# Current Topics in Microbiology and Immunology

103

# Retroviruses 1

Edited by P.K. Vogt and H. Koprowski

With 16 Figures

Springer-Verlag
Berlin Heidelberg New York 1983

Prof. Dr. P.K. Vogt
University of Southern California
School of Medicine
Dept. of Microbiology
2011 Zonal Avenue HMR 401
Los Angeles, California 90033
U.S.A.

Prof. Dr. Hilary Koprowski
The Wistar Institute
36th Street at Spruce
Philadelphia, PA 19104
U.S.A.

ISBN-13:978-3-642-68945-1 e-ISBN-13:978-3-642-68943-7
DOI: 10.1007/978-3-642-68943-7



Library of Congress Catalog Card Number 15-12910



Typesetting: Fotosatz Service Weihrauch, Würzburg.

2121/3321-543210

## Table of Contents



*Indexed in Current Contents*

# Avian Endogenous Viral Genes

UGO G. ROVIGATTI* AND SUSAN M. ASTRIN*



## 1 Introduction

Endogenous viruses are defined as germline genes that code for the components of a retrovirus. These genes are present in all cells of all tissues of an animal and are inherited by progeny in a Mendelian manner. Two general methods have been used to detect the presence of endogenous viral sequences in DNA of a given species. First, sequence homology with the genome of a known retrovirus can be demonstrated for the chromosomal DNA. Second, production of viral particles or viral components can be demonstrated for cells of the species. The first evidence for the existence of endogenous viruses came from studies on spontaneous leukemia in the murine system (for review see *Gross* 1958a). There it was also shown that in lymphoid tumors induced by X rays, a murine leukemia virus was produced which caused similar tumors when injected into unirradiated mice (*Gross* 1958b; *Lieberman* and *Kaplan* 1959). Endogenous viral genes have now been shown to be essentially ubiquitous in vertebrate species including man. The purpose of this article will be to describe and discuss the current state of information with respect to the endogenous viral genes of the domestic chicken, *Gallus gallus.* We

* Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

will discuss DNA structure, gene products, and some current ideas about the relationship between endogenous viral genes and disease.

An avian retroviral particle contains an RNA genome composed of two indentical 8-kilobase subunits, which are complexed with reverse transcriptase (a 154000-dalton protein composed of α and β subunits of 92000 and 62000 daltons, respectively). Four other proteins, termed group-specific antigens (gs), are internal to the viral particle (for reviews see *Vogt* and *Hu* 1977; *Eisenman* and *Vogt* 1978). The molecular weights of the four gs antigens are 27000, 19000, 15000, and 12000 daltons, and they are processed from a precursor of 76000 daltons. The viral envelope contains glycoproteins of 85000 daltons and 37000 daltons. In the avian system the envelope glycoprotein, which is the product of endogenous viral genes, is often termed chf or chick helper factor because production of envelope protein by cells complements or helps the production of the envelope-defective Bryan strain of Rous sarcoma virus (*Weiss* 1969; *H. Hanafusa* et al. 1970).

Upon infection of a susceptible cell, the viral RNA is transcribed into a double-stranded DNA provirus, which integrates into the host genome and serves as a template for the synthesis of messenger RNA and RNA to be encapsulated into virions. When such proviral sequences are present in DNA of the germ cells of a species, they are termed endogenous viral genes.

Endogenous viral genes were first demonstrated to be present in chicken cells in the 60s. A complement-fixation assay showed gs antigens to be present in uninfected cells from chicken embryos (*Dougherty* and *DiStefano* 1966). Subsequently, gs expression was shown to segregate as an autosomal dominant gene in crosses between a line which produced the gs antigen and a line which lacked the antigen (*Payne* and *Chubb* 1968). Another endogenous viral gene product, viral envelope or chf, was also identified in embryo cells and could be rescued by infection and recombination with sarcoma or leukosis viruses (*H. Hanafusa* et al. 1970; *Weiss* and *Payne* 1971). Both gs and chf antigen appeared to be present in several different flocks of white leghorn chickens and also in wild jungle fowls (*Weiss* and *Biggs* 1972). It was proposed that these antigens were the products of defective viral genomes which resided in the cellular DNA. However, it soon became clear that cells could also carry intact viral genomes, and that, in certain instances, viral particles could be released spontaneously or after induction by chemical and physical agents. A virus with a distinctive envelope and host range was shown to be the product of endogenous viral genes (*T. Hanafusa* et al. 1970; *Vogt* and *Friis* 1971). This virus was called RAV-0 and later shown to be encoded by a specific genetic locus, *ev 2* (*Astrin* et al. 1980a). A similar virus could be induced from normal chicken cells after treatment with ionizing radiation or chemical carcinogens or mutagens (*Weiss* et al. 1971).

Additional evidence for the presence of endogenous viral genes came from biochemical data that indicated several copies of viral DNA were present in the cellular genomes of cells expressing gs antigens ($gs^+$ cells), as well as of cells lacking these antigens ($gs^-$ cells) (*Rosenthal* et al. 1971; *Baluda* 1972; *Varmus* et al. 1972; *Neiman* 1973). However, RNAs specific for the gs and chf antigens were present only in $gs^+$ $chf^+$ cells, suggesting a transcriptional regulation over the expression of these genes (*Hayward* and *Hanafusa* 1973; *Wang* et al. 1977). However, gs and chf were not always coordinately regulated (*T. Hanafusa* et al. 1972). Another viral function, which does not appear to be expressed in a coordinate fashion with other viral genes, is the reverse transcriptase (*Weissbach* et al. 1972; *Eisenman* et al. 1978). The genetics of the induction and spread of RAV-0 virus have been studied extensively (*Crittenden* et al. 1974, 1977). RAV-0 propaga-

tion and horizontal transmission are complicated by the fact that the cells of only a very few lines of chickens are susceptible to its infection. The possibility of control by two different loci, one specifying the receptors for RAV-0 and the other acting as an epistatic inhibitor of its infectivity, has been studied (*Crittenden* et al. 1974), and recently elucidated (*Robinson* et al. 1981). It had been speculated that the genes controlling expression of gs, chf, and RAV-0 production were regulatory, rather than structural genes (for review see *Tooze* 1973). That this is not the case has been shown only recently (*Astrin* 1978; *Astrin* et al. 1980b).

## 2 The *ev* Loci

With the advent of restriction endonucleases and the Southern blot (*Southern* 1975), it first became possible to look at the structure of a set of unique sequences in a eucaryotic genome, providing one had a probe. Using radiolabeled avian leukosis or sarcoma virus genomic RNA or radiolabeled cDNA (complementary DNA made by reverse transcription), several groups began to look at the structure of the endogenous viral sequences in chickens of different phenotypes. It was immediately obvious that the situation was very complicated. The standard enzymes gave very complex patterns, often containing five or more bands, even with DNA from $gs^-$ $chf^-$ chickens, and there did not appear to be obvious correlations of pattern with phenotype. The situation was eventually elucidated by a combination of both a biochemical and a genetic approach. Using several restriction enzymes to screen a group of more than 150 birds of various phenotypes, *Astrin* (1978) clearly demonstrated that certain $gs^-chf^-$ birds gave the most simple pattern of bands (three bands with *Eco*RI, *Bam*HI, or *Hind*III). In addition, these bands were included in the pattern of all the other birds, and it was clear from the molecular weights that these fragments represented internal fragments of the viral genome, as well as junction fragments containing viral and cellular sequences. On the assumption that the DNA yielding the simple three-band pattern contained only a single endogenous provirus, more than 20 enzymes were screened and an enzyme, *Sac*I, was found which gave but a single band with the test DNA. When *Sac*I was used to cleave the other DNAs, simplified patterns emerged although, in most cases, multiple bands were still present. It was not clear whether each and every band represented an individual provirus, or whether some proviruses yield more than one fragment. However, it was possible at this stage to correlate individual bands with certain phenotypes, and it was also clear that virtually all the DNAs tested contained a common proviral element (*Astrin* 1978).

Further clarification of the situation required a genetic approach. Matings were set up in which homozygous $gs^-chf^-$ birds were mated with birds homozygous for the $gs^+chf^+$, $V^+$, or $gs^-chf^+$ phenotypes, and the $F_1$ progeny were backcrossed to the $gs^-chf^-$ parent. The phenotypes and genotypes (DNA restriction patterns) of the parents' $F_1$ progeny and backcross progeny were determined (*Astrin* et al. 1980b). From these experiments several major conclusions were possible. First, it was apparent that each *Sac*I fragment represented a separate genetic locus for endogenous viral sequences. In addition, there was an unexpectedly large number ($> 10$) of such loci, and one locus, designated *ev* 1, was common to all the white leghorns tested. Finally, it was possible to get an absolute correlation between segregation of a particular band (endogenous virus locus) and segregation of a particular phenotype. Thus, a good case could be made for the

Table 1. Distribution of endogenous retroviral loci in white leghorn chickens

| Chicken flock | No. of birds tested | % of birds containing specific *ev* loci | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *ev*1 | *ev*2 | *ev*3 | *ev*4 | *ev*5 | *ev*6 | *ev*7 | *ev*8 | *ev*9 | *ev*10 | *ev*11 | *ev*12 |
| Inbred lines | | | | | | | | | | | | | |
| $6_3$ | 24[a] | 100 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $7_2$ | 26[a] | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $15_B$ | 60[a] | 100 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| $15_1$ | 15 | 100 | 0 | 100 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 7 | 100 |
| C | 11 | 100 | 0 | 0 | 0 | 0 | 0 | 100[b] | 0 | 0 | 100[b] | 0 | 0 |
| $15_{14}$ | 12 | 100 | 0 | 0 | 0 | 0 | 89 | 0 | 0 | 0 | 100 | 100 | 0 |
| $15_{15}$ | 10 | 100 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 90 | 20 | 0 |
| Noninbred lines | | | | | | | | | | | | | |
| SPAFAS line 11 $gs^-chf^-$ | 200 | 99 | 0 | 0 | 56 | 25 | 0 | 0 | ND[c] | 0 | 0 | 0 | 0 |
| SPAFAS line 11 $gs^+chf^+$ | 10 | 100 | 0 | 100 | 30 | 70 | 20 | 0 | 0 | 0 | 0 | 0 | 0 |
| SPAFAS line 11 $gs^-chf^+$ | 6 | 100 | 0 | 0 | 83 | 0 | 33 | 0 | 17 | 67 | 0 | 0 | 0 |
| Heisdorf and Nelson $gs^+chf^+$ | 2 | 100 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Heisdorf and Nelson $gs^-chf^-$ | 3 | 100 | 0 | 0 | 33 | 67 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| K(−) | 34[a] | 100 | 0 | 0 | 12 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| K16 | 32[a] | 100 | 0 | 100 | 38 | 0 | 25 | 0 | ND | 0 | 0 | 0 | 0 |
| K18 | 24[a] | 100 | 0 | 0 | 71 | 0 | 100 | 0 | 79 | 58 | 0 | 0 | 0 |
| K28 | 39[a] | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

[a] Birds chosen so as to completely represent the gene in the breeding stock of these lines;
[b] Identified by *Sac*I and *Bam*HI digestion, not confirmed by genetic analysis;
[c] ND, not determined, additional restriction endonuclease analysis required for identification

Table 2. Associated phenotypes and identifying restriction fragments for 12 *ev* loci

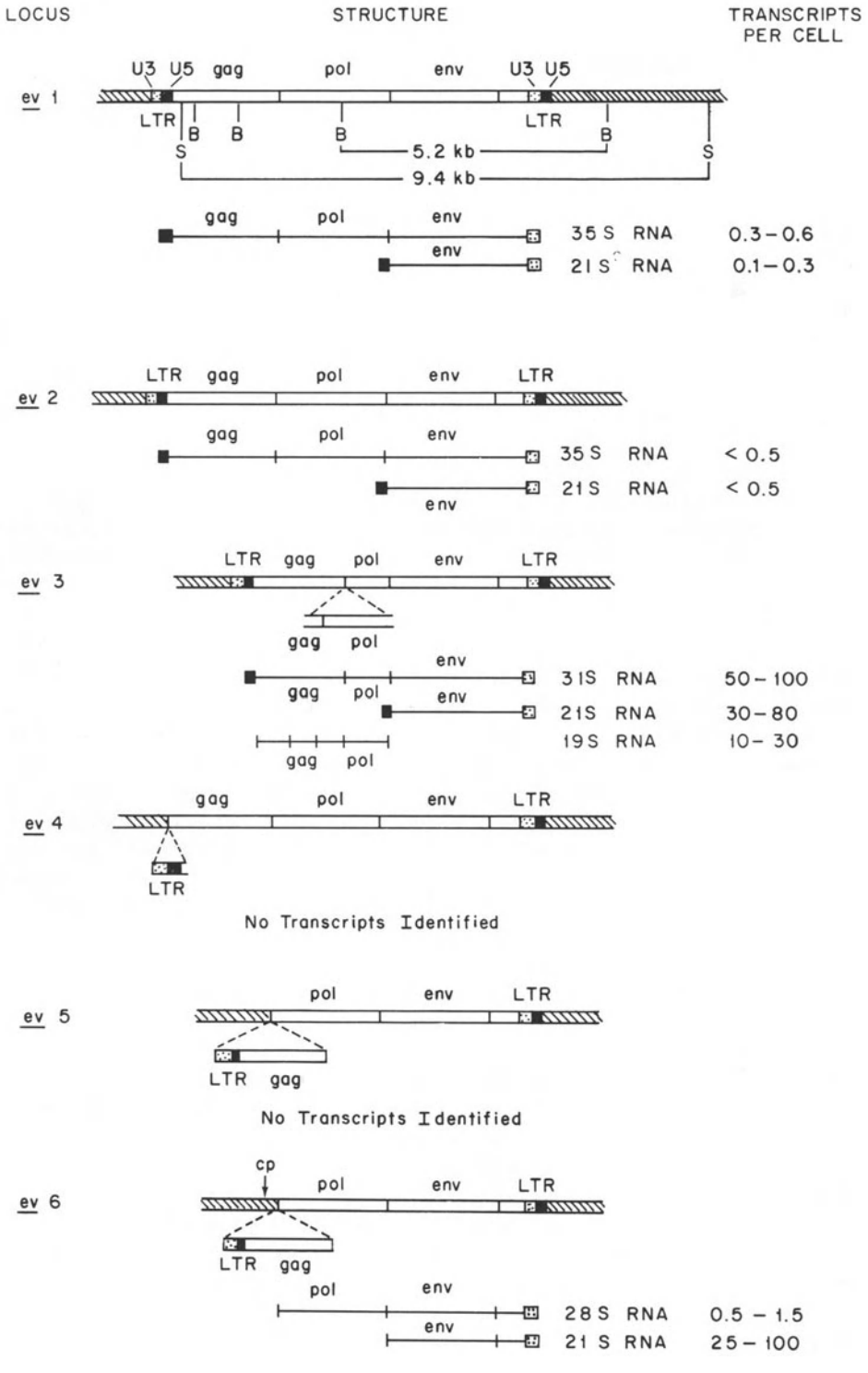
| Locus | Phenotype | Size of major *Sst*I fragment (kilobase pairs)[a] | Size of characteristic *Bam*HI fragment (kilobase pairs)[b] |
|---|---|---|---|
| *ev* 1 | none | 9.4 | 5.2 |
| *ev* 2 | V+ | 6.0 | 8.2 |
| *ev* 3 | gs+chf+ | 6.3 | 7.3 |
| *ev* 4 | none | 8.7 | 7.3 |
| *ev* 5 | none | 19.0 | 13.0 |
| *ev* 6 | gs−chf+ | 21.0 | 4.4 |
| *ev* 7 | V+ | 13.0 | 7.6 |
| *ev* 8 | none | 18.0 | 23.0 |
| *ev* 9 | gs−chf+ | 23.0 | 11.0 |
| *ev* 10 | V+ | 21.0 | 14.0 |
| *ev* 11 | V+ | 13.0 | NI[c] |
| *ev* 12 | V+ | 8.1 | NI[c] |

[a] This fragment is the right-hand virus cell junction fragment (see the restriction map for *ev* 1 in Fig. 1). It contains the majority of the proviral sequences as well as cellular sequences adjacent to the 3′ end of the provirus; [b] This fragment is also the right-hand virus cell junction fragment (see the *ev* 1 map in Fig. 1). It contains the proviral *env* sequences as well as cellular sequences adjacent to the 3′ end of the provirus; [c] Not identified

proposal that a particular locus coded for the proteins characteristic of a particular phenotype. Studies on the genetic content of the DNAs and RNAs associated with each locus (*Hayward* et al. 1980) strengthened these correlations, and it is now accepted that each of the genetic loci identified by the above approach does indeed code for a specific phenotype of endogenous viral gene expression. The individual genetic loci have been designated *ev* loci and numbered sequentially (*Astrin* 1978; *Astrin* et al. 1980b). Table 1 gives the frequency and distribution of the loci in flocks of white leghorns (data from *Tereba* and *Astrin* 1980). Table 2 gives the associated phenotypes and identifying restriction fragments associated with each locus (data from *Astrin* 1978; *Astrin* et al. 1980b). The biochemical and genetic properties of the loci are discussed in detail below.

