifest its suffering: thus, almost any metabolic abnormality will lead to accelerated destruction; that is, hemolysis. Fourthly, the anemia of chronic lead poisoning (which had been known for a long time to be associated with basophilic stippling) turns out to be the consequence of the fact that lead is a powerful inhibitor of 5′N; thus, in terms of its hematological effects, lead poisoning is a phenocopy of 5′N deficiency.

In 2001 a gene corresponding to a previously published 5′N cDNA was mapped to chromosome 7, and three different mutations in this gene were discovered in homozygosity in four subjects with 5′N deficiency. One was a missense mutation, one was a nonsense mutation, and one was a splicing mutation causing the loss of exon 9 from the mature mRNA. This important advance makes it now possible to carry out a molecular diagnosis of 5′N deficiency. Moreover, it will be possible to explore the implications of heterozygous 5′N deficiency, which is of special interest in view of previous suggestions that it may interact with Hb E disease to make its clinical expression more severe.

Conclusions

Hematologists know only too well that anemia is not a diagnosis but is the recognition of a sign for which we must find the cause and work out the pathogenetic mechanism, which ultimately must be explainable at the molecular level. Overall, for the majority of acquired anemias, things are by now pretty clear with respect to etiology but not necessarily with respect to molecular mechanisms. For instance, although we do understand that iron deficiency limits heme synthesis and consequently hemoglobin synthesis, we do not know exactly how the mean cell volume (MCV) is controlled by the supply of iron: for this reason we have omitted from this chapter a section on the commonest anemia of all—iron deficiency anemia. However, it is gratifying that we can at least offer a model for the molecular basis of megaloblastic anemia, another major public health problem in many countries. With respect to the majority of congenital anemias, and certainly for the commonest among them, molecular genetics has answered not all, but most of the questions with respect to etiology and pathogenesis.

However, lest we become complacent, we must admit that, as is often the case in medicine, the time lag between fundamental discoveries and therapeutic applications too often remains too long, and that advances in knowledge are not always used wisely. On the one hand, we saw recently a patient with PNH who had massive hemoglobinuria and had become severely iron-deficient as a result. He had been treated with erythropoietin, but not with iron: a few weeks of 400 mg of ferrous sulfate daily raised his hemoglobin level from 6.7 to 10.1 g/dl. On the other hand, for patients who have, for instance, severe chronic hemolytic disease due to PK deficiency, the outlook has improved only because we can offer better supportive treatment—not because we know the molecular basis. The latter will only become relevant once we learn to correct the PK deficiency by gene addition or by gene replacement. This is a major and worthy challenge for the next decade with respect to all the genetically determined anemias.

Further reading

Megaloblastic anemia and congenital dyserythropoietic anemias

Blount BC, Mack MM, Wehr CM *et al.* (1997) Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 3290–3295.

Chanarin I. (1979) *The Megaloblastic Anaemias*, 2nd edn. Oxford: Blackwell Scientific Publications.

Chui D, Oh-Eda M, Liao YF *et al.* (1997) Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell*, **90**, 157–167.

Dgany O, Avidan N, Delaunay J *et al.* (2002) Congenital dyserythropoietic anemia type I is caused by mutations in codanin-1. *American Journal of Human Genetics*, **71**, 1467–1474.

Fukuda MN. (1990) HEMPAS disease: genetic defect of glycosylation. *Glycobiology*, **1**, 9–15.

Ingram CF, Davidoff AN, Marais E *et al.* (1997) Evaluation of DNA analysis for evidence of apoptosis in megaloblastic anaemia. *British Journal of Haematology*, **96**, 576–583.

Luzzatto L, Falusi AO, Joju EA. (1981) Uracil in DNA in megaloblastic anemia. *New England Journal of Medicine*, **305**, 1156.

Savage DG, Lindenbaum J. (1995) Folate–cobalamin interactions. In: Bailey L (ed.) *Folate in Health and Disease*. New York: Marcel Dekker, pp. 237–285.

Wickramasinghe SN. (1997) Dyserythropoiesis and congenital dyserythropoietic anaemias. *British Journal of Haematology*, **98**, 785–797.

Wickramasinghe SN, Fida S. (1994) Bone marrow cells from vitamin B12- and folate-deficient patients misincorporate uracil into DNA. *Blood*, **83**, 1656–1661.

The sideroblastic anemias

Sideroblastic anemia

Bekri S, Kispal G, Lange H *et al.* (2000) Human ABC7 transporter: gene structure and mutation causing X-linked sideroblastic anemia with ataxia with disruption of cytosolic iron-sulfur protein maturation. *Blood*, **96**, 3256–3264.

Bottomley SS, May BK, Cox TC *et al.* (1995) Molecular defects of erythroid 5-aminolevulinate synthase in X-linked sideroblastic anemia. *Journal of Bioenergetics and Biomembranes*, **27**, 161–168.

Cox TC, Bawden MJ, Martin A *et al.* (1991) Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron responsive element in the mRNA. *EMBO Journal*, **10**, 1891–1902.

May A, Bishop DF. (1998) The molecular biology and pyridoxine responsiveness of X-linked sideroblastic anaemia. *Haematologica*, **83**, 56–70.

Pearson marrow-pancreas syndrome

Pearson HA, Lobel GS, Kocoshis SA *et al.* (1979) A new syndrome of refractory sideroblastic anemia with vacuolization of marrow precursors and exocrine pancreatic dysfunction. *Journal of Pediatrics*, **6**, 976–984.

Rotig A, Bourgeron T, Chretien D *et al.* (1995) Spectrum of mitochondrial DNA rearrangements in the Pearson marrow-pancreas syndrome. *Human Molecular Genetics*, **4**, 1327–1330.

Bone marrow failure syndromes

Aplastic anemia

Young NS. (2002) Acquired aplastic anemia. *Annals of Internal Medicine*, **136**, 534–546.

Young NS, Barrett AJ. (1995) The treatment of severe acquired aplastic anemia. *Blood*, **85**, 3367–3377.

Young NS, Maciejewski J. (1997) The pathophysiology of acquired aplastic anemia. *New England Journal of Medicine*, **336**, 1365–1372.

Paroxysmal nocturnal hemoglobinuria

Araten DJ, Nafa K, Pakdeesuwan K *et al.* (1999) Clonal populations of hematopoietic cells with paroxysmal nocturnal hemoglobinuria genotype and phenotype are present in normal individuals. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 5209–5214.

Karadimitris A, Luzzatto L. (2001) The cellular pathogenesis of paroxysmal nocturnal hemoglobinuria. *Leukemia*, **15**, 1148–1452.

Luzzatto L, Bessler M. (1996) The dual pathogenesis of paroxysmal nocturnal hemoglobinuria. *Seminars in Hematology*, **3**, 101–110.

Takeda J, Miyata T, Kawagoe K *et al.* (1993) Deficiency of the GPI anchor caused by a somatic mutation of the PIG-A gene in paroxysmal nocturnal hemoglobinuria. *Cell*, **73**, 703–711.

Fanconi's anemia

Alter B, Young NS. (1998) The bone marrow failure syndromes. In: Nathan DG, Orkin SH (eds). *Nathan and Oski's Hematology of Infancy and Childhood*, 5th edn. Philadelphia: W.B. Saunders, pp. 237–335.

Fanconi Anaemia/Breast Cancer Consortium. (1996) Positional cloning of the Fanconi anaemia group A gene. *Nature Genetics*, **14**, 324–328.

Foe JRLT, Roimans MA, Bosnoyan-Collins L *et al.* (1996) Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. *Nature Genetics*, **14**, 320–323.

Howlett NG, Taniguchi T, Olson S *et al.* (2002) Biallelic inactivation of BRCA2 in Fanconi anemia. *Science*, **26**, 606–609.

Joenje H, Patel KJ. (2001) The emerging genetic and molecular basis of Fanconi anaemia. *Nature Reviews in Genetics*, **2**, 446–459.

Levran O, Erlich T, Magdalena N *et al.* (1997) Sequence variation in the Fanconi anemia gene FAA. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 13051–13056.

Strathdee CA, Gavish H, Shannon WR *et al.* (1992) Cloning of cDNAs for Fanconi's anaemia by functional complementation. *Nature*, **356**, 763–767.

Dyskeratosis congenita

Collins K, Mitchell JR. (2002) Telomerase in the human organism. *Oncogene*, **21**, 564–579.

Heiss NS, Knight SW, Vulliamy JT *et al.* (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature Genetics*, **19**, 32–38.

Mitchell JR, Wood E, Collins K. (1999) A telomerase component is defective in the human disease dyskeratosis congenita. *Nature*, **402**, 551–555.

Vulliamy T, Marrone A, Goldman F *et al.* (2001) The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature*, **413**, 432–435.

Blackfan–Diamond anemia

Alter B, Young NS. (1998) The bone marrow failure syndromes. In: Nathan DG, Orkin SH (eds). *Nathan and Oski's Hematology of Infancy and Childhood*, 5th edn. Philadelphia: W.B. Saunders, pp. 237–335.

Draptchinskaia N, Gustavsson P, Andersson B *et al.* (1999) The gene encoding ribosomal protein S19 is mutated in Diamond–Blackfan anaemia. *Nature Genetics*, **21**, 169–175.

Willig T-N, Draptchinskaia N, Dianzani I *et al.* (1999) Mutations in ribosomal protein S19 gene and Diamond Blackfan anemia: wide variations in phenotypic expression. *Blood*, **94**, 4294–4306.

Red cell enzyme deficiencies

Beutler E. (1994) Glucose-6-phosphate dehydrogenase deficiency. *New England Journal of Medicine*, **324**, 169–174.

Hirono A, Kanno H, Miwa S *et al.* (2001) Pyruvate deficiency and other enzymopathies of the erythrocyte. In: Scriver CR, Beaudet AL, Sly WS *et al.* (eds). *The Metabolic and Molecular Basis of Inherited Disease*, 8th edn. New York: McGraw-Hill, pp. 4637–4664.

Luzzatto L, Mehta A, Vulliamy T. (2001) Glucose-6-phosphate dehydrogenase deficiency. In: Scriver CR, Beaudet AL, Sly WS *et al.* (eds). *The Metabolic and Molecular Basis of Inherited Disease*, 8th edn. New York: McGraw-Hill, pp. 4517–4553.

Mentzer WC. (1998) Pyruvate kinase deficiency and disorders of glycolysis. In: Nathan DG, Orkin SH (eds). *Nathan and Oski's Hematology of Infancy and Childhood*, 5th edn. Philadelphia: W.B. Saunders, pp. 665–703.

Chapter 13 The molecular basis of iron metabolism

Nancy C Andrews

Introduction

Iron is one of several metals that are essential for normal cellular processes. Its primary role in mammalian biology is to bind oxygen in hemoglobin and myoglobin. It is also involved in the enzymatic transfer of electrons by cytochromes, peroxidases, ribonucleotide reductases and catalases. However, the same properties that make iron useful for these functions can also lead to cellular damage when iron is present in excess. Normally, proteins constrain the activity of iron, but when their iron-binding capacity is exceeded, free iron promotes the formation of reactive oxygen species that attack cellular lipids, proteins and nucleic acids. Thus, iron balance must be carefully maintained to avoid the deleterious effects of iron deficiency and iron overload. All known disorders of iron metabolism can be considered abnormalities of iron balance.

There is no physiological excretion mechanism for iron: iron losses result only from bleeding and exfoliation of skin and mucosal cells. Under normal conditions, iron enters the body exclusively by dietary absorption, and absorption is meticulously regulated to balance the small losses. Iron balance is disrupted when intake and losses are not matched. Iron deficiency occurs when the dietary iron supply is inadequate, when losses are increased (primarily because of bleeding) or when both of these circumstances are present. Iron overload results when iron absorption is inappropriately increased due to genetic defects in iron regulatory proteins, or when repeated blood transfusions create a substantial iron burden.

Our understanding of the molecular processes of iron metabolism has advanced considerably over the last decade, as new techniques in genetics and molecular biology have been applied to problems in this field. Much of what we have learned has come from the study of animals with spontaneous and induced mutations in genes important for the transport and storage of iron. It is widely assumed that information gleaned from experiments in rodents can be directly extrapolated to humans. It is clear that iron metabolism is very similar among mammalian species, validating this approach.

Mechanisms of iron transport

General principles

Iron is a large, charged ion that cannot freely diffuse across cellular membranes. Transmembrane transfer requires specific carrier proteins. There are two general ways in which cells transport iron. Some cells, such as intestinal epithelial cells, hepatocytes and macrophages, are equipped both to take in (import) iron and to release (export) it. These cell types are involved in the acquisition, storage and mobilization of iron. Other cells, particularly erythroid precursors, import iron but do not release it unless the cells are destroyed.

Approximately 25 mg of iron are needed every day to support hemoglobin production in maturing erythrocytes. This amount is much greater than the 1–2 mg entering the body each day through the intestine. Obviously, the iron for erythropoiesis must be acquired from supplies already existing in the body. The source is reticuloendothelial macrophages, which serve the function of recycling iron from old erythrocytes and making it available to developing erythroid precursors.

Cells use at least three mechanisms to take up iron. Intestinal absorptive cells have cell surface transporters that carry ferrous (Fe^{2+}) ions directly across the membrane. Erythroid precursors have modified their transport pathway by adding a step to concentrate iron in a subcellular compartment, and then transferring it across the membrane of that compartment into the cytoplasm. Hepatocytes probably use both of

these mechanisms. Finally, recycling macrophages acquire iron through the phagocytosis of aged or damaged erythrocytes, lysing the cells and dissembling their hemoglobin.

Intestinal iron transport

Our current understanding of intestinal non-heme iron transport is illustrated in Figure 13.1. Iron absorption takes place in an acidic environment in the proximal small intestine, just distal to the gastric outlet. Most non-heme dietary iron is in the ferric (Fe^{3+}) form. It is reduced to Fe^{2+} by a brush border ferrireductase, most likely the recently identified duodenal cytochrome *b* (DCYTB). The Fe^{2+} ions pass through divalent metal transporter 1 (DMT1, formerly called Nramp2, DCT1), a membrane protein that allows iron to traverse the apical bilayer. DMT1 requires an acidic environment for its activity because it co-transports protons with iron atoms. Once it crosses the membrane, some of the iron is retained within the absorptive intestinal cells (enterocytes), and some is exported through the basolateral membrane through the action of a distinct transporter, ferroportin. A multicopper oxidase protein, hephaestin, facilitates basolateral transport, perhaps by oxidizing Fe^{2+} to Fe^{3+} to allow it to bind to the plasma iron carrier protein, transferrin.

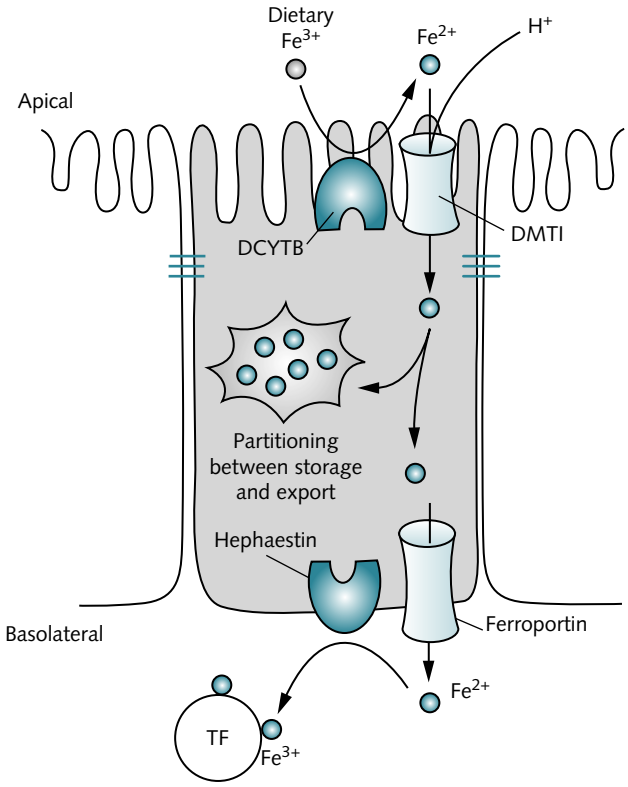


Fig. 13.1 Current understanding of intestinal non-heme iron absorption

The cartoon shows an absorptive enterocyte from the duodenal epithelium, joined to adjacent cells by iron-impermeable tight junctions. The apical brush border is at the top and the basolateral surface is at the bottom. Dietary Fe^{3+} iron is probably reduced by the ferrireductase DCYTB to produce Fe^{2+} ion for transport. Fe^{2+} crosses the apical membrane through the action of DMT1 to enter the cell. There, iron is partitioned between storage and export; stored iron is ultimately lost from the body when the epithelial cells senesce and exfoliate into the gut lumen. Meanwhile, a fraction of the iron is exported across the basolateral membrane, probably by ferroportin. Hephaestin is a ferroxidase-like protein that aids in iron export, probably by reducing the Fe^{2+} iron leaving ferroportin to the Fe^{3+} form, which binds to plasma transferrin (TF).

The mechanisms controlling the rate of iron flux through this transport system are not well understood. However, it appears that only some of the dietary iron taken up by enterocytes passes into the plasma. Iron that is retained within the enterocytes is lost from the body when these cells finish their short lifespan and slough into the gut lumen. The partitioning of iron (that is, the process that governs how much enters the plasma and how much is retained within cells) is probably regulated, and likely plays an important role in determining the overall efficiency of iron absorption.

Other iron uptake mechanisms

Upon entry into the plasma, iron attaches to transferrin, an abundant circulating protein that binds iron with extremely high affinity. Transferrin serves three important functions. Firstly, it keeps iron in solution. In an aqueous, neutral pH environment iron exists as Fe^{3+} ion, which is almost insoluble. Secondly, it renders iron non-reactive, and allows it to circulate in a safe, non-toxic form. Thirdly, transferrin facilitates the delivery of iron to cells bearing transferrin receptors on their surfaces.

Most differentiated cell types express few, if any, transferrin receptors, but there are three important exceptions: tumor cells, activated lymphocytes and erythroid precursors. Tumor cells presumably use transferrin receptors to optimize iron uptake to support rapid proliferation. The same may be true for activated lymphocytes. However, the greatest demand for iron is by erythroid precursors, to support the large-scale production of hemoglobin. In normal adults, about two-thirds of the total body iron endowment is found in hemoglobin, distributed among erythroid precursor cells and circulating erythrocytes.

Binding of iron-loaded transferrin to the transferrin receptor initiates receptor-mediated endocytosis, as shown in Figure 13.2. Portions of the cell membrane bearing liganded transferrin receptors invaginate into the cytoplasm, and bud off as intracellular vesicles (endosomes). Protons are pumped into the endosome to lower their internal pH, leading to the release of iron from transferrin. The liberated iron then leaves

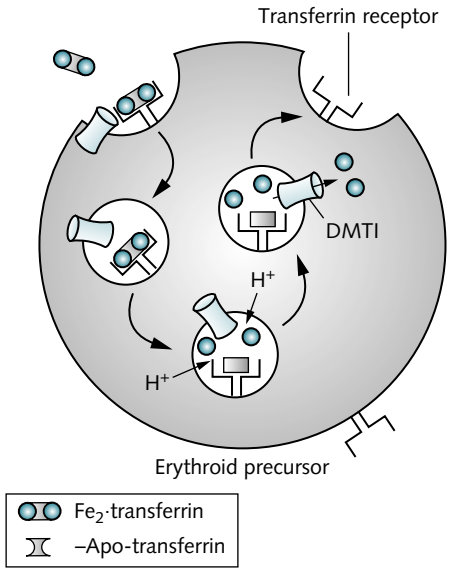


Fig. 13.2 The transferrin cycle
The transferrin cycle of receptor-mediated endocytosis is initiated by binding of diferric $(Fe_2)^{-}$ transferrin to a cell surface transferrin receptor. The ligand–receptor complex is internalized by invagination of clathrin-coated pits to form specialized endosomes. Influx of protons into the endosome decreases its pH to approximately 5.5, facilitating release of iron from transferrin. The iron is then transferred to the cytoplasm by DMT1. Apo-transferrin and transferrin receptor return to the cell surface for further cycles of iron uptake.

the endosome to enter the cytoplasm. This also requires a transmembrane transport step, which is probably mediated by DMT1 and facilitated by the low endosomal pH. Within the cytoplasm, iron is shuttled (by unknown mechanisms) to sites of use and storage. Meanwhile, transferrin and transferrin receptor proteins return to the cell surface, where they become available for further cycles of iron delivery.

Why should cells have evolved the complicated transferrin cycle when it is possible to take up iron directly? There are at least two likely answers. Firstly, tight binding of iron to transferrin is advantageous while iron is in the circulation, but it complicates the matter of bringing iron into cells. The pH-dependent release of iron, occurring in a controlled intracellular environment, solves the problem of liberating the iron. Secondly, binding of iron-loaded transferrin to transferrin receptors serves to concentrate iron in the vicinity of DMT1, probably achieving much higher local iron concentrations than would be possible without such a mechanism. This allows more efficient iron uptake by cells with large needs (erythroid precursors, tumor cells, activated lymphocytes) without exposing other cells to unnecessary iron.

Undoubtedly, other cell types use these and other schemes for assimilating iron. Hepatocytes and macrophages are particularly important in iron homeostasis and their iron uptake mechanisms, though not well understood, deserve mention. Hepatocytes express transferrin receptors and likely take up iron through the transferrin cycle. Hepatocytes are also capable of taking up non-transferrin-bound iron when the plasma iron concentration exceeds the binding capacity of transferrin. This is an abnormal situation, because there are usually about three times as many transferrin iron-binding sites as are needed (i.e. transferrin is normally about 30% saturated with iron). But patients with iron overload may have more iron than transferrin can accommodate, and that excess iron appears to be rapidly removed from the circulation by hepatocytes. The molecular mechanism for hepatic non-transferrin-bound iron uptake has not yet been identified. Furthermore, it is not known how hepatocyte iron stores are later mobilized when the iron they contain is needed elsewhere.

As discussed earlier, reticuloendothelial macrophages obtain iron by phagocytosing and breaking down erythrocytes. This probably takes place in discrete phagocytic vesicles within the cells, and likely involves the action of heme oxygenase, an enzyme that catalyzes the degradation of heme. Similar to intestinal cells and hepatocytes, reticuloendothelial macrophages partition their iron content into retained and released portions. This process is probably regulated in response to the iron needs of the body. While there is currently no direct way to measure how much iron is retained and how much is released, commonly used laboratory tests provide some information. The concentration of serum iron (and hence the transferrin saturation) is determined by two factors: macrophage iron release and erythroid iron utilization. When erythropoiesis occurs at a steady rate, transferrin saturation is determined primarily by the rate of macrophage iron release, increasing when there is increased iron export and decreasing when iron is retained or when less iron is being recycled from erythrocytes. In contrast, the concentration of serum ferritin roughly correlates with the amount of storage iron in the body. The origin(s) of serum ferritin is not known, but it appears to be derived primarily from hepatocytes and reticuloendothelial macrophages. Serum ferritin is not a very accurate indicator, however, because levels are increased by inflammation, tissue damage and rare congenital hyperferritinemia disorders. Nonetheless, a low serum ferritin value invariably indicates depleted iron stores.

Sites of iron storage

Iron is stored within cells in the cavities of ferritin protein multimers. There are two types of ferritin subunit (L and H), both approximately 20 kDa in size. These subunits assemble in varying proportions into 24-subunit cage-like structures. Up to several thousand iron atoms can be stored in each ferritin multimer. Like transferrin, ferritin serves the purpose of

preventing iron from reacting with other cellular constituents, and allows controlled iron release in response to increased cellular needs. The molecular details of iron incorporation into and release from ferritin are not well understood. Under some circumstances, ferritin and other cellular components are partially degraded and conglomerated to form hemosiderin, a heterogeneous iron-containing substance that probably serves little purpose but to keep iron from causing harm. Both ferritin and hemosiderin accumulate in iron-overloaded tissues.

The liver serves as the primary depot for iron in excess of immediate needs. It has a very large capacity for storing iron, though this capacity is ultimately exceeded in iron overload disorders. While other tissues (myocardium, pancreas) also fill up with iron in iron overload, the liver is frequently the first site where damage from iron overload becomes apparent. Hepatocytes avidly take up non-transferrin-bound iron from the plasma.

Reticuloendothelial macrophages are also important for iron storage, but their iron comes from degraded erythrocytes. Patients treated with frequent transfusions typically accumulate excess iron in macrophages first, and only later in other tissues. This pattern of iron accumulation has been referred to as 'siderosis' to distinguish it from hemochromatosis, which is primary iron loading of parenchymal cells.

Regulation of iron homeostasis

Iron homeostasis requires the coordinated regulation of iron transport and iron storage so that tissues will have adequate amounts to meet their needs but will not become overloaded with iron. Regulation must involve the control of cellular iron import, export and partitioning. Recently, several clues have emerged that should help lead to a comprehensive understanding of regulation at each of these steps.

There are at least four known regulators of intestinal iron absorption: iron stores, erythropoietic demand, hypoxia and inflammation. The *stores regulator* modulates absorption several-fold, increasing absorption in iron deficiency and decreasing absorption in iron overload. The *erythroid regulator* is more potent: it can increase iron absorption many-fold when erythropoiesis becomes iron-restricted. The *hypoxia regulator* is not well characterized, but its effects appear to be distinct from those of the erythroid regulator. This regulator increases iron absorption in response to hypoxia. Finally, emerging evidence suggests that an *inflammation regulator* also exists, decreasing iron absorption in response to inflammation from a variety of causes.

Recently, a peptide hormone has been discovered that is likely to be a common effector of the stores, erythroid, hypoxia and inflammation regulators. Hepcidin (also called LEAP, HAMP) is a 20- to 25-amino acid protein, produced by the liver, which is cleaved from a larger precursor molecule. Animal experiments have shown that hepcidin acts as a negative regulator of both intestinal iron absorption and macrophage iron recycling. The production of hepcidin is increased in animals given carbonyl iron to produce iron overload, presumably as part of a compensatory mechanism to decrease iron absorption and decrease plasma iron (primarily derived from recycling macrophages). This suggests that it mediates the effects of the stores regulator. The production of hepcidin is decreased by iron-restricted erythropoiesis, allowing more iron to enter the body through the intestine and more iron to enter the plasma from recycling macrophages. In this way, it also mediates the effects of the erythroid regulator. The production of hepcidin is also decreased in hypoxia, suggesting that it is an effector of the hypoxia regulator. Finally, hepcidin expression is induced by inflammation, probably through a direct action of the cytokine interleukin-6 on hepatocytes. In this case, induced hepcidin expression leads to decreased intestinal iron absorption and decreased macrophage iron release, acting as an inflammation regulator. There is growing evidence that, in response to the inflammatory regulator, increased hepcidin expression contributes to the abnormal iron homeostasis observed in the 'anemia of chronic disease' (also known as the anemia of chronic disorders). It is likely that a useful clinical assay for hepcidin levels in serum and/or urine will be available within the next few years. The actions of the various regulators are summarized in Figure 13.3.

Iron overload disorders

Iron deficiency is a major public health problem and the most common nutritional cause of anemia. However, most genetic

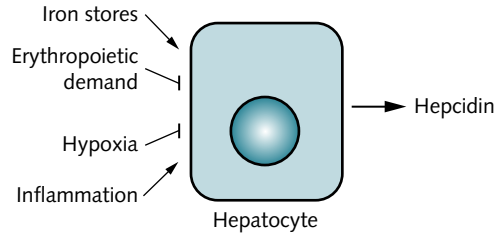


Fig. 13.3 Regulators of iron absorption
A general model for the regulation of intestinal iron absorption is depicted. Although this model is still somewhat speculative, recent data suggest that the four known regulators (iron stores, erythropoietic demand, hypoxia and inflammation) may all act by modulating hepatic production of hepcidin, an iron-regulatory hormone. Increased iron stores and inflammation both appear to increase hepcidin expression (→), whereas increased erythropoietic demand and hypoxia appear to decrease hepcidin expression (⊥).

disorders of iron metabolism result in iron overload rather than iron deficiency. They are attributable to mutations that affect the regulation of intestinal iron absorption and body iron distribution. The prominence of mutations that cause iron loading, rather than iron deficiency, can probably be explained by the fact that humans live in an iron-rich environment, and the body has developed strategies to limit iron absorption rather than to enhance it.

Several years ago, it was widely believed that there was only one major genetic disorder leading to iron overload. It is now clear that there are many genetic iron overload disorders. They can be generally classified as hemochromatosis disorders (i.e. iron deposition in parenchymal cells) or siderosis disorders (i.e. deposition of iron in reticuloendothelial macrophages). The known disorders and their genetic causes are listed in Table 13.1.

HFE-associated hemochromatosis

The most common form of hemochromatosis was classically described as the triad of cirrhosis, diabetes and skin melanosis ('bronze diabetes'). It is a late-onset disorder, which is inherited in an autosomal recessive pattern and characterized by iron deposition in parenchymal cells of the liver, pancreas and heart. Macrophages of the reticuloendothelial system are relatively spared. This disorder results from a small but chronic increase in intestinal iron absorption, averaging about two- to three-fold above the normal level. Over time, the presence of iron causes damage by promoting the formation of toxic oxygen radicals, which attack cellular structures and thereby cause reactive fibrosis. The earliest manifestation of HFE-associated hemochromatosis is increased transferrin saturation, often approaching 100% before tissue iron deposition is noted. The treatment for hemochromatosis is phlebotomy, and this has been used effectively for more than half a century. Initially, blood is removed frequently to rapidly decrease storage iron. Later, iron balance is maintained by periodic phlebotomy, titrated to meet the needs of the individual patient. This treatment apparently produces no significant morbidity, and it has been shown to normalize the life expectancy of affected patients.

Hemochromatosis has been recognized as an inborn error of iron metabolism since the 1930s. In 1976 a French physician, Marcel Simon, made the important observation that the genetic predisposition to hemochromatosis was linked to the human major histocompatibility complex on chromosome 6p, and was most frequently associated with an HLA-A3 haplotype. This insight laid the groundwork for the discovery of causative mutations in the *HFE* gene 20 years later. It is now known that most patients with classical hemochromatosis are homozygous for a unique mutation (cysteine 282 to tyrosine, or C282Y) in *HFE*.

HFE is an atypical HLA class I molecule, similar to its chromosomal neighbors. Although most members of this family are involved in immune regulation, HFE has no known function in the immune system. It interacts with the transferrin receptor on the cell surface, but there is controversy regarding the consequences of this interaction. It is still not clear how HFE functions in the regulation of intestinal iron absorption, but recent information suggests that it plays some role in regulating the expression of hepcidin.

The C282Y mutation is highly prevalent. In typical populations of individuals of European descent, the carrier frequency has been estimated to be between about 1 in 8 and 1 in 10. This means that about 1 in 200 individuals are homozygous, and at risk of iron loading. However, not all C282Y homozygotes will develop clinical hemochromatosis. There is a wide range in iron loading and its complications. Some individuals will have severe manifestations by the third decade of life, whereas others may never have signs or symptoms of hemochromatosis. This variability is probably explained by both genetic factors (modifying genes) and environmental factors (e.g. alcohol intake, dietary iron consumption, menstruation). The modifying genes have not yet been identified, but their characterization is the subject of active investigation in this field.

In addition to C282Y, other mutations and polymorphisms have been identified in the *HFE* gene. The most common of these is a histidine-to-aspartic acid substitution at amino acid 63 of the protein (H63D). This polymorphism is found in about one-fifth of the world's population. Although it may occasionally be associated with iron overload, particularly when it is found in individuals heterozygous for the C282Y mutation, its clinical significance remains poorly understood, though clinical laboratories test for it. Most other *HFE* mutations are quite rare and are not identified by routine screening tests.

In the past, the proportion of at-risk C282Y homozygous individuals who develop clinical hemochromatosis was estimated to be about 20–40%. Recently, Beutler and colleagues published the results of a large, questionnaire-based study of patients seen by a health maintenance organization, in which they concluded that fewer than 1% of C282Y homozygotes would have severe disease. However, their study design likely underestimated the number of affected individuals whom most clinicians would feel compelled to treat for hemochromatosis. This controversy is the subject of ongoing debate in the literature and has not yet been resolved. Until more data are available, it is probably prudent to assume that earlier estimates of prevalence may be correct. The consensus in the field is that this must still be considered a very prevalent disorder.

Juvenile hemochromatosis

Juvenile hemochromatosis is similar to HFE-associated he-

Table 13.1 Iron overload disorders.

Disorder	Chromosomal locations of defective genes	Gene (types of mutations)	Phenotype
Hemochromatosis disorders			
HFE-associated hemochromatosis (also called Type 1 hemochromatosis)	6p, near the HLA complex	*HFE* (missense and splicing mutations; C282Y is the most important)	Iron accumulation in the parenchymal cells of the liver, heart, pancreas; elevated transferrin saturation; relative paucity of iron in macrophages. Clinical manifestations include liver fibrosis, cirrhosis, markedly increased incidence of hepatocellular carcinoma, cardiomyopathy, diabetes
Juvenile hemochromatosis (also called Type 2 hemochromatosis)	1q 19q	Unknown Hepcidin (all known mutations prevent production of any hepcidin protein)	Similar to HFE-associated hemochromatosis, but greatly accelerated, leading to severe cardiac and endocrine complications in the second decade of life
TFR2-associated hemochromatosis (also called Type 3 hemochromatosis)	7q	Transferrin receptor-2 (*TFR2*; missense and nonsense mutations)	Similar to HFE-associated hemochromatosis
Type 5 hemochromatosis	? (Postulated to explain hemochromatosis in patients without mutations in known genes)	?	Similar to HFE-associated hemochromatosis but apparently not due to mutations in *HFE*, *TFR2* or hepcidin
Siderosis disorders			
Autosomal dominant siderosis (also called Type 4 hemochromatosis)	2q	Ferroportin (missense mutations)	Macrophage-predominant iron loading; parenchymal iron loading can occur later. Some patients have anemia early in their course (particularly women). Ferritin levels are markedly elevated, but serum transferrin saturation generally is not
African siderosis	?	Unknown	Similar to autosomal dominant siderosis but ferroportin mutations have not been reported. Thought to be a combination of genetic and environmental factors
Disorders of iron balance			
Atransferrinemia	3q	Transferrin (missense mutations)	Deficiency in serum transferrin leading to tissue iron overload and severe iron deficiency anemia
Aceruloplasminemia	3q	Ceruloplasmin (missense and null mutations)	Deposition of iron in the brain, liver and pancreas. Late-onset neurodegenerative disease, dementia and diabetes

mochromatosis, but very rare, and characterized by earlier onset of iron loading and its complications. The target organs are the same as those affected in HFE hemochromatosis, but cardiac and endocrine dysfunction are more problematic, and untreated patients typically die from cardiomyopathy by age 30 years. Liver cirrhosis and failure are uncommon. There are several possible explanations for this pattern. Firstly, experience with patients who develop siderosis from

chronic transfusion therapy suggests that rapid iron loading is especially toxic for the heart and endocrine tissues. Secondly, pathological iron deposition in the adolescent years may be particularly bad for young hearts which are growing to meet the demands of a larger body mass; this is analogous to the problems noted with doxorubicin cardiotoxicity in this age group. Furthermore, endocrine problems are probably more apparent in adolescents because they fail to go through normal pubertal development. Juvenile hemochromatosis is a particularly lethal disorder, but it can be effectively treated by phlebotomy.

Studies of families with juvenile hemochromatosis have shown that there are at least two genetic loci responsible. Some individuals are homozygous for mutations in the hepcidin gene on human chromosome 19q. To date, only two hepcidin mutations have been reported, both of which completely prevent production of the hepcidin protein. Other juvenile hemochromatosis families show linkage to a small region of chromosome 1q. The gene responsible for this form of juvenile hemochromatosis has not yet been identified, but it has been speculated that it may encode the as yet unknown hepcidin receptor.

TFR2-associated hemochromatosis

A third form of autosomal recessive hemochromatosis (type 3) has been reported to be indistinguishable from HFE-associated hemochromatosis in its clinical manifestations, but patients have no mutations in the *HFE* gene. Instead, their disease is caused by mutations in the transferrin receptor-2 (*TFR2*) gene on chromosome 7q. Transferrin receptor 2 is a protein of unknown function that is highly homologous to the transferrin receptor (also known as transferrin receptor-1), and highly expressed by hepatocytes and hematopoietic cells. TFR2 is capable of binding transferrin and bringing it into the cell, but it does so at much lower efficiency than does transferrin receptor. TFR2 does not form a complex with HFE, and its role in iron homeostasis is completely unknown. Some patients have been shown to have nonsense mutations, indicating that TFR2 is not an essential protein for survival. These patients are effectively treated by phlebotomy. It is likely that clinical tests for *TFR2* mutations will be developed within the next few years.

Type 5 hemochromatosis

Type 5 hemochromatosis is not yet a defined entity, but rather a disease that has been inferred to exist because there are patients with HFE-like hemochromatosis who have no apparent mutations in the *HFE*, *TFR2* or hepcidin genes. This designation probably includes a disorder due to at least one other gene that plays a role in the regulation of iron balance.

Autosomal dominant siderosis

Autosomal dominant siderosis, also called autosomal dominant (or type 4) hemochromatosis, has a distinct clinical picture. Patients with this disorder may have iron deficiency anemia early in life, but later present with increased serum ferritin concentration and macrophage iron accumulation. Ultimately, they can have parenchymal iron deposition in addition to macrophage iron loading, probably because their total body iron exceeds the storage capacity of their macrophages.

This disorder has a very interesting pathogenesis. It is due to missense mutations in ferroportin, the cellular iron exporter. This seems paradoxical at first, because the mutations alter ferroportin function, and ferroportin acts as the basolateral enterocyte transporter involved in intestinal iron absorption (Figure 13.1). However, it is important to consider that ferroportin also plays a major role in macrophage iron release. Apparently, loss of one functional ferroportin gene results in significant impairment of macrophage iron release, resulting in a decrease in the amount of plasma iron available to developing erythroid precursors. Iron-restricted erythropoiesis probably signals for a compensatory increase in intestinal iron absorption, which is not compromised by the loss of one ferroportin allele in the enterocytes. As a result, total body iron levels gradually rise. In this way, autosomal dominant siderosis exemplifies the meticulous balance involved in normal iron homeostasis.

African siderosis

The pathology of African siderosis is strikingly similar to that of autosomal dominant siderosis due to ferroportin gene mutations, but the genetic basis of African siderosis has not been described. Once called Bantu siderosis because of the population affected, this disorder was originally attributed to excessive dietary iron intake. It is common in sub-Saharan Africans, many of whom drink a traditional alcoholic beverage brewed in non-galvanized steel drums. The iron content of the brew is substantial, resulting in massive iron ingestion. However, the observations that not all drinkers develop iron overload and that some individuals develop similar iron overload without drinking the beverage support the notion that there is a genetic component to this disorder. It is not yet known whether European and American individuals of African descent are also more susceptible to iron overload as a result of the same iron-loading gene.

Abnormal iron distribution

There are two well-characterized (but very rare) disorders due

to mutations in plasma proteins important in iron metabolism. These mutations do not directly affect intestinal iron absorption. Rather, they perturb tissue iron distribution.

Atransferrinemia

Atransferrinemia is a severe deficiency of the plasma iron-binding protein transferrin, due to mutations that truncate or alter the coding sequence of the transferrin gene. As a result, erythroid precursors are iron-starved and severe anemia results. Paradoxically, all non-hematopoietic tissues are iron-loaded, probably because intestinal iron absorption is enhanced to try to provide more iron to erythroid precursors and because non-transferrin-bound iron is avidly taken up by many parenchymal cell types. This disorder can be treated by transfusion of packed red blood cells or, more appropriately, by infusion of human transferrin.

Aceruloplasminemia

Aceruloplasminemia is deficiency or absence of plasma ceruloplasmin. Ceruloplasmin was once thought to be a plasma copper carrier but it is now clear that its primary role is as a ferroxidase, aiding in the release of iron from macrophages, hepatocytes and cells of the central nervous system. Patients with this disorder are generally well early in life, but gradually develop tissue iron deposition in the liver, pancreas and brain. They typically present in middle age with retinal degeneration, dementia, hepatic iron deposition and diabetes. Treatment with deferoxamine is ineffectual; treatment with normal plasma may provide some benefit.

Iron deficiency disorders

It is striking that virtually all recognized inborn errors of iron metabolism result in tissue iron loading rather than iron deficiency. As mentioned earlier, this suggests that there has been strong evolutionary pressure to attenuate the intestinal absorption of iron in humans, perhaps to avoid the toxicity of a metal that is abundant in our environment. Consequently, genetic defects typically result in increased iron uptake. However, there are almost certainly inherited iron deficiency disorders as well. Although the genes affected have not yet been identified, there are several case reports of familial microcytic anemia associated with impaired intestinal iron absorption. It is likely that the mutations responsible for some of these will be discovered within the next few years.

Conclusions

Iron disorders are among the most common of human afflictions. They invariably result from abnormalities of iron balance. Iron overload is underdiagnosed because it produces signs and symptoms that are common in adult populations. However, iron overload disorders are usually easy to treat and clinicians should be vigilant in considering them.

Further reading

General

Aisen P, Enns C, Wessling-Resnick M. (2001) Chemistry and biology of eukaryotic iron metabolism. *International Journal of Biochemistry and Cell Biology*, **33**, 940–959.

Andrews NC. (1999) Medical progress: disorders of iron metabolism. *New England Journal of Medicine*, **341**, 1986–1995.

Regulation of iron homeostasis

Andrews NC. (2002) A genetic view of iron homeostasis. *Seminars in Hematology*, **39**, 227–234.

Finch C. (1994) Regulators of iron balance in humans. *Blood*, **84**, 1697–1702.

Nemeth E, Valore EV, Territo M *et al.* (2003) Hepcidin, a putative mediator of anemia of inflammation, is a type II acute-phase protein. *Blood*, epub ahead of print.

Nicolas G, Chauvet C, Viatte L *et al.* (2002) The gene encoding the iron regulatory peptide hepcidin is regulated by anemia, hypoxia, and inflammation. *Journal of Clinical Investigation*, **100**, 1037–1044.

Iron overload disorders

Ajioka RS, Kushner JP. (2002) Hereditary hemochromatosis. *Seminars in Hematology*, **39**, 235–241.

Beutler E, Felitti VJ, Koziol JA *et al.* (2002) Penetrance of 845G→A (C282Y) HFE hereditary hemochromatosis mutation in the USA. *Lancet*, **359**, 211–218.

Bulaj ZJ, Ajioka RS, Phillips JD *et al.* (2000) Disease-related conditions in relatives of patients with hemochromatosis. *New England Journal of Medicine*, **343**, 1529–1535.

Camaschella C, Roetto A, DeGobbi M. (2002) Genetic haemochromatosis: genes and mutations associated with iron loading. *Best Practice & Research Clinical Haematology*, **15**, 261–276.

Montosi G, Donovan A, Totaro A *et al.* (2001) Autosomal-dominant hemochromatosis is associated with a mutation in the ferroportin (SLC11A3) gene. *Journal of Clinical Investigation*, **108**, 619–623.

Roetto A., Papanikolaou G, Politou M *et al.* (2002) Mutant antimicrobial peptide hepcidin is associated with severe juvenile hemochromatosis. *Nature Genetics*, epub ahead of print.

Roy CN, Andrews NC. (2001) Recent advances in disorders of iron metabolism: mutations, mechanisms and modifiers. *Human Molecular Genetics*, **10**, 2181–2186.

Abnormal iron distribution

Beutler E, Gelbart T, Lee P *et al.* (2000) Molecular characterization of a case of atransferrinemia. *Blood*, **96**, 4071–4074.

Harris ZL, Takahashi Y, Miyajima H *et al.* (1995) Aceruloplasminemia: molecular characterization of this disorder of iron metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 2539–2543.

Chapter 14 Hemoglobinopathies due to structural mutations

Ronald L Nagel

Introduction

This chapter discusses those hemoglobinopathies that are caused by mutations in the exon (i.e. coding) portion of the α or β globin genes. Alterations of globin gene expression (thalassemias) are reviewed in Chapter 1. We will limit ourselves to the most frequent mutations that must be considered in the differential diagnosis of the common hemoglobinopathies.

Normal hemoglobin structure and function

The adult major hemoglobin molecule, Hb A, is a tetramer formed by four polypeptide chains: two α chains and two β chains. Each of these chains is attached to a prosthetic group (heme) formed by protoporphyrin IX in a complex with a single iron molecule (Figure 14.1).

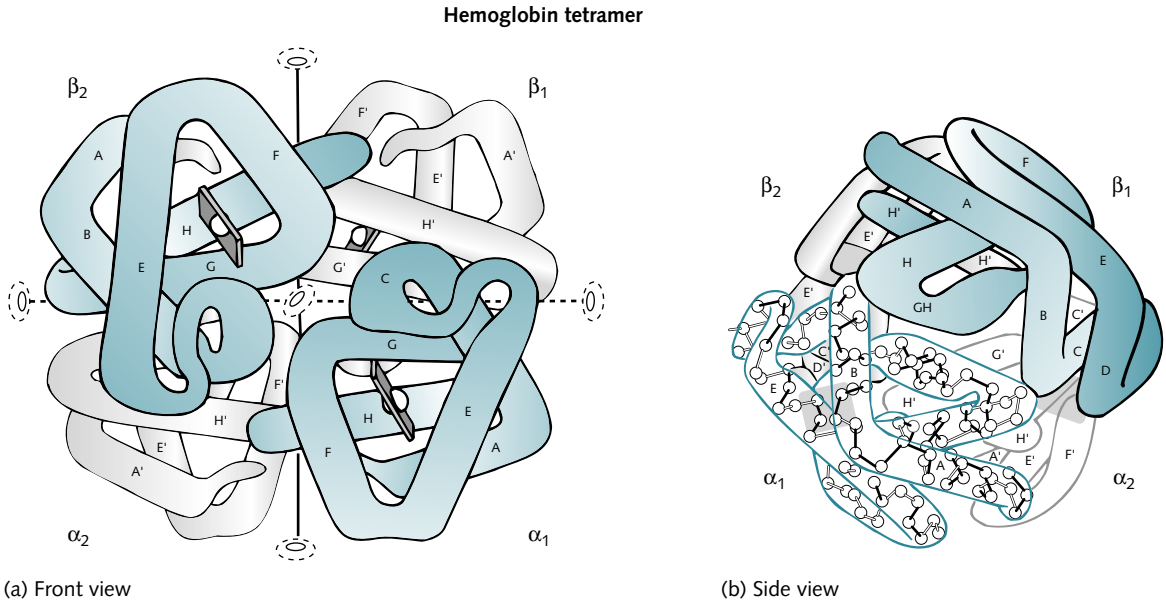


Fig. 14.1 Hemoglobin molecule
The *front view* depicts the hemoglobin tetramer and the three axes of symmetry. The vertical line (marked by a solid ellipse) tracks the true two-fold axis of symmetry (if you look down this axis you will see first the central cavity constituted by the β chains). The dashed ellipses and dashed lines mark the two pseudoaxes of symmetry, since the symmetry is only approximate. Only the 21 carbons are shown, with none of the side chains. Numbers in bold type depict the residues in direct contact between the α_1 and β_2 chains. In deep red is shown the $\alpha_1\beta_2$ dimer, which never dissociates and interacts with the $\alpha_2\beta_1$ dimer (light color) to change the conformation from T to R. Alterations in this area can produce high- or low-affinity hemoglobins. The *side view* of the tetramer depicts the stable dimer—the one that does not move or dissociate.

The heme is semi-buried in the globin, surrounded by a hydrophobic niche that favors the maintenance of the ferrous state of the iron. The heme pocket is large enough for oxygen to penetrate, but large ligands (molecules capable of binding to the iron), such as carbon monoxide and the family of isocyanates, have progressive difficulty in finding the iron.

Oxygen transport to tissues, the ultimate purpose of hemoglobin, is dependent on blood flow, which in turn is affected by cardiac output and by microcirculatory size and distribution, the hemoglobin concentration and O_2 extraction by the tissues, which in turn is dependent on the shape of the oxygen binding curve of the red cells and on tissue pO_2. The shape of the oxygen equilibrium curve for hemoglobin is sigmoid. This shape is determined by the extent of cooperativity. The initial portion of the curve has a very low slope, reflecting a low affinity for oxygen by hemoglobin at the beginning of the loading process. In other words, when hemoglobin is totally deoxygenated it has a rather poor avidity for oxygen (Figure 14.2). As the loading proceeds, and as the molecule binds more oxygen molecules, the slope of the reaction begins to change rapidly and becomes steep, indicating that the affinity for oxygen has markedly increased. After two molecules of oxygen have bound to two hemes of deoxyhemoglobin tetramers, the protein changes its avidity for oxygen. This property helps hemoglobin tetramers to promptly become fully oxygenated. Hence, in red cells that are exposed to sufficient oxygen to oxygenate only half of the hemes available, most molecules will either not be oxygenated at all or will be entirely oxygenated, with a very small compartment of partially oxygenated molecules.

At the molecular level, cooperativity is accounted for by the fact that hemoglobin can exist stably in only two different conformations, one for the oxygenated molecule (R state) and another for the deoxygenated molecule (T state), without intermediate conformations. The molecule of hemoglobin will bind two or three molecules of oxygen at low affinity (T state). While this concept has been challenged recently, postulated intermediate stable states have not been generally recognized.

The heme triggering mechanism for conformational change has been resolved. The iron in deoxyhemoglobin is slightly out of the plane of the heme (domed configuration) because the pyrrole rings are also slightly pyramidal. When the ligand binds the sixth coordinating position of the iron, significant steric stresses are introduced, and to relieve this strain the distal histidine moves 8° to become perpendicular to the heme, significantly decreasing the doming of the iron (the angle between iron and the heme decreases to 4°). There is also the displacement of FG5 in the same direction of the histidine F8. The configuration around the heme has now changed to the oxygenated (R state), and a chain of events

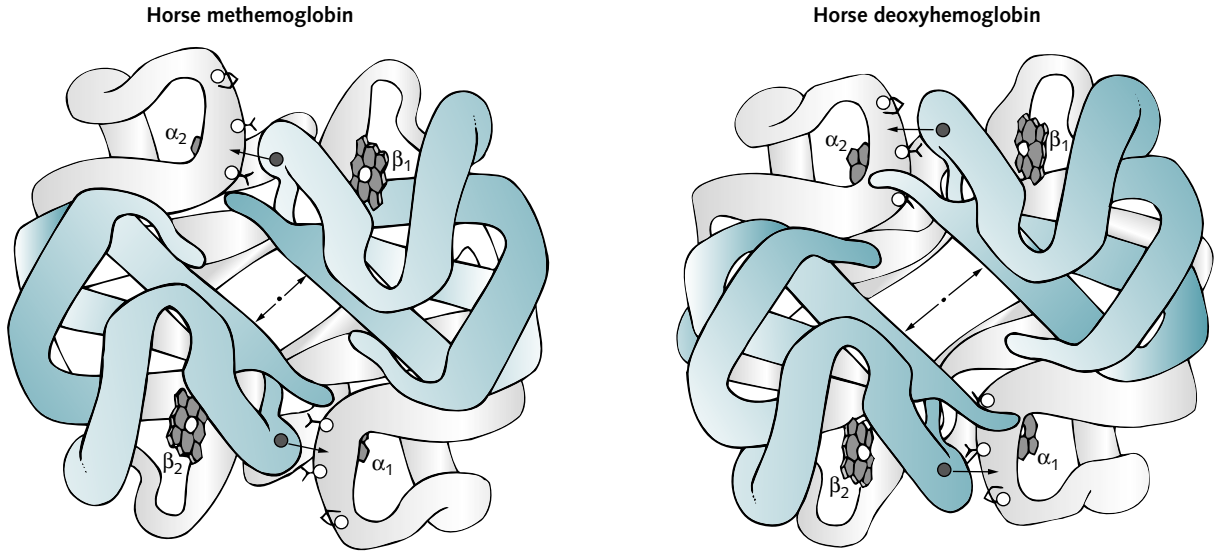


Fig. 14.2 The allosteric transition of hemoglobin: R→T

Right: Deoxygenated (T, tense) conformation of hemoglobin. Notice that the centre cavity (space between the two β chains, in blue) is larger than in the oxygenated tetramer (left). This space in the molecule is occupied by 2,3-diphosphoglycerate (2,3-DPG) when the tetramer is in the blue cell. The allosteric effector, 2,3-DPG, is in a little over equimolar concentration with the tetramer. The T form has *low* affinity for oxygen. *Left*: Oxygenated (R, relaxed) conformational state of hemoglobin: the central cavity has been reduced in size by the movement of the β chains (blue) towards each other. This tetramer cannot bind 2,3-DPG. The R form has *high* affinity for oxygen. The change in conformation forms and breaks bonds between the β and α chains (β dimers (blue) and α chains (white)) on each side of the central cavity. The title says 'methemoglobin' but this hemoglobin form is identical to oxyhemoglobin. Perutz did his crystallography for reasons of convenience.

takes place involving the critical interactions that change the conformation of the hemoglobin tetramer.

Hemoglobin binds CO_2 while it is delivering O_2 and releases CO_2 when it is binding O_2, helping to dissipate the increase in concentration of CO_2 in the tissues and conveniently delivering this metabolic end product to the alveoli of the lungs. It accomplishes this particular task with ease because carbon dioxide is an inhibitor of hemoglobin oxygen-carrying capacity by decreasing the oxygen affinity of the molecule.

Hemoglobin binds hydrogen ions efficiently in a low-pH environment and releases them when it encounters high pH (the *Bohr effect*). The Bohr effect describes the changes in oxygen affinity secondary to pH changes within a certain range : the lower the pH the lower the affinity or the higher the p50. This means that an increased concentration of protons favors a low-affinity state in hemoglobin. In other words, deoxyhemoglobin binds more protons than the oxy conformer does.

Sickle cell anemia

Genetics

The genetic basis of sickle cell anemia is central to the history of medical genetics. The disease was first described by Herrick (1910), a cardiologist, who observed sickle-shaped red cells (Figure 14.3) in the blood of a medical student from Grenada, who suffered from chronic hemolytic anemia. James V. Neel was the first to suggest that sickle cell anemia was a homozygous state and sickle trait (the asymptomatic carrier state) was a heterozygous state of a genetic character that had not yet been defined. Linus Pauling proposed that the sickling represented an abnormality of the hemoglobin molecule, based on the observation of the medical student that sickle cells, induced by deoxygenation, were birefringent. Birefringence indicated to Pauling that some type of molecular alignment or orientation existed inside these red cells, and since hemoglobin predominates overwhelmingly it had to be this particular protein which was involved in the pathology. Electrophoretic studies confirmed this interpretation and the concept of *molecular disease* was born.

The biochemical definition of Hb S was achieved by Vernon Ingram, who developed a technique capable of probing the primary sequence of a protein, and revealed that sickle hemoglobin differed from normal hemoglobin by a single peptide, which was later found to have a single amino acid change in position 6 of the β chain, in which a glutamic acid was replaced by a valine. At the gene level the change was A→T in the middle nucleotide of codon 6 of the β chain.

The globin genes were among the first to be located in the human genome, the β and β-like globin genes mapping to chromosome 11. Using 11 polymorphic sites located in the β gene cluster, it has been established that the β^S gene is associated with three distinctly different chromosomal haplotypes in Africa, identifiable by their specific array of DNA polymorphic sites (haplotypes), each one exclusively present in three separate geographical areas in Africa (Figure 14.4). Finally, a rather small ethnic group in southern Cameroon, the Eton people, have their own haplotype linked to the sickle gene.

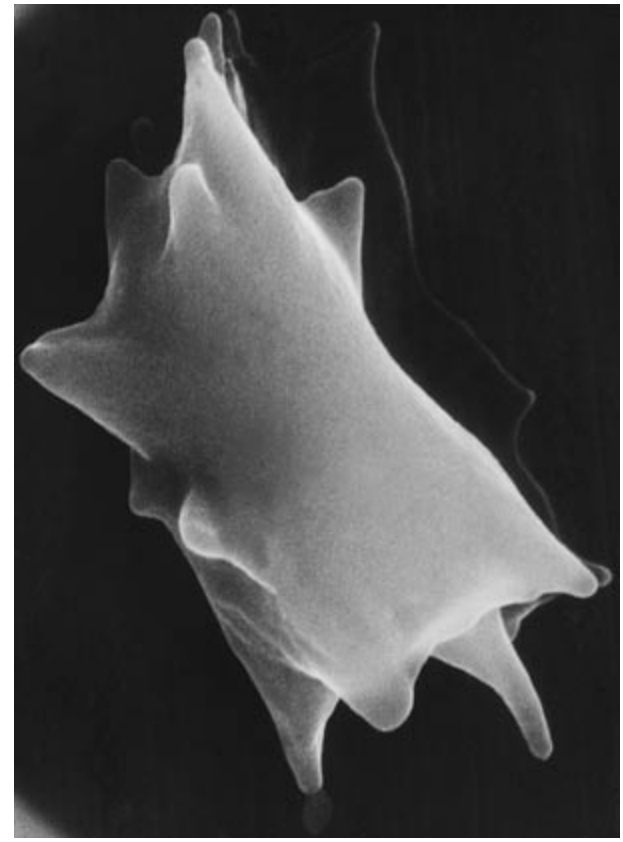


Fig. 14.3 Sickled cell
This is an Hb S-homozygous (SS) red cell that has been deoxygenated. Notice the digitations stemming in all directions. These digitations are the product of the presence of fascicles of fibers, which are the polymerized form of deoxyHb S.

The sickle gene has arisen around the world on at least five separate occasions, but the present-day gene frequencies demonstrate that the heterozygote is favored, compensating for the lower fitness of the homozygous state (balanced polymorphism). The selective pressure involved is the protective effect of Hb S in sickle carriers against *Plasmodium falciparum* malaria.

The multicentric origin of the sickle gene in the world expanded when a different haplotype was found to be associated with this gene in the eastern oasis of Saudi Arabia and among the 'tribals' of India. This Indo-European sickle mutation has been proposed to have originated in the Harappa of the

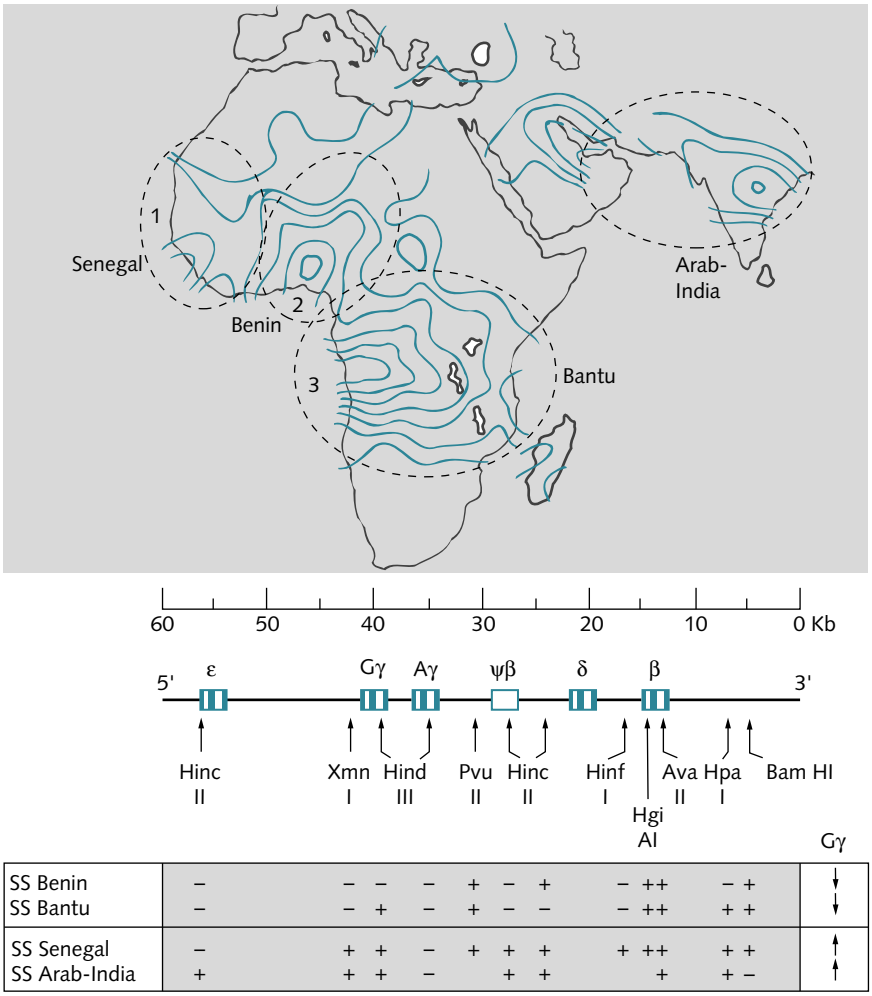


Fig. 14.4 β gene cluster haplotypes linked to β^S in Africa and the Middle East/Pakistan/India
At the top of the figure are shown the geographic distributions corresponding to the haplotypes described below. A haplotype is a particular array of polymorphic sites (that is, sites that vary among individuals), defined here by the capacity of endonuclease enzymes to recognize the short sequence and cut the DNA.

Indus Valley, and then was distributed, probably during the Sassanian Empire, to the present sites (eastern Saudi Arabia, Bahrain, Kuwait and Oman).

The sickle gene has also spread through gene flow; for example, the Benin haplotype-linked β^S has found its way to North Africa, Sicily, Greece, Turkey, most of the Arab world and the Americas, through the vagaries of wars of conquest (Sudanese troops in Sicily during the Arab conquest, for example) and the horrors of the Atlantic slave trade.

Clinical features

Sickle cell anemia refers to the homozygous state for the β^S gene, in which the majority of the hemoglobin in the red cells is sickle hemoglobin (Hb S). This induces sickling (marked changes in red cell shape produced by intercellular Hb S polymers) when the oxygen tension is reduced, increases in red cell viscosity, decreased pliability and, by consequence, lowering of the sickle cell rheological competence, and hemolysis.

In addition, there are myriad pleiotropic effects, such as increased adherence of sickle cells to endothelium, induction of red cell dehydration and irreversible sickled cell (ISC) status, autosplenectomy and urine concentration defects (Figure 14.5).

The primary events as well as the pleiotropic events are, in addition, affected by modifier or epistatic genes, which tend to be polymorphic, that is, different among individuals. The strong effect of all of these factors explains the great inter-patient differences in the intensity of the phenotype: some patients are severely affected whilst others have only mild disease, and the majority span these two extremes.

Anemia is present in all cases, with an average hemoglobin concentration of 8.4 g/dl, reticulocytosis varying between 5 and 20%, dense cell fraction between 5 and 40%, increased white cells ($10–20 \times 10^9$/L) and platelets in the upper limit of normal. Most of these features are modulated by the level of fetal hemoglobin (Hb F) in the sickle cells and the distribution function of the levels among red cells, since Hb F is het-

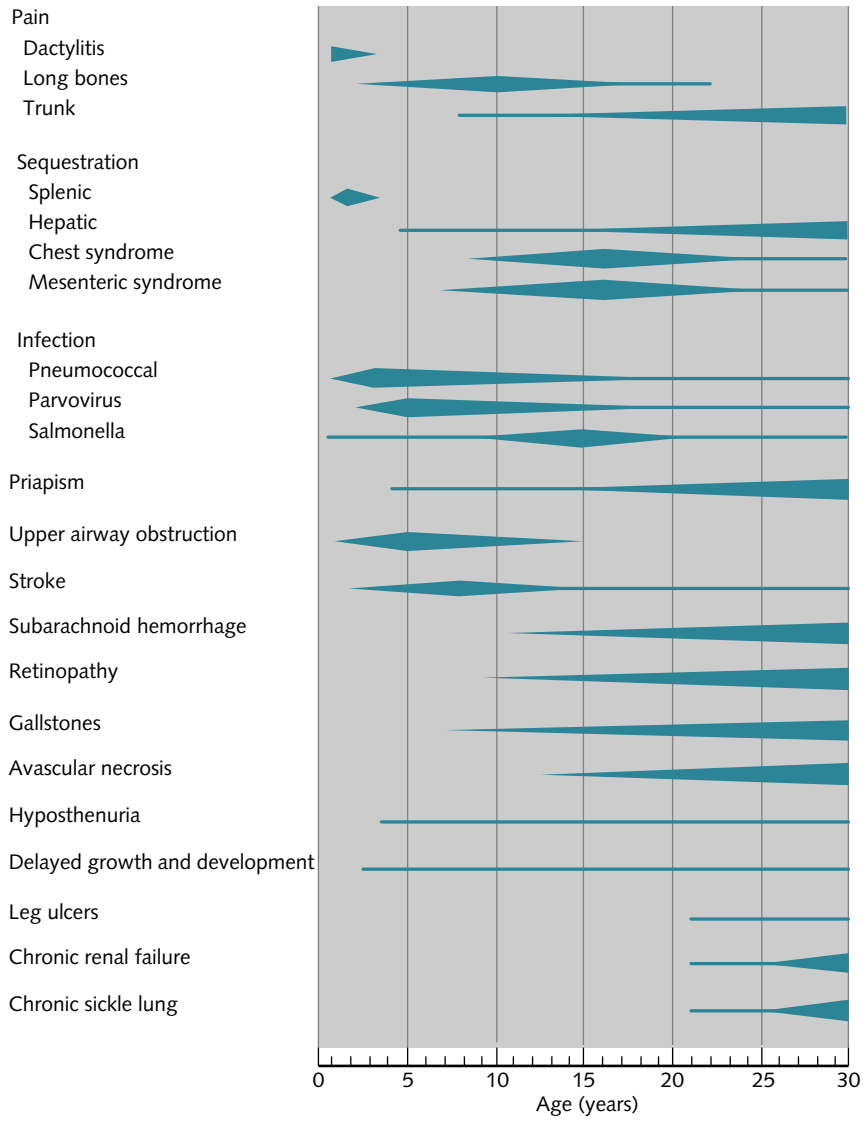


Fig. 14.5 Age-dependency of complications in sickle cell anemia
Reproduced from Davies SC, Oni L. (1997) Management of patients with sickle cell disease. *British Medical Journal*, **315**, 656–60, with permission.

erogeneously expressed in reticulocytes and mature red cells. They are also modulated by the co-presence of α thalassemia (−α/αα and particularly −α/−α), a feature quite common among sickle cell anemia patients (between 20 and 50% of cases, according to age). Gender is also a factor; for example, in females sickle cell anemia is associated with higher Hb F concentrations, particularly when associated with the Senegal haplotype.

The kidneys are particularly affected in this disease: there is an increase in glomerular filtration rate early in life, as well as urinary concentration defect, papillary necrosis with hematuria, renal insufficiency and nephrotic syndrome (Plate 14.1). The lungs are the site of the acute chest syndrome that can be caused by infection (viral, bacterial and atypical organisms) as well as fat embolism. The latter gives the most severe cases and has the highest mortality. This life-threatening complication may be arrested by exchange transfusion. Neurological complications also occur (for example, in children classical infarcts occur involving the major vessels) as well as small vessel infarcts. Some of these infarcts, particularly in the frontal lobe, have been associated with behavioral changes. The bones of patients with sickle cell anemia show a five-fold increase in red marrow and may suffer aseptic osteonecrosis, particularly at the site of terminal circulation (head of the femur and humerus) and also in the calcaneus. Peripheral retinopathy characterized by new vessel formation and choroidal infarcts is well recognized in patients with sickle cell disease. Patients with homozygous Hb SS, and some heterozygotes, may suffer hyphema secondary to eye trauma, since sickled cells cannot exit through the Schlemm canals (the vitreous is

very deoxygenated). The spleen is partially ineffectual during the first decade of life and essentially disappears (autosplenectomy), except in the co-presence of $-\alpha/-\alpha$ thalassemia.

The liver can harbor bilirubin stones in the gall bladder as well as the common duct, but fortunately the patients have infrequent cholecystitis. They may also suffer hepatic crises, characterized by increase in serum bilirubin, abnormal liver functional tests, pain and a fall in hemoglobin. Thick bile, due to suspended bilirubinate crystals, can be a factor, as well as sequestration of red cells, a hypermacrophage erythrophagia (Kupffer cells).

The heart may be affected by cardiomyositis, leading to cardiac insufficiency and, in a small number of patients, small vessel obstruction. The lack of more vaso-occlusion in the heart may be the consequence of the very high perfusion pressure of coronary vessels as well as the squeezing effect of the ventricular contraction. Finally, the ankles are the site for leg ulcers in sickle cell anemia, and these may be painful and reduce the quality of life.

In sickle cell anemia, pregnancy poses a significant risk to the mother and newborn (lower birth weight). Previous protocols of exchange transfusion during pregnancy have been replaced by putting the patient into 'high risk track' obstetric care, involving frequent follow-ups, with good results and fewer complications.

The most devastating complication of sickle cell anemia is painful crisis. This involves pain in the extremities, joints, lower back, abdomen, cranium and parotid glands. The pain may be insidious or rapidly progressive, and may begin in one site and extend over time to others. It involves infarction of the circulation in marrow, bone or muscle, or a combination of these. Further studies have demonstrated that the decrease in dense sickle cells is actually preceded by a decrease in light-density sickle cells. The latter event might record light-density sickle cell adhesion while the subsequent decrease in dense cells reflects their trapping in vessels bedecked with adherent cells. Fever is usually present even in the absence of intercurrent infection. Precipitating factors include dehydration, fever, infection, emotional stress, interference with the vascular circulation of limbs, intense exercise and the use of cocaine.

Another very serious complication is pulmonary hypertension. In sickle cell anemia patients with this complication, the average systolic, diastolic and mean pulmonary artery pressures and mean pulmonary capillary wedge pressure were almost double those in sickle cell anemia patients without pulmonary hypertension. The mean level of pulmonary artery pressure was the best predictor of survival, and each increase of 10 mmHg was associated with a 1.7-fold increase in death rate. The median survival of patients with pulmonary hypertension is close to 2 years, which indicates that this is one of the worst complications in sickle cell anemia.

Pathophysiology

There are five major contributors to the pathophysiology of sickle cell anemia.

Polymerization of Hb S

This chemical reaction underlying sickling is a nucleus-mediated reaction. That is, a nucleus of about 10 hemoglobin tetramers has to form first, before the polymerization can start its characteristic exponential course to form sickle fibers. The formation of a nucleus is a hit-and-miss affair, and it takes time to form the right structure. Hence, the reaction has a delay time, after deoxygenation, that is shortened by an increase in intracellular hemoglobin concentration (above the normal 33 g/dl), by lowering the pH and by increasing the temperature. This explains why red cell dehydration and infections are serious in sickle cell anemia patients. The final products of polymerization are double-stranded fibers which lead to red cell deformity with protuberances (Figure 14.6).

Adhesion of sickle red cells to the endothelium

Young sickle red cells have a propensity to adhere to small post-capillary venules. This has been demonstrated both in the mesenteric circulation (*ex vivo*) as well as *in vivo* in the cremaster muscle of sickle transgenic mice. Evidence exists that obstruction of the microcirculation occurs primarily in the venule side, preceded by adhesion of young sickle red cells, and completed by the trapping of dense cells and ISCs in areas bedecked with adhered red cells.

The mechanism of adhesion is not completely understood and is probably multifactorial, but von Willebrand factor and the vibronectin receptor ($\alpha_v\beta_3$) appear to be excellent candidates, since antibodies against either of the members of the pair inhibit adhesion in living circulatory preparations. Other proposed mechanisms have not been fully validated in microcirculatory preparations, nor have they been demonstrated to be operative in small venules.

Red cell dehydration

Some young sickle red cells undergo a fast process of dehydration after they enter the circulation. There are, in effect, intermediate dense cells and very dense cells. It is possible that different mechanisms are involved in each of these types. It is clear that Ca^{2+}-dependent K^+ efflux and probably deoxy-induced K^+ efflux are responsible for the latter, and that K:Cl co-transport overexpression, due to the relative youth of the red cells, is responsible for the former. K:Cl is no longer active in normal adult red cells, while reticulocytes retain a modicum of this system. Finally, cycles of oxy-deoxygenation seem

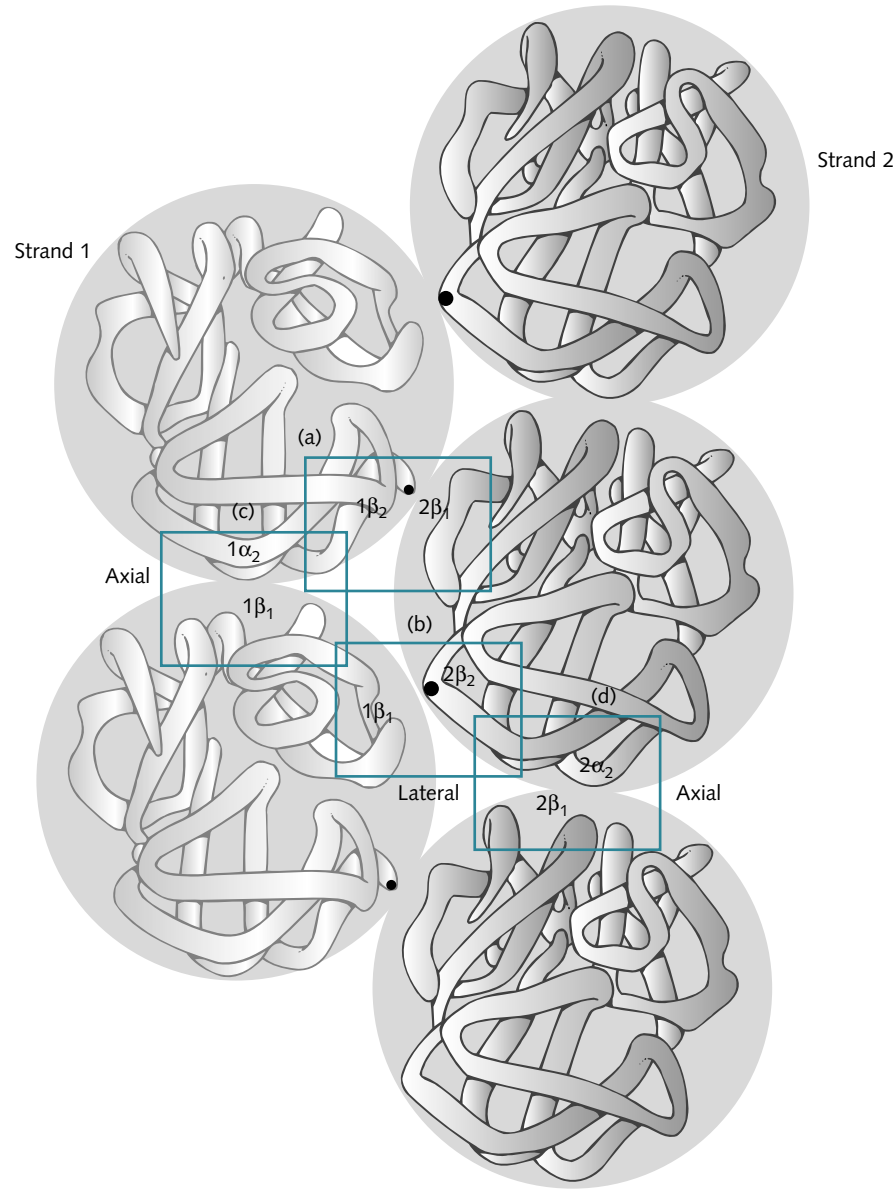


Fig. 14.6 Polymer Hb S structure
This is the assembly unit of the Hb S polymer: the Werner–Love double strand. This double strand is part of the Hb S crystal, but in the crystal it propagates in all directions to form a lattice. The sickle fiber wraps around to form a helical fiber. The squares depict the four different areas of contact that form the double strand.

to be required for dense cell formation. Patients with sickle cell anemia who have high Hb F levels and coexistent α thalassemia have fewer dense cells. Dense cells and ISCs both have very short lives (probably less than 5 days) and have markedly right-shifted oxygen equilibrium curves.

Anemia

Anemia is not a simple product of hemolysis in sickle cell anemia, as demonstrated by the marrow expansion of only five times normal, compared with the 10 times normal seen in thalassemia major. One reason for this is the right shift of the oxygen equilibrium curve described above: this type of oxygen binding curve results in increased oxygen delivery to the tissues. This explains why erythropoietin increases only below 9 g of hemoglobin. In addition, the response to erythropoietin is blunted (that is, it is inappropriately low for the anemia) below the threshold of 9 g of hemoglobin. The reason for this is not understood, but the fact that age increases the deficiency of erythropoietin suggests that 'silent' renal damage might

be involved in a reduction in the erythropoietin response. Another contributor to the anemia (a reduction of erythropoietin response) is the increase in 2,3-diphosphoglycerate (2,3-DPG) found in sickle red cells.

The hemolysis is also complex: actually there are three red cell populations in sickle cell anemia, the very dense red cells, which live only 4–5 days; the F cells (these are sickle cells containing Hb F), which have close to a normal lifespan; and the rest of the cells, which fall in between. Coexistence of α thalassemia and high Hb F reduces hemolysis considerably.

Interestingly, the level of Hb F also defines the properties of progenitors as well as those of the circulating hemopoietic cytokines. Low-Hb F sickle cell anemia patients have an increased number of circulating burst-forming units—erythroid (BFU-E), which are in active cycle. In addition, these patients have constitutively circulating granulocyte macrophage colony-stimulating factor (GM-CSF) and Steel factor, unlike normal individuals and high Hb F sicklers. In contrast, the low-Hb F patients have IL-3 circulating in addition to a hemopoietic inhibitory factor, TNF-α.

Sickle cell anemia patients are susceptible to an acute anemia superimposed on their chronic anemia, due to infection with parvovirus B19. This can produce a life-threatening drop in hemoglobin, which is generally self-limited (7–15 days). This results from its ability to infect BFU-E through the blood group P system that serves as receptor. Treatment is by transfusion of red cells in most cases.

Nitric oxide effects

Arginine, the substrate for nitric oxide (NO) synthase (NOs) in the generation of NO, is low in the plasma of sickle cell anemia patients and also in sickle transgenic mice, suggesting increased need in the production of NO. In transgenic mice, NOs inhibitors result in increases in endothelial NOs and NO activity, leading to lower blood pressure and diminished arteriolar response to NO-mediated vasodilators. Low peripheral resistance has also been observed in sickle cell anemia patients.

Interestingly, recent data suggest that hydroxyurea, used in the treatment of these patients, is involved in the induction of Hb F, one of the ameliorating mechanisms associated with this drug.

Also, NO might be involved in the gender differences in severity observed in this disease. Females live longer and have higher levels of Hb F. Recent data demonstrate that NO availability and NO responsiveness are greater in women than in men with sickle cell disease and regulate the expression of adhesion molecules. Endothelium-dependent blood flows are largely non-NO-mediated in males.

We will return to the subject of NO in the next section (Treatment).

Treatment

Hydroxyurea has been demonstrated to reduce the incidence of painful crises and acute chest syndrome by 50% in adult sickle cell (Hb SS) patients; in addition, hydroxyurea reduces their transfusion requirements. It has become apparent that the rise in Hb F alone does not explain the benefits of hydroxyurea fully, and it is felt that this drug must have other effects; for example, the reduction of sickle cell adhesion. Predictors of good results with hydroxyurea are high white cell count and the absence of the Bantu haplotype.

Bone marrow transplantation has been carried out in many patients with sickle cell anemia with reasonably good results, including the regeneration of the spleen in some cases. Mortality is between 5 and 20% depending on the protocol used. Selection of patients because of a history of central nervous system involvement in children initially produced strokes during convalescence, but adjustment of the protocol has reduced this risk. This procedure is generally restricted to severely ill patients because of the mortality of graft rejection, which occurs in up to 20%.

Preventive measures have been very successful in this disease, particularly in the reduction of infant mortality. These include prophylactic penicillin, pneumococcal vaccine and teaching patients about splenic sequestration.

The judicious use of exchange transfusion in acute chest syndrome, splenic sequestration, aplastic crises and liver crises has saved countless lives.

The treatment of painful crises remains a challenge. The best approach is a rapid assessment, preferably by quantitative instruments, of the intensity, quality and distribution of pain, the presence of comorbidities and the state of dehydration. Rapid, aggressive, individually tailored analgesia treatment, with the additional use of NSAIDs, has proved to be the best route. All of this is almost impossible in an emergency room; hence, a day hospital dedicated to sickle cell anemia is probably the best alternative service model.

Recent developments include therapy involving NO metabolism. As mentioned above, plasma arginine levels are low in sickle cell anemia and in the sickle transgenic mouse model ($\beta^{S+S\text{-}Antilles}$). Hence, an arginine diet supplementation, with a four-fold increase in arginine maintained for several months, was explored. Surprisingly, mean corpuscular hemoglobin concentration (MCHC) decreased and the percentage of high-density red cells was significantly reduced. These outcomes were found to be the consequence of the inhibition of the Ca^{2+}-activated K^+ (Gardos) channel. Clinical trials of arginine supplementation in sickle cell anemia are in the offing. Potential uses of inhaled NO in acute painful crises, acute chest syndrome and pulmonary hypertension are being actively explored.

Allogeneic transplantation is available to some of these patients. The mortality of the procedure is around 10%; hence, with the survival of these patients increasing dramatically the selection of patients for this procedure is not easy. In addition, lack of appropriate marrow donors has been found to be a problem. Miniablation, after initial enthusiasm, has been unsuccessful after the cessation of immune suppression.

Finally, the prospect of gene therapy for hemoglobinopathies has improved considerably due to the realization that the HIV-related lentivirus vector is an efficient means of delivering an anti-sickling or anti-thalassemia construct to the patient's stem cells and then allogeneically transplanting such cells in a partially ablated recipient. This approach has been tested successfully in transgenic mice models of sickle cell anemia and severe β-thalassemia.

For sickle cell anemia, the lentivirus vector contained a β^A globin gene that contains a change in amino acid (Gln→Thr) at position β87, which had previously been demonstrated to account for more than 90% of the Hb F anti-sickling potential. This lentivirus vector was optimized for transfer to hematopoeitic stem cells and also for high expression. Long-term expression was achieved, without preselection, with erythroid-specific accumulation of the anti-sickling protein in up to 52% of total hemoglobin and 99% of circulating red blood cells. In two mouse SCD models, BERK and SAD, inhibition of red blood cell dehydration and sickling was achieved by correction of hematological parameters, splenomegaly, and the prevention of the characteristic urine concentration defect.

Gene therapy of severe thalassemia in transgenic mice has also been achieved, with permanent, panerythroid correction by transplantation of syngeneic bone marrow transduced with an HIV-1-derived (β-globin gene/LCR) lentiviral vector. This was sustained for more than 7 months in both primary and secondary transplants, at which time approximately 95% of the red blood cells in all mice contained human β-globin, contributing to 32 ± 4% of all β-like globin chains. Hemoglobin levels, reticulocyte levels and red blood cell counts approached complete correction. Free β-globin chains were completely cleared from the thalassemic red cell membranes, splenomegaly was abated, and iron deposit was almost entirely eliminated from the liver.

These findings are encouraging for the future of gene therapy of hemoglobinopathies. Nevertheless, unknown challenges might lie ahead.

Sickle/β thalassemia

This syndrome is observed in locations in which both Hb S and β-thalassemia are frequent, such as Africa, Sicily, Greece, Turkey, the Arab countries and the regions of America with African and southern Mediterranean admixture.

The genotype may be S/β^+ thalassemia, in which the red cells contain between 20 and 40% Hb A, the remainder comprising Hb S and Hb F; or S/β^0 thalassemia, in which the red cells contain only Hb S and Hb F. The latter genotype can only be diagnosed by pedigree or genetic analysis.

The clinical picture of S/β^0 thalassemia is very similar to that of sickle cell anemia but it is milder. Anemia may be milder than Hb SS, and in some cases the mean cell volume (MCV) is lower. Retinopathy and osteonecrosis are more common, but autosplenectomy is less common.

The clinical picture of S/β^+ thalassemia is significantly milder than that of sickle cell anemia: there is less anemia, a lower MCV and a lower reticulocyte count but an increased risk of retinopathy and osteonecrosis. The spleen often enlarges during adulthood.

Hemoglobin CC disease

Homozygous CC individuals have a mild hemolytic anemia which is generally asymptomatic and rarely life-threatening. The anemia is moderate, the MCV is reduced (50–60 fl), there is minimal reticulocytosis and splenomegaly is common. The peripheral blood film is fairly characteristic and, with some training of the observer, diagnostic. The red cells are hypochromic, due to their flatness [rather than low mean corpuscular hemoglobin (MCH) or MCHC; in fact, they have a normal MCH and an increased MCHC]. Folded and target cells are prominent. Red cells may show intracellular tetragonal crystals of Hb C, best detected in reticulocyte smears, where they retain their red color.

Regarding the genetics of this disease, the β^C gene (β6 Val→Lys), a single base substitution, has a frequency one-quarter that of the β^S gene among African-Americans. The β^C gene most likely originated in Burkina Faso, where we find the highest frequency (up to 50% of the population), decreasing concentrically, and encompassing Mali, Ivory Coast and Ghana, all of which are east of the Niger river.

The pathophysiology of CC disease is dominated by the high tendency of Hb C to produce oxygenated tetragonal crystals (Plate 14.2). The CC red cells are uniformly denser than AA cells. This effect is associated with very high expression of K:Cl co-transport (higher than in SS cells, in spite of lower reticulocyte count), a transport system that, by extruding K^+ and water, can dehydrate red cells. There is evidence that the kinetics of this transport is altered in CC cells, with slower than normal turn-off, when the stimulus for activity, volume increase or low pH is removed. Whether this alteration is sufficient to explain the pathophysiology of CC cells and

the mechanistic basis for the interaction of Hb C with the membrane remains to be determined.

Recent epidemiological data from Burkina Faso reveals that homozygous CC individuals are particularly resistant to dying of malaria, which suggest that Hb C will eventually supersede the Hb S gene in the populations where the two mutants coexist. This is congruent with the previous findings that Hb CC red cells were not able to release the merozoites as normal red cells or AC red cells.

The multiplication rate of *P. falciparum* is measurably lower in CC cells than in normal red cells, probably due to the ring forms and trophozoites disintegrating within a subset of CC cells. In addition, the knobs present on the surface of infected CC cells are fewer in number and morphologically aberrant when compared with those on AA cells. It appears that only a subset of CC cells supports normal parasite replication. Hb C will eventually predominate due to the greater protection from malaria afforded to their carriers, who will be less likely to die before the reproductive period and hence will be able to transmit the Hb C gene to their descendants more efficiently than the sickle gene.

Hb C nucleates and grows by the attachment of Hb C molecules from the solution, but concurrent amorphous phases, spherulites, and microfibers are not building blocks for the crystal. Hb C crystallization is possible because of the huge entropy gain, likely stemming from the release of up to 10 water molecules per protein intermolecular contact–hydrophobic interaction. The higher crystallization propensity of oxyHb C is attributable to increased hydrophobicity resulting from the conformational changes that accompany the Hb C β6 mutation. The oxy ligand state is thermodynamically driven to a limited number of aggregation pathways with a high propensity to form the tetragonal crystal structure. This is in contrast to the deoxy form of Hb C, which energetically equally favors multiple pathways of aggregation, not all of which might culminate in crystal formation.

The presence of circulating tetragonal crystal-containing red cells in splenectomized CC patients does *not* result in vaso-occlusion, probably because the CC tetragonal crystals are crystals of oxyHb C, which melt as they approach the capillaries. DeoxyHb C has a different crystal form, but not enough of the large crystals to generate pathology.

Hemoglobin SC disease

SC disease is genotypically a double heterozygote; that is, a combination of sickle trait and Hb C trait. Since neither of these trait forms independently has a phenotype, the clinical picture of SC disease requires an explanation (*see below*).

SC disease is milder than sickle cell anemia, particularly in the first 20 years of life, in which the mortality approaches zero. Red cell survival is about 27 days compared with 17 days for sickle cell anemia red cells. Nevertheless, three complications are inappropriately severe: osteonecrosis, retinopathy and acute chest syndrome. SC patients frequently retain their spleen in adulthood, and have less anemia, lower MCV, lower reticulocyte count, fewer very dense and dense red cells and lower Hb F levels, on average, than sickle cell anemia patients. Their survival also seems to be better.

Like CC cells, the blood film in SC disease may be diagnostic. In addition to the cells common to CC disease (flat and apparently hypochromic cells), they exhibit target cells, and more specifically a few ISCs, as well as intracellular tetragonal crystals, better seen in reticulocyte smears (Plate 14.3).

The pathophysiology of SC red cells derives from the contribution that Hb C makes to the red cell: like CC cells (*see above*), SC cells are denser than AA or most of the sickle cell anemia red cells (Plate 14.4). The increase in MCHC implicit in this effect promotes the polymerization of Hb S, leading to more sickling than expected for a cell with only 50% Hb S. Interestingly, Hb S in turn favors Hb C crystallization; this explains, in addition to the differences in splenic activity, why crystals are generally more prominent in SC disease than in CC. Fortunately, for the reasons expressed above, this situation does not lead to an increase in vaso-occlusion beyond that which is predictable from the solubility of the available Hb S with the MCHC for SC cells.

Homozygous Hb E and Hb E/β thalassemia

The homozygote for Hb E has the phenotype of β thalassemia trait because this abnormal hemoglobin, common from the Eastern provinces of India to the Philippines, is both a mutation of the sequence of the β chains and a thalassemia, due to the generation of an alternative splicing site by the mutation.

Populations living near the common border of Cambodia, Laos and Thailand (the Khmer people) have the highest incidence of this abnormal hemoglobin, the selection pressure for which is resistance to infection by *P. falciparum* malaria.

The clinical syndrome is very mild, with no or minimal anemia (once the nutritional causes of anemia are treated) and low MCV, with target cells in the smear and hypochromia. Density gradients reveal that although the red cells are small they have a normal MCHC, as a consequence of a diminished MCH, due to the thalassemic component of this disease. People who inherit Hb E and β-thalassemia trait present a thalassemia major or intermedia picture (*see Chapter 1*).

The protective effect of Hb E against *P. falciparum* malaria was assessed in a mixed erythrocyte invasion assay demonstrating that the parasite preferentially invaded normal red cells compared with abnormal Hb AE, EE, E-β-thalassemia red cells. The heterozygote AE cells were the worst target, with invasion restricted to approximately 25% of the red cells. Hb AE might have an unidentified membrane abnormality that renders the majority of the red cell population relatively resistant to invasion by *P. falciparum*. This will reduce parasitemia and hence will reduce the lethality of the infection and is consistent with the Haldane hypothesis of heterozygote protection against severe malaria for Hb E.

Preliminary evidence of iron loading in affected patients with Hb β-thalassemia in Sri Lanka suggests variable but accelerated gastrointestinal iron absorption. The iron loading associated with chronic transfusions in patients with Hb E-β-thalassemia is similar to that observed in patients with β-thalassemia. These data, which represent the only cohort of patients with Hb-β-thalassemia to have undergone quantitative assessment of body iron burden, suggest that that guidance for the assessment of iron loading and initiation of chelating therapy in patients with β-thalassemia may be also applicable to those with Hb E-β-thalassemia. Further quantitative studies in both non-transfused and transfused patients will be necessary to settle this issue definitively.

Dominant sickle mutations

Dominant sickle mutations refer to β gene mutations that include the β^S mutation and exhibit a second mutation in the same chain. This turns the chains into super-Hb S, which produces symptoms in the heterozygote. Two instances have been characterized. Hb S-Antilles (β6 Glu→Val; 23 Val→Ile) is expressed in the heterozygote at about the 40% level and produces a syndrome resembling a mild sickle cell anemia phenotype. The mechanism is complex; the second mutation increases the solubility of deoxyHb S from 17 g/dl to about 11 g/dl, with the result that Hb S polymerizes more readily and more extensively. The second mutation has an oxygen equilibrium effect that favors sickling.