## 2.1 *ev* 1

*ev* 1 has been reported to be present in 506 out of 508 white leghorn embryos examined (*Tereba* and *Astrin* 1980). Since this locus is present in gs−chf− embryos (*Astrin* 1978), it apparently does not express detectable viral protein products. The structure and transcriptional products of *ev* 1 are shown in Fig. 1 (data from *Hayward* et al. 1980; *Hughes* et al. 1981a; *Baker* et al. 1981). As can be seen in the figure, the structure has no gross defects. The *gag*, *pol*, and *env* genes appear intact, and the genes are flanked by a terminal repeat of 250 nucleotides (*Skalka* et al. 1979; *Hishinuma* et al. 1981). Transcriptional activity of *ev* 1 is very low, however. An apparently normal 35S (8 kilobase) message containing 5′, *gag*, *pol*, *env*, and 3′ sequences is produced, but in extremely low abundance (0.3–0.6 copies per cell). Likewise, a normal 21S (3 kilobase) message containing 5′, *env*, and 3′ se-


LOCUS
STRUCTURE
TRANSCRIPTS PER CELL
ev 1
U3 U5 gag pol env U3 U5
LTR
B
B
B
B
S
5.2 kb
9.4 kb
S
gag pol env
35 S RNA
0.3–0.6
env
21 S RNA
0.1–0.3
ev 2
LTR gag pol env LTR
gag pol env
35 S RNA
< 0.5
21 S RNA
< 0.5
env
ev 3
LTR gag pol env LTR
gag pol
env
gag pol
31S RNA
50–100
env
21S RNA
30–80
19S RNA
10–30
gag pol
ev 4
gag pol env LTR
LTR
No Transcripts Identified
ev 5
pol env LTR
LTR gag
No Transcripts Identified
ev 6
cp
pol env LTR
LTR gag
pol env
28S RNA
0.5–1.5
env
21 S RNA
25–100

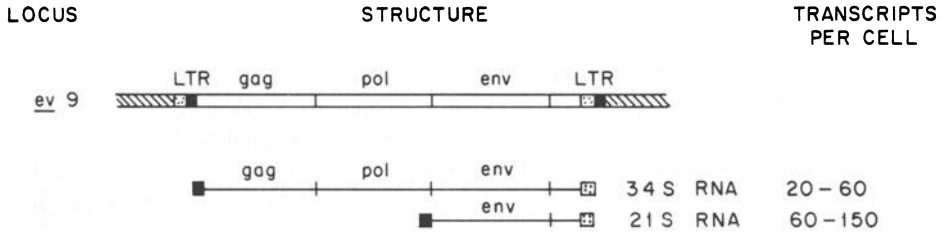



Fig. 1. Structure and transcriptional products for *ev* 1, *ev* 2, *ev* 3, *ev* 4, *ev* 5, *ev* 6, and *ev* 9. ▭, cellular sequences; ⊡■, long terminal repeat (LTR), composed of unique 3′ sequences (⊡ U3) and unique 5′ sequences (■ U5); *cp*, cellular promoter sequence. Restriction sites for *Sst* I (S) and *Bam* H1 (B) are indicated for *ev* 1. The 5.2- and 9.4-kb fragments are the identifying fragments listed for *ev* 1 in Table 2

quences is also produced in small quantities (0.1–0.3 copies per cell). Both messages have been found in nuclei and cytoplasm (*Hayward* et al. 1980; *Baker* et al. 1981). Recent work strongly suggests that methylation is in part responsible for the low level of expression of *ev* 1. An embryo has been identified which spontaneously expresses *ev* 1. Cells of this embryo produce a noninfectious particle containing *gag* polypeptides but lacking reverse transcriptase and envelope antigen (*Conklin* et al. to be published). In addition, other *ev*-1-containing embryos can be induced to express such particles by treatment with 5-azacytidine, an inhibitor of DNA methylation. The induced phenotype is stable for many generations in culture. The appearance of DNase-I-hypersensitive sites in the chromatin domain of *ev* 1 is correlated with its expression after 5-azacytidine treatment (*Groudine* et al. 1981). These results imply that the association of a $gs^-chf^-$ phenotype with *ev* 1 is a result of lack of efficient expression at the RNA level, due at least in part to methylation, as well as to probable structural defects in the *pol* and *env* genes.

The *ev* 1 provirus and flanking cellular sequences have been cloned using recombinant DNA techniques (*Hishinuma* et al. 1981). Using these clones, a comparison has been made between the structure of endogenous viral sequences and the structure of an exogenously acquired provirus. As mentioned above, the basic structure of the two types of provirus is very similar. The following common features are found. First, long terminal repeats (LTRs) flank the viral genes. The repeats flanking *ev* 1 are shorter than any other LTRs studied so far – only 273 base pairs, as compared to 325–1300 base pairs for other viruses. However, the *ev* 1 LTRs contain most of the regulatory sequences known to be present in other viral LTRs, such as a $tRNA^{trp}$ primer binding site, an AT-rich region similar to the Hogness box consensus sequence 32 nucleotides before the 5′ cap site, and polyadenylation signals about 20 nucleotides upstream from a CA dinucleotide. A second common feature is that a 6-base-pair sequence, present only once in the cellular DNA lacking *ev* 1, is repeated once at either end of the provirus in DNA containing *ev* 1. Such a repeat flanks exogenously acquired proviruses, as well as procaryotic and eucaryotic transposable elements. It is by virtue of these common structural features (LTRs and a repeat of flanking host sequences) that transposable elements and endogenous and exogenous retroviruses have been proposed to share a common origin and a common integration mechanism (*Hishinuma* et al. 1981; *Ju* and *Skalka* 1980; *Shimotohno* et al. 1980; *Dhar* et al. 1980; *Majors* and *Varmus* 1981; *Roeder* et al. 1980; *Dunsmuir* et al. 1980; *Levis* et al. 1980).

By the use of in situ hybridization techniques, *ev* 1 has been localized to the long arm of chromosome 1 (*Tereba* et al. 1979; *Tereba* and *Astrin* 1980). This result agrees with the results of studies using chromosome fractionation (*Padgett* et al. 1977). More recently, five other endogenous loci (*ev* 4, *ev* 5, *ev* 6, *ev* 8, and *ev* 13) have been localized to chromosome 1. Common structural features of these loci have led to the proposal that they have been generated by duplication and transposition of *ev* 1 sequences (*Tereba* 1981; *Tereba* and *Astrin* 1982). Cloning and sequencing studies now in progress should provide additional data with which to evaluate this interesting hypothesis.

## 2.2 *ev* 2

Locus *ev* 2 has been found exclusively in RPRL lines $7_2$ and 100 (Astrin 1978; *Tereba* and *Astrin* 1980). Genetic studies have demonstrated that this locus codes for RAV-0 produced by these lines (*Astrin* et al. 1980a). The proviral structure and transcriptional products for *ev* 2 are shown in Fig. 1. The locus appears structurally complete (*Hayward* et al. 1980; *Hughes* et al. 1981a), but is associated with an extremely low transcriptional activity. Between 0.1 and 0.5 copies per cell of apparently normal 35S and 21S messages are produced (*Hayward* et al. 1980). This leads to the production of extremely small quantities of RAV-0 virus. In the cultures of line $7_2$ embryos, the cells of which lack receptors for subgroup E virus and therefore cannot be infected by the RAV-0 they produce, no additional virus is obtained. However, in cultures of embryos such as line 100, which by virtue of being susceptible to infection acquire additional RAV-0 proviruses derived from *ev* 2, the quantity of virus produced is $10^3$- to $10^4$-fold greater (*Robinson* 1978; *Crittenden* et al. 1979). Thus, although *ev* 2 is extremely poor in transcriptional activity, the virus it produces forms transcriptionally active proviruses. Several proposals have been made to account for this phenomenon.

Early studies using a transfection assay to analyze infectivity of DNA from high and low producers of RAV-0 indicated that DNA infectivity parallels virus production. It was proposed that *cis*-acting regulatory sequences inhibited transcription of the *ev* 2 locus, but not of other RAV-0 proviruses acquired by virus infection of susceptible cells (*Cooper* and *Temin* 1976). These conclusions were supported by the observation that infectivity of the *ev* 2 locus could be increased by shearing the cellular DNA to approximately the size of the provirus (*Cooper* and *Silverman* 1978). In further experiments, clones of cells containing *ev* 2 as well as exogenously acquired RAV-0 proviruses were investigated for sites of proviral integration, DNA infectivity, and expression of RNA and virus (*Jenkins* and *Cooper* 1980; *Humphries* et al. 1979, 1981). Each exogenously acquired provirus had a different site of integration; RNA expression, virus production, and DNA infectivity varied 30- to 100-fold when different clones were compared. It was concluded that differences in flanking cellular sequences and/or in modifications such as methylation of proviral DNAs were responsible for the observed differences in expression. The methylation hypothesis is supported by recent experiments in which the *ev* 2 locus has been shown to be activated by 5-azacytidine to produce high levels of RAV-0 virus (*Eisenman* et al. to be published). The activation was shown to correlate with decreased methylation of the locus.

In situ hybridization to metaphase chromosomes has indicated that *ev* 2 is located near the middle of the long arm of chromosome 2 (*Tereba* et al. 1981). Thus far, *ev* 2 is the only endogenous viral locus localized to chromosome 2.

The genome of RAV-0, the gene product of the *ev* 2 locus, has been compared with the genomes of Rous sarcoma virus (RSV), the avian leukosis viruses (ALV), and other exogenous retroviruses of the avian system. The major techniques used in these comparisons have been oligonucleotide maps of the RNA genomes and restriction enzyme maps and sequencing studies of cloned proviral DNA. Of course, the acute transforming viruses such as RSV contain a gene which codes for a transforming protein. These transforming genes are not present in the genomes of the leukosis viruses, such as RAV-2, or of RAV-0 and other endogenous viruses. Comparison of the *gag* and *pol* regions reveals a strong homology for all the viruses (*Shank* et al. 1981). However, in the *env* region and the U3 region of the LTR, considerable divergence occurs between the endogenous and exogenous viruses (*Skalka* et al. 1979; *Tsichlis* and *Coffin* 1980; *Hishinuma* et al. 1981). The envelope glycoproteins of the exogenous viruses have been classified as A, B, C, or D by genetic and biochemical assays. RAV-0 and all other endogenous chicken viruses isolated to date have a subgroup E envelope (for review see *Vogt* 1969). Oligonucleotide and restriction enzyme maps show a corresponding divergence in the envelope regions of the genome. A second difference occurs in the region of the viral LTR, termed U3. This region corresponds to several hundred nucleotides at the 3′ end of the viral RNA and has also been called the C, or constant region. Whereas the U3 or C regions of all the exogenous retroviruses show a close homology, the C region of RAV-0 and other endogenous viruses is distinctly different from that of the exogenous viruses (*Tsichlis* and *Coffin* 1980). The sequence difference manifests itself biologically in two significant ways. First, the difference in C has been observed to be the major determinant for a difference in growth rate between RAV-0, which grows comparatively slowly, and the exogenous viruses which show a more rapid growth (*Tsichlis* and *Coffin* 1980). Second, the difference in C region has been implicated in the failure of the endogenous viruses to cause disease, whereas the exogenous viruses cause a wide variety of acute as well as long latent period neoplasms (*Crittenden* et al. 1980; *Robinson* et al. 1980). Acute disease is caused by the presence of transforming genes such as *myc*, *src*, *erb*, and *myb*, genes which are not present in the leukosis viruses or the endogenous viruses. However, the long latent period neoplasms, such as bursal lymphoma, are readily induced by the leukosis viruses and also by transformation-defective sarcoma viruses, but are never seen in connection with infection by RAV-0 or any other endogenous virus. This striking difference has been shown to be correlated with the difference in the C regions of the viruses (*Crittenden* et al. 1980; *Robinson* et al. 1980). Since the U3 region encodes a promoter for transcription by RNA polymerase II, it is possible to explain both the inefficient growth of RAV-0 and the lack of disease-producing capability by postulating that the RAV-0 promoter is markedly less efficient than that of the exogenous viruses. This hypothesis is made very plausible by the finding that diseases such as bursal lymphoma, which are associated with infection with leukosis viruses, are a result of activation of a cellular oncogene by insertion of the viral promoter sequences (*Neel* et al. 1981; *Payne* et al. 1981; *Hayward* et al. 1981). If the RAV-0 promoter were inefficient, it might lack the capacity to activate cellular oncogene expression to a level suitable for tumor formation.

## 2.3 *ev* 3

*ev* 3 codes for the proteins characteristic of the $gs^+chf^+$ phenotype (expressing *gag* and *env*). *ev* 3 was originally identified as being present in more than 60 $gs^+chf^+$ birds from

four different flocks (*Astrin* 1978). Genetic experiments using RPRL line $6_3$ and Kimber line K16, both of which are homozygous for *ev* 3, were used to analyze segregation of *ev* 3 and of the gs$^+$chf$^+$ phenotype (*Astrin* et al. 1979b; *Astrin* and *Robinson* 1979). $F_1$ progeny of a cross between a line $6_3$ or K16 parent and a gs$^-$chf$^-$ parent were backcrossed to the gs$^-$chf$^-$ parent. Seventy-nine progeny of the backcross matings were analyzed, and an exact correlation between the segregation of *ev* 3 and of the gs$^+$chf$^+$ phenotype was observed.

The proviral structure and transcriptional products of *ev* 3 are shown in Fig. 1. The *ev* 3 provirus is defective, lacking part of the *gag* region, as well as a portion of the *pol* region (*Hayward* et al. 1980; *Hughes* et al. 1981a). A 31S (6.5 kilobase) transcript is produced at levels of 50–150 copies per cell and is found in both nuclei and cytoplasm (*Hayward* et al. 1980; *Baker* et al. 1981). This transcript contains an internal deletion in the *gag-pol* region. This defect correlates with the fact that *ev* 3 cells do not contain the normal 180 000-dalton *gag-pol* precursor protein, but instead contain a 120 000-dalton polyprotein which lacks the *gag* determinants specific for P15, as well as some polymerase determinants. The 120 000-dalton polyprotein is not cleaved to yield functional P27, P19, and P12 gag components or functional reverse transcriptase (*Eisenman* et al. 1978). A second transcript of 3 kilobases (21S) is also found in *ev*-3-containing cells. This transcript appears identical to a normal *env* message, and is present at a level of 30–80 copies per cell in both nuclei and cytoplasm. A 19S (2.3 kilobase) transcript containing *gag* sequences is present at 10–30 copies per cell, but is restricted to the nucleus and may represent a residual product of processing of the *env* message (*Hayward* et al. 1980; *Baker* et al. 1981).

Work on the chromatin structure of the *ev* 3 locus (*Groudine* et al. 1981) has demonstrated that the *ev* 3 sequences are under-methylated as compared to *ev* 1 sequences. In addition, the locus is preferentially sensitive to DNAse 1 digestion, and contains nuclease-hypersensitive sites in each of its two LTRs. These features correlate nicely with the transcriptional activity of *ev* 3.

In situ hybridization experiments using metaphase chromosomes from line $6_3$ have indicated that *ev* 3 is located on a microchromosome (*Tereba* 1981).

## 2.4 *ev* 4, *ev* 5, and *ev* 8

Each of these three loci has been found in gs$^-$chf$^-$ cells (*Astrin* 1978), an indication that they do not express detectable viral protein products. Structures for the proviruses are shown in Fig. 1. No transcriptional products have been detected for *ev* 4 or *ev* 5 (*Hayward* et al. 1980). Lack of transcriptional activity of *ev* 4 and *ev* 5 is not surprising, since each locus has a deletion of 5′ sequences and lacks the 5′ LTR, the putative promoter for viral transcription (*Hayward* et al. 1980; *Hughes* et al. 1981a; *Baker* et al. 1981). No information on transcriptional activity or proviral defectiveness has been reported for *ev* 8.

*ev* 4 and *ev* 5 have been shown to be genetically linked in mating experiments where recombination between the two loci could be analyzed (*Astrin* et al. 1979b). Results of in situ hybridization experiments confirm this linkage and indicate that the two loci lie in proximity on the long arm of chromosome 1 (*Tereba* 1981; *Tereba* and *Astrin* 1982). As has been alluded to above, *ev* 4, *ev* 5, *ev* 6, *ev* 8, and *ev* 13 have all been localized to chromosome 1 by in situ hybridization. This finding has served as a basis for the hypothesis that *ev* 4, *ev* 5, *ev* 6 and *ev* 8 have been generated by multiple duplications of *ev* 1. Structural

similarities between retroviral proviruses, *ev* 1 included, and transposable elements lend this model a certain credence.

## 2.5 *ev* 6

*ev* 6 was originally identified in 14 birds of the gs$^-$chf$^+$ phenotype (expressing *env* but not *gag*) (*Astrin* 1978). It is one of two loci which code for this phenotype, the other being *ev* 9. Genetic experiments using Kimber line K18, which expresses the phenotype, have been used to analyze segregation of chf expression and *ev* 6. An examination of more than 30 progeny of backcross matings which were segregating for the expression of chf revealed an exact correlation between the segregation of *ev* 6 and segregation of the gs$^-$chf$^+$ phenotype (*Astrin* et al. 1980b).

The proviral structure and transcriptional products of *ev* 6 are shown in Fig. 1. The structure of the *ev* 6 provirus, like that of *ev* 4 and *ev* 5, is defective. *ev* 6 lacks both the 5′ LTR and *gag* sequences (*Hayward* et al. 1980; *Hughes* et al. 1981a). Surprisingly, although this locus lacks the viral promoter sequences present in the 5′ LTR, transcriptional products are observed. A low level (0.5–1.5 copies per cell) of a 28S (5.3 kilobase) RNA-containing *pol*, *env*, and U3 sequence is observed in the nucleus, and a moderately high level (25–100 copies per cell) of a 21S (RNA-containing *env* and U3 sequence is observed in both nucleus and cytoplasm (*Hayward* et al. 1980; *Baker* et al. 1981). The 21S RNA most likely serves as message for the production of *env* protein characteristic of *ev*-6-containing cells. Neither of these RNAs contains viral U5 sequences (*Hayward* et al. 1980; *Baker* et al. 1981). This finding is further evidence that the RNAs are not transcribed from a viral promoter. It has been proposed that the *ev*-6-encoded 21S *env* message is initiated within a cellular promoter located adjacent to the left end of the *ev* 6 provirus. Initiation within the cellular promoter and transcription of adjacent cellular sequences followed by transcription of viral sequences might provide a message which contains a cellular leader sequence covalently linked to viral information (*Hayward* et al. 1980; *Baker* et al. 1981). Further experimentation will be required to confirm this attractive hypothesis.

As mentioned above, *ev* 6 has been localized to chromosome 1 of the chicken by in situ hybridization (*Tereba* 1981; *Tereba* and *Astrin* 1982).

## 2.6 *ev* 7

*ev* 7 was originally identified in 15 embryos of RPRL line $15_b$ (*Astrin* 1978). This line of chickens has an interesting phenotype. Cells grown in the presence of bromodeoxyuridine are induced to express a noninfectious avian leukosis virus. The virus is detected as particles containing reverse transcriptase activity (*Robinson* et al. 1976; *Robinson* 1978). *ev* 7 segregates with this phenotype in backcross matings and apparently codes for the particles produced by line $15_b$ cells (*Robinson* et al. 1979b).

The *ev* 7 provirus does not appear to contain any gross deletions (*Baker* et al. 1981). No analysis of transcriptional products has been reported.

The genetic information of *ev* 7 has been reported to undergo recombination with *ev* 1 genetic information to produce infectious subgroup E virus. These viruses were shown to have P27 and P19 components which were characteristic of RAV-0 (*Robinson* et al.

1979a). This finding supports the conclusion that the viruses are formed by recombination of endogenous viral genetic information. The parents of the recombinant viruses were postulated to be the *ev*-7-encoded virus particles and *ev*-1-encoded RNA.

Genetic experiments utilizing line $15_b$ in crosses with lines $7_2$ and $6_3$ have shown that *ev* 7 segregates with the male (z) chromosome (*Smith* and *Crittenden* 1981). Independent experiments using in situ hybridization methods have also localized *ev* 7 to the z chromosome (*Tereba* et al. 1981). Thus far, *ev* 7 is the only endogenous provirus to be localized to the z chromosome.

## 2.7 *ev* 9

*ev* 9 is one of two loci, the other being *ev* 6, that code for envelope protein produced in $gs^-chf^+$ cells. *ev* 9 was first identified in several embryos of this phenotype and later observed to segregate with the phenotype in backcross matings (*Astrin* 1978; *Astrin* et al. 1980b).

The proviral structure and transcriptional products of *ev* 9 are shown in Fig. 1. The provirus has no apparent deletions; however, a 34S transcript is observed (*Hayward* et al. 1980). This size is slightly smaller than the transcript of an intact provirus which is 35S. The 34S transcript contains U5, *gag*, *pol*, *env*, and U3 sequences; is present in 20–60 copies per cell; but is confined to the nuleus (*Baker* et al. 1981). A second transcript of 3 kilobases (21S) contains U5, *env*, and U3 sequences; is present in 60–150 copies per cell; and is found in both nuclei and cytoplasm. This transcript is most likely the message for the production of envelope protein characteristic of *ev*-9-containing cells.

## 2.8 *ev* 10, *ev* 11, and *ev* 12

These three loci each code for a distinct subgroup E virus produced by a particular line of inbred white leghorn chickens (*Astrin* et al. 1980b). *ev* 10 is present in lines C, $15I_4$, and $15I_5$ (*Tereba* and *Astrin* 1980). It has been shown to segregate with the ability to produce infectious subgroup E virus ($V^+$ phenotype) in an analysis of 48 progeny of a backcross mating of line $15I_4$ chickens (*Crittenden* and *Astrin* 1981). *ev* 11 is present in line $15I_4$ and $15I_5$ (*Tereba* and *Astrin* 1980), and has similarly been shown to segregate with the $V^+$ phenotype in backcross matings (Crittenden and Astrin unpublished results). *ev* 12 is present in RPRL line $15_1$ (*Tereba* and *Astrin* 1980), and has been shown to segregate with the $V^+$ phenotype in matings of line $15_1$ chickens (*Smith* and *Crittenden* 1981).

*ev* 10, *ev* 11, and *ev* 12, like *ev* 2, produce only very small amounts of virus spontaneously. However, if the producer cells have receptors for subgroup E virus, additional proviruses are accumulated through infection, and much larger quantities of virus are produced. It is likely that the control mechanisms regulating expression of *ev* 2 (QV) are also operating in the cases of *ev* 10, *ev* 11, and *ev* 12.

## 2.9 *ev* 13, *ev* 14, *ev* 15, and *ev* 16

*ev* 13 is a locus which was identified in $gs^-chf^-$ cells, and shown by in situ hybridization to be present on chromosome 1 in a unique location (*Tereba* 1981; *Tereba* and *Astrin* 1982). No information other than its chromosomal location is available.

*ev* 14 is a $V^+$ locus identified in an embryo from Heisdorf and Nelson Farms. The pattern of *ev*-14-specific fragments produced after restriction enzyme digestion of embryo DNA is distinguishable from the pattern produced by any of the other 15 loci. *ev* 14 has been localized near the middle of chromosome 3 (*Tereba* et al. 1981).

*ev* 15 and *ev* 16 are elements identified in Heisdorf and Nelson and Kimber flocks (*Hughes* et al. 1981a; Astrin, unpublished results). They appear to consist only of a single copy of an LTR of the type associated with the endogenous proviruses. These elements could have been produced from the intact proviral elements by homologous recombination within the viral LTR.

## 3 Endogenous Viral Genes in Varieties of Chickens Other than the White Leghorn

Several studies have been conducted on *ev* loci in breeds of chicken other than the white leghorn (*Hughes* et al. 1979, 1981b; *Astrin* et al. 1980b). These studies have indicated that proviruses similar to those found in white leghorn are present in other chickens, that the number of proviruses per chicken is highly variable, and that there are *ev* loci present in other breeds that do not correspond to any of those identified in white leghorns.

In one study, DNAs from erythrocytes of 100 chickens representing 68 varieties were cleaved with endonuclease *Sac*I or endonuclease *Bam*H1, and the *ev* loci characterized by Southern blots (*Astrin* et al. 1980; Astrin and Wyban, unpublished results). This study revealed an unexpected multiplicity and diversity of fragments produced from the DNAs of individual chickens. Although there were fragments in the digestions which comigrated with bands representing already characterized *ev* loci, there were also many bands that did not comigrate with known loci. From these data it was estimated that there were probably more than 20 new loci present in the 100 chickens. Although each of the

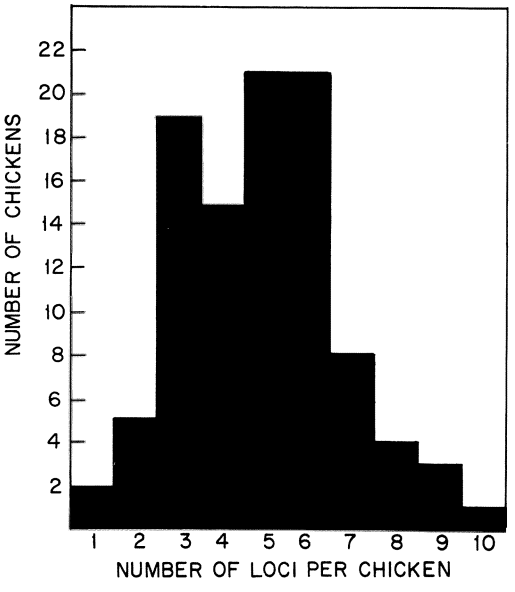



Fig. 2. Distribution of *ev* loci among 100 chickens

100 chickens had at least one *ev* locus, no single locus was common to all individuals. In fact, fewer than 20% of the chickens appeared to contain *ev* 1, the locus ubiquitous in white leghorns. A distribution of the *ev* loci is shown in Fig. 2. The number of loci per chicken ranged from one to ten with the average being 4.9.

An analysis of DNA of the red jungle fowl (*Gallas bankiva*) was also conducted to determine whether this putative ancestor of the domestic chicken had endogenous viral genes. In an analysis of 12 individuals from a single population, at least 4 *ev* loci appeared to be segregating in the population (Astrin and Wyban, unpublished results). These data, and the fact that virtually every chicken has at least one *ev* locus, indicate that the chicken genome probably contained viral sequences prior to domestication. A study conducted on other species of jungle fowl indicated that only the red jungle fowl has ALV-related endogenous viral genes (*Frisby* et al. 1979). These observations lead to the conclusion that viral genes first entered the germline after speciation of the jungle fowl and prior to domestication.

## 4 Origins of the Endogenous Viral Genes

For cellular genes multiple copies of a gene are generally considered to be the result of duplication of the ancestral gene via some unequal exchange between chromosomal regions and/or polyploidization. The findings that many of the *ev* loci in the white leghorn are very closely related in sequence to *ev* 1, and that six of these loci, including *ev* 1, are located on chromosome 1, suggest the process of gene duplication might be operating for these genes. In addition, it is extremely likely that at least some of the *ev* loci are the result of independent infections of the germline of chickens. This hypothesis is supported by the large variation in sites of residence, observed even among individuals of the same breed, and by the finding that organization of the *ev* loci and flanking sequences is identical to that observed for proviruses acquired by exogenous infection.

Although the endogenous viral genes of the chicken are structurally very similar to the genomes of exogenous viruses, two genetic differences distinguish the endogenous viruses. One of these differences resides in the gene for the viral envelope protein. The antigenic properties of the envelope protein and host range of the endogenous virus define it as a subgroup E virus, distinguishable from virus of other subgroups by its failure to interfere with their infection (for review see *Vogt* 1969). This structural and functional distinction between endogenous and exogenous chicken viruses seems at first to argue against the multiple germline infection hypothesis because this hypothesis predicts that exogenous subgroups would be found as endogenous virus. However, if the non-E subgroups have appeared only recently in evolutionary terms, they may not as yet have had time to become major residents of the germline. The second genetic difference between endogenous and exogenous retroviral genes of the chicken lies in the U3 or C region. This region of several hundred nucleotides is thought to encode the promoter for viral RNA synthesis. While the C regions of exogenous viruses are all highly homologous, the endogenous viral loci characterized thus far contain very different C regions (*Tsichlis* and *Coffin* 1980). The fact that the exogenous type C region appears to be absent or infrequent among endogenous viral loci may be a consequence of the association of the exogenous C region with oncogenicity in the nonacute leukosis viruses (*Crittenden* et al. 1980; *Robin-*

*son* et al. 1980). This association probably results in selection against acquisition of an endogenous viral genome containing the exogenous C region.

A somewhat novel mechanism for generating at least part of the multiplicity of *ev* loci documented above is suggested by the fact that the endogenous viral genome codes for an RNA-dependent DNA polymerase (reverse transcriptase). This enzyme can utilize transcribed material (RNA) as a template to synthesize genetic information (DNA). An ancestral *ev* locus may have been transcribed into RNA from which a complementary DNA sequence was produced by reverse transcription. This duplicated sequence may then have been inserted back into the chicken genome in a new location. This mechanism for generating multiple *ev* loci from a single locus can account for the genetic similarity in envelope and C region observed among the loci. It can also account for the deletions observed in many of the *ev* loci, deletions that could easily have arisen during reverse transcription or integration.

A scheme for the origin of *ev* sequences is presented in Fig. 3. This scheme takes into account the structural similarities between endogenous viruses and transposable elements by suggesting that the viruses have evolved from transposons. In this scheme it is also suggested that the subgroup E endogenous viruses are the ancestors of the current population of exogenous viruses. This proposal can be rationalized as follows: If an endogenous provirus evolved to the state of producing a viral particle (with a subgroup E envelope) these particles could then infect the germlines of other individuals. Eventually, all individuals would have subgroup E viruses in their germline (the current situation). The production of envelope glycoprotein by endogenous proviruses has been shown to interfere with exogenous infections by virus of the same envelope subgroup by reducing

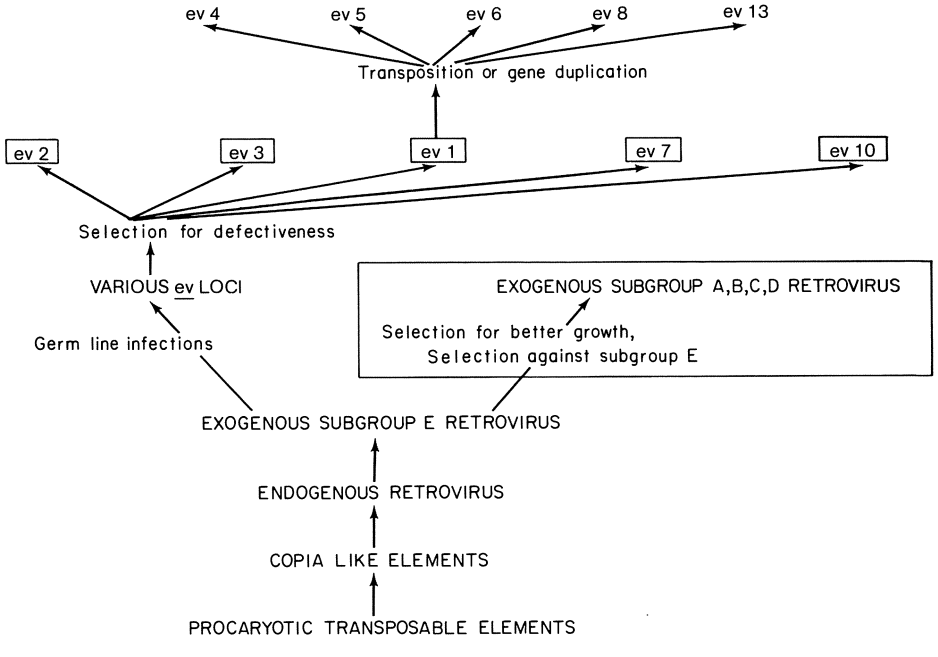



Fig. 3. Postulated evolution of *ev* loci

the availability of cellular receptors. Thus, as subgroup E endogenous proviruses accumulated, suitable hosts for growth of exogenous subgroup E viruses would have become fewer and fewer. Selective pressure for changes in envelope would have come into effect; hence, the appearance of subgroups A, B, C, and D, all of which may have evolved too recently to have entered the germline. Of course, the difference between the U3 region of exogenous and endogenous viruses must also be explained. It may be postulated that virus with an endogenous type U3 region acquired a different U3 region (exogenous type) by recombination with a host sequence. If the newly acquired sequences served as a more efficient promoter, the new virus would have been at a selective advantage. Entry into the germline of such a variant would have been selected against due to the association of an exogenous type U3 region with oncogenicity. The above considerations predict that the only reservoir remaining today for the ancestral virus (a subgroup E virus with an endogenous type of U3 region) would be as an endogenous provirus. Its survival as an exogenous virus would be severely restricted due to its relatively poor growth (U3 region) and lack of suitable hosts (those with active subgroup E receptors).