The other dominant Hb S mutation is Hb S-Oman (β6 Glu→Val; 121 Glu→Lys), which generates an even more powerful super-S because, even at expression levels of about 20% (resulting from concomitant $-\alpha/-\alpha$), it has a phenotype very similar to Hb S-Antilles. Since the two dominant forms of Hb S have the same solubility in the deoxy state, the more severe phenotype present implies that the 121 second mutation (identical to that in Hb O_{Arab}) must produce pathology of its own. This hypothesis is confirmed by the hemolytic anemia present in individuals homozygous for Hb O_{Arab} or Hb $G_{Philadelphia}$.

Unstable hemoglobins

There are about 100 mutations that render the hemoglobin molecule unstable (e.g. Hb Köln) but their incidence is low. They are generally inherited in an autosomal dominant manner. Mutations may render hemoglobin unstable and produce a hemolytic syndrome by four mechanisms:

- when the mutation introduces either a bulky or a charged side chain in the interior of this globular protein
- when the mutation introduces, in the α-helix portion of the molecule, a side chain that is not α-helix-friendly
- when the mutation destabilizes the heme attachment to the globin
- when the mutation interferes with the stability of the contact area of $\alpha_1\beta_1$, a dimer that normally does not dissociate.

Hb Köln (β 98Val→Met) is the most common of all the unstable hemoglobins, and affects all ethnic groups. It is difficult to diagnose by electrophoresis because of its instability and indistinct electrophoretic pattern. Two light bands are commonly observed because of the severe instability. The best test is isopropanol solubility, in which the hemolysate turns opaque, and a precipitate can be observed after centrifugation.

The patients, who can be of any ethnic origin, exhibit a mild to moderate hemolytic anemia and the presence of pigmenturia (mostly dipyrroles). The mechanism of the hemolytic anemia is the intracellular release of the hemes, which leaves very unstable tetramers, the formation of hemochromes and subsequently Heinz bodies, and the recognition of the abnormal red cells with hemoglobin/hemochromes attached to membranes by the spleen and other macrophages. Thrombosis is a recognized complication after splenectomy.

There is no need for treatment except for folic acid supplementation and attention to the possibility of aplastic crises due to parvovirus B19. Rarely, these patients require splenectomy.

High oxygen affinity hemoglobins

The sigmoid curve of the oxygen binding of hemoglobin is the product of the low affinity of the deoxy tetramers (T state) (e.g. Hb Chesapeake), which turn into the R state when two of the four hemes are oxygenated. After this molecular switch the hemoglobin acquires high affinity for oxygen. Mutations that stabilize the R state of the hemoglobin and interfere with the switch to T state will tend to have higher affinity than normal. Hence, the receptor mechanism for hypoxia in the kidney will interpret this situation as evidence of the presence of hypoxia and respond with increased secretion of erythropoietin, and

stimulate an increase in the number of red cells and hematocrit.

There are around 100 high-affinity hemoglobins which are, on the whole, rare. The first one described was Hb Chesapeake, a mutation of position 92 in the β chain. This mutation destabilizes the T state, so the conformational state of Hb Chesapeake is biased to the R state, and hence has increased affinity for oxygen. No serious clinical consequences of this mutation have been found apart from lifelong erythrocytosis.

The most common mutations responsible for high-affinity hemoglobins are those that interfere with the R→T transition: mutations in the $\alpha_2\beta_2$ contact area, and in the switch region of the molecule, particularly the C and N terminals, favoring the R state. In addition, mutations that interfere with the 2,3-DPG binding site in the central cavity will tend to increase the affinity of hemoglobin for oxygen.

Low oxygen affinity hemoglobins

Hemoglobins with low oxygen affinity (e.g. Hb Kansas, Hb Beth Israel) deliver oxygen more efficiently than normal hemoglobin, hence they tend to be detected as generating hyperoxia. The patient becomes anemic through a correction of the level of erythropoietin secretion. Nevertheless, if the shift to the right is far enough, there is a point at which the delivery of oxygen is normal again and no anemia exists. However, patients with Hb Kansas and Hb Beth Israel have clinically apparent cyanosis, since they have more than 5 g of circulating deoxyhemoglobin. Except for cyanosis, no other abnormalities have been found in these patients. The diagnosis is important in order to avoid unnecessary and sometimes invasive investigations. Diagnosis is by electrophoresis (helpful only if the amino acid substitution affects the overall charge of the molecule) and measurement of the oxygen dissociation curve.

The oxygen delivery properties of high- and low-affinity hemoglobins are discussed in Figure 14.7.

Hemoglobin Ms

These hemoglobins are the product of the substitutions of either the proximal or distal histidine by tyrosine. Also a nearby mutation, as in Hb M Milwaukee, may produce a similar picture. The patients appear slate grey in color, which may be confused with cyanosis (this is really pseudocyanosis), due to the increase in the deoxyhemoglobin (>5 g/dl) or methemoglobin status of the mutated chains as well as the change of the electrical environment, and hence of the visible absorption spectrum (Figure 14.8). The visible spectra of the hemolysate may help confirm the diagnosis, although electrophoresis can be helpful. Differential diagnosis includes genetic or acquired methemoglobinemia and sulfhemoglobinemia.

These mutations have been found worldwide, and they are rare, except in the Iwate prefecture in Japan where they are common.

Conclusions

Genetically abnormal hemoglobins present themselves to the

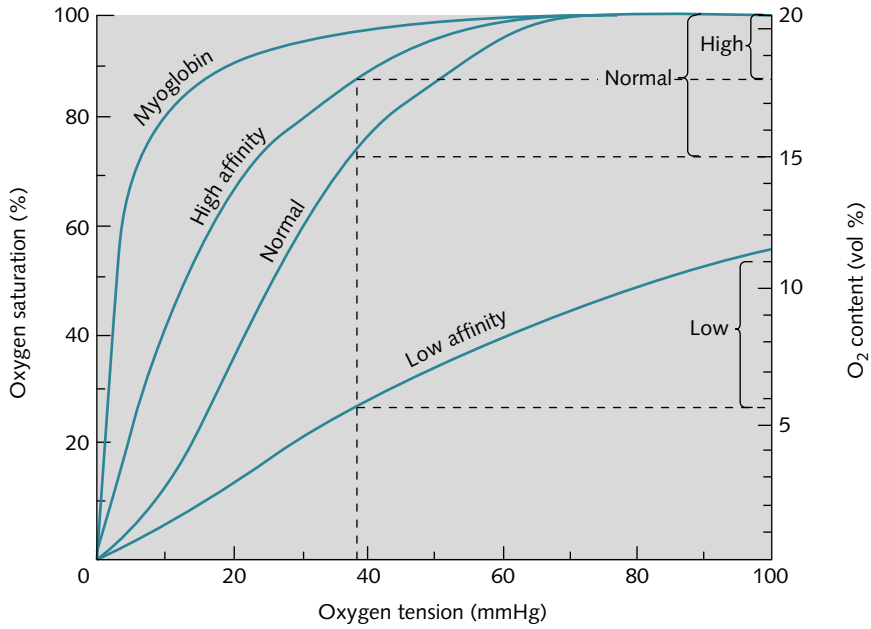


Fig. 14.7 Oxygen binding curves of hemoglobins with high and low oxygen affinity
Notice that the extraction of oxygen by the tissues, which is the difference between pulmonary oxygen pressure (100 mmHg) and capillary oxygen pressure (40 mmHg), is lower than normal in high-affinity hemoglobins (low p50) and higher than normal in low-affinity hemoglobins. This is why the former have erythrocytosis and the latter (most of the time) anemia.

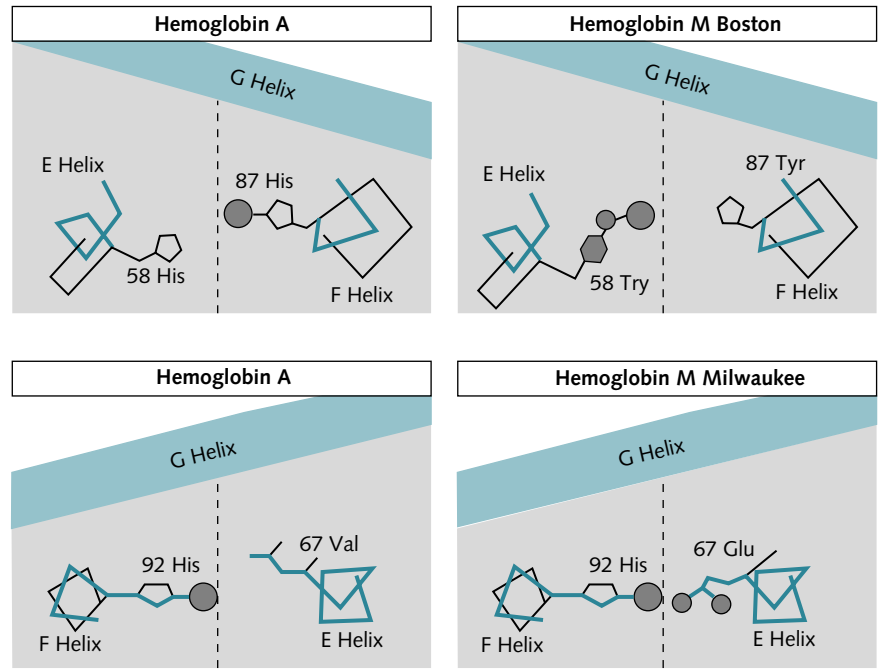


Fig. 14.8 Structure of Hb Ms
Heme environment of two Hb Ms compared with Hb A. Notice the diverse changes in the distal and proximal tyrosines (which have replaced the normal histidines) in each of the Hb Ms. These hemoglobins have a characteristic visible spectrum around 610 nm.

clinician usually in one or more of the following syndromes: (1) the presence of the homozygous or double heterozygous state of Hb S produces a picture of chronic anemia, chronic and insidious organ damage punctuated by painful crises and other complications; (2) the presence of Hb C in the homozygous state produces a mild chronic hemolytic syndrome; when doubly heterozygous with Hb S (SC disease), a syndrome similar to sickle cell anemia but milder, and characterized by microcytic anemia, is seen; (3) the presence of homozygous Hb E produces mild microcytic hemolytic anemia, but in combination with β-thalassemia produces a picture sometimes of very severe thalassemia intermedia; (4) heterozygous forms of *unstable hemoglobins* produce variable-intensity hemolytic anemias; (5) heterozygotes for *low oxygen affinity mutant hemoglobins* and Hb Ms may produce a mild hemolytic anemia syndrome, but always with cyanosis or pseudocyanosis as a background; (6) heterozygous forms of *high oxygen affinity mutant hemoglobins* often present with erythrocytosis.

This panoply of clinical syndromes is the product of the following alterations of the hemoglobin molecule:

- the creation of a new property for the hemoglobin molecule (Hb S, polymerization; Hb C, crystallization and microcytosis)
- changes in O_2 affinity (high and low affinity for ligands)
- changes in the environment of the heme (Hb Ms)
- changes in the stability of the molecule in solutions (unstable hemoglobins)
- some mutated hemoglobins are produced at lower rates and generate a thalassemic syndrome (Hb E).

Of all of these, only Hb S, Hb C and Hb E are frequent in populations at risk, due to their selection by malaria; the others are rare and generally represent private mutations.

Further reading

General

Aldrich T, Nagel RL. (1998) The pulmonary complications of sickle cell disease. In: Bone RC, Dantzker DR, George RB *et al.* (eds). *Pulmonary and Critical Care Medicine*, 5th edition. New York: Mosby.

Eaton WA, Hofrichter J. (1990) Sickle cell hemoglobin polymerization. [Review]. *Advances in Protein Chemistry*, **40**, 63–279.

Henry ER, Jones CM, Hofrichter J *et al.* (1997) Can a two-state MWC allosteric model explain haemoglobin kinetics? [Review]. *Biochemistry*, **36**, 6511–6528.

Henry ER, Bettati S, Hofrichter J *et al.* (2002) A tertiary two-state allosteric model of hemoglobin. *Biophysical Chemistry*, **98**, 198–164.

Nagel RL. (1995) Disorders of hemoglobin function and stability. In: Handon RI, Lux SE, Stossel TP (eds). *Blood: Principles and Practices of Hematology*. Philadelphia: J.B. Lippincott.

Nagel RL, Roth EF Jr. (1989) Malaria and red cell genetic defects. *Blood*, **74**, 1213–1221.

Steinberg M, Higgs D, Forget B *et al.* (eds). (2000) *Disorders of Haemoglobin*. Cambridge: Cambridge University Press.

Thein SL. (1998) Hematological diseases. In: Jameson JL (ed.). *Principles of Molecular Medicine*. Totowa, NJ: Humana Press.

Sickle hemoglobin

Castro O, Hoque M, Brown BD. (2003) Pulmonary hypertension in sickle cell disease: cardiac catheterization results and survival. *Blood*, **101**, 1257–1261.

Gaziev J, Lucarelli G. (2003) Stem cell transplantation for hemoglobinopathies. *Current Opinion in Pediatrics*, **15**, 24–31.

Halsey C, Roberts IA. (2003) The role of hydroxyurea in sickle cell disease. *British Journal of Haematology*, **120**, 177–186.

Herrick JB. (1910) Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. *Archives of Internal Medicine*, **6**, 517.

Imren S, Payen E, Westerman KA *et al.* (2002) Permanent and panerythroid correction of murine beta thalassemia by multiple lentiviral integration in hematopoietic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 14380–14385.

Kaul DK, Fabry ME, Nagel RL. (1996) The pathophysiology of vascular obstruction in the sickle syndromes. *Blood Reviews*, **10**, 29–44.

Kaul DK, Liu XD, Fabry ME *et al.* (2000) Impaired nitric oxide-mediated vasodilation in transgenic sickle mouse. *American Journal of Physiology. Heart and Circulatory Physiology*, **278**, H1799–H1806.

Morris CR, Morris SM Jr, Hagar W *et al.* (2003) Arginine therapy: a new treatment for pulmonary hypertension in sickle cell disease? *American Journal of Respiratory and Critical Care Medicine*, **168**, 63–69.

Nagel RL, Fleming AF. (1992) Genetic epidemiology of the β^S gene. In: Fleming AF. (ed.). *Baillière's Clinical Haematology, Volume 5*. London: Harcourt Brace Jovanovich, pp. 331–365.

Nagel RL, Johnson J, Bookchin RM *et al.* (1980) Beta-chain contact sites in the haemoglobin S polymer. *Nature*, **283**, 832–834.

Pawliuk R, Westerman KA, Fabry ME *et al.* (2001) Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science*, **294**, 2368–2371.

Prengler M, Pavlakis SG, Prohovnik I *et al.* (2002) Sickle cell disease: the neurological complications. [Review]. *Annals of Neurology*, **51**, 543–552.

Reiter CD, Gladwin MT. (2003) An emerging role for nitric oxide in sickle cell disease vascular homeostasis and therapy. *Current Opinion in Hematology*, **10**, 99–107.

Romero JR, Suzuka SM, Nagel RL *et al.* (2002) Arginine supplementation of sickle transgenic mice reduces red cell density and Gardos channel activity. *Blood*, **99**, 1103–1108.

Vichinsky E. (2002) New therapies in sickle cell disease. [Review]. *Lancet*, **360**, 629–631.

Hb C hemoglobin

Dewan JC, Feeling-Taylor A, Puius YA *et al.* (2002) Structure of mutant human carbonmonoxyhemoglobin C (betaE6K) at 2.0 A resolution. *Acta Crystallographica Section D, Biological Crystallography*, **58**, 2038–2042.

Fabry ME, Romero JR, Suzuka SM *et al.* (2000) Hemoglobin C in transgenic mice: effect of HbC expression from founders to full mouse globin knockouts. *Blood Cells, Molecules & Diseases*, **26**, 331–347.

Fairhurst RM, Fujioka H, Hayton K *et al.* (2003) Aberrant development of *Plasmodium falciparum* in hemoglobin CC red cells: implications for the malaria protective effect of the homozygous state. *Blood*, **101**, 3309–3315.

Lawrence C, Nagel RL. (2001) Compound heterozygosity for Hb S and HB C coexisting with AIDS: a cautionary tale. *Hemoglobin*, **25**, 347–351.

Modiano D, Luoni G, Sirima BS *et al.* (2001) Haemoglobin C protects against clinical *Plasmodium falciparum* malaria. *Nature*, **414**, 305–308.

Nagel RL, Lawrence C. (1991) The distinct pathobiology of SC disease: therapeutic implications. In: Nagel RL (ed). *Hematology/Oncology Clinics of North America*. Philadelphia: W.B. Saunders, pp. 433–451.

Olson JA, Nagel RL. (1986) Synchronized cultures of *P falciparum* in abnormal red cells: the mechanism of the inhibition of growth in HbCC cells. *Blood*, **67**, 997–1001.

Vekilov PG, Feeling-Taylor AR, Petsev DN *et al.* (2002) Intermolecular interactions, nucleation, and thermodynamics of crystallization of hemoglobin C. *Biophysics Journal*, **83**, 1147–1156.

Hb E hemoglobin

Chotivanich K, Udomsangpetch R, Pattanapanyasat K *et al.* (2002) Hemoglobin E: a balanced polymorphism protective against high parasitemias and thus severe *P falciparum* malaria. *Blood*, **100**, 1172–1176.

Fucharoen S, Winichagoon P. (2000) Clinical and hematologic aspects of hemoglobin E beta-thalassemia. [Review.] *Current Opinion in Hematology*, **7**, 106–112.

Olivieri NF, De Silva S, Premawardena A *et al.* (2000) Iron overload and iron-chelating therapy in hemoglobin E-beta thalassemia. *Journal of Pediatric Hematology Oncology*, **22**, 593–597.

Schrier SL. Pathophysiology of thalassemia. [Review.] (2002) *Current Opinion in Hematology*, **9**, 1231–1226.

Chapter 15 Molecular coagulation and thrombophilia

Björn Dahlbäck & Andreas Hillarp

Introduction

The risk of venous thrombosis is increased when the hemostatic balance between pro- and anticoagulant forces is shifted in favor of coagulation. If this is caused by an inherited defect, the resulting hypercoagulable state conveys a lifelong increased risk of thrombosis. Inherited resistance to activated protein C (APC) is the most common hypercoagulable state found associated with venous thrombosis. It is caused by a single point mutation in the factor V (FV) gene, predicting the substitution of Arg^{506} with Gln. The FV gene mutation (FV Leiden or FV:Q^{506}) confers a 5- to 10-fold increased risk of thrombosis and is found in 20–60% of Caucasian patients with thrombosis. Another common inherited risk factor for thrombosis is a point mutation (G20210A) in the 3′ untranslated region of the prothrombin gene. This mutation is present in approximately 1–4% of healthy individuals and is associated with an approximately three-fold increased risk of thrombosis. Other less common genetic risk factors for thrombosis are the deficiencies of natural anticoagulant proteins, such as antithrombin, protein C and protein S. Such defects are present in fewer than 1% of healthy individuals and together account for 5–10% of genetic defects found in patients with venous thrombosis. Owing to the high prevalence of FV Leiden and of the G20210A mutation in the prothrombin gene, combinations of genetic defects are relatively common in the general population. As each genetic defect is an independent risk factor for thrombosis, individuals with multiple defects have a highly increased risk of thrombosis, and multiple defects are consequently often found in patients with thrombosis.

Blood coagulation

At sites of vascular damage, circulating platelets adhere to subendothelial structures and undergo a series of reactions, which lead to primary hemostasis due to the formation of a platelet plug. Concomitant to these events, the subendothelial membrane protein tissue factor (TF) is exposed to blood. A small amount of activated factor VII (FVIIa), present in circulating blood, binds to TF and triggers a series of proteolytic reactions which culminate in the formation of thrombin and the conversion of fibrinogen to insoluble fibrin.

FVIIa bound to TF specifically cleaves and activates the two vitamin K-dependent plasma proteins, factor IX (FIX) and factor X (FX) (Plate 15.1). Activated FX (FXa) activates prothrombin to thrombin, whereas activated FIX (FIXa) activates FX. Both FIXa and FXa are poor enzymes that require protein cofactors, calcium ions and negatively charged phospholipid surfaces for the expression of their full biological activity. The protein cofactors for FIXa and FXa are the activated forms of factor VIII (FVIIIa) and factor V (FVa), respectively. As a result of multiple protein–protein and protein–phospholipid interactions, enzymatically highly efficient complexes are assembled on the phospholipid surface.

The initiation of blood coagulation by TF is usually referred to as the *extrinsic pathway* or the *TF pathway*. In association with injury, this is the physiologically most important mechanism of blood coagulation. However, coagulation can also be activated through the *intrinsic pathway*, which is triggered by the activation of the contact phase proteins (FXII, FXI, prekallikrein and high molecular weight kininogen) that follows upon exposure of blood to certain negatively charged surfaces. The intrinsic pathway does not appear to be physiologically important for injury-related coagulation *in vivo*; this is illustrated by the lack of bleeding problems in individuals with deficiency of FXII.

Thrombin generated at sites of vascular injury expresses a number of procoagulant properties. It amplifies the coagulation process by activating FXI and in addition it activates

platelets and converts fibrinogen to fibrin. Moreover, in a positive feedback reaction, thrombin converts the procofactors FV and FVIII into their biologically active counterparts (FVa and FVIIIa).

Regulation of blood coagulation

The efficient reactions of the coagulation system have considerable biological potential and strict regulation is required. For this purpose, several plasma proteins and protein–cell interactions are involved in the constant monitoring of the circulation. At each level of the coagulation pathway, membrane-bound molecules expressed on the surface of intact endothelial cells, circulating inhibitors and negative feedback mechanisms provide efficient control.

Antithrombin (AT) is the most important serine protease inhibitor (serpin) involved in the regulation of blood coagulation. AT inhibits thrombin as well as FXIa, FIXa and FXa and, under certain conditions, also FVIIa. AT forms a highly stable complex with the protease and, as a consequence, the protease is trapped and eliminated from the circulation. The activity of AT is stimulated by heparin, which accelerates the rate of formation of the AT–protease complexes. Under normal physiological conditions, heparan sulfate proteoglycans present on the endothelial cell surface stimulate the activity of AT, whereas heparin injections are used in clinical situations. During inhibition of thrombin, an important role of heparin is to function as a bridge between thrombin and AT. In addition, heparin induces conformational changes in AT, which are associated with the generation of a more efficient inhibitor. In the inhibition of FXa, the conformational change appears to be more important than the bridging mechanism.

The TF pathway is regulated by the TF pathway inhibitor (TFPI). TFPI is composed of three protease inhibitory domains belonging to the Kunitz type of inhibitors. TFPI has the unique capacity to inhibit the FVIIa–TF–FXa complex and is therefore highly efficient in turning off the TF pathway. The inhibition mediated by TFPI occurs in two steps. The first step is inhibition of FXa by the middle Kunitz domain; the first Kunitz domain then binds and inhibits FVIIa. Most of the TFPI is bound to glucosaminoglycans on endothelial cells (approximately 80%) and only a minor fraction of TFPI is present in plasma, where it is mainly associated with low-density lipoproteins.

The highly efficient procoagulant reactions of thrombin are physiologically adequate at sites of vascular injury and are instrumental in efficient hemostasis. However, the same reactions pose a threat to the organism as uncontrolled coagulation leads to thrombus formation. Nature has solved this dilemma in intricate and fascinating ways, one of which is the transformation of thrombin into an efficient initiator of a natural anticoagulant pathway, the protein C system. The conversion of thrombin from a procoagulant into an anticoagulant enzyme depends on the presence of intact endothelium. Thus, thrombin generated at sites of intact vasculature binds to the endothelial membrane protein thrombomodulin, which is a potent modulator of thrombin activity and a cofactor to thrombin in the activation of protein C (Plate 15.2). A recently discovered receptor for protein C, the endothelial protein C receptor (EPCR), has been shown to stimulate the activation of protein C by the thrombin–thrombomodulin complex. APC degrades membrane-bound FVa and FVIIIa by limited proteolysis in reactions which are potentiated by a cofactor protein designated protein S and, in the case of FVIIIa degradation, also by the non-activated form of FV (Plate 15.3).

Under physiological conditions, pro- and anticoagulant mechanisms are balanced in favor of anticoagulation, whereas the anticoagulant system is downregulated and procoagulant forces prevail at sites of vascular damage. Defects in this ingenious system are associated with increased thrombin generation, a hypercoagulable state, leading to an increased risk of thrombosis.

The protein C anticoagulant system

Protein C is a vitamin K-dependent plasma protein, which is synthesized mainly in the liver. It is homologous to FVII, FIX and FX and shares with them a common modular organization. From the N-terminus, these proteins contain a vitamin K-dependent γ-carboxyglutamic acid (Gla)-rich module, two epidermal growth factor (EGF)-like modules and a serine protease module. The Gla domains bind calcium ions and provide the vitamin K-dependent clotting proteins with phospholipid-binding properties. Upon activation by the thrombin–thrombomodulin complex, the serine protease module is converted to an active enzyme. APC is highly specific in its proteolytic activity, cleaving a limited number of peptide bonds in FVa and FVIIIa.

Intact FV is a high molecular weight protein and shares with the homologous FVIII molecule the modular arrangement A1, A2, B, A3, C1, C2. Upon activation of FV by thrombin or FXa, peptide bonds surrounding the B module are cleaved, and the B module is not part of FVa. FVIII is activated by thrombin in a similar fashion, which leads to release of the B module. APC cleaves three peptide bonds in FVa, at Arg^{306}, Arg^{506} and Arg^{679}, whereas FVIIIa is cleaved at Arg^{336} and Arg^{526}. As a consequence of the APC-mediated cleavages, FVa and FVIIIa lose their procoagulant properties.

APC alone has poor anticoagulant activity and it is only in the presence of its cofactors, protein S and FV, that efficient anticoagulant function is expressed. This was demonstrated in an experimental system based on the degradation of FVIIIa. In this system, it was found that full anticoagulant activity of

APC was obtained in the presence of the combination of FV and protein S. The synergistic APC cofactor activity of FV requires APC-mediated proteolysis of at least the Arg^{506} cleavage site. This is important for the understanding of the mechanism of APC resistance, being the result of the Arg^{506}→Gln mutation (FV Leiden; *see below*). The APC cofactor activity of FV appears specific for the degradation of FVIIIa, whereas the FVa degradation is unaffected by this FV activity. FV loses its APC cofactor activity upon proteolysis by thrombin, but it gains procoagulant properties as a cofactor to FXa. Thus, FV is similar to thrombin in being able to express both pro- and anticoagulant effects. However, whereas the anticoagulant effects of thrombin depend on its binding to thrombomodulin, the anticoagulant properties of FV are dependent on APC-mediated proteolysis of the non-activated form of FV. The detailed mechanisms by which FV functions as an APC cofactor remain to be elucidated.

Protein S is also a vitamin K-dependent plasma protein but, unlike the other vitamin K-dependent coagulation proteins, it is not a serine protease. It is a multimodular protein containing a Gla module, a thrombin-sensitive module, four EGF-like modules and a large module homologous to sex hormone-binding globulin (SHBG) (Figure 15.4). Protein S also has functions outside the protein C system and 60–70% of protein S in plasma circulates bound to C4b-binding protein (C4BP), a regulator of the complement system. The Gla-module of protein S provides both free protein S and the protein S–C4BP complex, with phospholipid-binding ability. This is important for the localization of coagulation and complement regulatory activities to certain cell membranes—for example, to the phosphatidyl serine that is exposed on apoptotic cells. Protein S binding to such cells has been shown to be involved in stimulation of phagocytosis of these cells.

During the degradation of free FVa (i.e. not bound to FXa) by APC, the cleavage at Arg^{506} is faster than that at Arg^{306}. The cleavage at Arg^{506} leads only to partial loss of FVa activity, whereas the cleavage at Arg^{306} leads to efficient inactivation of FVa. Protein S serves as cofactor for the cleavage at Arg^{306} but has minor effects on the Arg^{506} cleavage. This, together with a specific protection of the Arg^{506} site exerted by FXa, indicates that the Arg^{306} site is the most important site for the regulation of FVa activity in the prothrombinase complex. On the other hand, FVa, which is not part of a prothrombinase complex, is first cleaved at the Arg^{506} site, because the kinetics of this cleavage are more favorable than those for the cleavage site at Arg^{306}. In *in vitro* experiments, protein S has been shown to express an anticoagulant activity that is independent of the presence of APC. The exact mechanism is unknown but has been suggested to be related to inhibition of prothrombin activation through direct interactions of protein S with FVa, FXa and the phospholipid membrane. The *in vivo* physiological significance of this APC-independent anticoagulant activity is unclear. Regardless of its mode of action, protein S is an important anticoagulant protein *in vivo*, as demonstrated by animal studies and by the association between protein S deficiency and venous thrombosis.

Molecular genetics of venous thromboembolism

The annual incidence of venous thrombosis in Western societies is approximately 1–2 per 1000. Thrombotic episodes tend to occur in conjunction with surgery, fractures, pregnancy, the use of oral contraceptives and immobilization. In addition, genetic defects are frequently involved and many patients report positive family histories. Genetic defects known to predispose to thrombosis include inherited APC resistance due to FV Leiden, a point mutation in the prothrombin gene (G20210A) and deficiencies of anticoagulant protein C, protein S or AT.

Factor V gene mutation (FV Leiden) causing APC resistance

In 1993, APC resistance was described as a cause of inherited thrombophilia and it was soon demonstrated to be highly prevalent (20–60%) among thrombosis patients. In APC resistance, APC does not give a normal prolongation of the clotting time. In more than 95% of cases the molecular defect associated with APC resistance is a single point mutation in the FV gene. The mutation is a G→A substitution at nucleotide position 1691 in the FV gene, which predicts replacement of Arg^{506} with a Gln. The mutant FV is known as $FVR^{506}Q$, FV Leiden or F: Q^{506} (R and Q are one-letter codes for Arg and Gln, respectively).

The FV Leiden allele is found only in Caucasians, and the prevalence of the mutant FV Leiden allele in the general population of Western societies demonstrates considerable variation. High prevalence (up to 15%) is found in southern Sweden, Germany, Greece, Arab countries and Israel. In the Netherlands, the UK and the USA around 3–5% of the population carry the mutant allele. Lower prevalence (around 2%) is found in Hispanics. The high prevalence of the FV Leiden allele in certain populations suggests a possible survival advantage, and there is a reduced risk of bleeding after delivery in women carrying the mutation. In the history of mankind, the slightly increased risk of thrombosis associated with the FV Leiden allele has presumably not been a negative survival factor because thrombosis develops relatively late in life and does not influence fertility. In addition, many of the circumstantial risk factors for thrombosis, such as a sedentary life style, surgery and the use of oral contraceptives, did not affect our ancestors.

The high prevalence of the FV Leiden allele in Western societies is the result of a founder effect. It has been estimated that the mutation event was around 30 000 years ago, i.e. after the 'Out of Africa March', which took place 100 000 years ago and also after the separation of the Asians from the Europeans. This explains why the mutant FV allele is common among European populations but is not present among Japanese, Chinese, and the original populations of Africa, Australia and America.

A large number of studies have demonstrated relationships between the presence of APC resistance (FV Leiden) and an increased risk of venous thrombosis. Differences in the selection criteria of patients and in the prevalence of the mutant allele in the general population explain the wide variation in results obtained from different studies. However, the general consensus is that the FV Leiden allele is the most common genetic risk factor for venous thrombosis in Western societies. The odds ratio, describing the increased risk of thrombosis in affected individuals, has been calculated to be six- to eight-fold for those carrying the defect in a heterozygous form, whereas homozygous individuals are at 30- to 140-fold increased risk of thrombosis. The FV Leiden allele does not appear to be a strong risk factor for arterial thrombosis, such as myocardial infarction. Two mutations affecting the Arg^{306} site have recently been found in thrombosis cases, FV Cambridge and FV Hong Kong, but such mutations appear to be rare. They do not result in APC resistance and are not major risk factors for thrombosis.

The FV Leiden allele is associated with a hypercoagulable state, which is reflected in increased levels of prothrombin activation fragments in the plasma of individuals with inherited APC resistance. Two molecular mechanisms are involved (Figure 15.1). In one of these mechanisms, an APC cleavage site in FVa is lost, which impairs the normal degradation of FVa by APC. The other surprising observation is that FV Leiden is a poor APC cofactor in the degradation of FVIIIa because the cleavage at Arg^{506} is required for expression of APC cofactor activity of FV.

In the degradation of normal FVa, the APC cleavage at Arg^{506} has favorable kinetics compared with cleavages at other sites. The Arg^{506} cleavage is approximately 10-fold faster than the cleavage at Arg^{306} and the activity of FVa : Q^{506} (FVa Leiden) is therefore inhibited at an approximately 10-fold lower rate than FVa : R^{506}. Generated FVa Leiden persists longer than normal FVa and can form active prothrombinase complexes with FXa. However, degradation of free FVa (i.e. FVa not bound to FXa) is different from that of FVa, which is part of the prothrombinase complex. In the prothrombinase complex, the Arg^{506} site is protected by FXa from degradation by APC. In addition, protein S functions as an APC cofactor primarily for the Arg^{306} cleavage. As a consequence, APC-mediated degradation of FVa, which is part of the prothrombinase complex,

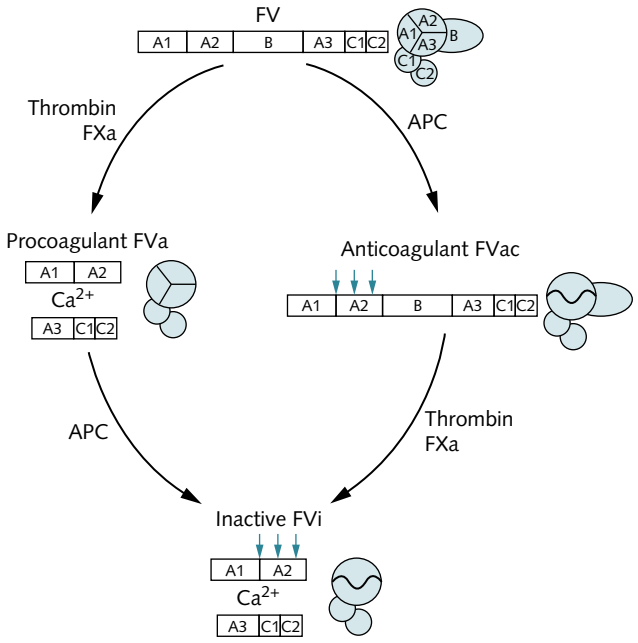


Fig. 15.1 Pro- and anticoagulant properties of factor V
Proteolytic modification of single-chain factor V (FV) results in the expression of either pro- or anticoagulant properties. Thrombin and FXa cleaves and activates FV to a procoagulant (FVa) that functions as a cofactor to FXa in the activation of prothrombin. Intact FV is sensitive to cleavage by APC, which recruits FV into an anticoagulant path. FV modified by APC (FVac) functions as a synergistic APC cofactor with protein S in the degradation of FVIIIa. The anticoagulant properties of FV are lost upon further proteolysis by thrombin or FXa. Likewise, the procoagulant effects of FVa are lost as a result of cleavage by APC. Thus, FV plays a crucial and central part in balancing pro- and anticoagulant forces. Arrowheads denote the three APC cleavage sites.

follows a different pathway compared with that of free FVa. Therefore, when FVa : R^{506} and FVa: Q^{506} are part of assembled prothrombinase complexes, the rates of their degradation by APC plus protein S are similar.

Laboratory investigation of inherited APC resistance due to the FV Leiden allele can be done both with a functional APC-resistance test and with molecular biology assays. A modified APC-resistance test involving dilution of the patient's plasma in FV-deficient plasma is highly sensitive and specific for the presence of the FV Leiden allele. The most commonly used molecular assay for FV Leiden involves polymerase chain reaction (PCR) amplification and restriction enzyme digestion.

Deficiency of antithrombin

Heterozygous AT deficiency is found in between 0.02 and 0.05% of the general population and in 1–2% of thrombosis patients, suggesting that the genetic defect is associated with a 10- to 20-fold increased risk of thrombosis—somewhat higher

than estimated for APC resistance. AT deficiency may be of either type I or type II. Type I deficiency is characterized by low levels of both immunological and functional AT, whereas type II denotes functional defects. Type II cases are divided into three subtypes: RS (reactive site mutants), HBS (heparin binding site mutants) and PE (mutants giving pleiotropic effects). A large number of AT deficiencies have been genetically analyzed (Figure 15.2). In most cases, the genetic defect is a point mutation, a small deletion or an insertion. Partial or whole-gene deletions are relatively uncommon causes of AT deficiency. Type II RS variants are defective in protease inactivation, and mutations in the vicinity of the reactive site have been found. The type II HBS deficiency carries mutations in the heparin binding site and type II PE AT variants are caused by a limited number of mutations between amino acids 402 and 429.

Protein C deficiency

Heterozygous deficiency of protein C is identified in 2–5% of thrombosis patients. The prevalence of protein C deficiency in the population is estimated to be approximately 0.3%. The 10-fold higher prevalence of protein C deficiency in thrombosis cohorts suggests that carriership is associated with a 10-fold increased risk of venous thrombosis, that is, a risk essentially similar to that associated with APC resistance. Protein C deficiency is not a risk factor for arterial thrombosis. Two types of protein C deficiency have been described. In type I, there is a parallel reduction in protein C antigen and functional activity. Type II is characterized by a functional defect in the protein, and its plasma concentration may be normal. The majority of reported cases of protein C deficiency are of type I. Homozygous or compound heterozygous protein C deficiency is a rare condition (1/200 000–1/400 000) that leads to severe and fatal thrombosis in the neonatal period. The clinical picture is that of purpura fulminans and the symptoms include necrotic skin lesions due to microvascular thrombosis. Other major symptoms are thrombosis in the brain and disseminated intravascular coagulation. Several cases have been successfully treated with fresh frozen plasma or with protein C concentrates.

Genetic analysis has been performed in a large number of cases with protein C deficiency (160 different mutations are known). Most genetic defects are missense mutations located within the region coding for the mature protein, which lead to single amino acid substitutions and type I deficiency (Figure 15.3). Mutations in the promoter region of the gene, which affect the plasma protein concentration, and mutations affecting the RNA splicing have also been found. In a minority of cases, the genetic defects lead to a type II deficiency. Mutations leading to type II deficiency have been found in almost all the modules of protein C, including the propeptide, the Gla module, EGF1, the activation peptide, and the serine protease domain.

Protein S deficiency

Heterozygous protein S deficiency is present in 1–5% of thrombosis patients. The prevalence of protein S deficiency

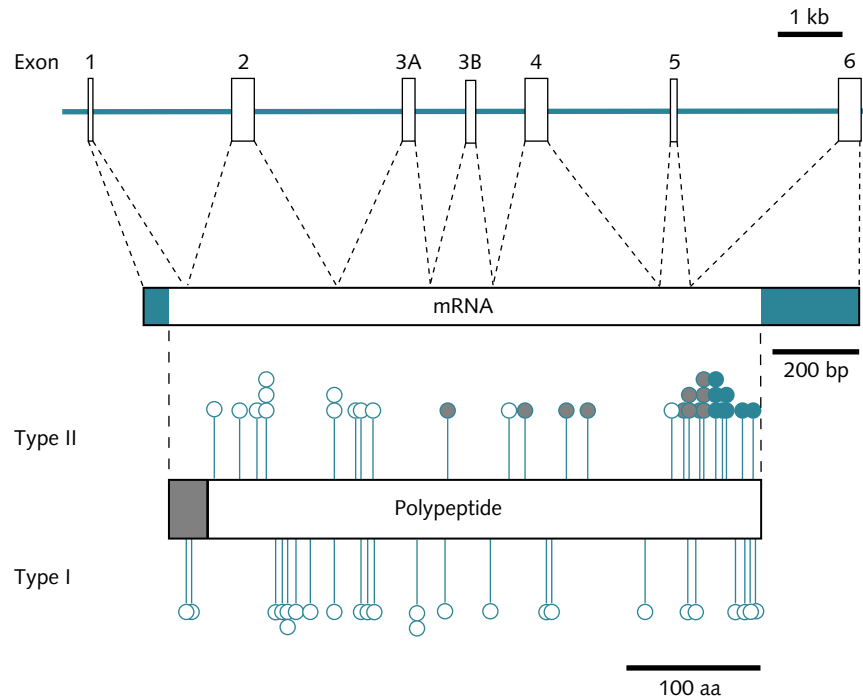


Fig. 15.2 Structure of the human antithrombin gene and locations of detrimental missense mutations in the antithrombin molecule
The gene for antithrombin (AT) is localized to chromosome 1 (1q23–q25) and spans 13.4 kb of DNA (upper part). It comprises seven exons and results in an mRNA of 1.7 kb (middle part). The AT molecule (lower part) is synthesized as a single polypeptide chain composed of a mature protein containing 432 amino acid residues, and a signal peptide (shaded) of 32 amino acid residues. Many mutations of different types causing AT deficiency have been described. Shown here are only missense mutations leading to amino acid substitutions associated with type I deficiency (indicated by open circles below the polypeptide chain) or type II deficiency. Open, shaded and filled circles denote type II HBS, type II RS and type II PE variants, respectively.

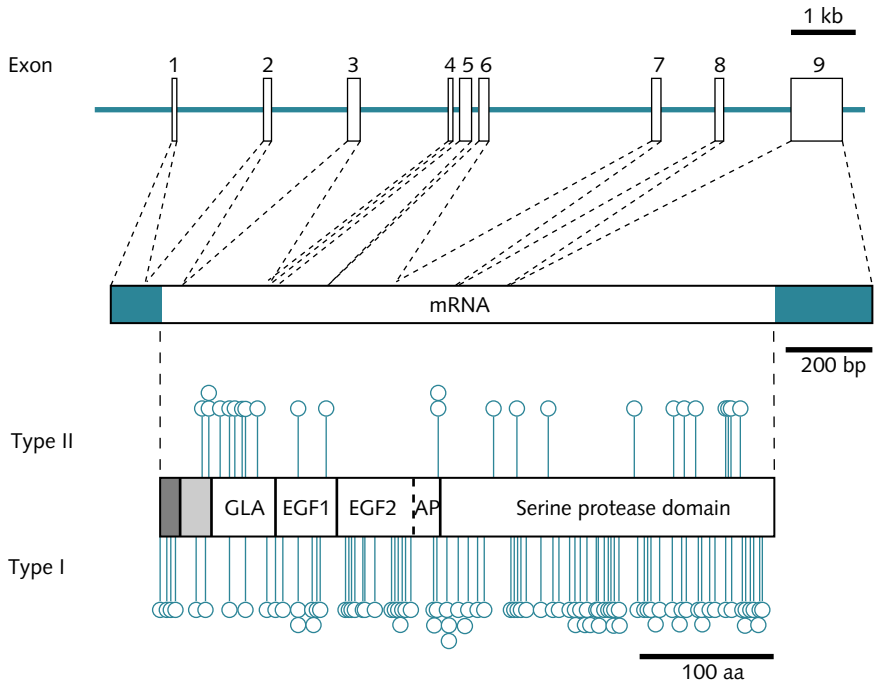


Fig. 15.3 Structure of the human gene and locations of detrimental missense mutations in the protein C molecule
Human protein C is encoded by the *PROC* gene, localized to chromosome 2 (2q13–q14), which spans approximately 11 kb of DNA (upper part). The gene comprises nine exons, which yield a ~1.8 kb mRNA transcript (middle part). The protein C mRNA encodes a prepro-protein C sequence of 461 amino acid residues (lower part). The pre-sequence (shaded) serves as a signal peptide and the pro-sequence (light shading) functions as the signal for proper γ-carboxylation of the protein. The mature protein consists of 419 residues and can be divided into a γ-carboxy glutamic acid (Gla) domain, two epidermal growth factor (EGF) domains and a serine protease domain. During processing of the protein, an internal dipeptide is removed from the protein and the mature protein circulates as a covalently linked two-chain molecule. Between the second EGF domain and the protease part of the molecule is an activation peptide (AP) region, which is released upon protein C activation. The circles indicate the locations of known missense mutations, leading to amino acid substitutions associated with type I deficiency (indicated below the polypeptide chain) or type II deficiency (above the polypeptide chain).

in the general population has been estimated to be between 0.03 and 0.13%. Family studies suggest that heterozygous carriers have a 5- to 10-fold increased risk of thrombosis compared with their healthy relatives, which is similar to that associated with protein C deficiency and APC resistance. The level of free protein S discriminates better between people with and without protein S deficiency than the level of total protein S. This is because the concentrations of protein S and C4BPβ$^+$, which is the protein S-binding isoform of C4BP, are equimolar in protein S-deficient individuals and most of the protein S is bound to C4BPβ$^+$. Protein S deficiency with low levels of both free and total protein S is called type I, whereas protein S deficiency with low free protein S and normal total protein S has been believed to constitute a separate genetic type (type III). However, coexistence of the two types in many protein S-deficient families demonstrates that they represent different phenotypic variants of the same genetic disease. Mutations in protein S leading to functional defective molecules are referred to as type II deficiency. To date, very few type II deficiencies have been found, which presumably is related to the poor diagnostic performance of available functional protein S assays. Homozygous protein S deficiency is extremely rare, but appears to give a picture similar to that of homozygous protein C deficiency, with purpura fulminans in the neonatal period. To date, more than 140 mutations in the protein S gene have been reported. Most of the gene defects are missense or nonsense mutations, and mutations affecting splicing or insertion/deletion defects are less common (Figure 15.4). Because of the large size of the protein S gene and the presence of a closely linked and highly similar pseudogene, the identification of mutations is not easy. With current molecular biology techniques, a genetic approach would be too costly for routine use. Furthermore, in some families with phenotypically established protein S deficiency, protein S gene mutations are not found. The reason for the difficulty in identifying protein S gene mutations in some families is unclear.

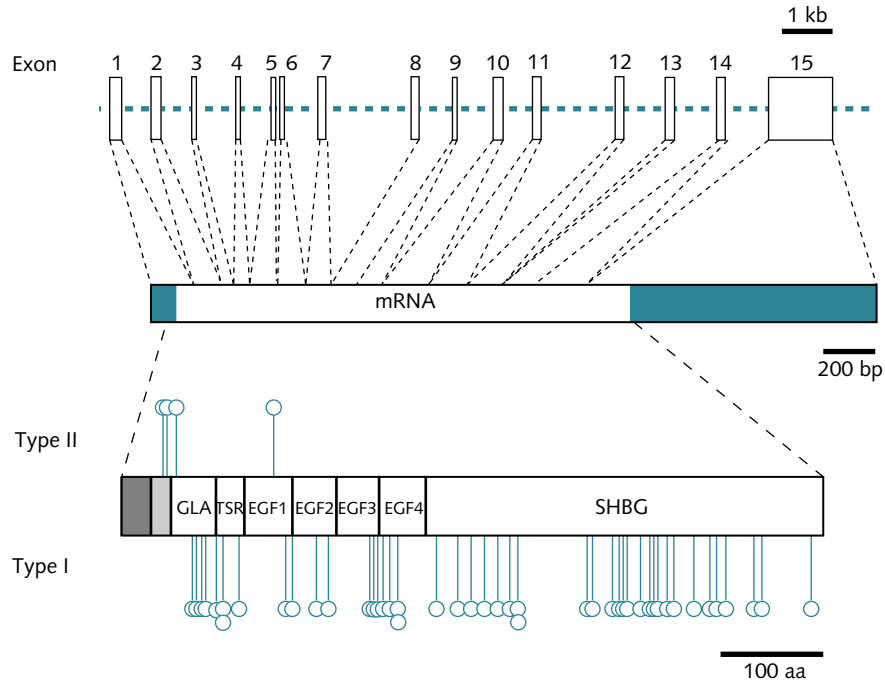


Fig. 15.4 Structure of the human protein S gene and location of detrimental missense mutations in the protein S molecule
The gene for human protein S (*PROS1*) comprises 15 exons (upper part of the figure), spans over 80 kb of DNA and is localized to chromosome 3 (3p11.1–q11.2). Exons are denoted by open bars and introns by lines. Introns denoted by dashed lines between exons indicate gaps and are not drawn to scale. The *PROS1* mRNA is approximately 3.5 kb in size (middle part of the figure). The mRNA is translated into a 676 amino acid residue prepro-protein S (lower part of the figure). The polypeptide chain can be divided into a signal peptide (dark grey), a pro-peptide (light grey), a thrombin-sensitive region (TSR), a γ-carboxyglutamic acid (Gla) domain, four epidermal growth factor (EGF)-like domains and a large carboxy-terminal domain homologous to sex hormone-binding globulin (SHBG). The circles indicate the localizations of known missense mutations that lead to amino acid substitutions associated with type I deficiency (indicated below the polypeptide chain) or type II deficiency (above the polypeptide chain).

Prothrombin mutation

A point mutation in the prothrombin gene (nucleotide 20210 G→A) has been identified as the second most common independent risk factor for venous thrombosis. The mutation is located in the 3′ untranslated region and the mechanism by which this mutation leads to an increased risk of thrombosis is not fully understood, even though it has been shown that the mutation is associated with increased plasma levels of prothrombin (Figure 15.5). The prevalence of the mutation in the general population is heterogeneous and dependent on geographical location and ethnic background. By analogy with the FV Leiden mutation, the prothrombin gene mutation is mainly found in populations of Caucasian origin. In southern Europe the prevalence is 2–4%, nearly as twice as high as that in northern Europe. Founder effects are the likely explanation for the differences in the distribution of the prothrombin 20210 G→A mutation. The mutation is found in 6–16% of patients with unselected deep venous thrombosis and carriers have an approximately three- to four-fold increased risk of thrombosis.

Severe thrombophilia is a multigenic disease

Venous thrombosis is a typical multifactorial disease, involving one or more environmental and/or genetic risk factors. In Western societies many individuals carry more than one genetic risk factor because the FV Leiden allele is so common. In contrast, in countries where the FV Leiden allele is rare few individuals carry more than one genetic defect. This may explain the difference in the incidence of thromboembolic disease between Japan and China on the one hand and Europe and the USA on the other. The frequency of individuals carrying two or more genetic defects can be calculated on the basis of the prevalence in the general population of the individual genetic defects. In a country where the prevalence of FV Leiden is 10%, combinations of protein C deficiency and FV Leiden are expected to be present in between 1 in 3000 and 1 in 10 000 individuals. A similar calculation for the combination of the prothrombin mutation and FV Leiden allele suggest the prevalence of combined defects to be 1–2 per 1000 individuals. Thus, a large number of people carry more

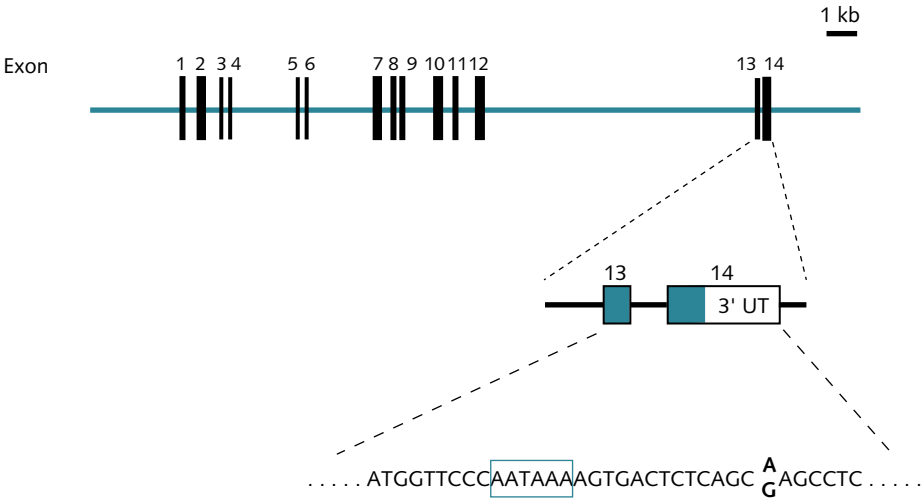


Fig. 15.5 Structure of the human prothrombin gene and localization of the prothrombin 20210 G→A mutation
The human gene for prothrombin (*F2*) comprises 14 exons and spans approximately 20 kb of DNA on chromosome 11 (11p11–q12). The nucleotide sequence flanking the G→A transition at nucleotide 20210 (indicated in bold) in the 3′ untranslated region of the *F2* gene is shown below. The putative polyadenylation signal is boxed. The 20210A allele has been reported to be associated with elevated levels of plasma prothrombin and an increased risk of venous thrombosis.

than one genetic defect and such individuals have considerably increased risk of thrombosis. The FV Leiden allele is thus found to be an additional genetic risk factor in certain thrombophilic individuals with deficiency of protein C, protein S or AT, as well as in cases with the prothrombin mutation (Figure 15.6). In a pooled analysis of the two most common defects, heterozygosity for the FV Leiden and prothrombin 20210 G→A mutations, the combined risk of venous thrombosis was 20-fold compared with a risk of 4.9 and 3.8 for the two single defects respectively.

The thrombotic tendency in individuals with inherited genetic defects is highly variable and some individuals never get thrombosis, whereas others develop recurrent severe thrombotic events at an early age. This depends on the particular genotype, the coexistence of other genetic defects and the presence of environmental risk factors, such as oral

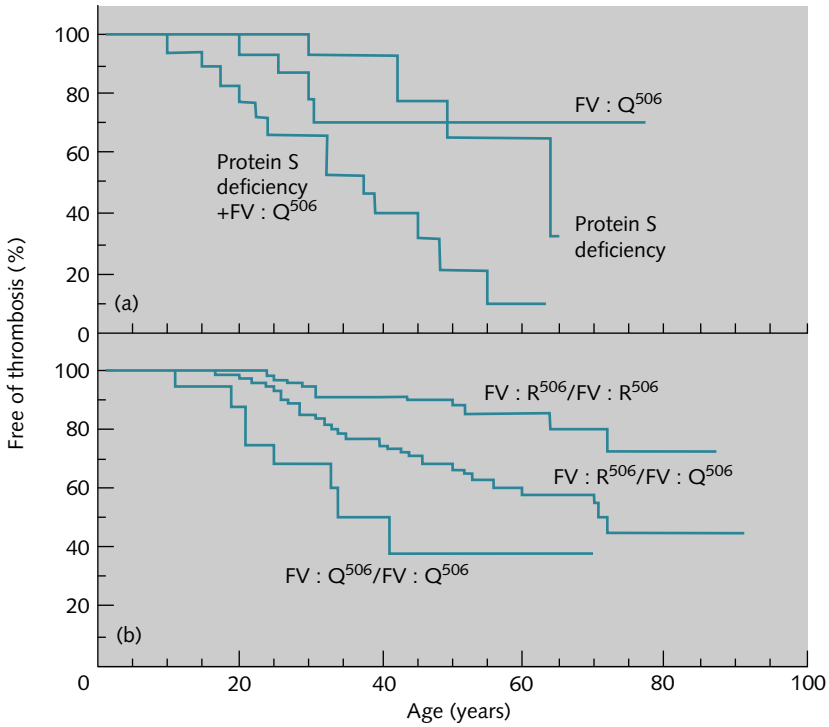


Fig. 15.6 Thrombosis-free survival curves for individuals with different *FV* genotypes and co-inherited protein S deficiency
(a) The increased risk of thrombosis with combined defects is illustrated by thrombosis-free survival curves for 21 individuals with single defects, FV Leiden or protein S deficiency, and 18 individuals with both defects. There was no significant difference between the two groups with single defects, whereas differences between the groups with either the FV Leiden allele or protein S deficiency and the group with combined defects were significant.
(b) Probability of being free of thrombotic events at a certain age for 146 normal individuals, 144 heterozygotes and 18 homozygotes for the FV Leiden allele (Kaplan–Meier analysis). Highly significant differences were observed between normals and heterozygotes and between heterozygotes and homozygotes.

contraceptives, trauma, surgery and pregnancy. Thus, women with heterozygosity for the FV Leiden allele who also use oral contraceptives have been estimated to have a 35- to 50-fold increased risk of thrombosis, whereas those with homozygosity have a several hundred-fold increased risk.

Management of thrombophilia

Decisions about medical intervention due to the presence of one or more genetic defects should be based on careful consideration of the clinical picture, including the patient's family history. The risk of bleeding complications due to anticoagulant therapy must always be weighed against the benefits of the anticoagulation effect, especially if an oral anticoagulant is used for periods exceeding 3–6 months, after which the risk of thrombotic recurrence probably declines. New clinical data are continually emerging and no general consensus regarding the screening, prophylaxis and treatment of symptomatic patients has yet been established.

When the FV Leiden allele is present in homozygous form, or when heterozygosity is combined with a second genetic defect, prophylactic treatment with heparin or oral anticoagulants is recommended in situations known to be associated with a high risk of thromboembolic complications, such as surgery or pregnancy, even if the patient has never experienced any thrombosis or has no family history of such complications. For heterozygous, asymptomatic carriers lacking a family history of thrombosis, short-term prophylaxis has been recommended in high-risk situations, but it remains to be established whether prophylaxis should be given in all situations associated with a risk of thrombosis.

Treatment of symptomatic heterozygous patients should be initiated as for any other patient with thrombotic events. It is not known whether the presence of a genetic defect is associated with an increased risk of recurrence, even though most studies on APC resistance tend to suggest that this is indeed the case. Patients with combined defects and probably also patients with single gene defects may be at increased risk of recurrence and should accordingly be given extended anticoagulation therapy beyond 6 months, even after an isolated thromboembolic event. However, more data are needed before these recommendations can be considered generally applicable.

The potential benefits of general screening for APC resistance and/or the FV Leiden allele prior to thrombotic events or in the presence of such circumstantial factors as oral contraceptive usage, pregnancy and surgery are obvious, but more prospective data are needed, not least in terms of cost–benefit ratios, before any general recommendations can be made.

Conclusions

Inherited APC resistance, caused by the FV Leiden mutation, is the most common genetic risk factor for thrombosis identified to date. The mutated FV has normal procoagulant properties, but the loss of the APC cleavage site at position 506 in FV results in impaired regulation of coagulation and a hypercoagulable state. The prevalence of FV Leiden in Caucasian populations varies between 2 and 15%. A genetic variant in the prothrombin gene (G20210A) is another common prothrombotic risk factor, with a prevalence of approximately 1–4% in the general population. Other, less common independent genetic risk factors include abnormalities in the genes for antithrombin, protein C and protein S. Families with thrombophilia present with variable penetrance of thrombosis that may be explained by different combinations of genetic defects and environmental risk factors. Patients with combined genetic defects are at higher risk of thrombosis than those with single gene defects. Thus, a detailed laboratory investigation is an important component of the evaluation process and needs to be performed in order to estimate the risk of thrombosis in each case.

Antithrombin, protein C and protein S mutation databases on the Internet

Antithrombin mutation database: *http://www.med.ic.ac.uk/divisions/7/antithrombin/default.htm*

Protein C mutation database: *http://www.xs4all.nl/~reitsma/Prot_C_home.htm*

Protein S mutation database: *http://archive.uwcm.ac.uk/uwcm/mg/search/120721.html*

Further reading

Blood coagulation: introduction and regulation

Dahlbäck B. (2000) Blood coagulation. *Lancet*, **355**, 1627–1332.

Dahlbäck B, Stenflo J. (2000) The protein C anticoagulant system. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW *et al.* (eds) *The Molecular Basis of Blood Diseases*, 3rd edn. Philadelphia: W.B. Saunders, pp. 614–656.

Esmon CT. (1992) The protein C anticoagulant pathway. *Arteriosclerosis and Thrombosis*, **12**, 135–145.

Venous thromboembolism: mechanisms and risk factors

Dahlbäck B, Carlsson M, Svensson PJ. (1993) Familial thrombophilia due to a previously unrecognized mechanism characterized by poor anticoagulant response to activated protein C: prediction of a cofactor to activated protein C. *Proceedings of National Academy of Sciences of the United States of America*, **90**, 1004–1008.

Franco RF, Reitsma PH. (2001) Genetic risk factors of venous thrombosis. *Human Genetics*, **109**, 369–384.

Lane D, Mannucci PM, Bauer KA *et al.* (1996) Inherited thrombophilia: Part 1. *Thrombosis and Haemostasis*, **76**, 651–662.

Lane D, Mannucci PM, Bauer KA *et al.* (1996) Inherited thrombophilia: Part 2. *Thrombosis and Haemostasis*, **76**, 824–834.

Reitsma PH. (2000) Genetic heterogeneity in hereditary thrombophilia. *Haemostasis*, **30** Supplement 2, 1–10.

Rosendaal FR. (1997) Risk factors for venous thrombosis: prevalence, risk, and interaction. *Seminars in Hematol*ogy, **34**, 171–187.

Zöller B, Garcia de Frutos P, Hillarp A *et al.* (1999) Thrombophilia as a multigenic disease. *Haematologica*, **84**, 59–70.

APC resistance and factor V Leiden

Bertina RM, Koeleman BP, Koster T *et al.* (1994) Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature*, **369**, 64–67.

Dahlbäck B. (1997) Resistance to activated protein C as risk factor for thrombosis: molecular mechanisms, laboratory investigation, and clinical management. *Seminars of Hematology*, **34**, 217–234.

Dahlbäck B. (1999) Procoagulant and anticoagulant properties of coagulation factor V: factor V Leiden (APC resistance) causes hypercoagulability by dual mechanisms. *Journal of Laboratory and Clinical Medicine*, **133**, 415–422.

Koster T, Rosendaal FR, de Ronde H *et al.* (1993) Venous thrombosis due to poor anticoagulant response to activated protein C: Leiden Thrombophilia Study. *Lancet*, **342**, 1503–1506.

Lindqvist PG, Svensson PJ, Dahlbäck B *et al.* (1998) Factor V R506Q mutation (activated protein C resistance) associated with reduced intrapartum blood loss—a possible evolutionary selection mechanism. *Thrombosis and Haemostasis*, **79**, 69–73.

Nicolaes GAF, Dahlbäck B. (2002) Factor V and thrombotic disease. Description of a Janus-faced protein. *Arteriosclerosis, Thrombosis, and Vascular Biology*, **22**, 530–538.

Svensson PJ, Dahlbäck B. (1994) Resistance to activated protein C as a basis for venous thrombosis. *New England Journal of Medicine*, **330**, 517–522.

Zivelin A, Griffin JH, Xu X *et al.* (1997) A single genetic origin for a common Caucasian risk factor for venous thrombosis. *Blood*, **89**, 397–402.

Antithrombin

Kottke-Marchant K, Duncan A. (2002) Antithrombin deficiency. Issues in laboratory diagnosis. *Archives of Pathology and Laboratory Medicine*, **126**, 1326–1336.

van Boven HH, Lane DA. (1997) Antithrombin and its inherited deficiency states. *Seminars in Hematology*, **34**, 188–204.

Protein C

Reitsma PH. (1997) Protein C deficiency: from gene defects to disease. *Thrombosis and Haemostasis*, **78**, 344–350.

Reitsma PH, Bernardi F, Doig RG *et al.* (1995) Protein C deficiency: a database of mutations, 1995 update. On behalf of the Subcommittee on Plasma Coagulation Inhibitors of the Scientific and Standardization Committee of the ISTH. *Thrombosis and Haemostasis*, **73**, 876–889.

Protein S

Aiach M, Borgel D, Gaussem P *et al.* Protein C and protein S deficiencies. *Seminars in Hematology*, **34**, 205–216.

Anderson HA, Maylock CA, Williams JA, Paweletz CP, Shu H, Shacter E. (2002) Serum derived protein S bind to the phosphatidylserine and stimulates the phagocytosis of apoptotic cells. *Nature Immunology*. Published online: doi:10.1038/ni871.

Gandrille S, Borgel D, Sala N *et al.* (2000) Protein S deficiency: a database of mutations—summary of the first update. For the Plasma Coagulation Inhibitors Subcommittee of the Scientific and Standardization Committee of the ISTH. *Thrombosis and Haemostasis*, **84**, 918.

Rezende SM, Lane DA, Zöller B *et al.* (2002) Genetic and phenotypic variability between families with hereditary protein S deficiency. *Thrombosis and Haemostasis*, **87**, 258–265.

Simmonds RE, Zöller B, Ireland H *et al.* (1997) Genetic and phenotypic analysis of a large (122-member) protein S-deficient kindred provides an explanation for the familial coexistence of type I and type III plasma phenotypes. *Blood*, **89**, 4364–4370.

Webb JH, Blom A, Dahlbäck B. (2002) Vitamin K-dependent protein S localizing complement regulator C4b-binding protein to the surface of apoptotic cells. *Journal of Immunology*, **169**, 2580–2586.

Prothrombin 20210 G→A mutation

Emmerich J, Rosendaal FR, Cattaneo M *et al.* (2001) Combined effect of factor V Leiden and prothrombin 20210 A on the risk of venous thromboembolism. Pooled analysis of 8 case-control studies including 2310 cases and 3204 controls. *Thrombosis and Haemostasis*, **86**, 809–816.

Poort SR, Rosendaal FR, Reitsma PH *et al.* (1996) A common genetic variation in the 3′-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood*, **88**, 3698–3703.

Rosendaal FR, Doggen CJM, Zivelin A *et al.* (1998) Geographic distribution of the 20210 G to A prothrombin variant. *Thrombosis and Haemostasis*, **79**, 706–708.

Chapter 16 The molecular basis of hemophilia

Paul L F Giangrande

Introduction: clinical features of hemophilia

Hemophilia is a congenital disorder of coagulation and affects approximately 1 in 10 000 males worldwide, with approximately 5000 patients with hemophilia in the United Kingdom. Hemophilia A is due to a deficiency of factor VIII in the circulating blood, and hemophilia B (also known as Christmas disease) is a clinically identical disorder caused by factor IX deficiency. Typical laboratory findings in hemophilia include a normal prothrombin time but prolonged activated partial thromboplastin time. The platelet count and bleeding time are normal. A specific factor assay is required to confirm the diagnosis.

The clinical severity (phenotype) is critically determined by the level of circulating factor VIII (or IX) in the blood, and severe hemophilia is defined by a clotting factor level of <1 IU/dl (Table 16.1). The hallmark of severe hemophilia is recurrent and spontaneous hemarthrosis. Typically, hinge joints such as the knees, elbows and ankles are affected, but bleeds may also occur in the wrist or shoulder. Bleeding into the hip joint is unusual. The affected joint is swollen and warm, and held in a position of flexion (Figure 16.1), with no external discoloration or bruising around the joint. It is unusual for an infant to suffer spontaneous hemarthroses in the first few months of life, and the first joint to be affected tends to be the ankle as the child learns to crawl. The first sign of a hemarthrosis in an infant will often be obvious discomfort and distress, accompanied by limping or reluctance to use a limb. Recurrent bleeds into a joint lead to synovitis and joint damage resulting in crippling arthritis (Figure 16.2). Bleeding into muscles is also a feature of hemophilia, but this is usually a consequence of direct injury, albeit often minor (Figure 16.3).

Table 16.1 The relation of blood levels of factor VIII (or IX) to the severity of hemorrhagic manifestations.

Level (IU/dl)	Hemorrhagic manifestations
50–100	Normal level, no bleeding problems
25–50	No problems in day-to-day life. Tendency to bleed after major surgery
5–25	Mild hemophilia. Bleeding typically occurs only after significant injury
2–5	Moderately severe hemophilia; occasionally apparently spontaneous bleeds. Most bleeds associated with injury, albeit often relatively minor
<1	Severe hemophilia with spontaneous and recurrent bleeding into muscles and joints

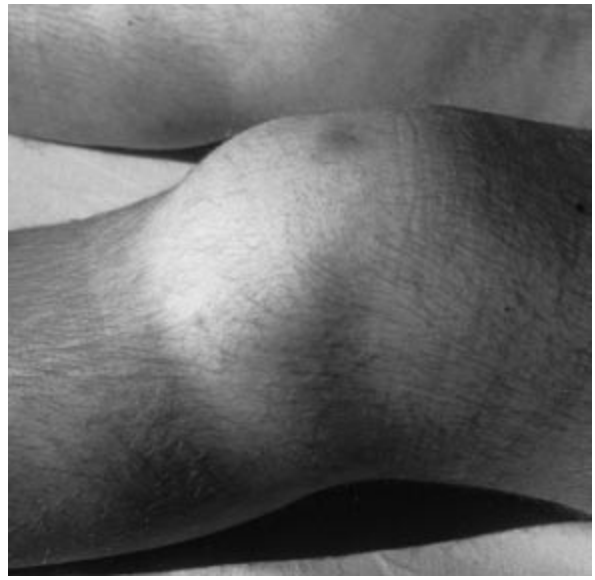


Fig. 16.1 Acute hemarthrosis in severe hemophilia
This usually arises in the absence of injury. The joints most frequently involved are the knees, elbows and ankles. The joint is swollen, warm and tender but there is no external bruising or discoloration.

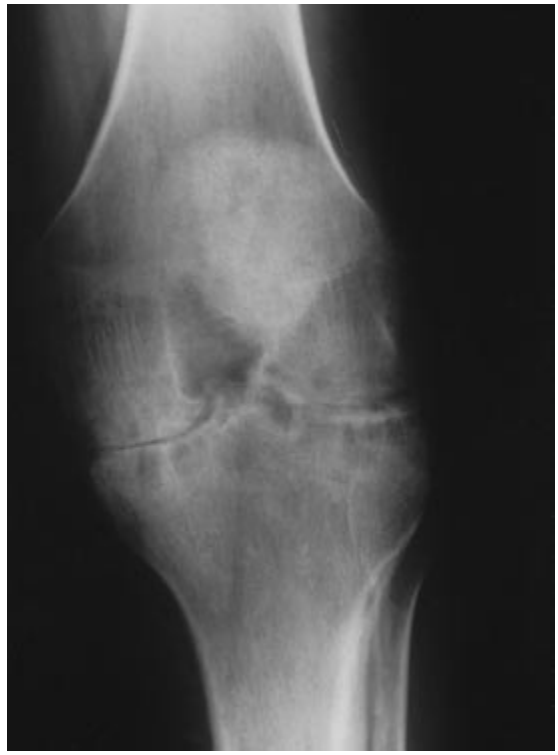


Fig. 16.2 Radiograph of the knee of a patient with severe hemophilic arthropathy
Joint replacement surgery was subsequently carried out in this case.

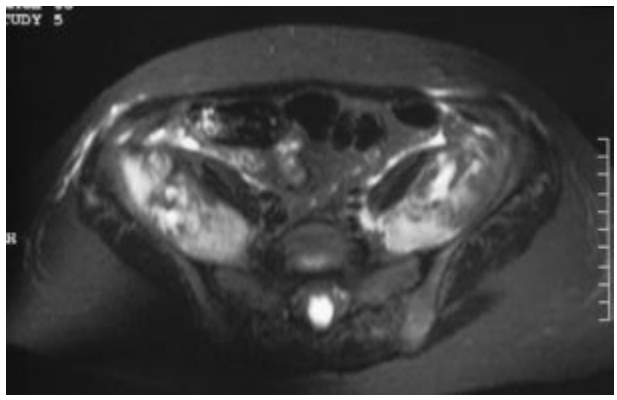


Fig. 16.3 Magnetic resonance imaging (MRI) scan showing bilateral iliopsoas hemorrhage
This bleed was associated with complete but transient paralysis in both legs, as the femoral nerve is located on the anterior surface of the muscle and may be compressed in such cases.

Bleeds into certain areas are particularly dangerous because of the risk of compression of neighboring structures. Patients with inhibitory antibodies are particularly at risk in this regard as bleeds may be more difficult to control. Bleeds in the tongue can obstruct the airway, and retroperitoneal bleeding within the iliopsoas muscle may result in femoral nerve compression, causing weakness and wasting of leg muscles (Figure 16.3). Bleeding from the gastrointestinal tract (melena) and bleeding into the urinary tract (hematuria) may also occur. There is also a significant risk of intracranial hemorrhage in severe hemophilia, which was a significant cause of mortality in the past when treatment was not so readily available. Higher levels of factor VIII (or IX) above 5 IU/dl are associated with a milder form of the disease, with no spontaneous joint bleeds but a definite risk of bleeding after even relatively minor injury.

Treatment of bleeding episodes involves the intravenous injection of coagulation factor concentrates; the total dose and frequency of treatment will be determined by the severity and site of bleeding. The great majority of joint bleeds will resolve with a single infusion of material if the bleed is recognized early and treated promptly. There is an increasing move to prophylactic therapy, in which the patient gives himself injections of coagulation factor concentrate two or three times a week to prevent bleeds rather than just treating on demand when bleeds occur. Patients on prophylactic therapy experience few or even no spontaneous bleeds, and thus progressive joint damage and arthritis can be avoided. The quality of life of patients on prophylaxis may be greatly enhanced, allowing them to lead much more independent lives.

Approximately 15% of patients with severe hemophilia A can be expected to develop inhibitory antibodies to factor VIII at some stage. Inhibitor development in hemophilia B is, by contrast, very rare and encountered in fewer than 1% of patients. The development of such antibodies poses considerable problems in treatment as these immunoglobulins (IgG) are capable of rapidly inactivating infused factor VIII; furthermore, the antibody titer may rise dramatically after a course of factor VIII. Very occasionally, acquired hemophilia may arise in a previously normal individual, due to the formation of autoantibodies directed against factor VIII, and both males and females may be affected. Hemarthrosis is unusual in acquired hemophilia, and the principal manifestations are usually extensive superficial purpura and muscle bleeds. Acquired hemophilia arises most often in the elderly, and there is an association with underlying malignant or autoimmune diseases.

Inheritance of hemophilia

The genes for factors VIII and IX are both located at the telomeric end of the X chromosome and thus hemophilia is inherited as an X-linked recessive condition (Figure 16.4). The daughters of affected males are obligate carriers but the sons are normal. The phenotype remains constant within a family, so the daughter of a man with only mild hemophilia may

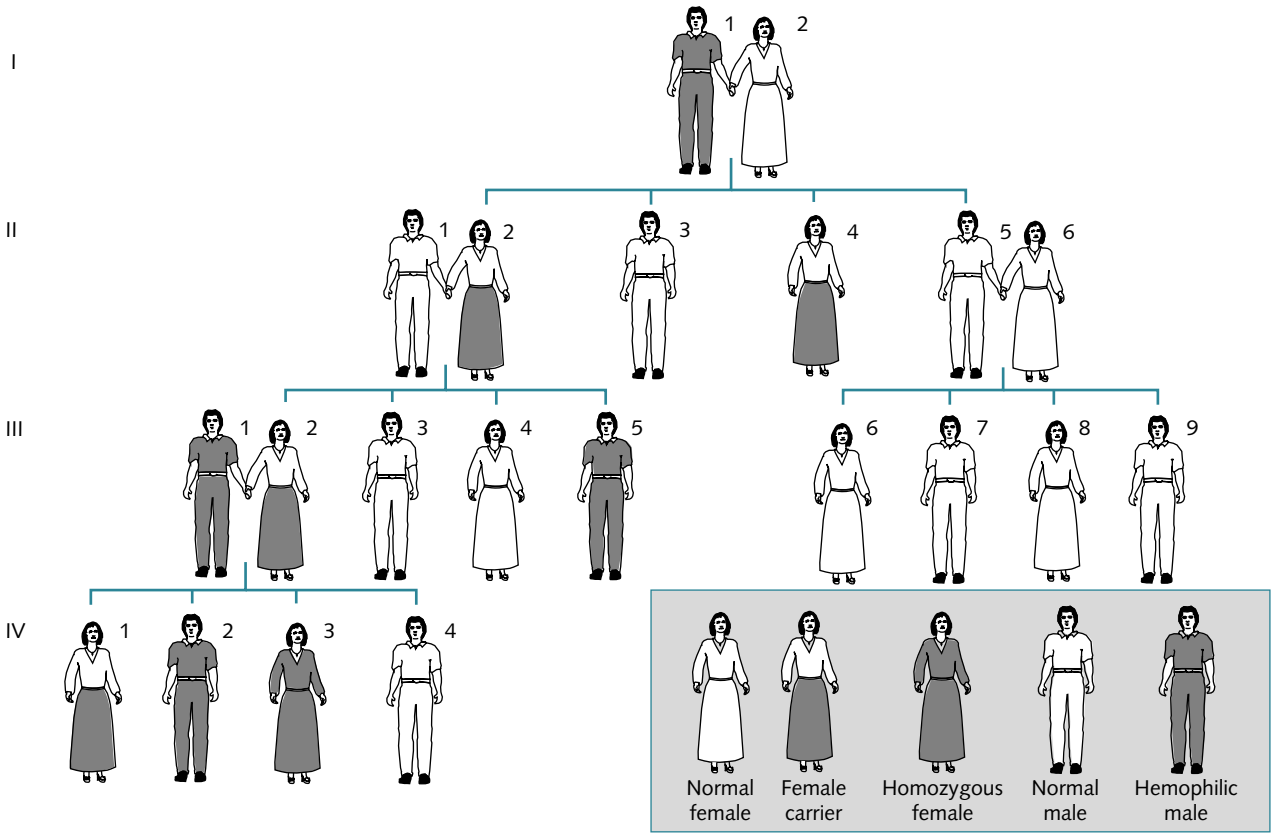


Fig. 16.4 The inheritance of hemophilia

The genes for factors VIII and IX are both on the X chromosome, and inheritance is thus sex-linked. The daughter of a man with hemophilia is an obligate carrier but the son of a hemophiliac will not be affected. Hemophilia may thus be transmitted to a grandson via a carrier daughter. Color blindness and Duchenne muscular dystrophy are other examples of X-linked disorders.

be reassured that she will not pass on a severe form of the condition. However, approximately one-third of all cases of hemophilia arise in the absence of a previous family history and are due to a new mutation. The most famous example is that of Queen Victoria, who had a hemophilic son (Leopold), and two daughters (Alice and Beatrice) who turned out to be carriers. There are instances of hemophilia affecting females due to inheritance of the defective gene from both parents, and there are also case reports of hemophilia in females with Turner's syndrome (XO karyotype) and androgen insensitivity syndrome (testicular feminization, XY karyotype).

Molecular basis of hemophilia A

Factor VIII is an essential cofactor for the activation of factor X by activated factor IXa (*see Chapter 15*). Factor VIII must itself undergo proteolytic cleavage at two distinct sites through the action of thrombin before it becomes physiologically active. It circulates in plasma as a large glycoprotein bound non-covalently to the larger protein, von Willebrand factor. The factor VIII gene was first cloned in 1984. It is 186 kb in length and is situated on the long arm of the X chromosome at Xq28 (Figure 16.5). The factor VIII gene consists of 26 exons, of which exon 14 is the largest, and 25 introns. The mature factor VIII protein is made up of 2332 amino acids.

By far the commonest single genetic defect causing severe hemophilia is an inversion in intron 22, which is encountered in as many as 45% of people with severe hemophilia in all ethnic groups (Figure 16.6). The inversion mechanism involves an intronless gene of unknown function, designated *F8A*. Two copies of this gene are located near the tip of the X chromosome and there is another copy within intron 22 of the factor VIII gene itself. During meiosis, either of the two telomeric copies may cross over with the intronic copy, resulting in a division of the gene into two halves facing in opposite directions and separated by approximately 500 kb. Cross-over with the distal copy is much more common than cross-over with the proximal copy, and accounts for approximately 80% of all inversions. It is now recognized that inversion almost

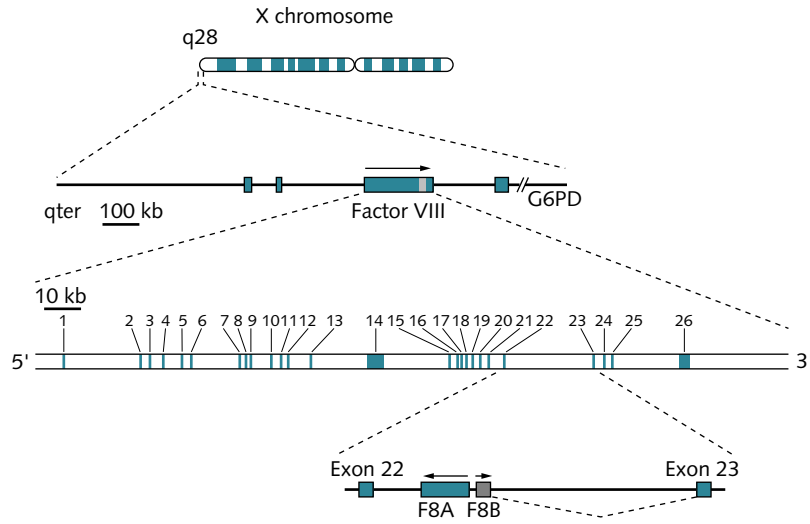


Fig. 16.5 The factor VIII gene
The factor VIII gene was cloned in 1984 and is located towards the telomeric end of the long arm of the X chromosome (Xq28). It maps distal to the gene encoding glucose-6-phosphate dehydrogenase (G6PD), about 1 Mb from the Xq telomere. The gene is composed of 26 exons, of which exon 14 is the largest. The large intron 22 contains a nested intronless gene termed *F8A* (the function of the gene and its transcript are unknown). There are two further copies of *F8A* located ~400 kb telomeric to the factor VIII gene. The spliced factor VIII mRNA is ~9 kb in length and the mature factor VIII molecule is composed of 2332 amino acids.

always occurs during a male meiosis. It is believed that the presence of a large region of non-homology between the X and Y chromosomes during meiotic pairing may favor a misalignment, and the presence of a second X chromosome with a complementary region may act as a stabilizing factor. An important clinical consequence of this observation is that, when an apparently new and spontaneous case of hemophilia is diagnosed in which the gene inversion is identified, it is likely that the defect arose in the maternal grandfather's allele; thus, the mother can generally be assumed to be a carrier and at risk of having another affected male child. The resulting truncated protein product is presumably unstable, resulting in severe hemophilia. The inversion is not found in individuals with mild forms of hemophilia.

The inversion is easily detected and the identification of this defect as the commonest cause of severe hemophilia has simplified both screening for carriers and the antenatal diagnosis of hemophilia (Figure 16.7). Using the restriction enzyme *Bcl*I, three bands are identified in normal individuals. The 21.5-kb band is contributed by intron 22 and the 14- and 16-kb bands are derived from the proximal and distal repeats, respectively. The inversion may be identified and classified according to the appearance of abnormal bands representing the products of recombination (Figure 16.7). More recently, inversions in intron 1 of the factor VIII gene have been identified as a cause of severe hemophilia. On the basis of the limited data currently available, this abnormality appears to be responsible for approximately 5% of all cases of severe hemophilia.

Developments in molecular biology have permitted the more rapid identification of defects in hemophilia. Southern blotting has been superseded by methods involving polymerase chain reaction (PCR) amplification of either patient DNA or material derived from the reverse transcription of mRNA (RT-PCR). Although automated sequence analyzers

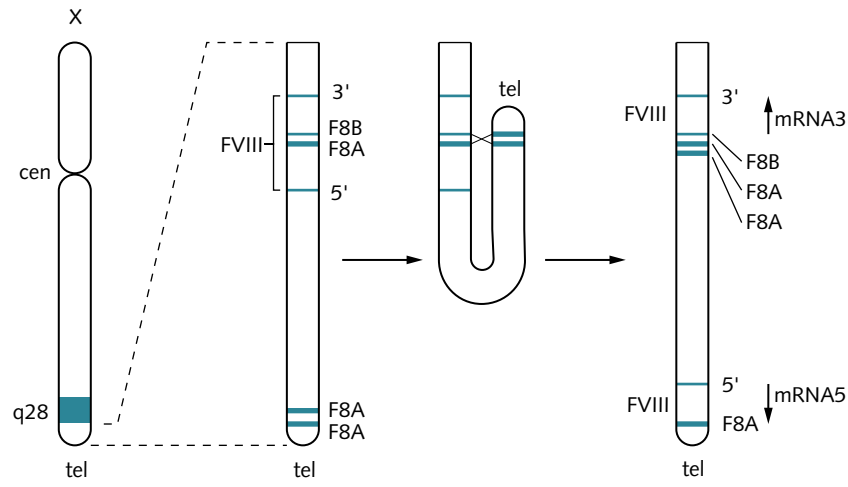


Fig. 16.6 'Flip tip inversion'
This unique inversion of the tip of the X chromosome is now recognized to be responsible for about half of all cases of severe hemophilia in all ethnic groups. Cross-over occurs between a copy of *F8A* within the factor VIII gene and one of the two telomeric copies. Cross-over with the distal copy is more common, occurring in approximately 80% of cases where an inversion is identified.

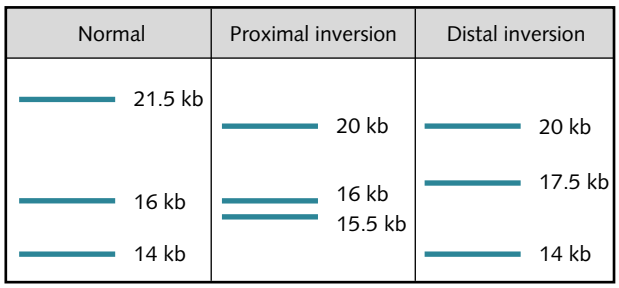


Fig. 16.7
Inversions within intron 22 of the factor VIII gene are by far the commonest cause of severe hemophilia, and are detected in almost half of all cases. This figure shows the results of Southern blot analysis using the *Bcl*I restriction enzyme. The normal pattern is shown on the left. A female carrier with an inversion will show a combination of the normal and either the proximal or the distal inversion pattern.

have been developed, gene sequencing of the entire factor VIII gene is both expensive and labour-intensive. Methods have been developed to identify restricted areas of abnormal DNA in patients with hemophilia, which may then be targeted for specific attention. These methods include *amplification and mismatch detection* (AMD), *conformation sensitive gel electrophoresis* (CSGE) and *denaturing gradient gel electrophoresis* (DGGE). Approximately 4% of cases of hemophilia are the consequence of gene deletions, which have been reported throughout the gene and which are very variable in size. As with the intron 22 inversion, most deletions are associated with a severe clinical phenotype. Frameshift mutations resulting from insertions or small deletions have also been identified as a cause of severe hemophilia. Most other cases of both severe and mild hemophilia are associated with single point mutations, and approximately 200 missense mutations have been identified as causing hemophilia A. A full list is outside the scope of this chapter, but a list of *Further reading* is provided at the end. Such mutations affect RNA processing, mRNA translation or the fine structure of factor VIII itself. Nonsense mutations result in the formation of stop conditions and the production of truncated factor VIII molecules devoid of any functional activity; for example, TGG(Trp)→TGA(Stop) at nucleotide 255 in exon 7, and CGA(Arg)→TGA at nucleotide 1941 in exon 18. Missense mutations that involve critical sites will also result in hemophilia A; for example, the Arg^{372}→His or Cys and Ser^{373}→Leu mutations, which involve a thrombin cleavage site. Approximately 40% of all missense mutations arise at CG dinucleotide sites, resulting in a change to TG or CA sequences. It is generally believed that CG nucleotides represent genomic hotspots. Cytosine is predominantly methylated in human DNA, but this is relatively unstable and 5-methylcytosine is prone to spontaneous deamination to yield a GT mismatch which is inefficiently repaired. It is also of interest that a missense mutation may be associated with varying degrees of clinical severity. Thus a C→T mutation at nucleotide 1689 has been reported in association with both severe and moderately severe phenotypes. Similarly, a Val^{326}→Leu substitution has been reported in individuals with either a severe or a moderately severe phenotype.

Molecular basis of hemophilia B

The factor IX gene is also located on the long arm of the X chromosome at band Xq27, and is encoded by a stretch of DNA approximately 34 kb long which contains eight exons (Figure 16.8). The basic structure of the gene is similar in organization to those of protein C and coagulation factors VII and X, and it is likely that they all originated in the distant past from a common ancestral gene by duplication. Factor IX is a polypeptide of 415 amino acids, and is made up of a glutamic acid-rich sequence (Gla domain) and two epidermal growth

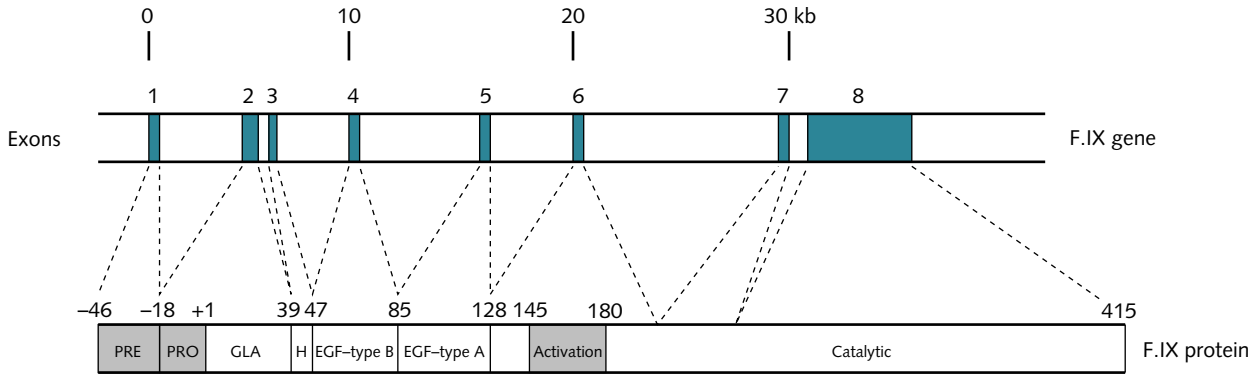


Fig. 16.8 The factor IX gene
The factor IX gene was first cloned in 1982. It is located on the long arm of the X chromosome at band Xq27. The gene spans 34 kb and contains eight exons (exons are shown as shaded boxes, and dotted lines between the gene and protein indicate protein domains encoded by each exon). The signal peptide and propeptide sequences are cleaved during the processing and activation of factor IX.

factor (EGF)-like domains separated from the serine protease domain by an activation region. The 12 glutamic acid residues in the Gla domain undergo post-translational γ-carboxylation, which is necessary for binding of calcium, and exon 2 encodes a recognition site for the carboxylase. Exon 1 encodes the signal peptide necessary for transport into the endoplasmic reticulum. Exon 6 encodes the activation peptide that is cleaved off during the activation of factor IX by either factor XI or a complex of tissue factor and factor VII. Exons 7 and 8 encode the catalytic regions of factor IX, which are responsible for the subsequent activation of factor X in the coagulation cascade. The gene is controlled by a promoter.

The gene for factor IX, which was cloned in 1982, is considerably smaller than that for factor VIII, and patients with hemophilia B have been studied more extensively than those with hemophilia A. The first defects identified in hemophilia B were gross deletions, detected by Southern blotting. However, it is now recognized that gene deletions account for only approximately 3% of all cases of hemophilia B. No equivalent of the factor VIII gene inversion has been encountered in hemophilia B and it is now clear that point mutations account for the vast majority of cases of hemophilia B; over 500 have been described from families around the world. The great majority involve single base changes, which have been identified in all domains of the protein. The unusually high frequency of mutations at CG dinucleotide sites in hemophilia B probably reflects the high number of CG dinucleotides at critical sites in the factor IX gene.

Mutations involving the glutamic acid residues within the Gla domain (residues 1–38) result in severe hemophilia, emphasizing their functional importance; for example, $Glu^{7}\rightarrow$ Val and Glu17→Lys. Mutations within exon 6 result in hemophilia through disruption of the activation of the factor IX molecule; for example, CGT→TGT at position 20 413 results in $Arg^{145}\rightarrow$Cys. Most cases of hemophilia B are attributable to mutations within exons 7 or 8, resulting in impaired catalytic activity; for example, a change from AGT to AGA at nucleotide 31 216 results in a change from Ser^{365} to Arg within the active site. The original case of Christmas disease has been identified as a TGT→TCT mutation at nucleotide 31 170 resulting in a change from Cys^{206} to Ser within exon 8. The creation of nonsense mutations will lead to severe hemophilia B due to the production of ineffective, truncated proteins; for example, CGA(Arg)→TGA(Stop) at nucleotide 30 863 and TGG(Trp)→TGA(Stop) at nucleotide 31 051 in exon 8.

A few patients with hemophilia B have been described in whom the level of factor IX rises significantly after puberty, and this is associated with loss of the bleeding tendency. Several point mutations have been reported in association with this interesting variant, referred to as the hemophilia B Leiden phenotype. All are located in the promoter region of the factor IX gene; for example, TTG→TAG at –20 and G→A at nucleotide –6. Most of these mutations have been shown to be located in regions which contain binding sequences for liver-enriched transcription factors, which are presumably influenced by androgenic steroids.