A third feature of the scheme shown in Fig. 3 is that it postulates a selective pressure against the acquisition of endogenous proviruses that spontaneously produce large amounts of infectious virus. To date, no endogenous provirus constitutively producing high levels of particles has been described. All the $V^+$ loci are transcriptionally repressed, producing very few particles. The reason for such a selection is not clear, since production of high titers of RAV-0 over long periods does not seem to be deleterious to the chicken.

Finally, the scheme in Fig. 3 provides two methods for the acquisition of new *ev* loci: germline infection and gene duplication. As discussed above, gene duplication mediated by a reverse transcription of a proviral transcript is an attractive hypothesis. In addition, a gene duplication mechanism may partially explain the accumulation of endogenous viral loci on chromosome 1.

## 5 Endogenous Viral Genes and the Biology of the Chicken

Are endogenous viral genes essential? The ubiquitous presence of endogenous viral genes in the various breeds of chicken, their presence throughout vertebrate species, and the association of their expression with development of the hematopoietic system in the mouse have led to the proposal that these genes are involved in ontogeny. However, the production by selective breeding of fertile chickens lacking ALV-related endogenous viral genes has been reported (*Astrin* et al. 1979a; *Astrin* et al. 1980b). These chickens were the progeny of matings of white leghorns of defined genotype. Two such matings had been identified, each of which gave rise to a small fraction of embryonic progeny in which no endogenous proviruses could be detected. When progeny from these two matings were hatched and analyzed for the presence of viral sequences, a single rooster was identified which lacked such genes (*Astrin* et al. 1979a). This rooster has been bred to its female sibling (known to be heterozygous for *ev* 1, *ev* 2, and/or *ev* 5), and additional birds lacking endogenous viral genes have been produced. A flock of over 100 such birds is now in existence (Crittenden and Astrin, unpublished results). The birds are healthy, normal, and fertile. Thus, endogenous viral genes are not essential in the chicken. How then can one account for their ubiquitous presence?

Three factors probably account for the fact that chickens lacking endogenous viral genes had not been previously identified. First, there was a large multiplicity of different endogenous viral loci in the total population of chickens; second, individual chickens usually possess several loci (an average of five); third, laboratory flocks of white leghorns appear to be uniformly heterozygous for locus *ev* 1. Given these factors, the production, by random matings, of a chicken lacking endogenous viral genes would be a rare event. If such genes are indeed nonessential, why are they so prevalent in the population? It seems likely that the presence of these genes offers a selective advantage to the chicken. Indeed, in recent experiments chickens with defined complements of *ev* loci were compared in terms of their response to infection by the leukosis virus RAV-1. The results of these experiments (discussed in detail in the last section of this article) lend considerable credence to the proposal that the presence of endogenous proviruses is beneficial to the chicken.

## 5.1 Endogenous Viral Genes as Inhibitors of Infection

It has long been known that cells infected by virus of a given subgroup (envelope specificity) are refractory to infection by other viruses of the same subgroup. It is presumed that penetration of the superinfecting virus is blocked or slowed by the binding to receptors of internally produced envelope glycoprotein. This interference phenomenon has also been documented for certain *ev* loci, namely *ev* 3, *ev* 6, and *ev* 9, all of which produce high levels of subgroup E envelope glycoprotein (chf$^+$ phenotype). Cells containing *ev* 3, *ev* 6, and/or *ev* 9 have been reported to have a much reduced susceptibility to infection with exogenous subgroup E virus. The mechanism of this phenomenon appears to be binding of internal envelope glycoprotein or of envelope glycoprotein shed by infected cells to the cellular receptors for subgroup E virus (*Robinson* et al. 1981).

## 5.2 Endogenous Viral Genes and the Response to Infection by Avian Leukosis Virus

Very recently, the effect of the presence of *ev* 3 (expressing gs and chf) on the response of chickens to infection with the avian leukosis virus, RAV-1, has been examined. Normally, this virus produces bursal lymphomas 4–12 months after inoculation. However, in birds lacking *ev* 3 and, therefore, not expressing gs or chf, an interesting nonneoplastic syndrome is seen 6–12 weeks after infection; the disease is not observed in birds that have *ev* 3. The syndrome is lethal and characterized by severe atrophy of lymphoid organs, an inflammatory response in the liver, and a lower immune response to particulate antigens (*Crittenden* et al. 1982).

What is the cause of the disease and why does the presence of *ev* 3 protect against it? Several lines of evidence indicate an immune response to infected organs may be involved. The timing and extent of the humoral immune response to the infecting virus is quite different in *ev* 3$^+$ and *ev* 3$^-$ birds. *Ev* 3$^+$ birds respond later, have lower antibody titers, and have a more prolonged viremia than do *ev* 3$^-$ birds (Crittenden et al. 1982). Thus, birds expressing *ev* 3 show evidence of immunologic tolerance to RAV-1. This finding is not surprising. Since *ev* 3$^+$ birds express chf and are, therefore, tolerant to sub-

group E envelope antigens, they would also be tolerant to those determinants shared between the subgroup A envelope of RAV-1 and subgroup E envelope.

What is the connection between the lack of immune tolerance and the disease? One possibility is that the disease is caused by a severe and prolonged cellular immune response directed against subgroup A antigen on the surface of RAV-1-infected cells. Such a response might be extremely destructive to infected organs and could result in the symptoms characteristic of the lethal syndrome. The presence of *ev* loci such as *ev* 3, *ev* 6, or *ev* 9, which express endogenous chf (envelope glycoprotein), would protect against the disease. We thus have the basis for a selective mechanism operating in favor of birds expressing endogenous chf. Indeed most flocks not selectively bred for particular *ev* loci consist of a majority of birds that are $chf^+$. In addition, biological assays on jungle fowl, the ancestor of the domestic chicken, indicate that they too are mainly $chf^+$ (*Weiss* and *Biggs* 1972). Thus, despite the fact that endogenous viral genes appear to perform no essential functions in the chicken, their presence appears to be extremely relevant to survival under normal circumstances.

## References

Astrin SM (1978) Endogenous viral genes of the white leghorn chicken: common site of residence and sites associated with specific phenotypes of viral gene expression. Proc Natl Acad Sci USA 75:5941–5945

Astrin SM, Robinson HL (1979) Gs, an allele of chickens for endogenous avian leukosis viral antigens, segregates with *ev* 3, a genetic locus that contains structural genes for virus. J Virol 31:420–425

Astrin SM, Buss EG, Hayward WS (1979a) Endogenous viral genes are nonessential in the chicken. Nature 281:339–341

Astrin SM, Crittenden LB, Buss EG (1979b) *ev* 3, a structural gene locus for endogenous virus, segregates with the $gs^+chf^+$ phenotype in matings of line $6_3$ chickens. Virology 99:1–9

Astrin SM, Crittenden LB, Buss EG (1980a) *ev* 2, a genetic locus containing structural genes for endogenous virus, codes for Rous-associated virus type 0 produced by line $7_2$ chickens. J Virol 33:250–255

Astrin SM, Robinson HL, Crittenden LB, Buss EG, Wyban J, Hayward WS (1980b) Ten genetic loci in the chicken that contain structural genes for endogenous avian leukosis viruses. Cold Spring Harbor Symp Quant Biol 19:1105–1109

Baker B, Robinson H, Varmus HE, Bishop JM (1981) Analysis of endogenous avian retrovirus DNA and RNA: viral and cellular determinants of retrovirus gene expression. Virology 114:8–22

Baluda MA (1972) Widespread presence in chickens of DNA complementary to the RNA genome of avian leukosis viruses. Proc Natl Acad Sci USA 69:576–580

Conklin KF, Coffin JM, Robinson HL, Groudine M, Eisenman R (to be published) Role of methylation in the induced and spontaneous expression of the avian endogenous virus *EV*-1: DNA structure and gene products. Mol Cell Biol

Cooper GM, Silverman L (1978) Linkage of the endogenous avian leukosis virus genome of virus-producing chicken cells to inhibitory cellular DNA sequences. Cell 15:573–577

Cooper GM, Temin HM (1976) Lack of infectivity of the endogenous avian leukosis virus-related genes in the DNA of uninfected chicken cells. J Virol 17:422–430

Crittenden LB, Astrin SM (1981) Independent segregation of *ev* 2 and *ev* 10, genetic loci for spontaneous production of endogenous avian retroviruses. Virology 110:120–127

Crittenden LB, Smith EJ, Weiss RA, Sarma PS (1974) Host gene control of endogenous avian leukosis virus production. Virology 57:128–138

Crittenden LB, Motta JV, Smith EJ (1977) Genetic control of RAV-0 production in chickens. Virology 76:90–97

Crittenden LB, Smith EJ, Gulvas FA, Robinson HL (1979) Exogenous virus expression in chicken lines maintained at the Regional Poultry Research Laboratory. Virology 95:434–444

Crittenden LB, Hayward WS, Hanafusa H, Fadly AM (1980) Induction of neoplasms by subgroup E recombinants of exogenous and endogenous avian retroviruses (Rous-associated virus type 0). J Virol 33:915–919

Crittenden LB, Fadley AM, Smith EJ (1982) Effect of endogenous leukosis virus genes on response to infection with avian leukosis and reticuloendotheliosis viruses. Avian Diseases 26:279–294

Dhar R, McClements WL, Enquist LW, Vande Woude GF (1980) Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. Proc Natl Acad Sci USA 77:3937–3941

Dougherty RM, DiStefano HS (1966) Lack of relationship between infection with avian leukosis virus and the presence of COFAL antigen in chick embryos. Virology 29:586–595

Dunsmuir P, Brorein WJ Jr, Simon MA, Rubin GM (1980) Insertion of the Drosophila transposable element *copia* generates a 5 base pair duplication. Cell 21:575–579

Eisenman R, Vogt VM (1978) The biosynthesis of oncovirus proteins. Biochem Biophys Acta 473:187–239

Eisenman R, Shaikh R, Mason WS (1978) Identification of an avian oncovirus polyprotein in uninfected chick cells. Cell 14:89–104

Eisenman R, Heater P, Robinson H, Conklin K, Coffin J, Goubin G, Cooper G, Groudine M (to be published) In: Scolnick EM, Levine AJ, Fox CF (eds) DNA methylation and the control of endogenous retrovirus gene expression. ICN-UCLA Symposia on Molecular and Cellular Biology, Tumor Viruses and Differentiation

Frisby DT, Weiss RA, Roussel M, Stehelin D (1979) The distribution of endogenous chicken retrovirus sequences in the DNA of galliform birds does not coincide with avian phylogenetic relationships. Cell 17:623–634

Gross L (1958a) Viral etiology of "spontaneous" mouse leukemia: a review. Cancer Res 18:371–381

Gross L (1958b) Attempt to recover a filterable agent from X-ray-induced leukemia. Acta Hematol 19:353–361

Groudine M, Eisenman R, Weintraub H (1981) Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. Nature 292:311–317

Hanafusa H, Miyamoto T, Hanafusa T (1970) A cell-associated factor essential for formation of an infectious form of Rous sarcoma virus. Proc Natl Acad Sci USA 66:314–321

Hanafusa T, Hanafusa H, Miyamoto T (1970) Recovery of a new virus from apparently normal chick cells by infection with avian tumor viruses. Proc Natl Acad Sci USA 67:1797–1803

Hanafusa T, Hanafusa H, Miyamoto T, Fleissner E (1972) Existence and expression of tumor virus genes in chick embryo cells. Virology 47:475–482

Hayward WS, Hanafusa H (1973) Detection of avian tumor virus RNA in uninfected chicken cells. J Virol 11:157–167

Hayward WS, Braverman SB, Astrin SM (1980) Transcriptional products and DNA structure of endogenous avian proviruses. Cold Spring Harbor Symp Quant Biol 44:1111–1122

Hayward WS, Neel BG, Astrin SM (1981) Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290:475–480

Hishinuma F, DeBona PJ, Astrin SM, Skalka AM (1981) Nucleotide sequence of acceptor site and termini of integrated avian endogenous provirus *ev* 1: integration creates a 6 bp repeat of host DNA. Cell 23:155–164

Hughes SH, Farhang P, Spector D, Schimke RT, Robinson HL, Payne GS, Bishop JM, Varmus HE (1979) Heterogeneity of genetic loci in chickens: analysis of endogenous viral and nonviral genes by cleavage of DNA with restriction endonucleases. Cell 18:347–359

Hughes SH, Bishop JM, Varmus HE (1981a) Organization of the endogenous proviruses of chickens: implications for origin and expression. Virology 108:189–207

Hughes SH, Vogt PK, Bishop JM, Varmus HE (1981b) Endogenous proviruses of random-bred chickens and ring-necked pheasants: analysis with restriction endonucleases. Virology 108:222–229

Humphries EH, Glover C, Weiss RA, Arrand JR (1979) Differences between the endogenous and exogenous DNA sequences of Rous-associated virus-0. Cell 18:803–815

Humphries EH, Allen R, Glover C (1981) Clonal analysis of the integration and expression of endogenous avian retroviral DNA acquired by exogenous viral infection. J Virol 39:584–596

Jenkins NA, Cooper GM (1980) Integration, expression and infectivity of exogenously acquired proviruses of Rous-associated virus-0. J Virol 36:684–691

Ju G, Skalka AM (1980) Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. Cell 22:379–386

Levis R, Dunsmuir P, Rubin GM (1980) Terminal repeats of the Drosophila transposable element *copia*: nucleotide sequence and genomic organization. Cell 21:581–588

Lieberman M, Kaplan HS (1959) Leukemogenic activity of filtrates from radiation-induced lymphoid tumors in mice. Science 130:387–388

Majors JE, Varmus HE (1981) Nucleotide sequences at host-viral junctions for mouse mammary tumor virus. Nature 289:253–258

Neel BG, Hayward WS, Robinson HL, Fang J, Astrin SM (1981) Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell 23:323–344

Neiman PE (1973) Measurement of endogenous leukosis virus nucleotide sequences in the DNA of normal avian embryos by RNA-DNA hybridization. Virology 53:196–204

Padgett TG, Stubblefield E, Varmus HE (1977) Chicken macrochromosomes contain an endogenous provirus and microchromosomes contain sequences related to the transforming gene of ASV. Cell 10:649–657

Payne GS, Courtneidge SA, Crittenden LB, Fadly AM, Bishop JM, Varmus HE (1981) Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell 23:311–322

Payne LN, Chubb RC (1968) Studies on the natural and genetic control of an antigen in normal chick embryos which reacts in the Cofal Test. J Gen Virol 3:379–391

Robinson HL (1978) Inheritance and expression of chicken genes that are related to avian leukosis sarcoma virus genes. Curr Top Microbiol Immunol 83:1–36

Robinson HL, Swanson CA, Hraska JF, Crittenden LB (1976) The production of unique C type viruses by chicken cells grown in Bromodeoxyuridine. Virology 69:63–74

Robinson HL, Eisenman R, Senior A, Ripley S (1979a) Low frequency production of recombinant subgroup E avian leukosis viruses by uninfected V-$15_B$ chicken cells. Virology 99:21–30

Robinson HL, Astrin SM, Salazar FH (1979b) V-$15_B$, an allele of chickens for the production of a noninfectious avian leukosis virus. Virology 99:10–20

Robinson HL, Pearson MN, DeSimone DW, Tsichlis PN, Coffin JM (1980) Subgroup-E avian-leukosis-virus-associated disease in chickens. Cold Spring Harbor Symp Quant Biol 44:1133–1142

Robinson HL, Astrin SM, Senior AM, Salazar FH (1981) Host susceptibility to endogenous viruses: defective glycoprotein-expressing proviruses interfere with infections. J Virol 40:745–751

Roeder GS, Farabaugh PJ, Chaleff DT, Fink GR (1980) The origins of instability in yeast. Science 209:1375–1380

Rosenthal PW, Robinson HL, Robinson WS, Hanafusa T, Hanafusa H (1971) DNA in uninfected and virus-infected cells complementary to avian tumor virus RNA. Proc Natl Acad Sci USA 68:2336–2340

Shank PR, Hughes SH, Varmus HE (1981) Restriction endonuclease mapping of the DNA of Rous-associated virus 0 reveals extensive homology in structure and sequence with avian sarcoma virus DNA. Virology 108:177–188

Shimotohno K, Mizutani S, Temin HM (1980) Sequence of retrovirus provirus resembles that of bacterial transposable elements. Nature 285:550–554