Inhibitor formation: etiology and clinical implications

A minority of patients with hemophilia will develop immunoglobulins directed against infused factor VIII (or IX) after exposure to these blood products for treatment of bleeding episodes. This is potentially very serious, as patients will be refractory to conventional doses of coagulation factor concentrates and bleeding will be difficult to control. Porcine factor VIII (Hyate:C), prothrombin complex concentrates (e.g. FEIBA) and recombinant activated factor VIIa (Novo Seven; Novo Nordisk) are valuable therapeutic materials in controlling bleeding in patients with inhibitory antibodies. Another important strategy in the management of patients who develop inhibitory antibodies is immune tolerance, which involves the daily administration of coagulation factor concentrate over a period of some months. This usually results in the eventual disappearance of the antibody, as the body becomes tolerant of the protein and inhibitor formation is suppressed.

Inhibitory antibodies interfere with the normal function of factor VIII in a number of different ways. The most frequent site of inhibitor binding occurs within the A2 and C2 domains. Inhibitors may thereby block the ability of activated factor VIIIa to bind and activate factors IXa and X, or inhibit the binding of factor VIII to von Willebrand factor or negatively charged phospholipid surfaces. Inhibitors may also hinder the activation of factor VIII by thrombin, or the subsequent release of factor VIII from von Willebrand factor. Proteolysis of factor VIII has recently been identified as a novel additional mechanism of inactivation in some cases.

Data from the UK registry show that 14% of all patients with severe hemophilia A have developed antibodies at some time, but it is quite likely that this figure underestimates the true prevalence as transient and low-titer inhibitors may not be detected. As a general rule, if an individual is susceptible to inhibitor development, this will become apparent at a fairly young age. Data from prospective studies involving recombinant factor VIII products suggest a median of approximately 10 exposure days for inhibitor development if this is to occur. It is now clear that the major factor that determines the predisposition to inhibitor development is the underlying molecular defect. However, there is also additional evidence from family and twin studies that other subtle genetic factors play a role, although no associations with specific human leukocyte antigens (HLA) or other linkages have been conclusively identified. Race may influence the risk of inhibitor development,

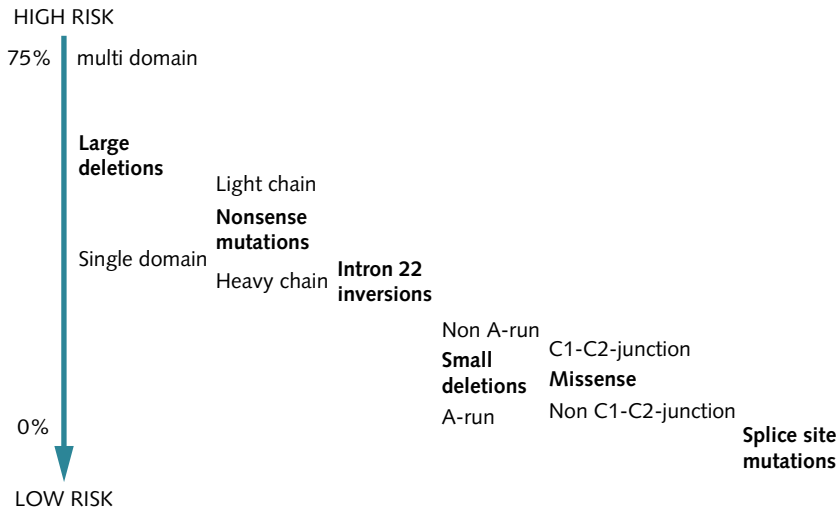


Fig. 16.9 Mutation types and the risk of inhibitor development

and several studies have shown that people of Afro-Caribbean origin are more susceptible to inhibitor formation.

Certain types of gene defects in hemophilia are undoubtedly associated with a significantly increased risk of inhibitor development (Figure 16.9). The risk of inhibitor development in patients with severe molecular defects, such as large deletions, nonsense mutations and the intron-22 inversion, is 7–10 times higher than in patients with other defects, such as missense mutations, small deletions and splice site mutations. The overall risk of inhibitor development in patients with the common intron-22 inversion is approximately 30%. Further studies will be required to determine the risk of inhibitor development in association with the newly reported intron-1 inversion, although 1 of 10 (10%) subjects in the original publication was known to have developed an inhibitor.

Inhibitor development in hemophilia B is a relatively rare event, occurring in probably fewer than 1% of patients, and these patients often have an underlying large gene deletion. By contrast with the immunoglobulin inhibitors in patients with hemophilia A, inhibitory antibodies in patients with hemophilia B are often capable of fixing complement proteins. The administration of coagulation factor concentrate in such cases may therefore trigger severe, often anaphylactic, allergic reactions. The development of nephrotic syndrome has also been reported in patients with hemophilia B and inhibitors who are undergoing immune tolerance induction treatment with factor IX concentrates.

Therapeutic applications of molecular biology to patient care

Carrier testing

Ideally, carriers of hemophilia should be identified before a pregnancy, and offered counseling. The inheritance of hemophilia is sex-linked, as it is for other disorders, such as color blindness and Duchenne muscular dystrophy. The daughters of men with hemophilia are thus obligate carriers of the condition, with a 50:50 chance of passing on the condition to a son, and there is a similar chance that a daughter of a carrier will also herself be a carrier of the condition. No special genetic tests are therefore required to determine the carrier status of daughters of men with hemophilia, although the results of DNA-based studies are likely to be useful for subsequent antenatal diagnostic procedures. The phenotype of hemophilia remains constant within a family, so that the daughter of a man with only mild hemophilia may be reassured that she can only transmit a similarly mild form of the condition. However, a more common problem is to be confronted with a woman with only a vague history of a bleeding disorder in a distant relative. National patient registers are particularly useful in such circumstances and may help to establish the type and severity of bleeding disorder of an affected relative. It may seem logical to initiate carrier testing to determine carrier status as soon as possible in girls with a family history of the condition, as this would facilitate the management of pregnancy in the case of an early and possibly unexpected pregnancy. It is clear that there are significant differences of opinion between geneticists and hematologists with regard to the timing of testing of children for genetic disorders. Clinical geneticists in the UK usually take the view that it is unethical to test healthy young children to determine carrier status for inherited disorders for conditions which have no immediate implications for their own health. There are also legal implications in the UK, as testing of young children ignores the rights of children with respect to the Children Act (1989), and testing cannot be considered to have been obtained with the informed consent of the individual child concerned. By contrast, guidelines from the Genetics Working Party of the UK Haemophilia Centre

Directors' Organisation (UKHCDO) conclude that carrier testing in young children may be appropriate, but that the issues must be discussed openly with the family.

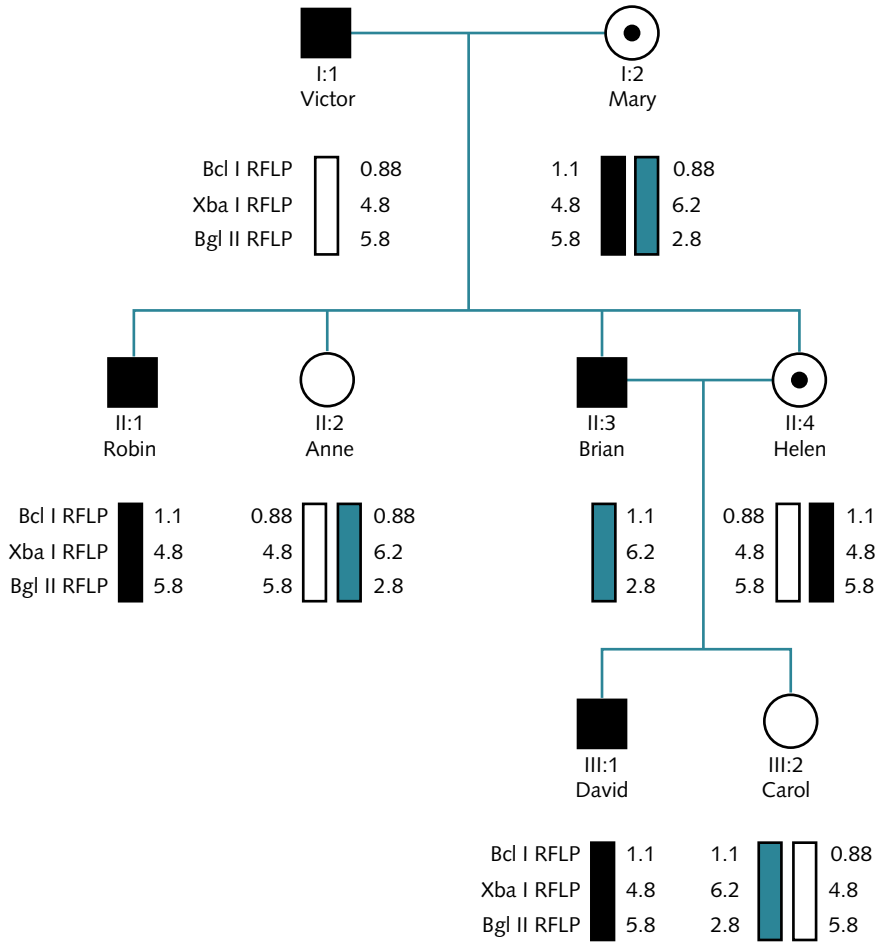
From a theoretical point of view, the ideal approach to identifying carriers is to characterize the precise genetic defect responsible for hemophilia within a family. Once the molecular defect has been identified in an individual with hemophilia A or B, direct screening for that defect could be applied in subsequent generations for both carrier testing and antenatal diagnosis. This direct method is clearly the method of choice but if the gene defect is unknown and it is not possible to establish this with the facilities available, indirect methods can still be used to determine carrier status. This involves the tracking of intragenic or extragenic polymorphisms in families which act as convenient markers for the molecular defect causing hemophilia within a family. The method is based on the fact that there are some genetic polymorphisms which represent natural variations of the genome sequence but have no adverse impact on the function of the molecule. Ideally, such polymorphic sequences should be intragenic, but closely linked extragenic polymorphisms may also be exploited. If extragenic markers are relied on, the small possibility of recombination may also result in erroneous diagnoses. The polymorphisms are detected by cleavage of patient DNA with restriction enzymes, which generate fragments of varying size according to the presence or absence of the polymorphism (Table 16.2). The fragments are detected by either PCR-based methods or Southern blotting. Figure 16.10 shows the results of restriction fragment length polymorphism (RFLP) analysis in one family, and illustrates how the method may permit identification of carriers. Limitations of this approach include the fact that samples have to be taken from several members to permit the tracking of the mutant gene responsible for hemophilia in the family. Blood from an affected family member will be required, and this may not be possible in some cases where early death from complications such as HIV infection has occurred. Furthermore, non-paternity may confound attempts to track the gene in families. It should also be appreciated that there is ethnic variation in the allelic frequencies of the various polymorphisms (more so with factor IX than factor VIII), and this may influence the choice of probes used in some family studies. For example, the allelic frequencies with *Bcl*I and *Hin*dIII are reversed in African-Americans compared with Caucasians. *Bgl*I is more likely to be informative in people of Afro-Caribbean origin but it is not likely to be informative in families of Chinese origin. RFLP analysis is likely to be informative in the majority of families, but occasionally carrier status cannot be determined even with multiple probes. The commonest problem is homozygosity for all markers in the proband's mother. However, the recent identification of hypervariable dinucleotide tandem repeats (VNTR) in introns 13 and 22 has also helped to identify a greater proportion of carriers. Once the carrier status has been determined and DNA markers have been identified, it is then possible to offer antenatal diagnosis of hemophilia to pregnant women. In the past, measurement of the ratio of the factor VIII level to that of von Willebrand factor antigen has been used to estimate the probability of a woman being a carrier of hemophilia. Such tests no longer have a role in the modern era of genetic testing, although it may certainly be useful to establish the level of factor VIII in a potential carrier; if it is low, this would have implications for management in the setting of surgery and other invasive procedures.

Antenatal diagnosis of hemophilia

As a general rule, it is the practice in the UK to perform antenatal procedures to determine whether a fetus has hemophilia *only* when a termination is being contemplated. The general experience in the UK has been that only a minority of women subsequently take up the offer of antenatal diagnosis with a view to termination if an affected fetus is identified. This may well reflect the fact that many women with affected relatives appreciate that major advances in treatment in recent years, such as the introduction of recombinant products for children

Table 16.2 Factor VIII intragenic DNA polymorphisms.

Site	Restriction enzyme	Southern blot allele	Allelic frequency
5′ flanking region	*Taq*I	9.5 kb	72%
		4.0 kb	28%
Intron 18	*Bcl*I	1.1 kb	29%
		0.88 kb	71%
Intron 22	*Xba*I	6.2 kb	41%
		4.8 + 1.4 kb	59%
3′ of exon 26	*Bgl*I	20 kb	10%
		5 kb	90%



Marker	Probe	Enzyme(s)	Alleles kb
Bcl I RFLP Intron 18	p114.12	Bcl I	1.1 + 0.88
Xba I RFLP Intron 22	p482.6	Xba I + Kpn I	6.2 + 4.8
Bgl II RFLP DXS15	DX13	Bgl II	5.8 + 2.8

Key

- Unaffected female
- Hemophilic male
- Carrier
- Unaffected male

Fig. 16.10 The use of restriction fragment length polymorphism (RFLP) analysis in a family to determine carrier status
Robin and David both have severe hemophilia A. Victor and Brian are normal. It is clear from the family tree that the disorder is associated with the 1.1/4.8/5.8 haplotype, and this permits the identification of Robin's sister Helen as a carrier. Anne and Carol are not carriers.

and the wider adoption of prophylaxis, offer the prospect of an essentially normal life for the younger generation of people with hemophilia.

Amniocentesis was the first technique employed for the antenatal diagnosis of hemophilia and other X-linked disorders, such as muscular dystrophy. Whilst amniocentesis is both technically simple and safe, an important limitation is the fact that it may only be employed in the second trimester of pregnancy, after approximately 15 weeks of gestation. Chorionic villus sampling (CVS) was first applied to the antenatal diagnosis of a number of genetic disorders in the early 1980s, but this technique is now the principal method used for the antenatal diagnosis of hemophilia and several other single-gene disorders. The principal advantage is that the method may be applied during the first trimester, so that if termination of the pregnancy is required this is easier to carry out. Furthermore,

the results of the test are often available within only a few days of the procedure, as (in contrast to amniocentesis) there is no need to culture cells before genetic analysis. A sample is obtained either by a transabdominal or a transvaginal route, under ultrasound guidance (Figure 16.11). However, CVS should not be undertaken before 11 weeks of pregnancy in order to minimize the risk of inducing congenital limb abnormalities. A disadvantage of CVS is that the procedure has to be carried out at a time when fetal sexing through ultrasound scanning is not feasible, so that a female fetus is unnecessarily exposed to risk.

Direct fetal blood sampling may be used for the antenatal diagnosis of hemophilia but this method is usually only offered as a last resort, either because it was not possible to carry out DNA-based family studies in time or because such studies were carried out but were not informative. In this technique, fetal blood is taken from fetal umbilical vessels under ultrasound guidance. The procedure requires considerable expertise and will thus not be available in all hospitals. It is usually carried out at a minimum of 18 weeks of gestation. The levels of factors VIII and IX in a normal fetus at around 19 weeks of gestation are significantly lower than those in an adult, at approximately 40 IU/dl and 10 IU/dl, respectively.

In future, it is likely that antenatal diagnosis of hemophilia (and other genetic disorders) will be based on the *isolation of fetal cells from the maternal circulation*. It is well documented that fetal lymphocytes may be isolated from the maternal circulation during pregnancy, and determination of the karyotype in fetal cells has been used to predict fetal sex. However, fetal lymphocytes may persist in the maternal circulation for many years after a pregnancy and this limits use of this method for second or subsequent pregnancies. More recently, fetal normoblasts have been isolated from the maternal circulation with a flow cytometer/cell sorter and fetal DNA extracted from these in order to probe for markers of hereditary genetic disorders. The attraction of this approach is that it is non-invasive and also offers the prospect of very early diagnosis, as early as 7 or 8 weeks of gestation. Preimplantation diagnosis is another technique that has been developed, involving the determination of embryonic sex using dual fluorescence *in situ* hybridization of blastomere cells with labeled probes specific for the sex chromosomes. The method may be particularly attractive to women who would not be prepared to undergo a conventional termination of a well-established pregnancy.

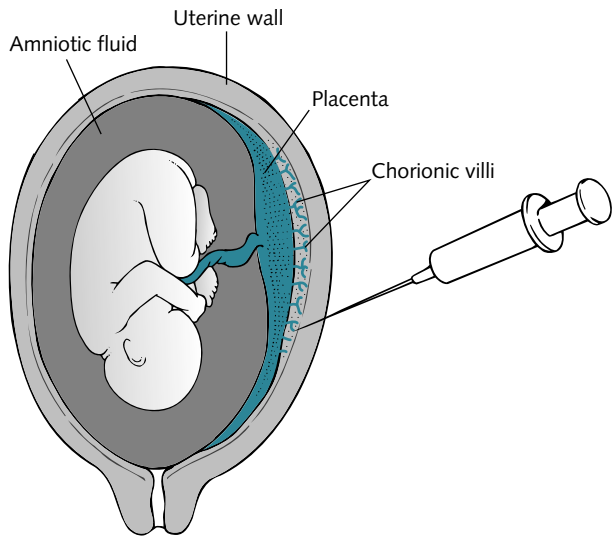


Fig. 16.11 Chorionic villus sampling (CVS)
A sample of trophoblastic tissue from the placental area is aspirated with a fine needle under general anesthesia. The procedure is usually carried out after 11 weeks of gestation. DNA isolated from the fetal tissue can then be analyzed to determine fetal sex and status with regard to hemophilia.

Recombinant blood products

The development of plasma-derived coagulation factor concentrates in the early 1970s dramatically improved both the longevity and quality of life of patients with hemophilia, and the demand for factors VIII and IX has risen steadily. The burgeoning global demand for factor VIII can no longer be met by products derived from volunteer blood donors. The manufacture of recombinant coagulation factor proteins offers the promise of unlimited supplies, albeit at increased cost. However, the most important advantage of recombinant products is safety with regard to the transmission of human pathogens. Many patients with hemophilia were infected with either HIV and/or hepatitis C before the introduction of physical methods of viral inactivation of plasma-derived coagulation factor concentrates in 1985. More recently, there has been concern about the possibility of transmission of variant Creutzfeldt–Jakob (vCJD) disease via blood products, as the prions believed to be the cause of this neurological disorder are extremely resistant to the usual viral inactivation procedures, such as heat treatment and exposure to a solvent/detergent mixture. Recombinant coagulation factor products offer the best possible protection from transmission of human blood-borne viruses and are regarded as the treatment of choice for all patients with hemophilia. However, the increased cost compared with conventional plasma-derived products has limited the availability of these products.

Recombinant coagulation factor concentrates are manufactured by insertion of the human gene into mammalian cell lines (such as Chinese hamster ovary cells or baby hamster kidney cells), which are then grown in culture on an industrial scale. Factor VIII (or IX) is then secreted into the growth medium, from which it is subsequently extracted by monoclonal or other immunoaffinity chromatography (Figure 16.12).

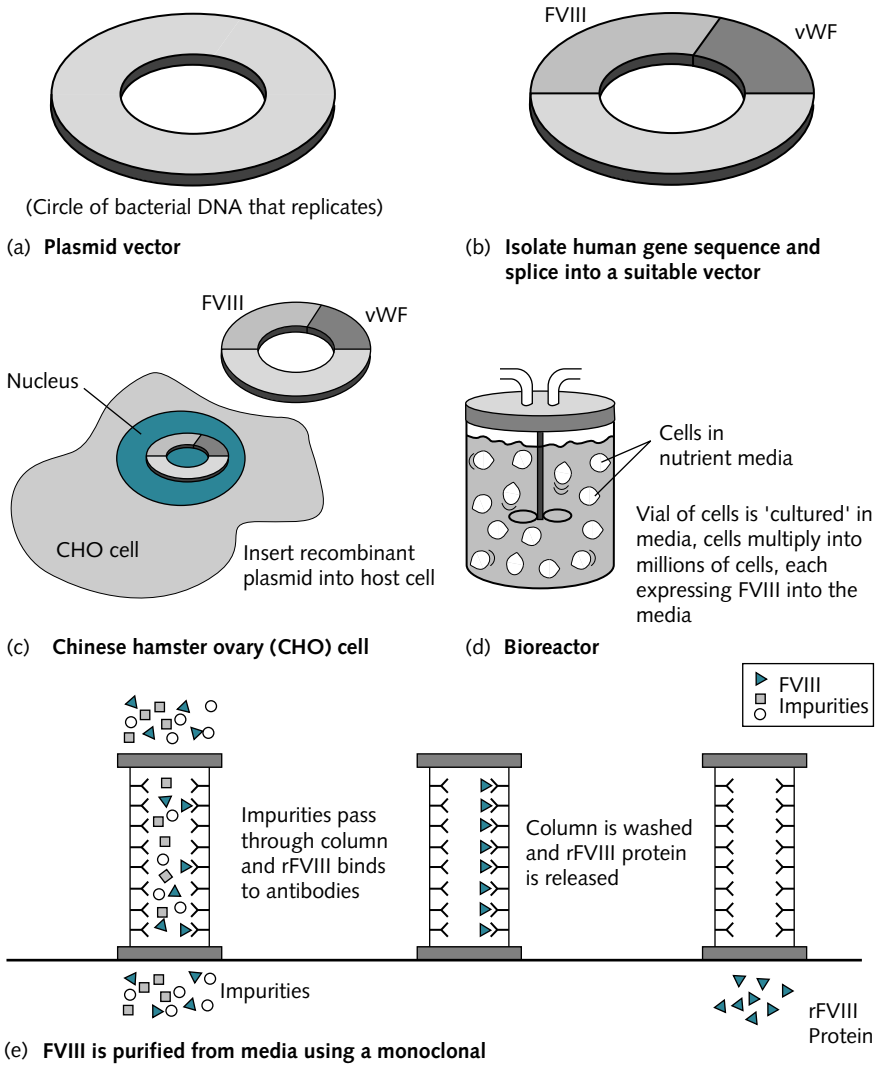


Fig. 16.12 Manufacture of recombinant factor VIII
The gene for human factor VIII is incorporated into a bacterial vector (a, b). Inclusion of the von Willebrand factor gene also enhances production of factor VIII. The vector is then inserted into a mammalian host cell (c). The cells grow and multiply in a nutrient medium, and then secrete factor VIII (d). Factor VIII can be extracted and purified by immunoaffinity chromatography (e). The final product contains no von Willebrand factor.

The original recombinant factor VIII products all contained added human albumin as a stabilizer. However, second- and third-generation products are now available in which alternative stabilizers are used and bovine proteins have also been eliminated from the culture media. These measures will further increase the margin of safety as regards viral and other infections. Recombinant factor VIII has an essentially identical structure [with the exception of one brand, Refacto (Wyeth), which has no B-domain] and glycosylation profile to natural plasma factor VIII. The pharmacokinetic profile and postinfusion recovery are also identical to that observed in plasma-derived concentrates.

Initial concerns about a potential increase in the incidence of inhibitors among people with hemophilia receiving recombinant products have proved unfounded. Although randomized trials comparing the incidence in previously untreated patients receiving recombinant products with those receiving plasma-derived products have never been conducted, it would be fair to say that the current consensus is that the incidence of inhibitor development is very similar for the two types of product.

Recombinant factor IX is also available. Recombinant factor IX is identical in amino acid sequence to the Ala^{148} (as opposed to the less common Thr^{148}) human polymorphic variant. Plasma factor IX is synthesized in the liver and undergoes post-translation glycosylation of a number of glutamic acid residues. Vitamin K is a vital cofactor for this process and is essential for its activity, but recombinant factor IX is not as effectively carboxylated. The postinfusion recovery does appear to be reduced when compared with plasma-derived products, although the plasma half-life is identical. It is a smaller molecule than factor VIII and requires no albumin or

other material to be added to the final product as a stabilizer. The cell line is grown in media that contain no animal or human-derived proteins but the product is subjected to nanofiltration to enhance its safety profile. There is no suggestion of an increased risk of inhibitor development associated with the use of recombinant factor IX.

Another useful recombinant product is recombinant factor VIIa (Novo Seven; Novo Nordisk). It is now recognized that factor VII plays a key role in the initiation of the coagulation cascade through contact with tissue factor released from damaged tissues, to form activated factor VII (VIIa) (*see also Chapter 15*). Recombinant factor VIIa is very useful in the clinical management of patients with hemophilia A or B and inhibitory antibodies, as well as patients with acquired hemophilia.

Looking to the future, it is likely that the direction of future research in genetic engineering will increasingly be applied to the production of modified molecules with more favorable properties. For example, it would obviously be useful to produce factor VIII molecules with a longer plasma half-life or reduced propensity to stimulate inhibitor development. Hybrid factor VIII molecules have been developed in which the A2 and C2 domains have been replaced by porcine equivalents. Since almost 90% of inhibitory antibodies bind to these two domains of the human factor VIII molecule, it is hoped that these new constructs may be of clinical use in the treatment of people with inhibitory antibodies. It has recently been discovered that factor VIII catabolism is mediated by low-density lipoprotein receptor-related protein (LRP), a hepatic clearance receptor with broad ligand specificity. Pharmacological blockade of these catabolic receptors is another potential target for prolonging the plasma half-life of infused factor VIII in patients with hemophilia. A further development has been the generation of transgenic livestock such as sheep, pigs and goats for the production of human coagulation proteins. Transgenic animals that secrete antithrombin, factor VIII or factor IX into their milk have been produced, and this approach is being explored with a view to the production of relatively cheap and unlimited supplies of biologically active products free of the risk of transmission of human pathogens (Figure 16.13). More recently, this work has been extended by the successful cloning of sheep. The production of transgenic animals by nuclear transfer may permit the establishment of large breeding colonies of livestock more quickly and efficiently than would be possible through the production of individual transgenic sheep by pronuclear microinjection.

Gene therapy for hemophilia

(*See also Chapter 23.*) Gene therapy offers the prospect of a cure for hemophilia in the long term, but it must be emphasized that there is no prospect of large-scale application for some years. The results of five recent clinical trials in the USA of gene therapy for hemophilia have yielded encouraging results, with some evidence of increased circulating levels of factor VIII or IX. However, all these studies have been Phase I studies, primarily designed to test safety rather than efficacy. There are two basic approaches to gene delivery into cells. The first technique involves the direct injection of transducing vector into the bloodstream or target tissue, with subsequent *in vivo* transformation of the cells that take up the gene. Alternatively, target cells may be modified by removal of cells from a patient, with subsequent modification *ex vivo* of these cells followed by reinfusion. Of five clinical studies initiated so far, three have targeted the liver (the natural site of factor VIII and IX synthesis), whilst two other groups have focussed on skeletal muscle cells and dermal fibroblasts.

Retroviruses and adenoviruses have been used extensively as vectors (Table 16.3). The principal advantage of using retroviruses as vectors is that the genetic material is actually integrated into the genome of the target cell, so expression of the transfected gene is permanent. However, integration is random, introducing the potential for oncogenesis through the disruption of oncogenes. A further problem with the use of retroviruses as vectors is that there is a physical limit of approximately 8 kb in the size of cassette that can be accommodated within the virus. The factor IX gene may be accommodated, but the full-length factor VIII gene cannot. Adenoviruses permit the transfer of larger genes and can transfect non-dividing cells but transferred DNA does not integrate permanently, so expression of the transfected gene is transient. A further limitation is that the immune response to adenoviral proteins, which are commonly encountered in everyday life, may limit the efficiency of transfer.

As there are a number of legitimate concerns relating to the use of viruses for gene transfer, other modes of gene transfer have been sought. For example, liposomes have a number of other advantages, in that there is no limit to the size or conformation of the DNA that may be incorporated and the liposomes themselves are composed of non-toxic materials that are easily degraded, thus facilitating repeated treatment. In another study, dermal fibroblasts were transfected with a plasmid containing sequences of the factor VIII gene. Cloned cells which were selected for their capacity to produce factor VIII were harvested and implanted into the omentum by laparoscopy. In four of six patients, plasma levels of factor VIII rose above baseline levels and were sustained for up to 10 months.

It is likely that gene therapy for hemophilia B will be achieved earlier than gene therapy for classical hemophilia A since the smaller size of the factor IX gene compared with the factor VIII gene permits the use of retroviral vectors; furthermore, factor IX (in contrast to factor VIII) may be absorbed

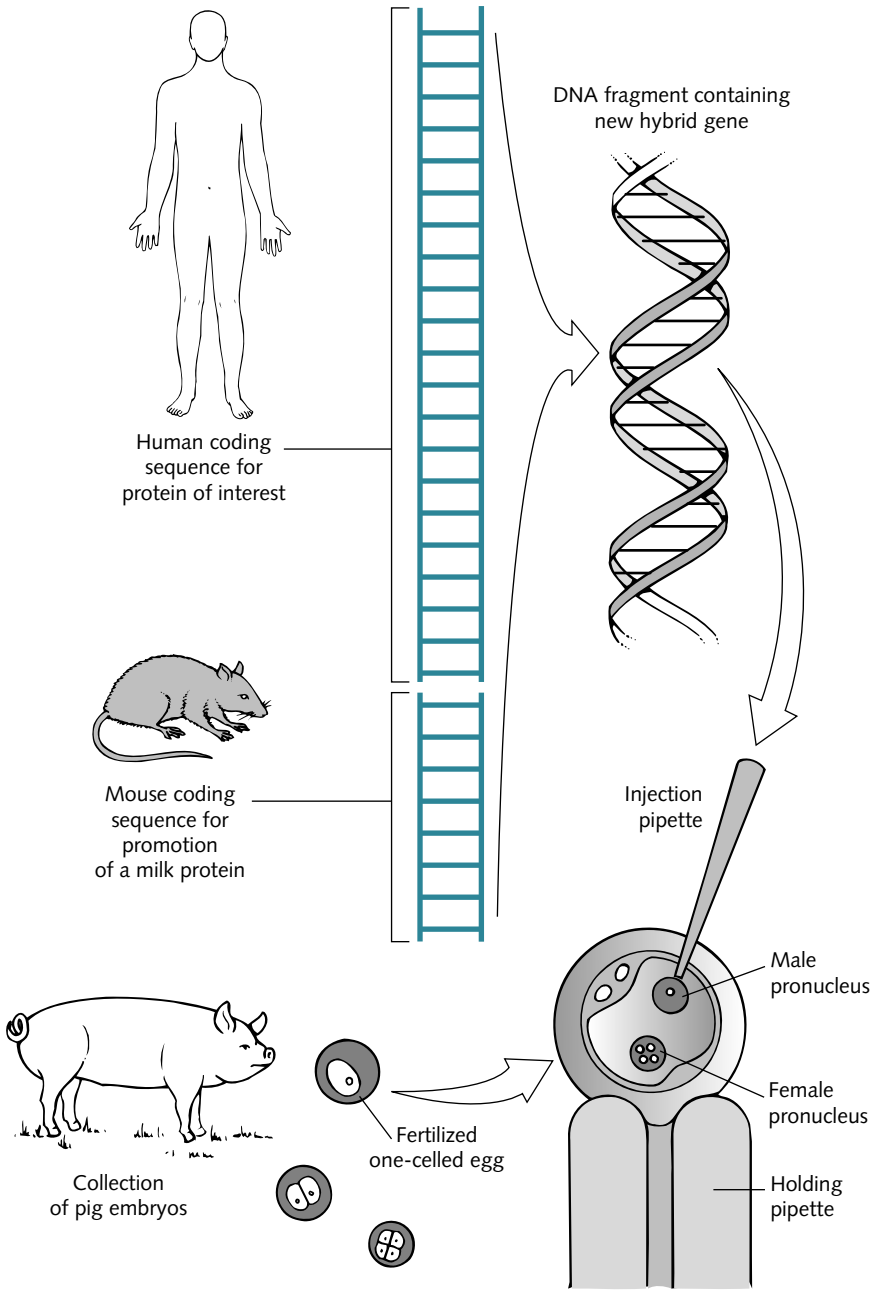


Fig. 16.13 Production of recombinant proteins using transgenic livestock ('pharming')

from subcutaneous tissues after local injection. Although the liver is the site of synthesis of factor IX, a number of other cells can produce factor IX very effectively after transfection with the human factor IX gene, even in the absence of vitamin K. Both human fibroblasts and keratinocytes can produce factor IX, but keratinocytes are particularly attractive cells for gene therapy as they are very accessible, grow well in culture and can be grafted with ease.

Gene therapy poses a number of ethical problems, particularly since effective and safe treatment with recombinant coagulation factors is now available for patients with either hemophilia A or hemophilia B. The use of viral vectors introduces risks such as oncogenesis and infection, or even modification of patient germlines. Patients will also need to be enrolled in clinical trials of very long duration, so that close follow-up will identify problems. Finally, in contrast to the usual form of clinical study, it is likely that children will have to be used in the initial clinical studies in preference to adults because the limited yield of coagulation proteins from current cell culture systems would not suffice for larger subjects.

Table 16.3 Comparison of adenoviruses and retroviruses as vectors.

Retroviruses	Adenoviruses
Physical limit to gene cassette (8 kb)	Accommodate much larger genes
Relatively low copy number	More efficient transfer
Can only infect dividing cells	Infect non-dividing cells
Permanent integration: sustained expression of transduced gene and potential for oncogenesis	Not integrated after infection: transient gene expression only
Poor immune response of host	Stimulate immune responses of host

Conclusions

Hemophilia is an inherited disorder of coagulation, associated with congenital deficiency of factor VIII (or IX). It is inherited in a sex-linked fashion, so that only males are affected. Approximately one-third of cases arise in families with no previous family history, and represent new mutations. The typical features of severe hemophilia include spontaneous bleeding into the joints, but in the absence of treatment more serious complications (such as intracranial hemorrhage) will lead to early death.

The first products used for the treatment of hemophilia were derived from human plasma, but unfortunately the use of pooled plasma products before 1985 resulted in the transmission of serious viral infections to many patients, such as HIV and hepatitis. In recent years, the development of recombinant blood products has eliminated the risk of transmission of these infections, and also offers the prospect of unlimited supplies. The life expectancy of the younger generation of hemophiliacs now approaches that of the normal population.

The commonest molecular defect in hemophilia A is an inversion in intron 22 of the factor VIII gene on the X chromosome, which accounts for approximately half of all cases. Genetic testing is now readily available in many centers to document the genetic defect in each family and to identify carriers within families. Antenatal diagnosis is now easily available, facilitating early termination of the pregnancy if hemophilia is identified. Recent trials of gene therapy for hemophilia have yielded encouraging results.

Hemophilia resources on the Internet

Factor VIII (and VII) mutation database: *europium.csc.mrc.ac.uk*

Factor IX mutation database: *www.kcl.ac.uk/ip/petergreen/haemBdatabase.html*

World Federation of Haemophilia: *www.wfh.org*

National Hemophilia Foundation (USA): *www.hemophilia.org*

Oxford Haemophilia Centre: *www.medicine.ox.ac.uk/ohc/*

Further reading

Introduction: clinical features of hemophilia

Rizza C, Lowe G (eds). (1997) *Haemophilia and Other Inherited Bleeding Disorders*. Eastbourne: W.B. Saunders.

Zeepvat C. (1998) *Prince Leopold: The Untold Story of Queen Victoria's Youngest Son*. Stroud, UK: Sutton Publishing.

Hemophilia A

Antonarakis SE, Rossiter JP, Young M *et al.* (1995) Factor VIII gene inversions in severe hemophilia A: results of an international consortium study. *Blood*, **86**, 2206–2212.

Bagnall RD, Waseem N, Green PM *et al.* (2002) Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood*, **99**, 168–174.

Kemball-Cook G, Tuddenham EGD, Wacey AI. (1998) The factor VIII structure and function resource site: HAMSTeRS version 4. *Nucleic Acids Research*, **26**, 216–219.

Lakich D, Kazazian HH, Antonarakis SE *et al.* (1993) Inversions disrupting the factor VIII gene as a common cause of severe haemophilia A. *Nature Genetics*, **5**, 236–241.

Rossiter J, Young M, Kimberland ML *et al.* (1994) Factor VIII gene inversions causing severe haemophilia A originate almost exclusively in male germ cells. *Human Molecular Genetics*, **3**, 1035–1039.

Hemophilia B

Briët E, Bertina RM, Van Tilburg NH *et al.* (1982) Hemophilia B Leyden: a sex-linked hereditary disorder that improves after puberty. *New England Journal of Medicine*, **306**, 788–790.

Crossley M, Ludwig M, Stowell KM *et al.* (1992) Recovery from hemophilia B Leyden: an androgen-responsive element in the factor IX promoter. *Science*, **257**, 377–379.

Giannelli F, Green PM, Sommer SS *et al.* (1998) Haemophilia B: database of point mutations and short additions and deletion. *Nucleic Acids Research*, **26**, 265–268.

Inhibitors

Giannelli F, Choo KH, Rees DJG *et al.* (1983) Gene deletions in patients with haemophilia B and factor IX antibodies. *Nature*, **303**, 181–182.

Hay CR, Baglin TP, Collins PW *et al.* (2000) The diagnosis and management of factor VIII and IX inhibitors: a guideline from the United Kingdom Haemophilia Centre Doctors' Organisation (UKHCDO). *British Journal of Haematology*, **111,** 78–90.

Lacroix-Desmazes S, Bayry J, Misra N *et al.* (2002) The prevalence of proteolytic antibodies against factor VIII in hemophilia A. *New England Journal of Medicine*, **346**, 662–667.

Oldenburg J, El-Maari O, Schwaab R. (2002) Inhibitor development in correlation to factor VIII genotypes. *Haemophilia*, **8** (Supplement 2), 23–29.

Saenko EL, Ananyeva NM, Kouiavskala DV *et al.* (2002) Haemophilia A: effect of inhibitory antibodies on factor VIII functional interactions and approaches to prevent their action. *Haemophilia*, **8**, 1–11.

Carrier testing and antenatal diagnosis

Clarke A and the Working Party of the Clinical Genetics Society (UK). (1994) The genetic testing of children. *Journal of Medical Genetics*, **31**, 785–797.

Giangrande PLF. (2002) Pregnancy in women with inherited bleeding disorders. In: Brenner B, Marder V, Conard J (eds). *Women's Issues in Thrombosis and Haemostasis*. London: Martin Dunitz.

Goodeve AC. (1998) Advances in carrier detection in haemophilia. *Haemophilia*, **4**, 358–364.

Tedgård U, Ljung R, McNeil TF. (1999) Reproductive choices of haemophilia carriers. *British Journal of Haematology*, **106**, 421–426.

Recombinant blood products

Brownlee GG, Giangrande PLF. (2001) Clotting factors VIII and IX. In: Buckel P. (ed.) *Recombinant Protein Drugs* (*Milestones in Drug Therapy* series). Basel: Birkhäuser, pp. 67–88.

Giangrande PLF. (2003) Treatment of hemophilia: recombinant products only? Yes. *Journal of Thrombosis and Haemostasis*, **1**, 214–215.

Lusher J, Ingerslev J, Roberts H, Hedner U. (1998) Clinical experience with recombinant factor VIIa. *Blood Coagulation and Fibrinolysis*, **9**, 119–128.

Mannucci PM. (2003) Treatment of hemophilia: recombinant products only? No. *Journal of Thrombosis and Haemostasis*, **1**, 216–217.

Soukharev S, Hammond D, Ananyeva NM *et al.* (2002) Expression of factor VIII in recombinant and transgenic systems. *Blood Cells, Molecules and Diseases*, **28**, 234–248.

Gene therapy

Kelley K, Verma I, Pierce GF. (2002) Gene therapy: reality or myth for the global bleeding disorders community? *Haemophilia*, **8**, 261–267.

Manno CS. (2002) Gene therapy for bleeding disorders. *Current Opinion in Hematology*, **9**, 511–515.

Roth DA, Tawa NE, O'Brien JM *et al.* (2001) Nonviral transfer of the gene encoding coagulation factors VIII in patients with severe hemophilia A. *New England Journal of Medicine*, **344**, 1735–1742.

Velander WH, Lubon H, Drohan WN. (1997) Transgenic livestock as drug factories. *Scientific American*, **276**, 70–74.

Chapter 17 The molecular basis of von Willebrand disease

Luciano Baronciani & Pier Mannuccio Mannucci

Introduction

Von Willebrand disease (vWD) is a common inherited bleeding disorder. Precise data regarding its prevalence are not available because of the extreme variability in clinical symptoms of mild vWD. However, population-based studies give an estimate of clinically significant vWD with a prevalence of at least 100 per million. vWD is caused by the deficiency or dysfunction of a multimeric plasma glycoprotein, von Willebrand factor (vWF). Because of its ability to bind to a number of ligands, vWF is involved in hemostasis via a variety of mechanisms but can be considered as having two main roles, in primary hemostasis and in intrinsic blood coagulation. vWF is directly involved in platelet binding to the subendothelium and in platelet-to-platelet interactions, and also acts as the carrier of procoagulant factor VIII (FVIII). Mutations at the *VWF* locus can affect vWF synthesis, its complex biosynthetic assembly, its stability in the circulation, and its binding interactions with specific ligands.

Function of von Willebrand factor in primary hemostasis

At the time of a hemostatic challenge, vWF acts as a bridge between platelets and the subendothelium of blood vessels, and is involved in the formation of the platelet plug. The role of platelets in hemostasis is to become irreversibly attached at sites of injury. The primary physical factor that affects platelet binding to the vessel wall is the rate of blood flow in the vessel, which is faster at the center and slower close to the wall. These variations in velocity create a shearing effect, or shear stress, between layers of fluid. Disruption of the vascular endothelial surface leads to exposure of the subendothelium and results in an alteration in the rate of blood flow and an increase in shear stress. Plasma vWF binds rapidly and tightly to subendothelial collagen. vWF does not constitutively bind platelets, but the immobilized vWF becomes activated. Even in high blood flow conditions, vWF is able to tether platelets and to expose thrombogenic surfaces through the interaction of its A1 domain and platelet receptor GpIb. However, the vWF–GpIb interaction does not provide irreversible platelet adhesion because of the fast dissociation rate, and platelets tethered to the vessel wall still move constantly in the direction of the flow, but at a much slower rate. A second molecule on the platelet surface is required to obtain firm platelet adhesion: the integrin GpIIb-IIIa ($\alpha IIb\beta_3$). This molecule is responsible for the platelet-to-platelet interaction, which is mediated once again by vWF and, under slow-flow conditions, by fibrinogen. $\alpha IIb\beta_3$ does not appear to be involved in the first events of platelet adhesion, probably because its rate of binding to vWF is too slow to mediate the initial platelet attachment to the vessel wall under high-flow conditions. However, when platelets become activated as a consequence of the vWF–GpIb interaction, $\alpha IIb\beta_3$ increases its affinity for its ligand, vWF. This event, together with the slow motion of the platelets due to the vWF–GpIb interaction, allows $\alpha IIb\beta_3$ to bind platelets irreversibly to the vessel wall (Figure 17.1). In vessels with a high shear rate, vWF is the primary mediator of platelet binding to the vessel wall and of platelet aggregation. This function of vWF is of great importance in primary hemostasis.

Function of von Willebrand factor in blood coagulation

vWF circulates in the plasma as a complex associated with

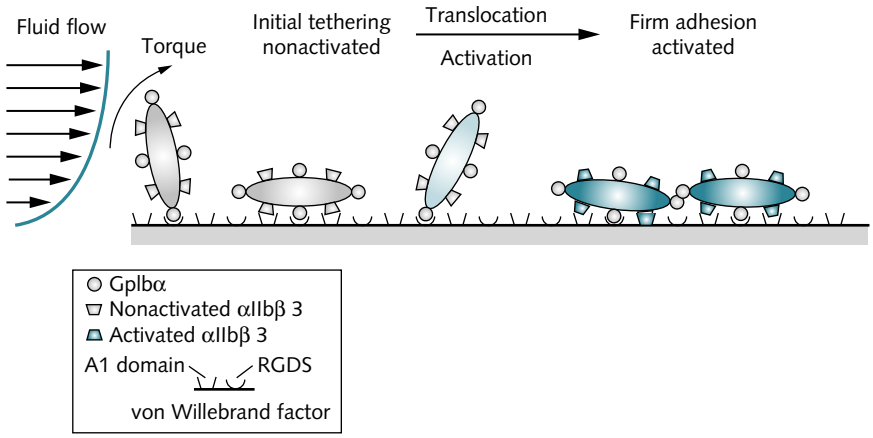


Fig. 17.1 Schematic representation of platelet adhesion to immobilized von Willebrand factor
At first, platelets are tethered to the subendothelium through the interaction of their GPIbα with von Willebrand factor (vWF; A1 domain), and the inactivated $\alpha IIb\beta_3$ does not bind to the RGDS sequence of vWF. Because of the torque imposed by the flowing fluid, the platelets begin to roll. New bounds are formed as different regions of the membrane of the rolling platelets come into contact with the surface, and the translocation continues until the $\alpha IIb\beta_3$ becomes activated and binds firmly to the RGDS of the vWF. Adapted with permission from Ruggeri ZM (1997). von Willebrand factor. *Journal of Clinical Investigation*, **99**, 559–564.

FVIII. The binding of vWF to FVIII is required to stabilize FVIII in the circulation, preventing cleavage by activated protein C (APC) or factor Xa. This interaction with vWF is crucial in prolonging the half-life of FVIII and concentrating FVIII at the point of bleeding. When patients with severe vWD (type 3) are treated with purified FVIII, this is cleared with a half-life of around 2 hours, whereas the vWF–FVIII complex infused in the same patient has a half-life of about 20–24 hours. vWF binds to FVIII via regions within the first 272 residues of the mature polypeptide, in the D′ and D3 domains of vWF.

Gene organization, synthesis and multimeric structure of von Willebrand factor

The gene encoding vWF has been mapped on chromosome 12, has a length of approximately 178 kb, with 52 exons, and transcribes a messenger RNA of about 8.2 kb. Analysis of the *VWF* gene is complicated by the existence of a partial unprocessed pseudogene on chromosome 22. The pseudogene extends from exon 23 to exon 34 and presents a high degree of homology with the gene (97%). vWF is synthesized in megakaryocytes and endothelial cells as a precursor of 2813 amino acids, the prepro-vWF. It is composed of a 22-residue signal peptide, a 741-residue propeptide and a 2050-residue mature subunit. More than 95% of the sequence accounts for structural domains that are arranged in the following order: D1-D2-D′-D3-A1-A2-A3-D4-B1-B2-B3-C1-C2-CK (Figure 17.2). The biosynthesis of vWF is a complex process, involving post-translational processing of the protein prior to storage or release into the circulation. The initial dimerization occurs by disulfide bonding between cysteine residues in the cystine knot (CK) carboxyl-terminal region of the monomer (tail-to-tail dimerization). The signal peptide is cleaved before it enters the Golgi apparatus. The tail-to-tail glycosylated dimers are then transported to the Golgi apparatus where multimerization and further glycosylation occur. The propeptide of vWF mediates the assembly of vWF multimers. This process requires the presence of the D1 and D2 domains (propeptide) and the D′ and D3 domains of the mature polypeptide. This is followed by cleavage of the propeptide sequence (Figure 17.3) and secretion of both the mature polypeptide and the propeptide into the circulation, or storage within the Weibel–Palade bodies of endothelial cells or the α-granules of platelets. The molecular weight of the mature subunit is 220 kDa but circulates as multimers of up to 20 000 kDa.

Von Willebrand disease and its classification

The most common symptoms of mild vWD are mucosal bleeding (epistaxis, gingival bleeding and menorrhagia) and prolonged bleeding after surgical procedures and dental extractions. Hemarthroses and soft-tissue hematomas are rare, but they occur in severely affected individuals. The diagnosis of vWD is suspected in individuals with these symptoms and a family history of bleeding. Several vWF assays are used in the diagnosis of vWD and its subtypes, such as those that measure the plasma levels of vWF antigen (vWF:Ag), vWF binding to type I or type III collagen (collagen binding activity, vWF:CB) and vWF interactions with the antibiotic ristocetin and platelet glycoprotein Ib (vWF ristocetin cofactor activity, vWF:RCo). This large number of measurements reflects the fact

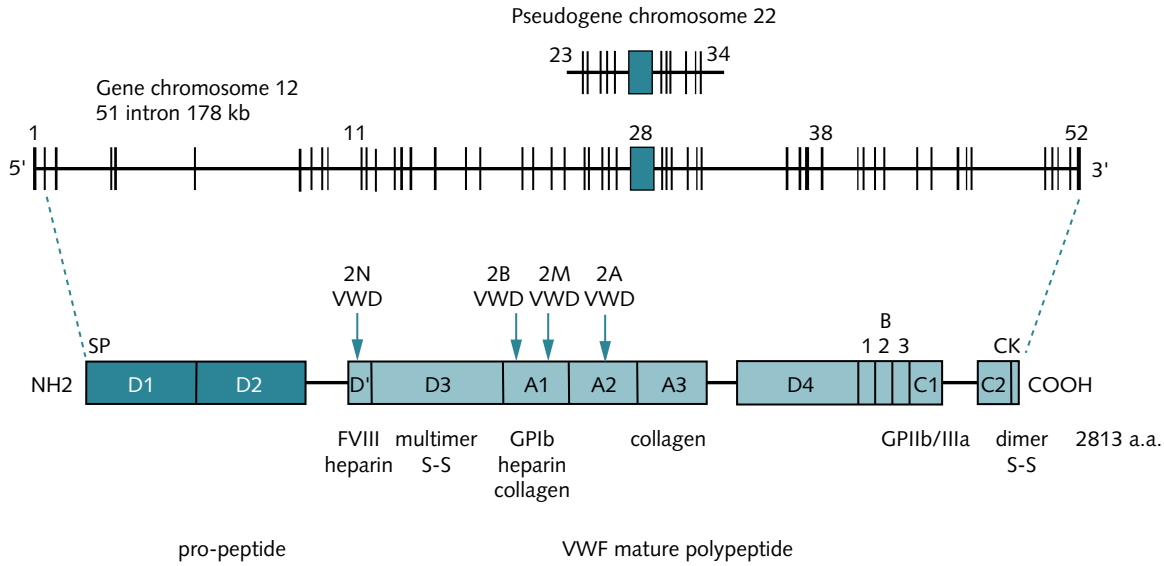


Fig. 17.2 Structure of the vWF gene, pseudogene and protein

The schematic structure of the prepro-vWF is shown, along with the homologous repeated domain. Also shown are the locations of intersubunit disulfide bonds involved in dimerization and multimerization, and the binding sites for several ligands.

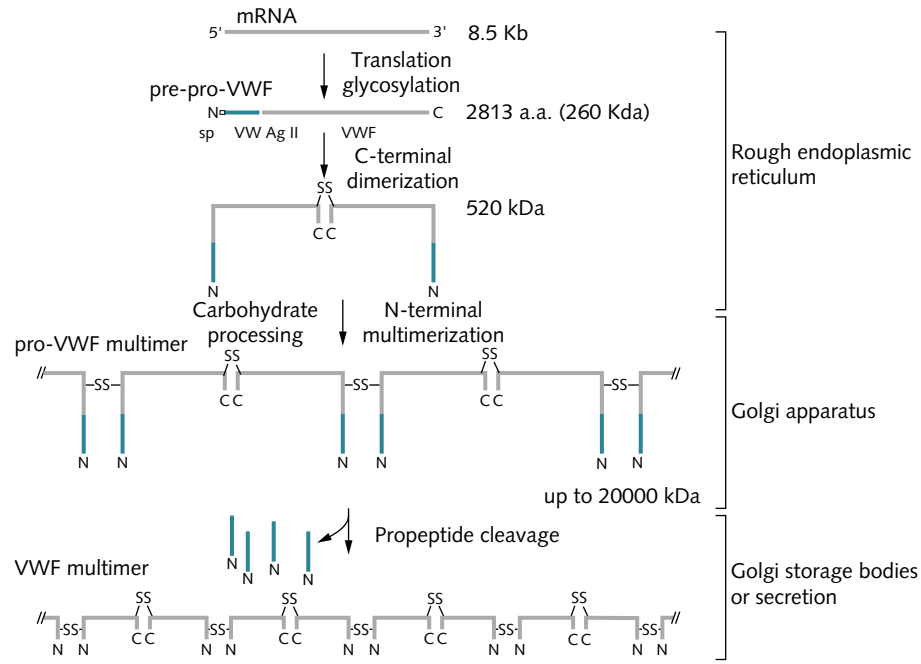


Fig. 17.3 Processing steps of vWF multimer synthesis and subcellular localization of the post-translational modification events

The propeptide is shown in blue and the mature subunit in gray. The N and C prefixes represent the amino- and carboxyl-terminal ends of the protein respectively. Monomers are linked together at the C-termini by disulfide bonds to form dimers (endoplasmic reticulum), which further multimerize in the Golgi apparatus. Cleavage of the propeptide occurs before secretion.

that none of them is by itself sensitive and specific enough for a diagnosis of vWD.

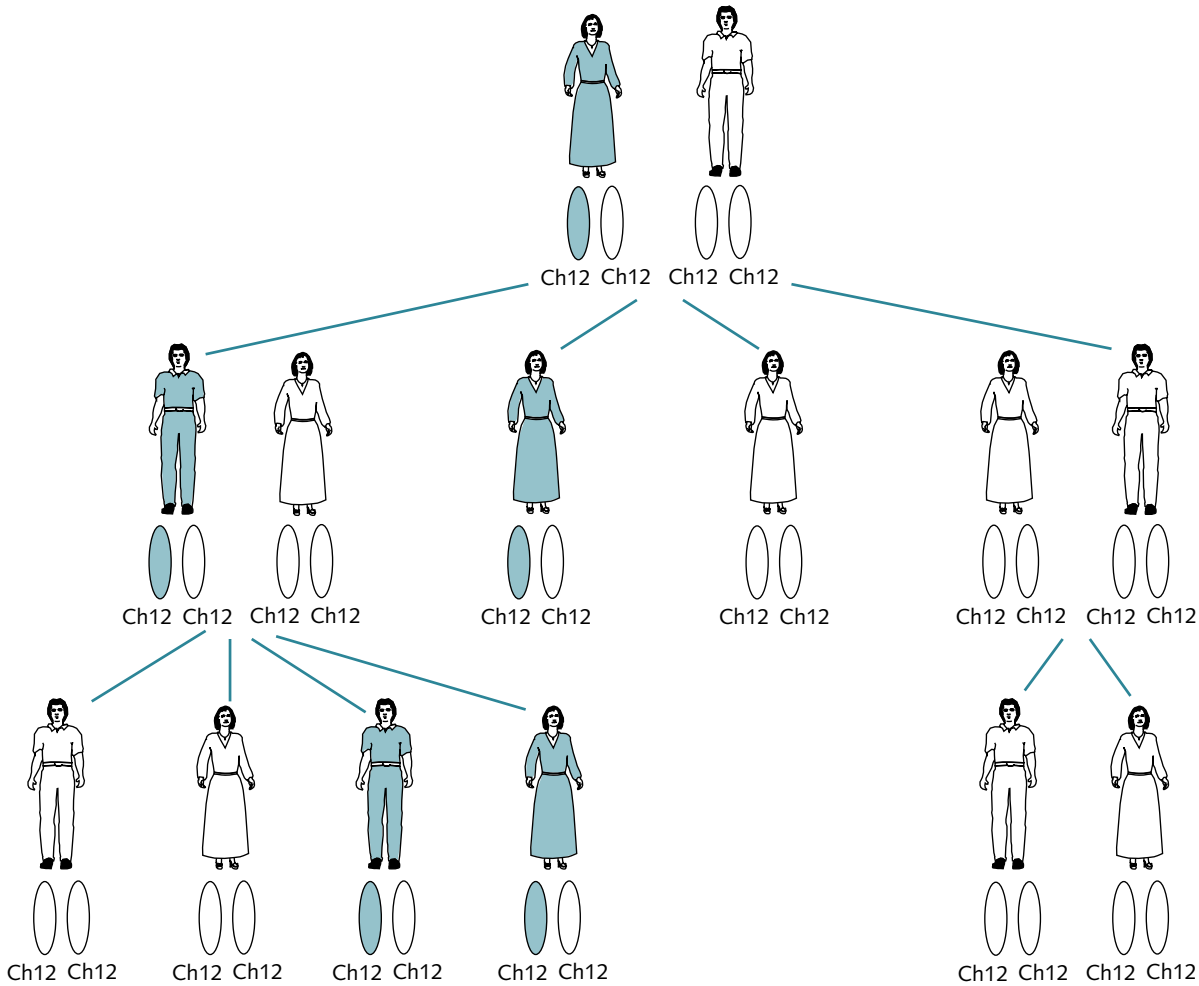
vWD is a highly heterogeneous disease in which there are quantitative and qualitative abnormalities of vWF, usually resulting from mutations at the *VWF* locus. The classification identifies three basic types (Table 17.1). Type 1 is characterized by the partial quantitative deficiency of vWF. Type 2 is characterized by qualitative abnormalities of the moiety and is further categorized into subgroups (2A, 2B, 2M and 2N) depending on the nature of the qualitative defect. Type 3 is the most severe form of the disease and is characterized by the complete absence of vWF in plasma and platelets. Transmission of type 1 vWD is usually dominant, type 2 is dominant in the majority of cases, whereas transmission of

Table 17.1 Current classification of von Willebrand disease.

Type 1	Partial quantitative deficiency of vWF
Type 2	Qualitative deficiency of vWF
2A	Decreased platelet-dependent vWF function, with lack of HMWM
2B	Increased vWF platelet-dependent vWF function, with lack of HMWM
2M	Decreased platelet-dependent vWF function, with normal multimeric structure
2N	Decreased vWF affinity for FVIII
Type 3	Complete deficiency of vWF

FVIII, factor VIII; HMWM, high molecular weight multimers; vWF, von Willebrand factor.

type 3 is recessive (Figure 17.4). The penetrance of disease due to the same mutation can be variable, even within the same family. In the last decade, many molecular defects of the *VWF* gene have been identified in vWD patients, mainly in the functional variants (types 2A, 2B and 2N). The unique phenotypes of these variants made it possible to restrict the genetic analysis to specific structural domains, such as the A1 domain for the variant 2B, A2 for the variant 2A and D′-D3 for the variant 2N. On the other hand, characterization of molecular defects in types 1 and 3 vWD requires extensive screening, since mutations are not restricted to specific regions and are usually scattered throughout the *VWF* gene. In most type 2 cases the mutations consist of amino acid substitutions (missense mutations), although small in-frame deletions or insertions have been reported. In type 3 vWD

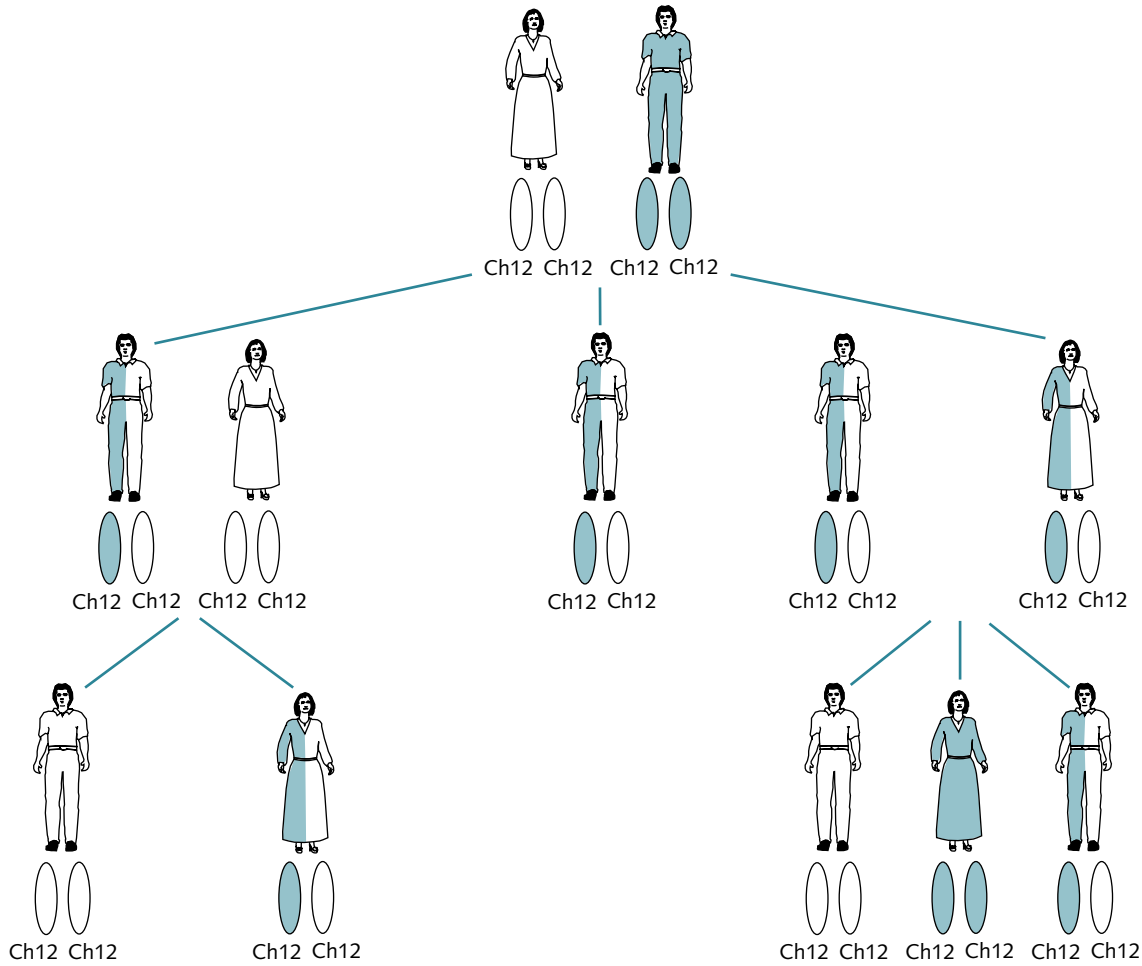


(a)

Fig. 17.4 Inheritance of vWD

(a) A family pedigree with autosomal dominant segregation, as found in vWD types 1, 2A, 2B and 2M.

(*Continued*.)



(b)

Fig. 17.4 (*Continued.*)
(b) A family pedigree with autosomal recessive segregation, as found in vWD types 2N and 3.

most of the molecular defects are responsible for null alleles (nonsense, splice site mutations, large gene deletions, small deletions and insertions).

Genetic defects in von Willebrand disease

Type 1 vWD

This is the most common form, accounting for approximately 70% of all cases of vWD. Sometimes type 1 vWD simply represents heterozygosity for a type 3 vWD allele. However, vWF levels in type 1 vWD patients are often considerably reduced below the expected 50% of normal. Individuals with type 1 vWD have very mild to moderate symptoms, a normal or variably prolonged bleeding time, decreased levels of vWF in plasma, and a multimeric structure of vWF. Type 1 vWD may be difficult to diagnose due to natural variations in circulating vWF levels related to the influence of environmental factors, such as exercise, thyroid hormone, estrogens and ABO blood type. vWF levels are lower by up to 30% in individuals with blood group O compared with individuals with other blood groups. Analysis of large kindreds previously classified as having type 1 vWD has shown the presence of phenotypic subtypes of type 1 vWD based upon the behavior of intraplatelet levels of vWF: platelet-normal, platelet-low and platelet-discordant. The mechanism of platelet-low vWD is thought to be defective production of vWF. Another mechanism, associated with platelet-normal vWD, is defective post-translational modification and release of vWF. Platelet-discordant vWD, characterized by normal platelet levels of vWF antigen but

low vWF activity, might be due to the latter mechanism or to enhanced clearance of the abnormal vWF.

Although type 1 is the most frequent form of vWD, few mutations have been identified; these are similar to those identified in type 3 vWD (deletions and nonsense and frameshift mutations). Several patients, clinically presenting as type 1, are compound heterozygotes for a null allele and a type 2N allele. Only two mutations have been shown to cause a clearly dominant type 1 vWD, with high penetrance and very low vWF levels. An *in vitro* study has clarified the molecular mechanism of the dominant negative missense mutations (C1149R and C1130F) found in the D3 domain. In a random dimerization process occurring in the endoplasmic reticulum, mutant and wild-type pro-vWF subunits form both homodimers and heterodimers. Retention of all mutant subunits in the endoplasmic reticulum will reduce the transport of wild-type subunits to the Golgi apparatus by about 50%, whereas subunits arriving in the Golgi apparatus assemble into large multimers and are normally secreted. However, dominant negative mutations appear to be unusual in type 1 vWD and do not always explain the incomplete penetrance and the high frequency of this disorder.

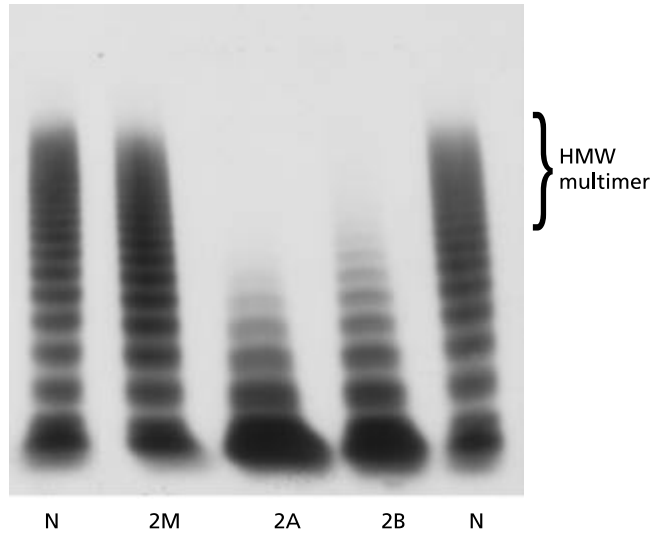


Fig. 17.5 Comparison of the multimeric structure of plasma vWF from vWD patients and normal subjects
Multimer analysis is performed with a low-resolution discontinuous buffer in a sodium dodecyl sulfate gel electrophoretic system with 1% agarose. Lanes 3 (vWD type 2A) and 4 (vWD type 2B) show the lack of high molecular weight multimers. Lane 2 (vWD type 2M) shows a normal multimeric pattern. Lanes 1 and 5 show a full multimeric structure from a normal control.

Type 2 vWD

Type 2 vWD is caused by qualitative abnormalities of vWF and accounts for approximately 20% of all vWD cases. Type 2 vWD is caused by a variety of mechanisms, and is associated with a variety of specific defects. This reflects the multifunctional role of vWF and the complicated post-translational processing of this moiety.

Type 2A

This is characterized by decreased platelet-dependent function due to the reduction or absence of high and intermediate molecular weight multimers (Figure 17.5). The bleeding diathesis is caused by the lack of the biologically active high molecular weight multimers. At least two mechanisms are known to produce type 2A vWD, subclassified as group 1 and group 2. Group 1 mutations lead to a vWF subunit that is presumably improperly folded, retained in the endoplasmic reticulum by the cell quality control machinery, and subsequently degraded. Multimers formed by the interactions between normal and mutant subunits are retained in the cell. The largest multimers, as a result of a greater likelihood of mutant subunit content, are more efficiently retained, accounting for the characteristic type 2A multimer pattern and autosomal dominant inheritance. The defect of group 2 does not interfere with biosynthesis and secretion, but instead renders the mutant subunit more susceptible to proteolytic cleavage in plasma by the metalloproteinase ADAMTS-13, which performs a similar function on wild-type vWF, although less efficiently. This enhanced plasma proteolysis results in the selective loss of the larger vWF multimers, very like what is seen in group 1. At least 24 mutations have been reported; they are responsible for both groups, and 20 of these are in the A2 vWF structural domain, the most frequent being R1597W and I1628T.

Type 2B

This uncommon subtype accounts for approximately 5% of all vWD cases. The 2B variant vWF shows increased affinity for platelet glycoprotein Ib and is usually associated with the absence of high molecular weight multimers in circulating plasma but not in platelets (Figure 17.5). The enhanced reactivity of the 2B vWF variant with its platelet receptor results in spontaneous binding to platelets and in thrombocytopenia which is thought to be secondary to the clearance of platelet–vWF complexes. Type 2B is inherited as a dominant trait. Mutations responsible for this subtype are nearly all contained in the A1 domain, which binds to the GpIb receptor. With a few exceptions, all mutations result in amino acid substitutions, and a few of them (R1306W, R1308C, V1316M and R1341Q) account for 80–90% of the reported cases. Type 2B vWD has sometimes been misdiagnosed and treated as autoimmune thrombocytopenia. Thrombocytopenia and the absence of the

larger multimers are not constantly present in type 2B vWD. A few variants have increased GpIb binding with no apparent loss of large multimers and thrombocytopenia, leading to earlier classification among type 1 variants (type 1 New York or type 1 Malmo). With the advent of molecular analysis, these mutations are now correctly classified as type 2B variants.

Type 2M

Type 2M vWD refers to qualitative variants with decreased platelet dependent function not caused by the absence of larger vWF multimers (Figure 17.5). With a pattern of vWF laboratory measurements similar to that of type 2A, type 2M shows a normal multimer distribution (hence M for multimer) or may even show the presence of ultralarge multimers (2M Vicenza). Type 2M is inherited as a dominant trait. The molecular defects are identified in the A1 domain, but in a region different from that of the 2B mutations. These defects, which in most cases are missense mutations, seem to downregulate the binding of the A1 domain to its platelet receptor. An important exception is the 2M Vicenza variant, the candidate mutation being in the D3 domain (R1205H). Furthermore, a second candidate mutation (M740I) has been identified recently in exon 17 in a few patients from the Vicenza area.

Type 2N

Type 2N vWD refers to all the qualitative variants characterized by decreased affinity for FVIII. The first description of this type of variant was that of a patient from Normandy (hence the N). This subtype is characterized by low levels of FVIII; however, vWF levels are usually normal and have an intact vWF multimeric pattern. It is difficult to distinguish type 2N from mild hemophilia A. Hemophilia is an X-linked disease, whereas type 2N vWD is an autosomal recessive disease. A definite diagnosis can be made by demonstrating the reduced binding of FVIII to native vWF with assays exploring this vWF property. Since the majority of mutations are located in the N-terminus of the mature vWF subunit (D′-D3), diagnosis can be confirmed by screening for mutations in exons 18, 19 and 20. Three mutations (T791M, R816W and R854Q) account for 90% of the type 2N mutations.

Type 3

Type 3 vWD is an autosomal recessive, clinically severe disorder characterized by the virtual absence of vWF. As in type 2N vWD, there is a secondary deficiency of FVIII and the patients therefore present a double defect of primary hemostasis and intrinsic coagulation. Type 3 vWD is rare and accounts for between 1 and 2% of all vWD cases, with a prevalence in the general population of 0.5–1 per million. The parents of type 3 patients are obligatory heterozygotes and in most cases are asymptomatic, but a minority of them have mild bleeding symptoms. Due to the large size of the gene, the presence of the pseudogene, the low prevalence of the disease and the absence of a specific localization of the mutations, the characterization of the molecular defects in type 3 has been relatively slow. Southern blot analysis first allowed the identification of large gene deletions, found in the majority of the patients who had developed alloantibodies to vWF after transfusion. However, the use of more specific and sensitive screening methods, such as single-strand conformational polymorphism, chemical cleavage mismatch analysis, and conformational sensitive gel electrophoresis, allowed the identification of most of the mutations in these patients. The most common mutations in type 3 vWD are nonsense mutations, small deletions, small insertions and splice-site mutations. Among the nonsense mutations, a few hotspot mutations at the arginine codons (R365X, R1659X, R1853X and R2535X) have been found repeatedly in different populations. A single cytosine deletion in a stretch of six cytosines in exon 18 appears to be particularly common in patients from Sweden and Germany. Although most of the mutations determine null alleles, a few in-frame deletions and several missense mutations have been identified. Patients with type 3 vWD may develop alloantibodies to vWF, which render replacement therapy ineffective. This complication is strongly associated with the presence of large gene deletions, although a few cases have been reported due to nonsense mutation.

Treatment of von Willebrand disease

While the treatment of patients with hemophilia A and B is facilitated by the close relationships existing between the content of FVIII or factor IX in the replacement material, the plasma levels attained after infusion and clinical efficacy, this model cannot be easily translated into the evaluation of products for the treatment of vWD, because it is still unclear which FVIII or vWF measurement in therapeutic products or in patient plasma correlates best with the severity of clinical bleeding and the efficacy of treatment. The situation is further complicated by the fact that vWD subtypes respond differently to treatment. Two main therapeutic agents are currently used to stop spontaneous bleeding and to prevent bleeding at the time of surgical procedures: the non-transfusional agent desmopressin and blood products that contain FVIII and vWF concentrated from plasma. Ancillary forms of treatment are platelet concentrates, synthetic fibrinolysis inhibitors and oral estrogen–progestogen preparations, which, in some clinical situations, are adjunctive or sometimes alternative to the two main treatments.

Desmopressin (1-deamino-8-D-arginine vasopressin, DDAVP) is a synthetic analog of the antidiuretic hormone vasopressin that, when administered to healthy volunteers or patients with mild hemophilia and vWD, increases FVIII and vWF transiently by releasing these moieties from storage sites into the plasma. Endothelial cell Weibel–Palade bodies appear to be the source of vWF, but the source of FVIII has not yet been determined. Desmopressin induces vWF release into plasma by binding to the vasopressin V2 receptor and thereby activating cyclic AMP-mediated signaling in vascular endothelial cells.

The advantage of this compound is that it is relatively inexpensive and carries no risk of transmitting blood-borne infectious agents. When infused intravenously over 30 minutes at a dose of 0.3 μg/kg, desmopressin is expected to increase plasma FVIII and vWF three- to five-fold above the basal levels. In general, high FVIII/vWF concentrations last for at least 8–10 hours in plasma. Patients with baseline plasma levels of FVIII/vWF measurements in the range of 10–20 IU/dl or more are those who are more likely to reach desmopressin levels sufficient to attain hemostasis after desmopressin treatment, taking into account variables such as the type and severity of the bleeding episode and the levels of FVIII/vWF that must be attained and maintained to secure hemostasis. Even though most patients with mild hemophilia A that is treated repeatedly with desmopressin become less responsive to therapy, this problem is less frequent and prominent in patients with type 1 vWD. The drug is also available in concentrated forms for subcutaneous and intranasal administration (at doses of 0.3 μg/kg and 150–300 μg, respectively), which can be convenient for home treatment.

Side effects of desmopressin are usually mild tachycardia, headache and flushing. Hyponatremia and volume overload due to the antidiuretic effect of desmopressin are relatively rare if fluid intake is not excessive during treatment. Even though no thrombotic episodes have been reported in vWD patients treated with desmopressin, this compound should be used with caution in elderly patients with cardiovascular disease, because a few cases of myocardial infarction and stroke have occurred in treated patients with hemophilia and uremia. Desmopressin has little or no oxytocic activity and has been used by us without mishap during the early period of pregnancy in 31 women with low FVIII levels (including carriers of hemophilia A and vWD patients) to prevent bleeding at the time of invasive diagnostic procedures such as chorionic villus sampling and amniocentesis.

Desmopressin is most effective in patients with type 1 vWD, particularly in those who have releasable vWF in storage sites, a condition usually reflected by normal vWF levels in platelets. In these patients FVIII, vWF and the bleeding time are usually corrected to normal values by desmopressin. In other vWD subtypes, responsiveness is varied (Table 17.2).

Table 17.2 Indications for desmopressin in different types of von Willebrand disease.

Type of disease	Response
1	Usually effective
2A	Usually ineffective
2B	May be contraindicated
2M	Predicted to be ineffective
2N	Partially effective
3	Ineffective

A poor and short-lasting response is seen in patients with the variant of type 1 vWD characterized by low levels of platelet vWF, perhaps because low levels in platelets are paralleled by low levels of releasable vWF in storage sites. In type 2A, FVIII levels are usually increased by desmopressin but the bleeding time is shortened in only a minority of cases. Desmopressin is contraindicated in type 2B, because of the transient appearance of thrombocytopenia. There is little experience in type 2M, but a poor response is predicted because vWF is dysfunctional in this subtype. In type 2N, levels of FVIII clotting activity increase after desmopressin, but released FVIII circulates for a relatively short time period in patients' plasma because the stabilizing effect of the abnormal vWF on FVIII is impaired. Therefore, plasma concentrates containing FVIII and vWF are preferable. Patients with type 3 vWD are usually unresponsive to desmopressin, because they lack releasable stores of vWF.

Transfusional therapy with plasma products containing both FVIII and vWF is the treatment of choice when bleeding occurs or must be prevented and the predicted response to desmopressin is considered suboptimal for hemostasis. FVIII and vWF may be infused as fresh frozen plasma (FFP) but the large volumes required limit its use. Cryoprecipitate contains five to ten times more FVIII and vWF than FFP (each bag contains approximately 80–100 IU). Early studies indicated that cryoprecipitate administered every 12–24 hours normalized plasma FVIII levels and stopped or prevented bleeding in vWD. On the basis of these observations, cryoprecipitate has been the mainstay of treatment for many years. However, virucidal methods cannot be applied to cryoprecipitate as currently produced by blood banks, so that this product carries a small but definite risk of transmitting blood-borne infectious agents. Therefore, virus-inactivated FVIII/vWF concentrates, originally developed for the treatment of hemophilia A, are currently perceived as safer and are preferred in the management of vWD patients unresponsive to desmopressin.

Two commercially available concentrates have been evaluated more extensively than others and clinical studies have demonstrated their efficacy in preventing or stopping bleed-

ing. One, licensed in the USA and in several European countries for the treatment of vWD, contains a relatively larger amount of vWF (measured as vWF:RCo) than of FVIII (approximately two to three times more in terms of IU). The virucidal method adopted is pasteurization. The other, licensed only in Europe so far, differs because it contains similar relative amounts of FVIII and vWF:RCo. Two virucidal methods, solvent/detergent and heating at high temperatures, are included in the manufacturing step, with the goal of inactivating both enveloped and non-enveloped virus. Other virally inactivated FVIII/vWF concentrates have been successfully employed in vWD patients, but clinical experience is more limited. Recently, a chromatography-purified concentrate that is particularly rich in vWF but has a low FVIII content has also been produced and evaluated. This concentrate was clinically efficacious when tested in a small number of type 3 vWD patients. Efficacy and safety are now under evaluation in Europe in larger series of patients.

The dosages recommended for the control or prevention of bleeding are summarized in Table 17.3. Dosages are expressed in IU/kg of FVIII because most of the available concentrates, being manufactured for the treatment of patients with hemophilia A, are labeled in terms of FVIII content only. Since FVIII has a longer half-life in vWD patients than in patients with hemophilia A (20–24 versus 12–14 hours), the infusion of one daily dose is sufficient to reach and maintain adequate plasma levels for the treatment of spontaneous bleeding episodes and to prevent excessive bleeding until healing is complete, depending on the site and extent of surgery. Since in the USA the Food and Drug Administration requires that plasma products licensed for treatment of vWD patients are labeled in terms of the actual defective protein to be replaced, the solvent/detergent, heat-treated concentrate is labeled in terms of vWF:RCo content. The doses of this concentrate recommended for their demonstrated efficacy in a large, prospective clinical trial are 40–60 IU/kg of vWF:RCo (50–75 IU/kg in children because of the lower *in vivo* recovery), which usually results in plasma levels of vWF:RCo of 80–120 IU/dl or higher. Since the plasma half-life of vWF:RCo is much shorter than that of FVIII:C (6–8 versus 20–24 hours), usually these doses should not be repeated more often than every 24 hours in order to avoid very high levels of FVIII:C, which may engender venous thromboembolic complications.

It is usually not necessary to carry out laboratory tests to monitor replacement therapy in patients with spontaneous bleeding episodes. For surgical procedures we recommend measuring FVIII every 12 hours on the operation day and then every 24 hours. The FVIII response can be predicted on the basis of pharmacokinetic data which indicate that 1 IU/kg will increase plasma FVIII levels by approximately 2 IU/dl (1.5 IU/dl in children). Those who use concentrates labeled in terms of vWF:RCo content may choose to monitor the plasma level of this moiety, although this is more complex to measure in the clinical setting and less standardized than the FVIII level. It remains to be demonstrated whether newer laboratory measurements, such as the collagen binding assay, will be simpler and more predictive of outcome.

Monitoring the bleeding time is usually not necessary. The prolonged bleeding time is frequently not normalized or even shortened in patients treated with FVIII/vWF concentrates. There are probably multiple reasons for the inconsistent effects of plasma products on the bleeding time. So far, no concentrate contains a fully functional vWF, as tested *in vitro* by evaluating the multimeric pattern and using functional assays. Despite no or partial correction of the bleeding time, major surgical procedures are successfully carried out and spontaneous bleeding episodes controlled following the infusion of FVIII/vWF concentrates. In the relatively rare instances when bleeding is not controlled and the bleeding time remains prolonged, platelet concentrates (given immediately after FVIII/vWF-containing preparations, at doses of 4 to 5×10^{11} platelets) are effective, particularly in patients with type 3 vWD, both in terms of bleeding time correction and the control of hemorrhages. Platelets from type 3 vWD patients

Table 17.3 Dosages of FVIII coagulant activity (FVIII:C) recommended in patients with von Willebrand disease treated with FVIII/vWF concentrates.*

Type of bleeding	Dose (IU/kg)	Number of infusions	Target
Major surgery	40–60	Once a day	Maintain plasma FVIII:C >50 IU/dl until healing is complete depending on the type of surgery
Minor surgery	30–50	Once a day or every other day	FVIII:C >30 IU/dl until healing is complete depending on the type of surgery
Dental extractions	20–30	Single	FVIII:C >30 IU/dl for at least 12 hours
Spontaneous bleeding episodes	20–30	Single	FVIII:C >30 IU/dl

*For concentrates labeled in terms of vWF:RCo, the recommended doses for adults, the number of infusions and the target plasma levels are the same as those for FVIII:C.

Table 17.4 Summary of management of different types and subtypes of von Willebrand disease.

	Treatment of choice	Alternative or adjunctive therapy
Type 1	Desmopressin	Antifibrinolytic amino acids
Type 2A	Factor VIII/vWF concentrates	Antifibrinolytic amino acids
Type 2B	Factor VIII/vWF concentrates	Antifibrinolytic amino acids
Type 2M	Factor VIII/vWF concentrates	Antifibrinolytic amino acids
Type 2N	Factor VIII/vWF concentrates	Antifibrinolytic amino acids
Type 3	Factor VIII/vWF concentrates	Antifibrinolytic amino acids, platelet concentrates
Type 3 complicated by alloantibodies	Recombinant factor VIII	Recombinant activated factor VII

lack vWF completely and there is no uptake of the protein from plasma after infusion of concentrates. The hemostatic effectiveness of the transfusion of normal platelets is likely to be due to the fact that these cells transport and localize vWF at sites of vascular injury. From a practical standpoint, it must be emphasized that in the largest prospective study carried out so far in vWD patients platelet concentrates became necessary to prevent or stop bleeding in one case only.

The different options available at the moment for the management of von Willebrand disease are summarized in Table 17.4. Treatment of spontaneous bleeding episodes and their prevention at the time of an invasive procedure is relatively simple and can certainly be tackled by the average clinical hematologist with access to a minimum of laboratory testing (FVIII assays). However, the patients need to be well characterized phenotypically because the choice of treatment must be tailored to the different types and subtypes of the disease. Such characterization is not simple, so that at most clinical centers it is probably not worthwhile setting up relatively complicated tests, such as multimer analysis and the vWF: RCo assay, when samples can be sent for analysis to more expert laboratories that have become proficient during the study of large series of patients.

Von Willebrand disease resources on the Internet

ISTH SSC vWF: *www.shef.ac.uk/vwf/index.html*

The Human Gene Mutation Database, Cardiff: *archive.uwcm.ac.uk/uwcm/mg/search/119125.html*

OMIM™ On line Mendelian Inheritance in Man: *www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=193400*

GeneCards™ is a database of human genes: *bioinfo.weizmann.ac.il/cards-bin/carddisp?vWF*

Further reading

General

Michiels JJ (ed.) (2001) Von Willebrand factor and von Willebrand disease. *Best Practice & Research. Clinical Haematology,* **14**, 235–462.

Sadler JE, Mannucci PM, Berntorp E *et al.* (2000) Impact, diagnosis and treatment of von Willebrand Disease. *Thrombostasis and Haemostasis*, **84**, 160–174.

Von Willebrand factor

Mancuso DJ, Tuley EA, Westfield LA *et al.* (1989) Structure of the gene for human von Willebrand factor. *Journal of Biological Chemistry*, **264**, 19514–19527.

Mancuso DJ, Tuley EA, Westfield LA *et al.* (1991) Human von Willebrand factor gene and pseudogene: structural analysis and differentiation by polymerase chain reaction. *Biochemistry*, **30**, 253–269.

Shelton-Inloes BB, Titani K, Sadler JE. (1986) cDNA sequences for human von Willebrand factor reveal five types of repeated domains and five possible protein sequence polymorphisms. *Biochemistry*, **25**, 3164–3171.

Wagner DD, Marder VJ. (1984) Biosynthesis of von Willebrand protein by human endothelial cells: processing steps and their intracellular localization. *Journal of Cell Biology*, **99**, 2123–2130.

Von Willebrand disease and its classification

Nichols WC, Ginsburg D. (1997) von Willebrand disease. *Medicine (Baltimore)*, **76**, 1–20.

Sadler JE. (1994) A revised classification of von Willebrand disease. *Thrombosis and Haemostasis*, **71**, 520–525.

Genetic defects in von Willebrand disease

Dent JA, Berkowitz SD, Ware J *et al.* (1990) Identification of a cleavage site directing the immunochemical detection of molecular abnormalities in type IIA von Willebrand factor. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 6306–6310.

Eikenboom JCJ, Matsushita T, Reitsma PH *et al.* (1996) Dominant type I von Willebrand disease caused by mutated cysteine residues in the D3 domain of von Willebrand factor. *Blood*, **88**, 2433–2441.

Keeney S, Cumming AM. (2001) The molecular biology of von Willebrand disease. *Clinical and Laboratory Haematology*, **23**, 209–230.

Lyons SE, Bruck ME, Bowie EJW *et al.* (1992) Impaired intracellular transport produced by a subset of type IIA von Willebrand disease mutations. *Journal of Biological Chemistry*, **267**, 4424–4430.

Mazurier C, Diéval J, Jorieux S *et al.* (1990) A new von Willebrand factor (vWF) defect in a patient with factor VIII (FVIII) deficiency but with normal levels and multimeric patterns of both plasma and platelet vWF. Characterization of abnormal vWF/FVIII interaction. *Blood*, **75**, 20–26.

Saba HI, Saba SR, Dent JA *et al.* (1985) Type IIB Tampa: a variant of von Willebrand disease with chronic thrombocytopenia, circulating platelet aggregates, and spontaneous platelet aggregation. *Blood*, **66**, 282–286.

Schneppenheim R, Federici AB, Budde U *et al.* (2000) von Willebrand disease type 2M 'Vicenza' in Italian and German patients: identification of the first candidate mutation (G3864A; R1205H) in 8 families. *Thrombostasis and Haemostasis*, **83**, 136–140.

Shelton-Inloes BB, Chehab FF, Mannucci PM *et al.* (1987) Gene deletions correlate with the development of alloantibodies in von Willebrand disease. *Journal of Clinical Investigation*, **79**, 1459–1465.

Zhang ZP, Blombäck M, Egberg N *et al.* (1994) Characterization of the von Willebrand factor gene (vWF) in von Willebrand disease type III patients from 24 families of Swedish and Finnish origin. *Genomics*, **21**, 188–193.

Treatment of von Willebrand disease

Castillo R, Escolar G, Monteagudo J *et al.* (1987) Hemostasis in patients with severe von Willebrand disease improves after normal platelet transfusion and normalizes with further correction of the plasma defect. *Transfusion*, **37**, 785–790.

Castillo R, Monteagudo J, Escolar G *et al.* (1991) Hemostatic effect of normal platelet transfusion in severe von Willebrand disease patients. *Blood*, **77**, 1901–1905.

Chang AC, Rick ME, Ross PL *et al.* (1998) Summary of a workshop on potency and dosage of von Willebrand factor concentrates. *Haemophilia*, **4** Supplement 3, 1–6.

Dobrkovska A, Krzensk U, Chediak JR. (1998) Pharmacokinetics, efficacy and safety of Humate-P in von Willebrand disease. *Haemophilia*, **4** Supplement 3, 33–39.

Goudemand J, Negrier C, Ounnoughene N *et al.* (1998) Clinical management of patients with von Willebrand's disease with a VHP vWF concentrate: the French experience. *Haemophilia*, **4** Supplement 3, 48–52.

Kaufmann JE, Oksche A, Wollheim CB *et al.* (2000) Vasopressin-induced von Willebrand factor secretion from endothelial cells involves V2 receptors and cAMP. *Journal of Clinical Investigation*, **106**, 107–116.

Mannucci PM. (1997) Desmopressin (DDAVP) in the treatment of bleeding disorders: the first 20 years. *Blood*, **90**, 2515–2521.

Mannucci PM, Federici AB. (1995) Antibodies to von Willebrand factor in von Willebrand disease. In: Aledort LM, Hoyer LW, Reisner HM *et al.* (eds). *Inhibitors to Coagulation Factors*, Volume 386. New York: Plenum Press, pp. 87–92.

Mannucci PM, Ruggeri ZM, Pareti FI *et al.* (1977) 1-Deamino-8-D-arginine vasopressin: a new pharmacological approach to the management of haemophilia and von Willebrand disease. *Lancet*, **i**, 869–872.

Mannucci PM, Bettega D, Cattaneo M. (1992) Consistency of responses to repeated DDAVP infusions in patients with severe vWD and alloantibodies to vWF. *European Journal of Haematology*, **82**, 87–93.

Mannucci PM, Tenconi PM, Castaman G *et al.* (1992) Comparison of four virus-inactivated plasma concentrates for treatment of severe von Willebrand disease: a cross-over randomized trial. *Blood*, **79**, 3130–3137.

Mannucci PM, Chediak J, Hanna W *et al.* and the Alphanate Study Group. (2002) Treatment of von Willebrand disease with a high-purity factor VIII/von Willebrand factor concentrate: a prospective multicenter study. *Blood*, **99**, 450–456.

Mazurier C, Gaucher C, Jorieux S *et al.* (1994) Biological effect of desmopressin in eight patients with type 2N ('Normandy') von Willebrand disease. *British Journal of Haematology*, **88**, 849–854.

Rodeghiero F, Castaman G, Meyer D *et al.* (1992) Replacement therapy with virus-inactivated plasma concentrates in von Willebrand disease. *Vox Sanguinis*, **62**, 193–199.

Chapter 18 Platelet disorders

Katherine A Downes & Keith R McCrae

Introduction

Hemostasis may be considered as consisting of primary and secondary phases. In primary hemostasis, platelets adhere to a site of vascular damage, initiating the formation of a platelet plug. Since this plug is unstable, it is subsequently stabilized through the process of secondary hemostasis, in which soluble coagulation factors assemble on the surface of adherent platelets, leading to activation of coagulation pathways and the formation of fibrin. The focus of this chapter will be on primary hemostasis and how genetic control of platelet glycoprotein expression affects this process. We will also review recent data concerning the impact of platelet glycoprotein polymorphisms on thrombotic disorders.

Normal platelet function

Primary hemostasis may be subdivided into three processes: the *adhesion* of platelets to exposed subendothelium; the *activation* of bound platelets; and the *aggregation* of additional platelets to form the platelet plug. Each of these involves specific platelet glycoproteins whose expression and function is genetically regulated.

Platelet adhesion involves the adherence of platelets to exposed subendothelial matrix. Under conditions of high shear stress, such as those in the arterial circulation, this process is critically dependent upon the interaction between the platelet glycoprotein (GP) Ib–IX–V complex and von Willebrand factor (vWF). Though vWF circulates in plasma in a largely inactive form, it adopts an active conformation capable of interacting with GPIb–IX–V after its incorporation into the subendothelial matrix and, in particular, its binding to collagen. Under high shear stress, the interaction between GPIb–IX–V and vWF leads to the 'tethering' of platelets to the exposed subendothelium, with more stable adhesion and spreading of adherent platelets promoted by additional interactions between platelet integrins, such as GPIa–IIa and GPIIb–IIIa, with their respective ligands (Figure 18.1).

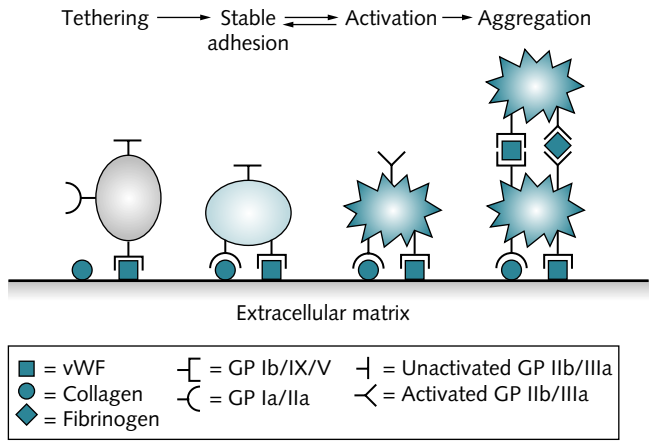


Fig. 18.1 Platelet thrombus formation
Following vessel wall injury, platelets are initially tethered to the subendothelial matrix. Under high-shear conditions, this interaction occurs primarily through binding of GPIb–IX–V to von Willebrand factor (vWF) incorporated into the matrix. Stable platelet adhesion requires additional binding of collagen to GPIa-IIa, as well as additional interactions such as the binding of GPIIb–IIIa to vWF and other ligands. Platelet activation is initiated upon tethering and is amplified upon binding to collagen. Platelet activation results in shape change, secretion of the contents of intracellular storage granules (vWF, fibrinogen, and agonists such as ADP), expression of the activated form of GPIIb–IIIa and other activation markers, and expression of platelet procoagulant activity (primarily anionic phospholipid). Binding of vWF and fibrinogen to activated platelets results in platelet aggregation, which involves interactions of vWF with GPIb–IX–V and of fibrinogen with GPIIb–IIIa on adjacent platelets. From Reiner AP, Siscovick DS, Rosendaal FR. (2001) Platelet glycoprotein gene polymorphisms and risk of thrombosis: facts and fancies. *Reviews in Clinical and Experimental Hematology*, **5**, 262–287, with permission.

GPIa-IIa (integrin $\alpha_2\beta_1$) is the major platelet collagen receptor, and GPs VI and perhaps IV may function as collagen receptors as well. The affinity of GPIa-IIa for collagen is increased during platelet activation (see below), perhaps through a mechanism involving protein disulfide isomerase activity on cysteine residues in the β_1 chain. GPIa-IIa may also transmit signals following engagement of collagen. Two female patients have been described who had mild bleeding disorders, with deficient collagen-induced platelet aggregation, and absent platelet GPIa-IIa; this deficiency resolved after the menopause, however, making a genetic basis uncertain. GPVI belongs to the immunoglobulin superfamily. Its primary role in normal platelet function is to signal through its Fcγ subunit, and although it binds collagen through its two immunoglobulin C2 loops, whether GPVI has a *direct* role in platelet adhesion or contributes to adhesion by promoting activation of GPIa-IIa remains uncertain. However, optimal platelet aggregation responses to collagen require both receptors, and patients with mild bleeding disorders associated with deficiencies of GPVI have been described. Finally, CD36, also known as GPIV or GPIIIb, may also function as a collagen receptor on platelets, though this receptor also binds other ligands, and its importance as a platelet collagen receptor is not well understood.

The second phase of primary hemostasis involves platelet activation, an energy-dependent process characterized by changes in the shape of the platelets and the secretion of platelet alpha and dense granule contents. Platelet activation depends upon the activation of signal transduction pathways following receptor ligation, and on the effects of thrombin on protease-activated receptors (PAR-1 and PAR-4) and GPIb. ADP released from dense granules also interacts with platelet purinergic receptors, further promoting platelet activation through a number of pathways, including inhibition of platelet adenyl cyclase. PAR-1 and -4 do not stimulate G(i) signaling pathways in the absence of secreted ADP, and cause human platelet aggregation independently of G(i) signaling.

The final step in platelet plug formation is platelet aggregation, the process by which adherent platelets aggregate with one another and with other platelets not bound to exposed subendothelium. Platelet aggregation is highly dependent upon platelet GPIIb–IIIa (integrin $\alpha IIb\beta_3$), and classically has been thought to involve the binding of symmetrical ends of the bipolar fibrinogen molecule to GPIIb–IIIa receptors on adjacent platelets. Recent studies, however, have demonstrated a role for other ligands, such as vWF and thrombospondin, as well as platelet GPIb–IX–V, in platelet aggregation (Ruggeri, 2000). The relative importance of these ligand–receptor systems depends on the degree of local shear and the thrombin activation state.

Evaluation of platelet function and identification of molecular defects

Most patients with clinically significant platelet dysfunction present with bleeding. 'Platelet-type' bleeding may consist of easy or excessive bruising, epistaxis, gingival bleeding, gastrointestinal bleeding, menorrhagia, or excessive intra- or postoperative bleeding. Patients with such a history (assuming normal coagulation studies) should be evaluated with a complete blood count (to exclude thrombocytopenia), a careful review of the peripheral blood film (to assess platelet morphology), and a screening test of platelet function, such as the bleeding time or a platelet function analyzer (PFA-100). For those whose screening studies suggest platelet dysfunction, measurement of platelet aggregation in response to platelet agonists, such as epinephrine, arachidonic acid, collagen, A23187 (a thromboxane analog) and adenosine diphosphate, as well as platelet agglutination in response to ristocetin, are employed. For patients in whom platelet aggregation studies suggest a functional defect potentially attributable to a specific receptor, flow cytometry using specific monoclonal antibodies can be used to determine the expression level of the glycoprotein in question.

Defective expression or function of platelet glycoproteins may be further characterized using molecular approaches. These may include reverse transcription of platelet mRNA followed by amplification of platelet glycoprotein cDNAs and analysis of the polymerase chain reaction (PCR) products. However, since the isolation of sufficient platelet RNA to perform such procedures may be difficult, coding regions (exons) within genes of individual platelet glycoproteins may also be amplified from genomic DNA. When searching for common or previously described mutations within a glycoprotein gene, procedures such as single-strand conformation polymorphism (SSCP), allele-specific PCR, and altered sensitivity of amplified DNA containing the mutation of interest to digestion by specific restriction endonucleases, can be employed. Once a mutation is defined in the proband, additional studies to determine its presence in family members may be indicated.

Disorders of platelet adhesion

The Bernard–Soulier syndrome

The Bernard–Soulier syndrome (BSS) is a rare bleeding disorder caused by abnormal expression or function of the GPIb–IX–V complex. BSS is usually inherited in an autosomal recessive manner, and when homozygous is often associated with consanguinity. Patients with BSS may present with giant

platelets (up to 20 μm in size), a prolonged bleeding time, and thrombocytopenia. Platelets from patients with BSS do not agglutinate in response to ristocetin, and their aggregation response to thrombin is delayed and reduced in magnitude due to the loss of a high-affinity thrombin-binding site within the N-terminal domain of GPIbα.

Genetics of the GPIb–IX–V complex

The four membrane-spanning polypeptides that constitute the GPIb–IX–V complex are encoded by different genes: GPIbα (chromosome 17 ptr-p12), GPIbβ (chromosome 22q11.2), GPIX (chromosome 3Q21) and GPV (chromosome 3Q29). The coding sequences of each of these are contained within single exons with the exception of the GPIbβ gene, which has an intron 10 base pairs after the transcriptional start site of the first exon. All four glycoprotein subunits are members of the leucine-rich repeat (LRR) family. The GPIb–IX–V ligand-binding domain resides largely within the N-terminal 282 residues of GPIbα, a region that contains seven tandem leucine-rich repeats, of which repeats 2–4 are critical for vWF binding. On the platelet surface, GPIbα is disulfide-linked to GPIbβ and non-covalently associated with GPIX and GPV in the ratio of 2:2:2:1 (Figure 18.2a).

Patients with BSS are usually homozygous or doubly heterozygous for point mutations or deletions in the genes encoding GPIbα, GPIbβ or GPIX. Single heterozygotes generally experience mild, if any, hemorrhage. Mutations may cause a synthetic defect resulting in absent or diminished amounts of a glycoprotein or a functional defect resulting in inability of the GPIb–IX–V complex to bind vWF. The resulting phenotype may range from essentially normal, through isolated giant platelets, to classic BSS with associated bleeding.

Surface expression of the GPIb–IX–V complex requires the presence of GPIbα, GPIbβ and GPIX. A mutation in any of the genes encoding these proteins may prevent proper post-translational modification, assembly, or membrane insertion of not only the affected protein but other proteins that constitute the complex. For example, a patient with BSS resulting from a G→A substitution at position 159 of the GPIbβ gene that caused premature polypeptide termination had undetectable levels of the GPIb–IX–V complex and absent GPIbα on the platelet surface, though GPIbα was present within the platelets and in plasma. Co-expression of the mutant GPIbβ with wild-type GPIbα and GPIX in HEK 293 cells was also associated with failure to express the GPIbα and GPIX proteins on the cell surface. Similarly, a homozygous missense mutation in the GPIbβ gene leading to a Tyr→Cys substitution at residue 88 diminished the expression of the GPIb–IX–V complex, with no detectable GPIX or GPIbβ, and diminished expression of GPIbα on the platelet surface. Transfection studies demonstrated that both Tyr88Cys and Tyr88Ala mu-

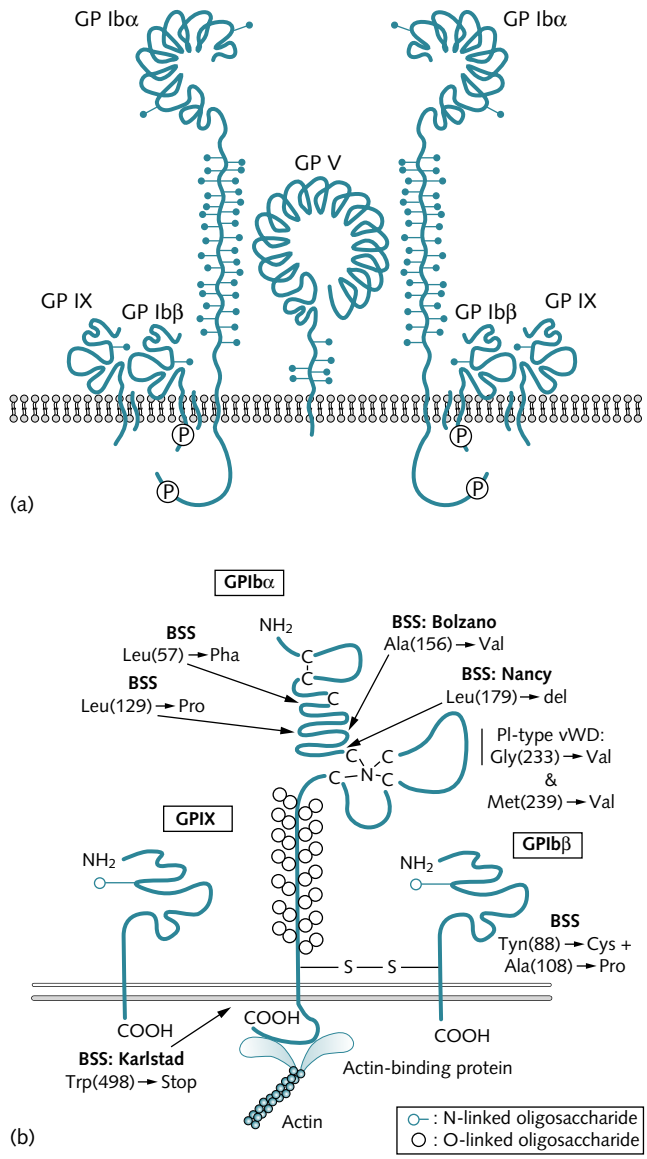


Fig. 18.2 Structure of the GPIb–IX–V complex (a) and locations of specific variant mutations (b)
(a) In this complex, GPIbα is disulfide-linked to GPIbβ and non-covalently associated with GPIX and GPV, in the ratio of 2:2:1. Phosphorylation sites in cytoplasmic tails of GPIbα (Ser 609) and GPIbβ (Ser 166) are also indicated. From Berndt MC, Shen Y, Dopheide SM *et al*. (2001) The vascular biology of the glycoprotein Ib–IX–V complex. *Thrombosis and Haemostasis*, **86**, 178–188, with permission. (b) Gene defects known to cause variant BSS. These mutations, primarily amino acid changes, lead to a loss of GPIbα function despite the expression of the GPIb–IX–V complex on the platelet surface at levels normally sufficient to support platelet adhesion. Also shown are the two mutations that lead to platelet-type von Willebrand disease. From Nurden AT, Nurden P. (2001) Inherited defects of platelet function. *Reviews in Clinical and Experimental Hematology*, **54**, 314–334, with permission.

tations in GPIbβ suppressed the expression of GPIb–IX–V. Interestingly, four individuals with giant platelets but normal platelet aggregation studies (Blood 1997) were heterozygous for Tyr88Cys in GPIbβ.

Mutations associated with the Bernard–Soulier syndrome

In the following sections, we will describe some of the better characterized mutations in the genes encoding GPIbα, GPIbβ and GPIX that have been associated with the BSS. This discussion is not designed to be inclusive, as an updated list of reported mutations associated with the BSS is maintained at http://bernard-soulier.org. As discussed, no GPV mutations leading to BSS have been described.

GPIbα. Mutations in the GPIbα gene are the most common cause of BSS. Homozygous defects leading to this disorder have included point mutations such as Leu129→Pro or Ala156→Val (Bolzano variant), nonsense mutations such as Trp498→stop, W343→stop or Ser444→stop (Kagoshima variant), and deletions such as a deleted T in codon 76 causing a frameshift and truncation after an additional 19 residues, or deletion of Leu179 in the seventh leucine-rich repeat of GPIbα (Nancy I variant). Platelets from a patient with a Leu129→Pro mutation manifested a 60% decrease in the surface expression of GPIb-IX and diminished binding of vWF, but normal expression of GPV. These platelets did not agglutinate with ristocetin or vWF, suggesting that the Leu129→Pro mutation affected not only the surface expression but function of GPIb–IX–V, perhaps through conformational changes in GPIb.

Several homozygous mutations within GPIbα have been localized through the use of a molecular model to an area containing the seven leucine-rich repeats of this glycoprotein. This model has allowed a structural analysis of the mechanisms by which these mutations disrupt receptor function (Figure 18.3), and provides a potential explanation for the species-specific interactions between vWF and GPIb–IX–V.

Compound heterozygous mutations in GPIbα that have been associated with BSS include a change in a Trp343 (TGG) codon to a nonsense (TGA) codon in one allele, accompanied by an unknown defect in the other. Kanaji *et al.* (1997) described a patient with BSS with an insertion of T at position 1418 of one allele, and a deletion of a single adenine from a seven-adenine repeat spanning positions 1438–1444 in the other, each mutation leading to a frameshift and premature polypeptide termination.

Mutations in GPIbα may also cause 'variant' BSS syndromes in which dysfunctional GPIb–IX–V complexes are expressed on the platelet surface (Figure 18.2b). Frequently these mutations occur in the leucine-rich repeat (LRR) sequence of the gene. Mutations responsible for variant BSS include missense mutations such as Leu57→Phe, homozygous defects such as deletion of Leu179 (Nancy I variant), a

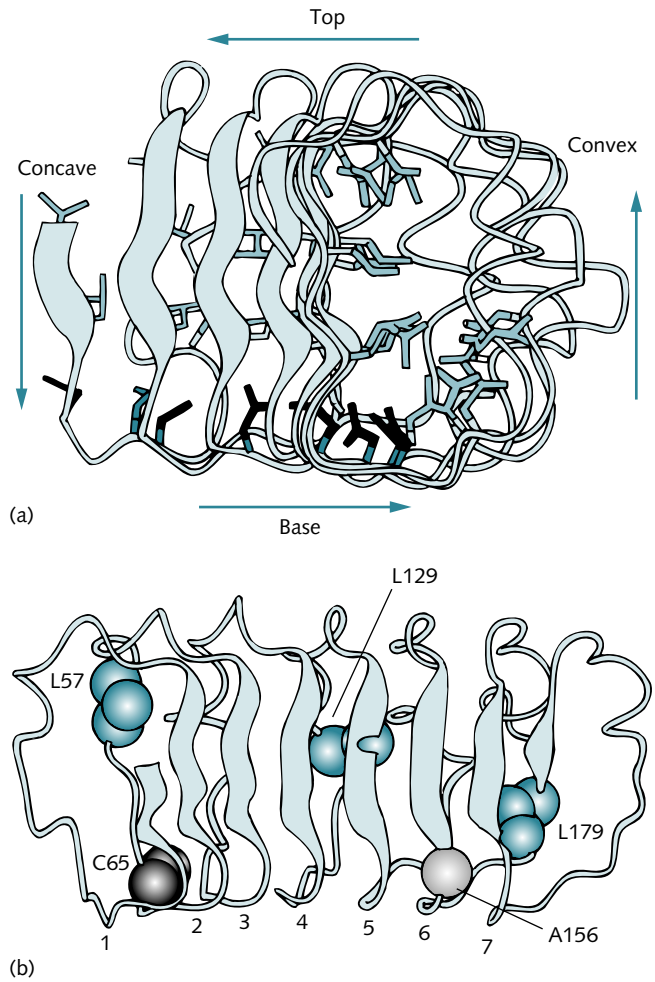


Fig. 18.3 Molecular model of the human GPIbα leucine-rich repeats

(a) Side view. Leucines of the repeat are shown in dark blue and conserved asparagine residues are shown in black. The four sides of the solenoid-like structure consisting of the seven leucine-rich repeats are labeled 'top', 'base', 'concave' and 'convex'. (b) The concave face of the solenoid, with the individual repeats numbered and the positions of mutations associated with BSS shown in van der Waal spheres: Leu57→Phe; Cys65→Arg; Leu129→Pro; Ala156→Val; Leu179 deletion. From Whisstock JC, Shen Y, Lopez JA *et al*. (2002) Molecular modeling of the seven tandem leucine-rich repeats within the ligand-binding region of platelet glycoprotein Ib alpha. *Thrombostasis and Haemostasis*, **87**, 329–333, with permission.

C→T substitution at nucleotide 515 leading to substitution of Ala for Val at codon 156 (Bolzano variant), and Trp498→stop (Karlstad variant). The Leu57→Phe mutation is characterized by enhanced proteolytic susceptibility of GPIbα, and is the only autosomally dominant inherited form of BSS. In the Nancy I variant, the deletion of a leucine residue in the leucine-rich region of GPIbα results in incomplete glycosyl-

ation. The Bolzano variant is characterized by giant platelets with variable recognition of platelet surface GPIbα by specific monoclonal antibodies, suggesting conformational changes in the protein. Platelets from a patient with this variant were able to bind thrombin but not vWF. The Trp498→stop mutation results in a truncated GPIbα polypeptide, though this variant contained a portion of the transmembrane domain as well as the juxtamembrane cysteines that form disulfide bonds with GPIbβ, demonstrating that GPIbα lacking the cytoplasmic tail may be expressed on the platelet surface if elements of the transmembrane domain are present.

The homozygous nonsense mutation Ser444→stop resulted in the variant BSS Kagoshima. Platelets from this patient lacked GPIbα on the surface, though a truncated GPIbα polypeptide was present within the platelets and circulating in plasma. The mutant GPIbα gene encoded a truncated polypeptide that lacked a transmembrane and cytoplasmic domain, but contained normal amounts of sialic acid.

GPIbβ. The GPIbβ gene is more complex than those encoding GPIbα or GPIX, and an alternative form of GPIbβ is expressed in endothelial cells. The megakaryocyte and endothelial cell mRNA species arise from different transcriptional start sites, and use different promoters.

The first patient described with a mutation in the GPIbβ gene also suffered from a developmental disorder known as the DiGeorge/velocardiofacial syndrome. This patient had a deletion in chromosome 22q11.2, a region containing the GPIbβ and TXB1 genes, the latter encoding a T-box family transcription factor. The second GPIbβ allele contained a point mutation (C→G) at position –133, a binding site for the GATA-1 transcription factor in the GPIbβ promoter. This patient had decreased GPIbβ mRNA, and his platelets lacked GPIbβ and had markedly reduced levels of GPIbα. A similar case has been described in a patient with a heterozygous 22q11.2 deletion accompanied by a single nucleotide deletion in the codon for Ala80, leading to a frameshift mutation and premature stop codon 15 amino acids short of the C-terminus. Other mutations in the GPIbβ gene have been described, including a G→A mutation at position 159, resulting in premature termination at amino acid 21.

GPIX. Like patients with BSS secondary to GPIbα or GPIbβ mutations, patients with BSS attributable to mutations in GPIX are usually homozygotes or compound heterozygotes. The first patient described with BSS attributable to a GPIX mutation was a compound heterozygote (A→G mutation in codon 21 leading to an Asp→Gly substitution, accompanied by an A→G substitution in codon 45 of the other allele leading to Asn→Ser). These mutations perturbed the single leucine-rich region in GPIX, and the mutant protein failed to form a normal complex in CHO cells when expressed with GPIbα and GPIbβ. A homozygous Asn45→Ser mutation was responsible for BSS in an additional patient, again apparently through alteration of the conformation of the GPIX leucine-rich domain. Another homozygous mutation in GPIX associated with BSS involved a Cys→Arg substitution at position 8, disrupting a putative disulfide bond between Cys8 and Cys12. A homozygous Phe55→Ser mutation also localized to the leucine-rich motif of GPIX, leading to impaired surface expression of GPIb–IX–V. Four BSS patients in two unrelated families with point mutations leading to Cys73→Tyr had absent GPIX and reduced surface expression of GPIbα. A homozygous Cys97→Tyr substitution has also been associated with BSS, absent GPIX and reduced expression of GPIbα and GPIbβ on the platelet surface.

A recent report characterized a novel mutation in the leader peptide of GPIX leading to BSS. This resulted from a missense mutation, leading to a Leu7→Pro mutation in GPIX in an α-helical hydrophobic core region of the leader peptide that caused incorrect insertion of GPIX into the endoplasmic reticulum and/or defective signal peptide cleavage. Immunoblotting revealed GPIbα, GPIbβ and GPV, but not GPIX, in platelets.

GPV. No mutations in the GPV gene responsible for the BSS have been described. While homozygous deletion of GPIbα in mice results in a phenotype similar to that of BSS, platelets from mice that lack GPV have normal size and function, and these animals may actually be more prone to thrombosis than to bleeding. Electron microscopy of bone marrow from GPV null mice reveals normal megakaryocyte ultrastructure, appropriate development of the megakaryocyte demarcation membrane system and normal expression of GPIb-IX on the cell surface.

Platelet type or pseudo-von Willebrand disease

A disorder related to BSS in which key mutations in GPIbα play a central role, is platelet-type von Willebrand disease. This is a rare, autosomal dominant disorder characterized by increased binding of vWF, particularly the larger multimers, to platelets, and increased agglutination of platelets in response to low doses of ristocetin. *In vivo*, vWF binding causes platelet agglutination leading to accelerated removal of agglutinated platelets in the spleen, thrombocytopenia, and reduction in high molecular weight vWF multimers in plasma. The molecular defects that have been identified in this disorder are Gly233→Val and Met239→Val mutations, both of which occur in the disulfide-bonded double-loop area between the leucine-rich and macroglycopeptide domains of GPIbα.

Disorders of platelet aggregation

Glanzmann thrombasthenia

Glanzmann thrombasthenia (GT) is a rare, congenital bleeding disorder caused by deficient expression or function of platelet GPIIb–IIIa (integrin $\alpha IIb\beta_3$). As in BSS, patients with GT present with platelet-type bleeding manifestations, including purpura, epistaxis, gingival bleeding and gastrointestinal hemorrhage. GPIIb–IIIa is the most abundant platelet surface glycoprotein and plays a key role in platelet spreading at sites of injury, and in platelet aggregation. On unactivated platelets, GPIIb–IIIa maintains a conformation in which spontaneous ligand binding does not occur; however, during the process of platelet activation the receptor becomes activated following phosphorylation of critical amino acid residues within its cytoplasmic tail (inside-out signaling), and adopts a conformation capable of binding ligands such as fibrinogen, thrombospondin and vWF. Binding of such ligands by GPIIb–IIIa molecules on adjacent platelets promotes platelet bridging, which leads to platelet aggregation. GPIIb–IIIa also mediates 'outside-in' signaling following ligand binding.

Structurally, GPIIb–IIIa is a Ca^{2+}-dependent heterodimer formed by non-covalent association of two transmembrane proteins—integrin αIIb (GPIIb) and β_3 (GPIIIa) (Figure 18.4). The N-terminal region of GPIIb contains seven homologous

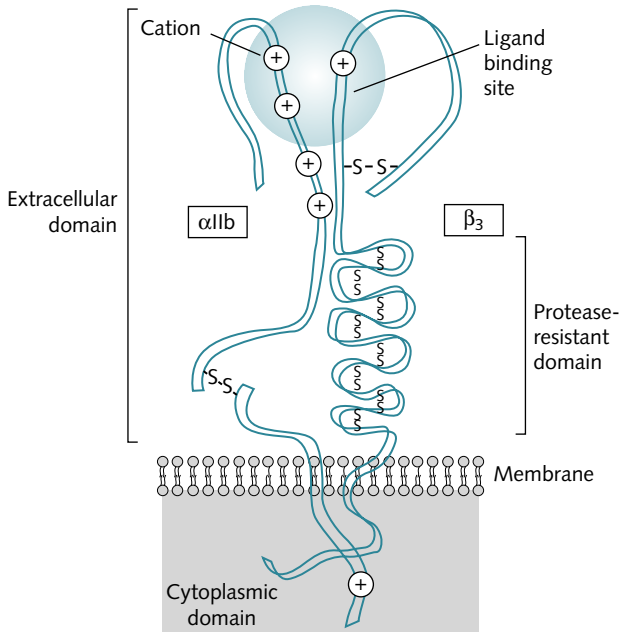


Fig. 18.4 Schematic model of $\alpha IIb\beta_3$
From Plow EF, Shattil SJ. (2000) Integrin αIIbβ3 and platelet aggregation. In: Colman RW, Hirsh J, Marder VJ *et al*. (eds). *Hemostasis and Thrombosis*, 4th edn. Philadelphia: Lippincott, Williams and Wilkins.

60 amino acid repeats in which phenylalanine-glycine (FG) and glycine-alanine-proline (GAP) consensus sequences reside. Ca^{2+}- and ligand-binding sites are contained within these regions, which form a seven-blade β propeller structure. Most of the GPIIb mutations associated with GT reside within or adjacent to these Ca^{2+}- and ligand-binding regions.

GPIIb–IIIa also plays an important role in mediating the uptake of fibrinogen by megakaryocytes for packaging into platelet α-granules. Platelets from patients who lack GPIb–IIIa do not take up fibrinogen. However, patients with diminished but detectable levels of GPIIb–IIIa contain normal amounts of fibrinogen, suggesting that reduced expression of the receptor is sufficient for normal fibrinogen uptake. Moreover, certain GT variants with molecular defects that impair inside-out signal transduction appear to mediate fibrinogen uptake normally, suggesting that fibrinogen uptake may not be an activation dependent process.

Genetics of the GPIIb–IIIa complex

The genes encoding integrins αIIb and β_3 are larger and more complex than those encoding members of the GPIb–IX–V complex, and unlike BSS, some cases of GT may result from splicing errors. The αIIb gene contains 30 exons and spans 20 kb, while that for β_3 contains 15 exons and spans 46 kb. These genes have been localized to 17q21–23. Homozygous or compound heterozygous molecular defects in these genes or their regulatory elements may result in GT, while single heterozygotes are usually asymptomatic. Like BSS, a mutation may lead to a synthetic defect resulting in absence of expression, or to a functional defect that results in failure of GPIIb–IIIa to bind fibrinogen or other ligands. GT mutations have been classified as either type I (less than 5% surface expression of $\alpha IIb\beta_3$) or type II (more than 5% surface expression of $\alpha IIb\beta_3$), though progress in molecular characterization of this disorder now allows patients with GT to be described by the nature of their mutation.

Mutations associated with Glanzmann thrombasthenia

Some of the more common mutations in the αIIb and β_3 genes that may cause GT are described below. Again, this discussion is not comprehensive; an updated listing of mutations associated with GT is maintained at http://med.mssm.edu/glanzmanndb.

αIIb. Mutations that affect the αIIb gene often cluster around one of the Ca^{2+}-binding phenylalanine-glycine or glycine-alanine-proline regions in the N-terminus of the protein, which are important in ligand binding and interactions with β_3 (Figure 18.5). An example is provided by one patient who was compound heterozygous for a Glu324→Lys mutation in one allele of αIIb and an Ile565→Thr mutation in the other.

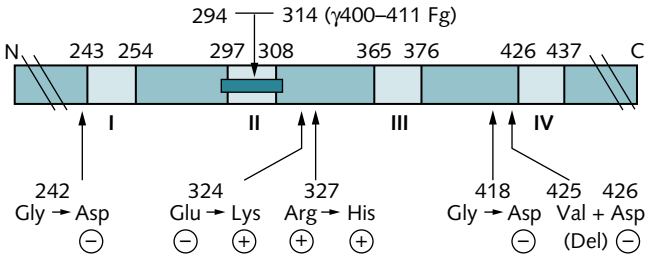


Fig. 18.5 Mutations detected within or close to Ca^{2+}-binding domains of GPIIb

The extracellular domain of αIIb has four Ca^{2+}-binding domains (labeled I–IV) that assure the conformation of the subunit both for its association with β_3 (GPIIIa) and for fibrinogen binding. That these regions are of importance in β_3 association is demonstrated by the presence of numerous missense mutations in this region as well as a deletion of two amino acids, all of which result in charge changes and an altered conformation with a severely reduced surface expression of the mature integrin. The blue bar indicates the binding site in αIIb for fibrinogen. From Nurden AT, Nurden P. (2001) Inherited defects of platelet function. *Reviews in Clinical and Experimental Hematology*, **54**, 314–334, with permission.

Other mutations that may affect αIIb expression involve the transmembrane domain, with such a mutation occurring in a French gypsy population in whom αIIb was undetectable on platelets. Analysis of the αIIb gene in these individuals revealed an 8 bp deletion at the 3′ end of exon 15 that resulted from abnormal splicing, causing a frameshift and premature truncation. In another case, a Glu324→Lys substitution resulting from a single base pair substitution in exon 12 of the αIIb gene resulted in less than 5% surface expression of GPIIb–IIIa; this mutation has been found in patients on at least three continents, suggesting that the codon for Glu324 may be a mutational hotspot for GT.

β_3. As with those in αIIb, mutations in the β_3 gene may impair expression or function of the protein. A common cause is a point mutation in the gene, resulting in a loss of cysteine residues and disruption of normal disulfide bond formation. Disulfides play a critical role in maintaining the conformation of β_3 and in its ability to complex with αIIb. Examples include a Cys374→Tyr mutation in one patient, which diminished the expression of β_3 on the platelet surface by ~90% and impaired the stability of $\alpha IIb\beta_3$. Another homozygous Cys542→Arg mutation led to the almost complete absence of $\alpha IIb\beta_3$ expression in platelets.

In addition to point mutations, a number of deletions, inversions and splice site mutations may induce premature termination of β_3 and lead to GT. A classic example reported in the Iraqi Jewish population is an 11 bp mutation in exon 13 leading to a frameshift, and protein truncation prior to the transmembrane domain.

Variant Glanzmann thrombasthenia

A number of patients with variant GT have been described, most of whom have mutations in the β_3 gene. The level of GPIIb–IIIa expression on the platelet surface in these patients is variably reduced, and functional abnormalities of the receptor are characteristic. The clinical phenotype is heterogeneous, and bleeding may range from severe to mild. Mutations in several patients with variant GT have been insightful in elucidating the structure–function relationships of $\alpha IIb\beta_3$. An Asp119→Tyr mutation was the first such mutation characterized, and identification of this mutation in subsequent patients led to the finding that Asp119 and Arg214 constitute a portion of a functionally important metal binding site within the β_3 polypeptide. Another mutation (Ser752→Pro) led to reduced surface expression of GPIIb–IIIa, with failure of platelets to become activated and bind fibrinogen after stimulation by ADP, though isolated GPIIb–IIIa and protease treated platelets bound fibrinogen normally. Further studies revealed that this mutation was associated with impaired inside-out signaling. Similar results were reported in platelets from a patient with a heterozygous C2268T mutation that led to a stop codon at Arg724, and truncation of the cytoplasmic domain after eight amino acids. These platelets did not bind fibrinogen after normal stimulation, but did so after exposure to an activating anti-GPIIb–IIIa antibody.

Platelet polymorphisms and thrombotic disease

An area that has received increasing attention in recent years has been the role of platelet glycoprotein polymorphisms in the predisposition to thrombotic disease. A number of polymorphisms that occur in platelet glycoproteins have been examined, though in most cases their relationship to thrombosis remains uncertain. This may reflect the fact that the pathogenesis of thrombotic disease is complex, diverse and dependent upon a number of other genetic and acquired factors in addition to platelet function. Given these considerations, the demonstration of a meaningful role for a platelet polymorphism in the development of thrombosis will require large numbers of patients who experience thrombotic events. Though a full discussion of this area is beyond the scope of this chapter, we will review some of the existing data on the more common polymorphisms.

GPIIb–IIIa

Given the importance of GPIIb–IIIa in platelet adhesion, spreading and aggregation, it is not surprising that common polymorphisms in these glycoproteins have been closely

examined for their role in thrombosis. Each of the subunits of GPIIb–IIIa contains a common amino acid dimorphism. GPIIIa consists of two common isoforms, PL^{A1} and PL^{A2}, which are due to a T→C nucleotide substitution involving nucleotide 1565 and resulting in a Leu33→Pro33 amino acid change (Pro in PL^{A2}). This substitution leads to conformational changes within GPIIb–IIIa, though whether these translate into significant functional differences remains an area of active investigation. The net result of multiple studies suggests that such changes are limited to a mild effect on platelet function involving signaling events that occur following ligand binding that potentially contributes to a mild prothrombotic predisposition among patients with the Pro33 allele. However, the clinical implications of these *in vitro* studies remain controversial, though some data suggest that the Pro33 allele may be a weak risk factor for myocardial infarction in young men.

Another polymorphism that has received considerable attention has been the Ile843Ser polymorphism in GPIIb. *In vitro*, most studies have not demonstrated an effect of this polymorphism on platelet function. As with the Leu33Pro polymorphism, clinical studies concerning the significance of this polymorphism have yielded controversial results. One study suggested that the Ser variant was associated with myocardial infarction in a cohort of 68 North American women less than 45 years of age with other cardiovascular risk factors. The same study suggested a possible association of the Ser843 polymorphism with stroke in this cohort.

GPIa/IIa

The importance of several polymorphisms within GPIa have been assessed. That which has received the most attention has been the C807T polymorphism, which involves the codon for Phe224 but does not result in an amino acid change. This polymorphism is in complete linkage disequilibrium with other polymorphisms, including G873A (a silent polymorphism at codon Thr246) and polymorphisms involving intronic and promoter regions within the GPIa gene. Platelets from patients with the 807T polymorphism have increased platelet glycoprotein Ia/IIa receptor density and increased collagen-dependent platelet adhesion under high shear stress conditions, though it is uncertain which of the associated polymorphisms is responsible for these differences. Clinical studies concerning the relevance of these polymorphisms have yielded conflicting results, though the cumulative evidence suggests that the 807T polymorphism is associated with increased susceptibility to arterial thrombotic disease in younger patients and microvascular disease among diabetics.

Another polymorphism of interest is the Glu505Lys polymorphism, which results from a G1648A change in the GPIa gene. This polymorphism is responsible for the platelet HPA-5 antigen system. It does not lead, however, to altered platelet adhesion to collagen under static conditions, and its clinical implications are uncertain. Most clinical studies examining this polymorphism have failed to demonstrate differences in the incidence of thrombotic events, and at this point it cannot be considered to be an important genotypic marker of thrombotic risk. A similar conclusion can be reached concerning the significance of the platelet GPIa-52C/T polymorphism, located in the 5′ regulatory region of the gene and involved in Sp1 and Sp3 transcription factor binding. The -52T variant is associated with decreased GPIa/IIa receptor density on the platelet surface.

GPIb–IX–V

The clinical relevance of several polymorphisms within GPIbα has been assessed. These include a Thr145Met polymorphism responsible for the platelet HPA-2 antigen system, a variable number of tandem repeats (VNTR) of a 13 amino acid sequence in the GPIbα macroglycopeptide region, and the GPIbα–5T/C polymorphism. The VNTR and Thr145Met polymorphism is in strong linkage disequilibrium in Asian and European populations, where the Thr145 allele is associated strongly with the VNTR-C and VNTR-D variants (which are smaller, containing only one or two of the 13 amino acid repeats), while the Met145 allele is associated with the larger VNTR-A and VNTR-B (four and three repeats, respectively) variants. The –5T/C polymorphism alters a consensus Kozac translation initiation sequence, and the –5T variant may decrease translational efficiency of the gene. However, the –5C allele is in strong linkage disequilibrium with the Thr145/VNTR-C haplotype.

Compared with polymorphisms affecting other platelet glycoproteins, there is less information available concerning the functional significance of these polymorphisms. Although the Thr145Met polymorphism induces conformational changes within GPIbα, there is no firm evidence that these changes alter ristocetin-induced platelet aggregation or the binding of vWF, nor are they associated with any changes in receptor density on the platelet surface.

Similar comments pertain concerning the clinical relevance of these changes. Case–control studies have suggested a two-fold increase in the risk of coronary heart disease and that of cerebrovascular disease in patients with the Met145 allele, with the greatest risk among women. However, in several of these studies the wide confidence intervals confound the conclusions. Studies concerning the significance of the VNTR polymorphisms have likewise generated conflicting results, as have those focussed on evaluation of the –5C/T polymorphism. The linkage disequilibrium between these genes makes their population-based study particularly difficult.

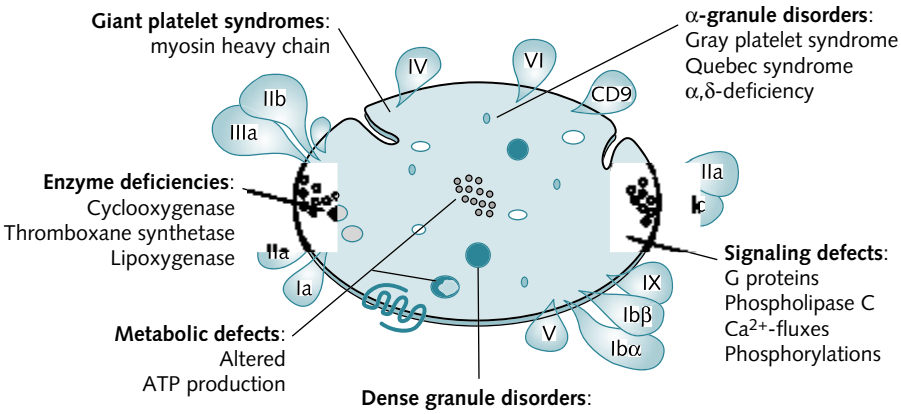


Fig. 18.6 Major inherited platelet function disorders that involve intracellular proteins and storage granules. From Nurden AT, Nurden P. (2001) Inherited defects of platelet function. *Reviews in Clinical and Experimental Hematology*, **54**, 314–334, with permission.

Summary

Platelet surface glycoproteins play critical roles in numerous aspects of platelet function. Mutations in these glycoproteins may have significant clinical importance, and may lead to marked deficiencies in primary hemostasis. Characterization of these mutations has had a profound impact on our understanding of platelet function, and the structure–activity relationships of the glycoproteins themselves. Likewise, additional platelet polymorphisms, though not associated with gross abnormalities in platelet function, have been implicated as a risk factor for thrombosis-related disorders in population-based studies.

Emerging data suggest that, like the function of platelet glycoproteins, the activity of platelet intracellular signaling proteins and the proteins that regulate platelet granule development and secretion may also be genetically regulated (Figure 18.6). Although this area remains comparatively unexplored, molecular data concerning the pathogenesis of prototypical disorders such as the Chédiak–Higashi syndrome and Hermansky–Pudlak syndrome have already become available. Though a description of these and related disorders is beyond the scope of this chapter, such information is likely to become increasingly relevant as additional data accumulate.

Further reading

Overview of platelet function

Berndt MC, Shen Y, Dopheide SM, Gardiner EE, Andrews RK. (2001) The vascular biology of the glycoprotein Ib–IX–V complex. *Thrombostasis and Haemostasis*, **86**, 178–188.

Calvete JJ. (1999) Platelet integrin GPIIb/IIIa: structure–function correlations. An update and lessons from other integrins. *Proceedings of the Society for Experimental Biology and Medicine*, **222**, 29–38.

Clemetson KJ, Clemetson JM. (2001) Platelet collagen receptors. *Thrombostasis and Haemostasis*, **86**, 189–197.

Kasirer-Friede A, Legrand C, Frojmovic MM. (2001) Complementary roles for fibrinogen, thrombospondin and vWF in mediating shear-dependent aggregation of platelets stimulated at threshold thrombin concentrations. *Thrombostasis and Haemostasis*, **86**, 653–659.

Kim S, Foster C, Lecchi A *et al.* (2002) Protease-activated receptors 1 and 4 do not stimulate G(i) signaling pathways in the absence of secreted ADP and cause human platelet aggregation independently of G(i) signaling. *Blood*, **99**, 3629–3636.

Kulkarni S, Dopheide SM, Yap CL *et al.* (2000) A revised model of platelet aggregation. *Journal of Clinical Investigation*, **105**, 783–791.

Phillips DR, Charo IF, Scarborough R. (1991) GPIIb–IIIa: the responsive integrin. *Cell*, **65**, 359–362.

Ruggeri ZM. (2000) Old concepts and new developments in the study of platelet aggregation. *Journal of Clinical Investigation*, **105**, 699–701.

Savage B, Cattaneo M, Ruggeri Z. (2001) Mechanisms of platelet aggregation. *Current Opinion in Hematology*, **8**, 270–276.

Sims PJ, Ginsberg MH, Plow EF. (1991) Effect of platelet activation on the conformation of the plasma membrane glycoprotein IIb/IIIa complex. *Journal of Biological Chemistry*, **266**, 7345–7352.

Soslau G, Class R, Morgan DA *et al.* (2001) Unique pathway of thrombin-induced platelet aggregation mediated by glycoprotein Ib. *Journal of Biological Chemistry*, **276**, 21173–21183.

Laboratory assessment of platelet disorders

Michelson AD. (2000) Flow cytometric analysis of platelets. *Vox Sanguinis*, **78**, 137–142.

Overview of inherited platelet dysfunction

Clemetson KJ. (2001) Platelet glycoproteins and their role in diseases. *Transfusion Clinique et Biologique. Journal de la Société Française de Transfusion Sanguine*, **8**, 155–162.

Nurden AT, Nurden P. (2001) Inherited defects of platelet function. *Reviews in Clinical and Experimental Hematology*, **54**, 314–334.

Bernard–Soulier syndrome

Overview

Hayashi T, Suzuki K. (2000) Molecular pathogenesis of the Bernard–Soulier syndrome. *Seminars in Thrombostasis and Haemostasis*, **26**, 53–59.

Lopez JA, Andrews RK, Afshar-Kharghan V *et al.* (1998) Bernard–Soulier syndrome. *Blood*, **91**, 4397–4418.

Lopez JA, Chung, DW, Fujikawa K *et al.* (1988) The alpha and beta chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 2135–2139.

Ruggeri ZM. (1991) The platelet glycoprotein Ib-IX complex. *Progress in Hemostasis & Thrombosis*, **10**, 35–68.

GPIbα gene mutations

De La Salle C, Baas M-J, Lanza F *et al.* (1995) A three base deletion removing a leucine residue in a leucine-rich repeat of platelet glycoprotein Ibα associated with a variant of Bernard–Soulier syndrome (Nancy I). *British Journal of Haematology*, **89**, 386–392.

Holmberg L, Karpman D, Nilsson I *et al.* (1997) Bernard–Soulier syndrome Karlstad: Trp 498→Stop mutation resulting in a truncated glycoprotein Ib alpha that contains part of the transmembranous domain. *British Journal of Haematology*, **98**, 57–63.

Kanaji T, Okamura T, Kuroiwa M *et al.* (1997) Molecular and genetic analysis of two patients with Bernard–Soulier syndrome: identification of new mutations in glycoprotein Ibα gene. *Thrombosis and Haemostasis*, **77**, 1055–1061.

Li C, Martin E, Roth GJ. (1995) The genetic defect in two well-studied cases of Bernard–Soulier syndrome: a point mutation in the fifth leucine-rich repeat of platelet glycoprotein Ib alpha. *Blood*, **86**, 3805–3814.

Li C, Pasquale DN, Roth GJ. (1996) Bernard–Soulier syndrome with severe bleeding: absent platelet glycoprotein Ib alpha due to homozygous one-base deletion. *Thrombosis and Haemostasis*, **76**, 670–674.

Lopez JA, Chung DW, Fujikawa *et al.* (1987) Cloning of the alpha chain of human platelet glycoprotein Ib: a transmembrane protein with homology to leucine-rich alpha 2-glycoprotein. *Proceedings of the National Academy of Sciences of the United States of America*, **16**, 615–619.

Miller JL, Lyle VA, Cunningham D. (1992) Mutation of leucine-57 to phenylalanine in a platelet glycoprotein Ibα leucine tandem repeat occurring in patients with an autosomal dominant variant of Bernard–Soulier disease. *Blood*, **79**, 439.

Noda M, Fujimura K, Takafuta T *et al.* (1995) Heterogeneous expression of glycoprotein Ib, IX, and V in platelets from two patients with Bernard–Soulier syndrome caused by different genetic abnormalities. *Thrombostasis and Haemostasis*, **74**, 1411.

Simsek S, Admiraal LG, Modderman PW *et al.* (1994) Identification of a homozygous single base pair deletion in the gene coding for human platelet glycoprotein Ibα causing Bernard–Soulier syndrome. *Thrombosis and Haemostasis*, **72**, 444–449.

Simsek S, Noris P, Lozano M *et al.* (1994) Cys209Ser mutation in the platelet membrane glycoprotein Ibα gene is associated with Bernard–Soulier syndrome. *British Journal of Haematology*, **88**, 839.

Ware J, Russell SR, Vicenete V *et al.* (1990) Nonsense mutation in the glycoprotein Ibα coding sequence associated with Bernard–Soulier syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 2026–2030.

Whisstock JC, Shen Y, Lopez JA *et al.* (2002) Molecular modeling of the seven tandem leucine-rich repeats within the ligand-binding region of platelet glycoprotein Ib alpha. *Thrombosis and Haemostasis*, **87**, 329–333.

Heterozygous forms

Kanaji T, Okamura T, Kuroiwa M *et al.* (1997) Molecular and genetic analysis of two patients with Bernard–Soulier syndrome. Identification of new mutations in the glycoprotein Ibα gene. *Thrombosis and Haemostasis*, **77**, 1055–1061.

Ware J, Russell SR, Vicenete V *et al.* (1990) Nonsense mutation in the glycoprotein Ibα coding sequence associated with Bernard–Soulier syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 2026.

Variant forms

Kunishima S, Miura H, Fukutani H *et al.* (1994) Bernard–Soulier syndrome Kagoshima: Ser 444→stop mutation of glycoprotein (GP) Ib alpha resulting in circulating truncated GPIB alpha and surface expression of GPIb beta and GPIX. *Blood*, **84**, 3356–3362.

Miller JL, Lyle VA, Cunningham D. (1992) Mutation of leucine-57 to phenylalanine in a platelet glycoprotein Ibα leucine tandem repeat occurring in patients with an autosomal dominant variant of Bernard–Soulier disease. *Blood*, **79**, 439–446.

GPIbβ gene mutations

Moran N, Morateck PA, Deering A *et al.* (2000) The surface expression of glycoprotein Ib alpha is dependent on GPIb beta: evidence from a novel mutation causing Bernard–Soulier syndrome. *Blood*, **96**, 532–539.

Kunishima S, Lopez JA, Kobayashi S *et al.* (1997) Missense mutations of the glycoprotein (GP) Ib beta gene impairing the GPIb alpha/beta

disulfide linkage in a family with giant platelet disorder. *Blood*, **89**, 2404–2412.

Kunishima S, Naoe T, Kamiya T, Saito H. (2001) Novel heterozygous missense mutation in the platelet glycoprotein Ib beta gene associated with isolated giant platelet disorder. *American Journal of Hematology*, **68**, 249–255.

Kurokawa Y, Ishida F, Kamijo T *et al.* (2001) Missense mutation (Tyr88 to Cys) in the platelet membrane glycoprotein Ib beta gene affects GPIb/IX complex expression. *Thrombosis and Haemostasis*, **86**, 1249–1256.

Deletions of 22q11.2

Kenny D, Morateck, PA, Gill, JC *et al.* (1999) The critical interaction of glycoprotein (GP) Ib beta with GPIX: a genetic cause of Bernard–Soulier syndrome. *Blood*, **93**, 2968–2975.

Velocardiofacial syndrome

Ludlow LB, Schick BP, Budarf ML *et al.* (1996) Identification of a mutation in a GATA binding site of the platelet glycoprotein Ibβ promoter resulting in the Bernard–Soulier syndrome. *Journal of Biological Chemistry*, **271**, 22076–22080.

GPIX

Clemetson JM, Kyrle PA, Brenner B *et al.* (1994) Variant Bernard–Soulier syndrome associated with a homozygous mutation in the leucine-rich domain of glycoprotein IX. *Blood*, **84**, 1124–1131.

Kunishima S, Tomiyama Y, Honda S *et al.* (1999) Cys97-Tyr mutation in the glycoprotein IX gene associated with Bernard–Soulier syndrome. *British Journal of Haematology*, **107**, 539–545.

Lanza F, De La Salle C, Bass MJ *et al.* (2001) A Leu7Pro mutation in the signal peptide of platelet glycoprotein (GP)IX in a case of Bernard–Soulier syndrome abolishes surface expression of the GPIb-V-IX complex. *British Journal of Haematology*, **118**, 260–266.

Noda M, Fujimura K, Takafuta T *et al.* (1996) A point mutation in glycoprotein IX coding sequence (Cys73{TGT} to Tyr{TAT}) causes impaired surface expression of GPIb/IX/V complex in two families with Bernard–Soulier syndrome. *Thrombostasis and Haemostasis*, **76**, 874–878.

Rivera CE, Villagra J, Riordan M *et al.* (2001) Identification of a new mutation in platelet glycoprotein IX (GPIX) in a patient with Bernard–Soulier syndrome. *British Journal of Haematology*, **112**, 105–108.

Suzuki K, Hayashi T, Yahagi A *et al.* (1997) Novel point mutation in the leucine-rich motif of the platelet glycoprotein IX associated with Bernard–Soulier syndrome. *British Journal of Haematology*, **99**, 794–800.

Wright SD, Michaelides K, Johnson DJD *et al.* (1993) Double heterozygosity for mutations in the platelet glycoprotein IX gene in three siblings with Bernard–Soulier syndrome. *Blood*, **81**, 2339–2347.

GPV

Kahn ML, Diacovo TG, Bainton DF *et al.* (1997) Glycoprotein V-deficient platelets have undiminished thrombin responsiveness and do not exhibit a Bernard–Soulier phenomenon. *Blood*, **94**, 4112–4121.

Lanza F, Morales M, del La Salle C *et al.* (1993) Cloning and characterization of the gene encoding the human platelet glycoprotein V. A member of the leucine-rich glycoprotein family cleaved during thrombin-induced platelet activation. *Journal of Biological Chemistry*, **268**, 20801–20807.

Michelson AD, Benoitt SE, Furman ME *et al.* (1996) The platelet surface expression of glycoprotein V is regulated by two independent mechanisms: proteolysis and the reversible cytoskeletal mediated redistribution to the surface-connected canalicular system. *Blood*, **87**, 1396–1408.

Ni H, Ramakrishnan V, Papalia JM *et al.* (2000) Increased thrombogenesis and embolus formation in mice lacking glycoprotein V. *Blood*, **98**, 368–373.

Poujol C, Ramakrishnan V, DeGuzman F *et al.* (2000) Ultrastructural analysis of megakaryocytes in GPV knockout mice. *Thrombosis and Haemostasis*, **84**, 312–318.

Ramakrisnana V, Reeves PS, DeGuzman F *et al.* (1999) Increased thrombin responsiveness in platelets from mice lacking glycoprotein V. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 13336–13341.

Ware J, Russell S, Ruggeri ZM. (2000) Generation and rescue of a murine model of platelet dysfunction: the Bernard–Soulier syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 2803–2808.

Platelet type or pseudo-von Willebrand disease

Miller JL. (1996) Platelet-type von Willebrand disease. *Thrombosis and Haemostasis*, **75**, 865–869.

Miller JL, Cunningham D, Lyle VA *et al.* (1991) Mutation in the gene encoding the α-chain of platelet glycoprotein Ib in the platelet type von Willebrand disease. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 4761–4765.

Takahashi H, Murata M, Moriki T *et al.* (1995) Substitution of Val for Met at residue 239 of platelet glycoprotein Ibα in Japanese individuals with platelet-type von Willebrand disease. *Blood*, **85**, 727–733.

Glanzmann thrombasthenia

Overview

Bennett JS, Vilaire G. (1979) Exposure of platelet fibrinogen receptors by ADP and epinephrine. *Journal of Clinical Investigation*, **64**, 1393–1401.

French DL, Coller BS. (1997) Hematologically important mutations: Glanzmann thrombasthenia. *Blood Cells, Molecules and Diseases*, **23**, 39–51.

French DL, Seligsohn U. (2000) Platelet glycoprotein IIb/IIIa receptors and Glanzmann's thrombasthenia. *Arteriosclerosis, Thrombosis, and Vascular Biology*, **20**, 607–610.

Kato A. (1997) The biologic and clinical spectrum of Glanzmann's thrombasthenia: implications of integrin alpha IIb beta 3 for its pathogenesis. *Critical Reviews in Oncology/Hematology*, **26**, 1–23.

GPIIb gene

Ambo H, Kamata T, Handa M *et al.* (1998) Novel point mutations in the alphaIIb subunit (Phe289Ser, Glu324Lys and Gln747Pro) causing thrombasthenic phenotypes in four Japanese patients. *British Journal of Haematology*, **102**, 829–840.

Arias-Salgade EG, Tao J, Gonzalez-Manchon C *et al.* (2002) Nonsense mutation in exon-19 of alphaIIb associated with thrombasthenic phenotype. Failure of GPIIb(del597–1008)to form stable complexes with GPIIIa. *Thrombosis and Haemostasis*, **87**, 684–691.

Basani RB, Vilaire G, Shattil SJ *et al.* (1996) Glanzmann thrombasthenia due to a two amino acid deletion in the fourth calcium-binding domain of alpha IIb: demonstration of the importance of calcium-binding domains in the conformation of alpha IIb. *Blood*, **88**, 167–173.

Basani RB, French DL, Vilaire G *et al.* (2000) A naturally occurring mutation near the amino terminus of alpha IIb defines a new region involved in ligand binding to alpha IIb beta 3. *Blood*, **95**, 180–188.

Burk CD, Newman PJ, Lyman S *et al.* (1991) A deletion in the gene for glycoprotein IIb associated with Glanzmann's thrombasthenia. *Journal of Clinical Investigation*, **87**, 270–276.

Ferrer M, Fernandez-Pinel M, Gonzalez-Manchon C *et al.* (1996) A mutant (Arg327→His) GPIIb associated to thrombasthenia exerts a dominant negative effect in stably transfected CHO cells. *Thrombosis and Haemostasis*, **76**, 292–301.

French DL, Coller BS, Usher S *et al.* (1998) Prenatal diagnosis of Glanzmann thrombasthenia using the polymorphic markers BRCA1 and THRA1 on chromosome 17. *British Journal of Haematology*, **102**, 582–587.

Grimaldi CM, Chen FP, Wu CH *et al.* (1998) Glycoprotein IIb Leu214Pro mutation produces Glanzmann thrombasthenia with both quantitative and qualitative abnormalities in GPIIb/IIIa. *Blood*, **91**, 1562–1571.

Gonzalez-Manchon C, Fernandez-Pinel M, Arias-Salgado EG *et al.* (1999) Molecular genetic analysis of a compound heterozygote for the glycoprotein (GP) IIb gene associated with Glanzmann's thrombasthenia: disruption of the 674–687 disulfide bridge in GPIIb prevents surface exposure of GPIIb–IIIa complexes. *Blood*, **93**, 866–875.

Gu JM, Xu WF, Wang XD *et al.* (1993) Identification of a nonsense mutation at amino acid 584-arginine of platelet glycoprotein IIb in patients with type I Glanzmann thrombasthenia. *British Journal of Haematology*, **83**, 442–449.

Honda S, Tomiyama Y, Shiraga M *et al.* (1998) A two-amino acid insertion in the Cys146-Cys167 loop of the alpha IIb subunit is associated with a variant of Glanzmann thrombasthenia. *Journal of Clinical Investigation*, **102**, 1183–1192.

Iwamoto S, Nishiumi E, Kajii E *et al.* (1994) An exon 28 mutation resulting in alternative splicing of the glycoprotein IIb transcript and Glanzmann's thrombasthenia. *Blood*, **83**, 1017–1023.

Kato A, Yamamoto K, Miyazaki S, Jung SM, Moroi M, Aoki N. (1992) Molecular basis for Glanzmann's thrombasthenia (GT) in a compound heterozygote with glycoprotein IIb gene: a proposal for the classification of GT based on the biosynthetic pathway of glycoprotein IIb-IIIa complex. *Blood*, **79**, 3212–3218.

Kirchmaier CM, Westrup D, Becker-Hagendorff K *et al.* (1995) A new variant of Glanzmann's thrombasthenia (Frankfurt I). *Thrombosis and Haemostasis*, **73**, 1058–1064.

Milet-Marsal S, Breillat C, Peyruchaud O *et al.* (2002) Analysis of the amino acid requirement for a normal alphaIIbbeta3 maturation at alphaIIbGlu324 commonly mutated in Glanzmann thrombasthenia. *Thrombosis and Haemostasis*, **88**, 655–662.

Mitchell WB, Li J, Singh F, Michelson AD *et al.* (2003) Two novel mutations in the alphaIIb calcium-binding domains identify hydrophobic regions essential for alphaIIbbeta3 biogenesis. *Blood*, **101**, 2268–2276.

Newman PJ, Seligsohn U, Lyman S *et al.* (1991) The molecular genetic basis of Glanzmann thrombasthenia in the Iraqi-Jewish and Arab populations in Israel. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 3160–3164.

Peretz H, Rosenberg N, Usher S *et al.* (1995) Glanzmann's thrombasthenia associated with deletion–insertion and alternative splicing in the glycoprotein IIb gene. *Blood*, **85**, 414–420.

Peyruchaud O, Nurden AT, Milet S. (1998) R to Q amino acid substitution in the GFFKR sequence of the cytoplasmic domain of the integrin alpha IIb subunit in a patient with a Glanzmann's thrombasthenia-like syndrome. *Blood*, **92**, 4178–4187.

Poncz M, Rifat S, Coller BS, Newman PJ *et al.* (1994) Glanzmann thrombasthenia secondary to a Gly273Asp mutation adjacent to the first calcium-binding domain of platelet glycoprotein IIb. *Journal of Clinical Investigation*, **93**, 172–179.

Ruan CG, Gu JM, Wang XD, Chu XH, Pan JL. (1993) Application of GPIIIa gene Taq I polymorphism to determination of carrier status

in Glanzmann's thrombasthenia families of Chinese origin. *Thrombosis and Haemostasis*, **68**, 64–69.

Ruan J, Peyruchaud O, Alberio L *et al.* (1998) Double heterozygosity of the GPIIb gene in a Swiss patient with Glanzmann's thrombasthenia. *British Journal of Haematology*, **102**, 918–925.

Schlegel N, Gayet O, Morel-Kopp MC *et al.* (1995) The molecular genetic basis of Glanzmann's thrombasthenia in a Gypsy population in France: identification of a new mutation on the alpha IIb gene. *Blood*, **86**, 977–982.

Scott JP III, Scott JP II, Chao Y-L *et al.* (1998) A frameshift mutation at Gly975 in the transmembrane domain of GPIIb prevents GPIIb–IIIa expression analysis of two novel mutations in a kindred with type I Glanzmann thrombasthenia. *Thrombosis and Haemostasis*, **80**, 546–550.

Tadokoro S, Tomiyama Y, Honda S *et al.* (1998) A Gln Pro substitution in the alphaIIb subunit is responsible for a moderate alphaIIbbeta3 deficiency in Glanzmann thrombasthenia. *Blood*, **92**, 2750–2758.

Tanaka S, Hayashi T, Hori Y *et al.* (2002) A Leu55 to Pro substitution in the integrin alphaIIb is responsible for a case of Glanzmann's thrombasthenia. *British Journal of Haematology*, **118**, 833–835.

Tao J, Arias-Salgado E.G., Gonzalez-Manchon C *et al.* (2000) A novel (288delC) mutation in exon 2 of GPIIb associated with type I Glanzmann's thrombasthenia. *British Journal of Haematology*, **111**, 96–103.

Tomiyama Y, Kashiwagi H, Shiraga M *et al.* (1994) Demonstration of a marked reduction in the amount of GPIIb in most type II patients with Glanzmann's thrombasthenia. *British Journal of Haematology*, **87**, 119–124.

Tomiyama Y, Kashiwagi H, Kosugi S *et al.* (1995) Abnormal processing of the glycoprotein IIb transcript due to a nonsense mutation in exon 17 associated with Glanzmann's thrombasthenia. *Thrombosis and Haemostasis*, **73**, 756–762.

Vinciguerra C, Khelif A, Alemany M *et al.* (1996) A nonsense mutation in the GPIIb heavy chain (Ser870→stop) impairs platelet GPIIb–IIIa expression. *British Journal of Haematology*, **95**, 399–407.

Wilcox DA, Wautier JL, Pidard D *et al.* (1994) A single amino acid substitution flanking the fourth calcium binding domain of alpha IIb prevents maturation of the integrin alpha IIb beta 3 complex. *Journal of Biological Chemistry*, **269**, 4450–4457.

Wilcox DA, Paddock CM, Lyman S *et al.* (1995) Glanzmann thrombasthenia resulting from a single amino acid substitution between the second and third calcium-binding domains of GPIIb. *Journal of Clinical Investigation*, **95**, 1553–1560.

Yatuv R, Rosenberg N, Dardik R *et al.* (1998) Glanzmann thrombasthenia in two Iraqi-Jewish siblings is caused by a novel splice junction mutation in the glycoprotein IIb. *Blood Coagulation and Fibrinolysis*, **9**, 285–288.

GPIIIa gene mutations

Ambo H, Kamata T, Handa M *et al.* (1998) Three novel integrin beta3 subunit missense mutations (H280P, C560F, and G579S) in thrombasthenia, including one (H280P) prevalent in Japanese patients. *Biochemical and Biophysical Research Communications*, **251**, 763–768.

Bajt ML, Ginsberg MH, Frelinger AL III *et al.* (1992) A spontaneous mutation of integrin alpha IIb beta 3 (platelet glycoprotein IIb-IIIa) helps define a ligand binding site. *Journal of Biological Chemistry*, **267**, 3789–3794.

Basani RB, Brown DL, Vilaire G *et al.* (1997) A Leu117→Trp mutation within the RGD-peptide cross-linking region of beta 3 results in Glanzmann thrombasthenia by preventing alpha IIb beta 3 export to the platelet surface. *Blood*, **90**, 3082–3088.

Bray PF, Shuman MA. (1990) Identification of an abnormal gene for the GPIIIa subunit of the platelet fibrinogen receptor resulting in Glanzmann thrombasthenia. *Blood*, **75**, 881–888.

Chen Y-P, Djaffar I, Pidard D *et al.* (1992) Ser752Pro mutation in the cytoplasmic domain of integrin beta 3 subunit and defective activation of platelet integrin alpha IIb beta 3 (glycoprotein IIb-IIIa) in a variant of Glanzmann thrombasthenia. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 10169–10173.

Chen Y-P, O'Toole TE, Ylanne J *et al.* (1994) A point mutation in the integrin beta 3 cytoplasmic domain (S752P) impairs bidirectional signaling through alpha IIb beta 3 (platelet glycoprotein IIb-IIIa). *Blood*, **84**, 1857–1865.

Coller BS, Seligsohn U, Little PA. (1987) Type I Glanzmann thrombasthenia patients from the Iraqi-Jewish and Arab populations in Israel can be differentiated by platelet glycoprotein IIIa immunoblot analysis. *Blood*, **69**, 1969–1703.

Coller BS, Cheresh DA, Asch E *et al.* (1991) Platelet vitronectin receptor expression differentiates Iraqi-Jewish from Arab patients with Glanzmann thrombasthenia in Israel. *Blood*, **77**, 75–83.

Djaffar I, Rosa J-P (1993) A second case of variant Glanzmann's thrombasthenia due to substitution of platelet GPIIIa (integrin beta 3) Arg214 by Trp. *Human Molecular Genetics*, **2**, 2179–2180.

Djaffar I, Caen JP, Rosa J-P. (1993) A large alteration in the human platelet glycoprotein IIIa (integrin beta 3) gene associated with Glanzmann's thrombasthenia. *Human Molecular Genetics*, **2**, 2183–2185.

Ferrer M, Tao J, Sanchez-Ayuso M *et al.* (1998) Truncation of glycoprotein (GP) IIIa (616–762) prevents complex formation with GPIIb: novel mutation in exon 11 of GPIIIa associated with thrombasthenia. *Blood*, **92**, 4712–4720.

Fournier DJ, Kabral A, Castaldi PA *et al.* (1989) A variant of Glanzmann's thrombasthenia characterized by abnormal glycoprotein IIb/IIIa complex formation. *Thrombosis and Haemostasis*, **62**, 977–983.

Fullard J, Murphy R, O'Neill S *et al.* (2001) A Val193Met mutation in GPIIIa results in a GPIIb/IIIa receptor with a constitutively high affinity for a small ligand. *British Journal of Haematology*, **115**, 131–139.

Grimaldi CM, Chen FP, Scudder LE *et al.* (1996) A Cys374Tyr homozygous mutation of platelet glycoprotein IIIa in a Chinese patient with Glanzmann's thrombasthenia. *Blood*, **88**, 1666–1675.

Hourdille P, Fialon P, Belloc F *et al.* (1986) Megakaryocytes from the marrow of a patient with Glanzmann's thrombasthenia lacked GPIb–IIIa complexes. *Thrombosis and Haemostasis*, **56**, 66–70.

Jackson DE, White MM, Jennings LK *et al.* (1998) A Ser162Leu mutation within glycoprotein (GP) IIIa (integrin beta 3) results in an unstable alpha IIb beta 3 complex that retains partial function in a novel form of type II Glanzmann thrombasthenia. *Thrombostasis and Haemostasis*, **80**, 42–48.

Kunicki TJ, Aster RH. (1978) Deletion of the platelet-specific alloantigen PlA1 from platelets in Glanzmann's thrombasthenia. *Journal of Clinical Investigation*, **61**, 1225–1231.

Lanza F, Stierle A, Fournier D *et al.* (1992) A new variant of Glanzmann's thrombasthenia (Strasbourg I). Platelets with functionally defective glycoprotein IIb-IIIa complexes and a glycoprotein IIIa 214Arg→214Trp mutation. *Journal of Clinical Investigation*, **89**, 1995–2004.

Loftus JC, O'Toole TE, Plow EF *et al.* (1990) A beta 3 integrin mutation abolishes ligand binding and alters divalent cation-dependent conformation. *Science*, **249**, 915–918.

Milet-Marsal S, Breillat C, Peyruchaud O *et al.* (2002) Two different beta3 cysteine substitutions alter alphaIIbbeta3 maturation and result in Glanzmann thrombasthenia. *Thrombosis and Haemostasis*, **88**, 104–110.

Morel-Kopp M-C, Kaplan C, Proulle V *et al.* (1997) A three amino acid deletion in glycoprotein IIIa is responsible for type I Glanzmann's thrombasthenia: importance of residues Ile325, Pro326, Gly327 for beta 3 integrin subunit association. *Blood*, **90**, 669–677.

Morel-Kopp, M-C, Melchior, C, Chen, P *et al.* (2001) A naturally occurring point mutation in the β3 integrin MIDAS-like domain affects differently $\alpha_v\beta3$ and αIIbβ3 function. *Thrombosis and Haemostasis*, **86**, 1425–1434.

Nair S, Li J, Mitchell WB, Mohanty D, Coller BS, French DL (2002) Two new beta3 integrin mutations in Indian patients with Glanzmann thrombasthenia: localization of mutations affecting cysteine residues in integrin beta3. *Thrombosis and Haemostasis*, **88**, 503–509.

Newman PJ, Seligsohn U, Lyman S *et al.* (1991) The molecular genetic basis of Glanzmann thrombasthenia in the Iraqi-Jewish and Arab populations in Israel. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 3160–3164.

Newman PJ, Weyerbusch-Bottum S, Visentin GP *et al.* (1993) Type II Glanzmann thrombasthenia due to a destabilizing amino acid substitution in platelet membrane glycoprotein IIIa. *Thrombosis and Haemostasis*, **69**, 1017.

Nurden A, Jacquelin B, Tuleja E *et al.* (2002) Reduced collagen-induced platelet aggregation in obligate heterozygotes of a Glanzmann thrombasthenia variant with a beta 3 mutation. *Thrombosis and Haemostasis*, **88**, 364–365.

Nurden P, Poujol C, Winckler J *et al.* (2002) A Ser752→Pro substitution in the cytoplasmic domain of beta3 in a Glanzmann thrombasthenia variant fails to prevent interactions between the alphaIIbbeta3 integrin and the platelet granule pool of fibrinogen. *British Journal of Haematology*, **118**, 1143–1151.

Nurden AT, Ruan J, Pasquet J-M *et al.* (2002) A novel 196Leu to Pro substitution in the beta3 subunit of the alphaIIbbeta3 integrin in a patient with a variant form of Glanzmann thrombasthenia. *Thrombosis and Haemostasis*, **13**, 101–111.

Rosenberg N, Yatuv R, Orion Y *et al.* (1997) Glanzmann thrombasthenia caused by an 11.2 kilobase-deletion in the glycoprotein IIIa (beta 3) is a second mutation in Iraqi-Jews that stemmed from a distinct founder. *Blood*, **89**, 3654–3662.

Ruan J, Schmugge M, Clemetson KJ *et al.* (1999) Homozygous Cys542→Arg substitution in GP IIIa in a Swiss patient with type I Glanzmann's thrombasthenia. *British Journal of Haematology*, **105**, 523–531.

Ruiz C, Liu CY, Sun QH *et al.* (2001) A point mutation in the cysteine-rich domain of glycoprotein (GP) IIIa results in the expression of a GPIIb–IIIa (alphaIIb beta3) integrin receptor locked in a high-affinity state and a Glanzmann thrombasthenia-like phenotype. *Blood*, **98**, 2432–2441.

Shpilberg O, Rabi I, Schiller K *et al.* (2002) Patients with Glanzmann thrombasthenia lacking platelet glycoprotein alphaIIbbeta3 (GPIIb/IIIa) and alphaVbeta3 receptors are not protected from atherosclerosis. *Circulation*, **105**, 1044–1048.

Simsek S, Heyboer H, de Bruijne-Admiraal LG *et al.* (1993) Glanzmann's thrombasthenia caused by homozygosity for a splice defect that leads to deletion of the first coding exon of the glycoprotein IIIa mRNA. *Blood*, **81**, 2044–2049.

Wang R, Shattil SJ, Ambruso DR *et al.* (1997) Truncation of the cytoplasmic domain of beta 3 in a variant form of Glanzmann thrombasthenia abrogates signaling through the integrin alpha IIb beta 3 complex. *Journal of Clinical Investigation*, **100**, 2393–2403.

Ward CM, Kestin AS, Newman P. (2000) A Leu262Pro mutation in the integrin beta3 subunit results in an alphaIIb-beta3 complex that binds fibrin but not fibrinogen. *Blood*, **96**, 161–169.

Yatuv R, Rosenberg N, Zivelin A *et al.* (2001) Identification of a region in glycoprotein IIIa involved in subunit association with glycoprotein IIb: further lessons from Iraqi-Jewish Glanzmann thrombasthenia. *Blood*, **98**, 1063–1069.

Ylanne J, Huuskonen J, O'Toole TE *et al.* (1995) Mutation of the cytoplasmic domain of the integrin beta 3 subunit. *Journal of Biological Chemistry*, **270**, 9550–9557.

Molecular diagnosis

Bray PF. (1994) Inherited diseases of platelet glycoproteins: considerations for rapid molecular characterization. *Thrombosis and Haemostasis*, **72**, 492–502.

French DL, Coller BS, Usher S. (1998) Prenatal diagnosis of Glanzmann's thrombasthenia using the polymorphic markers BRCA1 and THRA1 on chromosome 17. *British Journal of Haematology*, **102**, 582–587.

Rosenberg N, Dardik R, Rosenthal E. (1998) Mutations in the alphaIIb and beta3 genes that cause Glanzmann's thrombasthenia can be distinguished by a simple procedure using transformed lymphocytes. *Thrombosis and Haemostasis*, **79**, 244–248.

Platelet polymorphisms and thrombosis

Kunicki TJ. (2002) The influence of platelet collagen receptor polymorphisms in hemostasis and thrombotic disease. *Arteriosclerosis, Thrombosis, and Vascular Biology*, **22**, 14–20.

Reiner AP, Siscovick DS, Rosendaal FR. (2001) Platelet glycoprotein gene polymorphisms and risk of thrombosis: facts and fancies. *Reviews in Clinical and Experimental Hematology*, **5**, 262–287.

Congenital disorders of signaling and secretion

Rao AK. (2001) Congenital disorders of platelet function: disorders of signal transduction and secretion. *American Journal of Medical Science*, **316**, 69–76.

Chapter 19
The molecular basis of blood cell alloantigens

Willem H Ouwehand & Cristina Navarrete

Introduction

The transfusion of blood and fetal–maternal hemorrhage during pregnancy have provided unique models to study the immune response against a plethora of polymorphic blood cell surface markers, including the alloantigens of the human leukocyte antigen (HLA) system. Many blood cell membrane determinants show allelic variation, which can elicit the formation of alloantibodies. In nearly all transfusion situations and pregnancies, the recipient's immune system is challenged by blood cells mismatched for multiple alloantigen systems. As the difference between self and non-self is limited, alloantibodies are only formed by a subset of recipients. Red cell alloantibodies are detected in 1–1.5% of pregnant women and in 2–3% of transfused individuals, and can increase significantly in multitransfused patients. The HLA alloantigens are more immunogenic than those of the red cells, and 15–25% of multiparous women and 30–40% of patients on long-term prophylactic platelet transfusions are positive for HLA class I antibodies.

Alloantigens were initially defined as polymorphic membrane determinants identified by polyclonal alloantibodies in serum samples from alloimmunized patients or pregnant women, but the molecular basis of most alloantigens has now been resolved and it is possible to use DNA-based techniques to further characterize their polymorphism. Alloantigens can be categorized as those shared by all blood cells; for example, HLA class I and those unique to one cell type, such as Rh on red cells and human platelet antigen (HPA) on platelets (Table 19.1). When expression is limited to one type of blood cell, destruction of cells in the newborn by maternal blood cell alloantibodies may lead to the development of immune-mediated anemia, thrombocytopenia or neutropenia of the newborn. In contrast, HLA class I alloantibodies do not cause cytopenias in the newborn but may compromise the effectiveness of platelet transfusions, complicate organ

Table 19.1 Antigen expression on peripheral blood cells.

Antigens	Erythrocytes	Platelets	Neutrophils	B lymphocytes	T lymphocytes	Monocytes
A, B, H	+++	++/(+)	–	–	–	–
I	+++	++	++	–	–	–
Rh*	+++	–	–	–	–	–
K	+++	–	–	–	–	–
HLA class I	–/(+)	+++	+++	+++	+++	+++
HLA class II	–	–	–/+++**	+++	–/+++**	+++
GPIIb/IIIa	–	+++	(+)#	–	–	–
GPIa/IIa	–	+++	–	–	–	–
GPIb/IX/V	–	+++	–	–	–	–
FcRγIIIb	–	–	+++	–	–	–/+++†

*Non-glycosylated; **when activated; #inconclusive; †when differentiated to macrophages expressing FcγRIIIa.

transplantation, cause febrile non-hemolytic transfusion reactions or, on rare occasions, precipitate transfusion-related acute lung injury.

The formation of alloantibodies after an incompatible challenge in the form of blood transfusion is more the exception than the rule. In contrast to our detailed understanding of the molecular basis of blood cell alloantigens, we remain relatively ignorant about the mechanism of non-responsiveness. We have learned from animal experiments that restriction in the ability to mount an immune response is largely controlled by genes of the major histocompatibility complex (MHC) or HLA. However, the reason why, for example, some 25% of RhD-negative individuals fail to mount an anti-D response on repeated challenge with RhD-positive red cells remains elusive. An exception to this is our detailed understanding of the immune response against the HPA-1a alloantigen on platelets. There is a near-complete restriction on the ability to form HPA-1a antibodies by the HLA class II allele DRB3*0101. However, except for this example, our ability to identify the genes controlling the risk of alloimmunization remains limited and further research is needed to identify the genetic basis of this variability in responsiveness.

This chapter reviews the recent developments in the molecular aspects of blood cell alloantigens and will highlight their impact on clinical management. Recognizing the wide variety of clinical conditions in which the HLA alloantigens play a role, we have placed the main emphasis on:

- antibody-mediated cytopenias in the newborn by maternal blood cell-specific alloantibodies
- the complication of HLA class I alloimmunization in patients receiving prophylactic platelet transfusions.

The HLA antigens have also been used to introduce the molecular techniques currently used to identify alleles of genes.

Identification of HLA gene polymorphism

The impact of molecular biological techniques on our ability to scan and identify allelic variation of human genes is best exemplified by the HLA system. For decades the enormous diversity of the alloantigens of the HLA system has been a challenge in both the technical and the clinical sense. It is now obvious that the use of molecular techniques, including sequencing-based high-resolution typing, is contributing to improved outcome in transplant patients. Improved matching allows graft maintenance at lower levels of immunosuppression, which is of great importance with the emerging evidence that long-term use of potent immunosuppressive drugs is not without side effects. The following paragraphs review a number of current molecular techniques used to define alleles of the HLA genes. The same techniques may be used to define allelic variation in genes encoding other blood cell alloantigens.

Molecular typing techniques

Traditionally, the definition and characterization of the HLA molecules and polymorphisms have been carried out using serological and cellular techniques. With the development of gene cloning and DNA-based molecular techniques, it is now possible to perform a detailed analysis of these molecules at the single-nucleotide level. The result of this analysis has shown the existence of shared nucleotide sequences between alleles of the same and/or different loci. Similarly, it has been shown that there are certain locus-specific nucleotide sequences in both the coding regions (exons) and the non-coding (introns) regions of the various genes.

The DNA sequencing of a number of alleles at various MHC loci has demonstrated that most of the nucleotide substitutions are located in exons 2 and 3 of the HLA class I and exon 2 of the HLA class II molecules. These exons code for the distal membrane domains of the molecules, which form the peptide-binding groove (Figure 19.1a and b). On the basis of this information, a number of techniques have been developed to identify this polymorphism using the polymerase chain reaction (PCR) to amplify specific genes or regions to be analyzed. These include PCR-SSOP (sequence-specific oligonucleotide probing), PCR-SSP (sequence-specific priming) and conformational methods including *reference strand conformational analysis* (RSCA) and *sequencing-based typing* (SBT).

PCR-SSOP

In this technique the gene of interest is amplified by PCR using generic primers complementary to highly conserved gene segments. The PCR products are then immobilized onto support membranes (e.g. nylon membranes) and analyzed by probing the membranes with labeled oligonucleotides designed to anneal with polymorphic sequences present in the various alleles. By scoring the probes that bind to specific regions, it is possible to assign an HLA type (Figure 19.2). A recent modification includes the addition of PCR-amplified product to labeled probes immobilized on membranes or plates. This technique, which has been called reverse SSOP, is particularly useful when testing large number of samples, such as those required for bone marrow donor registries.

However, this technique does not determine whether the detected sequences are *in cis* or *in trans*; in order to resolve this ambiguity PCR-SSP must be used.

PCR-SSP (Figure 19.3)

This technique involves the use of allele-specific primers in

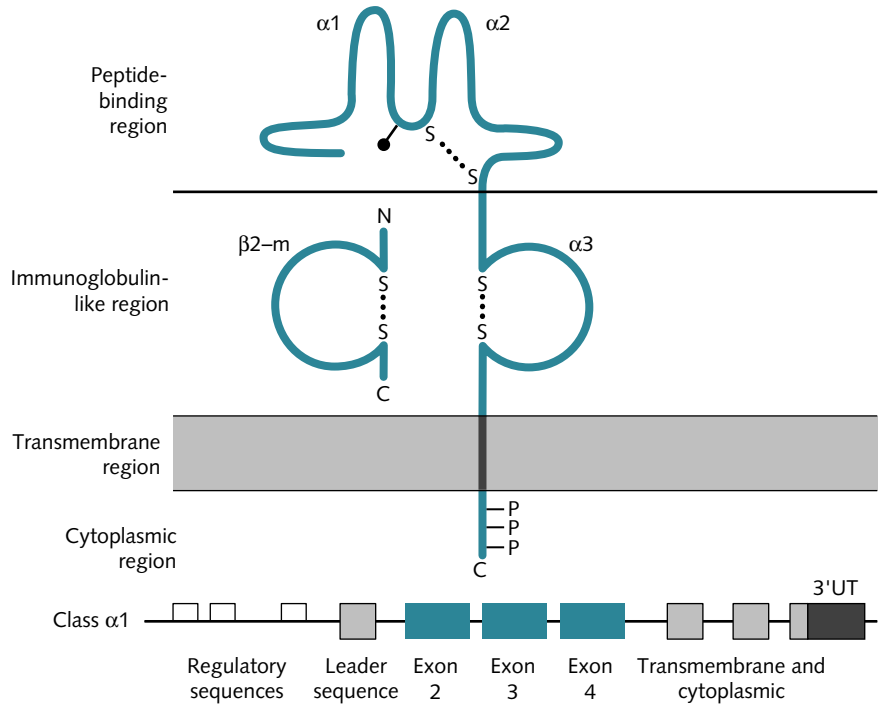


Fig. 19.1a Schematic presentation of HLA class I
The non-covalent association between the HLA class I protein (with three immunoglobulin-like domains, α_1, α_2 and α_3) and β_2-m is shown. The three α domains are encoded by three exons of the HLA class IA gene on chromosome 6.

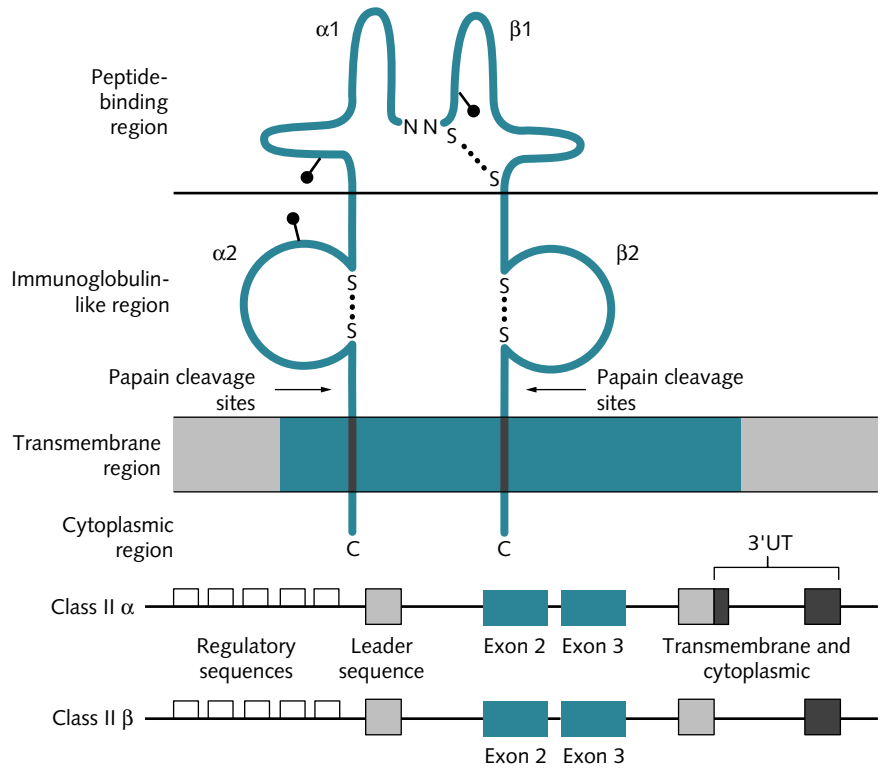


Fig. 19.1b Schematic presentation of HLA class II
The α and β chains of the HLA class II protein (each with two distinct immunoglobulin-like domains, α_1 and α_2, and β_1 and β_2) are non-covalently associated. Both domains of each chain are encoded by their respective exons of the α and β class II genes on chromosome 6.

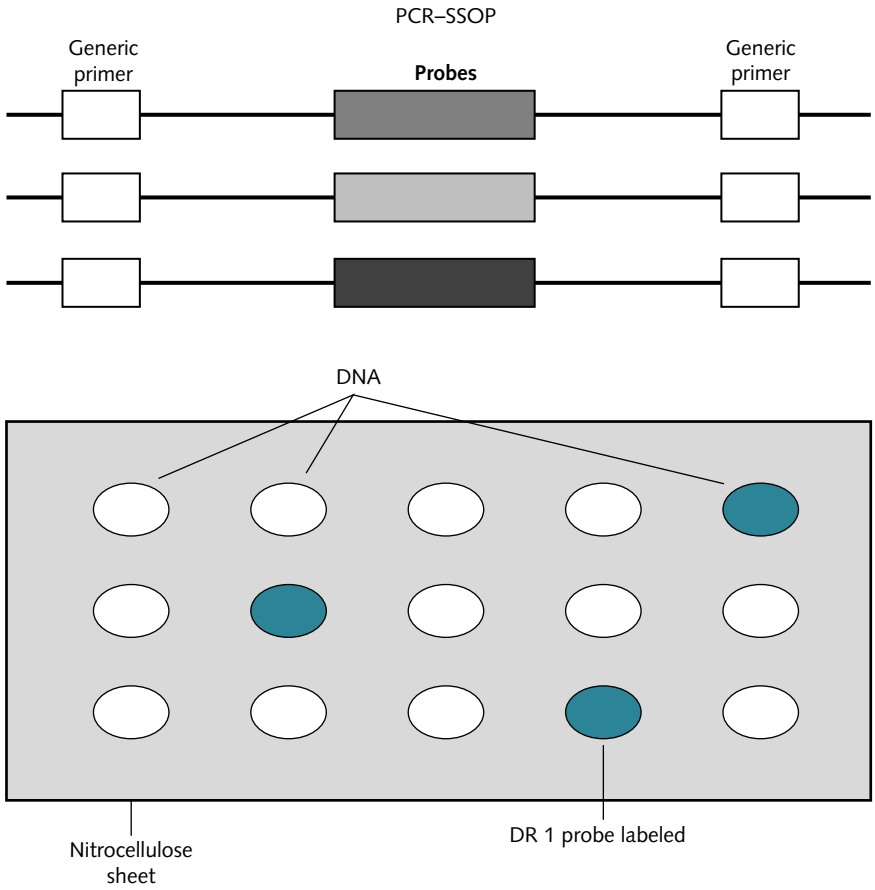


Fig. 19.2 Schematic presentation of PCR-SSOP
A schematic representation of the probing of PCR products with sequence-specific oligonucleotide probes. A segment of the HLA class II DR genes is amplified by PCR with a set of generic primers complementary to highly conserved sequences flanking the polymorphic site. The PCR product is immobilized on a nitrocellulose sheet and probed with sets of allele-specific DR probes. Binding or non-binding of the labeled probes is revealed (red and white spots, respectively).

the PCR. The amplified DNA is detected by gel electrophoresis and this allows the rapid identification of the HLA alleles in individual samples, since the read-out of this method is the presence or absence of an amplicon for which a specific primer was used. PCR-SSP was first developed to define the various HLA-DRB3 alleles. Although this is a rapid technique, many PCR reactions have to be set up for each sample, for example 24 reactions for low-resolution DR typing. An obvious advantage of PCR-SSP is that, as two sequences are detected *in cis*, ambiguities which may arise from PCR-SSOP typing can be resolved. For PCR-SSP typing, however, the target sequence of the alleles must be known, and novel unknown sequences may not always be detected. PCR-SSP is also the technique of choice for HPA typing (see *The molecular basis of platelet-specific or HPA antigens, p.* 235).

Conformational analysis methods

Conformational methods rely on the differential mobility of PCR products in gels. In single-stranded conformational polymorphism analysis (SSCA), PCR-generated DNA products are denatured by heating and rapid cooling to prevent re-annealing of the strands. The products are then run on a polyacrylamide gel, and the mobility depends upon the secondary structure of the single-stranded DNA. The major disadvantage of this technique for HLA typing is the tendency of single-stranded DNA to adopt many conformational forms under the same electrophoretic conditions, resulting in the presence of several bands from a single product.

A modification of this technique, which compares the mobilities in polyacrylamide gel electrophoresis of duplex molecules generated by mixing PCR products, is called double-stranded conformational analysis (DSCA). In this case, the mobility depends on the mismatching of the sequence and the formation of heteroduplex molecules. A modification of DSCA, known as reference strand conformational analysis (RSCA), has recently been developed and successfully applied to class I and II typing. In this technique, PCR is carried out on the DNA under test and on a reference DNA sample of known sequence using fluorescently labeled primers for the PCR of the reference DNA. The PCR products of 'tester' and 'reference' are melted and then mixed, re-annealing is allowed and

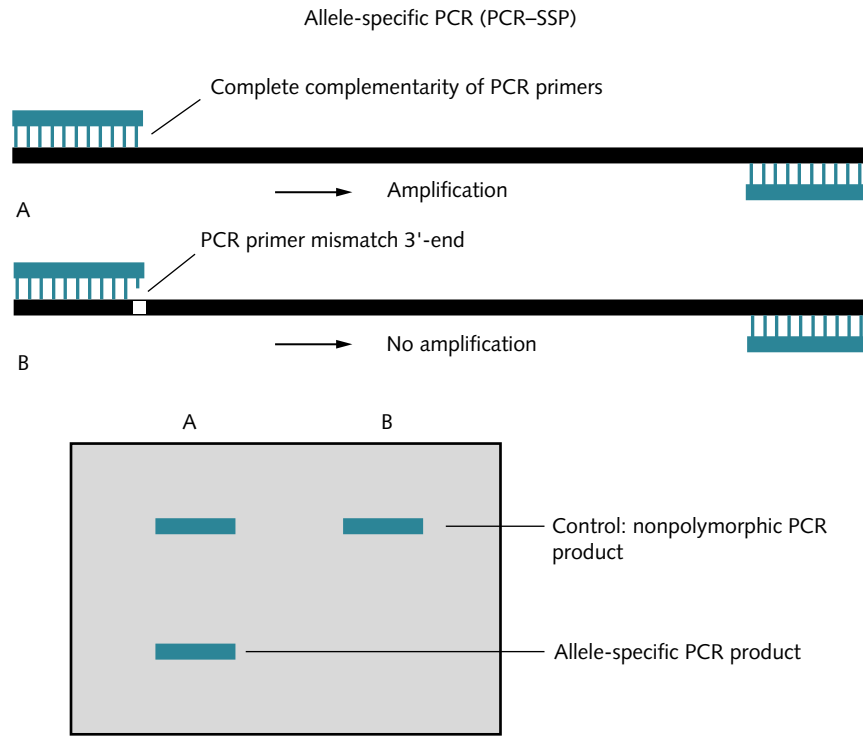


Fig. 19.3a PCR-SSP
With allele-specific PCR, specificity is obtained during the PCR. A single nucleotide mismatch at the 3′-end of the allele-specific primer (B primer in the figure) will prevent the polymerase from commencing DNA amplification. Therefore, no amplification of template DNA will occur with the B allele primer, whereas with the A allele primer a product is obtained. Ethidium bromide is used to reveal amplified DNA with ultraviolet light after DNA gel electrophoresis. Allele-specific DNA is obtained in the A reaction (lower band) but not in the B reaction. In both reactions a control PCR product is generated by amplification of a segment of the growth hormone gene (upper two bands).