Skalka A, DeBona P, Hishinuma F, McClements W (1979) Avian endogenous proviral DNA: analysis of integrated *ev* 1 and a related $gs^-chf^-$ provirus purified by molecular cloning. Cold Spring Harbor Symp Quant Biol 44:1097–1104

Smith EJ, Crittenden LB (1981) Segregation of chicken endogeneous viral loci *ev* 7 and *ev* 12 with the expression of infectious subgroup E avian leukosis viruses. Virology 112:370–373

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Tereba A (1981) 5′-Terminal deletions are a common feature of endogenous retrovirus loci located on chromosome 1 of White Leghorn chickens. J Virol 40:920–926

Tereba A, Astrin SM (1980) Chromosomal localization of *ev*-1, a frequently occurring endogenous retrovirus locus in White Leghorn chickens, by in situ hybridization. J Virol 35:888–894

Tereba A, Astrin SM (1982) Localization of five endogenous retrovirus loci on chromosome 1 of White Leghorn chickens by in situ hybridization. J Virol 43:737–740
Tereba A, Lei MMC, Murti KG (1979) Chromosome 1 contains the endogenous RAV-0 retrovirus sequences in chicken cells. Proc Natl Acad Sci USA 76:6486–6490
Tereba A, Crittenden LB, Astrin SM (1981) Chromosomal localization of three endogenous retrovirus loci associated with virus production in White Leghorn chickens. J Virol 39:282–289
Tooze J (1973) The molecular biology of tumor viruses. Cold Spring Harbor Laboratory Ed, Cold Spring Harbor, New York
Tsichlis PN, Coffin JM (1980) Recombination between endogenous and exogeneous avian tumor viruses: role of the c region and other portions of the genome in the control of replication and transformation. J Virol 33:238–249
Varmus HE, Weiss RA, Friis RR, Levinson W, Bishop JM (1972) Detection of avian-tumor-specific nucleotide sequences in avian cell DNAs. Proc Natl Acad Sci USA 69:20–24
Vogt PK (1969) Envelope classification of avian RNA tumor viruses. In: Dutcher RM (ed) Comparative Leukemia Research. Karger, Basel pp 153–167
Vogt PK, Friis RR (1971) An avian leukosis virus related to RSV(0): properties and evidence for helper activity. Virology 43:223–234
Vogt PK, Hu SSF (1977) The genetic structure of RNA tumor viruses. Annu Rev Genet 11:203–208
Wang S, Hayward WS, Hanafusa H (1977) Genetic variation in the RNA transcripts of endogenous virus gene in uninfected chicken cells. J Virol 24:63–73
Weiss RA (1969) The host range of BRYAN strain Rous sarcoma virus synthesized in the absence of helper virus. J Gen Virol 5:511–528
Weiss RA, Biggs PM (1972) Leukosis and Marek's disease viruses of feral red jungle fowl and domestic fowl in Malaya. J Nat Cancer Inst 49:1713–1725
Weiss RA, Payne LN (1971) The heritable nature of the factor in chicken cells which acts as a helper virus for Rous sarcoma virus. Virology 45:508–515
Weiss RA, Friis RR, Katz E, Vogt PK (1971) Induction of avian tumor viruses in normal cells by physical and chemical carcinogens. Virology 46:920–938
Weissbach A, Bolden A, Muller R, Hanafusa H, Hanafusa T (1972) Deoxyribonucleic acid polymerase activities in normal and leukovirus-infected chicken embryo cells. J Virol 10:321–327

# Synthesis, Integration, and Transcription of the Retroviral Provirus

STEPHEN H. HUGHES*



## 1 Introduction

The RNA genome of a retrovirus is copied into DNA by the viral enzyme, RNA-dependent DNA polymerase, and the DNA copy of the retrovirus genome is inserted into one of the host's chromosomes. This DNA copy, called a provirus, is transcribed into RNA by host enzymes, giving rise to both viral genomic RNA and viral messenger RNA.

It is the purpose of this review to explore the structure of the DNA provirus, and to examine the molecular events which create the provirus. Because the replication of the many species of retroviruses is quite similar and because the retrovirus literature is vast, only representative examples will be discussed and only representative references will be given.

The most basic problem in understanding the structure of the unintegrated forms of viral DNA and the integrated provirus is that all of these molecules are present in very

* Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, New York 11724, USA

small amounts, usually only a few copies per cell, even after an acute infection with high-titer virus. It is no surprise that progress in determining retrovirus DNA structure followed closely technical advances in nucleic acid biochemistry. The earliest experiments used solution hybridization which, together with the transfection protocols (*Hill* and *Hillova* 1972a, 1972b), demonstrated the presence both of integrated and unintegrated viral DNA (*Neiman* 1972; *Baluda* 1972; *Varmus* et al. 1973a, 1973b, 1975; *Guntaka* et al. 1976; *Weinberg* 1977), and established that a portion of the unintegrated viral DNA was supercoiled (*Guntaka* et al. 1976; *Weinberg* 1977). Analysis with restriction enzymes and the DNA transfer procedure of *Southern* (1975) permitted the elucidation of the structure of the unintegrated linear and circular viral DNAs (*Shank* et al. 1978a, 1978b; *Hsu* et al. 1978; *Yoshimora* and *Weinberg* 1979; *Keshet* et al. 1979; *Sherr* et al. 1979; *Bergmann* et al. 1980), as well as the structure of the provirus and its ability to integrate in many sites in the host genome (*Hughes* et al. 1978; *Steffen* and *Weinberg* 1978; *Sabran* et al. 1979; *Ringold* et al. 1979; *Martin* et al. 1979; *Cohen* et al. 1979; *Bacheler* and *Fan* 1979; *Gilmer* and *Parsons* 1979).

The first two technological tools, solution hybridization and the Southern transfer technique, made possible an increase in the sensitivity and specificity with which particular pieces of DNA could be detected. Recombinant DNA techniques (when finally permitted under the NIH guidelines) resolved the underlying problem: miniscule amounts of viral DNA contaminated with vast amounts of cellular DNA. Recombinant DNA, together with DNA-sequencing techniques, has brought our knowledge of the structure of retrovirus DNAs, both integrated and unintegrated, to its present level.

Before recombinant DNA technology was available, the only means for the preparation of significant amounts of pure viral DNA was in vitro DNA synthesis. This was an invaluable tool for the manufacture of cDNA probes for hybridization experiments, and could also be used to examine the mechanics of viral DNA synthesis.

The retroviruses are enveloped viruses; they are nonlytic and acquire an outer membrane when the immature virion core, containing the genomic RNA and the viral enzyme RNA-dependant DNA polymerase, is budded out through the outer membrane of the host cell. When a retrovirus virion penetrates a susceptible cell, viral DNA synthesis is initiated in the cytoplasm, presumably in some form of virion core structure. The virion core apparently contains all the machinery necessary to carry out viral DNA synthesis; virions treated with mild detergents which disrupt the virus membrane and permit the entry of deoxynucleotide triphosphates carry out extensive DNA synthesis when supplied only with the triphosphates and appropriate salts. In some cases, a complete and faithful DNA copy of the viral RNA can be synthesized (*Junghans* et al. 1975; *Rothenberg* et al. 1977; *Gilboa* et al. 1979a, *Benz* and *Dina* 1979; *Boone* and *Skalka* 1980; *Bosselman* and *Verma* 1980) and shown to be infectious (*Rothenberg* et al. 1977; *Gilboa* et al. 1979a; *Boone* and *Skalka* 1980). In most, but not all, instances the synthesis of viral DNA in detergent-activated virions faithfully parallels events in the infected cell, and has provided valuable insight into the precise molecular basis of retrovirus DNA synthesis. The principal advantages of in vitro DNA synthesis are that a large number of purified virions can be employed in a single in vitro reaction, yielding substantial amounts of viral DNA, and that the viral DNA which is synthesized is made in the absence of host cell DNA; it can be examined, either isotopically labeled or not, in the absence of the contaminating, and confusing, host sequences. Such experiments must be viewed carefully, however, since the final objective is to analyze events as they take place in vivo, and only

in those cases where the in vitro experiments correlate with and are documented by in vivo data can they be fully trusted. However, the correlations between in vitro and in vivo data have been very good, and the in vitro experiments can serve as a valuable guide even when in vivo data are absent.

## 2 How the RNA Genome is Copied into DNA

The integrated linear viral DNA is shown schematically in Fig. 1. The linear DNA is colinear with the genomic RNA, and at each end it has additional sequences derived from both ends of genomic RNA. These additional sequences are created during the two transfers of template which take place during viral DNA synthesis. Because the additional sequences are present at both ends of the viral DNA, the ends of the DNA are redundant, and form a direct repeat. The redundancies are called long terminal repeats (LTRs), and are made up of three segments: the $U_3$ region (unique sequences from the 3′ end of viral RNA), R (the redundant sequence found at both the 3′ and 5′ ends of viral RNA),

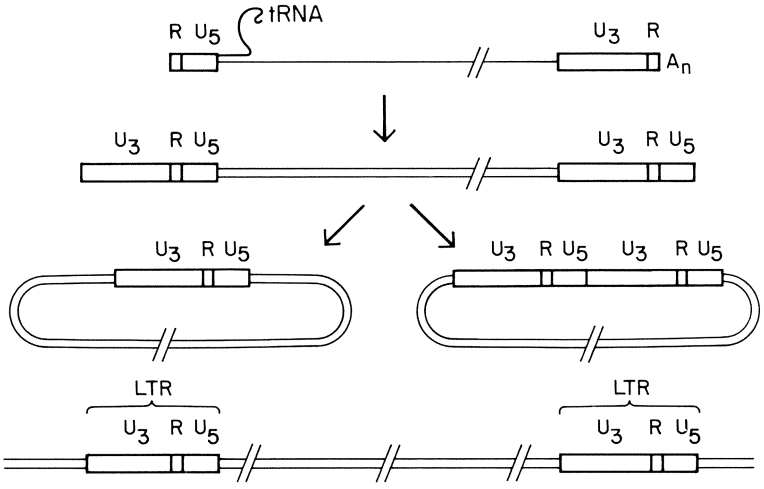



Fig. 1. The relationship of retrovirus genomic RNA, unintegrated viral DNA, and the provirus. *Top*, schematic representation of genomic RNA. The tRNA primer is hydrogen-bonded to viral RNA near the 5′ end of the genome. The redundancies found at both ends of the viral RNA are denoted *R*, and the $U_3$ and $U_5$ sequences are also marked. The relative size of $U_3$ and $U_5$ varies from one retrovirus species to another; *the drawing* shows the relationship for the ASV-ALV viruses. *Upper middle*, Unintegrated linear viral DNA, which contains, relative to the RNA from which it is derived, extra sequences at both ends. These extra segments form a direct repeat called the long terminal repeat (LTR), *Lower middle*, the two major closed circular supercoiled forms that derive from the unintegrated linear viral DNA. The smaller (*left*) contains a single copy of the LTR; the larger, (*right*) has two copies of the LTR. *Bottom*, the provirus closely resembles the unintegrated linear viral DNA in structure, except that it is attached to host DNA sequences at both ends. A comparison of the structure of the provirus with the structure of the genomic RNA demonstrates that the provirus is well designed as a template for viral RNA synthesis; signals for the initiation of viral RNA synthesis, as well as for poly-A addition, are carried within the LTRs

and $U_5$ (the unique sequences found at the 5′ end of the viral RNA). The structure of unintegrated linear viral DNA can be summarized: $U_3RU_5$ viral genes $U_3RU_5$. Thus, the virus has solved the problem of manufacturing a complete copy of its genome with an enzyme which requires a primer. The elegant solution depends on the small redundancies, R, which lie at the ends of the RNA, and gives rise to a DNA molecule in which the small redundancies in the RNA lie within the much larger redundancies, the LTRs, found at the ends of the linear DNA. The LTRs also provide the DNA copy of the virus with solutions to the problems of integration and of synthesizing RNA copies of the genome.

## 2.1 The Structure of Virion RNA and the Initiation of Viral DNA Synthesis

The genomes of retroviruses are single-stranded RNA, and range in length from just under 4 kilobases to just over 9 kilobases. The prototype genomes, those of the leukemia viruses, are about 7–8 kilobases in length, and closely resemble messenger RNA. The 5′ end of the RNA genome has an $m^7G^{5'}$ pppGm cap (*Furiuchi* et al. 1975; *Rose* et al. 1976), and the 3′ end of the RNA has a poly-A tail about 200 bases long (*King* and *Wells* 1976; *Bender* and *Davidson* 1976). The viral genome is plus-stranded (+); all of the messages have the same polarity as the genome (*Weiss* et al. 1977; *Hayward* 1977). The genome of a leukemia virus contains three known genes, which take up essentially the entire coding capacity. The genes are, in order, from the 5′ end of the RNA: *gag*, which encodes the group-specific antigens of the virion core; *pol*, the RNA-dependent DNA polymerase found in the virion, and *env*, the glycoprotein found in the outer membrane of the virus that plays a crucial role in the successful absorption of the virion by the host cell. In the virion genomic RNA is present as a dimer, the two subunits are hydrogen-bonded together near their 5′ ends (*Bender* and *Davidson* 1976; *Beemon* et al. 1976). Also found in the virion are several species of low-molecular-weight RNA (*Sawyer* and *Dahlberg* 1973; *Faras* et al. 1973), and, according to some reports, small DNA molecules as well (*Levinson* et al. 1972). Only one of these low-molecular-weight nucleic acids is known to have a specific function. A specific tRNA molecule is hydrogen-bonded to genomic RNA near the 5′ end of the viral RNA; the 3′ end of the tRNA is a perfect complement of the corresponding region in the retrovirus genome. The particular species of tRNA bound to genomic RNA is specific for the particular species of virus; tRNA trp is used in the avian sarcoma-avian leukosis viruses (ASV-ALV), and tRNA pro in in the murine leukemia viruses (*Sawyer* and *Dahlberg* 1973, 1974; *Taylor* et al. 1975; *Faras* and *Dibble* 1975; *Waters* et al. 1975; *Harada* et al. 1975, 1979; *Peters* et al. 1977). This tRNA plays a crucial role in the retrovirus life cycle; it serves as the point of initiation of retrovirus DNA synthesis. The presence of the primer underscores the experimental evidence that purified RNA-dependent DNA polymerase cannot initiate DNA synthesis de novo, a property it shares with all other known DNA polymerases (*Baltimore* and *Smoler* 1971; *Goodman* and *Spiegelman* 1971; *Leis* and *Hurwitz* 1972; *Hurwitz* and *Leis* 1972, *Wells* et al. 1972). The position of the tRNA primer precisely defines the site of initiation of viral DNA synthesis: the exact position depends on the particular retrovirus being studied, but is always 100–200 bases from the 5′ end of the virion RNA (*Taylor* and *Illmensee* 1975; *Peters* and *Dahlberg* 1979).

## 2.2 The First Transfer Between Templates

The first DNA to be synthesized is minus-stranded (-), that is, of opposite polarity to the RNA genome and the viral mRNAs. Since DNA synthesis would seem to run out of template almost immediately, there is a strong theoretical requirement to transfer DNA syn-

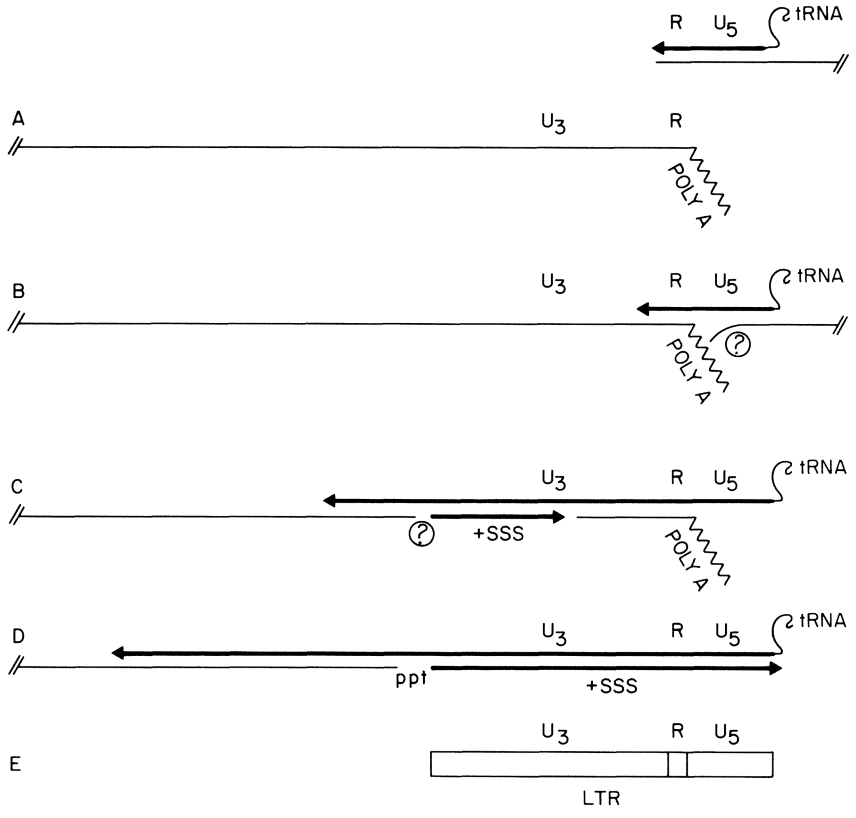



Fig. 2. The first transfer of DNA synthesis between templates, and the initiation of the specific (+) strand segment (+) strand strong stop. In *panel A*, a recently initiated (-) strand of DNA, still associated with its template RNA and linked to the tRNA primer, is shown on the *upper right*. The RNA strands are shown as *thin lines*, DNA strands by *thick lines*. This DNA-RNA hybrid has been positioned in the drawing over the 3′ end of a viral RNA molecule in such a fashion that the two identical R sequences are aligned. In *panel B*, the transfer between templates has taken place. Since the precise mechanism which exposes the R segment on the newly synthesized (-) strand DNA is not clearly understood this step is marked by a *question mark*. In *panel C*, the synthesis of the (-) strand DNA is proceeding along the template to which it was just transferred, and the specific (+) strand segment (+) strand strong stop (+SSS) has just been initiated. Since the mechanism of initiation of (+) strand strong stop is unknown, a *question mark* is placed at the site of (+) strand strong stop initiation. In panel *D*, synthesis of the (-) strand is proceeding along the RNA template, and (+) strand strong stop is now complete, having copied the first few bases of the tRNA primer (see text). The relationship of these structures to the LTR (panel *E*) is precise; the initiation site for the (-) strand and for (+) strand strong stop define the extent of the LTR. This is the LTR which will reside at the right end of the completed molecule, although the (+) strand strong stop will be transferred to he opposite end of the viral DNA in the second transfer of templates (see Fig. 3)

thesis to the 3′ end of either the same RNA molecule or to the second genomic RNA found in the virion. This transfer can be accomplished because the ends of the RNA molecule contain redundancies. The redundant region, R (see Fig. 1 and 2), is found both at the 5′ terminus of the RNA, and next to the poly-A tail at the 3′ terminus. These R sequences are 16–21 bases in length in the ASV-ALV viruses (*Stoll* et al. 1977; *Haseltine* et al. 1977; *Schwartz* et al. 1977), and about 50–60 bases in length in the MuLV viruses (*Coffin* et al. 1978). Although the transfer of the growing (-) DNA strand clearly takes place, molecular details are lacking. There are two reasonable models. The first suggest that part of the 5′ template RNA is degraded, presumably by the RNAse H activity of RNA-dependent DNA polymerase, which specifically degrades RNA found in RNA-DNA hybrids. The second model suggests that the DNA and RNA molecules are not tightly base-paired during synthesis, and that the transfer of templates can take place without degradation of the RNA genome. In vitro studies have, in general, favored degradation of the 5′ end of viral RNA, despite the presence of the $m^7G$ cap (*Darlix* et al. 1977; *Collett* et al. 1978; *Friedrich* and *Moelling* 1979). Whatever model ultimately proves correct, the transfer undoubtedly does take place; both in vitro and in vivo studies of the structure of the DNA portion of the replicative intermediates permit no other conclusion (*Haseltine* et al. 1979; *Dhar* et al. 1980; *Sutcliffe* et al. 1980; *Van Beveren* et al. 1980; *Amer* et al. 1981; *Swanstrom* et al. 1981a, 1981b), (see Fig. 2).

It is unclear whether the growing (-) strand DNA is transferred to the 3′ end of the same RNA molecule on which it was initiated, or to the second viral RNA present in the virion. Once the transfer takes place, the growing (-) strand DNA is free to proceed along the length of the viral RNA genome (see Fig. 2). Compared with rates of DNA synthesis by conventional DNA polymerases, viral DNA synthesis is relatively slow. Both in vivo and in vitro, it takes several hours to make a complete copy of the viral genome (*Rothenberg* and *Baltimore* 1977; *Varmus* et al. 1978). In the presence of melittin, significantly faster rates of polymerization can be achieved in vitro (*Boone* and *Skalka* 1980, 1981); however, since the rates of synthesis observed under these conditions are faster than in vivo rates, the significance of these observations is not clear. It is uncertain why the viral RNA-dependent RNA polymerase should be so much slower than the DNA-dependent DNA polymerases; however, the slow rate of polymerization has been most helpful, since it permits a much easier dissection of the steps in viral DNA synthesis.

## 2.3 The Initiation of the Specific (+) Strand Strong Stop DNA and the Second Transfer Between Templates

If the transfer of the growing DNA (-) strand between the 5′ and 3′ ends of viral RNA is the first crucial step in the synthesis of viral DNA, the second is the specific initiation and synthesis of a defined (+) strand DNA, which takes place long before the growing (-) strand is complete (*Varmus* et al. 1978; *Gilboa* et al. 1979b; *Mitra* et al. 1979). This (+) strand, now called (+) strand strong stop, is initiated at the end of a polypurine tract near the 3′ end of viral RNA (see Fig. 2). There is considerable conservation of the sequence near the site of initiation of this (+) strand, although the distance from the R sequence to the site of initiation for the (+) strand strong stop varies from about 160 bases in the avian endogenous virus, RAV-0 (*Hughes* 1982), to over 900 bases in the mouse mammary tumor viruses (MMTV) (*Majors* and *Varmus* 1981; *Donehower* et al. 1981). The primer for

(+) strand strong stop is usually assumed to be RNA, and conjectures have been made that the RNAse H activity of the RNA-dependent DNA polymerase leaves an appropriate RNA primer when it degrades viral RNA. All attempts to identify a specific primer for (+) strand strong stop have failed, and the mechanism of its precise initiation is unknown. The (+) strand strong stop DNA uses the previously synthesized (-) strand DNA as a template (Fig. 2). The 3′ end of this (-) strand DNA is attached to the rRNA primer. There is evidence, at least in vitro, that the (+) strand strong stop copies the first 15 bases of the tRNA primer up to the $m_1A$ which presumably cannot be copied (*Mitra* et al. 1979; *Taylor* and *Hsu* 1980) (see Fig. 2).

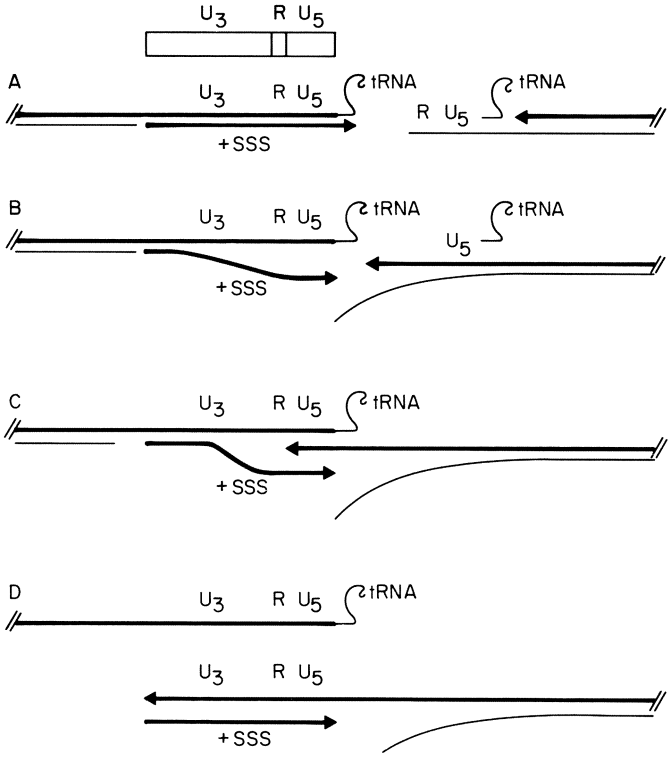



Fig. 3. The second transfer of DNA synthesis between templates. In *panel A*, at the left is the structure shown in the bottommost panel of Fig. 2; the viral (-) strand DNA with a complete (+) strand strong stop DNA (+SSS) attached to it. An abstract drawing of the LTR is given above it as a reference. On the right is shown the 5′ end of a viral RNA molecule, with the primer still attached and the growing (-) strand DNA. DNA is shown as *thick lines*, RNA as *thin lines*. In panel *B*, the (-) strand DNA has displaced the tRNA primer and is invading the double-stranded complex between (+) strand strong stop and its template. It must be emphasized that molecular details of these events are lacking; we do not know whether the tRNA primer in the complex on the left is removed first by RNAse H, how the (-) strand DNA in the complex on the right displaces the tRNA, or if the RNA template in the complex on the right is degraded by RNAse H. In *panel C*, the growing (-) strand displaces (+) strand strong stop from its template and copies it, giving rise to the completed double-stranded structure shown at the *bottom of panel D*. Again it should be clear that the position of initiation of (+) strand strong stop defines the end of the LTR, and in this case the left end of the molecule, since this is the LTR found at the left end of the unintegrated linear viral DNA (see Fig. 1)

Plus strand strong stop apparently has a role to play in the third crucial step of viral DNA synthesis, the second transfer of the (-) strand DNA (*Varmus* et al. 1978, 1979; *Shank* et al. 1978b; *Gilboa* et al. 1979b). As the growing (-) strand proceeds toward the 5′ end of viral RNA, it presumably encounters a barrier, the tRNA primer bound to virion RNA. Two obvious solutions exist: (a) the growing DNA simply displaces the tRNA molecule, or (b) the (+) strand strong stop may have already displaced the tRNA from the genomic RNA (Fig. 3). Once past the tRNA primer, the (-) strand could proceed to the end of genomic RNA which may be partially degraded by RNAse H. It is clear from the structure of viral DNA intermediates that a second transfer of the (-) strand takes place, and there is considerable indirect evidence that this transfer takes place to the (+) strand DNA and not to an RNA molecule (see Fig. 3). (*Shank* et al. 1978b; *Mitra* et al. 1979; *Gilboa* et al. 1979b; *Varmus* et al. 1979; *Baltimore* et al. 1979). The strongest arguments come from the structure of the completed linear viral RNA molecule (Fig. 1), a molecule whose (-) strand terminates at (or very near) a sequence complementary to the sequence where the (+) strand strong stop initiates. The major difficulty with the proposal that (+) strand strong stop is the template is that this proposal demands that RNA-dependent DNA polymerase be able to invade and copy a double-stranded DNA template, an activity it supposedly lacks (*Hurwitz* and *Leis* 1972). Strand displacement synthesis can take place, however, from a double-stranded DNA template in in vitro reactions in the presence of melittin (*Boone* and *Skalka* 1981), and the earlier reports that RNA-dependent DNA polymerase cannot strand displace DNA from a double-stranded DNA template may reflect inadequate reaction conditions. Although the issue is not fully resolved, the available data favor (+) strand DNA as a template for the growing (-) strand after the second transfer of templates (see Fig. 3). Once the growing (-) strand copies the tRNA primer binding site on genomic RNA, it contains sequences complementary to (+) strand strong stop. After the templates are exchanged for the second time the (-) strand can be completed by copying (+) strand strong stop (see Fig. 3). To complete the linear viral DNA, a (+) strand copy of this complete (-) strand is made. There is an apparent difference between the ASV-ALV virus in which these (+) strands are synthesized as relatively small segments (*Varmus* et al. 1978; *Shank* and *Varmus* 1978; *Kung* et al. 1981) and not rapidly ligated, and the MuLV, MMTV, and SNV viruses where the (+) strand is either made in a single piece, presumably primed by the (+) strand strong stop, or the nascent (+) strands are ligated so rapidly both in vivo and in vitro that small (+) strand segments cannot be seen (*Gilboa* et al. 1979b; *Chen* and *Temin* 1980; *Kung* et al. 1981).

The linear viral DNA is synthesized in the cytoplasm of infected cells (*Varmus* et al. 1974, 1978; *Guntaka* et al. 1976; *Fritsch* and *Temin* 1977a; *Shank* and *Varmus* 1978). (Visna is the exception; the linear viral DNA of visna is made in nucleus, (A. Haase, unpublished observation). Apparently only viral enzymes are required for these events; complete linear DNA can be made in detergent-activated virions supplied only with salts and deoxynucleoside triphosphates (*Junghans* et al. 1975; *Rothenberg* et al. 1977; *Gilboa* et al. 1979; *Benz* and *Dina* 1979; *Boone* and *Skalka* 1980; *Bosselman* and *Verma* 1980).

## 2.4 The Synthesis of Circular DNA Molecules

The linear viral DNA, synthesized in the cytoplasm of infected cells, migrates from the cytoplasm to nucleus (*Shank* and *Varmus* 1978). In the nucleus the linear is the precursor

for the two major forms of closed circular supercoiled viral DNA (Fig. 1). The smaller of the two supercoiled forms contains a single copy of the LTR, and the larger, two copies of the LTR (*Shank* et al. 1978b; *Hsu* et al. 1978). Dimeric circles have also been detected in small amounts; however, the function, if any, of the multimeric circles is unknown (*Goubin* and *Hill* 1979; *Kung* et al. 1981). The smaller supercoil is presumably formed by some sort of homologous recombination between the LTRs at the ends of the linear viral DNA. The circle which contains two LTRs is presumably formed by some sort of ligation event; however, the precise nature of this event is unknown. The main difficulty in describing the ligation event exactly is defining the ends of the linear DNA precisely. There is too little linear DNA present in infected cells (1–10 copies per cell in a standard acute infection) to analyze directly. Cloning the linear in *E. coli* and amplifying after cloning is not useful, since the ends of the linear must be altered to permit cloning. These strictures do not apply to the cloning of the circular forms of unintegrated viral DNA, and circular viral DNAs with both one and two copies of the LTRs have been cloned and sequenced (*Hager* et al. 1979; *Shoemaker* et al. 1980; *Ju* and *Skalka* 1980; *DeLorbe* et al. 1980; *Swanstrom* et al. 1981b). The region in the larger circle which joins the two LTRs together is of particular interest. In the complete form of the larger circle, the end of the $U_3$ region is joined exactly to the end of $U_5$. This joining creates a small imperfect inverted repeat. The size of the inverted repeat is variable from one retrovirus species to another; the variation among the carefully studied species is from 5 base pairs (SNV) in the inverted repeat to 18–23 base pairs in the MuLV-MSV viruses. These inverted repeats are found at the boundary between $U_3$ and $U_5$. Because $U_3$ and $U_5$ sequences are found within each LTR, these inverted repeats are also found at the ends of each LTR (*Dhar* et al. 1980; *Shimotohno* and *Temin* 1980; *Ju* and *Skalka* 1980; *Swanstrom* et al. 1981b). Although we now know a great deal about the sequence organization of the viral DNA intermediates, particularly the circular intermediates, we cannot answer a crucial question: which, if any, of the three major unintegrated DNA intermediates in the immediate precursor of the integrated provirus?

## 3 Integration of Viral DNA and Provirus Structure

Several groups have carefully examined the structure of the provirus and the flanking host DNA, first by the transfer procedure of Southern (*Hughes* et al. 1978, 1981c; *Steffen* and *Weinberg* 1978; *Martin* et al. 1979; *Sabran* et al. 1979; *Ringold* et al. 1979; *Cohen* et al. 1979; *Bacheler* and *Fan* 1979; *Gilmer* and *Parsons* 1979; *Canaani* and *Aaronson* 1979; *Jenkins* and *Cooper* 1980, and more recently by molecular cloning and DNA sequencing (*Dhar* et al. 1980; *Shimotohno* and *Temin* 1980; *Shimotohno* et al. 1980; *Majors* and *Varmus* 1981; *Hishinuma* et al. 1981; *Hughes* et al. 1981a).

We now know that proviruses can enter a large number of sites in the host genome, and that the same portion of viral DNA is always joined to host DNA. The integrated provirus resembles an unintegrated viral DNA molecule with the very ends of the linear viral DNA molecule joined to host sequences. Thus, the arrangement of the viral genome can be given: cell DNA $U_3RU_5$ viral genes $U_3RU_5$ cell DNA. The integration event does not lead to the loss of host sequences, rather the insertion of viral DNA causes a small (4–6 base pairs) duplication of host sequences at the site of integration. The structure of the provirus is reminiscent of the movable genetic elements found in prokaryotes and lower eukaryotes, and this structural similarity suggests that there may be functional

similarities, at least in the insertion of the movable genetic elements and retrovirus proviruses.