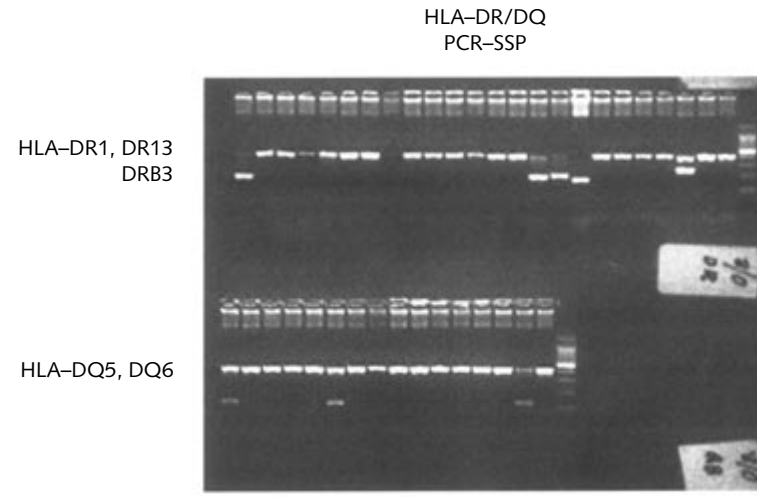


Fig. 19.3b PCR-SSP for DR and DQ alleles
Results of PCR-SSP with template DNA from a single donor. Each lane represents the result of gel electrophoresis of a single PCR reaction with one of the allele-specific DR or DQ primers. From the pattern of positive results, a DR (upper panel) and DQ (lower panel) type can be concluded; in this case DR1, DR13, DRB3 and DQ5, DQ6.

the products are run in an automated DNA sequencer. Only those duplexes containing a labeled strand are detected; that is, the reference DNA homoduplex and the heteroduplexes of reference and tester DNA. As the mobility of every known allele with the reference DNA has been established, the results obtained with unknown samples can be read. Test samples running at an as-yet unobserved mobility could represent possible new alleles, whose sequence can be determined by direct sequencing.

Sequencing-based typing

The principle of DNA sequencing is relatively straightforward. It involves the denaturation of the DNA to be analyzed to provide a single-strand template; a sequencing primer is then added and the extension is performed by the addition of polymerase in the presence of excess nucleotides. The sequencing mixture is divided into four tubes, each of which contains a specific dideoxyribonucleoside triphosphate (ddATP).

When this is incorporated into the DNA strand, elongation is interrupted, leading to chain termination. In each reaction there is random incorporation of the chain terminators and therefore products of all sizes are generated. The products of the four reactions are then analyzed by electrophoresis in parallel lanes of a polyacrylamide–urea gel and the sequence is read by combining the results of each lane using an automated DNA sequencer.

HLA antigens

These are a group of highly polymorphic cell surface molecules that play a central role in the induction and regulation of immune responses, and as such they are involved in self/non-self recognition, tolerance, rejection of allografts and graft-versus-host disease. The genes coding for these molecules form part of a complex genetic system called the MHC, located on the short arm of chromosome 6. This region spans a distance of approximately 4000 kb and is divided into HLA class I, class II and class III genes. Class III includes a group of non-MHC genes coding for proteins with various immunological functions, such as the complement factor 4 and tumor necrosis factor (TNF-α).

The development of recombinant DNA technology has led to increased understanding of the genetic complexity, structure and function of the HLA genes and molecules.

HLA class I genes

The HLA class I genes have been classified according to their structure and function as classical and non-classical, or class Ib genes. The classical HLA class I genes HLA-A, B and Cw code for heterodimers formed by a heavy (α) chain of approximately 43 kDa, non-covalently linked to the β_2 microglobulin (β_2-m) light chain of 12 kDa (Figure 19.1a). The latter is coded for by a gene located outside the HLA region on chromosome 15. The extracellular portion of the α chain has three domains (α_1, α_2 and α_3) encoded by exons 2, 3 and 4, respectively. Each domain is approximately 90 amino acids in length. The transmembrane and cytoplasmic domains are encoded by exons 5, 6 and 7, respectively. The β_2-m, which confers stability on the molecule, is non-covalently linked to the α_3 domain (Figure 19.1a).

The α_1 and α_2 domains are the most polymorphic regions of these molecules and form a groove consisting of two α helices, with an antiparallel-running β-pleated sheet forming the floor of the groove. This groove, which is approximately 25 Å long and 10 Å wide, can accommodate a variety of antigen-derived peptides of about eight to ten amino acids long to be presented to T cells.

In addition to the classical HLA class I genes, the non-classical HLA class I genes are also located in this region. They include HLA-E, -F and -G and their exon/intron organization is similar to that of the classical class I genes, but they have a more restricted polymorphism. The HLA class I genes are expressed on most tissues and blood cells, including T and B lymphocytes and platelets (Table 19.1). The non-classical class I genes HLA-E and -F are expressed on most tissues tested so far, whereas HLA-G has so far only been detected on trophoblasts and monocytes.

More recently, two MHC class I chain-related genes (MIC-A and MIC-B), located centromeric to HLA-B, have been described. Unlike the classical and non-classical HLA class I, MIC genes do not require binding to the β_2-microglobulin or peptide in order to be expressed on the cell surface. So far MIC expression has been detected on freshly isolated endothelial cells, fibroblasts, keratinocytes and monocytes. They have also been found to be expressed on intestinal epithelial cells as a result of stress, and on a variety of tumors of epithelial origin.

HFE is another non-classical class I gene, located 4 Mb telomeric of HLA-A. This gene has been found to be responsible for the development of hereditary hemochromatosis (HH). A single point mutation, 845A, replacing cysteine with tyrosine at position 282 (C282Y) is found in over 90% of HH patients in the UK. The other two mutations, replacing histidine by aspartate at amino acid position 63 (H63D) and serine by cysteine at amino acid 65 (S65C), appear to be associated with milder forms of HH. This gene does not have a direct immune function as it has lost the ability to bind antigenic peptides due to the closure of the antigen-binding groove. However, since HFE can bind to the transferrin receptor, and in this way regulate iron uptake and availability, it is possible that HFE may indirectly be involved in the regulation of immune responses.

HLA class II genes

The HLA class II genes DR, DQ and DP are all located within the HLA class II region. There is one non-polymorphic DRA and nine highly polymorphic DRB genes, of which DRB2, B6 and B9 are pseudogenes. The number of DRB genes expressed in each haplotype varies depending upon the DRB1 allele expressed; for example, HLA DR1, DR103, DR8 and DR10 haplotypes only express the DRB1 gene. DR15 and DR16 haplotypes additionally express the DRB5 gene, which codes for the DR51 product. The HLA DR17, DR18, DR11, DR12, DR13 and DR14 haplotypes also express the DRB3 genes, which code for the DR52 specificity, whilst the HLA DR4, DR7 and DR9 alleles also express the DRB4 gene, which encodes the DR53 product. There are a few exceptions to this gene distribution; for example, a DRB5 gene has been found

linked to a DR1 haplotype, and null DRB5 and DRB4 genes have been identified.

In contrast, there are two DQA and three DQB genes; of these, only A1 and B1 are expressed and both are polymorphic. Similarly, there are two genes, A1 and B1, for DPA and DPB, and both are polymorphic.

Other HLA-related genes located within the MHC class II region include the LMP2, LMP7, TAP1 and TAP2, which are genes involved in the transport and processing of peptides presented by class I molecules, and the HLA-DMA and DMB and HLA-DOA and -DOB genes, which participate in the loading of peptides in the HLA class II molecules.

The HLA class II genes (DR, DQ and DP) are constitutively expressed on B lymphocytes, monocytes and dendritic cells, and on activated T lymphocytes and granulocytes (Table 19.1). HLA class II expression can also be induced on non-hematopoietic cells, such as fibroblasts and endothelial cells, as the result of activation or by the effect of certain inflammatory cytokines, such as γ-interferon and TNF-α.

Function

The HLA molecules play a pivotal role in the induction and regulation of the immune response. Both the phenomenon of MHC restriction and the development of tolerance, learnt as T cells go through the thymus, result in the selection of a T-cell repertoire that will form the basis of an individual's capacity to respond to antigens. HLA class I molecules are primarily but not exclusively involved in the presentation of endogenous antigens to $CD8^+$ cytotoxic T cells, whereas the HLA class II molecules present primarily, but not exclusively, exogenous antigenic peptides to $CD4^+$ helper T cells. These cells, once activated, can initiate and regulate a variety of processes leading to the maturation and differentiation of cellular and humoral effector cells, including the secretion of cytokines. The presentation of antigenic peptides is a highly regulated process and requires fine interaction between the antigenic peptide, the antigen-binding groove of the HLA molecules and the T-cell receptor. Allelic variation of the HLA molecules can profoundly affect the ability to present certain peptides because of the presence or absence of critical contact residues in the peptide-binding groove.

HLA molecules on donor cells loaded with donor-derived peptides can also be recognized directly by T cells of the host by a mechanism called allorecognition. Two pathways of allorecognition, direct and indirect, have been identified, both of which lead to the strong alloimmunization seen in patients receiving blood transfusion or a solid organ or bone marrow/stem cell transplantation.

More recently, it has been shown that both classical and non-classical HLA class I molecules interact with two functionally distinct types of receptors, inhibitory and activating, present on natural killer (NK) cells. These receptors belong to two families, the Ig superfamily, also called killer Ig receptors (KIRs), and the C-type lectin superfamily CD94, which can covalently assemble with several members of the NKG2 family. The KIR receptors interact with products of the HLA-A, -B, -Cw and -G loci, whereas CD94-NKG2 recognize the non-classical HLA-E molecule presenting peptides derived from several HLA-class I, A, B or C alleles and from HLA-G. MIC-A and MIC-B gene products are recognized by receptors present on both NK and γδT cells.

Thus, HLA molecules have become increasingly relevant in a variety of clinical situations, such as susceptibility to certain autoimmune and infectious diseases, solid organ and stem cell transplantation, and in blood cell alloimmunization. With regard to the latter, two examples will be discussed: refractoriness for prophylactic platelet transfusion by HLA alloimmunization (Figure 19.4), and HLA class II restriction of the formation of anti-HPA-1a antibodies.

Prophylactic platelet transfusions

Prophylactic platelet transfusions are essential to prevent bleeding during intensive chemotherapy or other myeloablative therapies. Increments in the platelet count after the infusion of an adult dose of donor platelets ($>250 \times 10^9$/L) are frequently disappointing. Non-immune factors, such as splenomegaly, bleeding, sepsis, fever and certain drugs (e.g. amphotericin), can compromise the beneficial effect of donor platelet infusions. In 10–20% of patients the problem of poor increments is further compounded by antibody-mediated destruction of donor platelets. Despite this, the clinical definition of refractoriness remains much disputed; clinically, the picture of an increased frequency of platelet transfusions to maintain satisfactory platelet counts and effective hemostasis requires further laboratory investigations for HLA class I and for HPA alloantibodies, platelet autoantibodies or high-titer anti-A or anti-B antibodies.

The ability to type donors and patients for HLA class I A and B and HPA by molecular techniques using genomic DNA (see *Molecular typing techniques, p. 226*, and *The molecular basis of platelet-specific or HPA antigens, p. 235*) has resulted in more accurate matching of donors with patients, with an improved platelet recovery. The algorithm currently used for the management of the alloimmunized patient with poor increments is shown in Figure 19.4.

Traditionally, HLA matching for the provision of compatible platelets for patients who have become immunologically refractory to random platelet transfusion has been based on the serological definition of these antigens at the specificity level. More recently a new approach based on the identification of linear or conformational epitopes present in each HLA

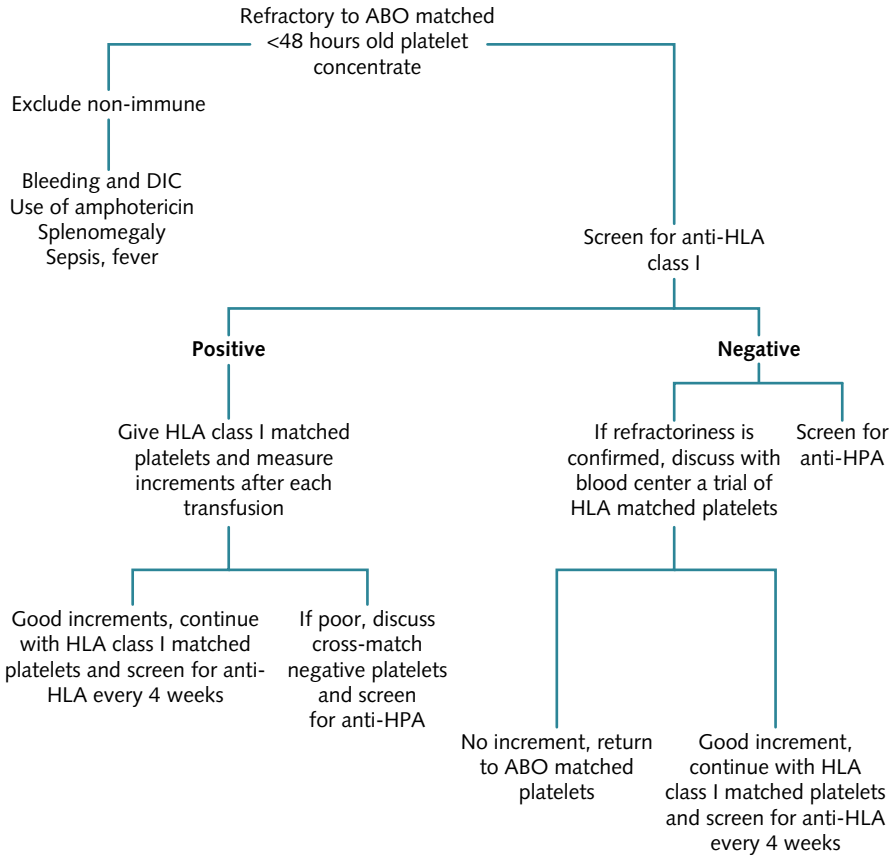


Fig. 19.4 Platelet transfusions in alloimmunized patients
An algorithm outlining the decision process for the management of alloimmunized patients refractory for random donor platelets. After confirmation of refractoriness for random donor platelets, patients are screened for HLA class I alloantibodies and, if positive, HLA class I matched platelets are transfused. In 20–30% of patients, increments with HLA class I matched platelets are poor and screening for HPA antibodies should follow. Also, the possible presence of potent anti-A or anti-B should be excluded since platelets do carry ABO blood group antigens. If there are no detectable HLA class I antibodies, a trial of HLA-matched platelets and screening for HPA antibodies should be considered (right arm of algorithm).

allele has been described. According to this strategy, each HLA antigen is converted into a string of potentially immunogenic epitopes, which are represented by amino acid triplets on exposed parts of the HLA chains accessible to alloantibodies. Consequently, it is possible to determine the number of triplets which are either shared or different between the donor and the recipient. The algorithm also performs intra- and inter-locus comparisons of polymorphic triplets in amino acid sequence positions, to determine the spectrum of non-shared triplets between HLA antigens of the donor and the patient. This approach relies on the use of the HLAMatchmaker computer program developed by Dr Rene Duquesnoy (2000, *Human Immunology*, **63**, 3395–2).

Hemolytic disease of the newborn (HDN)

The Rh system

The Rh system is the most immunogenic red cell blood group system after the ABO system. Whilst the antigens of the latter are of carbohydrate nature, the Rh antigens are non-glycosylated proteins. Two-dimensional gel electrophoresis of trypsin digests of Rh proteins, which could be performed once human monoclonal antibodies specific for Rh alloantigens were available, provided evidence of a high level of homology between the proteins carrying the RhD and the RhC/c and RhE/e alloantigens.

The Rh genes

Cloning of the *RhD* and *RhCE* genes (carrying the RhD antigen and the RhC/c and RhE/e antigens, respectively) revealed a high level of sequence homology (Table 19.2), confirming the observation made by gel electrophoresis. These molecular studies suggested that the C/c and E/e antigens are localized on a single protein, the RhCE protein encoded by the *RhCE* gene and its allelic variants. The recent expression of the *RhD* and *RhCE* genes in K562 cell lines by retroviral transfection provided the ultimate proof that the RhD alloantigen and the RhC/c and E/e alloantigens are localized on two distinct proteins encoded by these two genes and their respective allelic variants.

The 75-kb *RhD* gene, with 10 exons, encodes a 30-kDa non-glycosylated protein of 417 amino acids of unknown function

Table 19.2 Amino acid sequence of the RhD and RhCE proteins.

```
CE    MSSKYPRSVRRCLPLWALTLEAALILLFYFFTHYDASLEDQKGLVASYQVGQDLTVMAALGLGFLTSNFRRHSWSS
D     ------------------------------------------------------------I-------S--------

      VAFNLFMLALGVQWAILLDGFLSQFPPGKVVITLFSIRLATMSAMSVLISAGAVLGKVNLAQLVVMVLVEVTALGT
      -------------------------S-----------------L-----VD----------------------N

      LRMVISNIFNTDYHMNLRHFYVFAAYFGLTVAWCLPKPLPKGTEDNDQRATIPSLSAMLGALFLWMFWPSVNSPLL
      ----------------MM-I---------S----------E----K--T----------------I----F--A--

      RSPIQRKNAMFNTYYALAVSVVAISGSSLAHPQRKISMTYVHSAVLAGGVAVGTSCHLIPSPWLAMVTLGLVAGLI
      ----E----V------V-----------------G---K-----------------------------------

      SIGGAKCLPVCCNRVLGIHHISVMHSIFSLLGLLGEITYIVLLVLHTVWNGNGMIGFQVLLSIGELSLAIVIALTS
      -V----Y--G--------P-S-I-GYN----------I-------D--GA-------------------------

      GLLTGLLLNLKIWKAPHVAKYFDDQVFWKFPHLAVGF
      -----------------E------------------
```

Amino acids are given in single-letter code. Identicity has been underlined.

which traverses the red cell membrane 12 times (Figure 19.5). The gene is deleted in RhD-negative individuals (15% of the Caucasian population). The RhC/c and E/e alloantigens are carried on the highly homologous 30-kDa RhC/cE/e protein, which differs by only 36 of the 417 amino acids from RhD (Table 19.2). The difference between RhC and Rhc is associated with six nucleotide substitutions, of which four result in a replacement. One of these, residue 103, is exofacial and seems to be the most critical one of the C/c polymorphism. The difference between RhE and Rhe is defined by a single proline–alanine-226 replacement in the fourth extracellular loop of the RhCE protein (Figure 19.5).

Immunogenicity of RhD and prevention of immunization

The absence of the RhD protein in RhD-negatives is the most plausible explanation of its relatively high immunogenicity, as the immune system has not been tolerized by a lookalike RhD protein encoded by an allele of the *RhD* gene. Until the late 1960s, the formation of anti-D in RhD-negative pregnant women had a high incidence and was associated with significant neonatal morbidity and mortality because of hydrops fetalis and kernicterus. The discovery that the formation of anti-D could be prevented by the administration of passive IgG anti-D combined with a much reduced family size has led to a steep decline in HDN-associated mortality. It is estimated that, per annum in the UK, approximately 150–200 cases of severe RhD immunization require intrauterine treatment. The officially reported annual mortality for RhD HDN is nine cases, but estimates from screening laboratories suggest that the figure is closer to 50.

Knowledge of the structure of the *RhD* gene now makes it possible to give better support to RhD-immunized women, as the fetal *RhD* genotype can be determined in the first trimester from amniocyte DNA or from maternal serum. The former is now routinely used in cases of severe RhD immunization with an *RhD* (*D*/*d*) heterozygous partner, whereas the latter technique, which is non-invasive, will need further validation before routine application in genetic counseling. The immunogenicity of RhE in RhE-negative individuals is low and the change in severe HDN because of anti-E is infinitely small when compared with anti-D. Its poor immunogenicity is best explained by the relatively small difference between self and non-self. The greater sequence difference between RhC and Rhc is reflected in a greater risk of potent alloantibody formation during pregnancy, although severe HDN by anti-C and anti-c antibodies is relatively uncommon.

The Kell system

The K antigen of the Kell system is, after RhD, the second most immunogenic protein blood group alloantigen system on the red cell membrane. The Kell alloantigens K and k are carried on a 93-kDa, type II transmembrane protein of 732 amino acids. The *KEL* gene spans about 21.5 kb, with 19 exons, and has sequence homology with genes of the protein family of zinc-binding endopeptidases, like the CALLA (CD10) antigen and human common acute lymphoblastic leukaemic antigen. The difference between K and k is a single base change in exon 6, causing a methionine→threonine-193 replacement. It is of interest to note that, despite this minor difference between the two forms, K is one of the most immunogenic blood group

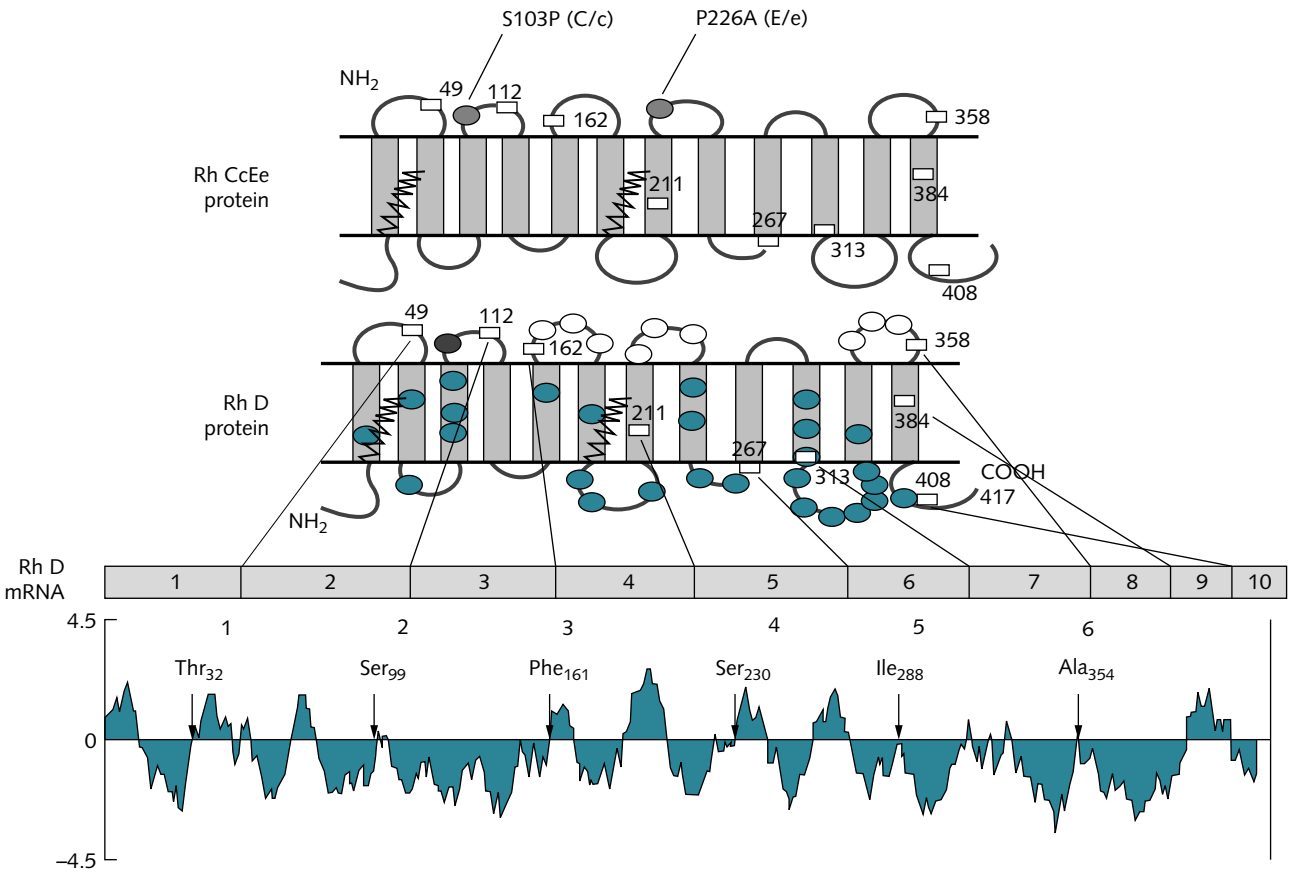


Fig. 19.5 RhD and RhCE proteins

Schematic representation of the topography of the RhCE (upper figure) and RhD (lower figure) proteins. Both 30-kDa non-glycosylated proteins traverse the red cell membrane 12 times. The difference between RhD and RhCE is defined by 36 of the 417 amino acids (black circles are RhD-specific residues; red circles indicate C/c and E/e polymorphisms). The rectangular boxes (1–10) indicate the positions of exon–exon boundaries in the *RhD* mRNA, and a hydropathy plot of the RhD protein is shown at the bottom.

substances. This is possibly explained, although not proven, by the absence of an N-glycosylation site in the k protein at residue 191.

Immunogenicity of K and the mechanism of fetal anemia

Transfusion frequently triggers the formation of anti-K. As anti-K antibodies can cause severe HDN, it is recommended that K-negative blood should be transfused to all girls and women of childbearing age. In K alloimmunization, fetal anemia seems to be mainly caused by antibody-mediated inhibition of erythroid maturation. The question of why anti-K downregulates erythropoiesis has not been answered, but it points to a possible regulatory role of the K-carrying endopeptidase in erythroid differentiation and maturation. Owing to the aplastic nature of the anemia, the degree of fetal anemia cannot be assessed by the measurement of bilirubin levels in the amniotic fluid, but should be based on the obstetric history, maternal anti-K level, and paternal K status. If there is a high-risk scenario, with a K-positive partner, the fetal K type should be determined by PCR on amniocyte DNA. If the fetus is K-positive, fetal anemia should be assessed and treated by peri-umbilical blood sampling (PUBS).

Red cell alloantigens other than Rh and K

As is the case for the Rh and K systems, the molecular structure of most other blood group alloantigens has been resolved over the last two decades. Detailed discussion of these systems is beyond the scope of this chapter, but PCR-based typing can now be performed for most clinically relevant blood group systems (e.g. Duffy, Jk, M/N and S/s). This knowledge is of use when blood group alloantibodies are associated with severe HDN. These users include patients with thalassemia, in whom blood transfusion has been given before a complete red cell phenotype is established.

The high level of sequence homology between the *RhD* and *RhCE* genes and the occurrence of crossing-over events, which are often associated with reduced expression of the Rh antigens, requires ample expertise in the use of PCR for prenatal diagnosis. Therefore, such tests should be performed in laboratories with a scientific interest in blood group genetics.

Frequency

Of all pregnant women, 1–1.5% screen positive for red cell alloantibodies. In roughly half of them, the antibodies are of possible clinical significance, and antibody potency needs monitoring during pregnancy to determine the risk of significant hemolysis requiring therapy. Severe disease with intrauterine intravascular transfusion therapy occurs in approximately 1 in 5000 pregnancies and is mainly due to anti-D or anti-K.

Pathology

Maternal alloantibodies can be formed against a blood group alloantigen present on the fetal red cells but absent from the maternal ones. RhD, K, Rhc and RhC are the main culprits with respect to HDN. Red cell alloantibodies of the IgG class can cross the placenta, bind to the fetal red cells and shorten their survival. In K immunization, hyperbilirubinemia is a less reliable parameter for the prediction of the severity of fetal anemia, and these pregnancies can be complicated by early fetal loss.

Treatment of HDN

An increased concentration of unconjugated bilirubin in the neonate poses the risk of kernicterus. Treatment of severe hemolytic anemia and hyperbilirubinemia in the postdelivery setting is by exchange transfusion with compatible donor red cells. When a pregnancy is preceded by one with a history of severe HDN, intrauterine intravascular transfusion of compatible donor red cells by PUBS is the treatment of choice and has a good outcome in 90% of cases. Genetic counseling of couples with a heterozygous partner has been greatly helped by the discovery of the *Rh* and *K* genes (see *The Rh genes*, *p. 232*, and *The Kell system*, *p. 233*).

Prevention of HDN

Before the introduction of anti-D prophylaxis, most HDN cases were caused by RhD immunization. The routine RhD testing of all pregnant women, combined with anti-D prophylaxis for RhD-negatives carrying an RhD-positive infant, has been extremely successful in lowering HDN-associated morbidity and mortality. Screening for clinically significant red cell alloantibodies in pregnant women is the standard of care in most European countries and cases of possible severe disease should be identified early in pregnancy to allow the prevention of morbidity or mortality.

Neonatal alloimmune thrombocytopenia

Platelet-specific or HPA alloantigen systems

Besides the alloantigens shared with other blood cells (e.g. HLA class I A and B alloantigens), platelets also express alloantigens which are carried on proteins uniquely expressed on platelets but not on other blood cells (Table 19.1). Alloimmunization against platelet-specific antigens or HPA is associated with three clinical syndromes:

1 Neonatal/fetal alloimmune thrombocytopenia (NAITP).
2 Post-transfusion purpura (PTP).
3 Refractoriness for platelet transfusions (PR).

NAITP was first described by van Loghem in 1959 and was initially thought to be a rare disorder. Prospective screening studies in pregnant (Caucasian) women have shown that 1 in 1100 neonates have severe thrombocytopenia ($<50 \times 10^9$/L) due to maternal anti-HPA-1a, confirming the notion that the most frequent cause of severe thrombocytopenia in the term newborn is maternal alloantibodies against a fetal HPA alloantigen.

This serious clinical condition is caused by the destruction of fetal/neonatal platelets by maternal HPA alloantibodies of the IgG class. Cerebral bleeds in the perinatal period are the most concerning complication, which either occur *in utero* or during delivery. In cases of severe thrombocytopenia ($<20 \times 10^9$/L), there remains a small but definite risk of this serious complication in the first days of life, warranting treatment. For proper clinical management, the cause of severe thrombocytopenia in an otherwise healthy term neonate should be determined with urgency and correction of a count less than 20×10^9/L by platelet transfusion is of utmost importance. This should precede the outcome of platelet antibody investigation, as this can be time-consuming.

The molecular basis of platelet-specific or HPA antigens

So far, 19 platelet-specific alloantigen systems have been described. All are bi-allelic and have been mapped to certain membrane proteins. Eleven of the 19 alloantigen systems are on the integrin heterodimer $\alpha_{IIb}\beta_3$ or GPIIb/IIIa (Table 19.3 and Figure 19.6); of the remaining eight, three are on GPIb/IX/V, two on the integrin $\alpha_2\beta_1$ or GPIa/IIa and one each on GPIV, GPV and CD109. With the advent of PCR, the molecular basis

Table 19.3 Platelet-specific alloantigen systems.

System	Antigen	Alternative names	Glycoprotein	Nucleotide change	Amino acid change
HPA-1	1a	Zw^a, Pl^{A1}	GPIIIa	T^{196}	$Leucine^{33}$
	1b	Zw^b, Pl^{A2}		C^{196}	$Proline^{33}$
HPA-2	2a	Ko^b	GPIbα	C^{524}	$Threonine^{145}$
	2b	Ko^a, Sib^a		T^{524}	$Methionine^{145}$
HPA-3	3a	Bak^a, Lek^a	GPIIb	T^{2622}	$Isoleucine^{843}$
	3b	Bak^b		G^{2622}	$Serine^{843}$
HPA-4	4a	Yuk^b, Pen^a	GPIIIa	G^{526}	$Arginine^{143}$
	4b	Yuk^a, Pen^b		A^{526}	$Glutamine^{143}$
HPA-5	5a	Br^b, Zav^b	GPIa	G^{1648}	$Glutamic\ acid^{505}$
	5b	Br^a, Zav^a, Hc^a		A^{1648}	$Lysine^{505}$
HPA-6w			GPIIIa	G^{1564}	$Arginine^{489}$
	6bw	Ca^a, Tu^a		A^{1564}	$Glutamine^{489}$
HPA-7w			GPIIIa	C^{1267}	$Proline^{407}$
	7bw	Mo		G^{1267}	$Alanine^{407}$
HPA-8w			GPIIIa	T^{2004}	$Arginine^{636}$
	8bw	Sr^a		C^{2004}	$Cysteine^{636}$
HPA-9w			GPIIb	G^{2603}	$Valine^{837}$
	9bw	Max^a		A^{2603}	$Methionine^{837}$
HPA-10w			GPIIIa	G^{281}	$Arginine^{62}$
	10bw	La^a		A^{281}	$Glutamine^{62}$
			GPIIIa	G^{1946}	$Arginine^{633}$
		Gro^a		A^{1946}	$Histidine^{633}$
		Iy^a	GPIbβ	G^{141}	$Glycine^{15}$
				A^{141}	$Glutamic\ acid^{15}$
		Oe^a	GPIIIa		
		Va^a	GPIIb/IIIa		
		Gov^a	CD109		
		Gov^b			
		Pl^T	GPV		
		Vis	GPIV		
		Pe^a	GPIbα		
		Sit^a	GPIa		

of all but five of the systems has been resolved in the last decade. With the exception of one system, the difference between the two alleles is a single-nucleotide substitution in the gene encoding the relevant glycoprotein (Table 19.3). Amplification of genomic DNA by PCR-SSP (see *PCR-SSP*, *p. 226*, and Figure 19.7) can be used to type donors and patients, even when the latter are thrombocytopenic.

Immunogenicity and immune response restriction

As the difference between self and non-self is defined by a single amino acid substitution, the immunogenicity of the HPA alloantigens is relatively poor when compared with that of some of the other blood cell alloantigens (e.g. RhD and HLA class I). The two most clinically relevant HPA antigens are HPA-1a and HPA-5b on the β_3 and α_2 integrins, that is, those on platelet glycoproteins GPIIIa and GPIa, respectively. Alloantibodies against other HPA alloantigens are observed infrequently in pregnancy but do occur, albeit at a low frequency, in hemato-oncological and other patients on long-term prophylactic platelet transfusion.

The HPA-1a and 1b alloantigens are based on a cytosine for guanidine mutation in the GPIIIa gene, resulting in a leucine–proline-33 mutation. Why the immunogenicity of HPA-1a (leucine-33) is magnitudes higher than that of its antithetical antigen HPA-1b (proline-33) was initially not well understood. In the early 1980s, it was discovered that the formation of anti-HPA-1a in pregnancy was positively associated with the HLA haplotype A1, B8, DR3. Further study revealed that nearly all antibody formers were positive for the 0101 allele of the DRB3 gene (DRB3*0101 or DR52a; see *HLA class II*

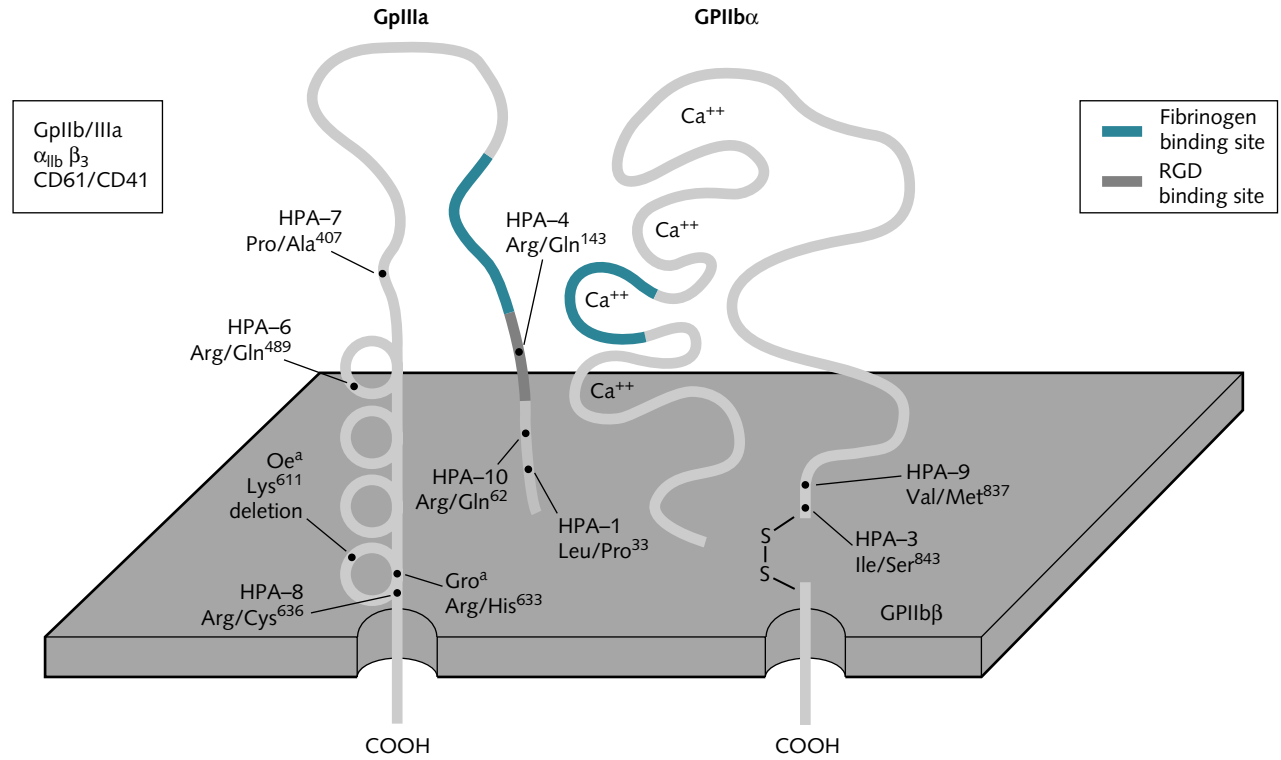


Fig. 19.6 Schematic presentation of GPIIb/IIIa

Schematic presentation of platelet GPIIb/IIIa or the $\alpha_{IIb}\beta_3$ integrin. GPIIIa is recognized by murine monoclonal antibodies of the CD61 cluster and the heterodimer by antibodies of the CD41 cluster. The amino acid substitutions arising from the allelic variation of the GPIIb and GPIIIa genes are depicted by black dots and the name of the HPA system is noted. Amino acids are given as three-letter acronyms. The fibrinogen-binding site is shown in red and the Arg/Gly/Asp (RGD)-binding site in dark gray. The RGD peptide is the minimal fibrinogen-derived peptide which binds GPIIb/IIIa.

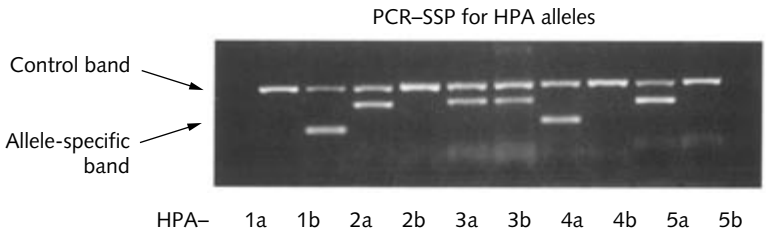


Fig. 19.7 PCR-SSP for HPA alleles

Results of the agarose gel electrophoresis of PCR products obtained by amplification of segments of the *GPIIIa*, *GPIbα*, *GPIIb* and *GPIa* genes using allele-specific primers for the 'a' and 'b' alleles of the HPA-1, -2, -3, -4 and -5 systems. The results for this donor are HPA-1b1b, 2a2a, 3a3b, 4a4a and 5a5a. Products of the allele-specific amplification are the lower bands. In all reactions a set of control primers has been included to produce an amplicon (upper band) derived from the growth hormone gene.

genes, p. 230, and Figure 19.8). A prospective study in 25 000 pregnant women showed that this class II marker has an odds ratio of 140, which makes it one of the most reliable HLA associations reported to date, with negative predictive power equal to that of HLA B27 in ankylosing spondylitis. A difference in the efficiency of presentation of the GPIIIa-leucine-33 (HPA-1a)-derived oligopeptide between DRB3*0101-positive and -negative antigen-presenting cells to CD4-positive T cells is the most likely explanation of this restriction in allo-immunization. The frequency of the HLA DRB3*0101 allele in Caucasians is 33%, and this marker therefore has a high negative predictive value but a low positive one for anti-HPA-1a formation.

Allele frequencies

In Caucasians, the allele frequency for most HPA systems is skewed towards the 'a' allele. The allele frequencies vary

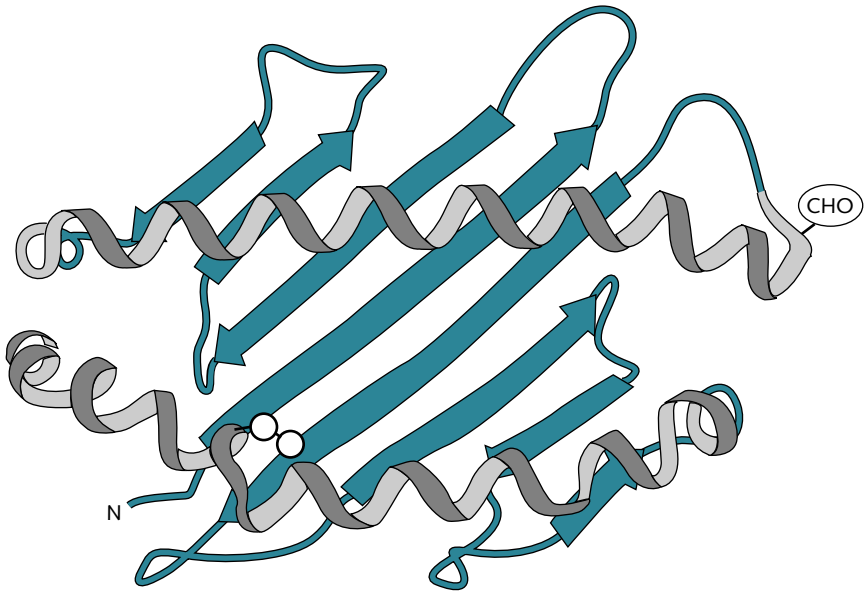


Fig. 19.8 Structure of HLA DRB3*0101 molecule
A view of the peptide-binding site of the HLA class II DRB3*0101 molecule as seen by the T-cell receptor of a T lymphocyte. Two α helices (see Figure 19.1b) are resting on a base of antiparallel-running β-pleated sheets. The antigen-derived peptide is not shown, but in our example of immune response restriction an oligopeptide derived from fetal GPIIIa-Leu33 will be juxtaposed between the two helices. The presence of this peptide defines the difference between self and non-self and triggers the proliferation of GPIIIa-Leu33 (HPA-1a) alloantigen-specific helper T cells.

between populations; for example, GPIIIa-proline-33 (HPA-1b) is extremely rare or absent in the Far East, whilst the opposite is the case for the GPIIIa-glutamine-143 form (HPA-4b).

Incidence of NAITP

Alloantibodies against the HPA-1a (Pl^{A1}, Zw^{a}) alloantigen occur in 1 in 365 pregnancies and cause severe thrombocytopenia with a neonatal platelet count of less than 50×10^9/L in 1 in 1100 term neonates.

Pathology

Maternal IgG alloantibodies against a fetal HPA alloantigen can cross the placenta and bind to the fetal platelets, thus reducing platelet survival. HPA-1a and HPA-5b are the two clinically most relevant alloantigens.

Characteristics

Neonatal alloimmune thrombocytopenia presents in the otherwise healthy newborn as petechiae or ecchymoses, or is found coincidentally by a whole blood count. Severe cases can present with neurological symptoms because of cerebral bleeds or with hydrops fetalis or cerebral cysts. Diagnosis is based on the detection of HPA alloantibodies in the maternal serum combined with a parental HPA incompatibility, as determined by PCR-SSP.

Inheritance

All systems described to date are bi-allelic with codominant expression of both alleles. Homozygous women can produce alloantibodies against a paternally inherited platelet alloantigen present on the fetal platelets but absent from the maternal ones. In cases where the partner is homozygous for the relevant alloantigen, there is a 100% chance that his future offspring will be at risk, whilst in the case of heterozygosity 50% will be affected.

Treatment

A neonatal platelet count $<20 \times 10^9$/L should be corrected immediately, preferably with HPA-1a- and 5b-negative donor platelets, as these will be compatible with the maternal HPA alloantibody in over 95% of cases. In cases of absence of HPA-compatible platelets, initial transfusion of random ABO/RhD-compatible platelets should be considered, followed by the transfusion of compatible ones if the platelet count dips again. In a typical case the platelet count should recover to normal within a week, although a more protracted recovery can occur.

In a subsequent pregnancy a decision needs to be taken on whether treatment of the fetus, the mother or both is indicated, or whether conservative management is acceptable. When conservative management is acceptable, the pregnancy should be closely monitored and the mother should be advised to avoid any non-steroidal anti-inflammatory drugs and aspirin. The delivery needs careful planning between obstetric and pediatric teams in close consultation with the consultant hematologist. Treatment during pregnancy should be reserved for cases in which the estimated risk of severe fetal/neonatal thrombocytopenia is considerable, and the treatment should be in collaboration with a fetal medi-

cine unit. The available treatments during pregnancy are (1) intrauterine intravascular transfusion of compatible platelets by PUBS at weekly intervals or just before delivery, (2) intravenous IgG or corticosteroids, or a combination of both, to the mother. As no randomized trials have been performed on either therapy, firm evidence of efficacy is lacking. However, weekly platelet transfusion by PUBS, although invasive and technically demanding, has shown good outcome in families with previously severely affected children. Repeated infusion of intravenous IgG to the mother remains highly controversial, as the initial report of its possible effectiveness made use of a historical control group. The costs are high and there remains a small but definite risk of transmission of infectious agents by a protein concentrate obtained from large plasma pools. The precise mechanism of action, if any, of corticosteroids administered to the mother on the severity of fetal disease is poorly understood.

Counseling

Counseling of couples with an index case about the risks of severe fetal/neonatal thrombocytopenia in a next pregnancy needs to be based on the severity of disease in the index case and the outcome of immunological investigations. The following should be taken into account:

- Thrombocytopenia in subsequent cases is as severe or generally more severe.
- Antibody specificity and titer have some correlation with severity; for example, HPA-5b antibodies generally cause mild thrombocytopenia, which rarely results in a cerebral bleed. The latter is generally associated with HPA-1a antibodies.
- A high-titer HPA antibody is more likely to be associated with severe thrombocytopenia, but cerebral bleeds also occur with low titers.

Neonatal alloimmune neutropenia

Maternal alloimmunization against neutrophil-specific alloantigens on fetal/neonatal neutrophils is rare, although there are no precise prevalence figures. Clinical presentation is one of mainly bacterial infection with a selective neutropenia on a whole blood cell count. The number of well-characterized neutrophil-specific alloantigen systems is limited.

The NA system

The most immunogenic, bi-allelic NA alloantigen system is localized on the neutrophil FcRγIIIb (CD16), one of the two low-affinity receptors (R) for the constant domain (Fc) of human IgG(γ). The difference between the two alleles is based on four amino acids with an arginine→serine, asparagine→serine, aspartate→asparagine and valine→isoleucine at positions 36, 65, 82 and 106. This polymorphism results in a shift in molecular weight between the two isoforms because of the loss of a glycosylation site in the NA1 protein (50–65 kDa, FcγRIIIb-NA1 and 65–80 kDa, FcγRIIIb-NA2). Two other FcγRIIIb-associated alloantigens have been reported. The LAN antigen, which was associated with neonatal alloimmune neutropenia in an Aboriginal family, and SH, which is based on an alanine→asparagine-78 mutation of an FcγRIIIb isoform encoded by an additional *FcRγIIIB* gene, which is present in 10% of Caucasians. The latter is associated with increased expression of the FcγRIIIb.

The FcγRIIIb null phenotype is rare and is based on a double deletion of the *FcγRIIIb* gene, and is in some associated with a deletion of the *FcγRIIc* gene. The deficiency for the most abundant Fc receptor on neutrophils does not seem to be associated with an obvious clinical phenotype. This is in contrast with a mutation in the *FcγRIIIa* gene, which encodes a leucine→histidine-48 substitution in the first extracellular domain of the NA cell FcγRIIIa which, although only described in one infant, was associated with recurrent and serious respiratory tract viral infection from birth. The FcγRIIIb null phenotype can cause immune neutropenia in the newborn due to maternal anti-FcγRIIIb isoantibodies.

PCR techniques can be used to determine the NA, SH and FcγRIIIb null genotypes, and transfectants expressing the FcγRIIIb NA allotypes are useful probes for alloantibody detection.

NB alloantigens

Of the two NB alloantigens NB1 and NB2, the former is localized on a 58- to 64-kDa secondary granule protein and is expressed as a glycosyl–phosphatidyl–inositol anchored surface membrane glycoprotein on neutrophils. The percentage of neutrophils expressing the NB antigens varies between individuals, and anti-NB alloantibodies can give a typical bimodal fluorescence profile when granulocytes from a donor with partial NB expression are used. The gene encoding the NB protein has not been cloned and NB alloantigen typing is based on the use of human immune antisera and immunofluorescence.

Alloantigens on CD11a and CD11b

The genes encoding the αL and αM subunits of the β_2 integrin or CD11a and CD11b are polymorphic. The alloantiserum Ond^a defines an arginine→threonine-766 mutation in αL (CD11a) and the Mart^a alloantiserum an arginine→histidine-61 mutation in αM (CD11b). Alloantibody formation against these two polymorphisms has been observed in transfusion

recipients, but no neonatal neutropenia due to anti-CD11a or CD11b has been reported. This is best explained by the wide tissue distribution of these proteins.

In addition to neonatal neutropenia, neutrophil-specific antibodies are implicated in non-hemolytic febrile transfusion reactions, transfusion-related acute lung injury and autoimmune neutropenia. Severe but reversible neutropenia in the newborn might require treatment with antibiotics to control bacterial infection. So far, there is minimal or no evidence that the mutations in the FcγRIIIb protein have any functional consequences, exemplified by the recently described FcγRIIIb null phenotype, which is not linked with an obvious pathological phenotype. In sharp contrast, the single amino acid mutation in the NK cell *FcγRIIIa* gene, which might be rare, might be associated with a more grave clinical condition.

Conclusions

The molecular basis of most blood cell alloantigens, including those of the HLA system, has been discovered over the last two decades. The use of molecular techniques allows their reliable definition at high resolution, replacing previous less reliable techniques which were based on the use of polyclonal antibody reagents. This new body of knowledge has already made a significant contribution to current clinical management. On the one hand, first-trimester fetal typing for blood cell-specific alloantigens can now be used in cases of severe maternal alloimmunization to prevent fetal and neonatal morbidity, and improved selection of HLA- and HPA-matched donor platelets aids the management of hemato-oncological patients. On the other hand, it is envisaged that in the near future bone marrow and solid organ transplants across apparent mismatches can progress because of the identification of so-called permissive mismatches, and better HLA matching will allow accommodation of the graft at lower levels of immune suppression. Finally, more complete understanding of the molecular rules defining a subject's immune response status should lead to the identification of high responders, in order to target these for more specific manipulation of their immune system with the aim of establishing alloantigen-specific non-responsiveness.

Further reading

Bennett PR, Le Van Kim C, Colin Y *et al.* (1993) Prenatal determination of fetal RhD type by DNA amplification. *New England Journal of Medicine*, **329**, 607–610.

Deo YM, Graziano RF, Repp R *et al.* (1997) Clinical significance of IgG Fc receptors and Fc gamma R-directed immunotherapies. *Immunology Today*, **18**, 127–135.

Gruen JR, Weissman SM. (1997) Evolving views of the major histocompatibility complex. *Blood*, **90**, 4252–4265.

Howell WM, Navarrete C. (1996) The HLA system: an update and relevance to patient–donor matching strategies in clinical transplantation. *Vox Sanguinis*, **71**, 6–12.

Huizinga TW, Roos D, von dem Borne AE. (1990) Neutrophil Fc-gamma receptors: a two-way bridge in the immune system. *Blood*, **75**, 1211–1214.

Kuijpers RW, von dem Borne AE, Kiefel V *et al.* (1992) Leucine33–proline33 substitution in human platelet glycoprotein IIIa determines HLA-DRw52a (Dw24) association of the immune response against HPA-1a (Zwa/PIA1) and HPA-1b (Zwb/PIA2). *Human Immunology*, **34**, 253–256.

Le van Kim C, Mouro I, Cherif Zahar B *et al.* (1992) Molecular cloning and primary structure of the human blood group RhD polypeptide. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 10925–10929.

Madrigal JA, Arguello R, Scott I *et al.* (1997) Molecular histocompatibility typing in unrelated donor bone marrow transplantation. *Blood Reviews*, **11**, 105–117.

Maslanka K, Yassai M, Gorski J. (1996) Molecular identification of T cells that respond in a primary bulk culture to a peptide derived from a platelet glycoprotein implicated in neonatal alloimmune thrombocytopenia. *Journal of Clinical Investigation*, **98**, 1802–1808.

Mouro I, Colin Y, Cherif Zahar B *et al.* (1993) Molecular genetic basis of the human Rhesus blood group system. *Nature Genetics*, **5**, 62–65.

Newman PJ, Valentin N. (1995) Human platelet alloantigens: recent findings, new perspectives. *Thrombosis and Haemostasis*, **74**, 234–239.

Smythe JS, Avent ND, Judson PA *et al.* (1996) Expression of RhD and RhCE gene products using retroviral transduction of K562 cells establishes the molecular basis of Rh blood group antigens. *Blood*, **87**, 2968–2973.

van Loghem JJ, Dorfmeyer H, van de Hart M *et al.* (1959) Serological and genetical studies on the platelet antigen (zw). *Vox Sanguinis*, **4**, 161–169.

Vaughan JI, Manning M, Warwick RM *et al.* (1998) Inhibition of erythroid progenitor cells by anti-Kell antibodies in fetal alloimmune anemia. *New England Journal of Medicine*, **338**, 798–803.

von dem Borne AE, Ouwehand WH. (1989) Immunology of platelet disorders. *Baillière's Clinical Haematology*, **2**, 749–781.

von dem Borne AE, de Haas M, Roos D, Homburg CH *et al.* (1995) Neutrophil antigens, from bench to bedside. *Immunological Investigation*, **24**, 245–272.

Warkentin TE, Smith JW. (1997) The alloimmune thrombocytopenic syndromes. *Transfusion Medicine Reviews*, **11**, 296–307.

Williamson LM, Hackett G, Rennie J. (1998) The natural history of fetomaternal alloimmunization to the platelet-specific antigen HPA-1a (PlA1, Zwa) as determined by antenatal screening. *Blood*, **92**, 2280–2287.

Chapter 20 Functions of blood group antigens

John R Pawloski & Marilyn J Telen

Introduction

Erythrocyte blood group antigens are, by definition, polymorphic epitopes of red cell surface structures that are recognizable by antibodies. These epitopes generally represent relatively small changes, such as the substitution of one amino acid, in molecules that are largely the same from one individual to another. After the initial discovery of ABO blood groups at the beginning of the 20th century, red cell serologists went on to discover hundreds of such polymorphic epitopes and gradually succeeded in organizing them into blood groups, that is, antigens that arise from polymorphisms of a single parent molecule. During the late 20th century, most of these blood groups were localized to carrier structures (proteins or polysaccharides). Discovery of the biochemical and genetic basis for blood group antigens has now led to exploration of the function of these proteins beyond their importance in transfusion medicine.

This chapter reviews current knowledge about erythrocyte membrane proteins that bear blood group antigens and whose functional importance has been characterized. Primary attention is paid to two of the most interesting of these proteins: the anion exchanger protein, which bears antigens of the Diego blood group system, and the Rh proteins. Functional information about other proteins is summarized briefly.

Anion-exchanger protein 1 (band 3 protein)

Anion-exchanger protein 1 (AE1) is the most prevalent integral protein of the red blood cell (RBC) membrane (~1 million copies/cell), constituting 25–30% of the protein mass. Previously referred to as band 3 protein, on the basis of its original electrophoretic migration profile, AE1 is a 95–102 (variable glycosylation) kDa glycosylated protein constructed of 911 amino acids, yielding three distinct structural and functional domains: the N-terminal 45 000 Da cytoplasmic domain, the 55 000 Da C-terminal hydrophobic transmembrane domain, and a small 28 amino acid acidic cytoplasmic C-terminus (Figure 20.1). Extracellular domains of AE1 express the Diego blood group system, composed of two sets of antithetical antigens, Diego (Di^a/Di^b) and Wright (Wr^a/Wr^b). AE1 also expresses ABH and li antigens on a single multibranched N-linked oligosaccharide attached to the fourth extracellular domain.

Transmembrane domain

The transmembrane segment of AE1 (amino acids 404–882) is responsible for the electroneutral co-transport of HCO_3^- for Cl^-, accommodating the tide of HCO_3^- generated by the activity

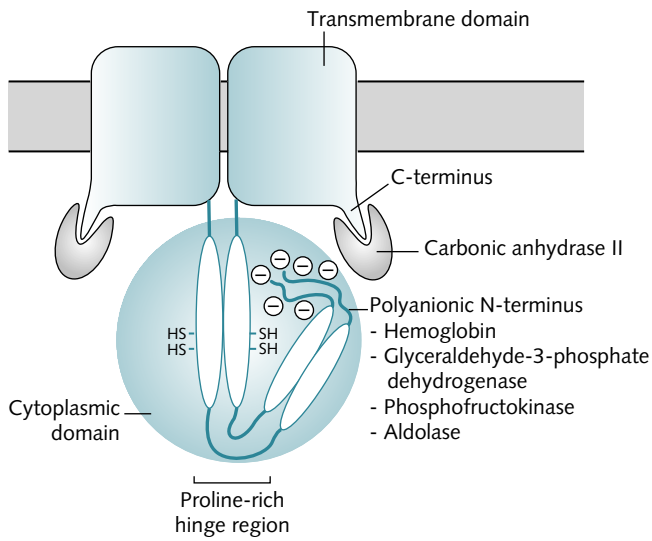


Fig. 20.1 Schematic representation of an anion-exchanger 1 protein dimer
The cytoplasmic domains (CDAE1) are enclosed by the tinted circle. CDAE1 cysteine thiols are indicated by –SH, and ⊖ indicates the negative charges of acidic amino acids in the N-terminus. Proteins known to bind specifically to the N-terminus are also listed.

of erythrocytic carbonic anhydrase on CO_2 entering the RBC, thus increasing the capacity of the blood to carry CO_2 from the tissues to the lungs. In the peripheral tissues, CO_2 taken up by RBCs is converted to H^+ and HCO_3^- by the action of carbonic anhydrase. The transmembrane domain of AE1 exports HCO_3^- in exchange for Cl^- to maintain physiological pH in the plasma, while H^+ binds to hemoglobin (Hb), right-shifting its oxygen affinity to facilitate O_2 unloading. The reverse occurs in the lungs. The anion exchange domain is not specific for HCO_3^-/Cl^- and can also transport other anions (SO_4^{2-}, $H_2PO_4^-$), anionic phospholipids, and even divalent cations. Of particular interest, AE1 also transports NO_2^-, NO_3^- and $OONO^-$, all products of nitric oxide (NO) metabolism (see below). An N-terminal truncated form of AE1 is expressed by the basolateral membrane of the α-intercalated cells in the renal collecting ducts and mediates acid secretion.

C-terminal cytoplasmic domain

The C-terminus of the transmembrane domain of AE1 concludes in a 33-residue cytoplasmic protrusion consisting predominantly of acidic amino acids and has recently been shown to possess a binding site for the carbonic anhydrase II isoform. Using glutathione-*S*-transferase (GST) fusion proteins with normal and mutated constructs, Vince and colleagues demonstrated that the critical binding site for carbonic anhydrase II is the acidic residue 887–890 region of AE1. A GST fusion protein containing the C-terminal 33 amino acids of AE2 could also bind carbonic anhydrase II, suggesting this tethering of carbonic anhydrase II to anion exchangers is a general mechanism promoting bicarbonate transport across membranes. AE2 is the most widely expressed form of AE, located in the basolateral membrane of nearly all epithelial cells.

N-terminal cytoplasmic domain

The primary function of the N-terminal cytoplasmic domain (amino acids 1–403) of AE1 (CDAE1) is to maintain erythrocyte integrity and stability by bracing the cytoskeleton to the plasma membrane. This large hydrophilic domain projects into the RBC cytoplasm with a spatial orientation that is strongly pH-dependent. Under alkaline conditions (pH 10) the appendage is fully extended, while a slightly acidic milieu (pH 6.5) induces flexion such that the N-terminus is brought close to the cell membrane. Flexibility in this part of AE1 is conferred by a proline-rich hinge region between amino acids 175 and 190. The actual orientation of CDAE1 is likely to be a semi-flexed state because the internal pH of RBCs is closer to 7.0–7.2. However, given that carbonic anhydrase II binds to the C-terminus of AE1 in a juxtamembrane location and that each AE1 channel can export 2000–3000 molecules of HCO_3^- per second, the micromilieu immediately adjacent to CDAE1 may experience much lower pHs under certain conditions, inducing greater degrees of CDAE1 flexion. These regions were once thought to be discrete operational units, but it is now clear that remote alterations in membrane domain structure can influence cytoplasmic domain activity, suggesting that these regions are functionally linked.

CDAE1 contains high-affinity binding sites for three peripheral membrane proteins that link AE1 to the underlying cytoskeletal meshwork. Ankyrin is a 206-kDa pyramidal protein that interacts with several sites throughout the length of CDAE1, hinting that ankyrin binds to a flexed conformation of CDAE1 *in vivo*. Ankyrin also possesses three structural–functional domains: an N-terminal AE1 and tubulin-binding domain, a middle spectrin-binding domain, and a C-terminal domain containing sequences that regulate ankyrin binding to AE1 and spectrin. CDAE1 also binds protein 4.2, or pallidin, a 77-kDa protein that binds to ankyrin as well, acting to reinforce the AE1–ankyrin interaction. Pallidin shows striking homology to transglutaminases but lacks enzymatic activity owing to loss of a necessary conserved active site residue. Finally, CDAE1 interacts (at least *in vitro*) with protein 4.1, a 66-kDa moiety that firmly binds spectrin near the actin-binding site, thus intensifying the spectrin–actin interaction. The spectrin-binding activity of protein 4.1 is concurrently regulated by phosphorylation/dephosphorylation, Ca^{2+}–calmodulin association, and phosphatidylinositol pathways.

The N-terminus of CDAE1 also includes specific, high-affinity binding sites for the glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphofructokinase (PFK), and aldolase. Binding to CDAE1 results in a decrease in enzymatic activity (90% for aldolase and 100% for GAPDH), while at the same time enzyme substrates, products and cofactors displace the bound enzyme, suggesting that specific CDAE1-enzyme interactions take place at the catalytic site or a linked allosteric site. Aldolase and PFK mutually inhibit each other from binding CDAE1, but neither enzyme can displace bound GAPDH, indicative of a unique binding site for GAPDH. Phosphorylation of tyrosine 8 of AE1 also regulates glycolytic enzyme binding, and thus activity. Binding affinities of the glycolytic enzymes for CDAE1 are strongly dependent on H^+ and salt concentration, decreasing dramatically as the pH and ionic strength are raised into the physiological range, calling into question the extent of binding *in vivo*. Several studies, however, support the notion of significant enzyme–CDAE1 binding in the intact erythrocyte.

Nitric oxide export

The role of nitric oxide (NO) as a critical third gas in the human respiratory cycle [in addition to oxygen (O_2) and carbon dioxide (CO_2)], is gaining rapid acceptance. NO is gener-

ated chiefly by the activity of endothelial NO synthase and is a principal vasodilator and inhibitor of platelet activation in the cardiovascular system. A portion of vascular NO penetrates circulating RBCs, where it interacts with hemoglobin (Hb) at the heme prosthetic groups, or with β-chain cysteine thiols (cysβ93) to form biologically active *S*-nitrosohemoglobin (SNO-Hb). Under low oxygen tension or oxidative conditions, SNO-Hb can release NO to effect vasodilation or inhibit platelet activation (Figure 20.2). However, effective NO export from the RBC requires that the reactive NO species be shielded from the intense scavenging activities of the heme groups. Recent investigations reveal that erythrocytes protect bioactive NO by sequestering SNO in the membrane following addition of physiological amounts of aqueous NO. Further, these 'SNO-RBCs' elicit relaxation of vascular tissue *in vitro* under tissue pO_2 conditions (5–10 mmHg), while contracting smooth muscle under normoxic arterial conditions (120 mmHg).

Compartmentalization of SNO in the erythrocyte membrane by SNO-Hb requires close interaction between the hemoprotein and the inner membrane surface. Hemoglobin binds to the cytoplasmic face of the RBC membrane via both non-specific binding and through specific protein–protein interactions. The predominant site of specific interaction between Hb and the RBC membrane is through high-affinity binding to the N-terminal cytoplasmic domain of AE1. More specifically, the site of Hb binding within CDAE1 has been localized to the polyanionic extreme N-terminus, which contains 18 acidic amino acids, thus forming an anionic milieu at physiological pH. Analysis of co-crystals of Hb and a CDAE1 N-terminal peptide has shown that this stretch of acidic residues inserts deeply into the 2,3-diphosphoglycerate binding pocket formed between β-globin subunits of tetrameric Hb. Consistent with this binding mechanism, CDAE1 binds with higher affinity to deoxyHb than to oxyHb (in which the β-cleft is occluded), and consistent with the requirements of thermodynamic linkage, the oxygen affinity of Hb is reduced in the presence of isolated CDAE1. Interestingly, the AE1 of kidney intercalated cells is a truncated isoform, lacking the N-terminal 65 residues of the CDAE1, and does not bind Hb, suggesting that these residues perform an AE1-related function specific to the RBC. Binding of Hb to CDAE1 is non-cooperative but causes a rightward shift in the oxygen dissociation curve. Simply stated, binding of Hb to CDAE1 stabilizes the 'T' or deoxy conformation of Hb. CDAE1 also binds oxidized [Fe^{3+}] or metHb more avidly than oxyHb. Because AE1 is present in insufficient quantity to bind a significant fraction of the total Hb in an RBC, the influence of AE1 on overall Hb oxygen delivery is probably negligible. This raises the possibility that AE1 may select a subpopulation of Hb molecules (e.g. SNO-Hb) to interact with, subserving an alternate function sufficiently executed by a small fraction of total Hb molecules. Specifically, SNO-Hb binding to AE1 would promote the T structure and enhance NO release. Thus, the CDAE1/SNO–Hb interaction would appear designed to accelerate NO release from SNO-Hb (Figure 20.3). In support of this assertion, sequestration of SNO in RBC membranes following treatment with NO is markedly reduced by prior treatment with AE1 inhibitors [e.g. DIDS (4,4′-diisothiocyanatostilbene-2,2′-disulfonic acid), phenylglyoxal and niflumic acid]. Further, SNO-Hb can transnitrosylate inside-out vesicles (IOV) of RBC membranes, an effect blocked in IOVs purified from RBCs treated with AE1 inhibitors. Conveniently, CDAE1 contains two cysteine residues (cys_{201} and cys_{317}) with reactive thiols that are clustered at the subunit interface of dimeric AE1. These thiols are removed by chymotrypsin (which cleaves CDAE1 from the transmembrane domain) and are surrounded by amino acids that fit an *S*-nitrosylation motif. Accordingly, transnitrosylation of RBC IOVs by SNO-Hb is blocked by thiol-modifying agents or enzymatic digestion of the IOVs with chymotrypsin. In crosslinking studies,

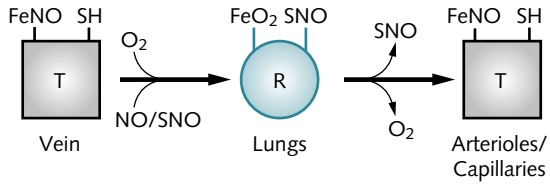


Fig. 20.2 Simplified allosteric model of NO-hemoglobin interactions

Nitric oxide (NO) binds to the heme groups of deoxygenated or T-state hemoglobin (Hb). Upon oxygenation, Hb assumes the R conformation and NO is transferred to the cysteine thiol in position 93 of the β-chain (cysβ93) to form *S*-nitrosohemoglobin (SNO-Hb). During transit through the physiological oxygen gradient, pO_2 falls, SNO-Hb converts to the T state, and NO is released from cysβ93 to evoke vasodilation (thus regulating blood flow) or to inhibit platelet activation.

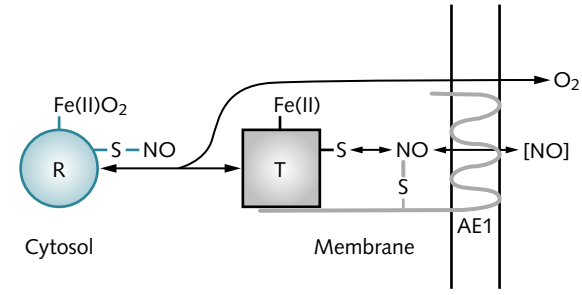


Fig. 20.3 Interaction between *S*-nitrosohemoglobin (SNO-Hb) and the cytoplasmic domain of erythrocyte anion-exchanger 1 (CDAE1)

Binding to CDAE1 stabilizes the T state of SNO-Hb and reduces NO affinity. SNO-Hb transfers NO to vicinal thiols in CDAE1, which can then release NO bioequivalents to effect vasodilation or inhibit platelet activation.

Hb forms a disulfide bond (–S–S–) between its cysβ93 thiols and the cytoplasmic domain cysteines of AE1, suggesting close spatial interplay in this region. Hemichromes, which are native Hb molecules with altered conformation near the heme, have a very high affinity for CDAE1. Moreover, hemichromes formed by the covalent modification of the cysβ93 thiol (a SNO-Hb mimetic) bind preferentially to AE1. Therefore, transnitrosylation of a vicinal thiol within the cytoplasmic domain of AE1 is a strong candidate mechanism for transfer of the NO group from cysβ93 of SNO-Hb to the RBC membrane (Fig. 20.3). In fact, like Hb, SNO-Hb forms disulfide links with AE1 under oxidative conditions, confirming the proximity of SNO-Hb and CDAE1 thiols. Further, *S*-nitrosylated AE1 can be immunoprecipitated from the membranes of RBCs treated with NO and IOVs treated with SNO-Hb. Thus, a role for the anion exchange function of AE1 in transmembrane movement of NO bioactivity is enticing. While it has been demonstrated that RBCs can release small *S*-nitrosylated thiols, it is appealing to consider that NO bioequivalents transit the RBC membrane, at least in part, by means of AE1-mediated transport of an NO congener or NO_x derived in the immediate juxtamembrane locale from CDAE1-liganded (S)NO.

Deficiencies and mutations of AE1

Clinical conditions resulting from mutations or deficiencies in AE1 range from asymptomatic to severe, with manifestations predictably reflecting perturbation of one or both of the two primary functions of AE1, HCO_3^-/Cl^- exchange and cytoskeletal membrane stability. A significant segment of the population (5–20%) are heterozygous for an AE1 polymorphism called 'band 3 Memphis', involving a lysine to glutamic acid substitution in codon 56 resulting in a functionally normal AE1.

Southeast Asian ovalocytosis (SAO), characterized genotypically as a deletion of amino acids 400–408 in AE1, results in oval RBCs with increased rigidity, decreased osmotic fragility, and reduced expression of certain red cell antigens. Natural selection dramatically increases the prevalence of SAO in areas with endemic malarial infection, given that this defect confers reduced parasitemia and disease severity. However, homozygosity for the SAO mutation appears to be incompatible with life, as homozygotes have never been observed (even in areas with 40% prevalence), and the incidence of miscarriage in heterozygous parents is twice that for normals (20 versus 9.8%).

Over 40 mutations of the AE1 protein have been found to be associated with a subset of individuals with dominant hereditary spherocytosis (HS), a condition characterized by loss of erythrocyte surface area and a mild to severe hemolysis almost invariably improved by splenectomy. Despite being scattered throughout the entire protein, many mutations in the membrane spanning domain involve substitutions of highly conserved arginine residues thought to be important in maintaining proper membrane orientation. Other mutations causing HS, including those in the cytoplasmic domain, lead to decreased binding affinities for the cytoskeletal bridging proteins ankyrin and pallidin.

Complete AE1 deletion in a murine model results in a deadly hypercoagulable state in which only 5–10% of AE1 null mice survive the neonatal period. Almost uniformly, animals were found at necropsy to have large thrombotic lesions of the heart, subcapsular liver necrosis suggesting arteriolar ischemia, and large vein thrombi. Inaba and colleagues similarly described a total deficiency of AE1 in Japanese black cattle homozygous for a nonsense mutation (CGA→TGA) in codon 646 causing an early stop signal. These animals exhibit moderate uncompensated anemia, spherocytosis, and mild acidosis at rest that is exacerbated by exercise. The only reported human case of total absence of AE1 occurred in a Portuguese female delivered by Caesarian section at 36 weeks of gestation, who was hydropic and severely anemic and had massive hepatosplenomegaly. PCR demonstrated homozygosity for the Coimbra mutation (V488M), complete absence of AE1 and protein 4.2, and reduction of spectrin and ankyrin by 43 and 57%, respectively. Spherocytes and erythroblasts dominated the peripheral blood morphology, while hyperchloremic acidosis was identified by age 3 months. The patient is being treated with an intensive transfusion regimen, chelator therapy, and oral base replacement to counteract the acidosis. Other than slightly delayed psychomotor development and stable nephrocalcinosis, her condition remains stable.

Rh antigens

The rhesus (Rh) blood group antigens constitute the most polymorphic (45 independent antigens) family of erythrocyte antigens and are also the most frequent inducers of transfusion- and parturition-related alloantibodies. Three separate proteins make up the Rh family of proteins, including RhD and RhCE, which are unglycosylated proteins of 417 amino acids with 97% homology, and Rh-associated glycoprotein (RhAG), which shows ~40% homology to RhD/CE. While scores of individual Rh antigens are recognized, only five epitopes (D, C, c, E and e) carried on the RhD/CE proteins are most routinely identified; RhAG is thought to carry only the high-frequency Duclos antigen. RhCE expresses the antithetical Cc/Ee antigens, while no antithetical antigen exists for D. However, the presence (Rh-positive) and absence (Rh-negative) of the D antigen have sometimes been designated as D and d, respectively. Thus, the eight most common haplotypes associated with the Rh system are Dce, dce, Dce,

dCe, DcE, dcE, DCE and dCE. The D antigen is a collection of conformation-dependent epitopes along the entire RhD protein. Deletion of the *RHD* gene characterizes most D-negative Caucasians, while in other populations (mainly Japanese and African blacks) the D-negative phenotype appears to arise from a variety of missense and nonsense mutations in otherwise intact *RHD* genes. The related RhAG protein, a glycosylated 409 amino acid 50-kDa unit, appears to function as a chaperone for RhD and RhCE to the RBC membrane; Rh antigens are only expressed if RhAG is also present. The comprehensive review by Avent and Reid provides additional genetic and immunological details regarding the Rh genes and proteins.

Using known amino acid sequences, hydropathy studies suggest the Rh proteins span the erythrocyte membrane 12 times, suggesting a transporter function for the Rh family. While the physiological role(s) of these proteins have remained elusive, a spate of recent studies indicates that the Rh proteins may be involved in mammalian ammonium transport. Marini and colleagues found sequence similarity between RhAG, and to a lesser extent RhCE, and the Mep1 protein in yeast and the Amt1 protein in the plant *Arabidopsis thaliana*, both of which represent NH_4^+ transporters. The same group provided further support for this hypothesis by showing that RhAG expressed in a yeast mutant with deletions in the three endogenous ammonium transporters (*mep1Δ mep2Δ mep3Δ*) could restore growth defects elicited by low-ammonium medium. However, this study has been criticized because the expressed RhAG was not glycosylated and the authors failed to show actual NH_4^+ uptake. In contrast, Westhoff and colleagues showed that *Xenopus* oocytes expressing solely human RhAG protein were able to specifically transport NH_4^+ in a saturable, non-electrogenic, and pH-dependent manner, suggesting that RhAG may function as an NH_4^+/H^+ antiporter. This analysis is complicated by the coincident findings that expression of RhCE alone did not alter NH_4^+ transport, nor did co-expression of RhAG/RhCE, suggesting that RhAG-mediated NH_4^+ transport may be silenced in human erythrocytes.

Others argue that the data suggest the Mep/Amt channels are NH_3 gas permeases rather than NH_4^+ transporters. However, since NH_3 transport by mammalian erythrocytes would appear to be physiologically irrelevant, they conclude that erythrocytic Rh proteins represent CO_2 transporters. Supporting this assertion, Soupene and colleagues found that Rh protein expression in the green alga *Chlamydomonas reinhardtii* was significantly higher under elevated CO_2 (air + 3% CO_2 for 3 h) conditions than in air alone (0.035% CO_2). Further, Rh mRNA and protein expression remained low under conditions is which methylammonium uptake was high (nitrogen-limiting, with arginine as the sole nitrogen source instead of NH_4Cl). Conversely, when Rh expression was high (high CO_2 + NH_4Cl), methylammonium uptake was undetectable. A CO_2 transport function for the Rh proteins is made more appealing by the fact that carbonic anhydrases are known to bind to the C-terminus of AE1, a protein recognized to associate with Rh in the RBC membrane. Further studies in mammalian systems will be needed to confirm this notion.

Rh_{null} disease is characterized by the complete absence of Rh proteins in the RBCs of affected subjects. The resulting erythrocytes exhibit stomatocytic and spherocytic changes, increased osmotic fragility, and deviations in ionic fluxes and cell volume. Two variations of Rh_{null} disease have been identified, termed 'amorph' and 'regulator', based on the underlying molecular defect from which they arise. The amorph variety involves a genetic change in the *RhCE* gene coupled with deletion of *RHD*, while the regulator phenotype results from two mutant *RHAG* genes (homozygote or compound heterozygote). These null phenotypes also display marked reductions in Rh-associated membrane proteins, including the LW glycoprotein (ICAM-4), integrin-associated protein (IAP) and glycophorin B (a type I membrane glycoprotein that bears the Ss antigens and may serve to couple Rh with larger complexes containing AE1 and GPA).

Glycophorins C and D

Glycophorins C and D (GPC/D) are 32- and 23-kDa membrane glycoproteins arising from the same four-exon gene (2q14-q21) via alternative in-frame mRNA translation initiation sites, giving rise to the Gerbich erythrocyte antigens, Ge:2 and Ge:3. GPC and D are identical except that D has been truncated by 21 amino acids at the N-terminus and there are two or three copies per cell of GPC for each copy of GPD. Despite being infrequent targets of alloantibodies, GPC/D may represent vital membrane-stabilizing components by serving as high-affinity binding sites for the peripheral membrane protein 4.1, which acts to affix actin and spectrin elements. In addition to AE1–ankyrin–spectrin, the GPC/D-protein 4.1-spectrin/actin bridge is believed to constitute the other major connection site between the erythrocyte membrane and its spectrin skeleton. Evidence for the importance of GPC/D in membrane/cytoskeletal stability originally arose from observations in cells containing natural defects in one or more members of this bridge. For instance, complete absence of erythrocyte GYPC/D (termed the Leach phenotype) results in a rare form of hereditary elliptocytosis characterized by impaired red cell mechanical properties. However, these same GPC/D-deficient elliptocytic cells were also found to contain reduced amounts of protein 4.1, and reconstituting these deficient cells with the spectrin–actin binding domain of protein 4.1 fully restored the membrane's mechanical properties without re-establishing the connection of protein 4.1 to GPC

or the lipid bilayer. Evidence from a recent investigation further questions the physiological significance of this putative membrane-stabilizing interaction by showing that conditions promoting the dissociation of GPC from protein 4.1 in the intact RBC have little or no impact on the mechanical properties of the cell, as assayed by ektacytometry and nickel mesh filtration. Thus, while the GPC/D–protein 4.1–spectrin/actin assembly probably occurs physiologically, its role as a regulator of membrane stability remains to be determined.

GPC/D also exist in variant forms, with preservation of protein but not Gerbich antigen expression, usually the result of an internal exon deletion or duplication. One particular variant, the Melanesian Gerbich negativity phenotype, has been shown to reach a high frequency (46.5%) in coastal areas (the Wosera) of Papua New Guinea, where *Plasmodium falciparum* malaria is hyperendemic, suggesting that it may confer protection against erythrocyte invasion by the parasite. This polymorphism is characterized genotypically by a deletion of exon 3 (thus GPCΔex3) and phenotypically by ovalocytic RBCs. Recently, Maier and colleagues offered convincing evidence that the *P. falciparum* erythrocyte-binding antigen 140 (EBA-140) binds with high affinity to GPC, and that this interaction mediates a principal invasion pathway into human erythrocytes. They showed that EBA140 does not bind to GPC in Ge-negative erythrocytes from Papua residents homozygous for *GPCΔex3*, nor can *P. falciparum* infect such cells using this invasion pathway. Interestingly, while the invasion of *GPCΔex3* erythrocytes by *P. falciparum* parasites is less efficient *in vitro*, no differences in infection rates for either *P. falciparum* or *P. vivax* in Ge-negative subjects have been previously observed. However, subjects with other RBC polymorphisms that are found in high frequencies in endemic malaria regions, notably sickle cell hemoglobin (hemoglobin S), also demonstrate infectivity rates similar to those of subjects lacking the sickle phenotype. Nevertheless, these individuals with HbS trait show reduced parasite density and experience fewer cases of cerebral or severe malaria. Thus, it may be that the GPCΔex3 phenotype is selected for by reducing erythrocyte invasion rates by malaria parasites, thus lessening disease severity. Further studies are required to determine whether there has been natural selection of the *GPCΔex3* allele as a safeguard against severe malaria.

Blood group antigen proteins that function as adhesion molecules

A number of erythrocyte membrane proteins have been identified as adhesion molecules. As shown in Table 20.1, these include the proteins that bear the Lutheran, LW and Indian blood group antigens.

CD44, which bears the Indian antigens and was first described on erythrocytes as In(Lu)-related p80, was the first to be characterized as an adhesion molecule. CD44 bears homology to the cartilage link and proteoglycan core proteins, which are known to interact with hyaluronan. This similarity led to the demonstration that CD44 also bound to this component of extracellular matrix and basement membranes. On leukocytes, CD44 appears to be involved in the interaction of leukocytes with endothelial cells, including the homing of lymphocytes to peripheral lymphoid organs and sites of inflammation. Ligand binding to CD44 can also induce cytokine release and T-cell activation. However, on erythrocytes the functional significance of CD44 remains unclear, although one report has indicated that CD44 contributes to the ability of early erythroid precursors to bind to bone marrow fibronectin. CD44 may also play an important role in the metastatic behavior of tumor cells.

The molecule that bears Lutheran blood group antigens has been identified as a laminin receptor. Lutheran antigens

Table 20.1 Adhesion molecules of erythrocytes.

Blood group antigen	Alternative name(s)	Ligand/adhesive function
In^a/In^b	CD44, In(Lu)-related p80	Hyaluronan, possibly also fibronectin
JMH	CD108, semaphorin K1 (SEMA7A)	Possible role in adhesion of activated lymphocytes
Lutheran	B-CAM/LU, CD239	Laminin receptor
LW	CD242, ICAM-4, LW	Leukocyte and endothelial integrins ($\alpha_4\beta_1$, $\alpha_4\beta_3$, $\alpha_v\beta_1$,$\alpha_v\beta_3$)
Nak^a (platelets)	CD36 (reticulocytes only), platelet glycoprotein IV, Nak^a (platelets)	Thrombospondin (platelets), LDL
Ok^a	CD147, neurothelin	Type IV collagen, fibronectin. Laminin in other tissues
None known	VLA-4 (reticulocytes only), $\alpha_4\beta_1$ integrin (CD49d/CD29)	Thrombospondin, VCAM-1, possibly fibronectin
None known	CD47, integrin-associated protein (IAP)	Thrombospondin
None known	CD58, lymphocyte associated antigen-3 (LFA-3)	CD2
RAPH1	CD151, MIC2 gene product	Proposed role in kidney and skin morphogenesis, possibly through adhesion to laminin or collagen

are expressed by two proteins which arise from alternate splicing of a single gene. These two proteins (sometimes called B-CAM and LU) are identical except for their cytoplasmic domains, and they both appear to function equally well as laminin receptors. Interestingly, B-CAM/LU expression and function are increased on red cells from patients with sickle cell disease, and this function can be activated further by exposure of sickle red cells to epinephrine and other reagents that lead to increased intracellular cAMP. However, the contribution of this activity to the pathophysiology of vaso-occlusion remains unproven.

The LW protein also appears to be an adhesion receptor, although, unlike CD44 and B-CAM/LU, LW interacts with cell surface proteins rather than extracellular matrix components. LW has also been called ICAM-4 because it is highly homologous to other members of the ICAM family of adhesion molecules. Like other ICAMs, LW can bind to the leukocyte integrins CD11a/CD18, CD11b/CD18 and CD11c/CD18. Recently, it has also been shown that LW can bind to αv integrins, and that LW may contribute to the adhesion of sickle red cells to the endothelium via $\alpha_v\beta_3$. In addition, some investigators speculate that LW is also important in facilitating the adhesion of erythroid precursors to bone marrow macrophages in erythroid islands during erythropoiesis.

Blood group antigen proteins with other functions

Table 20.2 lists blood group antigens associated with functions other than adhesion. As indicated, proteins bearing blood group antigens have a broad diversity of functions. Some, such as the proteins that bear the Kidd and Colton blood group antigens, are transporters. Others, such as those that bear the Cartwright and Kell antigens, are ectoenzymes. In addition, erythrocytes bear receptors for complement components and chemokines.

Summary

Proteins that bear blood group antigens have diverse functions, and some proteins, such as AE1, encompass several functions within a single protein molecule. Abnormalities of these proteins, in the form of either deficiencies or mutations, can lead to red cell disorders, such as hemolytic anemia, or have more far-reaching effects, as in the association of Kx deficiency and neuroacanthocytosis. Finally, these proteins undoubtedly contribute both to normal physiology and to the pathophysiology of certain diseases, most notably sickle cell anemia.

Further reading

AE1

Alda JO, Garay R. (1990) Chloride (or bicarbonate)-dependent copper uptake through the anion exchanger in human red blood cells. *American Journal of Physiology*, **259**, C570–C576.

Anderson RA, Marchesi VT. (1985) Regulation of the association of membrane skeletal protein 4.1 with glycophorin by polyphosphoinositide. *Nature*, **318**, 295–298.

Bromberg PA, Theodore J, Robin ED *et al.* (1965) Anion and hydrogen ion distribution in human blood. *Journal of Laboratory and Clinical Medicine*, **66**, 464–475.

Bruce LJ, Kay MMB, Lawrence C *et al.* (1993) Band 3 HT, a human red-cell variant associated with acanthocytosis and increased anion transport, carries the mutation pro 868 leu in the membrane domain of band 3. *Biochemical Journal*, **293**, 317–320.

Table 20.2 Diverse functions of blood group antigen proteins.

Blood group	Alternative name(s)	Function
Cartwright	ACHE	Acetylcholinesterase
Colton	Aquaporin-1 (AQP-1)	Water channel
Cromer	Decay accelerating factor, CD55	Promotes the degradation of C3 and C5 convertases
Dombrock	DOK	ADP-ribosyltransferase ectoenzyme
Duffy	DARC	Chemokine receptor
Kell	KEL	Zinc-binding neutral endopeptidase; cleaves big endothelin-3 to endothelin-3
Kidd	UT1	Urea transporter
Knops/McCoy	C3b/C4b receptor (CD35), complement receptor type 1	Binds C3b and C4b and facilitates immune clearance
Kx		Possibly a neurotransmitter transporter; deficiency causes neuroacanthocytosis or McLeod syndrome

Cassoly R, Salhany JM. (1983) Spectral and oxygen-release kinetic properties of human hemoglobin bound to the cytoplasmic fragment of band 3 protein in solution. *Biochimica et Biophysica Acta*, **745**, 134–139.

Cohen CM, Dotimas E, Korsgren C. (1993) Human erythrocyte membrane protein 4.2 (pallidin). *Seminars in Hematology*, **30**, 119–137.

Davis L, Lux SE, Bennett V. (1989) Mapping the ankyrin-binding site of the human erythrocyte anion exchanger. *Journal of Biological Chemistry*, **264**, 9665–9672.

Davis LH, Davis JQ, Bennett V. (1992) Ankyrin regulation: an alternatively spliced segment of the regulatory domain functions as an intramolecular modulator. *Journal of Biological Chemistry*, **267**, 18966–18972.

Funder J, Wieth JO. (1966) Chloride and hydrogen ion distribution between human red cells and plasma. *Acta Physiologica Scandinavica*, **68**, 234–245.

Genton B, Al-Yaman F, Mgone CS *et al.* (1995) Ovalocytosis and cerebral malaria. *Nature*, **378**, 564–565.

Hassoun H, Wang Y, Vassiliadis J *et al.* (1998) Targeted inactivation of murine band 3 (AE1) gene produces a hypercoagulable state causing widespread thrombosis *in vivo*. *Blood*, **92**, 1785–1792.

Higashi T, Richards CS, Uyeda K. (1979) The interaction of phosphofructokinase with erythrocyte membranes. *Journal of Biological Chemistry*, **254**, 9542–9550.

Inaba M, Yawata A, Koshino I *et al.* (1996) Defective anion transport and marked spherocytosis with membrane instability caused by hereditary total deficiency of red cell band 3 in cattle due to a nonsense mutation. *Journal of Clinical Investigation*, **97**, 1804–1817.

Jarolim P, Palek J, Amato D *et al.* (1991) Deletion in erythrocyte band 3 gene in malaria-resistant Southeast Asian ovalocytosis. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 11022–11026.

Jarolim P, Palek J, Rubin HL *et al.* (1992) Band 3 Tuscaloosa: Pro327→arg327 substitution in the cytoplasmic domain of erythrocyte band 3 protein associated with spherocytic hemolytic anemia and partial deficiency of protein 4.2. *Blood*, **80**, 523–529.

Jarolim P, Rubin HL, Zhai S *et al.* (1992) Band 3 Memphis: a widespread polymorphism with abnormal electrophoretic mobility of erythrocyte band 3 protein caused by the substitution AAG GAG (lys→glu) in codon 56. *Blood*, **80**, 1592–1598.

Jarolim P, Rubin HL, Brabec V *et al.* (1995) Mutations of conserved arginines in the membrane domain of erythroid band 3 protein lead to a decrease in membrane-associated band 3 and to the phenotype of hereditary spherocytosis. *Blood*, **85**, 634–640.

Jenkins JD, Madden DP, Steck TL. (1984) Association of phosphofructokinase and aldolase with the membrane of the intact erythrocyte. *Journal of Biological Chemistry*, **259**, 9374–9378.

Jennings ML. (1989) Structure and function of the red blood cell anion transport protein. *Annual Review of Biophysics and Biophysical Chemistry*, **18**, 397–430.

Jia L, Bonaventura C, Bonaventura J *et al.* (1996) *S*-nitrosohaemoglobin: a dynamic activity of blood involved in vascular control. *Nature*, **380**, 221–226.

Kindt JT, Woods A, Martin BM *et al.* (1992) Covalent alteration of the prosthetic heme of human hemoglobin by $BrCCl_3$ cross-linking of heme to cysteine residue 93. *Journal of Biological Chemistry*, **267**, 8739–8743.