## 3.1 Host Sequences Chosen for Provirus Integration

Proviruses, like transposable elements, join a defined pair of sites in their own DNA to a large number of different sites in the genomes of their hosts. Proviruses have been shown to enter sites on many different chromosomes (*Hughes* et al. 1981c), as does the yeast transposon, TY-1 (*Calos* and *Miller* 1980), and the mobile genetic element of *Drosophila, copia* (*Calos* and *Miller* 1980). Although proviruses can enter many sites, it is not clear whether all potential integration sites in the genome are equally favored. Some of the prokaryotic transposons which can enter many sites still have preferences: for example, the *E. coli* transposon, Tn-9, has a distinct preference for T-A-rich sequences. The available data do not yet allow us to rigorously exclude the possibility that a provirus has a preference for integrating into certain regions or into certain sequences. It can be said, however, that if a specificity does exist, large numbers of acceptable sites are well distributed about the genomes of most organisms. It is also clear that if there is specificity it is of such a nature that direct comparisons of the flanking host sequences from several proviruses of a particular retrovirus species integrated into the same host cell DNA have not produced a pattern that can be readily discerned (*Shimotohno* and *Temin* 1980; *Shimotohno* et al. 1980; *Majors* and *Varmus* 1981; *Hughes* et al. 1981a; *McClements* et al. 1981).

There are two special cases of apparent specificity of integration worth mentioning. In ALV-induced tumors, the integrated provirus frequently resides in the immediate vicinity of the cellular oncogene c-*myc* (*Neel* et al. 1981; *Payne* et al. 1981, 1982; *Hayward* et al. 1981). This does not reflect specific integration, but rather the selection, from a large number of integration events, of those few which induce neoplastic transformation. The relative inefficiency of ALV-induced leukemogenesis reflects the fact that only a tiny fraction of the total integration events are near c-*myc*. (The activation of c-*myc* by an integrated provirus will be discussed in a later section.) The question that is raised in these experiments is why other known oncogenes, c-*src* for example, are not activated by ALV. Not all the ALV tumors have the provirus next to c-*myc*; however, none are known with the provirus next to any other known oncogene, including c-*src*. Several explanations are possible; the possibility which is relevant for this discussion is simply that the ALV provirus integrates next to c-*myc* much more frequently than it integrates next to other endogenous oncogenes, which, if correct, argues for some kind of preference for integration sites near c-*myc*.

The only direct evidence for a specific site or region for integration is the studies of baboon endogenous virus proviruses in human cells. The baboon endogenous virus provirus integrates into several different sites in baboon cells (*Cohen* et al. 1980, 1981); however, in human cells it apparently always inserts into sites on the short arm of chromosome 6 (*Lemons* et al. 1977, 1978, R. Lemons, F. O'Brien, and J.C. Cohen, unpublished observations). Southern transfer experiments suggest that there are, at a minimum, several integration sites on the short arm of chromosome 6. These multiple integration sites may be related, however. In all the integration events, the end of all of the proviruses lies very close to a Pst I site in host DNA (R. Lemons, F. O'Brien, and J.C. Cohen, unpublished observations).

## 3.2 Host Sequences are Duplicated During Integration

The insertion of a provirus does not result in the loss of a cellular DNA sequence; rather, as is the case of *copia* (*Dunsmuir* et al. 1980), TY-1 (*Farabaugh* and *Fink* 1980; *Gafner* and *Phillipsen* 1980), Mu (*Allet* 1979; *Kahmann* and *Kamp* 1979), and several bacterial transposons (*Calos* and *Miller* 1980), the insertion event is accompanied by the duplication of a small region of cellular DNA (*Majors* and *Varmus* 1981; *Hughes* et al. 1981a; *Hishinuma* et al. 1981; *McClements* et al. 1981). Because of this duplication, the host sequences immediately flanking the provirus form a small direct repeat (*Dhar* et al. 1980; *Shimotohno* and *Temin* 1980; *Shimotohno* et al. 1980; *Majors* and *Varmus* 1981; *Hishinuma* et al. 1981; *Hughes* et al. 1981a). The most likely explanation for these observations is that during the integration event a staggered cut is made in host DNA, and the repair of this staggered cut gives rise to the duplication of the host sequences. The size of the direct repeat appears to be directly related to the strain of virus; MMTV and the ASV-ALV viruses induce 6-base-pair repeats, SNV a 5-base-pair repeat, and the MULV-MSV viruses a 4-base-pair repeat. Unfortunately, these experiments were done in a number of different cell types from several different host species, and for this reason an absolute correlation between virus species and repeat size cannot be made.

If, as is likely, the size of the induced repeat of host sequences does correlate with the virus species, it implies the intimate participation of a viral protein in the integration of proviral DNA.

## 3.3 The Provirus Sequences that are Joined to Host DNA

Since proviruses always join the same sites in their own DNA to host DNA, the same viral DNA sequences (the LTRs) are always found at the ends of a provirus. The joining of host and virus DNAs results in the apparent loss of 2 base pairs from each end of the viral DNA (*Dhar* et al. 1980; *Shimotohno* and *Temin* 1981; *Majors* and *Varmus* 1981; *Hughes* et al. 1981a), comparing the sequences present in the provirus LTRs with the joint between the LTRs in the larger unintegrated circular form of viral DNA. The ends of retrovirus LTRs all show sequence homology. The last two bases at each end of all known proviruses are $\frac{TG}{AC}$ at the left, and $\frac{CA}{GT}$ on the right. It is interesting to note that these base pairs are also found at the ends of eukaryotic transposable elements (*Temin* 1982). The basic structural organization of the provirus is also reminiscent of some of the prokaryotic and eukaryotic transposons. Like proviruses, whose ends form a substantial direct repeat (the LTRs), the ends of TY-1, *copia*, and Tn-9 also form substantial direct repeats (*Calos* and *Miller* 1980). The very ends of each LTR form short, imperfect inverted repeats (*Dhar* et al. 1980; *Ju* and *Skalka* 1980; *Sutcliffe* et al. 1980; *Van Beveren* et al. 1980; *Swanstrom* et al. 1981b; *Majors* and *Varmus* 1981; *Hughes* et al. 1981a). Both *copia* and Tn-9 have similar small inverted repeats at the ends of the long direct repeat (*Calos* and *Miller* 1980; *Levis* et al. 1980). This is not universal among the movable elements flanked by long direct repeats; the long direct repeats of the yeast movable element, TY-1, lack the small inverted repeats (*Farabaugh* and *Fink* 1980; *Gafner* and *Phillipsen* 1980) (unless a tiny inversion of 2 base pairs is counted).

In addition to direct examination of the integrated provirus, studies on cloned unintegrated circular proviral DNA produced evidence for the kinds of rearrangements

which would be produced by inter- or intramolecular integration events (*Shoemaker* et al. 1980; *Ju* et al. 1980). The molecular consequences of these events closely parallels the molecular events during recombination into host DNA (*Shoemaker* et al. 1980; *Ju* et al. 1980).

All of these observations suggest that the mechanism of integration of retrovirus proviruses is probably closely related to the mechanism of integration of the transposable elements. How similar is not yet clear; however, recent analysis of DNA from *Drosophila* cells reveals unintegrated closed circular supercoiled copies of the *copia* elements which contain either one or two copies of the 276-base-pair sequence which forms the direct repeats at the end of integrated *copia* DNA (*Flavell* and *Ish-Horowicz* 1981).

## 3.4 Which of the Unintegrated Retrovirus DNAs is the Immediate Precursor of the Provirus?

The similarities between the integration of transposable elements and retrovirus proviruses does not help solve one of the remaining mysteries of retrovirus integration: what is the direct molecular precursor of the provirus, since little more is known about the integration of transposable elements than about the integration of retrovirus DNA.

Although none of the three forms of integrated viral DNA found in nuclei of infected cells can be ruled out, the arguments pro and con for each form as a precursor can be considered. In bacterial systems a plasmid carrying a single insertion element (which would correspond to a smaller unintegrated viral DNA with one LTR) can give rise through recombination to a structure which superficially resembles a provirus (*Shapiro* 1979; *Calos* and *Miller* 1980). This has led to proposals that the small circle with a single LTR is the precursor of the provirus. However, all the simplest forms of such a model predict that the large circle with two LTRs would produce proviruses bounded at one or both ends by tandem LTRs. Proviruses of this type have not been observed, suggesting that models of provirus integration which are close homologs of *Shapiro's* model (1979) for the insertion of transposons should be viewed with caution.

Since no cellular information is lost when viral DNA is inserted into the genome the viral precursor is probably either an actual circle, or a linear in which the ends are held close together, presumably by a protein complex. Unfortunately, either of the two viral DNAs most likely to serve as direct precursors of the provirus fit this description. The larger closed circular form is an aesthetically pleasing candidate, being already circular; however, the linear molecule could easily be held as a circle in vivo by interactions with proteins. If the unintegrated linear viral DNA is the immediate precursor of the provirus, the large circle with two LTRs could be explained as the product of a side reaction, since the ends of the linear would be active in recombination. A portion of the unintegrated circular DNAs present after an acute infection appear to be products of abnormal integration events, some apparently inter- and some intramolecular in nature (*Shoemaker* et al. 1980; *Ju* et al. 1980). We cannot, with the current data, rationally choose which of the unintegrated viral DNAs is the immediate precursor of the provirus.

## 3.5 Host Factors Involved in Integration

If host factors are not required for synthesis of unintegrated viral linear DNA, and in some instance perhaps not for the formation of circular viral DNA either (*Dina* and *Benz*

1980; *Guntaka* 1980), what host factors, if any, are needed for integration? There is apparently a direct or indirect requirement for the host cells to divide; cells held in $G_0$ by serum starvation do not successfully synthesize or integrate viral DNA, although the processes of viral DNA replication and integration are resumed when the cells are refer with medium containing serum (*Varmus* et al. 1977; *Fritsch* and *Temin* 1977b). Although the precise nature of the block is not resolved, there are reports that synthesis of linear viral DNA is started, but not completed (*Fritsch* and *Temin* 1977b; *Varmus* et al. 1977). This could reflect a general depletion of cellular resources by serum starvation, in particular, a lack of the deoxynucleotide triphosphates necessary for viral DNA synthesis.

Better evidence for a direct requirement for host DNA replication comes from experiments in which the infection was performed coordinately with BudR density labeling of host DNA. Even when host DNA replication was minimal, the vast majority of viral DNA entered the newly replicated heavy-light host DNA, strongly implying a preference of the provirus for replicating, or recently replicated, host DNA (*Varmus* et al. 1979).

The replication of certain murine viruses is strongly restricted on certain strains of mouse cells (FV-1 restriction). This also indicates the involvement of a host factor or factors. Viral DNA is made under the restrictive condition but apparently not integrated efficiently (*Sveda* and *Soeiro* 1976; *Jolicoeur* and *Baltimore* 1976). There is evidence that in at least certain of the restrictive strains that the absolute amount of supercoiled DNA made during an acute infection is markedly reduced (*Yang* et al. 1980; *Jolicoeur* and *Rassart* 1980, 1981), and it has been reported that the linear DNA, which is present in normal or near-normal amounts in the restricted cells, is significantly less infectious in transfection assays than that made in parallel in permissive cells (*Yang* et al. 1980). FV-1 restriction has been linked with a particular site in the viral genome, a site within the coding region of the *gag* protein p30 (*Schindler* et al. 1977; *Hopkins* et al. 1977; *Gautsch* et al. 1978; *Tress* et al. 1979). The *gag* proteins form the virion core, and have an intimate relationship with virion RNA and RNA-dependent DNA polymerase (*Bandyopadhyay* and *Levy* 1978). The FV-1 system suggests that at least the p30 protein also interacts with a host factor or factors, and this interaction has some significant effect on the successful synthesis of unintegrated viral DNA. Such a proposal should be weighed against evidence suggesting that it is apparently possible to synthesize circular viral DNA in vitro, in detergent-activated virions, in the absence of all host factors (*Dina* and *Benz* 1980; *Guntaka* 1980).

## 3.6 Excision, Deletion, and Transposition of the Provirus

While the majority of proviruses are complete and contain two LTRs, there are examples of incomplete proviruses (*Hughes* et al. 1978, 1981b; *Martin* et al. 1979; *Hayward* et al. 1980). Incomplete proviruses are especially common among the endogenous avian viruses (*Hayward* et al. 1980; *Hughes* et al. 1981b). While some of the incomplete proviruses are missing only internal sequences, some (the more interesting class) have lost one of the LTRs. Such proviruses could arise in two ways: by the integration of an incomplete provirus, or, alternatively, by the integration of a complete provirus followed by the subsequent loss of a portion of the viral genome. We do not know which event is more likely; however, loss of only a portion of the provirus is not a principal pathway for the reversion of RSV-transformed cells that initially contain a complete RSV provirus (*Varmus* et al. 1981a). The entire coding region and one of the two LTRs of the provirus are lost

relatively readily, apparently by a homologous recombination event between the LTRs (*Donner* et al. 1980; *Hughes* et al. 1981b; *Varmus* et al. 1981b). Such events leave behind in the genome an LTR associated with no other viral sequences. A "free" LTR could profoundly affect the transcription of a particular region of the host genome, without adding any burdensome viral genes. Retrovirus LTRs are not the only elements usually associated with a mobile genetic element which are also found as "free repeats" in the host genome. The direct repeats associated with TY-1 and Tn-9 repeats are found "free" without other TY-1 or Tn-9 sequences (*Chow* and *Bukhari* 1977; *McHattie* and *Jackowski* 1977; *Cameron* et al. 1979); however, the only copies of the 276-base-pair sequence which flank the *Drosophila* element, *copia*, found in the *Drosophila* chromosome, are those found in direct association with *copia* sequences (*Levis* et al. 1980).

Can the integration event be reversed? Does a retrovirus genome ever excise precisely from the host chromosome? No evidence has ever been presented which documents such an event; events of this type, if they occur at all, are rare. It is also unlikely that there are any kind of frequent "transposition" events of proviruses which involve only DNA intermediates, whether dependent on provirus excision, or on synthesizing a DNA copy of the provirus directly as an intermediate in transposition (*Steffen* and *Weinberg* 1978; *Bacheler* and *Fan* 1979; S. Hughes, unpublished observations). There is considerable evidence that virions are matured, and rendered infectious, only after being budded from an infected cell (*Sarkar* et al. 1971; *Canaani* et al. 1973; *Stoltzfus* and *Synder* 1975; *Shapiro* et al. 1976; *Yoshinaka* and *Luftig* 1977; *Opperman* et al. 1977; *Cooper* and *Okenquist* 1978). The only well-documented exception to these rules involves the intercisternal A particles, which can apparently successfully infect cells after fusion of previously uninfected cells with cytoplasts containing intercisteral A particles. This successful infection during cell fusion, without the host cell budding infectious virions, suggests that, at least in this system, an intercellular reinfection might be possible (*Malech* and *Wivel* 1976). Even in this system, the mediating particle probably resembles a virion core, and, although direct evidence is lacking, it seems very likely that there will be an RNA intermediate in the process. It would seem more fruitful, in light of all the evidence, for those interested in transposons to look carefully for RNA intermediates in transposition, than for retrovirologists to search for proviruses giving rise directly to DNA intermediates for transposition.

## 3.7 Prospects for Further Experiments

One way to resolve precisely the various factors involved in integration would be to recapitulate the integration event either in vivo using cloned DNA, or better yet, in a defined in vitro system. Because of the difficulties in preparing pure unintegrated linear viral DNA, all the experiments have been done with cloned DNAs that contain LTRs present either in one or two copies. Such molecules have been reintroduced into cells by a variety of protocols, including calcium phosphate precipitation and microinjection (*Copeland* et al. 1981; Capecchi, unpublished observations). In all cases, the sequences at the ends of the LTRs used in normal retrovirus DNA integration are ignored, and the cloned DNA behaves with respect to integration like any other DNA inroduced into cells by calcium phosphate or microinjection. The failure of these experiments should be viewed, at least in retrospect, in light of the probable involvement of viral proteins in the

integration process. Bacterial transposons encode genes responsible for their transposition (*Calos* and *Miller* 1980). The complexity of the illegitimate recombination event of retrovirus DNA integration suggests a special mechanism, and the probable involvement of a viral protein or proteins. This argument is strengthened by the fact that the various species of retroviruses induce repeats of 4, 5, or 6 base pairs in cellular DNA, and, although the data is not absolutely definitive, since many different cell types are involved, there seems to be a direct correlation between the length of the induced repeat of cellular DNA and the species of retrovirus involved. MMTV and the ASV-ALV viruses induce a 6-base-pair repeat (*Majors* and *Varmus* 1981; *Hishinuma* et al. 1981; *Hughes* et al. 1981a), SNV a 5-base-pair repeat (*Shimotohno* and *Temin* 1980; *Shimotohno* et al. 1980), and the MuLV-MSV viruses a 4-base-pair repeat (*Dhar* et al. 1980; *Shoemaker* et al. 1980; *McClements* et al. 1981). This suggests that the virus, and by implication, a viral enzyme, determines the length of the cellular DNA repeats induced during integration into cellular DNA. Taken in this context, the failure of the previous experiments is understandable. It may be necessary to provide viral proteins, in addition to the correct viral precursor (which could be the linear). Using infected cells will probably not suffice, since many of the viral proteins are synthesized as polyproteins which are only matured in budded virions (*Yoshinaka* and *Luftig* 1977; *Oppermann* et al. 1977). The true DNA precursor may even be part of a residual virion core complex. This possibility is underscored by the observation that in certain circumstances, closed circular viral DNA molecules are synthesized in detergent-activated virions (*Dina* and *Benz* 1980; *Guntaka* et al. 1980). While it may eventually be possible to use an entirely in vitro system for integration, it must be borne in mind that retrovirus DNA has been shown to integrate only into newly replicated DNA (*Varmus* et al. 1979), a requirement which may be difficult to meet in vitro.