Kliman HJ, Steck TL. (1980) Association of glyceraldehyde-3-phosphate dehydrogenase with the human red cell membrane. *Journal of Biological Chemistry*, **255**, 6314–6321.

Kollert-Jons A, Wagner S, Hubner S *et al.* (1993) Anion exchanger 1 in human kidney and oncocytoma differs from erythroid AE1 in its NH_2 terminus. *American Journal of Physiology*, **265**, F813–F821.

Korsgren C, Lawler J, Lambert S *et al.* (1990) Complete amino acid sequence and homologies of human erythrocyte membrane protein 4.2. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 613–617.

Ling E, Danilov YN, Cohen CM. (1988) Modulation of red cell band 4.1 function by cAMP-dependent kinase and protein kinase C phosphorylation. *Journal of Biological Chemistry*, **263**, 2209–2216.

Liu SC, Zhai S, Palek J *et al.* (1990) Molecular defect of the band 3 protein in Southeast Asian ovalocytosis. *New England Journal of Medicine*, **323**, 1558–1560.

Liu SC, Jarolim P, Rubin HL *et al.* (1994) The homozygous state for the band 3 protein mutation in Southeast Asian ovalocytosis may be lethal. *Blood*, **84**, 3590–3591.

Low PS. (1986) Structure and function of the cytoplasmic domain of band 3: center of erythrocyte membrane-peripheral protein interactions. *Biochimica et Biophysica Acta*, **864**, 145–167.

Low PS, Allen DP, Zioncheck TF *et al.* (1987) Tyrosine phosphorylation of band 3 inhibits peripheral protein binding. *Journal of Biological Chemistry*, **262**, 4592–4596.

Mallozzi C, Di Stasi AMM, Minetti M. (1997) Peroxynitrite modulates tyrosine-dependent signal transduction pathway of human erythrocyte band 3. *FASEB Journal*, **11**, 1281–1290.

McMahon TJ, Stone AE, Bonaventura J *et al.* (2000) Functional coupling of oxygen binding and vasoactivity in *S*-nitrosohemoglobin. *Journal of Biological Chemistry*, **275**, 16738–16745.

McMahon TJ, Moon RE, Luschinger BP *et al.* (2002) Nitric oxide in the human respiratory cycle. *Nature Medicine*, **8**, 711–717.

Ohanian V, Wolfe LC, John KM *et al.* (1984) Analysis of the ternary interaction of the red cell membrane skeletal proteins spectrin, actin, and 4.1. *Biochemistry*, **23**, 4416–4420.

Pawloski JR, Stamler JS. (2002) Nitric oxide in RBCs. *Transfusion*, **42**, 1603–1609.

Pawloski JR, Swaminathan RV, Stamler JS. (1998) Cell-free and erythrocyte *S*-nitrosohemoglobin inhibits human platelet aggregation. *Circulation*, **97**, 263–267.

Pawloski JR, Hess DT, Stamler JS. (2001) Export by red blood cells of nitric oxide bioactivity. *Nature*, **409**, 622–626.

Rauenbuehler PB, Cordes KA, Salhany JM. (1982) Identification of the hemoglobin binding sites on the inner surface of the erythrocyte membrane. *Biochimica et Biophysica Acta*, **692**, 361–370.

Ribeiro ML, Allosio N, Almeida H *et al.* (2000) Severe hereditary spherocytosis and distal renal tubular acidosis associated with the total absence of band 3. *Blood*, **96**, 1602–1604.

Rybicki AC, Qiu JJH, Musto S *et al.* (1993) Human erythrocyte protein 4.2 deficiency associated with hemolytic anemia and a homozygous 40 glutamic acid lysine substitution in the cytoplasmic domain of band 3 (band $3^{Montefiore}$). *Blood*, **81**, 2155–2165.

Saul A, Lamont G, Sawyer WH *et al.* (1984) Decreased membrane deformability in Melanesian ovalocytes from Papua New Guinea. *Journal of Cell Biology*, **98**, 1348–1354.

Sayare M, Fikiet M. (1981) Cross-linking of hemoglobin to the cytoplasmic surface of human erythrocyte membranes. Identification of band 3 as a site for hemoglobin binding in Cu^{2+}-*o*-phenanthroline catalyzed cross-linking. *Journal of Biological Chemistry*, **256**, 13152–13158.

Shingles R, Roh MH, McCarty RE. (1997) Direct measurement of nitrite transport across erythrocyte membrane vesicles using the fluorescent probe, 6-methoxy-*n*-(3-sulfopropyl) quinolinium. *Journal of Bioenergetics and Biomembranes*, **29**, 611–616.

Stamler JS, Jia L, Eu JP *et al.* (1997) Blood flow regulation by *S*-nitrosohemoglobin in the physiological oxygen gradient. *Science*, **276**, 2034–2037.

Stamler JS, Toone EJ, Lipton SA *et al.* (1997) (S)NO signals: translocation, regulation, and a consensus motif. *Neuron*, **18**, 691–696.

Takakuwa Y, Mohandas N. (1988) Modulation of erythrocyte membrane material properties by Ca^{2+} and calmodulin: implications for their role in regulation of skeletal protein interactions. *Journal of Clinical Investigation*, **78**, 80–85.

Thevenin BJM, Bicknese SE, Park J *et al.* (1996) Distance between cys-201 in erythrocyte band 3 and the bilayer measured by single-photon radioluminescence. *Biophysical Journal*, **71**, 2645–2655.

Tsuneshige A, Imai K, Tyuma I. (1987) The binding of hemoglobin to red cell membrane lowers its oxygen affinity. *Journal of Biochemistry*, **101**, 695–704.

Vince JW, Reithmeier RAF. (1998) Carbonic anhydrase II binds to the carboxyl terminus of human band 3, the erythrocyte Cl^-/HCO_3^- exchanger. *Journal of Biological Chemistry*, **273**, 28430–28437.

Vondenhof A, Oslender A, Deuticke B *et al.* (1994) Band 3, an accidental flippase for anionic phospholipid? *Biochemistry*, **33**, 4517–4520.

Walder JA, Chatterjee R, Steck TL *et al.* (1984) The interaction of hemoglobin with the cytoplasmic domain of band 3 of the human erythrocyte membrane. *Journal of Biological Chemistry*, **259**, 10238–10246.

Wang CC, Moriyama R, Lombardo CR *et al.* (1995) Partial characterization of the cytoplasmic domain of human kidney band 3. *Journal of Biological Chemistry*, **270**, 17892–17897.

Waugh SM, Low PS. (1985) Hemichrome binding to band 3: nucleation of Heinz bodies on the erythrocyte membrane. *Biochemistry*, **24**, 34–39.

Weaver DC, Marchesi VT. (1984) The structural basis of ankyrin function. I. Identification of two structural domains. *Journal of Biological Chemistry*, **259**, 6165–6169.

Weaver DC, Pasternack GR, Marchesi VT. (1984) The structural basis of ankyrin function. II. Identification of two functional domains. *Journal of Biological Chemistry*, **259**, 6170–6175.

Zhou J, Low PS. (2001) Characterization of the reversible conformational equilibrium in the cytoplasmic domain of human erythrocyte membrane band 3. *Journal of Biological Chemistry*, **276**, 38147–38151.

RH

Avent ND, Reid ME. (2000) The Rh blood group system: a review. *Blood*, **95**, 375–387.

Beckmann R, Smythe JS, Anstee DJ *et al.* (1998) Functional cell surface expression of band 3, the human red blood cell anion exchange protein (AE1), in K562 erythroleukemia: band 3 enhances the cell surface reactivity of Rh antigens. *Blood*, **92**, 4428–4438.

Beckmann R, Smythe JS, Anstee DJ *et al.* (2001) Coexpression of band 3 mutants and Rh polypeptides: differential effects of band 3 on the expression of the Rh complex containing D polypeptide and the Rh complex containing CcEe polypeptide. *Blood*, **97**, 2496–2505.

Marini AM, Urrestarazu A, Beauwens R *et al.* (1997) The Rh (rhesus) blood group polypeptides are related to NH_4^+ transporters. *Trends in Biochemical Sciences*, **22**, 460–461.

Soupene E, Ramirez RM, Kustu S. (2001) Evidence that fungal MEP proteins mediate diffusion of the uncharged species NH_3 across the cytoplasmic membrane. *Molecular and Cellular Biochemistry*, **21**, 5733–5741.

Soupene E, King N, Feild E *et al.* (2002) Rhesus expression in a green alga is regulated by CO_2. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 7769–7773.

Vince JW, Reithmeier RAF. (2000) Identification of the carbonic anhydrase II binding site in the Cl^-/HCO_3^- anion exchanger AE1. *Biochemistry*, **39**, 5527–5533.

Westhoff CM, Ferreri-Jacobia M *et al.* (2002) Identification of the erythrocyte Rh blood group glycoprotein as a mammalian ammonium transporter. *Journal of Biological Chemistry*, **277**, 12499–12502.

Glycophorins C and D

Chang SH, Low PS. (2001) Regulation of the glycophorin C-protein 4.1 membrane-to-skeleton bridge and evaluation of its contribution to erythrocyte membrane stability. *Journal of Biological Chemistry*, **276**, 22223–22230.

Colin Y, Le Van Kim C, Tsapis A *et al.* (1989) Human erythrocyte glycophorin C. Gene structure and rearrangement in genetic variants. *Journal of Biological Chemistry*, **264**, 3773–3780.

Gascard P, Cohen CM. (1994) Absence of high-affinity band 4.1 binding sites from membranes of glycophorin C- and D-deficient (Leach phenotype) erythrocytes. *Blood*, **83**, 1102–1108.

Hemming NJ, Anstee DJ, Mawby WJ *et al.* (1994) Localization of the protein 4.1-binding site on human erythrocyte glycophorins C and D. *Biochemical Journal*, **299**, 191–196.

Maier AG, Duraisingh MT, Reeder JC *et al.* (2003) *Plasmodium falciparum* erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. *Nature Medicine*, **9**, 87–92.

Patel SS, Mehlotra RK, Kastens W *et al.* (2001) The association of the glycophorin C exon 3 deletion with ovalocytosis and malaria susceptibility in the Wosera, Papua New Guinea. *Blood*, **98**, 3489–3491.

Pinder JC, Chung A, Reid ME *et al.* (1993) Membrane attachment sites for the membrane cytoskeletal protein 4.1 of the red blood cell. *Blood*, **82**, 3482–3488.

Takakuwa Y, Tchernia G, Rossi M *et al.* (1986) Restoration of normal membrane stability to unstable protein 4.1-deficient erythrocyte membranes by incorporation of purified protein 4.1. *Journal of Clinical Investigation*, **78**, 80–85.

Telen MJ, Le Van Kim C, Chung A *et al.* (1991) Molecular basis for elliptocytosis associated with glycophorin C and D deficiency in the Leach phenotype. *Blood*, **78**, 1603–1606.

Adhesion proteins

Aruffo A, Stamenkovic I, Melnick M *et al.* (1990) CD44 is the principal cell surface receptor for hyaluronate. *Cell*, **61**, 1303–1313.

Bailly P, Hermand P, Callebaut I *et al.* (1994) The LW blood group glycoprotein is homologous to intercellular adhesion molecules. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 5306–5310.

Bailly P, Tontti E, Hermand P *et al.* (1995) The red cell LW blood group protein is an intercellular adhesion molecule which binds to CD11/CD18 leukocyte integrins. *European Journal of Immunology*, **25**, 3316–3320.

Campbell IG, Foulkes WD, Senger G *et al.* (1994) Molecular cloning of the B-CAM cell surface glycoprotein of epithelial cancers: a novel member of the immunoglobulin superfamily. *Cancer Research*, **54**, 5761–5765.

El Nemer W, Rahuel C, Colin Y *et al.* (1997) Organization of the human LU gene and molecular basis of the Lu(a)/Lu(b) blood group polymorphism. *Blood*, **89**, 4608–4616.

Goldstein LA, Zhou DF, Picker LJ *et al.* (1989) A human lymphocyte homing receptor, the hermes antigen, is related to cartilage proteoglycan core and link proteins. *Cell*, **56**, 1063–1072.

Hines PC, Zen Q, Burney SN *et al.* (2003) Novel epinephrine and cyclic AMP mediated activation of B-CAM/LU-dependent sickle (SS) RBC adhesion. *Blood*, **101**, 3281–3287.

Lee G, Parsons SF, Mankelow TJ *et al.* (2003) Novel secreted isoform of adhesion molecule ICAM-4: potential regulator of membrane-associated ICAM-4 interactions. *Blood*, **101**, 1790–1797.

Parsons SF, Mallinson G, Holmes CH *et al.* (1995) The Lutheran blood group glycoprotein, another member of the immunoglobulin superfamily, is widely expressed in human tissues and is developmentally regulated in human liver. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 5496–5500.

Spring FA, Dalchau R, Daniels GL *et al.* (1988) The In^a and In^b blood group antigens are located on a glycoprotein of 80,000 MW (the CDw44 glycoprotein) whose expression is influenced by the In(Lu) gene. *Immunology*, **64**, 37–43.

Spring FA, Parsons SF, Ortlepp S *et al.* (2001) Intercellular adhesion molecule-4 binds alpha(4)beta(1) and alpha(V)-family integrins through novel integrin-binding mechanisms. *Blood*, **98**, 458–466.

Telen MJ, Eisenbarth GS, Haynes BF. (1983) Human erythrocyte antigens. Regulation of expression of a novel erythrocyte surface antigen by the inhibitor Lutheran In(Lu) gene. *Journal of Clinical Investigation*, **71**, 1878–1886.

Telen MJ, Udani M, Washington MK *et al.* (1996) A blood group-related polymorphism of CD44 abolishes a hyaluronan-binding consensus sequence without preventing hyaluronan binding. *Journal of Biological Chemistry*, **271**, 7147–7153.

Udani M, Zen Q, Cottman M *et al.* (1998) Basal cell adhesion molecule/lutheran protein: the receptor critical for sickle cell adhesion to laminin. *Journal of Clinical Investigation*, **101**, 2550–2558.

Verfaillie CM, Benis A, Iida J *et al.* (1994) Adhesion of committed human hematopoietic progenitors to synthetic peptides from the C-terminal heparin-binding domain of fibronectin: cooperation between the integrin alpha 4 beta 1 and the CD44 adhesion receptor. *Blood*, **84**, 1802–1811.

Zen Q, Cottman M, Truskey G *et al.* (1999) Critical factors in basal cell adhesion molecule/Lutheran-mediated adhesion to laminin. *Journal of Biological Chemistry*, **274**, 728–734.

Zennadi R, Hines PM, De Castro LM *et al.* (2004) Epinephrine acts via erythroid signaling pathways to activate sickle cell adhesion to endothelium via LW-$\alpha_v\beta_3$ interactions. *Blood*, in press.

Chapter 21 Autoimmune hematological disorders

Drew Provan, James B Bussel & Adrian C Newland

Introduction

Autoimmune diseases are disorders in which there are antibodies or cells that react against self antigens causing disease, when there is an adaptive immune response mounted against a self antigen or antigens. This results in clearance of the antigen from the body. The normal adaptive response results in complete removal of the non-self antigen, whereas in autoimmune disease there is incomplete clearance of the antigen, which leads to perpetuation of the immune response. Autoimmune disorders occur in about 5% of the population, although many individuals have no symptoms. In all, there are more than 70 different disorders, most of which are uncommon, apart from rheumatoid disease and autoimmune thyroiditis. Autoimmune diseases are clinical syndromes mediated by activation of T or B lymphocytes or both, in the absence of infection or other discernible cause. Until recently, although we could describe the pathological features of autoimmune disease, we had little idea as to their actual cause. Through the development of animal models and the identification of target genes, we have gained considerable insight into the pathogenetic basis of these complex diseases. Autoreactive cells may affect virtually any body tissue, including blood, and blood disorders in which autoantibodies are found include cytopenias [autoimmune hemolytic anemia, idiopathic thrombocytopenic purpura (ITP) and autoimmune neutropenia], in addition to coagulation disorders such as acquired hemophilia.

Although autoimmune disease is clinically and pathologically diverse, the common end result is damage to antigen, which may result in disease. Factors that play a role in this process include immune dysregulation, genetic factors and triggering events. We will discuss all of these after briefly reviewing the structure and function of the immune system.

The immune system

The immune system comprises cells and molecules whose main role is defense against invading pathogens. The two principal components are the *innate* immune system, comprising skin, mucous membranes, neutrophils, macrophages and other scavenging cells, in addition to the complement system and natural killer cells; and the *adaptive* immune system, which involves B and T lymphocytes in addition to antigen-presenting cells (APCs) (Figure 21.1). The B cells are responsible for secreting antibody, aided by T cells. Key features of the adaptive system include antigen receptor diversity, antigen specificity and immunological memory. This is in sharp contrast to the innate system, which lacks these features.

The innate immune system

Despite varied challenges by many antigens, because the innate system lacks the ability to develop immunological memory the responses remain the same throughout life. In evolutionary terms, the innate system probably developed before the adaptive system. In the adaptive system, the cells responsible possess surface receptors for antibody (immunoglobulin, Ig) and complement. Microorganisms opsonized by antibody or complement are recognized by these receptors, phagocytosed and broken down within the interior of the phagocyte. Within cells such as neutrophils, killing and digestion of the pathogen involves the generation of superoxide and hydroxyl radicals, nitric oxide and proteolytic enzymes. In addition to the removal of pathogens, the innate system also plays a role in the removal of dead cells and remodeling of tissue during healing. Cells undergoing programmed cell death (apoptosis) express molecules such as phosphatidyl serine on their surface, targeting their removal. Dendritic cells also play a key role in innate

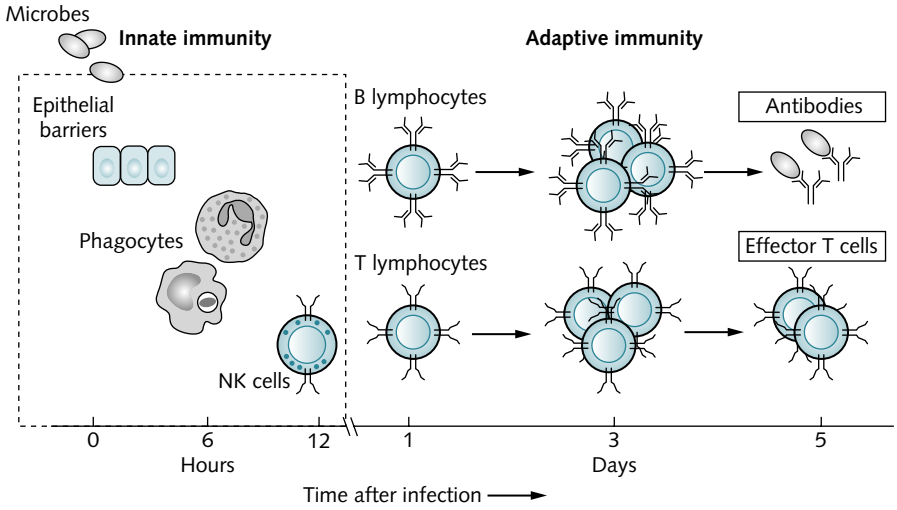


Fig. 21.1 Innate and adaptive immune systems
The innate system comprises physical barriers (e.g. skin) along with scavenger cells, whilst the adaptive system is made up of B and T lymphocytes. Temporally, the innate system is the immediate first line of defence but lacks specificity; the adaptive system comes into play later and possesses immunological specificity and memory. Reprinted from Abbas AK, Lichtman AH, Pober JS. (2000) *Cellular and Molecular Immunology*, 4th edition. Copyright 2000, with permission from Elsevier.

immunity, and activation of dendritic cells occurs following exposure to heat-shock proteins, interferon alpha (IFN-α) and other stimuli. Dendritic cells are professional APCs that are able to migrate to lymph nodes, process antigen and present this to T cells in conjunction with major histocompatibility molecules, of which there are two classes, MHC (major histocompatibility complex) class I and II.

The adaptive immune system

Two requirements of an effective immune system are (1) the ability to recognize millions of potential antigens and (2) the prevention of self-reacting lymphocytes from causing tissue damage. The former is achieved through irreversible somatic recombination of Ig and T-cell receptor (TCR) genes, generating many millions of different antibody and TCR molecules.

B and T cells possess antigen receptors on their surface

The B-cell antigen receptor is an Ig and that of T cells is the TCR. These molecules are expressed on their respective cell surfaces and interact with antigen, either as native (Ig) or processed (TCR) antigens. The TCR is a transmembrane protein and consists of a heterodimer of either α:β or γ:δ subunits. TCRs, like Igs, contain hypervariable regions and in evolutionary terms the two receptors are probably related. Unlike TCRs, antibody molecules are both expressed on the B-cell surface and secreted into body fluids. One feature that both receptors share is the ability to generate enormous diversity through irreversible recombination of germline variable (V), diversity (D) and joining (J) region segments in addition to random mutations within the rearranged genes. Immunoglobulin molecules possesses two key regions: the hypervariable region (antigen binding) and the Fc portion at the C terminal end, which, as outlined above, is recognized by the Fc receptor (e.g. FcγR) on macrophages. Antibodies may be of the IgG, IgA, IgM, IgD or IgE class, with subclasses within some of the groups (e.g. there are four IgG subclasses and two IgA subclasses). Having such enormous diversity ensures that there is an antibody for every potential antigen, but the downside of this extreme diversity is that antibodies are generated that recognize self-antigens (*autoantibodies*), and it is likely that in normal healthy subjects autoantibodies are generated against a wide variety of antigenic targets. Since autoimmune disease is not common, there must exist a mechanism for removing self-reacting antibodies. In effect, an immunological lack of responsiveness or *tolerance* must exist, whereby self-reactive cells are prevented from causing damage. Recent research has shown this to be the case.

The MHC

Class I MHC molecules comprise human leukocyte antigens (HLA)-A, -B and -C, and class II molecules consist of HLA-DP, -DQ and -DR. Class II molecules are responsible for presentation of antigen to the TCR on helper T cells.

Natural killer (NK) cells have receptors for the Ig Fc region and are responsible for antibody-dependent cellular cytotoxicity following FcγR linkage of NK cells and antibody-opsonized targets. In addition, NK cells can effect killing using killer-activating receptors, which recognize specific molecules on nucleated cells. An inhibitory molecule (killer inhibitory) recognizes MHC class I on nucleated cells, preventing killing, but if MHC class I is lost (for example, during infection of the cell by virus or after malignant transformation) the nucleated cell is recognized as being abnormal and is therefore targeted for destruction.

Soluble molecules: cytokines orchestrate the immune response

The innate system relies on a complex system of soluble molecules such as cytokines and complement components, which coordinate the entire immune response. We will discuss these briefly here since they are implicated in the pathogenesis of autoimmune disease.

Cytokines are mediators secreted by one cell that influence the behavior of other cells. Most cytokines are soluble, apart from interleukin (IL)-1 and TNF-α, which have membrane-bound forms. Cytokines are low molecular weight molecules of around 15–25 kDa whose actions include promotion of cell growth, inflammation, immunity and repair of tissues. These molecules are responsible for the regulation and orchestration of the entire immune response. Their effects are short-lived and their actions are largely local. In order to exert their effect, cytokines interact with specific receptors and promote signal transduction. Their main routes of action are via the Janus kinase (JAK)-STAT and Ras-MAP pathways. The cytokine profile may be pro- or anti-inflammatory and the cytokine balance will dictate whether a helper T-cell clone engages in a Th1- or Th2-type response. In general, Th1 responses are effective against intracellular pathogens and Th2 responses aid B cells.

Cytokine profiles of Th1 and Th2 responses

(Table 21.1)

Antigen-presenting cells, and in particular dendritic cells, are responsible for T-cell differentiation towards the Th1 or Th2 phenotype, and the cytokine IL-12 plays a key role in the Th1 response, IL-4 effecting the Th2 response.

Table 21.1 Cytokine profiles of Th1 and Th2 responses.

Th1 response (activates macrophages)	Th2 response (deactivates macrophages)
IL-2	–
IL-3	–
–	IL-4
–	IL-5
–	IL-6
–	IL-10
–	IL-13
TNF-α	–
TNF-β	TNF-β
IFN-γ	–
GM-CSF	GM-CSF

The table shows the principal cytokines involved in generation of Th1 and Th2 responses. Imbalance in Th1 or Th2 cytokines may play a role in the development of autoimmune disease, allergy and other disorders.

Since cytokines play such a key role in orchestrating an effective immune response, it is likely that dysregulation of cytokine levels may induce an autoimmune response in some disorders. This has been shown to be the case in experimental models and also in human disease. For example, transfection of the gene for IFN-γ on the insulin promoter has been shown to induce inflammation within the pancreas, with aberrant expression of MHC class II and the development of diabetes. In addition, proinflammatory cytokines, such as IL-12, TNF and IFN-γ, can induce organ-specific autoimmunity. Because of space limitations complement will not be discussed in detail.

Following infection, cells of the immune system migrate towards the affected site. Complement component C3b coats the pathogen surface. The molecules C3b, C3a, C4a and C5a, in addition to neutrophil chemoattractant, trigger mast cells to release histamine. This induces smooth muscle contraction and increased blood vessel permeability, allowing neutrophils to pass through the blood vessel walls more easily, an essential requirement for an effective innate response.

T cells

(Figure 21.2)

These develop within the thymus. T cells bearing α:β recognize processed antigen presented to them by APCs including dendritic cells. Within the thymus, T cells are selected in order to prevent autoreactivity; that is to say, mechanisms exist whereby T cells are prevented from reacting with self antigens. Positive selection occurs when T cells are able to express TCRs that are able to interact with MHC complexes on thymic cortex epithelial cells. The effect of positive selection is to prevent apoptosis. T cells expressing TCRs with high or low affinity for self molecules are negatively selected and undergo apoptosis.

Immunological tolerance prevents damage to self antigens

Tolerance defines the body's ability to recognize, but not react with, self antigens, while retaining the ability to respond to non-self antigens. This process involves the negative selection of T cells, as outlined above. In addition, the process involves the control of autoreactive B cells. The process of selecting T cells and B cells in the thymus and bone marrow respectively is known as *central tolerance*. For autoreactive lymphocytes that escape into the periphery there are additional *peripheral tolerance* mechanisms to provide a safety net for unwanted autoreactivity.

T-cell tolerance: central mechanisms

Immature T cells from the bone marrow migrate to the

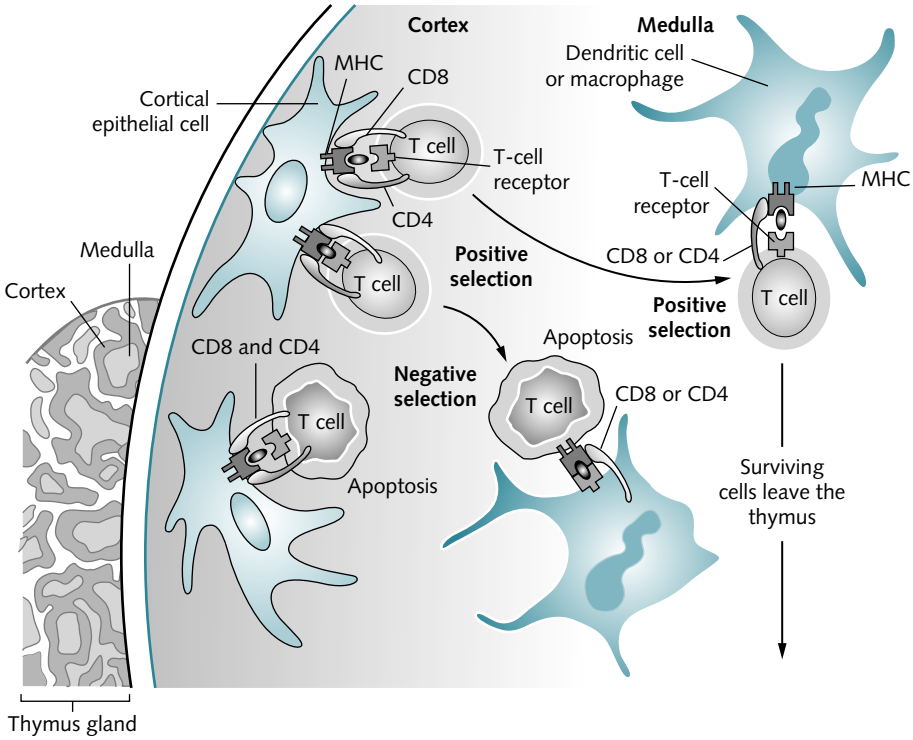


Fig. 21.2 T-cell selection within the thymus
T cells undergo positive and negative selection. $CD4^+CD8^+$ T cells interact with MHC–peptide complexes. Depending on the strength of the interaction, the T cells either undergo apoptosis (the majority) or survive and leave the thymus (the minority). Reproduced with permission from the *New England Journal of Medicine*.

thymus, where they complete their development. The T cells within the thymus interact with MHC molecules in low- or high-affinity interactions. If the TCRs have low affinity for the peptide (e.g. self peptide) they receive apoptotic signals and die within the thymus. T cells participating in high-affinity interactions have a similar fate and it is only when the interaction is of intermediate affinity that the T cells survive and migrate to the periphery, a process termed *positive selection*. In general, positive selection occurs when $CD4^+CD8^+$ T cells interact with TCR–MHC–peptide complexes. For most T cells the interaction is of low avidity and the T cells die before leaving the thymus. A minority of $CD4^+CD8^+$ T cells have intermediate avidity reactions and hence these T cells survive, after which they mature into $CD4^+CD8^-$ and $CD4^-CD8^+$ cells. $CD4^+$ T cells are the main effectors of autoimmune disease.

T-cell tolerance: peripheral mechanisms

From studies of animal models it is apparent that self-reactive lymphocytes are present peripherally, having escaped central tolerance checkpoints. Furthermore, some form of immunological ignorance is invoked whereby such circulating autoreactive T cells fail to respond to the specific self antigen. The mechanisms for such ignorance are not fully characterized but may involve low levels of circulating antigen (i.e. below a critical threshold), or the antigen may be in a separate compartment from the autoreactive cells (e.g. the blood–brain barrier), or there may be an absence of the costimulatory signals required for T-cell costimulation and activation.

T-cell activation: role of costimulatory molecules

Antigen reaches lymphoid tissue via the lymphatic system or, in some cases, within dendritic cells that have ingested and processed antigen. In general, antigen in blood is taken to the spleen and tissue antigen is taken to the lymph nodes. Antigen is then processed and presented on MHC molecules through two major routes. In one route, the antigen may be produced within the cell itself; for example, viral antigens may be expressed within cells and complexed with MHC class I molecules. Alternatively, APCs may take up antigen exogenously and they then phagocytose, process and express the antigen on MHC class II molecules.

Recognition of antigen by T cells

Antigen recognition differs between $CD4^+$ and $CD8^+$ T cells. For example, $CD4^+$ cells can only recognize antigen on class II molecules and $CD8^+$ cells recognize antigen in association with MHC class I molecules. In this way the MHC dictates the type of response generated.

TCR signaling

TCR molecules are found on the T-cell surface complexed with CD3 molecules. CD3 transmits signals intracellularly, with tyrosine phosphorylation of residues in the CD3 cytoplasmic tail. This transmits signals to the nucleus, resulting in T-cell proliferation. Co-receptor molecules on T cells are important since they play a major role in T-cell activation following engagement of the TCR. The generation of a specific immune response to a pathogen requires the recognition of the foreign agent by specialized cells in the body, presentation of the foreign antigen to T cells, and the orchestration of the subsequent immune response. T cells must receive two signals from the APCs for a complete immune response to occur. One signal involves the presentation of antigens in context of the MHC on the surface of the APC to the TCR complex on the T cells. This interaction provides the specificity of the immune response. This signal alone is not sufficient to trigger an optimal immune response and a second signal is required. The second, or costimulatory, signal involves the interaction of B7 on the APC with CD28 receptors on the T cells. The interaction of B7 with CD28 is required for the T cells to proliferate, produce lymphokines, and provide help for the induction of antibody- and cell-mediated immune responses directed at the original antigen. The B7-induced CD28 signal is delivered to the T-cell nucleus via several intermediate steps and results in the direct induction of the IL-2 gene and the receptor for this lymphokine. T cells stimulated via their TCR (antigen-specific signal) and the costimulatory signal proliferate, produce many lymphokines and direct the specific immune response against the presented antigen.

The principal co-receptors on APCs are CD80 (B7-1), CD86 (B7–2) and CD40, binding respectively to CD28, cytotoxic T-lymphocyte antigen 4 (CTLA-4) and CD40 ligand on T cells. Dendritic cells express large quantities of both B7 and CD40, are professional APCs and are the most effective stimulators of naive T cells. Following activation, the T cells undergo clonal expansion generating effector cells which move towards the site of inflammation after leaving the lymphoid tissue (Figure 21.3).

Mechanisms for controlling autoreactive T cells

Harmful autoreactive T cells may be eliminated by several mechanisms, including apoptosis, anergy, inhibition, clonal deletion and suppression. Anergy refers to T cells that do not produce IL-2 after their encounter with antigen. Such cells are not activated and may produce IL-10 which further suppresses T-cell activation. Inhibition of T cells is mediated by CTLA-4, also known as CD152. This molecule binds to the B7 family of receptors (B7-1 and B7-2) on APCs, including B cells and dendritic cells, with higher affinity than to CD28, a counter-receptor for B7. CTLA-4 mediates suppression of the T cell.

Loss of immunological tolerance leads to autoimmunity

In order for autoimmune disease to occur, there must be loss of tolerance to self antigens. This occurs despite central and peripheral T-cell tolerance controls. Possible mechanisms leading to loss of tolerance include variability in intrathymic T-cell deletion and peripheral activation of harmful T cells in the periphery; for example, by infection. The latter has been postulated to cause type I diabetes mellitus (insulin-dependent diabetes mellitus, IDDM) in some individuals (Figure 21.4).

Beneficial effects of autoimmunity

The concept of clonal deletion has been central to immunology for many years. This process deletes self-reactive lymphocytes in order to prevent the harmful effects that these might otherwise cause. We now know that low-level autoreactivity is found in normal healthy individuals and autoantigens play

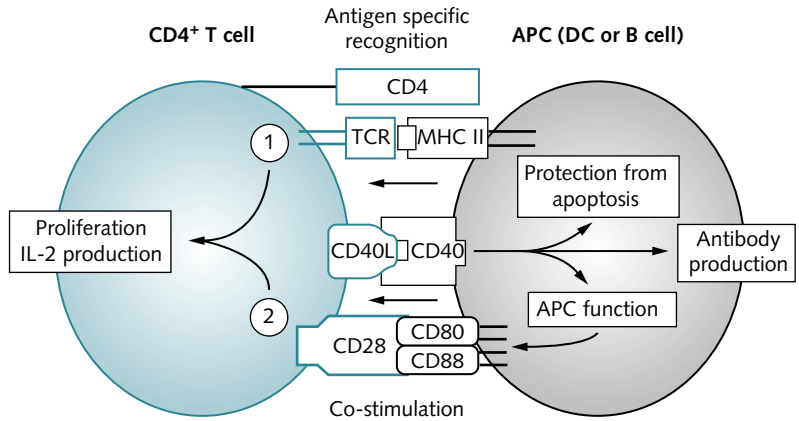


Fig. 21.3 Interaction between APCs, e.g. dendritic cells or B cells and CD4+ T cells
The first signal to the T cell is via MHC T-cell receptor. This signal is not sufficient for proliferation and IL-2 production, and a second signal, such as B7-1 or B7-2-CD28, is required. If no second signal is received the T cells undergo apoptosis. Figure kindly supplied by Dr Masataka Kuwana, Institute for Advanced Medical Research, Keio University School of Medicine, Tokyo.

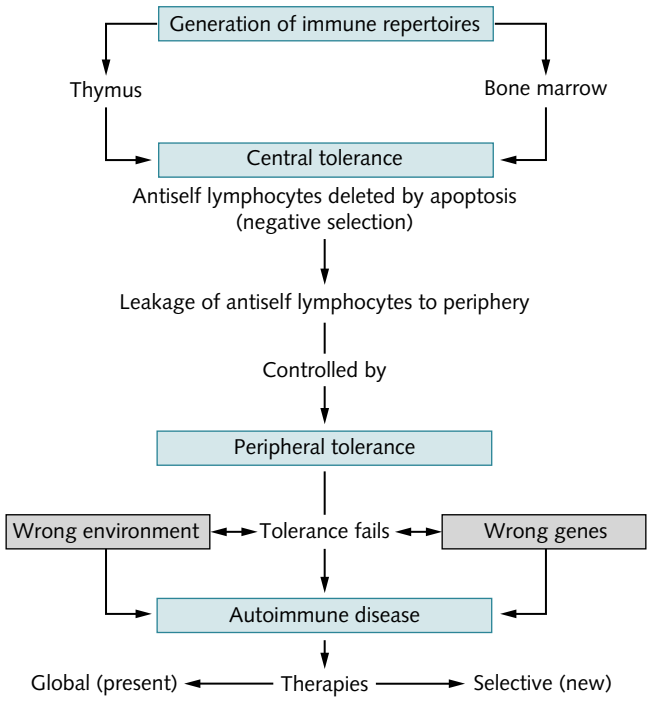


Fig. 21.4 Autoimmunity is multifactorial
Central tolerance is achieved within the bone marrow or thymus but some self-reacting cells leak out into the periphery, where peripheral tolerance checkpoints exist. Self-reactive lymphocytes escaping central *and* peripheral tolerance may cause autoimmune disease in predisposed individuals (the 'wrong' genes) when they encounter a specific trigger (the 'wrong' environment). Reproduced with permission of BMJ Publications.

a role in generating our normal immune repertoire. Self and non-self antigens have few differences and lymphocytes have probably not evolved in order to distinguish between the two; their role is rather to respond to antigen under specific circumstances; for example, as directed by the cytokine network. In addition, autoreactive cells are also responsible for the remodeling of tissues, wound healing and other physiological processes. *Autoimmunity* therefore differs from *autoimmune disease*. In fact, transient autoimmune attack of self antigens has been shown to occur when there is tissue damage, but this type of autoimmune response is generally not sustained.

The spectrum of autoimmune diseases

For clinical convenience, autoimmune diseases are subdivided arbitrarily into those that are *organ-specific* (e.g. Hashimoto's thyroiditis) and those that are *non-organ-specific* (e.g. systemic lupus erythematosus, SLE). Many diseases lie between these two extremes. Conditions in which there are circulating immune complexes tend to be systemic, while conditions associated with autoantibodies or autoreactive T-cell responses are organ-specific. However, in general this classification tells us nothing about the causes of disease and it may be better to classify them into those diseases in which there is disturbed selection, regulation or death of T and B cells (mediated by Fas or Fas receptor abnormalities, for example); a second group might be those in which there is abnormal expression of an antigen; for example, the demyelination syndrome that occurs following gut infection with the bacterium *Campylobacter jejuni*. This method of classification may help guide treatment.

Factors that may play a role in the development of autoimmune disease

These include
- failure of tolerance to self antigens
- infection
- tissue injury
- abnormalities of APCs
- imbalance between pro- and anti-inflammatory cytokines
- genetic factors
- variable effector mechanisms, e.g. production of immune complexes, autoantibodies or autoreactive T cells.

Role of genetic factors

It has long been recognized that genetic factors play a key role in autoimmune disease, the strongest correlation being with MHC genes, especially MHC class II. Support for a genetic basis comes from a variety of observations. For example, autoimmune diseases often cluster within families and twin studies show that there is much higher concordance between monozygotic than dizygotic twins: in monozygotic twins the chance of both twins having autoimmune disease is higher than in dizygotic twins. Examples of diseases with high concordance rates between monozygotic twins include type I diabetes, rheumatoid arthritis and SLE; the concordance rate for monozygotic twins is around 20% and for dizygotic twins it is less than 5%. Some disorders are due to mutations in a single gene; for example, autoimmune polyglandular endocrinopathy with candidiasis and ectodermal dysplasia (APECED) and the autoimmune lymphoproliferative disorder. In APECED there is a mutation in the gene for an autoimmune regulator protein which is active within the thymus. The mutation leads to autoimmunity and immune deficiency. In the autoimmune lymphoproliferative disorder the autoreactivity is due to an inability to induce apoptosis of activated lymphocytes following encounter with antigen. The disorder is autosomal dominant and involves Fas or its receptor, both of which are

involved in the downregulation of activated cells following antigen exposure.

Most autoimmune diseases, however, are not caused by single gene mutations but rather they are multigenic, several genes acting together to cause the disease phenotype. IDDM is a prime example of a disease in which multiple genes are implicated. For example, 95% of Caucasians with IDDM possess HLA-DR3, -DR4 or both (only found in 40% of normal subjects), and 40–50% of patients are heterozygous for -DR3/-DR4 (found in 5% of normal subjects). Other non-HLA genes are also implicated in IDDM, including IL-2 polymorphism and another region that maps close to *CTLA-4*.

Surprisingly, some of these genetic polymorphisms occur in normal individuals who have normal immune function with no evidence of autoimmune disease. So it would appear that there is a need for specific genetic polymorphisms in conjunction with other susceptibility polymorphisms. The MHC is one candidate for this role.

Most autoimmune diseases described to date have linkage with specific MHC class I or II polymorphisms, but again these do not produce the disease phenotype in isolation; rather they require the presence of polymorphisms within other genes, such as those for cytokines (e.g. TNF-α). For diseases such as ankylosing spondylitis, insulin-dependent diabetes and rheumatoid disease, certain HLA subtypes induce disease susceptibility. The link between MHC genotype and disease susceptibility is logical given the fact that autoimmune diseases involve autoreactive T cells. The ability of a T cell to respond to a particular antigen depends largely on the MHC since the MHC determines how the antigen is presented to the autoreactive T cells. In addition, the MHC has a powerful influence on determining the body's T-cell repertoire. Other HLA alleles appear to be able to offer *protection* from autoimmune disease. For example, when the disease-causing HLA allele HLA-DQB1*0301 or 0302 is present together with HLA-DQB1*0602, the latter appears to offer protection from disease. The mechanism underlying such protection is not known. Some HLA alleles cause disease only within specific populations. For example, HLA-DRB1*0401 and *0402 are associated with rheumatoid disease in Europeans but not in Blacks or Hispanic individuals. Again, the mechanism here is not clearly understood.

Some autoimmune diseases have strong association with cytokine gene polymorphisms

Since cytokines, through their interaction with their ligands, orchestrate the immune response, it is possible that dysregulation of cytokines may play a role in autoimmune disease. Indeed, many single-nucleotide polymorphisms (SNPs) have been described within cytokine or cytokine receptor genes and such SNPs may alter cytokine structure or may alter their expression. This may result in over- or underexpression. In the blood disorder ITP and other autoimmune diseases, abnormal cytokine profiles have been described. In chronic ITP, IL-2, IFN-γ and IL-10 levels are increased, suggesting a Th1 activation profile (Table 21.2).

Table 21.2 Autoimmune diseases in which strong associations with cytokine single-nucleotide polymorphisms have been described.

Disease	Implicated cytokine polymorphism
Juvenile chronic arthritis	IL-6
Myasthenia and multiple sclerosis	TGF-β
Rheumatoid disease	IL-10

Associations have now been reported for many autoimmune disorders and a full database is provided at *http://www.bris.ac.uk/pathandmicro/services/GAI/cytokine4.htm.*

Not only do cytokine gene SNPs predispose to disease, but data from animal models such as the rheumatoid rat indicate that specific polymorphisms may determine disease chronicity and severity. In addition, there are data to suggest that specific genetic polymorphisms may correlate with responses to specific treatments, but it is probably too early to draw firm conclusions from such studies.

Antibody Fc receptor gene polymorphisms

For autoimmune blood diseases in which autoantibody-opsonized cells are sequestered and destroyed within the reticuloendothelial system there are data from a number of studies showing that polymorphisms within the Fcγ receptors (FcγRs) may influence disease. Antiplatelet antibodies, as with other antibodies, possess two distinct functional components: namely a Fab end, which binds to the targeted antigen, and an Fc end, which is recognized by scavenger cells (e.g. macrophages) in the spleen, liver and bone marrow. The receptors on the macrophages that sense the presence of antibodies are termed Fc receptors (FcR). How well the FcRs recognize and bind antibody attached to platelets determines, in part, how aggressively the platelets are destroyed.

Polymorphisms of the *FcγRIIA* gene have been implicated in heparin-induced thrombocytopenia, SLE and childhood recurrent bacterial infections.

Intravenous Ig (IVIg) is effective in ITP, although its mechanisms of action are not fully understood. However, one of the actions of IVIg is to attach to FcRs on the macrophages, thereby blocking the receptor. Once the receptor is blocked it cannot

bind antibody bound to platelets and the antibody-coated platelets are spared destruction by the immune system.

Recent reports have highlighted genetic variations in FcR genes which alter their ability to recognize and bind antibody. The genetics of the FcR genes is complex, but essentially there are three major types: FcγRIA, FcγRIIA and FcγRIIIA. A recent study of FcγRIIA suggests that a polymorphism in the *FcγRIIA* gene may be responsible for causing refractory ITP (ITP that responds poorly to treatment). The polymorphism influences the efficiency of the receptor to bind with antibody molecules. A similar polymorphism in the *FcγRIIA* gene has been shown to alter the function of the receptor and may predispose individuals to autoimmune disease such as ITP, and in those individuals developing ITP the disease is more likely to be chronic. Human *FcγRIIA*-transgenic mice have been shown to have a more severe form of ITP than normal mice, which provides further evidence that FcγRIIA contributes to platelet destruction.

Mouse models of autoimmune disease

These have provided much insight into autoimmune disease and to date around 25 genes have been shown to play a role in the development of disease, either when the genes are deleted or overexpressed. The genes encode cytokines or their receptors, costimulatory molecules or proteins involved in apoptotic pathways. Whether a mutation within one of these critical genes induces disease depends on the genetic background of the animal, which would tend to imply that other genes can influence the phenotype. One other point worth noting is that mutations within a gene may be involved in more than one clinical disorder.

Human studies

A number of candidate genes have been studied in human autoimmune disease, including variants of CTLA-4. This molecule is a natural downregulator of activated T cells. One polymorphism within this gene causes a reduction in the inhibitory signal normally induced by the CTLA-4 molecule and this polymorphism has been found to be associated with human IDDM, thyroid disease and primary biliary cirrhosis.

Critical events in the generation of autoreactivity

Since autoimmune diseases are multifactorial, there are obviously triggering events within genetically predisposed individuals that lead to the development of autoimmune disease. Animal models show that if animals genetically susceptible to autoimmune disease are injected with self antigens from genetically identical animals with an appropriate adjuvant (e.g. bacterial), the animal will mount an immune response against the self antigen. In humans autoimmune disease usually arises spontaneously although it is accepted that there must be triggering events giving rise to the disease phenotype. What factors might be responsible for inducing autoimmune disease? Several have been proposed, including infectious and non-infectious triggers.

Environmental factors in the development of autoimmune disease

If the development of autoimmune disease were entirely genetic then we would expect complete or near-complete concordance in monozygotic twins. However, the concordance rate for the second twin if the first twin has autoimmune disease is less than 50%, which suggests that additional, possibly external, factors are required for autoimmune disease. Possible external influences might include infection. Evidence for a role for infection in the development of autoimmune disease is provided by disorders such as Lyme disease and its associated arthritis. In this condition there is cross-reactivity between the pathogen and host tissue antigens. The self protein targeted is leukocyte function-associated antigen 1 (LFA-1; also called CD11a and CD18). LFA-1 shares determinants with protein antigens of *Borrelia burgdorferi*. Another example is Epstein–Barr virus (EBV), which appears to have a role in the development of SLE, from case–control studies using stored serum from patients with SLE. Whether significant or not, there is one antigenic determinant on EBV that is shared with SLE. However, this evidence is fairly circumstantial at present.

How might infection cause autoimmunity?

Pathogens may be able to induce autoimmunity through a variety of mechanisms:

- production of local inflammation
- production of neoantigens
- molecular mimicry.

Local inflammation can expose costimulatory molecules on APCs and lead to the breakdown of T-cell anergy and the development of autoimmune disease. Tissue injury may result in the production of neoantigens for which an autoantibody may have specificity. Lastly, in molecular mimicry, antigens on microorganisms may resemble those on the host tissues such that antibodies produced against the pathogen will cross-react with the host tissue. Examples may include multiple sclerosis, IDDM and childhood acute ITP, if the antibody produced in a childhood viral infection by chance

cross-reacts with antigen(s) on the platelet surface. Although elegant, there is little apart from circumstantial evidence at present to support the existence of molecular mimicry in man.

Non-infectious triggers

Infection may play a role in autoimmune disease, but what non-infectious triggers might there be which could induce autoimmune disease in susceptible individuals? Hormones may play a role since we know that most autoimmune disorders are commoner in women than men, and in mice with SLE the administration of estrogens worsens the disease; this effect is believed to be due to an alteration of the B-cell repertoire. Also in SLE, complement deficiency (e.g. C1 or C4 components) may worsen the disorder. Haptens such as drugs can also induce autoimmune disease. For example, penicillin acts as a hapten when it binds to the red cell membrane, inducing an autoimmune response that leads to autoimmune hemolytic anemia. Complexes formed between two proteins may trigger disease. For example, celiac disease is an autoimmune disorder in which transglutamase and its substrate gliadin form a complex that induces autoantibody generation against both gliadin and transglutamase.

Epitope spread

This phenomenon may play a role in some autoimmune diseases, and involves the initial generation of autoreactivity with subsequent development of chronic autoimmune disease. During the process there is a gradual increase in the number of autoantigens targeted by T cells; hence there is spread to epitopes beyond the initial triggering epitope (Figure 21.5).

Hormonal influences

Many autoimmune diseases are more common in females, particularly in the child-bearing years, when hormone levels are at their highest. The mechanism whereby the hormonal profile of females influences the development of autoimmune disease is unknown.

Autoimmune disease is complex and multifactorial

From studies of autoimmune disease, it is clear that ITP and other autoimmune disorders are multifactorial. It is likely that loss of tolerance to a self antigen alone is insufficient to generate the autoimmune disorder; instead patients probably require

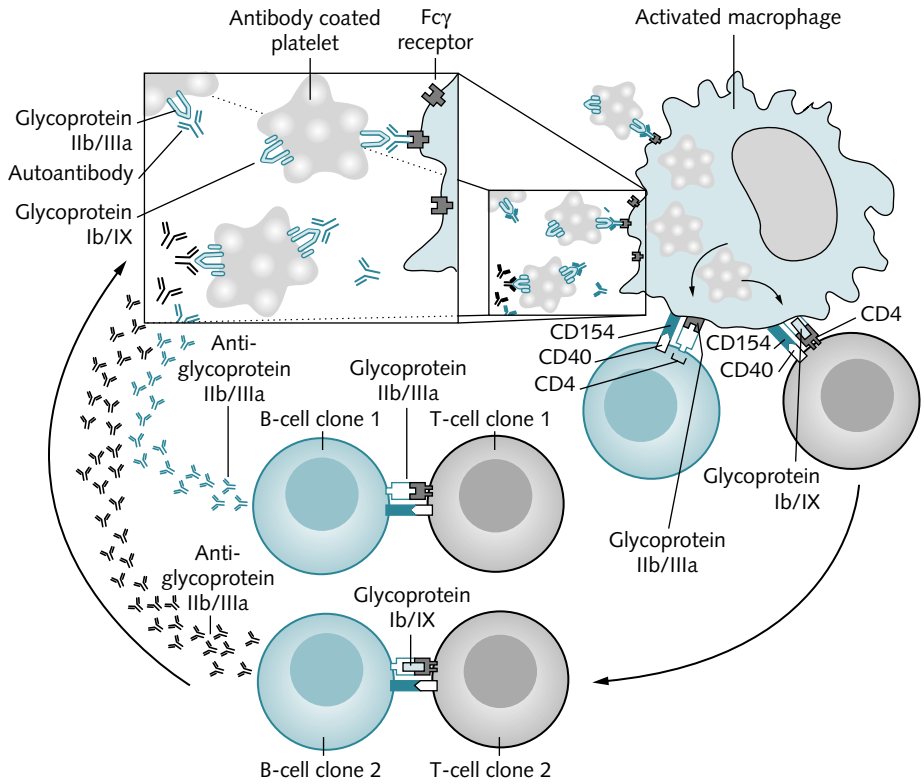


Fig. 21.5 Epitope spread
In ITP, for example, platelet glycoproteins are processed and presented to T cells by antigen-presenting cells (APCs). A variety of different peptides may be presented by the APCs, resulting in multiple T-cell clones, each targeting distinct epitopes on different glycoprotein molecules. Reproduced with permission from the *New England Journal of Medicine*.

(1) a specific set of genetic determinants (e.g. polymorphisms within MHC, CTLA4, or other genes); (2) dysregulation of the immune response (involving dendritic cells, T or B cells, or all three); (3) and an environmental trigger. Autoimmune disease arises only when all these determinants are present in an individual at one particular time. This is reinforced by the observation that self-reactive lymphocytes are commonly found in normal individuals. For example, siblings of patients who have autoimmune disorders are more likely to have autoantibodies themselves, albeit at lower titers than their affected sibs, but no overt evidence of autoimmune disease *per se,* perhaps because they have not been exposed to the environmental trigger required to tip the balance toward autoimmune disease.

How do autoreactive cells induce tissue damage in autoimmune disease?

This is a large topic which will be discussed only briefly here. In autoimmune disease, since the autoantigen is not cleared completely from the affected individual the process tends to be prolonged because there is always a supply of autoantigen available to keep the process going. Autoantigens, especially IgG and IgM, may attach to cell membrane antigens on cells and cause local tissue damage. Where the antigens are soluble a more systemic disease profile results. In addition to autoantibody-mediated disease, T cells can themselves directly cause disease. When autoantibodies such as IgG and IgM attach to blood cells premature destruction of the cells results. Cells involved include red cells, resulting in autoimmune hemolytic anemia, white cells, causing neutropenia (any white cell may be involved but neutropenia is commonest) and platelets, inducing thrombocytopenia. Red cell attack by IgG antibodies results in premature red cell destruction by the macrophages of the reticuloendothelial system, via the Fcγ receptors. When the autoantibody is IgM, complement activation occurs and there may be intravascular lysis, although more commonly the complement cascade does not reach the lytic stage but generally stops at the C3d stage because of the presence of complement regulatory proteins. However, even with incomplete complement activation, the presence of molecules such as C5a generates local inflammation and tissue damage through the activation of cytokines. Autoimmune blood cell disorders are discussed below.

Rather than cause direct cellular damage, some autoantibodies may activate or block receptors. For example, Graves disease is caused by an autoantibody that stimulates the thyroid-stimulating receptor, leading to elevated levels of thyroid hormones. In myasthenia gravis the autoantibody blocks the acetyl choline receptors, preventing neuromuscular transmission.

Other mechanisms of autoantibody and T-cell damage also occur but there is insufficient space to discuss these here.

Idiopathic thrombocytopenic purpura as a model of autoimmune blood disease

ITP is usually an acquired disorder in which platelets are coated (*opsonized*) with antiplatelet autoantibodies and removed prematurely by the reticuloendothelial system, predominantly the spleen, leading to a reduced peripheral blood platelet count. The etiology of ITP is unknown and the clinical course is variable and unpredictable. ITP has an incidence of around 60 new cases per million population per year in the USA. Childhood ITP is generally termed 'acute' since the illness is seasonal, typically follows a trivial viral infection or vaccination, is transient in most cases and requires no treatment, with spontaneous recovery in 80% of cases. One proposed mechanism invoked in childhood ITP is molecular mimicry, in which the antibody directed against an invading pathogen, by chance, cross-reacts with one of the platelet glycoprotein epitopes. As discussed earlier, the normal adaptive immune response ceases once the offending pathogenic antigen is destroyed, and this might account for the acute nature of ITP in childhood. That is to say, once the pathogen is eradicated the source of non-self antigen is removed and cross-reacting antibody levels fall.

In the adult (chronic) form there is usually no obvious antecedent illness and most patients have chronic thrombocytopenia; spontaneous recovery is uncommon. In most cases of adult ITP the platelet glycoprotein (GP) antigen targets are GPIIb/IIIa and GPIb/IX.

Etiology

It is believed that ITP is most likely due to an inappropriate immune response to an environmental trigger; the nature of this trigger has not yet been identified. The disorder may represent an abnormality of APCs, with an increase in the numbers of CD4- and CD8-positive cells. The platelets are rapidly destroyed by the immune complexes that bind to the Fc receptors on the platelets, or due to autoantibodies that bind to the antigenic site on the platelets. Platelets that are coated with antibody or immune complexes are rapidly cleared by the reticuloendothelial system.

As regards causal genetic abnormalities, most attention has focussed on the identification of MHC susceptibility genes, given their role in determining the nature and specificity of the adaptive immune response. However, the results of HLA association studies in ITP have not produced clear results. A large analysis focussing on HLA-DR4 gene variations in more than 100 Japanese patients with ITP reported that the DRB1*0410 allele was significantly increased in ITP patients compared with controls. Moreover, this allele was significantly decreased in patients who showed a good response to

prednisolone. MHC may therefore play a role in some cases of ITP but there are clearly other genes implicated. These include genetic polymorphisms within cytokine and other immune regulatory genes.

Clinical features of ITP

The ITP phenotype is heterogeneous: some patients suffer major bleeding from the outset, while others have few problems apart from an increased tendency to bruise. This may partly be explained by the acquired platelet dysfunction that is seen in some patients with ITP, which in turn may be related to the target antigen involved in the autoimmune process (this is discussed later; see *Antibodies and their target antigens*). Autoantibodies reacting with GPIIb/IIIa affect platelet aggregation, and anti-GPIb/IX autoantibodies impair platelet adhesion to the subendothelial matrix, causing unexpectedly severe bleeding for the level of the platelet count. In general, however, in contrast with thrombocytopenia due to marrow infiltration (e.g. leukemias, lymphomas and other malignancies) or aplasia, patients with ITP are able to tolerate remarkably low platelet counts and are able to maintain an adequate quality of life. The degree of bleeding is largely dependent on the platelet count, and patients with platelet counts below 10×10^9/L (and usually below 5×10^9/L) are at greatest risk of bleeding, including intracranial bleeding.

Diagnosis

The diagnosis of ITP remains clinical, and one of exclusion. Secondary causes include SLE, lymphoproliferative disease and HIV infection. Standard investigation includes a full blood count (isolated thrombocytopenia), blood film (to ensure no red cell fragments, leukemia, parasitic infections) and autoimmune profile (to exclude secondary cause). A bone marrow examination is often carried out in adults but not usually in children, and will usually show normal or increased megakaryocytes in an otherwise normal marrow. Immunological assays have been devised, including platelet-associated IgG or IgM and monoclonal antibody immobilization of platelet antigens, but these do not alter the management and are of debatable value.

Antibodies and their target antigens

Antiplatelet antibodies

Many patients with ITP show elevated levels of platelet-associated IgG, which is believed to be the autoantibody, but for unknown reasons platelet-associated IgG may be elevated in other non-immunological causes of thrombocytopenia. The autoantibodies involved in ITP are generally IgG, but IgA and IgM autoantibodies have been reported. Opsonized platelets are removed prematurely by the reticuloendothelial system through an Fc-dependent mechanism. However, many patients fail to respond to therapies aimed at inactivation of the reticuloendothelial system, suggesting that other mechanisms of platelet destruction exist.

Antigenic targets

Using antigen-specific assays such as the monoclonal antibody-specific immobilization of platelet antigens, platelet-associated IgG and antigen capture assays, several platelet antigens have been characterized. These include GPIIb/IIIa ($\alpha_{IIb}\beta_3$, the fibrinogen receptor) and GPIb/IX (the von Willebrand receptor), which appear to be the most frequently involved. Less commonly, GPIa/IIa, GPIV and GPV are involved. Recent reports suggest that possibly 40% of autoantibodies are reactive to both GPIIb/IIIa and GPIb/IX, possibly due to the serum in some patients with ITP containing two different IgG antibodies. In terms of disease chronicity, GP-specific autoantibodies may be important in the pathogenesis of chronic ITP; from available data GPIIb/IIIa appear to play a major role in the development of chronic ITP in 30–40% of cases.

Which epitopes are involved?

Previous investigators have looked for autoantigenic epitopes on the GPIIb/IIIa molecule using competitive binding between human autoantibodies and mouse monoclonal antibodies. In addition, enzyme-cleaved IIb or IIIa fragments and synthesized peptides corresponding to different sequences of GPIIIa have been used to localize epitopes on the respective glycoprotein.

Kekomaki and colleagues have shown that the 33-kDa chymotryptic core fragment of IIIa is a frequent target. Fujisawa and colleagues, using synthetic peptides corresponding to IIIa sequences, showed that in five of 13 sera from patients with chronic ITP binding was to residues 721–744 or 742–762, corresponding to the carboxy terminal of IIIa. Recently, Nieswandt and colleagues have looked at the pathogenic effects of IgG monoclonal antibodies of different IgG subclasses against murine GPIIb/IIIa, Ibα, Ib/IX, V and CD31. Their data suggest that, at least in mice, the antigenic specificity of the antiplatelet antibodies determines the pathogenic effects rather than the IgG subclass. They also demonstrated that antibodies against GPIb/IX caused thrombocytopenia through an Fc-independent mechanism, whilst that from autoantibodies against GPIIb/IIIa involved the Fc system. Further work is clearly needed in order to determine the significance of all of these findings, which may translate into stratification

of patients into those in whom Fc receptor blockade or inactivation is a useful option, and those in whom it is not.

Standard treatment

There is a lack of clinical trial data to help guide treatment and our energies should now be focussed on constructing high quality randomized trials in order to determine the most effective therapy in this disorder. Therapy is seldom necessary for patients whose platelet counts exceed 20–30 $\times$ 10^9/L and in whom there are few spontaneous bleeding episodes, unless they are undergoing any procedure likely to induce blood loss. Standard treatments, including oral prednisolone, IVIg and splenectomy, will elevate the platelet count sufficiently in most adults (Figure 21.6). However, some 20–25% of adults with ITP are refractory to first-line therapy.

Chronic refractory ITP

This defines those patients who fail to respond to first-line treatment or require unacceptably high doses of corticosteroids to maintain a safe platelet count. A number of agents have been used as second-line therapy for ITP, including high-dose steroids, high-dose IVIg, intravenous anti-D, vinca alkaloids, danazol, azathioprine, combination chemotherapy and dapsone.

Targeted versus untargeted therapies for autoimmune disease

Until now most of our treatments for autoimmune disease have been untargeted and unselective in their modes of action. In disorders such as ITP the therapeutic aim has been to induce global immunosuppression in the hope that, as part of this process, the ITP-related component of the immune system will be suppressed, and this will help reduce the quantity of autoantibody produced. For antibody-mediated autoimmune diseases, what remains unclear is whether the B-cell population that is generating the antiplatelet autoantibody is the primary problem, or whether events downstream, such as those involving antigen presentation or T-cell regulation, are disturbed, and simply driving the passive B cells, resulting in the autoantibody phenotype.

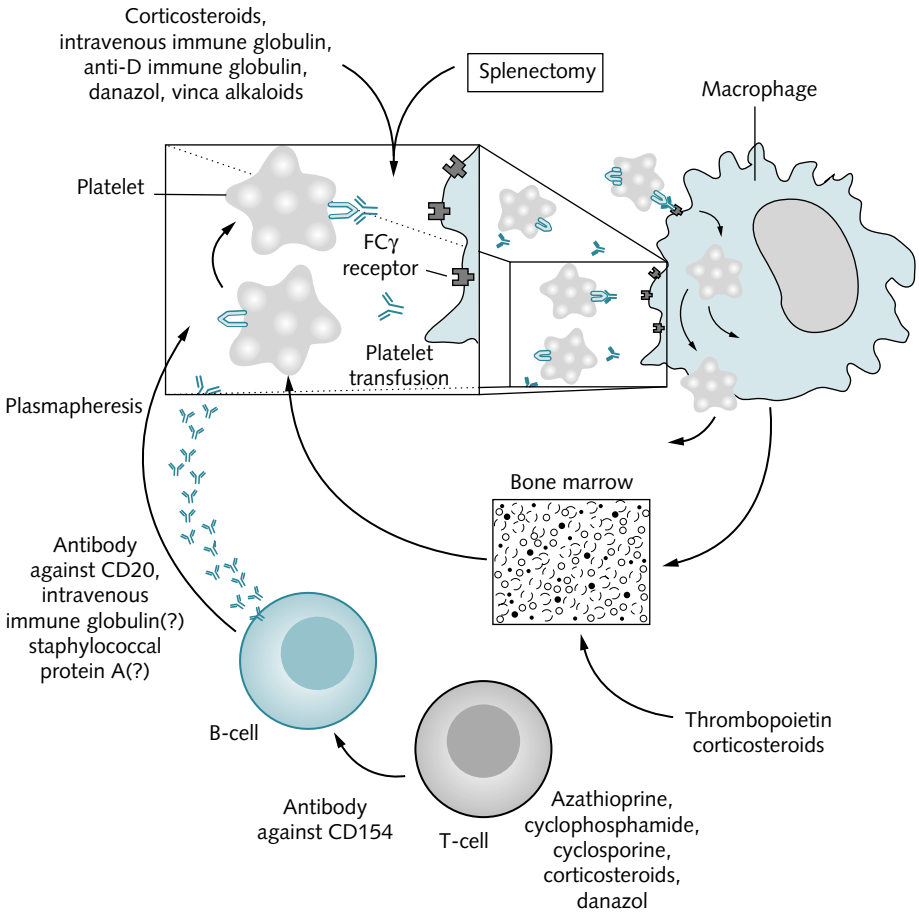


Fig. 21.6 Standard and novel treatment strategies in ITP
Standard treatments include corticosteroids, intravenous immunoglobulin, danazol and vinca alkaloids. Their sites and modes of action are illustrated. New therapies include monoclonal antibodies against CD154 (CD40 ligand) and CD20 (on B cells, leading to transient B-cell depletion). Reproduced with permission from the *New England Journal of Medicine*.

Now that we have a clearer understanding of the immunological mechanisms involved in autoimmune disease, we have started to develop more targeted therapies. We are now developing treatments designed to target T cells, B cells and other effectors within the immune system. For ITP these include Campath-1H and anti-CD20. Although these agents are not entirely specific because they deplete the B-cell compartment, they should reduce the quantity of autoantibody produced. Other therapies which may have been shown to be of benefit in ITP are mycophenolate mofetil and anti-CD40 ligand.

Campath-1H

Campath-1H is a humanized IgG monoclonal antibody which targets the CD52 antigen, present on mature human lymphocytes (T and B cells) and monocytes. Campath-1H is effective in the treatment of malignant B-cell disorders, especially B-cell chronic lymphocytic leukemia (B-CLL), in which it has been shown to be effective in clearing lymphocytes from both blood and bone marrow.

Campath-1H has been used in a variety of autoimmune diseases, including rheumatoid arthritis, vasculitis and Wegener's granuloma. There is ongoing interest in the use of Campath-1H for the treatment of autoimmune hematological disease that is refractory to first- and second-line therapies. One recent study of the use of Campath-1H in autoimmune neutropenia, autoimmune hemolytic anemia, pure red cell aplasia, immune thrombocytopenia and combined hemolytic anemia and ITP (Evans syndrome) has shown responses in 15 out of 21 patients treated; in six patients the response was sustained. Campath-1H therefore appears to be an effective agent in severe refractory autoimmune disease. The drug is well tolerated, but because it can precipitate bleeding during administration it should not be given in the presence of active bleeding (or active infection).

Anti-CD20 monoclonal antibody therapy

Rituximab, a genetically engineered chimeric human/mouse anti-CD20 monoclonal antibody, has been developed as a treatment for B-cell lymphoproliferative disease (non-Hodgkin's lymphoma). The antibody is an IgG κ immunoglobulin comprising murine light- and heavy-chain variable-region sequences and human constant-region sequences. The antigen-binding domain binds to the CD20 antigen on B cells while the Fc domain mediates B-cell lysis, through recruitment of immune effector cells. Because of its specificity for B cells, rituximab has been viewed as a potential treatment for autoimmune disease, the rationale being the reduction or elimination of autoantibody-producing B cells with concomitant improvement of the autoimmune disease. A recent study by Stasi and colleagues reports on the efficacy of rituximab used to treat 25 patients with chronic refractory ITP. Patients were treated if their platelet counts were below 20×10^9/L irrespective of symptoms, or at higher platelet counts if bleeding or bruising was problematic. All patients had received between two and five previous treatments; eight had failed splenectomy. Rituximab was administered in the same manner and dose as that used in non-Hodgkin's lymphoma. After four courses, 40% of patients achieved a platelet count of at least 50×10^9/L; five achieved complete remission (platelets $>100 \times 10^9$/L) and five partial remission (platelets $50–100 \times 10^9$/L). Responses were seen during treatment with rituximab, with a peak response up to 4 weeks after the end of treatment. Twenty-eight percent had responses that lasted for more than 6 months.

The results suggest that the use of rituximab resulted in responses similar to those given by other second-line agents used in ITP (including vinca alkaloids, cyclophosphamide and azathioprine), around 40–50%, but sustained responses to these agents is usually seen in fewer than 20% of patients; that is, lower than for rituximab. Rituximab would appear to be useful for some patients with chronic symptomatic refractory ITP in whom there is a definite need to elevate the platelet count to a safe level.

Costimulatory blockade

Therapies such as Campath-1H and anti-CD20 may not produce lasting remission if the autoimmune B cells are driven by dysregulated T cells, and a novel agent, CTLA-4-Ig, has been evaluated in psoriasis in an attempt to block T-cell costimulation, thereby inducing anergy in the T-cell compartment. CTLA-4-Ig, a fusion protein between CTLA-4 and the Fc portion of human Ig, binds to B7-1 and B7-2, blocking T-cell costimulation (Figure 21.7). This small trial showed that, at least within this group of patients, CTLA-4-Ig was able to improve the disorder and was shown to be safe. CTLA-4-Ig may have applications within other autoimmune disorders, including ITP. If a drug such as CTLA-4-Ig were shown to be effective in ITP, not only would it provide an additional targeted treatment modality, but it would also provide useful evidence of T-cell dysfunction in this disease. Interestingly, the *CTLA-4* gene has been mapped as a susceptibility gene in autoimmune thyroid disease and IDDM in humans.

Other options: *Helicobacter pylori* eradication

This bacterium is the main cause of gastritis and peptic ulcer disease. It has also been implicated in the development of gastric adenocarcinoma and mucosa-associated lymphoid tumors and in some autoimmune disorders. Previous studies of *H. pylori* in ITP showed improvement in platelet counts

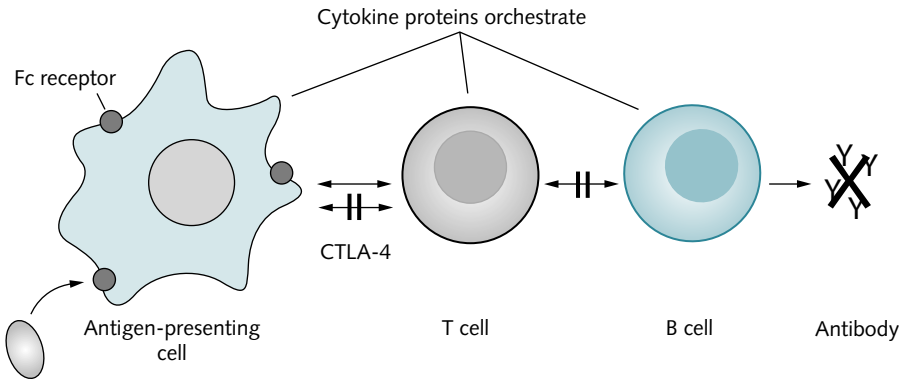


Fig. 21.7 Costimulatory blockade may be beneficial in some autoimmune diseases
Cytotoxic T-lymphocyte antigen 4 linked to human Ig Fc (CTLA-4-Ig) blocks the critical second signal between antigen-presenting cells and T cells, resulting in T-cell anergy. This should result in a reduction in antibody production and amelioration of disease if the autoimmune disease is antibody-mediated. CTLA-4-Ig treatment has been shown to be of benefit to patients with psoriasis. Similarly, anti-CD40 ligand also blocks the second signal with similar results, and has been shown to be of value in refractory ITP.

after eradication of the bacterium in patients shown to be positive for *H. pylori*. More recently, Emilia and colleagues looked for the presence of *H. pylori* in 30 patients with chronic refractory ITP. *H. pylori* was found in 13 of the 30 patients (43.3%). Standard triple therapy for *H. pylori* eradication resulted in a complete response in four of 12 patients in whom the bacterium was eradicated, and partial response in two of 12 (16.6%). The responses were maintained for a median of 8.33 months. In addition, there are other anecdotal reports of improvements in platelet counts in adults and children with ITP after eradication of *H. pylori*. Larger studies are required to confirm these earlier findings, but from the available data triple therapy appears to offer a non-immunosuppressive therapy for patients with refractory ITP and possibly other autoimmune diseases.

Novel therapies for the treatment of other autoimmune diseases

Human and animal studies have been helpful in terms of learning how the immune system works in both health and disease, but can such information be translated into better patient care? Until recently the treatment strategy for autoimmune diseases has been to induce global immunosuppression in the hope that the autoimmune process may be abrogated or stopped. In some cases treatment is effective but it is clear from longitudinal follow-up studies of patients with ITP that there is very significant morbidity and mortality associated with our current treatments. Infection plays a major role in the death of patients with autoimmune disease, and such fatalities are usually induced by immunosuppression. Now that we have better understanding of the components of the immune system and how these interact in disease we should be able to develop targeted therapies that aim to modify specific components of the immune system while leaving most of the immune system intact and able to fight infection.

Such therapeutic advances are in fact being made, and many of these have been developed through knowledge concerning specific components of the immune system in disease. We have now been able to develop targeted therapies for rheumatoid disease, multiple sclerosis, psoriasis, diabetes and systemic lupus. We will discuss each of these briefly in turn.

Rheumatoid arthritis and blockade of TNF-α

Disease-modifying therapies, designed to reduce deformity in rheumatoid disease, have been around for a number of years. Methotrexate is one of the main agents in this category. Bone destruction in rheumatoid disease is partly mediated by macrophages through an inflammatory process. With our increased understanding of cytokines and their interactions, specific modulators of the immune response have been developed. Because TNF-α plays a key role in the inflammatory response, antibodies against TNF-α have been developed and shown to be of value. To date, two antibodies appear promising: the first is a TNF-α receptor–IgG1 fusion protein, etanercept. The second, infliximab, is a monoclonal antibody directed against TNF-α itself. This form of therapy also appears to have a place in the management of other autoimmune diseases, such as Crohn's disease, psoriatic arthropathy and ankylosing spondylitis.

Another cytokine involved in the pathogenesis of rheumatoid disease is IL-1. Blockade of the IL-1 receptor using a recombinant antagonist may slow down the development of erosive bone disease but from available data this appears to be less effective than TNF-α.

Multiple sclerosis

IFN-β has recently been introduced for the treatment of multiple sclerosis and the available study data indicate that IFN-β may delay the onset of the disease if started immediately after the patient's first attack of optic neuritis.

Psoriasis

This autoimmune skin disorder has been treated by TNF-α blockade with and without methotrexate. Other agents used in small clinical trials include IL-10 and CTLA-4-Ig; the latter molecule is a recombinant protein comprising the extracellular domain of CTLA-4 linked to the constant region of IgG1. CTLA-4-Ig downregulates activated T cells and prevents activation of naive T cells. CTLA-4-Ig was well tolerated in a small trial conducted by Bristol-Myers-Squibb and the patients' skin condition improved by more than 50%. Further studies are required to confirm safety and efficacy.

There are studies of other agents used in the treatment of autoimmune disease but there is insufficient space to discuss these here. However, it would appear that our knowledge of immunity and autoimmunity is now being used to design more subtle and specific treatments for patients with autoimmune disease. It is expected that these treatments will reduce immunosuppression, morbidity and mortality and offer modern solutions to these otherwise intractable diseases. No doubt the future will see even more designer drugs being developed for this fascinating group of diseases.

Conclusions

Autoimmune diseases are complex immunological disorders affecting 5% of the population. Until recently our understanding of the pathogenesis and treatment of these disorders was severely limited. However, with greater understanding of the immune system in health and autoimmune disease we are able to identify underlying abnormalities leading to the develop of autoimmunity. With this new knowledge we have been able to modify our therapies by replacing non-selective immunosuppressive treatment with more subtle targeted therapies.

Further reading

General

Janeway CA, Travers P, Walport M, Capra JD. (1999) *Immunobiology: The Immune System in Health and Disease*, 4th edition. New York: Current Biology.

Roitt IM, Delves PJ. (2001) *Roitt's Essential Immunology*, 10th edition. Oxford: Blackwell Science.

Immune system and HLA

Chien YH, Gascoigne NR, Kavaler J *et al.* (1984) Somatic recombination in a murine T-cell receptor gene. *Nature*, **309**, 322–326.

Delves PJ, Roitt IM. (2000) The immune system. First of two parts. *New England Journal of Medicine*, **343**, 37–49.

Delves PJ, Roitt IM. (2000) The immune system. Second of two parts. *New England Journal of Medicine*, **343**, 108–117.

Klein J, Sato A. (2000) The HLA system. First of two parts. *New England Journal of Medicine*, **343**, 702–709.

Klein J, Sato A. (2000) The HLA system. Second of two parts. *New England Journal of Medicine*, **343**, 782–786.

Parkin J, Cohen B. (2001) An overview of the immune system. *Lancet*, **357**, 1777–1789.

Tonegawa S. (1983) Somatic generation of antibody diversity. *Nature*, **302**, 575–581.

Tolerance

Kamradt T, Mitchison NA. (2001) Tolerance and autoimmunity. *New England Journal of Medicine*, **344**, 655–664.

Mackay IR. (2000) Science, medicine, and the future: tolerance and autoimmunity. *British Medical Journal*, **321**, 93–96.

Sarzotti M. (1997) Immunologic tolerance. *Current Opinion in Hematology*, **4**, 48–52.

Autoimmunity

Davidson A, Diamond B. (2001) Autoimmune diseases. *New England Journal of Medicine*, **345**, 340–350.

Maloy KJ, Powrie F. (2001) Regulatory T cells in the control of immune pathology. *Nature Immunology*, **2**, 816–822.

Sinha AA, Lopez MT, McDevitt HO. (1990) Autoimmune diseases: the failure of self tolerance. *Science*, **248**, 1380–1388.

Taneja V, David CS. (2001) Lessons from animal models for human autoimmune diseases. *Nature Immunology*, **2**, 781–784.

Role of genetic and other factors in autoimmune disease

Albert LJ, Inman RD. (1999) Molecular mimicry and autoimmunity. *New England Journal of Medicine*, **341**, 2068–2074.

Benoist C, Mathis D. (2001) Autoimmunity provoked by infection: how good is the case for T cell epitope mimicry? *Nature Immunology*, **2**, 797–801.

Choy EH, Panayi GS. (2001) Cytokine pathways and joint inflammation in rheumatoid arthritis. *New England Journal of Medicine*, **344**, 907–916.

Criswell LA, Moser KL, Gaffney PM *et al.* (2000) PARP alleles and SLE: failure to confirm association with disease susceptibility. *Journal of Clinical Investigation*, **105**, 1501–1502.

Ermann J, Fathman CG. (2001) Autoimmune diseases: genes, bugs and failed regulation. *Nature Immunology*, **2**, 759–761.

Gianani R, Sarvetnick N. (1996) Viruses, cytokines, antigens, and autoimmunity. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 2257–2259.

Marron MP, Zeidler A, Raffel LJ *et al.* (2000) Genetic and physical mapping of a type 1 diabetes susceptibility gene (IDDM12) to a 100-kb phagemid artificial chromosome clone containing D2S72-CTLA4-D2S105 on chromosome 2q33. *Diabetes*, **49**, 492–499.

Spielman RS, Ewens WJ. (1996) The TDT and other family-based tests for linkage disequilibrium and association. *American Journal of Human Genetics*, **59**, 983–989.

Wanstrat A, Wakeland E. (2001) The genetics of complex autoimmune diseases: non-MHC susceptibility genes. *Nature Immunology*, **2**, 802–809.

Idiopathic thrombocytopenic purpura

Bussel JB. (2002) Novel approaches to refractory immune thrombocytopenic purpura. *Blood Reviews*, **16**, 31–36.

Cines DB, Blanchette VS. (2002) Immune thrombocytopenic purpura. *New England Journal of Medicine*, **346**, 995–1008.

de Luis DA, Varela C, de La Calle H *et al.* (1998) *Helicobacter pylori* infection is markedly increased in patients with autoimmune atrophic thyroiditis. *Journal of Clinical Gastroenterology*, **26**, 259–263.

Emilia G, Longo G, Luppi M *et al.* (2001) *Helicobacter pylori* eradication can induce platelet recovery in idiopathic thrombocytopenic purpura. *Blood*, **97**, 812–814.

Fujisawa K, Tani P, O'Toole TE *et al.* (1992) Different specificities of platelet-associated and plasma autoantibodies to platelet GPIIb-IIIa in patients with chronic immune thrombocytopenic purpura. *Blood*, **79**, 1441–1446.

Gasbarrini A, Franceschi F, Tartaglione R *et al.* (1998) Regression of autoimmune thrombocytopenia after eradication of *Helicobacter pylori*. *Lancet*, **352**, 878.

Kekomaki R, Dawson B, McFarland *et al.* (1991) Localization of human platelet autoantigens to the cysteine-rich region of glycoprotein IIIa. *Journal of Clinical Investigation*, **88**, 847–854.

Nieswandt B, Bergmeier W, Rackebrandt K *et al.* (2000) Identification of critical antigen-specific mechanisms in the development of immune thrombocytopenic purpura in mice. *Blood*, **96**, 2520–2527.

Portielje JE, Westendorp RG, Kluin-Nelemans *et al.* (2001) Morbidity and mortality in adults with idiopathic thrombocytopenic purpura. *Blood*, **97**, 2549–2554.

Provan D, Newland A. (2002) Fifty years of idiopathic thrombocytopenic purpura (ITP): management of refractory ITP in adults. *British Journal of Haematology*, **118**, 933–944.

Stasi R, Pagano A, Stipa E *et al.* (2001) Rituximab chimeric anti-CD20 monoclonal antibody treatment for adults with chronic idiopathic thrombocytopenic purpura. *Blood*, **98**, 952–957.

Zentilin P, Savarino V, Garnero A *et al.* (1999) Is *Helicobacter pylori* infection a risk factor for disease severity in rheumatoid arthritis? *Gastroenterology*, **116**, 503–504.

Therapeutic advances

Abrams JR, Lebwohl MG, Guzzo CA *et al.* (1999) CTLA4Ig-mediated blockade of T-cell costimulation in patients with psoriasis vulgaris. *Journal of Clinical Investigation*, **103**, 1243–1252.

Bell S, Kamm MA. (2000) Antibodies to tumour necrosis factor alpha as treatment for Crohn's disease. *Lancet*, **355**, 858–860.

Brandt J, Haibel H, Cornely D *et al.* (2000) Successful treatment of active ankylosing spondylitis with the anti-tumor necrosis factor alpha monoclonal antibody infliximab. *Arthritis and Rheumatology*, **43**, 1346–1352.

Kremer JM. (2001) Rational use of new and existing disease-modifying agents in rheumatoid arthritis. *Annals of Internal Medicine*, **134**, 695–706.

Kuwana M, Kawakami Y, Ikeda Y. (2003) Suppression of autoreactive T-cell response to glycoprotein IIb/IIIa by blockade of CD40/CD154 interaction: implications for treatment of immune thrombocytopenic purpura. *Blood*, **101**, 621–623.

Maini RN, Taylor PC. (2000) Anti-cytokine therapy for rheumatoid arthritis. *Annual Review of Medicine*, **51**, 207–229.

Mease PJ, Goffe BS, Metz J *et al.* (2000) Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. *Lancet*, **356**, 385–390.

Chapter 22 Hematopoietic growth factors

Graham Molineux

Introduction

It has been known since ancient times that the blood system responds to specific stimuli in order to preserve homeostasis. To provide for ongoing replenishment of short-lived cellular blood elements, as well as the prompt and specific differentiation of target cells in response to stress states such as blood loss or infection, is the process of *hematopoiesis*. Hematopoiesis is regulated on several levels: the cells themselves have differing maturation potentialities (Figure 22.1), the extracellular matrix within which hematopoiesis occurs is known to have more than a merely permissive role, and there exists also a series of glycoprotein hormones termed the hematopoietic growth factors (HGFs). The mechanisms used to effect regulation of hematopoiesis involve complex cellular interactions as well as responses to circulating HGF and may result in apoptosis, mitosis and the differentiation of specific subsets of hematopoietic progenitor cells.

The development of semisolid culture systems more than three decades ago allowed normal hematopoietic progenitor cells to proliferate *in vitro* and form recognizable colonies when conditioned media (containing various sera or extracts) were used. Because the specific protein factors that induce colony growth were initially identified in the sera through these colony formation assays, they were named 'colony-stimulating factors' (CSFs). Considerable progress followed quickly to determine the cell types contained within the hematopoietic colonies and to characterize the different CSFs in both murine and human systems. The *in vitro* assays continue to be refined and have been used to identify the lineage hierarchy of normal hematopoietic progenitor cells and the CSFs involved at each level.

In addition to the CSFs, many interleukins (ILs, originally immune cell-to-immune cell signaling molecules) have been identified which are active in the hematopoietic cascade. Some ILs can be, or originally were, considered to be single-lineage HGFs, such as IL-2 (T-cell growth factor), IL-5 (eosinophil differentiation factor), IL-7 (B-lymphocyte progenitor cell growth factor) and IL-12 (natural killer cell stimulatory factor), although in general they are now believed to be quite pleiotropic.

Many of these cytokines are produced by marrow stromal cells and are undoubtedly involved in hematopoietic stem cell development at specific sites within the marrow. Indeed, several examples of interactions with specific extracellular matrix materials have emphasized that CSFs often are localized very specifically and may effect their actions via locally determined mechanisms. We now know that the CSFs induce proliferation of hematopoietic progenitor cells, activate mature blood cells, enhance mature effector cell function and initiate the production of other HGFs.

Molecular cloning

The use of improved biotechnology tools led to the isolation, cloning and expression of these factors to produce large quantities for pharmacological use, but the challenges early investigators faced were substantial. We will illustrate a few examples below.

Murine IL-3, the first HGF to be cloned, was also known as 'multi-CSF' from its demonstrated activity as a T-cell stimulatory factor, mast cell growth factor, eosinophil CSF, erythroid megakaryocyte-stimulating factor, and granulocyte- and macrophage-stimulating activities in murine systems. When murine cDNA clones were used to probe for the human gene, they were unable to hybridize human DNA, leading to the conclusion that there was only minor homology between the two molecules. It was later learned that there is only 29% homology between murine and human IL-3 proteins, with no cross-reactivity, prohibiting their use in assays for each other.

The cloning of erythropoietin (EPO), the first HGF to become commercially available, provides interesting insight into the challenges involved in the field of biotechnology research. Following its purification from the urine of aplastic anemia patients, there was a limited amount known about

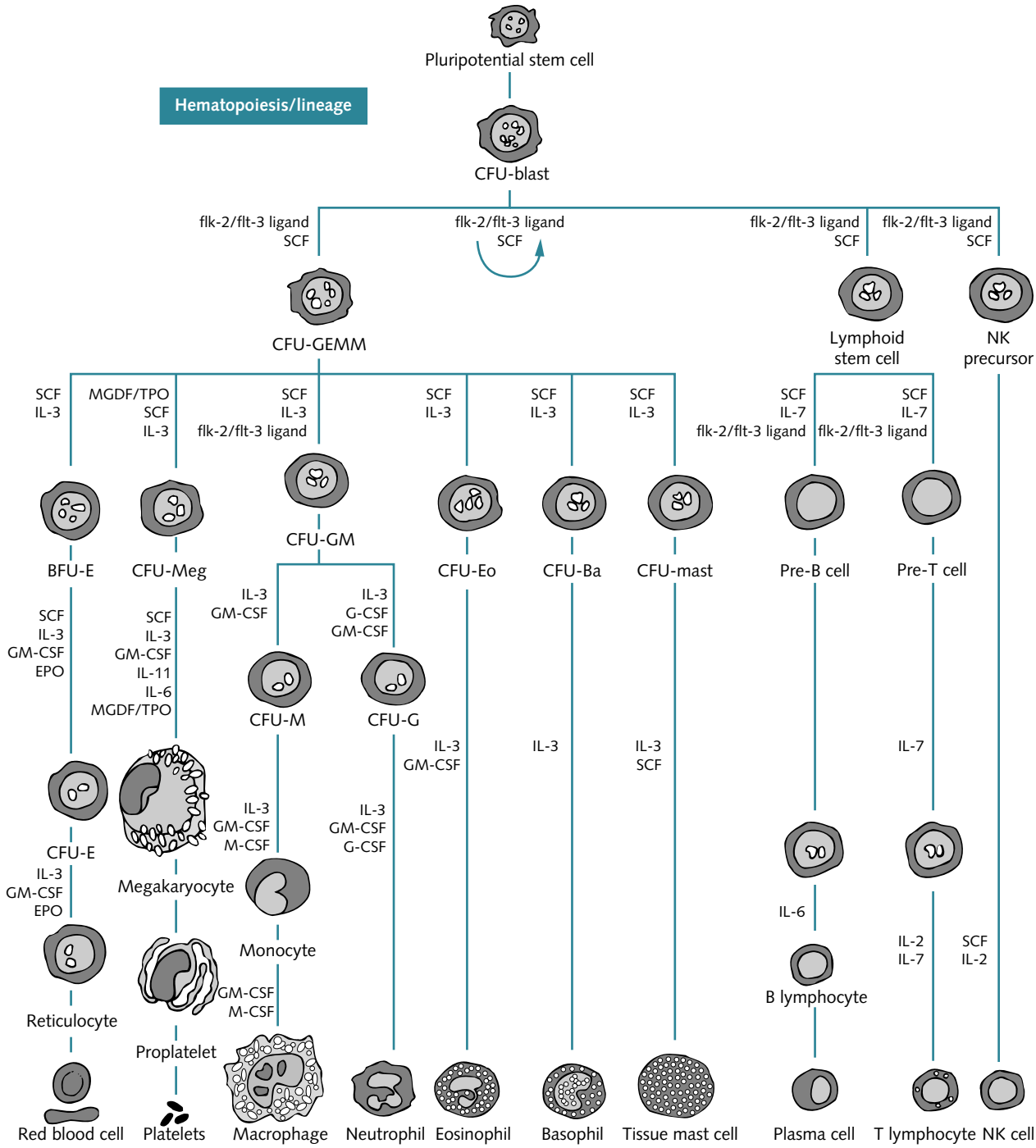


Fig. 22.1 Scheme of hematopoiesis, including some of the growth factors that influence the production of blood cells

CFU-GEMM, multilineage colony-forming cell; BFU-E, erythroid burst-forming cell; CFU-GM, granulocyte-macrophage colony-forming cell; CFU-M, macrophage colony-forming cell; CFU-Eo, eosinophil colony-forming cell; CFU-Ba, basophil colony-forming cell; IL, interleukin; SCF, stem cell factor; G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte-macrophage colony-stimulating factor; M-CSF, macrophage colony-stimulating factor; TPO, thrombopoietin; MGDF, megakaryocyte growth and development factor; EPO, erythropoietin.

the primary structure of human EPO and no known source of erythropoietin mRNA. No cell lines had been described that produced significant amounts of EPO that could be useful as a source of mRNA.

The human *EPO* gene was cloned in two laboratories using similar techniques. It was necessary to use several methods to isolate a single-copy gene from a genomic library with multiple DNA probes. Once the technology was developed and the appropriate DNA probe sequences were identified, however, it was possible to clone the genes for EPO from the human, monkey, mouse and many other species. The choice of cell line (Chinese hamster ovary, CHO) to produce the glycosylated product necessary for *in vivo* activity of EPO complicated the final stages of drug development.

Another successful commercialized use of biotechnology is granulocyte colony-stimulating factor (G-CSF). The presence of this factor in the serum of endotoxin-treated mice was first recognized as having hematopoietic regulatory activity and was able to induce differentiation of a mouse leukemic cell line. The human G-CSF molecule was purified from the bladder carcinoma cell line 5637 or from a squamous cell carcinoma cell line and the gene was cloned. Luckily for the early investigators, G-CSF has biological activity across species (i.e. human G-CSF is biologically active in mice), probably because of close homology in functional domains (73% overall homology at the protein level between human and murine molecules) (Figure 22.2).

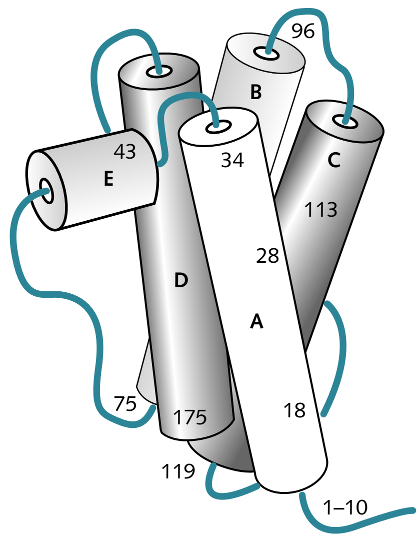


Fig. 22.2 G-CSF as an example of the structure of a cytokine Note the four antiparallel helices. Amino acid numbers are shown.

In contrast, granulocyte-macrophage colony-stimulating factor (GM-CSF), another myeloid lineage factor which is also commercially available, was purified from the human monocyte cell line U937 and the mouse T-lymphoma cell line ATCC-CRL-8080. Unlike G-CSF, there is no close homology between human GM-CSF and murine GM-CSF, although the molecular structure of GM-CSF is relatively highly conserved, with 56% homology at the protein level between human and murine forms; however, the hematopoietic activity does not cross these species barriers (Figure 22.3). The lack of an appropriate animal model for testing may have resulted in some delays in drug development.

Thrombopoietin was long thought to exist, but demonstration of the lineage-restricted regulator of platelet production was not completed until relatively recently. The redundancy apparent in the regulators of this lineage may also manifest itself in other lineages when a fuller understanding is gained. However, on a practical level the overlap of many cytokines in the regulation of thrombopoiesis led to the development of numerous putative platelet factors, such as IL-3, GM-CSF, IL-6 and IL-11, before the lineage-restricted regulator was found. Thus, blind alleys also existed to mislead the intrepid early researcher.

The results from molecular cloning have not only allowed the production of large quantities of purified factors

```
                                         -------A-------                  -S
                   10         +20         30          40          50         60
Human   MWLQSLLLLGTVASCISAPARSPSPSTQPWEHVNAIQEARRLLNLSRDTAAEMNETVEVI
Mouse   MWLQNLLFLGIVVYSLSAPTRSPITVTRPWKHVEAIKEALNLLD---DMPVTLNEEVEVV
Bovine  MWLQNLLLLGTVVCSFSAPTRPPNTATRPWQHVDAIKEALSLLNHSSDTDAVMNDT-EVV

        -             -----B-----          ------C------              -S--
                   70         80          90          100         110        120
Human   SEMFDLQEPTCLQTRLELYKQGLRGSLTKLKGPLTMMASHYKQHCPPTPETSCATQIITF
Mouse   SNEFSFKKLTCVQTRLKIFEQGLRGNFTKLKGALNMTASYYQTYCPPTPETDCETQVTTY
Bovine  SE KFDSQEPTCLQTRLKLYKNGLQGSLTSLMGSLTMMATHYEKHCPPTPETSCGTQFISF

        -----D-------
                  130         140
Human   ESFKENLKDFLLVIPFDCWEPVQE
Mouse   ADFIDSLKTELTDIPFECKKPSQK
Bovine  KNFDEDLKEFLFIIPFDCWEPAQK
```

Fig. 22.3 Divergence in amino acid sequence of GM-CSF between mouse and human

for further experiments, but have often provided clues about how cytokine networks interact or how disease states may be further studied. It is interesting to note that chromosome 5q contains a family of genes involved in hematopoietic regulation. The genes for GM-CSF, IL-3, M-CSF and the M-CSF receptor (c-fms), IL-4, IL-5 and the receptor for platelet-derived growth factor (PDGF) all map to the long arm of chromosome 5, prompting speculation that additional genes involved in hematopoiesis will be discovered in this region. It is noteworthy that deletions involving chromosome 5q have been described in various hematopoietic disease states. The gene encoding G-CSF is on chromosome 17, near the t(15;17) translocation breakpoint characteristic of acute promyelocytic leukemia.

Hematopoiesis

Expression

In the majority of mature cell types investigated, transcription of the *GM-CSF*, *G-CSF* and *M-CSF* genes is constitutive even in the absence of specific stimuli, although the transcripts are rapidly degraded in the cytoplasm and do not become useful protein unless the stabilization of mRNAs leads to induction. There is little evidence for constitutive production of some CSFs, except that fibroblast and endothelial cell lines spontaneously produce M-CSF *in vitro* (whether this occurs *in vivo* is as yet unclear), implying regulation via production in response to unknown factors. Contrasting with the myeloid growth factors are the recently described thrombopoietin, which would appear to be produced constitutively and regulated primarily via consumption rather than production, and the classically inducible EPO. Production of EPO from cells around the proximal tubule of the kidney has been shown in a series of elegant studies to be stimulated by hypoxia via intermediaries such as transcription factors [e.g. hypoxia-inducible factor (HIF), HNF4 and the COUP family] through to post-transcriptional events. Less is known about the mechanisms required to shut off CSF production, although there is evidence for downregulation of transcription after induction. The biological effects of cytokines are mediated through binding to a low number of high-affinity receptors on target cells. These cell surface receptors also appear to be widely expressed by abnormal hematopoietic cells and by non-hematopoietic cells. The potential significance of this ligand binding will be discussed later.

Receptor binding

Receptors consist of extracellular binding domain(s) and an intracellular domain that serves to activate cytoplasmic kinases, leading to cell differentiation and proliferation. Defects in the internal domain of this receptor can alter cell behavior to favor proliferation without differentiation, a remarkable discovery and one that has led to the suggestion of a specific lesion which predisposes severe congenital neutropenia patients to the development of acute myeloid leukemia. G-CSF exerts its function via the activation of a membrane receptor that belongs to the superfamily of hematopoietin receptors, also referred to as class I cytokine receptors. A characteristic structural feature of members of this family is the presence of four highly conserved cysteine residues and a motif of tryptophan–serine–*x*–tryptophan–serine within an region of approximately 200 amino acids in the extracellular domain. This region is referred to as the cytokine receptor homology (CRH) region and is crucial for ligand binding (Figure 22.4).

There appear to be common pathways but also some unique mechanisms for signaling by these individual factors, which helps to explain the synergistic activities observed in *in vivo* and *in vitro* studies. Recent work by Duarte and Frank has illustrated for the first time the nature of the synergy between a pair of cytokines, in this case G-CSF and stem cell factor (SCF). They showed that G-CSF, but not SCF, induces the phosphorylation of tyrosine residues on STAT1 and STAT3 (STAT = signal transducer and activator of transcription), but that SCF induces phosphorylation of the serine at position 727 of STAT3, which is required for full activation of STAT3. Therefore, there would appear to be complementary phosphorylation events which can be effective singly but which together lead to full activation of the downstream pathway.

Some factors have been identified only as a ligand that binds to a receptor identified with a clinical consequence, as in the case of the MPL ligands, a family of closely related HGFs that bind to the thrombopoietin receptor, c-MPL. The name MPL comes from the initial finding that a truncated form of this receptor, v-MPL, was part of the transforming gene in a murine myeloproliferative leukemia retrovirus (MPLV). The proto-oncogene (m-*MPL*) was later isolated and found to be present on platelets, megakaryocytes and a few earlier hematopoietic precursor cells. Subsequently, the ligand binding to it (c-MPL ligand) was purified from thrombocytopenic plasma and identified as the physiological regulator of platelet production, historically called *thrombopoietin*.

Endogenous production

The most straightforward situation for understanding clinical disease as a result of deficient HGF production is the anemia of renal disease and the lack of EPO production by the diseased kidney. It was known that nephrectomized rats would not produce EPO in response to hypoxic stimuli, but that ureter-ligated rats would still respond normally. Subsequent studies showed that anemic patients with chronic kidney disease also had low EPO levels, leading to the realization that

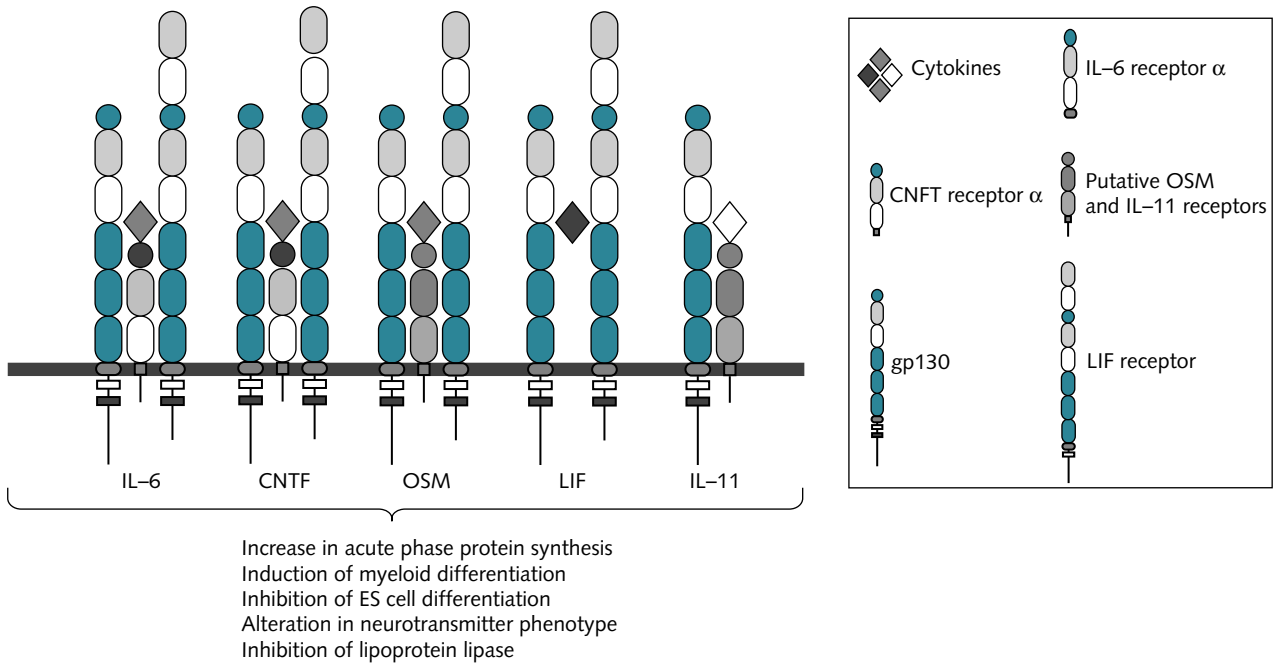


Fig. 22.4 Receptors for various cytokines, showing the structural similarities between family members

an EPO deficiency state existed and that replacement therapy may be beneficial.

In another situation where endogenous levels correlate well with the clinical scenario, thrombopoietin (TPO or c-MPL ligand) regulates most of the normal production of platelets. Thrombopoietin is the only endogenous ligand for c-MPL and is the key regulator of platelet production. As outlined above, this lineage-restricted cytokine is probably regulated via consumption rather than induction of production.

In the case of myelopoiesis, the situation is not quite as clear. Endogenous GM-CSF is rarely found in the circulation in any clinical situation, as it is believed to be primarily a locally acting factor. In contrast, the levels of circulating endogenous G-CSF increase in a variety of pathological conditions, such as gram-negative and fungal infections, and exposure to endotoxins. The highest G-CSF levels are found in neutropenic patients and are correlated with fever. In patients with cyclic neutropenia there are marked fluctuations in serum levels, and in patients with autoimmune neutropenia the peripheral absolute neutrophil count changes in parallel with the serum level of endogenous G-CSF, with a delay of 4–5 hours.

Although both G-CSF and GM-CSF are potent stimulators of myelopoiesis, the existing data suggest that G-CSF is the primary factor for the upregulation of neutrophils in infection and in various pathological conditions where the absolute neutrophil count is decreased. One possible explanation for this localization is that both G-CSF and GM-CSF are elevated during infection but that G-CSF circulates in the blood and stimulates neutropoiesis, while GM-CSF remains localized at the site of infection to help retain and activate arriving cells. The concept of regulation by consumption has been exploited in the evolution of so-called designer cytokines, such as PEG-rHuMGDF [an artificial c-MPL ligand comprising a truncated thrombopoietin-like polypeptide and a covalently attached poly(ethylene glycol) to increase the hydrodynamic size above the threshold for renal clearance] and similarly derived G-CSFs. These materials are effectively self-regulating inasmuch as the target cells they induce can in turn remove the drug itself via receptor-mediated clearance.

Clues from knockout animals

Technological advances in the production of transgenic and gene deletion or 'knockout' animals have further elucidated the constitutive and reactive roles for these factors *in vivo*.

The physiological roles of IL-2, IL-4, IL-7 and IL-10, as shown in gene knockout experiments, illustrate the essential roles of some growth factors in maintaining immune competent cells; somewhat surprisingly, mice deficient in IL-7 may have much more severe immune impairment than animals deficient in IL-2 or IL-4.

Mice deficient in the growth factor GM-CSF have relatively normal hematopoiesis but those deficient in G-CSF are severely neutropenic, with impaired neutrophil function. This suggests that G-CSF is required for maintaining the normal

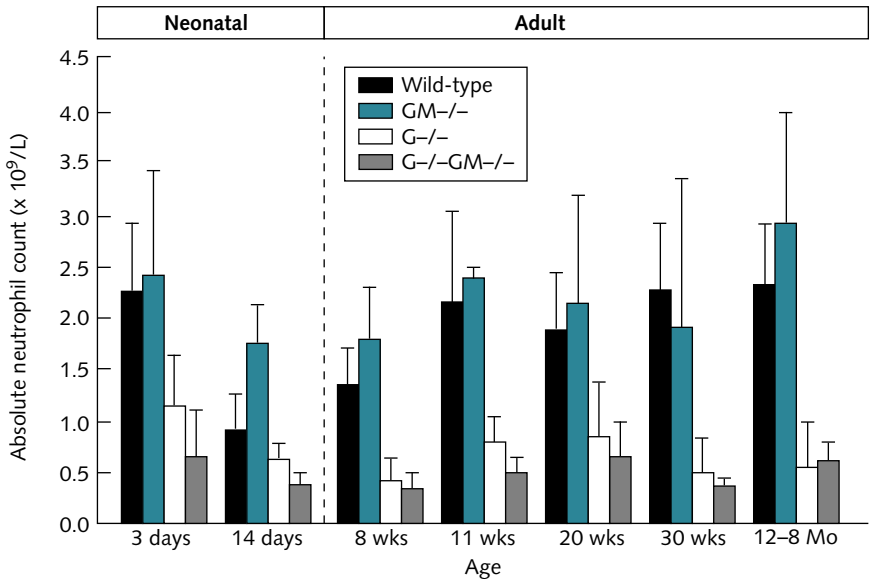


Fig. 22.5 Baseline neutrophil numbers in mice lacking G-CSF, GM-CSF or both factors
This indicates the reduction in basal neutrophil numbers in the G-CSF knockouts and the relatively normal neutrophil numbers in GM-CSF knockout mice. The ability to respond to infection by elevating neutrophil numbers differs between the G-CSF and GM-CSF knockouts.

quantitative balance of neutrophil production during steady-state granulopoiesis *in vivo*, and indicates that GM-CSF has a role in emergency granulopoiesis during infection. Data reported from double knockout (*G-CSF*$^{-/-}$,*GM-CSF*$^{-/-}$) mice confirmed this hypothesis (Figure 22.5).

It is interesting to note that, in mice in which the gene for thrombopoietin has been eliminated by homologous recombination (*Tpo*$^{-/-}$), the number of platelets and megakaryocytes is less than 20% of that in normal mice (*Tpo*$^{+/+}$) (reviewed in Murone *et al.*, 1998). This paper describes a classic knockout experiment in which heterozygotes had platelet counts 60% of normal. This gene dosage effect is the strongest evidence yet that the thrombopoietin gene is not regulated by the platelet count. These thrombopoietin-deficient mice grew and bred normally and had normal hemostasis.

Because megakaryocytes were present and produced platelets, it appears that thrombopoietin is not necessary for megakaryocyte differentiation or platelet shedding. Rather, it serves as an amplification system for the production of platelets.

The constitutive expression of thrombopoietin is confirmed by measured thrombopoietin mRNA levels in liver and kidneys of thrombocytopenic animals. These measurements show that there is no change in transcription rate despite a greater than 20-fold rise in the concentration of thrombopoietin in the circulation. This principle was confirmed in mice in which the thrombopoietin receptor had been eliminated by homologous recombination (m-*MPL*$^{-/-}$). Despite having a platelet count 15–20% of normal and markedly elevated thrombopoietin levels, there was no change in the thrombopoietin mRNA levels in liver and spleen in these animals. The constitutively produced thrombopoietin is bound by the avid thrombopoietin receptors on platelets and rapidly cleared from the circulation. Proof of this comes from the recent demonstration that platelets from normal, but not c-MPL-deficient, mice bound radioactively labeled thrombopoietin, which was rapidly internalized and degraded. When normal murine platelets were infused into c-*MPL* knockout mice, the elevated thrombopoietin levels dropped to normal values within 2 hours.

This type of gene-targeting experimentation will undoubtedly provide many further advances in our understanding of hematopoietic regulation in the years ahead.

Drug development

The CSFs are appealing as therapeutic agents because they are used to recapitulate, on a pharmacological basis, what they are known to effect on a physiological scale. The limiting feature of HGFs is that other, less-defined processes may also be triggered.

There is often great species specificity in *in vitro* and *in vivo* effects. For this reason, preclinical development of HGFs may rely increasingly on the use of non-human primates and SCID (severe chronic immune deficiency) mice. As mentioned earlier, one of the major barriers in HGF research is the species specificity of these proteins, in addition to the antigenicity that they, or the products associated with their expression vector, may elicit.

The design of preclinical studies, especially for cytokines with pleiotropic actions, must carefully parallel the design of the proposed initial clinical trial, as results obtained with a

recombinant human HGF administered to other species may not be transferable to the clinic.

An additional challenge as a result of the administration of HGFs as pharmacological agents is the possibility of eliciting both the clinical effects of the direct action of the factor and indirect actions due to the activation of other cytokines or networks. As stated, over the last few years a large number of *HGF* genes have been cloned and their products produced in bacteria or yeast. Although all of these molecules merit discussion, we will focus our attention on the CSFs in clinical use.

Erythropoietin

The recombinant form of EPO was one of the first marketed recombinant HGFs. Native EPO is a glycoprotein hormone that regulates the production of red blood cells by acting on committed erythroid progenitor cells. There are two forms of EPO marketed in Europe: epoetin alpha and epoetin beta, which have different carbohydrate content and diluent constituents. The greatest benefit of EPO is its ability to correct the anemia of chronic renal failure and the treatment of chemotherapy-induced anemia. Though it is not understood precisely what the relationship is, it would appear that one of these forms of EPO has been recently associated with the development of a serious side effect. Pure red cell aplasia (PRCA) is a rare disease historically associated with thymoma. However, PRCA has been documented in over 140 European patients receiving Eprex® for anemia secondary to renal insufficiency. The development of PRCA has been shown to be associated with the development of antibodies to Eprex which cross-reacted with and neutralized endogenous EPO, resulting, paradoxically, in a progressive anemia much more serious than that which the patient had had originally, and one that is now refractory to EPO therapy. These patients are likely to remain dependent upon blood transfusions for the rest of their lives.

A second-generation derivative of EPO, darbepoetin alfa, has been developed and is now available for patient use. Darbepeotin alfa is an analog of EPO into which new glycosylation sites have been engineered. The resulting molecule contains five N-linked carbohydrate chains (in contrast to three in the prototype molecule) and remains resident in the circulation for three times longer, resulting in prolonged stimulation of erythropoiesis; consequently fewer injections are required (Figure 22.6). It remains to be seen what further advances will be made with what is the most mature of all growth factor therapeutics.

Granulocyte colony-stimulating factor

G-CSF is the major growth factor involved in the production of neutrophilic granulocytes. Severe congenital neutropenia and acute myeloblastic leukemia are both characterized by arrest of maturation in the myeloid lineage. Although marrow cells from patients with severe congenital neutropenia frequently show reduced responsiveness to G-CSF, neutrophilic colony formation *in vitro*, as well as neutrophil production *in vivo*, can usually be induced with G-CSF. Patients with severe

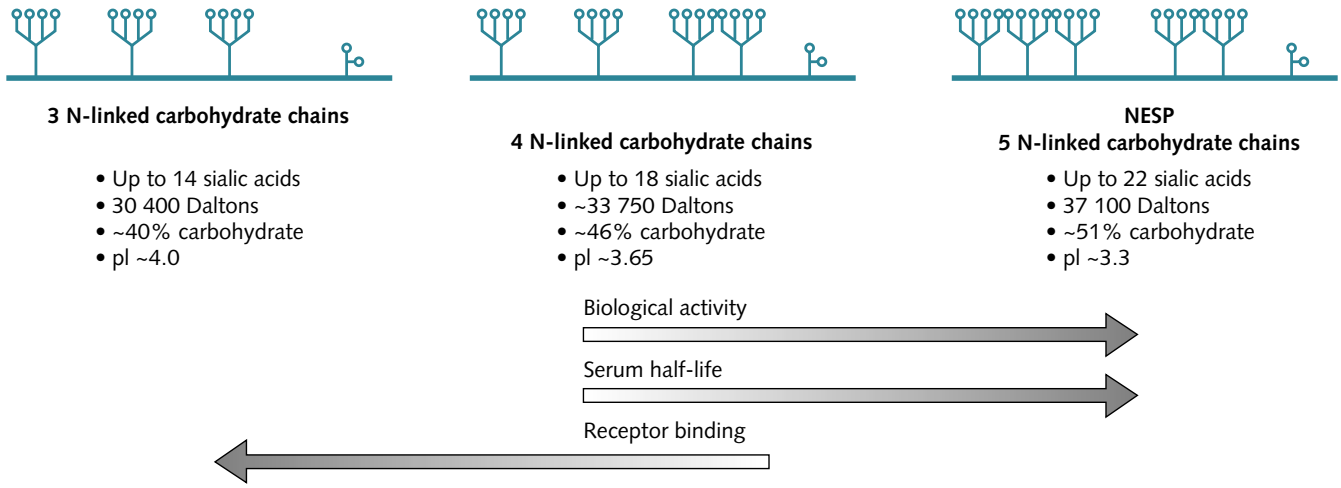


Fig. 22.6 Biochemical and biological properties of rHuEPO and rHuEPO analogs containing four and five N-linked carbohydrate chains
Redrawn from Egrie JC, Browne JK. (2001) Development and characterization of novel erythropoiesis stimulating protein (NESP). *British Journal of Cancer* **84** (S1), 3–10. With permission.

congenital neutropenia have an increased risk of developing acute myeloblastic leukemia, which suggests that the same functionally related defects may be involved in the pathogenesis of the disease.

G-CSF has been extensively studied *in vitro* and in clinical trials. Many benefits to cancer patients are attributed to the use of this factor, including reductions in febrile neutropenia, documented infections, intravenous antibiotic use and hospitalizations due to infections. Other benefits are reduction in mucositis, which hampers eating, reduction in antibiotic-induced diarrhea, the ability to administer full-dose chemotherapy on time, completion of chemotherapy sooner, and the ability to intensify chemotherapy (Neupogen® package insert; Amgen, Thousand Oaks, California, USA).

Neupogen (rmetHuG-CSF) is approved for the support of myelosuppression in severe congenital neutropenia and as a result of chemotherapy in acute myeloid leukemia, non-myeloid malignancies and bone marrow transplantation. In addition, it is approved for the mobilization of peripheral blood progenitor cells (PBPC). More than 2.2 million patients have received this product and it is approved in 46 countries.

Recently, a novel form of Neupogen has been developed for use in a similar patient population: pegfilgrastim or Neulasta®. G-CSF is eliminated from the body by two major processes: renal filtration and neutrophil-mediated processes. It was hypothesized that if the hydrodynamic size of a G-CSF conjugate were to be increased to a level above the threshold for renal clearance, then the only processes left by which the conjugate could be cleared from the body would be neutrophil-mediated. Pegfilgrastim represents such a conjugate, where a 20-kDa poly(ethylene glycol) (PEG) has been chemically attached to the N-terminus of the protein. This relatively large molecule is resistant to renal elimination yet retains the ability to stimulate neutrophil production. Thus, in cancer patients made neutropenic by cytotoxic chemotherapy it has been shown that pegfilgrastim can remain in the body for a substantial time, during which it stimulates the recovery of infection-fighting neutrophils. As neutrophil recovery proceeds and the risk of infection recedes, the neutrophils themselves accelerate destruction of pegfilgrastim (Figure 22.7). The drug is therefore unusual among therapeutic agents in that it has the ability to correct the deficiency for which it is administered, only then to be removed by the products of that very recovery. It thus represents a self-regulating drug; it was suggested above that such a drug may be possible, but pegfilgrastim represents the first proof that such a drug could indeed be made. It is now approved for use in several countries, including EU countries, the USA and Australia.

Granulocyte-macrophage colony-stimulating factor

Despite the similarity in name between G-CSF and GM-CSF (sargramostim, Leukine®) and primary clinical indications that do not differ grossly between the two, it is perhaps surprising that these molecules are very distinct from one another. For instance, in preclinical work G-CSF crosses species easily, whereas GM-CSF does not—the protein and species have to be matched. In addition, animals in which the GM-CSF gene, or its receptor, have been disabled or knocked out show normal neutrophil numbers, unlike the 20% of normal number seen in G-CSF knockout animals. GM-CSF knockout animals also show a lung pathology typified by the accumulation of a lipid-rich proteinaceous material in the alveoli due to the lack of macrophage-derived surfactants. Thus, it would appear that GM-CSF is more a regulator of mature phago-

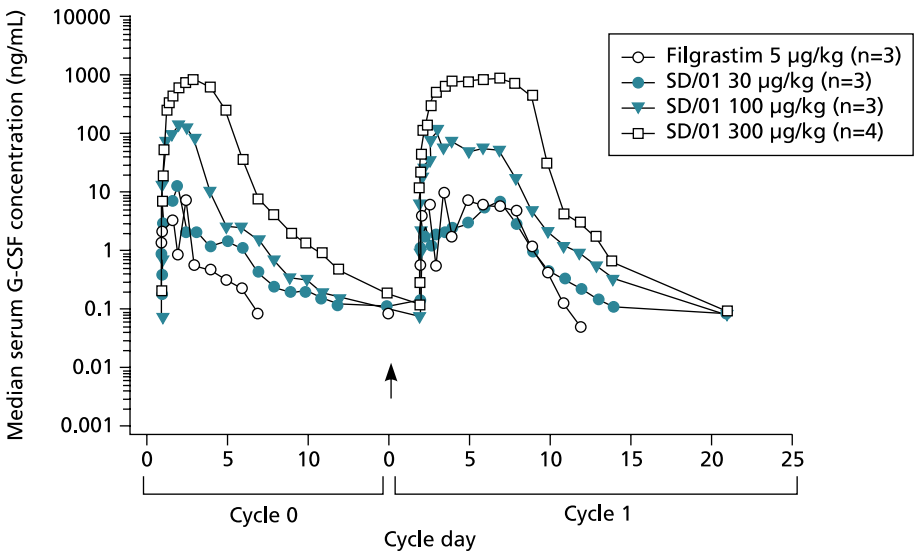


Fig. 22.7 Median G-GSF concentration in patients receiving G-CSF or polyethylene glycol-conjugated G-CSF
Redrawn from Johnston E, Crawford J, Blackwell S *et al.* (2000) Randomized, dose-escalation study of SD/01 compared with daily filgrastim in patients receiving chemotherapy. *Journal of Clinical Oncology*, **13**, 2522–2528. With permission.

cytic cell function than of leukocyte production. However, clinically GM-CSF is used to increase neutrophil numbers in cancer patients undergoing chemotherapy in much the same way as G-CSF. It also mobilizes PBPC like G-CSF does, and may add benefits such as modulation of immune function via its effects on monocytes.

c-MPL ligand or thrombopoietin

The MPL ligands are a family of closely related HGFs that bind to the thrombopoietin receptor, c-MPL. In addition to the endogenous MPL ligand, thrombopoietin, two recombinant MPL ligands, recombinant thrombopoietin and pegylated megakaryocyte growth and development factor (PEG-MGDF), are under investigation.

When recombinant thrombopoietin or PEG-MGDF is administered to normal animals or humans, there is a dose-dependent increase in the platelet count, but no effect on leukocytes or erythrocytes. When administered after chemotherapy in animal models of humans, MPL ligands reduce the duration, and sometimes the degree, of thrombocytopenia. The MPL ligands may also be effective in reducing the thrombocytopenia in patients with HIV infection, liver disease or myelodysplasia, or for plateletpheresis.

The c-MPL ligand has moved rapidly through its preclinical studies and initial trials in human subjects. It appears to be a very powerful agent for the treatment of thrombocytopenia. The preclinical animal models clearly demonstrate that MPL ligands have modest multilineage effects as well as their major thrombopoietic effect; this response is, in part, at the level of the marrow. Numerous *in vitro* studies have demonstrated that recombinant thrombopoietin can enhance the proliferation of erythroid and myeloid progenitors.

Detailed examination of both TPO-l-deficient mice and c-MPL-deficient mice also demonstrated the cross-lineage effects of thrombopoietin. This may be the best demonstration of the multilineage effect of thrombopoietin, as the bone marrow precursors of megakaryocytes were decreased in all of the mice, but precursors of all other lineages were also decreased in these mice despite no effect on the leukocyte and erythrocyte counts. Administration of recombinant thrombopoietin to the thrombopoietin-deficient mice, but not to the c-MPL-deficient mice, corrected the precursor cell defect.

Interleukin-11

IL-11 is a good example of a growth factor with pleiotropic actions which has been developed for clinical use in thrombocytopenia (Figure 22.8).

Human and mouse IL-11 share 88% amino acid identity and are reactive across species. Similar in sequence and gene structure to IL-6, G-CSF, oncostatin-M (OSM), leukemia inhibitory factor (LIF) and growth hormone, IL-11 signals through gp130, which was originally identified as a component of the IL-6 receptor.

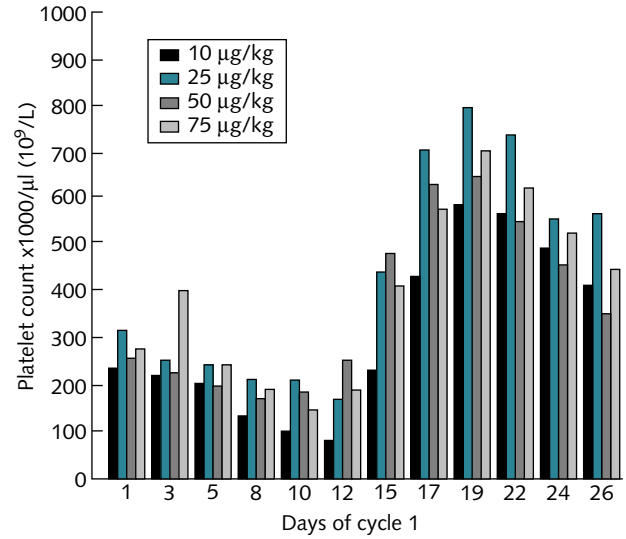


Fig. 22.8
The ability of interleukin-11 to raise platelet counts after cancer chemotherapy has led to its approval for this use. IL-11 is a pleiotropic cytokine and has many systemic effects in addition to improving platelet numbers.

Administration of IL-11 to normal mice causes an increase in circulating platelet levels, which is probably mediated via its demonstrated effects on megakaryocyte maturation alone or in combination with other growth factors, such as IL-3 and SCF. This ability has led to a development program culminating in Food and Drug Administration (FDA) approval for IL-11 in thrombocytopenia. Despite the appearance of relative lineage fidelity implied by its effects upon predominantly a single circulating cell type, the discovery of IL-11 was facilitated by its actions upon a different cell type: plasmacytoma cells. Among the other properties of IL-11 are abilities to weakly or indirectly influence B-cell development, the augmentation of antibody responses, the induction of hepatic acute phase proteins, the promotion of erythroid colony growth *in vitro*, the protection of intestinal mucosa from cytotoxic damage *in vivo*, influences upon bone metabolism and the inhibition of adipogenesis. Among these myriad effects, boosting platelet numbers led most quickly to registration; whether other properties of this material will eventually lead to other indications is as yet unclear. Nonetheless, IL-11 is a good example of a growth factor which was isolated and purified on the basis of one activity yet finally found a therapeutic niche in a different setting. How the growth factor is used clinically may also be expected to change in the face of the development of what is widely held to be the true physiological regulator of platelet production — thrombopoietin. Perhaps the other distinctive

properties of IL-11 will support continuation of its use, or possibly its lifetime as a front-line platelet factor will be limited by the emergence of thrombopoietin-like regulators.

Stem cell factor

SCF enhances the proliferation of early hematopoietic cells, especially in combination with a variety of agents, particularly MPL ligands such as thrombopoietin and MGDF. It works in concert with other HGFs to potentiate the effects of these cytokines, with the resultant hematopoietic phenotype determined by the precise combination.

Mobilization of stem cells

Although SCF in combination with other growth factors can promote the proliferation and development of primitive multipotent and putative stem cells *in vitro*, it was somewhat surprising to find that SCF *in vivo* (even by itself) could induce significant emigration of these and other progenitor cells from the bone marrow to the blood (Figure 22.9). Moreover, data also suggest that the total body load of primitive progenitor cells increases in both mice and humans treated with SCF.

However, perhaps much more significant from the therapeutic viewpoint was the finding that the interaction between SCF and G-CSF that was observed for progenitor cell development *in vitro* also occurred *in vivo* in terms of mobilization of hematopoietic progenitor cells into the blood. In man, low doses of SCF (≥10 mg/kg per day) in combination with G-CSF mobilize approximately three-fold more $CD34^+$ cells than G-CSF alone, confirming data obtained in primates (Figure 22.10). Furthermore, the quality of these cells in respect to their capacity to engraft lethally irradiated animals appears to show an improvement over that seen with G-CSF alone. In other words, the combination of SCF and G-CSF preferentially mobilizes the most primitive hematopoietic cells. This has been confirmed in humans, at least with respect to multipotential and megakaryocyte precursors.

The therapeutic potential of these findings is obvious, particularly since equivalent cells, induced by more conventional agents such as chemotherapy and/or G-CSF or GM-CSF, have proved useful in transplant settings, both as a means of hematopoietic support for myelosuppressed patients and as a source of cells capable of re-establishing hematopoiesis in myeloablated animals and patients. In fact, the speed of recovery of, for example, platelets after transplantation either with PBPC alone or in support of marrow transplantation, has indicated that PBPC offer superior performance to equivalent marrow populations, particularly in the short term.

Clinical adoption

HGFs can reduce the morbidity and possibly the mortality from some types of cancer treatment and from other HGF-deficiency states. Reductions in hospitalization and supportive care (transfusion requirements and antibiotics) have been documented in several trials and can reduce total care costs. However, the cost is substantial, and these drugs are one of the many expensive new pharmacy purchases, prompting payers to ask whether the cost is warranted by the benefit.

Economic and clinical analyses of CSFs have mostly been complementary. The use of CSFs is justified on economic grounds in those cases in which the most clinical benefit has

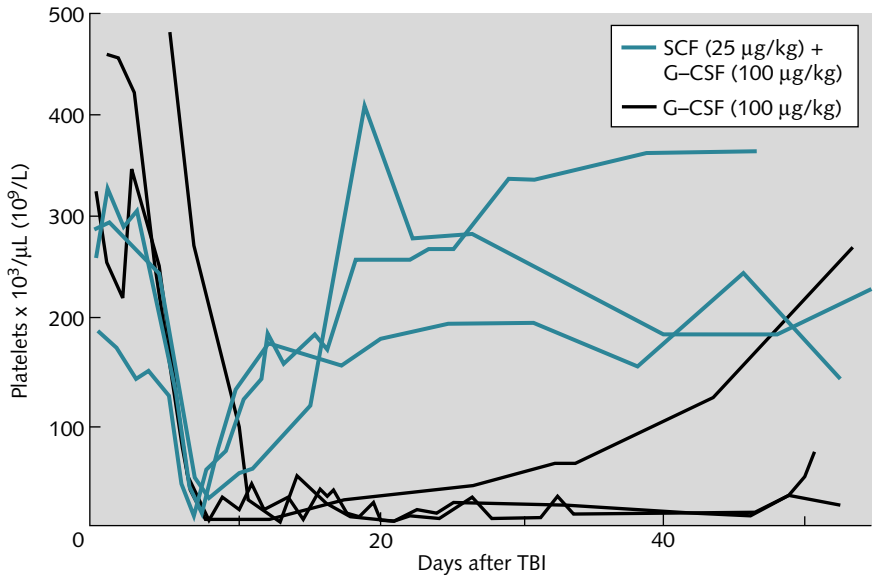


Fig. 22.9
The ability of SCF to enhance the *in vivo* response to G-CSF has been documented in several preclinical models. Here, the ability of G-CSF to mobilize hematopoietic progenitor cells into the blood is increased by the inclusion of SCF in baboons.

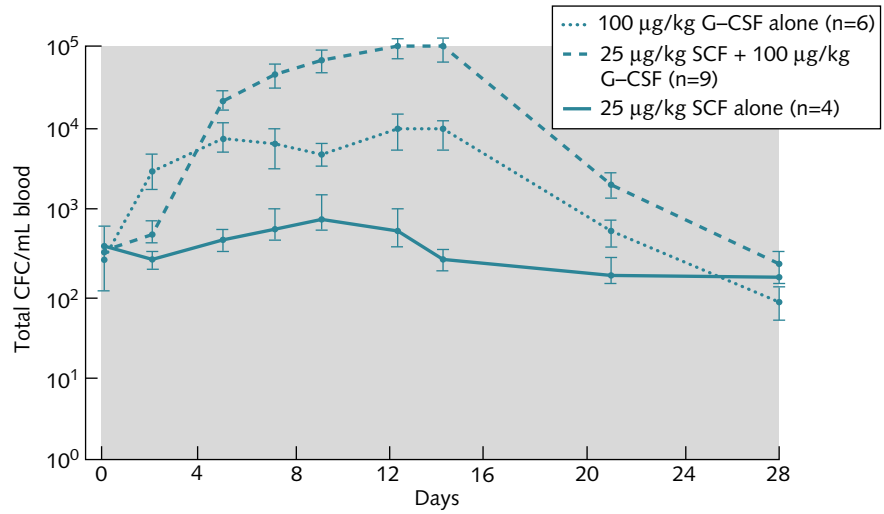


Fig. 22.10
The progenitor cells mobilized by G-CSF can successfully engraft a transplanted baboon. The cells mobilized by G-CSF + SCF promote more rapid recovery, probably as a result of the increased numbers seen in Figure 22.9, but there is also indirect evidence that the cells are of higher quality.

been observed; that is, for primary prophylaxis of febrile neutropenia when the rate is high, for peripheral stem cell mobilization, and for hematopoietic reconstitution after stem cell transplantation.

The high acquisition costs of HGFs have attracted substantial attention from purchasers, providers and health economists. Initial economic studies evaluating the use of growth factors demonstrated that their use is cost-saving when the risk of febrile neutropenia is >40%. In these cases the savings in hospitalization and treatment costs are larger than the drug acquisition costs. At lower risk levels the use of HGF has been shown to partially offset the drug acquisition costs, the level of the offset depending on the level of risk for febrile neutropenia, the daily charge for hospitalization and the duration of febrile neutropenia.

The American Society for Clinical Oncology (ASCO) published clinical practice guidelines for the use of growth factors in 1994 which recommended the use of HGF when the risk of febrile neutropenia was >40%, coinciding with the results of the economic analysis. Revised guidelines from ASCO issued in 1996 did not change the recommendation for HGF use and have not been updated since that date.

Lyman revised his initial cost estimates in 1996 (*see* American Society of Clinical Oncology, 1996). The new analysis added to the direct costs of the treatment of febrile neutropenia the indirect and intangible costs. When indirect costs are included, HGF costs are offset when the risk of febrile neutropenia is at about 30%, and when intangible costs are added HGF costs are offset when the risk of febrile neutropenia is approximately 20%. Recent economic analysis of HGF use has shifted the emphasis from cost-offset to cost-effectiveness studies, in which costs of treatment are balanced against clinical outcomes. Silber and colleagues in 1998 modeled the use of growth factors in early breast cancer patients, in which standard-dose chemotherapy is associated with survival benefit. The use of growth factors for the 50% of patients most at risk of chemotherapy dose modification was associated with a cost per year of life saved of about $35 000, well within the range of cost-effectiveness ratios considered to provide good economic value in the delivery of conventional medical care.

Conclusions

The development of recombinantly produced endogenous proteins for pharmacological delivery remains one of the most important medical advancements of the decade. In the present millennium, the promise of exploiting knowledge of the human genome for additional novel proteins with therapeutic potential looms large. Additional proteins or small-molecule mimics may require some engineering to optimize their use as therapeutics, especially those whose natural function may not be the same therapeutic target. In addition, evolution of protein therapeutics is likely to continue towards 'designer cytokines', as it is recognized that the recombinant equivalent of a cytokine can invariably be improved upon when employed as therapeutic. Continued mining of the genome combined with molecular techniques should provide therapeutics for many important diseases in the near future.

Further reading

General

American Society of Clinical Oncology. (1996) Update of recommendations for the use of hematopoietic colony stimulating factors:

evidence-based clinical practice guidelines. *Journal of Clinical Oncology*, **14**, 1957–1960.

Burgess A, Metcalf D. (1980) Characterization of a serum factor stimulating the differentiation of myelomonocytic leukemic cells. *International Journal of Cancer*, **319**, 415–418.

Cebon J, Layton JE, Maher D *et al.* (1994) Endogenous haemopoietic growth factors in neutropenia and infection. *British Journal of Haematology*, **86**, 265–274.

LeBeau M, Lemons R, Espinosa R. (1989) Interleukin-4 and interleukin-5 map to human chromosome 5 in a region encoding growth factors and receptors and are deleted in myeloid leukemias with del(5q). *Blood*, **73**, 647–650.

Erythropoietin

Besarah A, Flaharty KK, Erslev AJ *et al.* (1992) Clinical pharmacology and economics of recombinant human erythropoietin in end-stage renal disease: the case for subcutaneous administration. *Journal of the American Society of Nephrology*, **2**, 1405–1416.

Casadevall N, Nataf J, Viron B *et al.* (2002) Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. *New England Journal of Medicine*, **346**, 469–475.

Egrie JC, Browne JK. (1991) The molecular biology of erythropoietin. In: Erslev AJ, Adamson JW, Eschbach JW (eds). *Erythropoietin: Molecular, Cellular and Clinical Biology*. Baltimore: Johns Hopkins University Press, pp. 21–40.

Elliott S, Lorenzini T, Asher S *et al.* (2003) Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nature Biotechnology*, **21**, 414–421.

Eschbach JW, Abdulhadi MH, Browne JK *et al.* (1989) Recombinant human erythropoietin in anemic patients with end-stage renal disease; results of phase III multicenter clinical trial. *Annals of Internal Medicine*, **111**, 992–1000.

Eschbach JW, Kelly MR, Haley NR *et al.* (1989) Treatment of the anemia of progressive renal failure with recombinant human erythropoietin. *New England Journal of Medicine*, **321**, 158–163.

Evans RW, Rader B, Mannienen DL, Cooperative Multicenter IPO Clinical Trial Group. (1990) The quality of life of hemodialysis patients treated with recombinant human erythropoietin. *Journal of the American Medical Association*, **263**, 825–830.

Jacobs K, Shoemaker C, Rudersdorf R *et al.* (1985) Isolation and characterization of genomic and cDNA clones of human erythropoietin. *Nature*, **313**, 806–810.

Lin F, Suggs S, Lin C *et al.* (1985) Cloning and expressions of the human erythropoietin gene. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 7580–7584.

G/GM-CSF

Burgess AW, Begley CG, Johnson GR *et al.* (1987) Purification and properties of bacterially synthesized human granulocyte-macrophage colony-stimulating factor. *Blood*, **69**, 43–51.

Cantrell MA, Anderson D, Cerretti DP *et al.* (1985) Cloning, sequence, and expression of a human granulocyte/macrophage colony-stimulating factor. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 6250–6254.

Ernst T, Ritchie A, Demetri G, Griffin J. (1989) Regulation of granulocyte and monocyte colony stimulating factor mRNA levels in human blood monocytes is mediated primarily at a post-transcriptional level. *Journal of Biological Chemistry*, **264**, 5700–5703.

Johnston E, Crawford J, Blackwell S *et al.* (2000) Randomized, dose-escalation study of SD/01 compared with daily filgrastim in patients receiving chemotherapy. *Journal of Clinical Oncology*, **18**, 2522–2528.

LeBeau MM, Westbrook CA, Diaz MO *et al.* (1986) Evidence for the involvement of GM-CSF and FMS in the deletion (5q) in myeloid disorders. *Science*, **231**, 984–987.

Molineux G, Kinstler O, Briddell B *et al.* (1999) A new form of Filgrastim with sustained duration in vivo and enhanced ability to mobilize PBPC in both mice and humans. *Experimental Hematology*, **27**, 1724–1734.

Nagata S, Tauchiya M, Asano S *et al.* (1986) Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor. *Nature*, **319**, 415–418.

Rapoport AP, Abboud CN, DiPersio JF. (1992) Granulocyte-macrophage colony-stimulating factor (GM-CSF) and granulocyte colony-stimulating factor (G-CSF): receptor biology, signal transduction, and neutrophil activation. *Blood Reviews*, **1**, 43–57.

Trapnell BC, Whitsett JA. (2002) GM-CSF regulates pulmonary surfactant homeostasis and alveolar macrophage-mediated innate host defense. *Annual Review of Physiology*, **64**, 775–802.

Zhan Y, Lieschke GJ, Grail D *et al.* (1998) Essential roles for granulocyte-macrophage colony-stimulating factor (GM-CSF) and G-CSF in the sustained hematopoietic response of Listeria monocytogenes-infected mice. *Blood*, **91**, 863–869.

Stem cell factor

Bodine D, Seidel N, Zsebo K *et al.* (1993) *In vivo* administration of stem cell factor to mice increases the absolute number of pluripotent hematopoietic stem cells. *Blood*, **82**, 445–455.

Duarte RF, Frank DA. (2000) SCF and G-CSF lead to the synergistic induction of proliferation and gene expression through complementary signaling pathways. *Blood*, **96**, 3422–3430.

Molineux G, Migdalska A, Smitkowski M *et al.* (1991) The effect on hematopoiesis of recombinant stem cell factor (ligand for c-kit) administered *in vivo* to mice either alone or in combination with granulocyte colony-stimulating factor. *Blood*, **78**, 961–966.

Okada S, Nakauchi H, Nagayoshi K *et al.* (1992) *In vivo* and *in vitro* stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood*, **80**, 3044–3050.

Tong J, Gordon M, Srour E *et al.* (1993) *In vivo* administration of recombinant methionyl human stem cell factor expands the number of human marrow hematopoietic stem cells. *Blood*, **82**, 784–791.

Thrombopoietin

Bartley TD, Bogenberger J, Hunt P *et al.* (1994) Identification and cloning of a megakaryocyte growth and development factor that is a ligand for the cytokine receptor mpl. *Cell*, **77**, 1117–1124.

de Sauvage FJ, Hass PE, Spencer SD *et al.* (1994) Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand. *Nature*, **369**, 533.

Kaushansky K, Lok S, Holly RD *et al.* (1994) Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. *Nature*, **369**, 568.

Lok S, Kaushansky K, Holly RD *et al.* (1994) Cloning of murine thrombopoietin cDNA and stimulation of platelet production *in vivo*. *Nature*, **369**, 565.

Murone M, Carpenter DA, De Sauvage FJ. (1998) Hematopoietic deficiencies in c-mpl and TPO knockout mice. *Stem Cells*, **1**, 1–6.

Silber JH, Fridman M, Shpilsky A *et al.* (1998) Modeling the cost-effectiveness of granulocyte colony-stimulating factor use in early-stage breast cancer. *Journal of Clinical Oncology*, **16**, 2435–2444.

Wendling F, Maraskovsky E, Debill N *et al.* (1994) c-Mpl ligand is a humoral regulator of megakaryopoiesis. *Nature*, **369**, 571.

Chapter 23 Molecular therapeutics in hematology

A Keith Stewart & Jeffrey A Medin

Introduction

Clinical investigation of gene transfer in humans began with the seminal trial of Rosenberg and colleagues (1990) involving transplantation of genetically altered lymphocytes into patients, and was closely followed by studies of therapeutic gene transfer using human bone marrow. Studies involving the blood system have therefore been pivotal in the early development of human gene therapy. Gene marking or gene therapy protocols are now under intensive investigation worldwide, with more than 250 trials by the end of year 1997 and more than 500 trials by the end of 2001. Importantly, some of the first successes in the entire field of gene therapy have recently been realized in this area. Because hematology has thus contributed so much to the genesis of human gene therapy, it is the purpose of this chapter to reiterate both the inherent promise and the remaining obstacles posed by the application of gene transfer using hematopoietic cells.

Gene transfer strategies

An ever-increasing array of techniques has been described to facilitate gene transfer into blood cells. These techniques can be broadly grouped as *physical methods* and *viral vectors*.

Physical means of gene delivery

Physical methods of gene transfer are generally of low efficiency and provide only transient gene expression in the absence of selection. The advantages of these systems are that genes are transferred without viral sequences, which may affect the biology of the target cell or the host. In addition, large or multiple genes can be transferred and gene transfer is independent of the proliferative status and cell surface receptor profile of the target cell. Those methods of physical gene transfer which have been used in the clinical setting include *electroporation* (using electric fields which create channels in cell membranes, allowing passage of DNA into the cell), *particle bombardment* (microscopic gold beads labeled with DNA which are forced through the cell membrane by CO_2-driven pressure) and *liposomal encapsulation* of DNA. Liposomal-mediated gene delivery has gained most acceptance in the clinical arena, driven by the higher gene transfer efficiencies obtained with newer liposomal formulations and by safety considerations as an alternative to viral vector-mediated gene delivery. Other nonviral gene delivery systems include the use of plasmid DNA alone (so-called naked DNA), the use of synthetic polymers, and the adaptation of bacterial gene delivery systems. The use of plasmid DNA alone is particularly efficacious in specific tissues such as muscle. Thus, DNA-based gene delivery is gaining in popularity for some functions in malignancy and infectious disease, and in applications directed towards the vascular bed. In cancer, DNA-based vaccines are in clinical trials in which the immunogenic antigen is encoded by the plasmid itself with or without an added immunostimulatory gene. Yet, while this approach is promising, clearly more work on modulating the biology of the immune response is needed. For example, in one very recently published report, Phase I/II clinical trials were performed using plasmids constructed to express chimeric idiotype proteins to upregulate immunotherapy for B-cell lymphomas. Unfortunately, while the naked DNA vaccination approach was well tolerated, little specific immune response was generated even when a potent costimulatory cytokine was also introduced.

Physical methods such as electroporation, direct DNA transfer and liposome-mediated DNA transfer have been used with varying levels of success in transferring genes into

hematopoietic cells. However, genes rarely integrate with physical methods of transfer and the vectors are lost in multiplying cells, and thus are generally of little value in long-term stable gene transfer applications. Nevertheless, when transient gene expression is sufficient, such as in cancer immunotherapy applications, physical methods of gene transfer may find a niche.

Viral vectors for gene transfer

Viral gene transfer methods take advantage of the normal virus life cycle to transfer genetic material into a host cell. For the synthesis of gene transfer vectors, the viral genome is modified largely by deleting most viral genes and inserting therapeutic or marker genes in their place. The viral gene products necessary for the production of recombinant viral particles are then provided *in trans* by packaging cell lines which have been transfected with plasmids carrying the missing viral genes. Recombinant viral vectors are then transfected into the packaging cells and replication-incompetent recombinant viral particles are produced. Replication-incompetent viruses are safe for clinical applications because they transduce target cells only once, and are unable to cause a secondary infection because of the absence of secondary expression of viral genes required for replication (Figure 23.1).

The most commonly used viral backbones in clinical trials are based on the murine onco-retroviruses and human adenoviruses; however, herpes, pox, vaccinia, adeno-associated viral (AAV) and lenti-retroviral (lentivirus) vectors are also being developed and evaluated. The relative merits and disadvantages of each vector system are outlined in Table 23.1.

Retroviral gene transfer systems

Retroviruses (including onco-retroviruses and lenti-retroviruses) are double-stranded RNA viruses. The viral RNA genome is reverse transcribed into double-stranded DNA, which integrates into the genome of target cells in a fairly random manner, although regions of active gene transcription are preferred. Currently, retroviral vectors derived from the Moloney murine leukemia virus (MMLV) and other murine onco-retroviruses are used for clinical gene transfer protocols targeting hematopoietic cells. Recently, however, recombinant lenti-retroviruses (such as those derived from 'gutted' HIV, for example) have been used in the laboratory for reasons discussed below; to date no clinical trials have yet been approved using these alternative retroviral delivery agents.

In gene transfer vectors the wild-type onco-retroviral genome is modified by deleting most of the *gag* and all *pol* and *env* sequences. Viral sequences which are retained in the onco-retroviral vector include the packaging signal and the long terminal repeats (LTRs) which are necessary for viral integration and often drive transcription of the marking or therapeutic transgene in the absence of a heterologous promoter. These modifications render the vectors replication-incompetent while making 6–8 kb (in recombinant onco-retroviruses at least) available for the insertion of desired genes.

Onco-retroviral packaging cell lines have been engineered to minimize the chance of accidental replication competent retrovirus (RCR) production. Plasmids used to create stable packaging cell lines that express key components of the virus are based on the wild-type onco-retroviral genome but have extensive deletions and split genes. These modifications have minimized recombination-prone homologous sequences between the gene transfer vector and the stable packaging cells, thereby significantly reducing the chance of RCR. Another measure of safety is naturally present in onco-retroviral gene transfer systems as these vectors, when packaged in murine-based packaging cell lines, are rapidly inactivated by human serum. Lastly, in addition to the safety features described above, all viral supernatants and patient cell samples that have been exposed to clinical grade onco-retroviral supernatant are extensively tested for the presence of RCR and other possible contaminants prior to infusion into patients.

In addition to the relative safety of onco-retroviral vectors, another advantage of this delivery system includes the stable integration of the gene delivery vector into the host cell genome, allowing long-term transgene expression in the host cell and its progeny. Disadvantages include the requirement of the target cells to be in cycle for integration, the randomness of the integration event itself, the specificity of receptor binding, which can reduce infection rates into some target cells, and the potential silencing of transgene expression. Since long-term expression of transgenes is a requirement for most hematopoietic stem cell (HSC)-based gene therapies, many viral vectors use viral promoters, such as the endogenous MMLV/LTR, which can direct high levels of expression. Nevertheless, some studies, especially those using ES cells, have shown that such promoters can be silenced *in vivo*. There are several potential mechanisms of such transgene downregulation in HSCs and their progeny. These include cytokine-mediated promoter suppression, DNA methylation, an unstable integration environment, and contextual position of the transgene within the onco-retroviral vector (internal promoter competition). Another potentially overriding possibility is that perhaps culture conditions for *ex vivo* cell transductions have caused primate HSCs to fully differentiate, therefore transgene expression is lost by the maturation and death of provirus-positive cells over time. Thus, even if efficient gene transfer is achieved, transgene expression, and thus the therapeutic impact, may be low. Newer-generation onco-retroviral vectors and improved *ex vivo* transduction methods, however, have led to significantly higher and longer

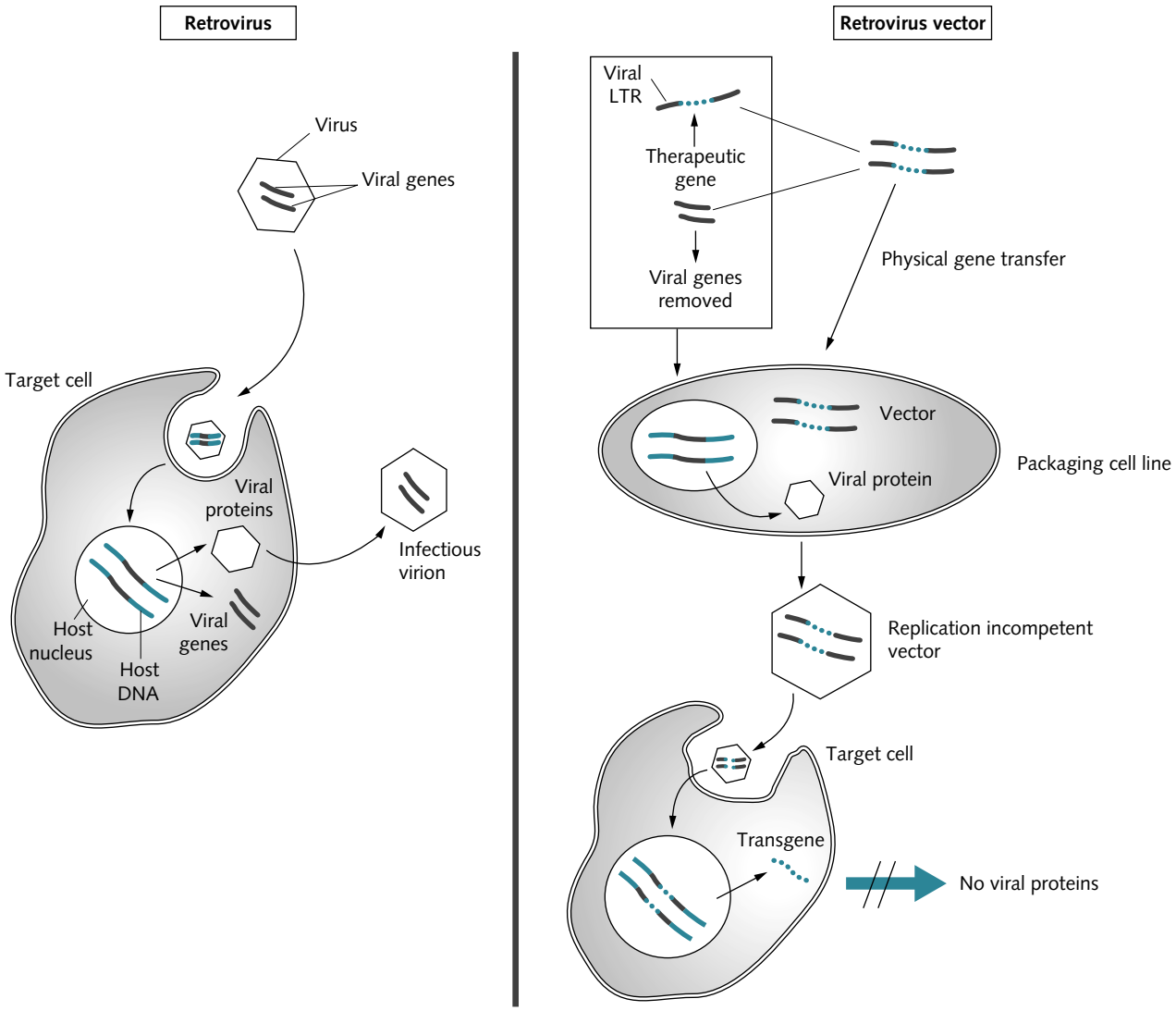


Fig. 23.1 Biology of retroviruses and packaging cell lines
After random integration into the host genome, viral proteins are produced which encapsulate viral RNA. In a retroviral vector the *gag*, *pol* and envelope-encoding genes have been removed, allowing space for cloning of the gene of interest. To make an infectious virus, the vector is introduced into a packaging cell line expressing *gag*, *pol* and *env* genes. The packaged vector can then bind to, and be internalized by, a target cell, where integration into the host genome occurs and the vector transgene is produced. Since the target cell lacks the genes necessary to make a new virion, the vector is not infectious.

transgene expression than was observed previously. For example, MSCV, MND and MFG vectors have been designed with multiple modifications to the MMLV backbone, which enable genes to be expressed more efficiently in murine, non-human primate and human hematopoietic cells.

Retrovirally encoded marking and therapeutic transgenes can also be expressed in a cell-specific manner using specific internal cellular promoters, which allow regulated expression. For example, expression in leukocytes can be directed by CD11b, CD18 or major histocompatibility complex (MHC) I promoters, while expression in erythroid cells can be greatly enhanced by the β-globin promoter and locus control region. Furthermore, viral transgene expression may be regulated by inducible promoter systems such as the *tet* operon system, which uses exogenous tetracycline to specifically transactivate a promoter, leading to ordered expression of the gene of interest. Such regulated and cell-specific systems may prove of utility in therapeutic applications by providing an additional control mechanism for the expression of potentially toxic gene products, for example.

Table 23.1 Some viral vectors used in the application of gene therapy.

	Onco-retrovirus	Adenovirus	Helper-dependent adenovirus	Adeno-associated virus (AAV)	Herpes virus	Lenti-retrovirus	Pox viruses
Advantages	Well studied	Very efficient gene transfer	Large cloning capacity	Cell cycle-independent	Very efficient gene transfer	Integrates	Efficient gene transfer
	Integrates	High transgene expression	Not immunogenic	Safe	Cell cycle-independent	Efficient gene transfer Large capacity	
Disadvantages	Cell cycle-dependent	Immunogenic	Technically cumbersome	Limited cloning capacity	Helper virus contamination	Safety issues	Transient expression
	Promoter silencing	Transient expression	Helper virus contamination	Helper virus contamination			Immunogenic
	Safety issues	Existing immunity	Transient expression	Existing immunity Production	Cytotoxic		
Potential applications	Stem cell gene therapy, e.g. hemophilia, Gaucher, Fabry, immunodeficiencies	Cancer immunotherapy Cystic fibrosis	Gene augmentation therapy	Cancer immunotherapy Gene augmentation therapy via muscle	Cancer immunotherapy	Gene replacement therapy targeting non-dividing cells	Cancer immunotherapy

As discussed above, MMLV-based onco-retroviral vectors only integrate into cells that are actively undergoing mitosis, which precludes their use for delivery to target cells that are non-dividing, such as most HSCs. This has led to the development of replication-incompetent viral vector systems based on lenti-retroviruses (lentiviruses) and human foamy virus, for example, which can integrate into these key non-dividing target cells. Foamy viruses also have a large genome (and hence transgene-carrying capacity), do not cause any known disease, and can be used for *in vivo* gene transfer because they are not effectively inactivated by human serum. Lentiviral vectors also show significant promise as gene transfer vehicles and many laboratories around the world are adopting this alternative delivery system. Current lentiviral vectors contain less than 25% of the HIV-1 genome and are therefore relatively safe. To further increase the safety, many laboratories now use third-generation recombinant lentiviruses that have a self-inactivating 3′-LTR. To further minimize the chance of recombination leading to the production of replication-competent lentivirus, some laboratories even use up to seven separate plasmids in transient transfections of packaging cell lines to generate recombinant virions. Lenti-retroviruses have also been pseudotyped with alternative env proteins, such as vsv-g, which expands the tropism of the virus and even allows concentration of the effective viral titer by ultracentrifugation. There is thus considerable optimism that recombinant lentiviral and/or foamy viral vectors will overcome some of the shortcomings of MMLV-based onco-retroviruses by facilitating delivery of genes of interest to a wide spectrum of cell types, including quiescent HSCs. Preliminary studies show that this may indeed be the case, although this needs to be further substantiated by a number of independent laboratories in long-term, large-animal studies. Furthermore, due to an observed increase in lentivirus-mediated transduction efficiencies over previous methods, it may even be possible to modulate other secondary components of the *ex vivo* gene therapy procedure, such as enhancing engraftment of transduced cells in recipients by minimizing exposure to differentiation-inducing cytokines, for example. Alterations have also been made to lentiviral backbones, such as the addition of a central polypurine tract and a WPRE element (the post-transcriptional regulatory element of woodchuck hepatitis virus) to possibly enhance nuclear import or enhance transgene expression levels, respectively.

Other viral gene transfer systems

Recombinant adenoviral gene transfer systems are commonly used for some gene therapy applications. Adenoviruses infect cells efficiently *in vitro* and *in vivo*, express high levels of

transgene, can infect both cycling and stationary cells, and exhibit wide tissue tropism. Adenoviral vectors have been used to efficiently transfer genes into a variety of human cancers and into hematopoietic progenitor and malignant cells. Nevertheless, circulating B and T lymphocytes have generally proven relatively resistant to adenoviral-mediated gene transfer. The receptor for adenoviruses has been identified as the Coxsackie-and-adenovirus receptor protein and has been useful in determining the mechanism of adenoviral binding and cell entry. Adenoviral vectors do not integrate into the genome and are lost from cycling target cells within weeks of transduction. First-generation adenoviral vectors may incite potent immune responses due to viral genes encoded by the vectors resulting in the elimination of transduced cells and preventing successful re-administration. Adenoviruses are therefore ideal vectors if an immune response and/or short-term high-level transgene expression is desired; however, they are not suitable for therapies that require long-term expression or integration of the transgene. In an attempt to overcome these limitations, newer generations of adenoviral vectors have either retained immunosuppressive sequences in E3, eliminated all viral sequences except the necessary inverted terminal repeats, or are based on a helper-dependent system with low levels of contaminating helper virus. Lytic adenovirus vectors, which replicate only in permissive cells, have been developed and hold promise for the treatment of solid tumors; for example, one such vector system only replicates in cells defective in the tumor suppressor gene *p53*.

The human parvovirus AAV is also being used as a vehicle for therapeutic gene transfer. AAVs are non-pathogenic to humans and establish a latent infection in the absence of adenovirus. Wild-type AAV integrates into human chromosome 19q13.3-qter, which appears to be a benign genomic integration site. Only the AAV inverted terminal repeats are required for recombinant AAV vectors, which allows approximately 4 kb for inserts. However, both the efficiency and the specificity of integration are lost without the added wild-type genes. A number of studies have focussed on determining the sequences responsible for the specific integration site, and results suggest that the rep protein is necessary for integration. AAVs have titers of approximately 10^6/ml and can be concentrated up to 10^9/ml. Recombinant AAV particles can integrate into both stationary and cycling cells, although they do not do so very efficiently. Integration of a recombinant AAV vector carrying the human β-globin cDNA into murine HSCs was demonstrated by the presence of the transgene and expression of human β globin in primary and secondary murine bone marrow transplant recipients for up to 9 months. Human $CD34^+$ cells have also been stably transduced with recombinant AAV vectors, with up to 80% of colony-forming units (CFU) carrying the transgene. Interestingly, the optimal use of recombinant AAV-based gene transfer vectors to affect the hematopoietic system may actually be in secondary manifestations. In a landmark study, Kay and colleagues (2000) injected recombinant AAV vectors that engineered expression of factor IX into skeletal muscles of severely afflicted hemophilia B patients. Long-term vector persistence was observed, as were slight increases in the circulating levels of the corrective pro-enzyme.

Herpes simplex-based viral vectors or amplicons are also now being developed for human use. Like adenovectors, herpes virus-based vectors exhibit a high efficiency of gene transfer into a wide variety of cell types but are not integrated. These vectors are distinguished by their wide tissue tropism and large cloning capacity. Such vectors have been demonstrated to be highly efficient in transducing HSCs, with gene transfer efficiency of 80–100%.

Stem cell gene delivery in hematology

Pluripotent HSCs are attractive targets for gene therapy in humans because of their capacity for self-renewal and the systemic multilineage distribution of their progeny (Figure 23.2). Sustained expression of transgenes at clinically relevant levels in the progeny of HSCs would result in novel and potentially curative treatments for a wide range of blood diseases, including hemophilia A and B, hemoglobinopathies, hereditary immune deficiencies and lysosomal storage diseases. Even the partial correction of such blood disorders would have a substantial impact on the transfusion needs of the affected populations. Nevertheless, despite the inherent promise of HSC gene transfer, successful long-term engraftment of genetically modified HSCs in humans has proved to be a difficult and often elusive goal.

Gene marking of hematopoietic stem cells

Genetic marking of HSCs facilitates study of the long-term distribution and survival of transplanted cells *in vivo*, although in the absence of clinical trials the field has suffered somewhat from a paucity of models with which to test long-term experimental outcomes. In most applications, the cells to be tagged are incubated *ex vivo* with a replication-defective retrovirus bearing a reporter gene. The reporter gene used most often in clinical studies is the bacterial neomycin phosphotransferase (neo^R) gene, which, when expressed, confers resistance to the neomycin analog G418. Other marker genes are now available and include the murine heat-stable antigen, the human CD24 and CD25 antigens, the truncated nerve growth factor receptor, modified CD4 and CD34 antigens, and the green fluorescent protein (Plate 23.1). The stable and unique integration pattern of proviral DNA in the genome of marked cells can provide a permanent marker for individual hemat-

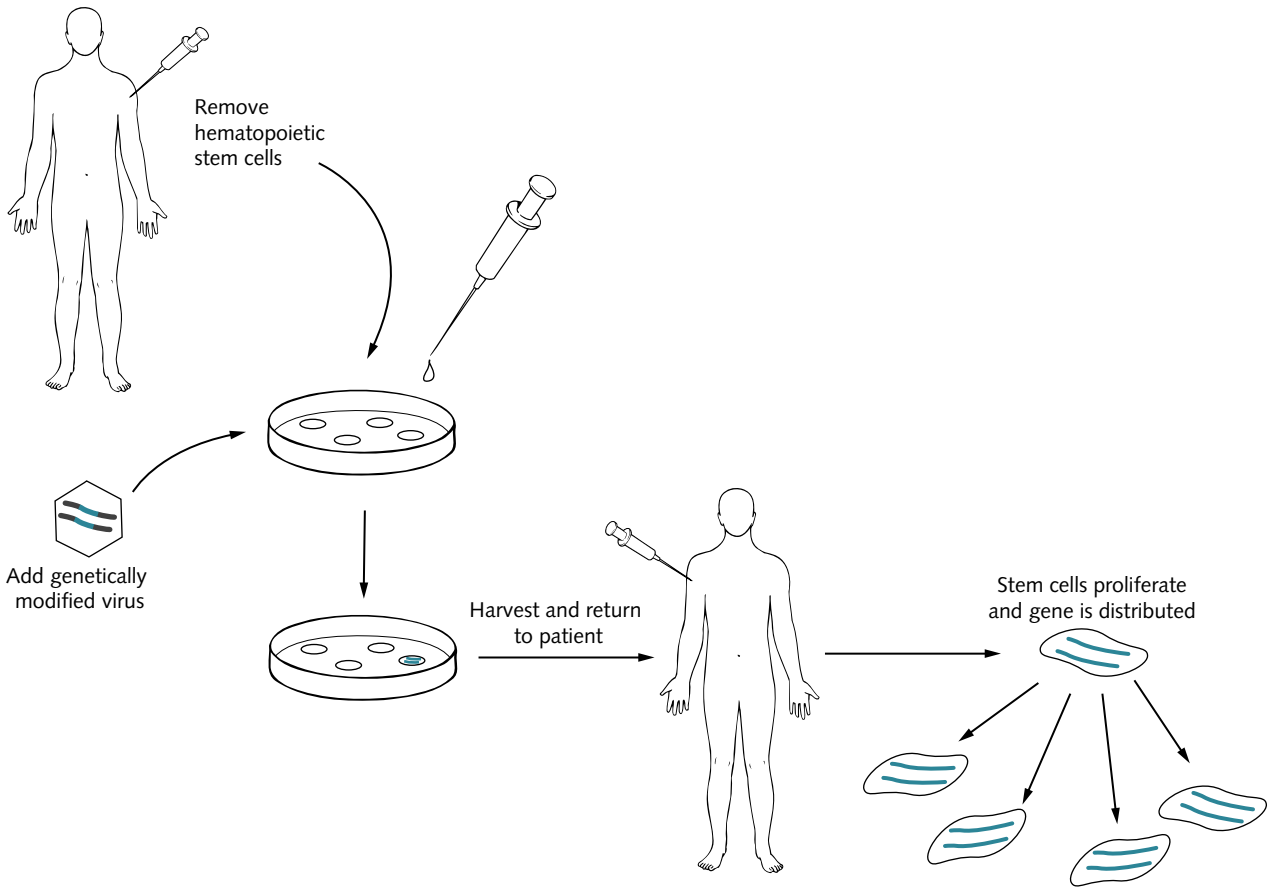


Fig. 23.2 *Ex vivo* transduction of hematopoietic stem cells
Hematopoietic stem cells can easily be harvested from patients and retrovirally transduced. The transduced cells are returned to the patient, where all blood cells maturing from the gene-modified stem cells retain and express a copy of the transgene. In theory, this gene transfer method could help distribute the gene product throughout the body at clinically relevant levels.

opoietic or malignant cells and their clonal descendants. This marking pattern is established using detailed PCR analyses that have recently been developed, which reflect the actual site of integration of the provirus. Clinical applications in which gene marking has provided new and important information include the infusion of onco-retrovirally marked, autologous, tumor-infiltrating lymphocytes (TIL) into patients with advanced melanoma and the infusion of onco-retrovirally marked bone marrow or peripheral blood into patients with myeloid leukemia, myeloma, and neuroblastoma. The study of such patients offers three important lines of investigation: (1) is retroviral-mediated gene transfer relatively safe? (2) do genetically altered bone marrow or blood stem cells contribute to long-term hematopoiesis? and (3) do malignant cells or their precursors contribute to the high relapse rates observed after myeloablative therapy and autologous HSC transplantation?

A number of groups have reported the consequences of infusing gene-marked bone marrow cells into humans and the contribution of contaminating tumor cells in the graft to disease recurrence. The first important observation from these studies is that retroviral-mediated gene transfer as currently practiced appears relatively safe. No detrimental effects, either on the autograft or in patients, have been reported in these studies. That said, in a very recent clinical gene therapy protocol for an inherited disorder, the possibility exists that T-cell leukemias in two patients resulted from the random integration of a therapeutic provirus into an oncogene (see below). Replication-competent virus has also not been detected at appreciable levels in patients participating in clinical trials. The second important observation from earlier gene marking studies is that HSCs contribute to long-term hematopoiesis, albeit at relatively low levels (Table 23.2).

In the first reported gene marking studies in children, 2–15% of clonogenic hematopoietic progenitor cells were marked after ABMT (*see papers by* Brenner *et al.*, 1993). The marker gene was detectable for up to 4 years after transplant and was found in granulocytes, B cells and T cells, at least by

Table 23.2 Summary of some reported results from earlier gene marking trials.

Center	Cells	Protocol	*In vitro* CFU-GM gene transfer	*In vivo* results
St Jude Children's Hospital	BM	1 exposure 6 hours	5–20%	2–15% bone marrow CFU at 1 year
NIH	MPB $CD34^+$	3 exposures 3 days	21%	0.01–0.001% blood cells +ve at >18 months
	BM $CD34^+$	+ cytokines		
University of Toronto	BM	2 exposures 21 days stroma	37%	3% CFU +ve at 2 years 0.01% blood +ve at 2 years

MPB, mobilized peripheral blood; BM, bone marrow.

PCR. In the earlier adult gene marking studies reported, however, retroviral transduction of marrow or peripheral blood HSCs has resulted in the detection of integrated vector in only 0.01–0.1% of peripheral blood cells (Figure 23.3). Although the marker gene persisted for up to 2 years, neo^R-positive cells could be detected only intermittently with analyses employing a sensitive PCR reaction.

Work by a number of groups suggested that modification of the transduction protocols would result in a significant improvement in the engraftment of genetically modified HSCs. A second generation of gene-marking trials therefore attempted to increase levels of gene transfer in adult human HSCs. For example, since preclinical experience indicated that the use of bone marrow stroma enhances gene transfer into HSCs, this approach was evaluated in clinical trials. However, the results of this adaptation have proved to be no better than those discussed above and, in general, are very similar to those of the pediatric gene transfer studies. Although in one study there is a high level of gene-marked progenitors in the early post-transplant period, with a mean of 12% gene-marked bone marrow colonies (CFU-GM), the level drops to 3% by 2 years after transplantation (Figure 23.4), and only 0.01% of the blood cells in the periphery are marked. Recently, a third generation of primate gene-marking studies has been performed. Here, for example, novel cytokine combinations have been employed to induce HSC cycling and yet maintain some level of the primitiveness and engraftability of the transduced cell population. In addition, a fragment of the fibronectin protein has been used to co-localize virus and target cells and alternatively pseudotyped onco-retroviruses have shown promise in infecting more primitive hematopoietic cells. Indeed, levels of marking are now more encouraging and are reaching a point where some hematological disorders may actually be cured. For example, one recent study in baboons by Horn and colleagues (2002), using vectors that engineer the expression of fluorescence proteins, demonstrated stable, multilineage,

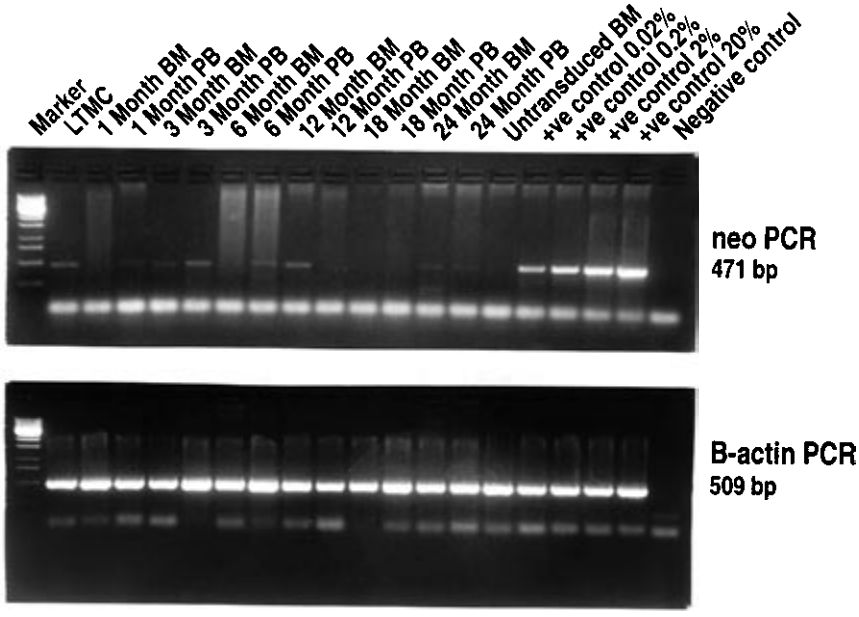


Fig. 23.3 PCR analyses of hematopoietic cells of a patient transplanted with marked bone marrow (BM) cells
Gene marking of peripheral blood (PB) BM is shown over a 2-year period for a patient who received gene-marked BM cells grown on stroma during retroviral infection. Although the *neo* resistance marker gene is readily detectable for 2 years, the levels of gene transfer are low.

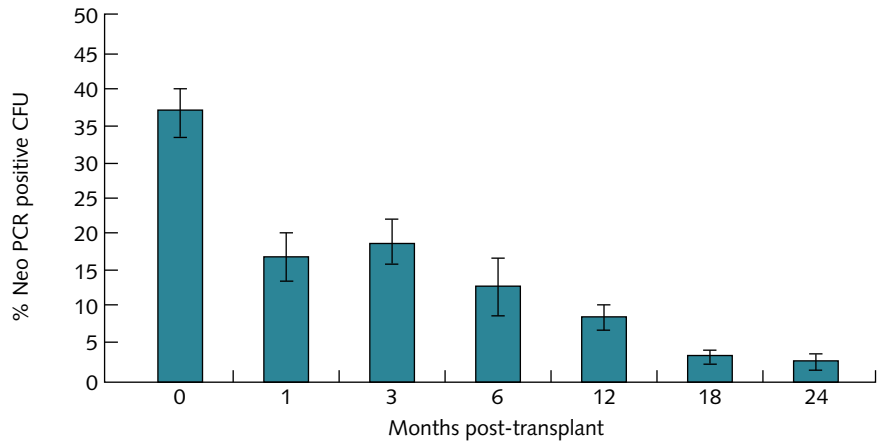


Fig. 23.4 Transgene-positive colonies over time in patients in a gene-marking trial
The number of colonies arising from bone marrow that contain the *neo*^{R} marker gene falls with time after gene transfer. Shown here are the average numbers of PCR-positive colonies from 15 patients in a gene-marking trial. By 2 years after transplantation, fewer than 5% of colonies are gene-marked. Note that the number of positive colony-forming units far exceeds the percentage of cells in the blood (Figure 23.3) that are gene-marked, raising the possibility that gene-marked bone marrow progenitors do not differentiate normally in the patient or are immunologically cleared.

functional marking of up to 25% of peripheral blood cells derived from transduced and transplanted CD34$^+$ hematopoietic cells. This, along with recent clinical results mentioned below, has led to a renewed sense of optimism in the field and suggests that the successful application of human therapeutic gene transfer will eventually be achieved.

Gene marking and relapse

In total, data on more than 40 patients enrolled in gene-marking studies during bone marrow transplant for acute myeloid leukemia, neuroblastoma, chronic myeloid leukemia (CML), breast cancer, myeloma, acute leukemia or non-Hodgkin's lymphoma have been presented (Table 23.3). Gene-marked tumor cells have been detected in a high percentage of relapsed patients. In one patient with acute myeloid leukemia, the simultaneous detection of a cytogenetic marker along with the *neo*^{R} gene confirmed that gene-marked cells contributed to relapse. A second critical observation was made when Rill and colleagues (1994) reported that a multiplicity of neuroblastoma cells in the graft contributed to relapse. The high frequency of gene-marked relapse, despite the very low frequency of transfused malignant cells, strongly suggests that a large percentage of tumor cells in the graft contribute to relapse, or that tumor cells susceptible to retroviral gene marking are uniquely capable of engraftment and clonal expansion.

Table 23.3 Gene marking studies.

	Patients	Relapse	Marked
AML	12	4	2
Neuroblastoma	9	5	4
CML	2	2	2
Myeloma	6	0	0
Breast cancer	5	3	0
Acute leukemia	5	2	2
NHL	3	1	3

AML, acute myeloid leukemia; CML, chronic myeloid leukemia; NHL, non-Hodgkin's lymphoma.

In adults, Deisseroth and colleagues (1994) at the University of Texas MD Anderson Cancer Center have examined tumor relapse in the HSC malignancy CML. Again, by demonstrating that clonogenic cells at relapse contain both the *neo*^{R} marker gene and the tumor-specific *BCR-ABL* oncogene rearrangement, this group demonstrated that infused gene-marked tumor cells can contribute to relapse. Another important, if somewhat contrasting, published report of gene marking hematopoietic cells in cancer patients is from Dunbar and colleagues (1995) at the National Institutes of Health. In 11 patients with breast cancer or myeloma, CD34$^+$ bone marrow or peripheral blood hematopoietic progenitor cells were isolated and marked. None of the six myeloma patients have demonstrated gene-marked relapse, nor have CD38 bright (plasma cells) in the marrow of two patients contained the marker gene, as determined by PCR 1 year after transplantation. Similarly, in three breast cancer patients who have relapsed no evidence for the *neo*^{R} gene has been found in tumor biopsies.

Future directions

The gene-marking studies mentioned above set the stage for the investigation of a multiplicity of manoeuvers designed to increase the efficiency of gene transfer and the duration of its effect using HSCs. Some strategies being pursued are described in Table 23.4 and include increasing the cell-to-virus contact by incorporating HSC selection, altering retroviral envelope targeting or co-localization of vector and cell

Table 23.4 Strategies for optimization of gene transfer into HSCs.

Strategy	Method
Inducing recipient cells to cycle	Optimization of ***ex vivo*** cytokine stimulation Collection of cells during recovery phase after myeloablation or mobilization Culture on stromal layers
Increased cell–virus contact	Centrifugation of cells and virus during transduction (spinoculation) Viral supernatant flow-through systems Coat dishes with fibronectin fragment Higher viral titers and multiple exposures
Increase viral receptor levels on target cells	Increase levels of amphotropic receptor by phosphate depletion Transfer viral receptor into cell by adenovirus or adeno-associated virus Target subpopulations of cells that have high levels of receptors
Alternatively pseudotyped recombinant retroviruses	Exploiting GALV, RD114, 10A1 receptors for entry VSV-G envelope to expand tropism and allow virion concentration
Modified retroviral vectors to target non-cycling cells	Lentivirus and foamy virus vectors
Positive selection of transduced cells	Add positive selectable marker to vector: metabolic, fluorescent, cell-surface

in gene transfer protocols. New vector systems improving on both retroviral and lentiviral backbones and incorporating alternative envelope pseudotyping may increase gene transfer efficiency, while refinement of growth factor combinations may induce HSC cycling more efficiently, thus improving retroviral integration. Many of these approaches have now been incorporated into clinical protocols and have likely contributed to the recent successes in this field, as described below.

Chemoprotection

The above results suggest that normal bone marrow cells could be removed, transduced with a recombinant retroviral vector containing a drug-resistance gene, and returned to untreated patients with no toxicity. Theoretically, such protected cells would expand clonally after chemotherapy treatment and confer relative resistance on the cytotoxic or cytostatic agent, allowing further dose escalation and a potential cure of some patients. Furthermore, such drug-resistance genes, if expressed in a multicistronic format with a second therapeutic gene product, may even serve as selectable markers that allow *ex vivo* or *in vivo* enrichment or preselection of functionally gene-marked cells, which may improve clinical outcomes.

Gene products may confer chemoprotection

Numerous potential mechanisms of cellular resistance to chemotherapy agents have been described. Candidate gene products for chemoprotection (and their substrates) include P-glycoprotein (anthracyclines, taxol, etoposide, vinca alkaloids), multidrug resistance-associated protein (MRP; anthracyclines, etoposide, vinca alkaloids), cytidine deaminase (cytosine arabinoside), ribonucleotide reductase (hydroxyurea), topoisomerase II (etoposide), aldehyde dehydrogenase (cyclophosphamide), O^6-alkylguanine-DNA-alkyltransferase (nitrosoureas), glutathione S-transferase (melphalan), dihydrofolate reductase (methotrexate), thymidylate synthase (floxuridine) and tubulin (vinca alkaloids).

The cDNA for the human *MDR-1* gene (P-glycoprotein) has been subcloned and expressed in both mouse and human cells in culture. Transfer of the *MDR-1* gene into mouse and human bone marrow cells using onco-retroviral constructs leads to drug resistance *in vitro*. Transduced and transplanted cells engraft and confer drug resistance to bone marrow *in vivo* and also allow positive selection for *MDR-1*-transduced cells, for example, by chemotherapy (Figure 23.5). For human bone marrow cells, using a short-term culture with supernatants obtained from a high-titer producer line, 10–50% of the bone marrow-derived progenitor cells express the transferred *MDR-1* gene. With the best *MDR-1* retroviruses developed so far, drug concentrations up to ten times the 50% lethal dose can be used. It should be noted, however, that a caveat to these *MDR-1* studies has been observed *in vivo* when strong metabolic selective pressure was used to select for transgene positive cells. Bunting and colleagues (1998) reported the adverse development of a myeloproliferative syndrome in mice transplanted with *MDR-1*-transduced grafts.

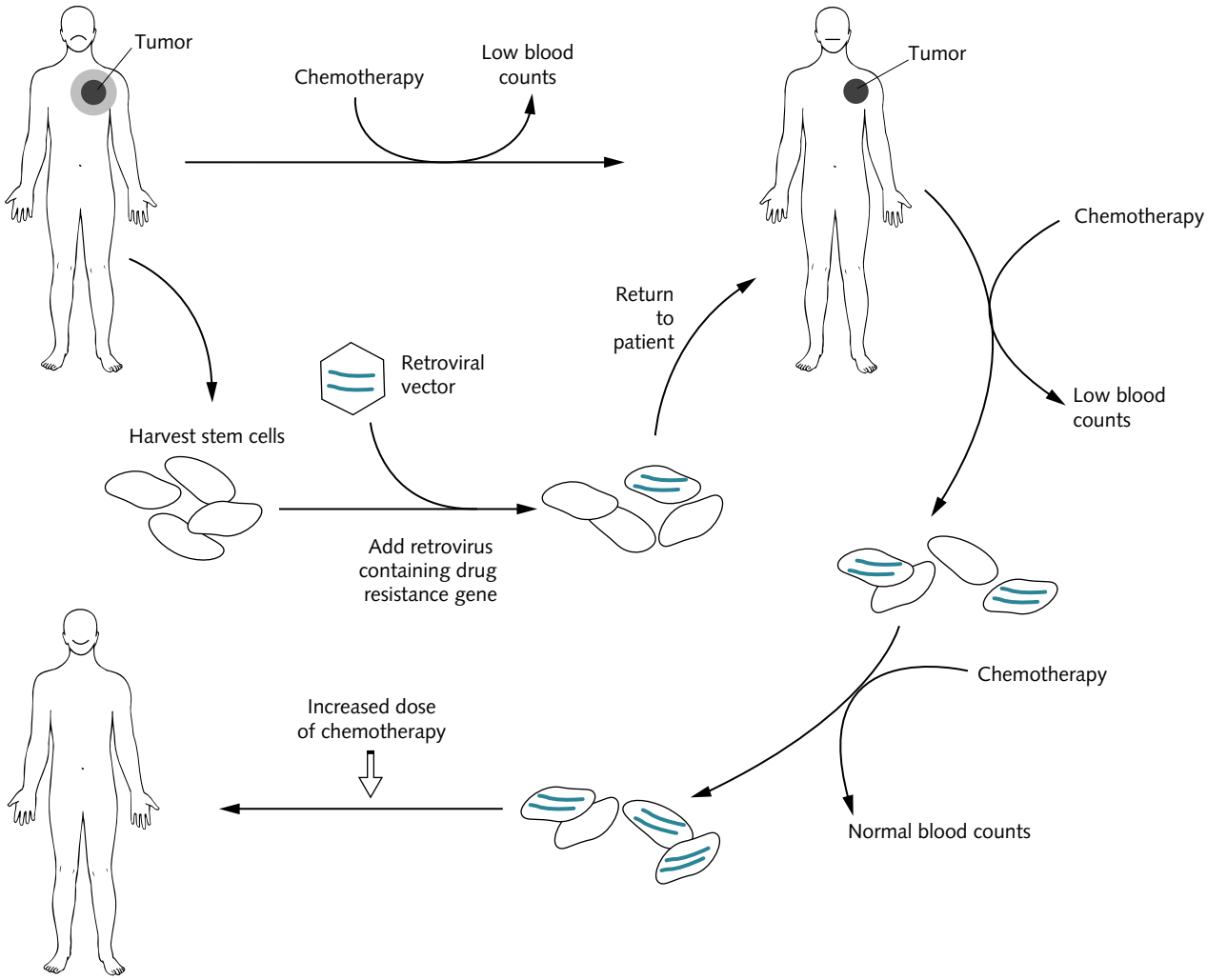


Fig. 23.5 Hematopoietic stem cells removed from the patient can be genetically engineered to express a drug resistance gene
Theoretically, after return to the patient, the cells are selected by chemotherapy (normal stem cells survive chemotherapy less well than the cells expressing the drug resistance gene). At a certain point the number of drug-resistant bone marrow cells rises to a degree that protects bone marrow from the effects of chemotherapy, and thus higher doses of chemotherapy can be given without all the usual side effects. Note, however, that systemic side effects not related to bone marrow will persist. This method could also be used to select for a second gene product expressed in concert with the *MDR-1* gene.

Clinical trials of chemoprotection

Clinical trials of *MDR-1* gene transfer into hematopoietic cells for chemoprotection in cancer therapy have been performed. The approach used for gene transfer is generally similar to that used for human gene-marking trials. Preliminary results indicated that infusion of *MDR-1* gene-modified cells is safe and no deleterious effects of infusing gene-modified cells have been observed. Neutrophil and platelet recovery times were normal following infusion of gene-modified cells. In a recent landmark study by Abonour and colleagues (2000) concerning this approach, a combination of some of the incremental improvements in gene transfer protocols (mentioned above) yielded fairly efficient transfer of the human *MDR-1* gene into long-term repopulating hematopoietic cells. Indeed, appreciable levels of gene marking were observed 1 year after transplantation in PCR analyses of colonies from six recipients out of the 12 subjects enrolled and transplanted in this protocol, at levels up to 15% positive. Furthermore, in three of these patients functional expression of the *MDR-1* gene was demonstrated by growth of colonies in the presence of paclitaxel.

Gene augmentation therapy for immunodeficiencies

The first human clinical gene-transfer trials for inherited single-gene disorders focussed on adenosine deaminase (ADA) deficiency and are summarized in Table 23.5. Recombinant onco-retroviruses containing the normal human *ADA* cDNA were transferred into peripheral blood T lymphocytes, bone marrow cells or cord blood cells from ADA-deficient patients. T-lymphoid cells expressing the normal *ADA* gene have a selective growth and survival advantage over ADA-deficient cells, even though patients are maintained on polyethylene glycol-modified bovine ADA (PEG-ADA) enzyme replacement therapy (ERT) for ethical reasons.

In the early studies, patients who had received multiple infusions of autologous *ADA*-transduced blood cells had increased levels of enzyme in their serum, and up to 20% of their peripheral blood T cells were found to carry provirus. In two of three trials in which patients received autologous marrow or cord blood cells transduced with *ADA*-containing onco-retroviruses, between 12 and 40% of CFU were transduced, and genetically marked cells were found for more than 1 year after infusion. In the one study in which provirally marked cells were not maintained for longer than 6 months, there was lower *in vitro* gene transfer efficiency of 5–12% CFU prior to transplantation (*see* Hoogerbrugge *et al.*, 1996). In the cord blood study there was evidence for a selective growth advantage of T cells, as there were higher levels of marked T cells than myeloid cells even while the patients were maintained on progressively lower doses of PEG-ADA enzyme therapy. ERT was withdrawn from one patient and the number of T cells carrying the provirus increased to 30%; however, the total number of B lymphocytes and natural killer cells dropped and the patient had reduced immune function. The patient subsequently resumed PEG-ADA treatment.

Inherited immunodeficiencies (both ADA-SCID, as mentioned above, and X-linked SCID, below) have become focal points for the potential benefits and some of the potential hazards of gene therapy. The above studies in ADA-SCID likely represent the first tangible correction of an inherited disorder by stable transfer of a therapeutic gene into primitive hematopoietic cells and their progeny. Note that these initial results have been surpassed recently in a study published by Aiuti and colleagues (2002), in which the *ADA* gene was transferred into $CD34^+$ cells of non-myeloablated recipients for whom ERT was not available. Many of the incremental improvements in the gene transfer protocols mentioned above were adapted in this protocol, and high levels of gene marking (up to 25% of CFU-C initially) and functional correction were demonstrated. Long-term, multilineage, hematopoietic cell marking was also found in both patients. This study demonstrated, without the confounding implications of ERT, that stable long-term correction of this immune deficiency could be accomplished using this therapeutic approach. Indeed, these patients are presently at home, have normal growth and development parameters, and respond effectively to immune challenge from a variety of agents.

Impressive results have also been demonstrated in efforts to correct another inherited immunodeficiency, X-linked SCID (Cavazzana-Calvo *et al.*, 2000). X-linked SCID (SCID-X1) is caused by a deficiency of the common γ chain subunit (γc) of cytokine receptors for IL-2, IL-4, IL-7, IL-9 and IL-15. Expression of γc was also expected to offer a growth advantage to productively transduced cells. This was indeed the case, and in this first report in 2000 two patients were shown to have fully corrected immune function as a result of the gene therapy. Since this first description of this benefit, other X-linked SCID patients have also been treated by this method. In fact, to date ten patients have received this gene therapy and seven patients remain with improved immune function. Yet even with these

Table 23.5 Early gene therapy trials for ADA deficiency.

Center	Cells	Protocol	*In vitro* gene transfer	*In vivo* results
NIH	PB supernatant	9–12 h		Increased immune repertoire Number of T cells normalized after 2 years
Italy	PB BM	Co-culture 72 h or supernatant	2.5–50% cells 30–40% CFU	Multilineage repopulation with marked cells Increased immune repertoire for 2 years
The Netherlands	$CD34^+$ BM	Co-culture + IL-3	5–12% CFU	Transduced cells detected 3–6 months after transplant
St Jude Children's Hospital	CB $CD34^+$	Supernatant + IL-3 + IL-6 + SCF	12.5–21.5% CFU	Multilineage repopulation with transduced cells
Japan	PB	Supernatant + IL-2	3–7% cells	Improvement in immune function; 10–20% of blood cells carrying provirus >1 yr

BM, bone marrow; CB, cord blood; NIH, National Institutes of Health; PB, peripheral blood.

impressive results this study highlights how much further the field must progress in its understanding before this type of therapy can be the broad therapeutic panacea it was hoped. This is because of recent findings that two of the patients receiving the corrective γc gene have gone on to develop leukemias approximately three years after the transplantation. In one of these patients the development of this proliferative disorder has been ascribed possibly to a deleterious integration event that caused higher levels of expression of a proto-oncogene. Yet while these findings are worrisome and are causing regulatory agencies and others to re-examine protocols, the position of most seems to be that more study is definitely required and that risks and benefits for each protocol must be evaluated in detail. For example, because these vectors and transduction conditions have been used for other studies and no adverse events have been reported, the possibility exists that this leukemic effect is a specific consequence of overexpression of the γc gene itself, since it affects many diverse signaling pathways that influence a number of cellular functions *in vivo*.

Other inherited diseases of the hematopoietic system

Gaucher disease, which results from a deficiency in the enzyme glucocerebrosidase and is manifested mainly in macrophages, has also been proposed as being especially amenable to treatment by gene transfer into HSCs. Results from two earlier clinical gene transfer studies targeting this disorder have been reported. Mobilized blood or marrow $CD34^+$ cells from Gaucher patients were transduced with a glucocerebrosidase-expressing onco-retroviral vector and infused into non-myeloablated autologous recipients. In both studies, transduced cells were detected at low levels in blood and/or marrow leukocytes. One patient who received cells transduced with an MFG-based onco-retroviral vector had an increase in levels of enzyme corresponding to 50% of normal, which was maintained for 12 months after infusion. No therapeutic benefit or increased enzyme was detected in other patients or in the other study.

Collectively, the results from the gene transfer studies for genetic disease described above (and others not directly mentioned) illustrate several points which will likely affect the clinical success of gene transfer protocols for other single inherited gene disorders. (1) The presence in patients of cells carrying the provirus for longer than 1 year has demonstrated the feasibility of gene therapy. (2) The transfer of genes that provide a growth or survival advantage can provide long-term expression and maintenance of transduced cells. (3) However, a selective advantage of transduced cells cannot compensate for poor *ex vivo* gene transfer efficiency and, if too strong, may lead to amplification of deleterious transformation events. (4) Successful gene therapy of deficiency disorders could be severely limited by immune responses to the transgene product, which may be considered foreign, in immunocompetent patients. (5) Using the incrementally optimized protocols, gene delivery and expression levels may at present be sufficient to correct a number of disorders that are directly or indirectly manifested in the hematopoietic system. Clinical gene therapy trials for other inherited single-gene disorders, such as Fanconi's anemia type C, mucopolysaccharidosis type I, Fabry disease and chronic granulomatous disease, are completed, ongoing or proposed.

Gene therapy of hematological malignancy

A number of potential approaches to the gene therapy of hematological malignancies can be considered (Table 23.6) and are discussed here.

Genetic immunotherapy

Many tumor cells express unique peptide antigens on their cell

Table 23.6 Approaches to gene therapy targeting hematological malignancies.

Strategy	Genes employed	Problems and merits
Gene replacement	*p53, p16, Rb*	Gene delivery to every cell required Specific to defective cancer cell
Gene inhibition	*BCR-ABL, c-MYC, cyclin D1, BCL-2*	Gene delivery to every cell required
Suicide genes	Thymidine kinase, cytosine deaminase	Immunogenicity and bystander effect contribute
Drug-resistance genes	*MDR-1, DHFR*	Stem cell gene delivery required
Immunotherapy	*IL-2, IL-12, B7-1, CD40L, GM-CSF,* tumor-associated antigens	Systemic response, autoimmunity a theoretical problem Requires that tumour express foreign antigen

surface in association with major histocompatibility (MHC) class I molecules, and the specificity of these antigens may allow immune effector cells to distinguish tumor from normal tissue (Figure 23.6). Some of these tumor-associated antigens have been isolated and shown to be recognized by human cytotoxic T-lymphocyte (CTL) cell lines. Nevertheless, tumor cells known to express potentially antigenic peptides manage to evade host immunosurveillance and proliferate *in vivo*. Thus, tumor cells may lack the necessary accessory signals required to induce the expression, by immune effector cells, of the cytokines necessary for the activation and directed *in vivo* expansion of CTLs. The end result is anergy, a failure of the T cells to respond to the tumor antigen. In addition to optimal presentation of antigen to the T-cell receptor, efficient activation of naive T cells requires a second co-stimulatory signal. It is now appreciated that molecules of the B7 family (B7-1/CD80, B7-2/CD86) on antigen-presenting cells engaging CD28/CTLA-4 receptors on T cells play a key role in this process, inducing autocrine IL-2 production and T-cell proliferation. Murine models have demonstrated that T-cell-mediated rejection of tumors can be induced by transduction of tumor cells with such co-stimulatory molecules. In the absence of co-stimulatory signals, it is possible to bypass this requirement by ectopic expression of cytokines, the down-

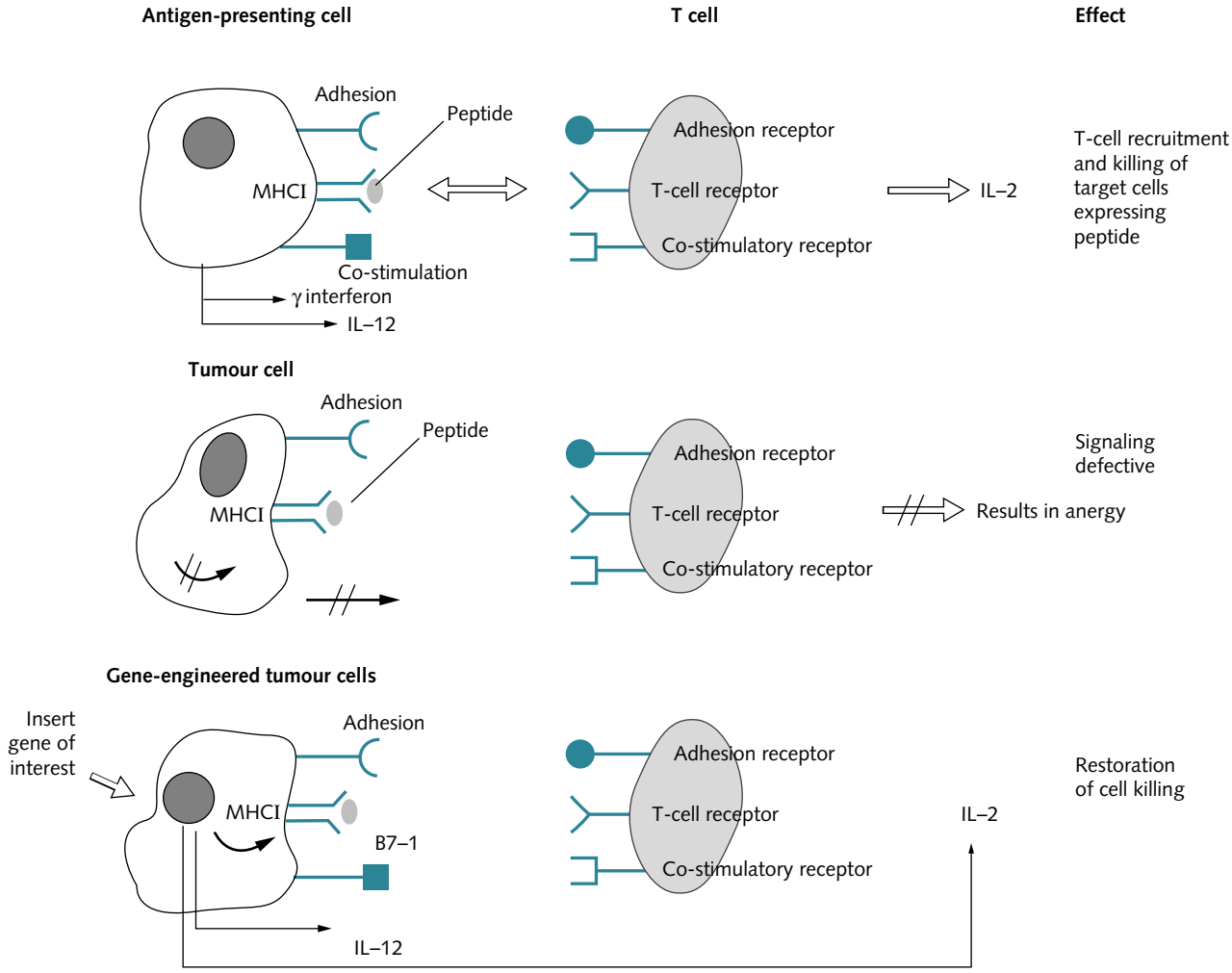


Fig. 23.6
Efficient signaling of T cells occurs after adhesion, after provision of the major histocompatibility complex (MHC class I)/T-cell receptor signal, and after a costimulatory signal is received. Under normal circumstances professional antigen-presenting cells present processed peptides in the groove of the MHC I complex to T cells. Tumor cells are often deficient in one or all of the components required to function as antigen-presenting cells because they lack an appropriate tumor antigen, cannot process the antigen or are deficient in the adhesion molecules, MHC class I or costimulatory molecules required to generate a T-cell response. These missing components can be provided or enhanced by the use of gene transfer techniques. By overcoming the deficiencies of the tumor cell, the gene-engineered cells can serve as autologous cancer vaccines presenting foreign antigen to the host T cells.

stream products of co-stimulation, and thus overcome or prevent anergy of the immune effector cells. Thus, proof of the principle that cytokine gene-transduced tumor cells can prevent tumor engraftment has been obtained. Such models have also shown that transduction of the genes for various cytokines, such as IL-2, IL-4, IL-6, IL-7, IL-12, IFN-γ, GM-CSF and TNF-α, into murine tumors not only leads to primary rejection of the modified cells but often elicits protective immunity against subsequent tumor challenge with unmodified tumor cells. Furthermore, in such models, synergy has been demonstrated between molecules with varying mechanisms of action, for example IL-2, IL-12 and B7-1 (Figure 23.7).

Proof of the concept of immunotherapy as a valid approach to the treatment of hematological malignancy has now been provided by clinical studies that have demonstrated that infusion of allogeneic T cells can eradicate minimal disease in patients relapsing after allogeneic transplantation. Unfortunately, these encouraging allogeneic responses require a haploidentical T-cell donor, which is not available to the vast majority of patients. Another, more widely applicable, approach is the use of autologous tumor cells which have been genetically engineered to express one or more of the immune-stimulatory cytokines mentioned above. Trials in myeloma, low-grade lymphoma, leukemia and chronic lymphocytic leukemia which use this strategy are being pursued.

Cellular immunotherapy

As outlined above, the object of some cancer immunotherapies is to efficiently present tumor-associated antigens to the immune system. The most immunologically powerful (so-called professional) antigen-presenting cells are bone marrow-derived dendritic cells (DCs). DCs express MHC class I and II, B7-1, B7-2, CD40, ICAM-1 and LFA-3, are capable of presenting processed antigen for days, and are potent stimulators of immunity when administered as a vaccine to animals. DCs modulate immune responses in part by the secretion of IL-12 (hence the notion of converting tumor cells into antigen-presenting cells by the introduction of IL-12 and B7-1 genes). DCs can be readily expanded from bone marrow progenitors *in vitro* using cytokine-supplemented medium (useful cytokines include Flt3L, TNF-α, GM-CSF and IL-4) and can also be propagated using GM-CSF-expressing tumor cells as a continuous source of cytokine (presumably explaining

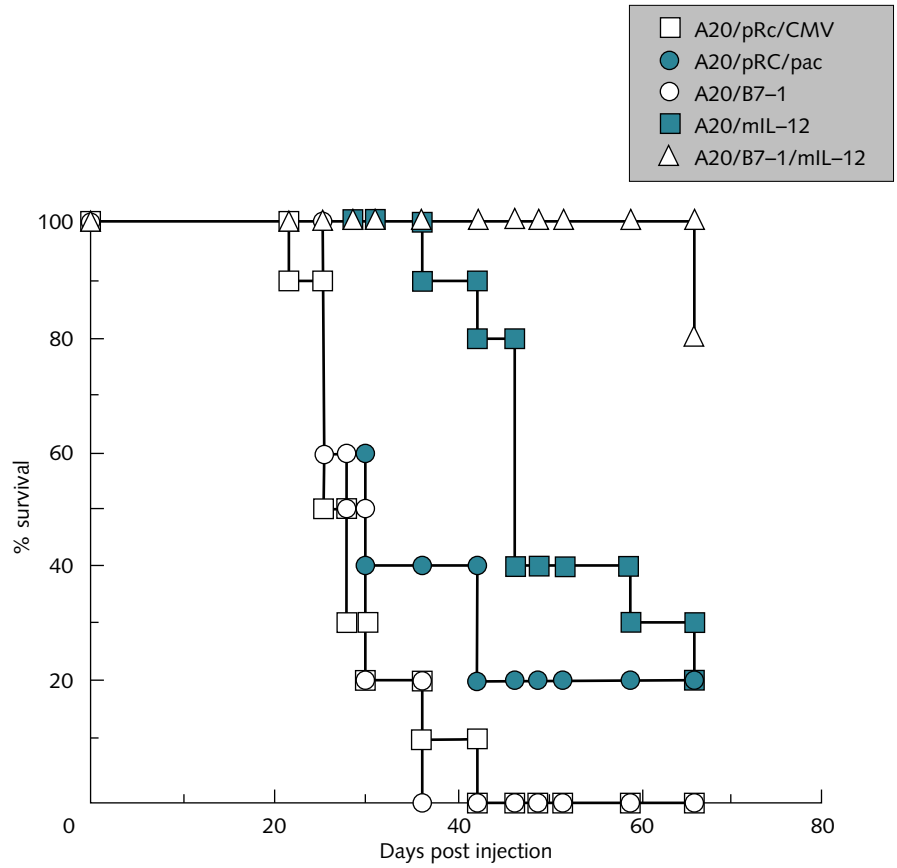


Fig. 23.7
The A20 lymphoma cell line was genetically engineered to express B7-1, IL-12, or a combination of both. The survival of mice inoculated with a mixture of 10^5 parental and 10^5 variant cells at the same site on day 0 is shown. Although there is a marginal increase in survival for the IL-12 cell line, clearly most of the protected mice were vaccinated with the A20 cell line co-expressing B7-1 and IL-12. Reproduced from Pizzoferrato *et al*. (1997) with permission.

the *in vivo* anti-tumor properties of GM-CSF). Such cells may also be genetically engineered to express tumor-associated antigens and thus hold great promise for the immunotherapy of hematological malignancy.

Suicide gene therapy

The so-called suicide genes are enzymes which, when expressed by transduced target cells, confer susceptibility to drug-induced cell death by specifically converting normally non-toxic prodrugs into potent cytolytic or cytostatic molecules. The gene most commonly employed to date in clinical trials in this context is the human herpes virus type 1 thymidine kinase gene (*Tk*), which, when expressed, confers sensitivity to the drug ganciclovir. Use of *Tk* is further enhanced for some applications by diffusion of the converted prodrug into neighboring cells, and thus a *bystander effect* occurs, which is mediated largely via the gap junctions connecting cells (Figure 23.8). The obvious limitations of this treatment approach are that not all cells targeted will be successfully gene-modified and thus, even with the bystander effect, only a fraction of malignant cells will be destroyed. Nevertheless, applications of this type are in clinical trials for solid tumors. Another proven clinical application of this approach, which may have a higher probability of success, is in the prevention of graft-versus-host disease (GVHD) following transplantation. In this approach, gene-modified lymphocytes are infused into patients relapsing after an allogeneic bone marrow transplant, for example. If signs of GVHD are noted, ganciclovir treatment is started and the specifically transduced and transplanted lymphocytes are aborted. Other approaches being explored as part of the therapeutic arsenal of cancer gene therapy include combining thymidine kinase with potent effector cytokines, the use of novel cell suicide genes, and the use of cell death genes, which are host-derived and thereby non-immunogenic.

Genetic inhibition: antisense and ribozymes

A number of oncogenes, such as *RAS*, c-*MYC*, *cyclin D1* and

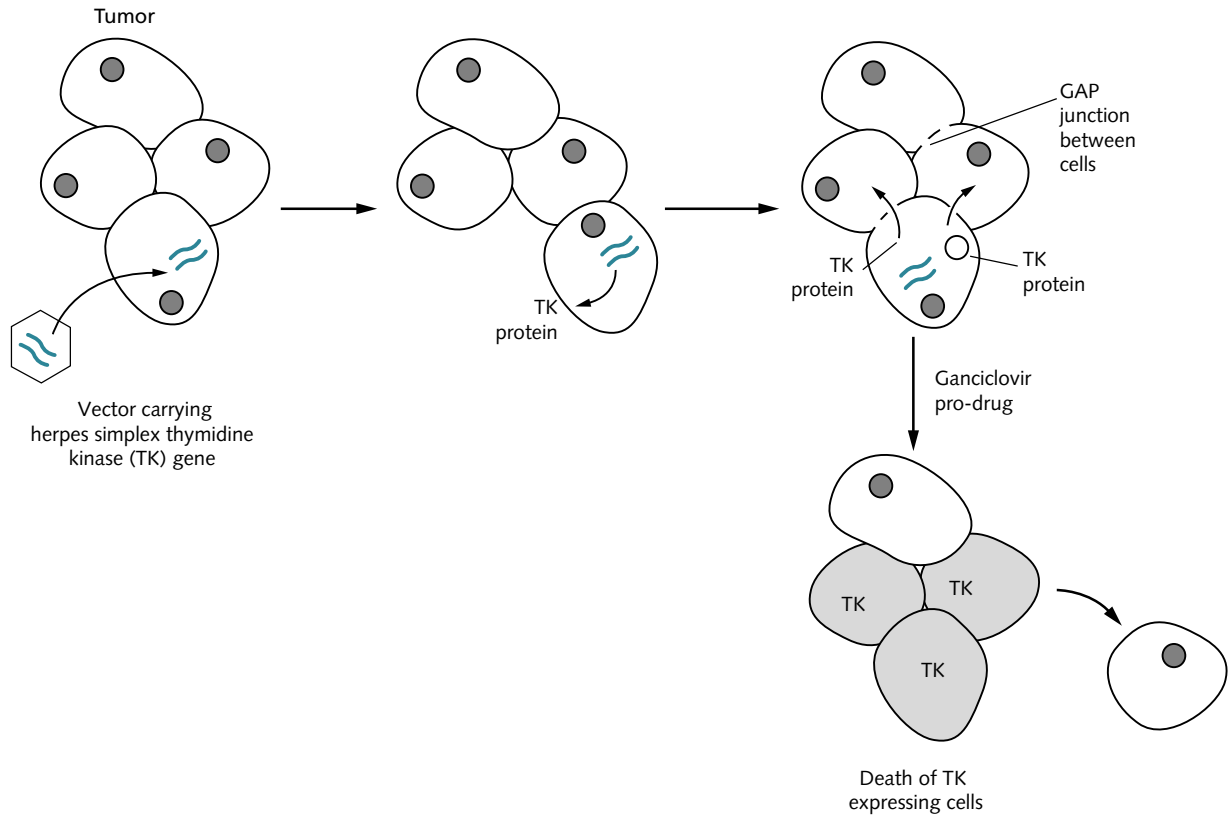


Fig. 23.8 Basis of suicide gene therapy

The metabolic product of the activity of the thymidine kinase suicide gene appears capable of diffusing into neighboring cells via the gap junctions that join the cells together. This non-specific diffusion allows a greater effective cell-killing percentage than might be predicted using direct gene transfer efficiency alone. Because cells are dying and releasing tumor antigen into the local milieu of the tumor, it is possible that such suicide gene therapies will synergize with immune-based treatment strategies.

BCR-ABL, are overexpressed in hematological malignancy and in many cases are pivotal in the progression of malignant disease. One gene therapy approach to such malignancies is therefore the use of antisense cDNAs to bind to and prevent the transcription of the sense target cDNA. Alternatively, RNAs with enzymatic activity (*ribozymes*) which possess highly sequence-specific RNA cleaving potential have been used. Lastly, RNA interference (RNAi), which entails the use of small double-stranded RNAs to elicit the downmodulation of target gene expression, could also be used in this framework. The failure of current gene transfer technology to introduce genetic material into all target cells will, however, likely limit such therapeutic strategies *in vivo* for the time being. *In vitro* strategies are now being pursued using such molecules for the purging of blood or bone marrow HSC transplant products for return to patients after myeloablation.

Tumor suppressor gene replacement therapy

A number of tumor suppressor genes (for example, *p53*, retinoblastoma and *p16*) are commonly deleted in hematological malignancies. Many of these tumor suppressor genes have been inserted into viral vectors for the purpose of gene replacement therapy. However, most hematological malignancies are disseminated at presentation and current gene transfer technology is inadequate for widespread gene replacement *in vivo* in most hematology patients. Use of gene replacement vectors *in vitro* is, however, being explored for the purpose of purging bone marrow or stem cell autografts free of tumor prior to autologous transplantation. In the future, gene transfer technology may allow the use of such genes in a more therapeutically relevant manner.

Vascular-mediated anti-tumor therapy

Since the lack of gene delivery to every cancer cell is currently limiting the application of successful gene therapy in some applications, attention has recently turned to the use of the supporting vascular epithelium as a target to inhibit tumor proliferation and dissemination. One very promising strategy is the sustained delivery of angiostatic genes with viral vectors. In this strategy, inhibitors of blood vessel growth are expressed within the tumor for sustained periods, leading to tumor regression and the inhibition of metastasis. Preclinical evaluation of such strategies is under way, targeting a variety of factors and using many model systems.

Conclusions

Concepts and approaches to the molecular therapy of diseases of the blood system which are currently popular have been described in this chapter but many other therapeutic strategies are also being developed. Indeed, the advent and recent success of human gene therapy trials combined with rapid advances in gene transfer technology and the full sequencing of the human genome are setting the stage for a new era in human medicine. The field of human molecular medicine will surely continue to expand and make inroads in a therapeutically relevant way. It seems likely that study of the blood system, both in the laboratory and in patients, will continue to play a pivotal role in the further development of this field.

Further reading

Introduction

Anderson WF. (1992) Human gene therapy. *Science*, **256**, 808–813.

Brenner MK. (2001) Gene transfer and the treatment of hematological malignancy. *Journal of Internal Medicine*, **249**, 345–358.

Brenner MK, Rill DR, Holladay MS *et al.* (1993) Gene marking to determine whether autologous marrow infusion restores long-term haemopoiesis in cancer patients. *Lancet*, **342**, 1134–1137.

Brenner MK, Rill DR, Moen RC *et al.* (1993) Gene-marking to trace origin of relapse after autologous bone-marrow transplantation. *Lancet*, **341**, 85–86.

Miller AD. (1992) Human gene therapy comes of age. *Nature*, **357**, 455–460.

Rosenberg SA, Aebersold P, Cornetta K *et al.* (1990) Gene transfer into humans — immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *New England Journal of Medicine*, **323**, 570–578.

Williams DA, Smith FO. (2000) Progress in the use of gene transfer methods to treat genetic blood diseases. *Human Gene Therapy*, **11**, 2059–2066.

Zanjani ED, Anderson WF. (1999) Prospects for in utero human gene therapy. *Science*, **285**, 2084–2088.

Gene transfer strategies

Bergelson JM, Cunningham JA, Droguett G *et al.* (1997) Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science*, **275**, 1320–1323.

Bischoff JR, Kirn DH, Williams A *et al.* (1996) An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. *Science*, **274**, 373–376.

Cantwell MJ, Sharma S, Friedmann T *et al.* (1996) Adenovirus vector infection of chronic lymphocytic leukemia B cells. *Blood*, **88**, 4676–4683.

Chen HH, Mack LM, Kelly R *et al.* (1997) Persistence in muscle of an adenoviral vector that lacks all viral genes. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 1645–1650.

Crystal RG. (1995) Transfer of genes to humans: early lessons and obstacles to success. *Science*, **270**, 404–410.

Danos O, Mulligan RC. (1988) Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 6460–6464.

Dilloo D, Rill D, Entwistle C *et al.* (1997) A novel herpes vector for the high-efficiency transduction of normal and malignant human hematopoietic cells. *Blood*, **89**, 119–127.

Gunzburg WH, Salmons B. (1996) Development of retroviral vectors as safe, targeted gene delivery systems. *Journal of Molecular Medicine*, **74**, 171–182.

Hallek M, Wendtner CM. (1996) Recombinant adeno-associated virus (rAAV) vectors for somatic gene therapy: recent advances and potential clinical applications. *Cytokines and Molecular Therapy*, **2**, 69–79.

Hawley RG, Lieu FHL, Fong AZC *et al.* (1994) Versatile retroviral vectors for potential use in gene therapy. *Gene Therapy*, **1**, 136–138.

Horn PA, Topp MS, Morris JC *et al.* (2002) Highly efficient gene transfer into baboon marrow repopulating cells using GALV-pseudotype oncoretroviral vectors produced by human packaging cells. *Blood*, **100**, 3960–3967.

Kotin RM, Siniscalco M, Samulski RJ *et al.* (1990) Site-specific integration by adeno-associated virus. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 2211–2215.

Lever AM. (1996) HIV and other lentivirus-based vectors. *Gene Therapy*, **3**, 470–471.

Marconi P, Krisky D, Oligino T *et al.* (1996) Replication-defective herpes simplex virus vectors for gene transfer *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 11319–11320.

Miller AD, Buttimore C. (1986) Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Molecular and Cellular Biology*, **6**, 2895–2902.

Miller DG, Adam MA, Miller AD. (1990) Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Molecular and Cellular Biology*, **10**, 4239–4242.

Mulligan RC. (1993) The basic science of gene therapy. *Science*, **260**, 926–932.

Naldini L, Blomer U, Gallay P *et al.* (1996) *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, **272**, 263–267.

Neering SJ, Hardy SF, Minamoto D *et al.* (1996) Transduction of primitive human hematopoietic cells with recombinant adenovirus vectors. *Blood*, **88**, 1147–1155.

Pawliuk R, Bachelot T, Wise RJ *et al.* (1999) Long-term cure of the photosensitivity of murine erythropoietic protoporphyria by preselective gene therapy. *Nature Medicine*, **5**, 768–773.

Qin G, Takenaka T, Telsch K *et al.* (2001) Preselective gene therapy for Fabry disease. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 3428–3433.

Rollins SA, Birks CW, Setter E *et al.* (1996) Retroviral vector producer cell killing in human serum is mediated by natural antibody and complement: strategies for evading the humoral immune response. *Human Gene Therapy*, **7**, 619–626.

Temin HM. (1993) Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 6900–6903.

Yang Y, Nunes FA, Berencsi K *et al.* (1994) Cellular immunity to viral antigens limits E1-deleted adenoviruses for gene therapy. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 4407–4411.

Stem cell gene delivery

Brenner MK. (1996) Gene transfer to hematopoietic cells. *New England Journal of Medicine*, **335**, 337–339.

Cheng L, Du C, Murray D *et al.* (1997) A GFP reporter system to assess gene transfer and expression in human hematopoietic progenitor cells. *Gene Therapy*, **4**, 1013–1022.

Deisseroth AB, Zu Z, Claxton D *et al.* (1994) Genetic marking shows that Ph+ cells present in autologous transplants of chronic myelogenous leukemia (CML) contribute to relapse after autologous bone marrow transplant in CML. *Blood*, **83**, 3068–3076.

Dunbar CE, Cotter-Fox M, O'Shaughnessy JA *et al.* (1995) Retrovirally marked CD34-enriched peripheral blood and bone marrow cells contribute to long term engraftment after autologous transplantation. *Blood*, **85**, 3048–3057.

Rill DR, Santana VM, Roberts WM *et al.* (1994) Direct demonstration that autologous bone marrow transplantation for solid tumors can return a multiplicity of tumorigenic cells. *Blood*, **84**, 380–383.

Stewart AK, Sutherland DR, Nanji S *et al.* (1999) Engraftment of gene-marked hematopoietic progenitors in myeloma patients after transplant of autologous long-term marrow cultures. *Human Gene Therapy*, **10**, 1953–1964.

Chemoprotection

Abonour R, Williams DA, Einhorn L *et al.* (2000) Efficient retrovirus-mediated transfer of the multidrug resistance 1 gene into autologous human long-term repopulating hematopoietic stem cells. *Nature Medicine*, **6**, 652–658.

Bunting KD, Galipeau J, Topham D *et al.* (1998) Transduction of murine bone marrow cells with an MDR1 vector enables ex vivo stem cell expansion, but these expanded grafts cause a myeloproliferative syndrome in transplanted mice. *Blood*, **92**, 2269–2279.

Guild BC, Mulligan RC, Gros P *et al.* (1988) Retroviral transfer of a murine cDNA for multidrug resistance confers pleiotropic drug resistance to cells without prior drug selection. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 1595–1599.

Hesdorffer C, Ayello J, Ward M *et al.* (1998) Phase I trial of retroviral-mediated transfer of the human MDR1 gene as marrow chemoprotection in patients undergoing high-dose chemotherapy and autologous stem-cell transplantation. *Journal of Clinical Oncology*, **16**, 165–172.

Sorrentino BP, Brandt SJ, Bodine D *et al.* (1992) Selection of drug-resistant bone marrow cells *in vivo* after retroviral transfer of human MDR1. *Science*, **257**, 99–103.

Ward M, Richardson C, Pioli P *et al.* (1994) Transfer and expression of the human multiple drug resistance gene in human CD34+ cells. *Blood*, **84**, 1408–1414.

Gene augmentation therapy

Aiuti A, Slavin S, Aker M *et al.* (2002) Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science*, **296**, 2410–2413.

Blaese RM, Culver KW, Miller AD *et al.* (1995) T lymphocyte-directed gene therapy for ADA-SCID: initial trial results after 4 years. *Science*, **270**, 475–480.

Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G *et al.* (2000) Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science*, **288**, 669–672.

Hoogerbrugge PM, van Beusechem VW, Fischer A *et al.* (1996) Bone marrow gene transfer in three patients with adenosine deaminase deficiency. *Gene Therapy*, **3**, 179–183.

Kay MA, Manno CS, Ragni MV *et al.* (2000) Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nature Genetics*, **24**, 257–261.

Kohn DB, Weinberg KI, Nolta JA *et al.* (1995) Engraftment of gene-modified umbilical cord blood cells in neonates with adenosine deaminase deficiency. *Nature Medicine*, **1**, 1017–1023.

Kohn DB, Weinberg KI, Nolta JL. (1997) PEG-ADA reduction in recipients of ADA gene transduced autologous umbilical cord blood CD34+ cells. *Blood*, **90**, 404a.

Onodera M, Ariga T, Kawamura N *et al.* (1998) Successful peripheral T-lymphocyte-directed gene transfer for a patient with severe combined immune deficiency caused by adenosine deaminase deficiency. *Blood*, **91**, 30–36.

Gene therapy of hematological malignancies

Bonini C, Ferrari G, Verzeletti S *et al.* (1997) HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia. *Science*, **276**, 1719–1724.

Guinan EC, Gribben JG, Boussiotis VA *et al.* (1994) Pivotal role of the B7: CD28 pathway in transplantation tolerance and tumor immunity. *Blood*, **84**, 3261–3282.

Nanda NK, Sercarz EE. (1995) Induction of anti-self-immunity to cure cancer. *Cell*, **82**, 13–17.

Pardoll D. (1992) Immunotherapy with cytokine gene-transduced tumor cells: the next wave in gene therapy for cancer. *Current Opinion in Oncology*, **4**, 1124–1129.

Pardoll DM. (1993) Cancer vaccines. *Immunology Today*, **14**, 310–316.

Pizzoferrato E, Chu NR, Hawley TS *et al.* (1997) Enhanced immunogenicity of B cell lymphoma genetically engineered to express B7-1 and interleukin-12. *Human Gene Therapy*, **8**, 2217–2228.

Chapter 24 Gene expression profiling in the study of lymphoid malignancies

Ulf Klein & Riccardo Dalla-Favera

Introduction

Lymphoid malignancies have long been studied using immunohistochemical, biochemical and genetic approaches to dissect their phenotype and genotype and identify their cellular derivation. The recent advent of DNA microarray analysis, or gene expression profiling, provides another potent tool to analyze the pathogenesis of these tumors. By allowing simultaneous screening of thousands of parameters in the form of expressed genes, microarray analysis has already had a major impact on our understanding of lymphoid neoplasia.

DNA microarrays contain a large number (tens of thousands) of either polymerase chain reaction (PCR) products or shorter oligonucleotides representing distinct mRNA sequences, which are spotted on glass slides. RNA transcripts from the cells under study are labeled with fluorescent tags and hybridized to the microarrays, and the respective fluorescent signals are acquired by scanning the gene chip. The normalized data generated from the hybridizations of various samples are then comparatively analyzed by applying biostatistical methods to determine differences in gene expression.

The enormous number of data points resulting from microarray experiments, especially in a large study, called for the development of hierarchical clustering algorithms that can identify patterns of gene expression among the cell types under study (Figure 24.1). With these algorithms, unsupervised clustering can be used to identify cell types which have not been classified *a priori*, whereas supervised analysis allows the identification of genes expressed differentially between samples defined *a priori* according to a given criterion; for example, cell type or genotype.

The pattern of expressed genes identified by supervised or unsupervised clustering that specifies a cell type is conventionally called a 'signature'. These signatures are then used for a number of goals in the study of disease pathogenesis, and eventually for the development of novel clinical approaches for these diseases. Thus, gene expression signatures may help to distinguish novel tumor subtypes that are not currently identifiable by other diagnostic means based on single-marker analysis. It would be especially desirable to identify subtypes of tumors that cannot be distinguished by current methods but differ in their clinical outcome. Secondly, the comparative analysis of gene expression profiles derived from lymphoid malignancies versus those from healthy lymphocytes may help to identify the normal cellular counterparts of the tumors, and provide insight into the molecular mechanisms responsible for malignant development. Thirdly, gene expression profiling may allow the activity of certain signaling pathways to be monitored in the cell type under study. To this end, either the respective profiles are scanned for concerted changes in the expression of genes known to be associated with a particular signaling pathway or transcriptional response, or the gene expression signatures specific for a certain signaling or transcriptional activation are generated *in vitro* and then tracked for their occurrence in the respective *in vivo* signatures. These results may provide insights into the specific stimuli to which a cell is subjected. Finally, the comparative analysis of a large number of samples, especially when using supervised analysis tools, allows the identification of genes that are specifically up- or downregulated in B-cell malignancies. Following validation of the corresponding gene products in tumor biopsies, such molecules may represent potential diagnostic or therapeutic targets.

Identification of tumor subtypes

Certain categories of lymphoid malignancies, such as diffuse large B-cell lymphomas (DLBL), exhibit marked clinical heterogeneity, suggesting that what is presently recognized as a

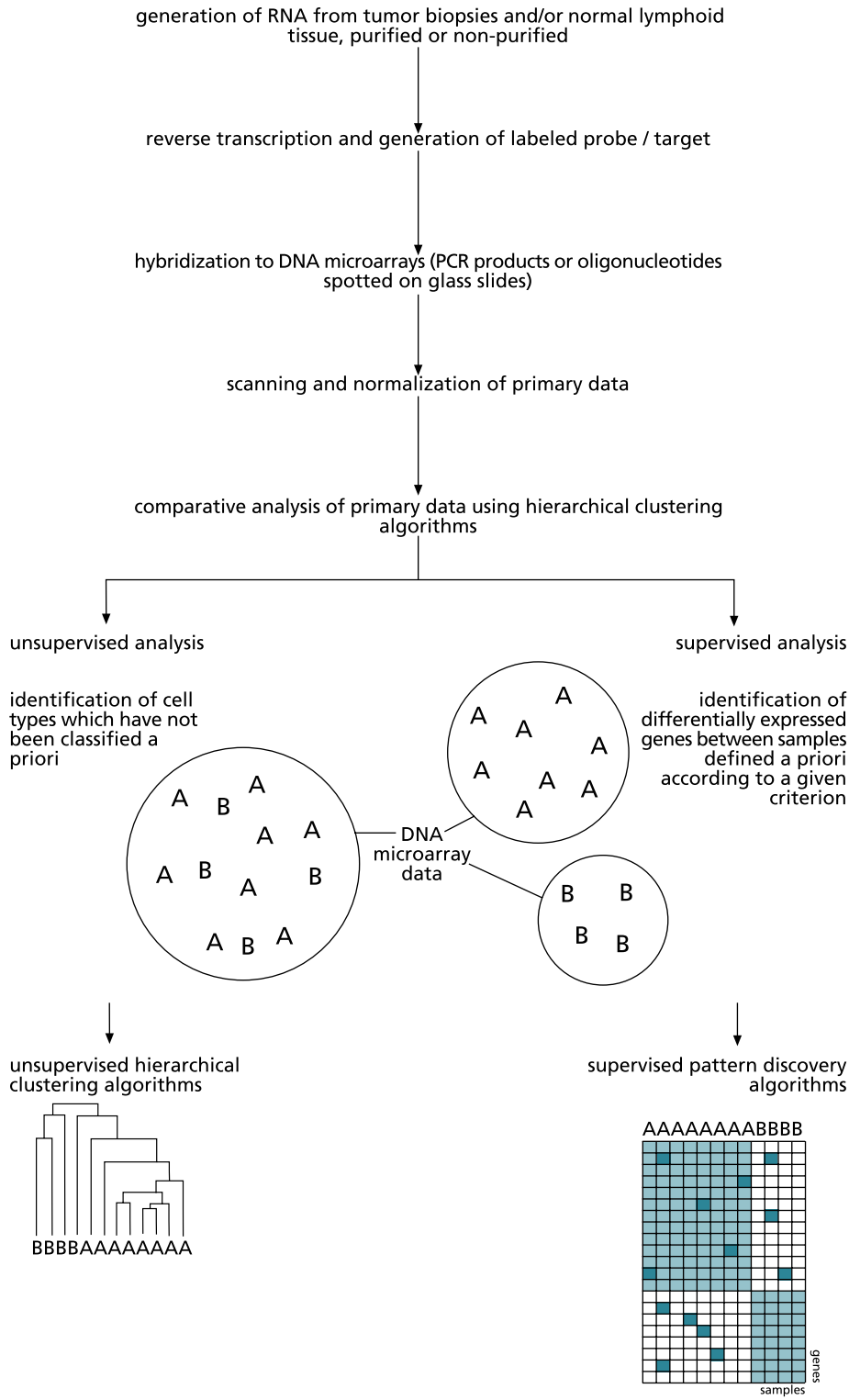


Fig. 24.1 General strategy for the identification of tumor subtypes and genes expressed differentially among tumor subtypes by DNA microarray analysis

In the matrix (lower right), blue and white identify upregulated and downregulated genes, respectively. Columns represent individual samples and rows correspond to genes.

single disease entity may in fact include distinct pathological conditions. By enabling the simultaneous screening of thousands of parameters, DNA microarray analysis substantially enhances the sensitivity of analyses that are aimed at identifying tumor subtypes, which may eventually lead to improved diagnostic approaches that can predict clinical outcome with higher confidence.

The feasibility of the molecular classification of cancer by gene expression profiling was first demonstrated by its ability to distinguish acute myeloid leukemia (AML) from acute lymphoblastic leukemia (ALL). Without previous knowledge about the origin of tumor specimens, gene profiling could determine whether new leukemia cases are derived from ALL or AML. The principal steps of this approach are summarized in Figure 24.2a.

As mentioned above, DLBL represent a category of lymphoid malignancies that appears very heterogeneous. Using a gene chip enriched for genes and expressed sequence tags (ESTs) expressed in B lymphocytes (a 'lymphochip'), a panel

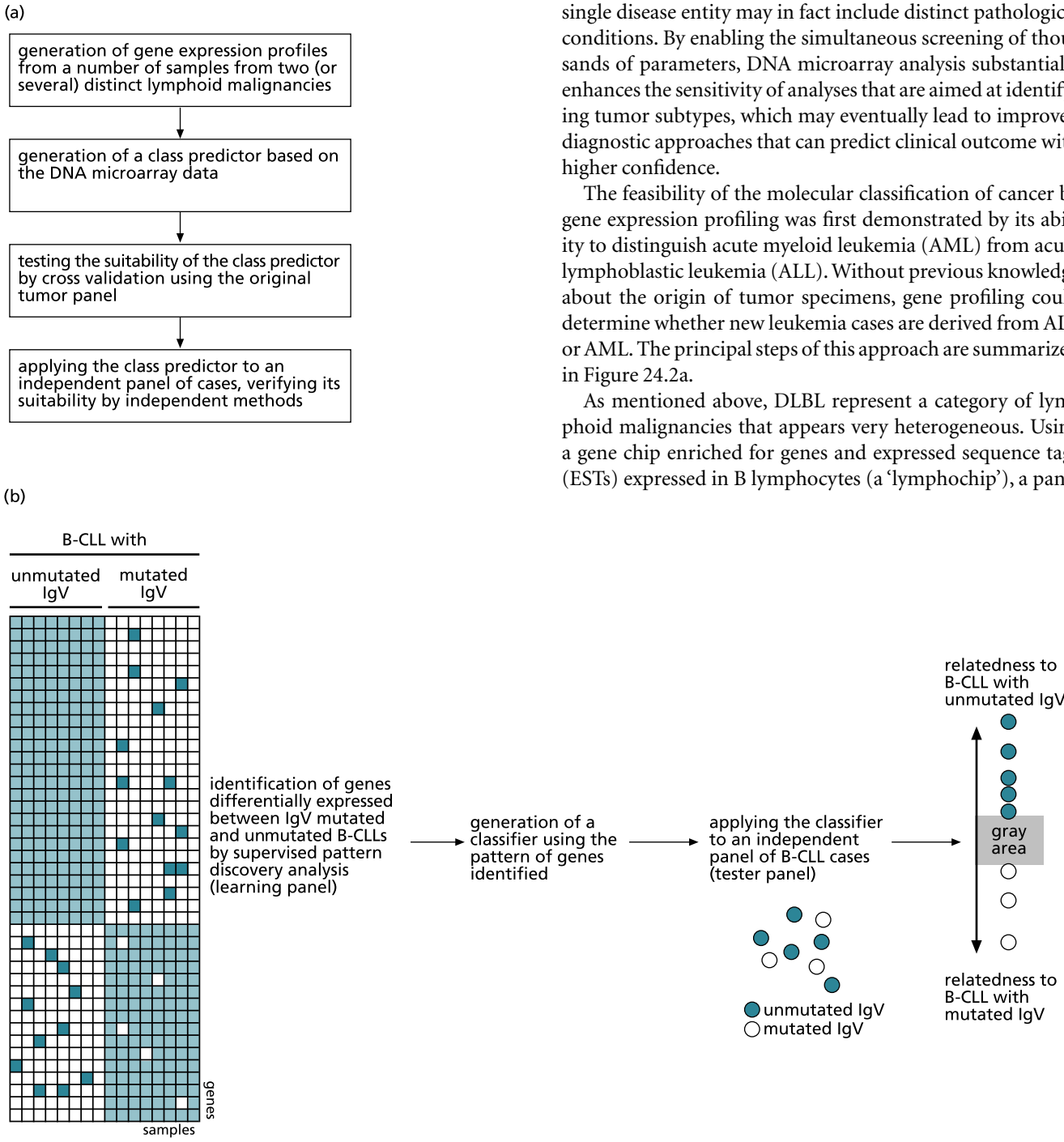


Fig. 24.2 Class predictors (classifiers)

(a) Generation and testing of a class predictor by gene expression profiling to identify tumor subtypes.

(b) Generation of a class predictor by supervised pattern discovery. Genes differentially expressed between two tumor subtypes that differ in the level of IgV gene somatic hypermutation are identified by supervised analysis. A classifier is generated using these genes and is applied to an unclassified panel. The relatedness of the respective cases to either subtype is scored and quantified. Upregulated genes are shown in blue and downregulated genes in white. Columns represent samples and rows correspond to genes. The arrows in the diagram to the right denote increasing relatedness to either B-CLL subtype. The relatedness of cases outside the gray area to either of the B-CLL subtypes is statistically significant. B-CLL, B-cell chronic lymphocytic leukemia.

of DLBL cases could be separated into two subgroups, one that has characteristics of germinal center B cells and another that resembles *in vitro*-activated B cells. Significantly, the clinical data from these two subgroups suggested that the germinal center-like DLBL have a more benign clinical course. However, subsequent studies, using either the same or a different type of DNA microarray, found that DLBLs can be subdivided into at least three groups by gene expression profiling, and even hinted at the existence of several additional subgroups. Consequently, separate research groups may yield slightly different conclusions, especially when studying a complex disease. This may reflect (1) slight differences in the individual diagnostic criteria applied by the respective laboratories for clinical clarification of the cases, (2) differences in the preparation of the samples, (3) the type of microarrays, or (4) the biostatistical analysis. With regard to DLBL, although gene expression profiling may allow DLBL cases to be broadly subdivided into phenotypic groups, further studies are needed to establish a firm correlation between DLBL subtypes identified by gene expression profiling on the one hand and by their clinical behavior on the other.

Compared with other B-cell malignancies, B-cell chronic lymphocytic leukemia (B-CLL) was long considered a relatively homogeneous disease, primarily because of the uniform morphology of the tumor cells. Genetic analyses could identify subgroups of B-CLLs that show particular genomic lesions, which most frequently represent deletions. Furthermore, unlike other tumor types, B-CLLs can be subdivided into those that harbor somatic hypermutations in their rearranged immunoglobulin variable (IgV) genes and those with germline IgV sequences. Notably, the two subgroups also differ in their clinical prognosis, the IgV-mutated B-CLLs showing a more benign disease course. On this background, it was surprising to find that global gene expression analyses performed independently by two laboratories revealed few gene expression differences among the two B-CLL subgroups (approximately 100 among more than 10 000 genes). Nonetheless, the small set of genes distinguishing IgV-mutated and unmutated B-CLL cases could assign unclassified panels of B-CLL cases into either of the two subgroups with high confidence. Based on the example of B-CLL, Figure 24.2b summarizes the use of a set of genes identified by pattern discovery analysis in the class prediction of unrelated cases. At present, the application of DNA microarray analysis for distinguishing B-CLL subtypes is unlikely to be used widely for the clinical diagnosis of B-CLL cases because it is more laborious and costly than determining the IgV gene hypermutational status. Eventually, however, this set of genes may be helpful in clarifying the diagnosis of ambiguous cases (e.g. those showing a low level of IgV somatic hypermutation), either by performing microarray experiments or by establishing biochemical or immunohistochemical assays to detect the most informative B-CLL subtype-specific molecules.

Normal cellular counterparts and the mechanism of transformation

Ever since it became possible to classify lymphoid tumors as B-cell- or T-cell-derived malignancies, the focus has shifted to further elucidate the stage of differentiation from which each tumor type derives within each lymphoid lineage. To understand the precise developmental stage of a B-cell malignancy has obvious implications for understanding its pathogenesis: it may become possible to dissociate the components of the malignancy that represent the normal cellular stage of development from which the tumor arose from truly disease-associated genes, providing potential insights into the mechanism of transformation. Also, tumor-specific gene products thus identified can be studied further for their involvement in the pathogenesis of the malignancy.

Early studies relied previously on morphological and histological comparison with normal lymphoid tissue to determine the derivation of lymphoid tumors. Subsequently, these analyses were refined by the use of monoclonal antibodies against lymphoid-derived cell surface antigens. This phase was followed by the effort to determine the presence of somatic hypermutation in the rearranged IgV genes of B-cell malignancies as well as B-cell subpopulations isolated from healthy individuals as a marker of mature, antigen-experienced B cells. The combination of cell surface markers used to isolate the phenotypically diverse normal B-cell subsets and the concomitant analysis of the IgV genes led to the establishment of a B-cell developmental scheme that was used to assign various B-cell tumor subtypes to defined differentiation stages.

These undertakings clarified the derivation of several tumor entities, as exemplified by the case of Burkitt lymphoma, which, based on the expression of markers associated with bone marrow development, was long thought to be derived from an early B cell with actively rearranging antigen receptor genes, but was eventually classified as a germinal center-derived B-cell tumor on the basis of the presence of hypermutated IgV genes. Still, these methods could not conclusively determine the derivation of various tumor entities, most notably DLBL and B-CLL. However, since gene expression profiling substantially increases the number of parameters available for comparing tumor and normal B cells, it should permit a decidedly more conclusive assignment of a tumor to a normal B-cell subpopulation, also by way of outnumbering ectopically expressed genes whose expression could mar such comparisons. A notable advantage of gene profiling conducted with a

common DNA microarray is that comparisons are relative, so that the actual identities of the gene products are irrelevant; genes are simply scored as up- or downregulated. Instead, it is the quantity of the gene expression changes between cell subsets that is measured and used to determine developmental relatedness (Figure 24.3).

On the basis of IgV gene mutation studies, DLBL was shown early on to be derived from B cells residing in the germinal center of the peripheral lymphoid organs. However, as detailed in the previous section, this tumor group is extremely heterogeneous, both in morphological/histological presentation and clinical prognosis. Using genome-wide gene expression profiling, DLBL cases could be classified into three main subtypes, one being related to germinal center B cells, one related to *in vitro*-activated B cells, and a third, less definable subtype. While the relationship between the germinal center B-cell-related DLBL subtype and germinal center B cells is strongly suggested by the comparison with *ex-vivo* cells, the precise normal cellular counterpart is not so evident for the DLBL tumors related to *in vitro*-activated B cells. This is because the existence and localization of such cells in the lymphoid tissues is unclear as no B cell with such a phenotype has yet been described. Nevertheless, although questions remain open regarding the cellular derivation of DLBL subtypes, this example represents a first demonstration of the suitability of gene expression profiling for assigning lymphoid malignancies to normal B-cell developmental stages.

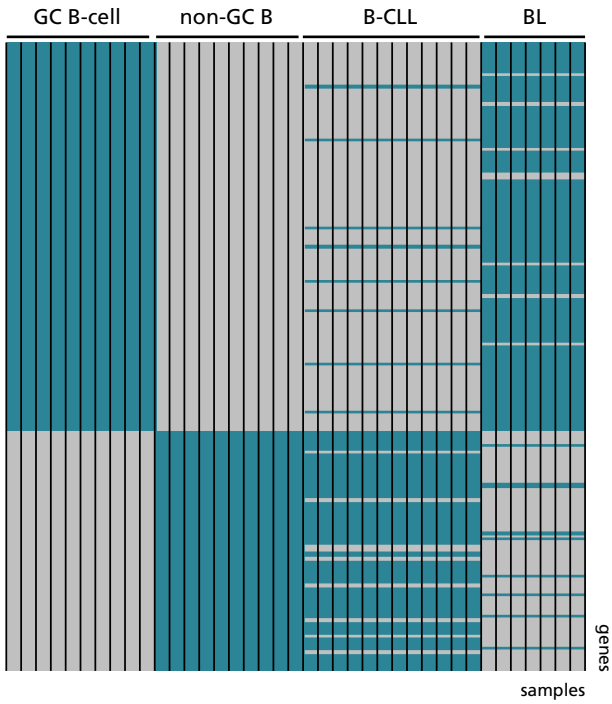


Fig. 24.3 Identification of the histological derivation of B-cell malignancies

Genes expressed differentially between normal B-cell subpopulations are identified by supervised pattern discovery analysis. The expression level of the respective genes in different subtypes of B-cell malignancy [B-cell chronic lymphocytic leukemia (B-CLL) and Burkitt lymphoma (BL)] is shown along with the differentially expressed genes. The relatedness of the B-CLL cases to the non-germinal center (GC) B-cells and that of the BL cases to the GC B-cells is visible from the color-coded expression values and can be expressed quantitatively by statistical analysis. Upregulated genes are shown in blue and downregulated genes in gray.

A methodologically different approach was employed in an attempt to identify the normal cellular counterpart of B-CLL. This malignancy is unique among B-cell derived tumors in that B-CLL cases carry either somatically mutated or unmutated IgV genes in almost equal fractions, suggesting that only a subset of cases is derived from cells that have passed the germinal center (see previous section). Also, the surface marker expression ($CD5^+CD23^+CD27^+$) of B-CLL tumors is inconsistent with that of any known normal B-cell subpopulation. To define the normal counterpart of B-CLL, specific expression profiles were first established by supervised pattern discovery analysis for each of the developmentally distinct B-cell subpopulations, followed by the tracking of the respective profiles in the B-CLL gene expression data. The relatedness of a given B-CLL biopsy to the subset-specific profile can be expressed quantitatively by a measure of the statistical significance (*P* value). Using this approach, a relationship of B-CLL to a particular post-germinal center developmental stage, the antigen-experienced (memory) B cell, was uncovered, a result that is consistent with other phenotypic and genotypic properties, suggesting that B-CLL is derived from the malignant transformation of an antigen-experienced B cell.

A subtype of ALL that carries a translocation involving the mixed-lineage leukemia gene (*MLL*) has a particularly poor prognosis. Gene expression profiling of this subgroup versus ALLs without the respective translocation and AML showed that the ALL subgroup with *MLL* translocation displays a unique gene expression profile with characteristic features of a distinct early hematopoietic progenitor. This cell type is clearly distinct in its phenotype from both the presumed normal cellular counterparts of ALL without translocations and that which gives rise to AML. Therefore, ALL with *MLL* translocation has been proposed to represent a distinct disease.

Another finding of the *MLL* analysis has possible implications for the mechanism of transformation of this tumor entity. Among the genes appearing in the *MLL*-specific signature, a subset, namely the *HOX* genes, represent known target genes of the *MLL* fusion gene product that has been generated by the chromosome translocation. Since one of the *HOX* genes, *HOXA9*, is known to induce leukemia in mice upon overexpression, its gene product (or other family members) might represent an important component of *MLL* tumorigenesis. While future studies need to address this hypothesis

more directly, the example of the *MLL* demonstrates that a chromosomal translocation can determine a unique gene expression program.

Tracking cellular pathways

Because of the large number of genes whose expression in a particular cell type can be monitored simultaneously, DNA microarray experiments are bound to have a profound impact on the analysis of cellular signaling pathways and the cellular response to the activation of transcriptional activators or repressors. The most straightforward way to assess the activity of certain signaling pathways in a cell population is to examine microarray data for changes in the expression of known downstream targets of the respective pathway (Figure 24.4). While this approach relies on previous knowledge about the genes activated in the pathway under study, experimental *in vitro* or *in vivo* systems that recapitulate the activation of a particular signaling pathway can also be used to identify new target genes. For example, cells can be activated through a surface receptor *in vitro*, or cells differing genetically at defined loci (e.g. corresponding wild-type and gene knockout cells) can be studied for differences in their response following activation (Figure 24.4). Likewise, pharmacologically active compounds can be tested in a cell type or in an *in vitro* system in which the relevant signaling pathway has been activated. Transcription factors can be expressed ectopically in a cell, in either a constitutive or an inducible fashion. Regardless of the approach taken, the effects of the stimuli (i.e. the gene expression changes) are determined by comparison with a negative or uninduced control. The identified gene expression profiles can then be tracked within the gene expression data derived from normal or malignant cell populations, potentially providing information about the activity of a particular signaling pathway or transcriptional response in a given cell population. However, the interpretation might be obscured by the connectivity of cell signaling pathways *in vivo*, and tackling this problem clearly represents a new task for the field of computational biology.

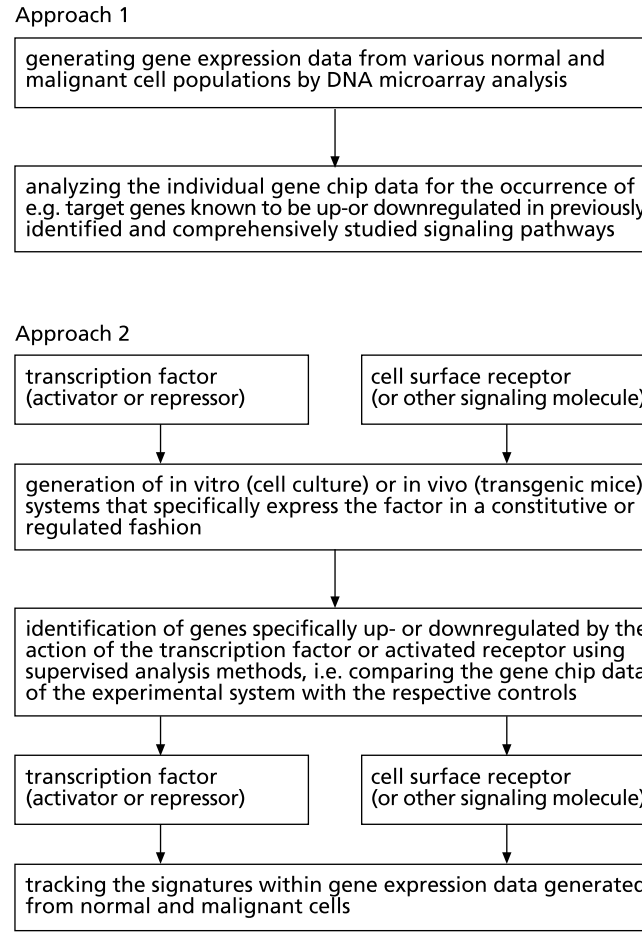


Fig. 24.4 Tracking cellular pathways in normal and malignant cells using two gene expression profiling approaches
Approach 1. DNA microarray data of normal and malignant cells are screened for characteristic gene expression changes of a known transcriptional response or signaling pathway.
Approach 2. A factor-specific gene expression signature is generated *in vitro* or *in vivo*, followed by tracking of the respective profile in normal or malignant cells. Approach 2, in contrast to approach 1, also results in the identification of previously unidentified genes associated with the activation of the particular response or pathway.

Identification of tumor-specific genes

The comparative DNA microarray analysis between a certain lymphoid malignancy and various normal and malignant cell populations (Figure 24.5) truly represents an unbiased way to identify tumor-specific genes, provided that all samples are processed in the same way. The obvious advantage of this approach over large-scale sequencing strategies, such as EST sequencing and serial analysis of gene expression (SAGE), lies in the circumstance that tumor-specific genes are directly identified by the comparative nature of the microarray approach; that is, the simultaneous analysis of a large number of gene chips representing the various subsets. The other approaches require laborious and costly generation of EST or SAGE profiles from the respective cell populations, or specific screening for potential tumor-specific genes by traditional analysis methods in a large panel of normal and malignant cell populations. Clearly, the parallel use of multiple samples of each cell population in a comparative DNA microarray analysis, as opposed to pooling of RNA derived from several samples, allows the identification of differences that would otherwise be lost in procedures involving the pooling. On the

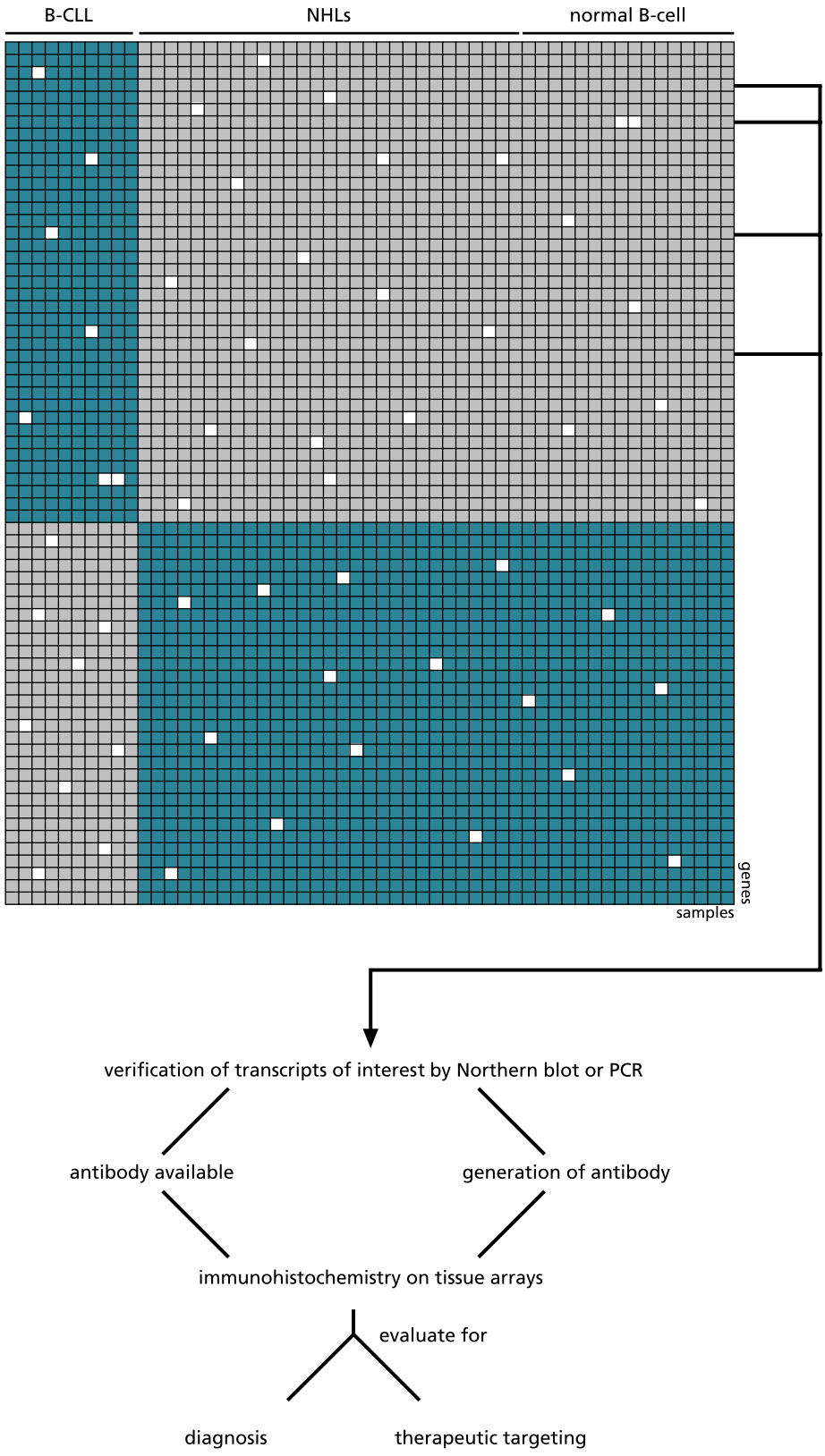


Fig. 24.5 Identification of tumor-specific genes by supervised analysis Genes specifically up- or downregulated in a B-cell malignancy [B-cell chronic lymphocytic leukemia (B-CLL)] relative to the various subtypes of non-Hodgkin lymphoma (NHL) and normal B cells are identified by supervised pattern discovery analysis. The expression of genes of potential interest for diagnosis or therapy, such as cell surface receptors and molecules involved in signaling, is verified by Northern or PCR analysis and eventually by immunohistochemistry. Upregulated genes are shown in blue and downregulated genes in gray. Columns represent individual samples and rows correspond to genes.

negative side, gene chip-based expression analyses, in contrast to EST sequencing and SAGE, are limited by the identity of the gene segments represented on the DNA microarray. Given the pace of development, however, it seems likely that gene chips covering the whole transcribed genome of relevant organisms, including mice and humans, will soon be available.

The expression of potential tumor-specific genes identified by comparative DNA microarray analyses should be verified by PCR or, preferably, Northern blot analysis. This is absolutely required, not only because the regions representing a specific gene on the DNA microarray cover only part of that gene, but also because [this is especially the case for the oligonucleotide-based gene chips (Affymetrix)] these sequences in general have not usually been tested for the detection of their corresponding RNA transcript. The existence of cell type-specific splice variants of many genes adds another level of complexity to this issue. Ideally, an attempt should also be made to detect the presumed gene product in tissue sections by immunohistological analysis or flow-cytometric analysis of cell suspensions.

Acknowledgments

We thank Andrea Califano, Gustavo Stolovitzky, Yuhai Tu, Katia Basso and Vladan Miljkovic for their constant input in some of the gene expression profiling experiments described here, and Richard Baer for comments on the manuscript. Work in the authors' laboratory was supported by the National Cancer Institute, the Leukemia Research Fund, and the Human Frontiers Science Program.

Further reading

Introduction

Brown PO, Botstein D. (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genetics*, **21**, 33–37.

Califano A, Stolovitzky G, Tu Y. (2000) Analysis of gene expression microarrays for phenotype classification. *Proceedings of the International Conferences on Intelligent Systems in Molecular Biology*, **8**, 75–85.

Eisen MB, Spellman PT, Brown PO *et al.* (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14863–14868.

Lockhart DJ, Dong H, Byrne MC *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnology*, **14**, 1675–1680.

Tamayo P, Slonim D, Mesirov J *et al.* (1999) Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 2907–2912.

Identification of tumor subtypes

Alizadeh AA, Eisen MB, Davis RE *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, **403**, 503–511.

Golub TR, Slonim DK, Tamayo P *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531–537.

Klein U, Tu Y, Stolovitzky GA *et al.* (2001) Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *Journal of Experimental Medicine*, **194**, 1625–1638.

Rosenwald A, Alizadeh AA, Widhopf G *et al.* (2001) Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *Journal of Experimental Medicine*, **194**, 1639–1647.

Rosenwald A, Wright G, Chan WC *et al.* (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *New England Journal of Medicine*, **346**, 1937–1947.

Shipp MA, Ross KN, Tamayo P *et al.* (2002) Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nature Medicine*, **8**, 68–74.

Normal cellular counterparts and mechanism of transformation

Alizadeh AA, Eisen MB, Davis RE *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, **403**, 503–511.

Armstrong SA, Staunton JE, Silverman LB *et al.* (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics*, **30**, 41–47.

Klein U, Tu Y, Stolovitzky GA *et al.* (2001) Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *Journal of Experimental Medicine*, **194**, 1625–1638.

Rosenwald A, Wright G, Chan WC *et al.* (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *New England Journal of Medicine*, **346**, 1937–1947.

Shipp MA, Ross KN, Tamayo P *et al.* (2002) Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nature Medicine*, **8**, 68–74.

Tracking cellular pathways

Coller, HA, Grandori C, Tamayo P *et al.* (2000) Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 3260–3265.

Feske S, Giltnane J, Dolmetsch R *et al.* (2001) Gene regulation mediated by calcium signals in T lymphocytes. *Nature Immunology*, **2**, 316–324.

Glynne R, Akkaraju S, Healy JI *et al.* (2000) How self-tolerance and the immunosuppressive drug FK506 prevent B-cell mitogenesis. *Nature*, **403**, 672–676.

Shaffer AL, Yu X, He Y *et al.* (2000) BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity*, **13**, 199–212.

Identification of tumor-specific genes

Alizadeh AA, Eisen MB, Davis RE *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, **403**, 503–511.

Klein U, Tu Y, Stolovitzky GA *et al.* (2001) Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *Journal of Experimental Medicine*, **194**, 1625–1638.

Rosenwald A, Alizadeh AA, Widhopf G *et al.* (2001) Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *Journal of Experimental Medicine*, **194**, 1639–1647.

Appendix 1 Glossary

Adenine (A) Nitrogenous (purine) base, one member of the base pair A–T (adenine–thymine).

Allele Alternative forms of a gene occupying the same locus on homologous chromosomes; segregate at meiosis.

Allogeneic BMT *Allogeneic bone marrow transplantation*. Marrow from an HLA-matched donor (related or unrelated) is administered to the patient following 'conditioning' (usually chemotherapy with radiotherapy). *Cf.* autologous BMT.

Alu sequences DNA sequences recognized by the restriction enzyme *Alu I*; Alu sequences are represented ~300 000 times in the human genome.

Anergy Immunological unresponsiveness to specific antigenic re-challenge.

APC *Antigen presenting cell*. Cells capable of triggering naive T cell response to antigen, leading to clonal T cell expansion. APCs include dendritic cells, activated B cells and macrophages.

Apoptosis Programed cell death involving nuclear DNA fragmentation.

ARMS *Amplification refractory mutation system*—a PCR method that allows amplification of a single specific allele due to specific primers used, i.e. the primers bind only to the mutant allele.

Autologous BMT Patient's own bone marrow is collected following treatment (presumed disease-free) and reinfused in order to reconstitute the hematopoietic system.

Autoradiography Detection of radioactively labeled molecules on radiographic film. For analysis of nucleic acids, single-stranded DNA is immobilized onto a membrane before adding radiolabeled probe. Allows analysis of length and number of DNA fragments after separation by gel electrophoresis.

Autosome Non-sex chromosome.

BAC Bacterial artificial chromosome; allows propagation of large DNA molecules (up to 300kb).

Base pair (bp) A pair of complementary nucleotide bases in a duplex DNA or RNA molecule.

bcl-2 Protein that inhibits apoptosis; overexpressed in, e.g., follicular NHL carrying t(14; 18).

Cap Structure at the 5′ end of mRNAs containing a methylated guanine residue.

CDR *Complementarity determining region*. Those components of Ig and TCR molecules that make contact with specific ligand; determine specificity. The CDR regions are the most variable portions of the Ig and TCR molecules.

Centromere Chromosome region to which spindle fibers attach during cell division.

Chromosome painting Labeling of whole chromosome using FISH. Involves mixture of different sequences from single chromosome.

Clonality Determination of clonal origin of cells.

Clone Population of cells or DNA molecules arising from a single progenitor.

Cloning The process of asexually producing a group of cells (clones), all genetically identical, from a single ancestor. In recombinant DNA technology, the use of DNA manipulation procedures to produce multiple copies of a single gene or segment of DNA is referred to as cloning DNA.

Cloning vector DNA molecule from a virus, a plasmid, or the cell of a higher organism into which another DNA fragment of appropriate size can be integrated without loss of the vector's capacity for self-replication; vectors introduce foreign DNA into host cells, where it can be reproduced in large quantities.

Codon DNA or corresponding RNA sequence of three base pairs that codes for a particular amino acid or termination signal.

Complementary DNA (cDNA) DNA copy of a messenger RNA generated by reverse transcriptase.

Consensus sequence DNA or amino acid sequence that specifies the most commonly found DNA base or amino acid at each position in a sequence of similar DNA or amino acid sequences.

Conserved sequence DNA or protein sequence that has remained essentially unchanged throughout evolution.

Contigs Groups of clones representing overlapping genomic regions.

Cosmid Plasmid DNA containing Cos sites to enable it to be packaged into phage particles. For example, cosmids can be packaged in λ phage particles for infection into *E. coli.* May harbor inserts of ~40kb.

Cyclins Family of proteins that are positive regulators of cell cycle.

Cytosine (C) Pyrimidine base of RNA and DNA; one of cytosine–guanine (C–G) pair.

Diploid Cell or organism containing two complete sets of homologous chromosomes (*cf.* haploid).

DNA polymerase Enzyme that adds nucleotides to a growing chain of DNA in a 5′ to 3′ direction during DNA replication using a DNA strand as a template to copy from.

Domain In a protein this refers to a discrete region of the protein which has a specific function associated with it.

Electrophoresis Method for separating large molecules (nucleic acids or proteins) from a mixture of similar molecules using an electric

current. Molecules travel through agarose or polyacrylamide gel at varying rates depending on their electrical charge and size.

Endonuclease Enzyme that cleaves nucleic acid at internal sites in the nucleotide sequence.

Epitope That part of an antigen that binds to the antigen-binding region of an antibody.

EST Expressed sequence tag.

Exon Segment of a gene which is present in the fully mature RNA after transcription, i.e. contains coding sequences (*cf.* introns).

FACS *Fluorescence activated cell sorting*. Technique used to separate cells from a population on the basis of their expression of specific antigens.

FISH Fluorescence *in situ* hybridization. A technique used to visualize locations on chromosomes which hybridize to specific nucleotide probes. Biotin, digoxigenin or fluorescent dyes are incorporated into cDNA probes using nick-translation. Probes are then hybridized to metaphases or interphase nuclei thus defining gene number and location.

Flow cytometry Analysis of biological material by detection of the light-absorbing or fluorescing properties of cells or subcellular fractions (i.e. chromosomes) passing in a narrow stream through a laser beam. An absorbance or fluorescence profile of the sample is produced. Automated sorting devices, used to fractionate samples, sort successive droplets of the analyzed stream into different fractions depending on the fluorescence emitted by each droplet.

Flow karyotyping Use of flow cytometry to analyze and/or separate chromosomes on the basis of their DNA content.

Frameshift Mutation that shifts the reading frame of triplet codons in a gene during translation of mRNA.

Gardos channel Red cell transport system accounting for selective loss of K^+ in response to increase in intracellular ionized Ca^{2+}.

G banding Technique used to visualize the band patterns of chromosomes on staining with Giemsa.

Gene Unit of heredity that specifies an RNA or mRNA. A gene will also contain intronic regions and regions that control transcription.

Gene expression Process by which a gene's coded information is converted into the structures present and operating in the cell. Expressed genes include those that are transcribed into mRNA and then translated into protein, and those that are transcribed into RNA but not translated into protein (e.g. transfer and ribosomal RNAs).

Gene mapping Construction of a map of different genetic loci based on their physical position with respect to each other.

Gene therapy Insertion of normal DNA directly into cells to correct a genetic defect.

Genetic code The sequence of nucleotides, coded in triplets (codons) along the mRNA, that determines the sequence of amino acids in protein synthesis. The DNA sequence of a gene can be used to predict the mRNA sequence, and the genetic code can, in turn, be used to predict the amino acid sequence.

Genomic library Collection of clones made from a set of randomly generated overlapping DNA fragments representing the entire genome of an organism.

Genotype The genetic make-up of a cell or organism responsible for its appearance.

Grb-2 *Growth factor receptor-bound protein 2*. Participates in downstream signaling after activation of a variety of cellular receptors. Binds to EDGF and PDGF.

Guanine (G) Purine base of RNA or DNA. One member of the base pair G–C (guanine and cytosine).

Haploid A single set of chromosomes (half the full set of genetic material), present in the egg and sperm cells of animals, and in the egg and pollen cells of plants (*cf.* diploid).

Haplotype Combination of closely linked genes. Usually inherited together as a 'block', arising from one chromosome.

Hemizygote Diploid cell or organism that contains only one allele of a gene due to loss of one chromosome of a homologous chromosome pair.

Heterodimer A complex of two non-identical moieties, e.g. proteins such as the T-cell receptor.

Heterozygote Diploid cell or organism that contains different alleles of a gene at one locus on homologous chromosomes.

Homeobox A short stretch of nucleotides whose base sequence is virtually identical in all the genes that contain it. It has been found in many organisms from fruit flies to human beings. In the fruit fly, a homeobox appears to determine when particular groups of genes are expressed during development.

Homodimer Complex of two identical moieties.

Homotetramer Tetramer comprising four identical subunits.

Hybridization The ability of complementary single-stranded DNA or RNA molecules to form a duplex.

***In situ* hybridization** Use of DNA or RNA probe to detect the presence of the complementary DNA sequence in cloned bacterial or cultured eukaryotic cells.

Interphase Period of the mitotic cell cycle between one mitosis and the next.

Intron Intervening sequence. Segments within the coding region of a gene which are not present in the fully mature RNA after transcription due to removal by splicing.

In vitro Taking place outside a living organism (literally '*in glass*'; *cf. in vivo*).

In vivo Taking place within a living organism ('*in life*').

Karyotype Photomicrograph of an individual's chromosomes arranged in a standard format showing the number, size and shape of each chromosome type.

Kilobase (kb) 1 000 nucleotides.

Kinase An enzyme that phosphorylates a substrate.

Knockout The ability to remove a specific gene in a cell or organism by molecular techniques.

Leucine zipper Leucine-rich domain of a protein that allows protein–protein interaction.

Library An unordered collection of cloned DNA whose relationship to each other can be established by physical mapping.

Linkage The tendency for two genes in close proximity on a chromosome to be inherited together.

Locus The position of a gene on a chromosome.

LOH *Loss of heterozygosity*. Deletions believed to occur during tumor development. Identified as loss of an allele at a specific locus.

Lyonization X chromosome inactivation in mammals. One X is randomly inactivated in each cell.

Maternal inheritance Preferential carriage of a gene by the maternal parent.

Megabase (Mb) 10^6 nucleotides.

Messenger RNA (mRNA) The mature transcript from a gene transcribed by RNA polymerase which specifies the order of amino acids during mRNA translation to protein.

Metaphase Stage in mitosis when the parental and newly synthesized chromosomes are maximally condensed but prior to their segregation to opposite spindle poles.

MHC *Major histocompatibility complex*. A polymorphic family of genes that are involved in mediating T cell immune responses; present peptides to TCR.

Minimal residual disease Low level disease present following therapy; generally not detected using standard techniques such as light microscopy, but requires molecular techniques such as PCR for detection.

Missense mutation Single-base substitution causing incorporation of inappropriate amino acid into a protein.

Monosomy Condition in which one member of a chromosome pair is missing.

Multidrug resistance (MDR) Cell mechanism conferring drug resistance to a wide variety of chemotherapeutic agents.

Mutation A transmissible change in nucleotide sequence which leads to a change or loss of normal function encoded by that nucleotide sequence.

Nonsense mutation Mutation resulting in the premature termination during protein synthesis.

Northern blot Technique which transfers RNA molecules after size fractionation on gels to filter papers for hybridization to specific probes.

Nucleotide A subunit of DNA or RNA consisting of a nitrogenous base (adenine, guanine, thymine or cytosine in DNA; adenine, guanine, uracil or cytosine in RNA), a phosphate molecule and a sugar molecule (deoxyribose in DNA and ribose in RNA).

Oncogene Mutated gene which is normally involved in the correct control of cell division such that disruption of the normal gene function leads to cell immortalization and transformation.

Open reading frame Series of triplet codons in the coding region of a gene that lie between the signals to start and stop translation.

PBSC *Peripheral blood stem cell*. Following administration of chemotherapy ± growth factor, e.g. G-CSF, stem cells enter the peripheral circulation and are collected by leucopheresis.

PCR *Polymerase chain reaction*. Technique to amplify a target DNA sequence by multiple rounds of DNA synthesis. Involves a heat-stable DNA polymerase (e.g. Taq, isolated from hot spring bacterium, *Thermus aquaticus*) and two oligonucleotide primers (generally ~20 bases), one complementary to the (+) strand at one end of the sequence to be amplified and the other complementary to the (−) strand at the other end. Newly synthesized DNA can subsequently serve as additional templates for the same primer sequences, allowing a million-fold increase in DNA sequence after 30 cycles of primer annealing, strand elongation, and DNA melting.

PFGE *Pulsed field gel electrophoresis*. An electrophoretic technique to separate very large molecules of DNA by periodically altering the direction of the electric field through which the samples are migrating.

Phenotype The observable characteristics of a cell or organism resulting from the expression of the cell's genotype.

Plasmid A circular autonomously replicating, extrachromosomal DNA molecule. Some plasmids are capable of integrating into the host genome.

Pleckstrin homology domain First noted in platelet protein pleckstrin, 100 residues, function unknown.

Polymorphism DNA sequence variation among individuals. May be used as linkage marker.

Positional cloning Method for identifying the gene responsible for a genetic disease in the absence of a transcript or protein product; relies on the use of markers tightly linked to the target gene.

Primer Short oligonucleotide sequence that provides the starting point for polymerases to copy a nucleotide sequence and make a double strand.

Probe DNA or RNA molecules (single stranded) of specific base sequence, may be labeled radioactively (e.g. ^{32}P) or using non-radioactive means (e.g. digoxigenin) to detect complementary sequence on Southern blot, etc.

Promoter DNA sequence that targets RNA polymerase to a gene for transcription.

Proto-oncogene Refers to a normal gene involved in controlling cell division which, when mutated, becomes an oncogene.

Pseudogene A duplicated gene that has become non-functional.

Purine base Organic base containing two heterocyclic rings that occurs in nucleic acids (adenine and guanine in DNA and RNA).

Pyrimidine base Organic base containing one heterocyclic ring that occurs in nucleic acids. (Cytosine (C) and thymine (T) in DNA; cytosine (C) and uracil (U) in RNA.)

Q banding Banding technique to visualize the banding patterns of chromosomes using quinacrine stain.

Recombination The process by which progeny derive a combination of genes different from that of either parent. In higher organisms, this can occur by crossing over. Also occurs in non-germline cells, e.g. B cell and T cell during production of Ig and TCR genes, respectively.

Reporter gene A gene encoding a product which can be easily measured when introduced into cell, e.g. by transfection.

Restriction enzyme An endonuclease (usually bacterial) that recognizes specific, short nucleotide sequences and cuts DNA at those sites.

Restriction enzyme cutting site Specific nucleotide sequence of DNA at which a particular restriction enzyme cuts the DNA. Some sites occur frequently in DNA (e.g. every several hundred base pairs), others much less frequently (rare-cutter, e.g. every 10 000 base pairs).

Retrovirus RNA virus that replicates by first converting its RNA genome to a double-stranded DNA copy using the enzyme reverse transcriptase.

Reverse transcriptase Enzyme used to make a DNA copy of RNA. Found in retroviruses allowing conversion of their RNA genome into double-stranded DNA which may then integrate into the host genome.

RFLP *Restriction fragment length polymorphism*. Heritable differences in the length of DNA fragments from a specific region of DNA generated by restriction enzymes due to DNA sequence differences.

Ribozyme An RNA molecule with properties of RNAse; cleaves single-stranded RNA.

RNA polymerase Enzyme that makes an RNA copy of a DNA template.

RNA splicing Removal of introns from transcribed RNA to generate a mature mRNA.

RT-PCR *Reverse transcriptase polymerase chain reaction*. Amplification of RNA by PCR after copying of the RNA→cDNA by reverse transcription.

Sequencing Determination of the order of nucleotides (base sequences) in a DNA or RNA molecule or the order of amino acids in a protein.

Somatic mutation Mutation arising in a somatic cell.

Southern blot Technique which transfers DNA molecules after size fractionation on gels to filter papers for hybridization to specific probes.

Src homology 3 (SH3) domain Protein domain on oncoprotein Src; binds to proline-rich domains on other proteins.

Syngeneic Genetically identical (e.g. identical twins).

Taq polymerase Heat-stable DNA polymerase used for DNA amplification (*see* PCR). More recent heat-stable polymerases include VENT and Tth.

TATA box Also called Hogness Box; DNA sequence found in many eukaryotic promoters that binds the TATA binding protein in order to recruit RNA polymerase for transcription. Consensus sequence TATAAAA; specifies position where transcription is initiated.

T-cell receptor Membrane protein complexes that are expressed on T-lymphocytes and recognize specific antigens when associated with MHC molecules.

Telomere Specialized structures at the ends of chromosomes involved in the replication and stability of linear DNA molecules.

Thymine (T) Nitrogenous base, one member of the base pair A–T (adenine–thymine).

Tolerance Reduced ability to mount an immune response to specific antigens *in vivo*.

Transcription Process by which a DNA template (a gene) is copied to RNA by RNA polymerase.

Transcription factor Proteins, other than RNA polymerase, that are required for transcription of all genes.

Transfection The introduction of DNA into cells in culture.

Translation The mechanism by which mRNA is used as a template to synthesize protein on the ribosome.

Tumor suppressor gene A gene that negatively regulates cell division such that mutation in these genes results in uncontrolled cell division and tumor progression, e.g. Rb, p53 genes.

Tyrosine kinase Enzymatic activity catalyzing the attachment of phosphate group to tyrosine residue within a protein molecule.

Uracil (U) Pyrimidine base that replaces the DNA base thymine in RNA molecules; forms base pair with adenine (A–U).

Vector DNA molecule in which DNA sequences can be cloned; agent used to deliver genes for gene transfer.

VNTR Variable number of tandem repeats. Variations in numbers of tandem repeat DNA sequences found at specific loci in different populations. May be used as genetic markers.

Western blot A technique which transfers protein molecules after size fractionation on gels to filter papers for analysis with antibodies.

X inactivation Inactivation of one of the X chromosomes in female somatic cells.

YAC Yeast artificial chromosome. Plasmid DNA which contains DNA sequences that allow plasmid maintenance in yeast cells permitting cloning of very large regions of DNA.

Appendix 2 Cytogenetic glossary

add Addition of chromosomal material of unknown origin to another chromosome
der Derivative of a chromosomal rearrangement
del Deletion of chromosomal material
dup Duplication of chromosomal material; extra copy of part of a chromosome
cen Centromere
Hypo/hyperdiploid Human karyotype with <46 or >46 chromosomes, respectively
i Isochromosome; chromosome whose arms are mirror images of each other
ins Insertion of chromosomal material
inv Rotation of chromosome segment by 180°
mar Marker chromosome. Signifies any structurally rearranged chromosome
p Short (petit) arm of chromosome
q Long arm of chromosome
p+/q+ Addition of DNA to short/long arms, respectively
p–/q– Deletion of DNA from short/long arms, respectively
t Translocation; chromosome segment moves from one chromosome to another. May or may not be reciprocal
ter Terminus (end); e.g. pter, qter = short arm and long arm ends, respectively
+/– Gain/loss of chromosomal material, respectively

Appendix 3 Cluster designation (CD) antigens used in this book

CD	Expression	Comments
CD1a, b, c	Thymocytes, dendritic cells	Ligand for some γδ T cells
CD3	T cells	Complex of molecules that transduce signals from T-cell receptor
CD4	T cells, monocytes, tissue macrophages, microglial cells, EBV-transformed B cells	Binds HLA class II Receptor for HIV env gp120
CD5	Thymocytes and T cells, most T cell malignancies, B-CLL	Binds to CD72
CD8	T cell subset	Binds MHC class I
CD10	T and B cell precursors, BM stromal cells, pre-B ALL	Zinc metalloproteinase, marker for pre-B cell ALL (also termed common ALL antigen, cALLa)
CD15	Neutrophils, eosinophils, monocytes	Also called Lewis-x (Le^x)
CD19	B cells, B cell malignancies	Function unknown
CD20	B cells, B cell malignancies	Function unknown
CD21	Mature B cells, follicular dendritic cells	Complement control protein. Receptor for complement C3d, EBV
CD23	Mature B cells, activated macrophages, eosinophils, follicular dendritic cells, CLL	Low affinity receptor for IgE
CD28	T cell subsets, activated B cells	Receptor for costimulatory signal, binds CD80 (B7-1) and CD86 (B7-2)
CD34	Hematopoietic precursors, capillary endothelium	Ligand for CD62L (L-selectin)
CD38	Early B- and T cells, activated T cells, germinal centre B cells, plasma cells	Function unknown
CD43	Leucocytes (except resting B cells)	Binds CD54 (ICAM-1)
CD55	Hematopoietic and non-hematopoietic cells	Decay accelerating factor (DAF). Binds C3b
CD56	NK cells	Isoform of neural cell adhesion molecule (N-CAM)
CD59	Hematopoietic and non-hematopoietic cells	Binds complement C8 and C9
CD79a, b	B cells	Component of B cell antigen (analogous to CD3)
CD80	Antigen presenting cells	Costimulation. Ligand for CD28 and CTLA- 4
CD86	Antigen presenting cells	Costimulation. Ligand for CD28 and CTLA- 4

Abbreviations: see list on p. xvii

Index

Page numbers in *italics* refer to figures and those in **bold** to tables; note that figures and tables are only indicated when they are separated from their text references.