## 4 Transcription of the Provirus

The provirus is the template from which both virion RNA and viral messages are normally synthesized. (The possibility that unintegrated DNA can act as a template for transcription is not resolved.) The normal cellular transcriptional machinery transcribes the provirus and processes the mRNAs; viral components are not known to be involved. However, virion RNA must, on both theoretical and experimental grounds, interact with viral proteins, and it is possible that such interactions could have a role in regulating viral RNA processing. The transcriptional start site and the poly-A addition site reside with the LTR, although there is evidence which suggests that either the particular location of the provirus in the host genome, or the modification of the proviral sequences, or both, plays some role in regulating the level of provirus transcription.

### 4.1. Signals for Initiation of RNA Synthesis and Poly-A Addition are Encoded in the Provirus

The elucidation of the structure of the provirus made it clear that the initiation site for genomic viral synthesis and the poly-A addition site would both be within the LTR (*Hughes* et al. 1978; *Sabran* et al. 1979). In retrospect the advantages of having all the

signals for synthesizing and processing viral RNAs encoded within the provirus seem simple and obvious, particularly when it is known that the provirus can enter many sites in the host genome. It must be remembered, however, that the starting material from which the provirus is copied is the genomic RNA, and the complexities of linear DNA synthesis exist, at least in part, to provide an adequate DNA template that can be used in the subsequent synthesis of virion RNA and mRNA.

Most retroviruses appear to have only one transcriptional initiation start site and one poly-A addition site; both lie within the LTR. The start site and the poly-A addition site define the redundancy, R, found at both ends of viral RNA genome (*Stoll* et al. 1977; *Haseltine* et al. 1977; *Schwartz* et al. 1977). The start site coincides with the 5′ end of both genomic RNA and the viral mRNAs (*Mellon* and *Duesberg* 1977; *Cordell* et al. 1978). The position of initiation has been confirmed in vitro; the same site is used as the initiation site by RNA polymerase II in in vitro transcription systems (*Yamamoto* et al. 1980a; *Fuhrman* et al. 1981).

The transcriptional start site within the LTR looks very much like the transcriptional start site for a normal cellular gene; the kinds of sequences which are found associated with cellular promoters (TATAA boxes, etc.) are also found, appropriately located, within retrovirus LTRs (*Dhar* et al. 1980; *Shimotohno* and *Temin* 1980; *Sutcliffe* et al. 1980; *Yamamoto* et al. 1980b, *Van Beveren* et al. 1980; *Swanstrom* et al. 1981b). Cloned LTRs have been used to drive the transcription of genes linked to them in vitro (*Blair* et al. 1980; *Chang* et al. 1980; *Oskarsson* et al. 1980). When these chimeras are reintroduced into cells, the transcriptional start site within the LTR causes the linked sequences to be expressed at moderate to high levels.

The transcriptional start site is present in both LTRs, and in some cases both these start sites can be active, giving rise to both viral RNA and to transcripts of those host sequences which lie to the right of the provirus (*Quintrell* et al. 1980; *Hayward* et al. 1981; *Payne* et al. 1981, 1982). (The importance of these "downstream" transcripts will be discussed in a later section.)

The LTR also contains a poly-A addition signal (AATAAA); poly-A is added to the transcript at the end of the R sequence (Fig. 1). The relative positions of the transcriptional start and poly-A addition sites would make it theoretically possible to synthesize a transcript beginning at one side of an R sequence and finishing, with poly-A addition, on the other (Fig. 1). Presumably, the proximity of the start signal and the poly-A addition site prevents the poly-A site from being used efficiently until the entire viral genome has been copied. The primary nuclear transcript appears to closely resemble mature virion RNA. Virion RNA must contain, for obvious reasons, all the sequences present in the provirus, and cannot, therefore, be spliced. A subset of the viral messenger RNAs are known to be spliced (*Weiss* et al. 1977; *Mellon* and *Duesberg* 1977; *Rothenberg* et al. 1978; *Cordell* et al. 1978). The remaining mRNAs appear to be the same size as genomic RNA; however, indirect evidence now suggests that at least some of the "full length" viral mRNA probably has a small region removed by splicing (D. Schwartz, unpublished observations).

Mouse mammary tumor virus (MMTV) LTRs appear to contain additional sequences important in the regulation of transcription. The level of transcription of a subset of MMTV proviruses is enhanced by glucocorticoid hormones. This inducibility is now known to reside within the LTR; several groups have now made constructs in which a cloned MMTV LTR is linked to another gene. When the chimera is reintroduced into an

appropriate eukaryotic host, transcription of the sequences attached to the MMTV LTR is responsive to glucocorticoids (*Lee* et al. 1982; J. Majors, unpublished observations; G. Hager, unpublished observations). Experiments are currently underway to precisely localize the region within the MMTV LTR (which is the largest retrovirus LTR, around 1200 base pairs) that is involved in the glucocorticoid induction.

## 5 Influences of the Host on Provirus Expression

The provirus forms a complete multigenic transcriptional unit, and contains all the information necessary for the expression of the full complement of viral genes. The provirus, however, is inserted into the host genome, and there is evidence that the flanking host sequences, or modification of proviral sequences by the host, have significant effects on viral gene expression. Because a direct alteration of context of a particular provirus is usually not possible, the arguments concerning the effects of host sequences are, for the most part, indirect. The species of the host cell can affect the absolute level of mRNA transcripts, as well as the proportions of spliced and unspliced RNAs. Rous sarcoma virus RNA is present at a much higher level in chick cells than in mammalian cells (*Coffin* and *Temin* 1972; *Quintrell* et al. 1980); in all likelihood, this results from enhanced rates of transcription. However, the proportion of *src* mRNA (which is spliced) is much higher, relative to the other RNAs, in mammalian cells than in chick cells (*Quintrell* et al. 1980).

Proviruses present in different cell clones derived from infection of a particular cell line also show differences in expression (*Ringold* et al. 1979; *Berkower* et al. 1980; *Quintrell* et al. 1980). An interesting special case occurs when MMTV proviruses are introduced into a mink lung cell line by infection. A portion of the proviruses are transcriptionally inactive, even in response to glucocorticoids (J. Majors, unpublished observations). Since the MMTV LTR is known to contain sequences which permit MMTV transcription to be induced by glucocorticoids, the lack of induction shown by some MMTV proviruses presumably represents some host constraint on MMTV expression.

There is also an in vivo counterpart to these observations. Moloney murine leukemia virus (MoMuLV) can be introduced into early mouse embryos, and consequently into the germ line of mice (*Jaenisch* 1976). Mice were obtained with MoMuLV proviruses in different sites in the genome; mice with proviruses in different locations exhibit differences in which tissues express the experimentally introduced MoMuLV provirus, and differences in the time during development at which the provirus is first expressed (*Jaenisch* et al. 1981). These differences are heritable, again, some sort of host constraint or influence is probably involved.

There are also examples, particularly among the endogenous viruses, in which a particular provirus is expressed very poorly, yet when the provirus is induced and the cells reinfected, they give rise to a virus which replicates well. The endogenous chicken provirus, ev-2, is expressed at low levels in a small proportion of cells potentially capable of expression (*Crittenden* et al. 1974). The infectious virus which apparently derives from ev-2, RAV-0, is much more efficiently expressed after its proviruses are introduced to new locations in the genome by infection. These data can be explained in several ways: for example, that the RAV-0 virus derives from ev-2 only after some mutation or recombination event. However, taken together with the other available data, some effect of the host on the expression of the endogenous provirus seems more likely.

Two hypotheses (not necessarily mutually exclusive) have been used to explain the differences in levels of expression seen among proviruses. The first hypothesis is that the cellular sequences linked to the provirus exert a profound negative effect on the rate of provirus transcription. The second hypothesis is that the effects are not due to the sequence context per se, but to methylation of the provirus and its immediate surroundings.

Which data argue in favor of each hypothesis? In favor of the first hypothesis are experiments which apparently show that removing the host sequences by shearing the DNA results in increases in infectivity scored in transcription assays (*Cooper* and *Temin* 1976; *Cooper* and *Silverman* 1978). In favor of the second hypothesis are the observations that certain methylated proviruses that are essentially inactive in transfection experiments, become infectious after the proviral DNA is molecularly cloned and propagated in bacteria (*Harbers* et al. 1981). A simple interpretation of these data would be that the loss of the methyl groups during the propagation of the provirus in *E.coli* results in the increases seen in transfection infectivity. The chief criticism of both of these sets of experiments is that what is measured is based upon transfection, and only indirectly, at best, upon transcription.

Further evidence suggestive of effects on transcription due to methylation comes from experiments in which methylation inhibitors (5-azacytidine) have been shown to enhance the transcription of normally inactive methylated proviruses (*Groudine* et al. 1981); however, it is possible that the drug has effects on transcription other than those directly related to its effects on methylation.

These criticisms will, in all likelihood, be answered by more and better experiments; the data currently available strongly suggest that the host, through DNA context or DNA modification or some mechanism not yet fully understood, exerts, at least in some cases, a profound influence on the expression of a provirus. To turn the problem around: if the primary objective is to study the influence of neighboring sequences and/or host modification on the expression of cellular genes, the provirus can be viewed as an extremely useful tool. The provirus represents a set of genes with defined initiation and termination signals which can be inserted at a great variety of places in the genome, whose expression is controlled by cellular machinery. What better way to probe for the effects of context and/or modification?

## 6 Effects of Provirus Integration on Host Gene Expression

The integration of retrovirus provirus can profoundly influence the expression of genes in the region near the integration site: introduction of a provirus can have either a positive or negative effect. The simplest negative case would be the insertion of provirus directly into the gene or its regulatory elements. Two cases have been reported in which a provirus has a negative effect on gene expression. In DBA/2J mice, a particular endogenous MLV provirus is associated in genetic crosses with the coat color marker dilute; reversion of the dilute mutation is accompanied by the loss of the particular endogenous provirus (*Jenkins* et al. 1981). Although there is as yet no proof in this case that the endogenous provirus resides exactly at the dilute locus, this seems the most reasonable explanation.

The other case is more fully documented (*Varmus* et al. 1981b). Rat cells transformed by Rous sarcoma virus were reverted by a subsequent infection with the Maloney strain of MuLV. In two different revertant lines, the RSV provirus was shown to contain an MuLV provirus. Neither of the MuLV proviruses (each of which has the same orientation as the RSV provirus it resides within) entered the *src* gene itself; both have integrated into different regions of the RSV provirus upstream of *src*. In both cases, the site where the MuLV virus resides is normally eliminated from *src* mRNA by splicing. Direct analysis of the mRNA in the revertant cells demonstrates that the MuLV provirus interferes with *src* mRNA production. Transformed cells can be isolated from the MuLV reverted cell lines; these have lost most of the MuLV provirus, apparently by homologous recombination between the LTRs. The residual LTR is not sufficient to block expression of the *src* gene, even though the MuLV LTR contains a polyadenylation site. These observations are somewhat puzzling: why does the presence of a complete retrovirus prevent *src* expression that the presence of the LTR permits? The simplest answer would be that either sheer size of the MuLV insert or the RNA processing signals contained in the provirus interferes with *src* mRNA production. Analysis of the *src* mRNA does not distinguish between these possibilities. What is also unclear, since LTRs can in some cases serve as sites of initiation for downstream mRNAs (to be discussed below), is why the MuLV LTR does not itself induce *src* mRNA synthesis.

These questions aside, it is clear that the insertion of a provirus can block the expression of a nearby gene; although no fully documented examples are known, it seems likely that proviruses can, at least in some cases, integrate directly into a coding region, thereby destroying gene function.

Proviruses can also enhance the transcription of nearby genes (*Quintrell* et al. 1980; *Neel* et al. 1981; *Hayward* et al. 1981; *Payne* et al. 1981, 1982). Provirus LTRs carry sequences which function as promoters, as well as polyadenylation signals. Since there are two identical LTRs, it is logical to assume that the downstream LTR can initiate transcripts of the cellular sequences lying to the right of the provirus. Such transcripts have been described in RSV-transformed mammalian cells (*Quintrell* et al. 1980), where the induced transcripts are presumably of small importance to the cell, and in ALV-induced tumors where a transcript originating in the LTR and copying part or all of the endogenous oncogene, c-*myc*, is deeply involved in the oncogenic event (*Neel* et al. 1981; *Hayward* et al. 1981; *Payne* et al. 1981, 1982).

The first hints that the site of integration, and by implication, the presence of the provirus, might play a crucial role in leukemogenesis came from studies which demonstrated that in many ALV-induced tumors the ALV provirus seemed to reside at or near a particular site in the host genome (*Neel* et al. 1981; *Payne* et al. 1981). Since ALV proviruses can, and do, integrate at a large number of sites in nontumorous tissue, location of the provirus takes on special significance, especially when taken together with the knowledge that the ALV provirus does not contain an oncogene. The full significance of the initial observations became clear only after it was shown that the special site in the host genome occupied by the ALV proviruses in tumors was adjacent to the cellular oncogene, c-*myc*, and that the presence of the provirus was coupled with a large increase in transcription from the c-*myc* locus (*Neel* et al. 1981; *Hayward* et al. 1981; *Payne* et al. 1982).

In some of the tumors, an ALV LTR lies upstream of the c-*myc* region, and the promoter in the LTR is in position to initiate a transcript which begins at the normal site

within the LTR (at the beginning of R) and continues through c-*myc*. These events have been documented in two ways: by cloning the appropriate DNA fragments from the tumor, and by direct demonstration in RNA from the tumor of a single transcript that contains the $U_5$ region from the viral DNA and c-*myc* sequences (*Neel* et al. 1981; *Hayward* et al. 1981; *Payne* et al. 1982).

The effects of the integrated provirus on the transcription of neighboring sequences is not confined to promoter insertion. Two other types of activation events are known, although both could be manifestations of a single phenomenon. c-*myc* can be activated by a provirus inserted upstream in the "wrong" orientation, that is, with the promoter in the LTRs pointed away from c-*myc*. In this case, the level of c-*myc* RNA is clearly elevated, but the c-*myc* transcript contains no detectable viral sequences (*Payne* et al. 1982). A provirus downstream of c-*myc* can also elevate the level of c-*myc* transcription. In the well-studied case, the provirus has the same transcriptional orientation as c-*myc* and the c-*myc* transcript ends within the retrovirus LTR, probably at the normal polyadenylation site at the junction between R and $U_5$ (*Payne* et al. 1982).

That proviruses are capable of action-at-a-distance effects on transcription is not surprising when taken in the context of other examples of action-at-a-distance effects. The SV40 72-base-pair repeats have significant effects on transcription over more than 2 kilobases (*Banerji* et al. 1981; *Moreau* et al. 1981); sequences which are functionally related to the SV40 72-base-pair repeats are found in the MuLV LTR (*Levinson* et al. 1981). The yeast transposable element, TY-1, which is structurally similar to a provirus, also shows the same kinds of action-at-a distance enhancement of the transcription of several cellular genes, and can exert such effects from several distinct locations. However TY-1 is always found upstream of the genes it activates, and is always found in one of the two possible orientations (*Errede* et al. 1980; *Williamson* et al. 1981; and G. Fink, personal communication).

In many cases, the ALV proviruses found in the tumor are extensively deleted, and lack most of the sequences which encode the viral replicative genes. Several explanations are possible; a simple suggestion is that in the animal the immune system selects against those cells which manufacture (and express on their surfaces) viral proteins. Immune surveillance would then favor the growth of tumor cells which contain only a deleted provirus, and select against those cells which contain an intact, and actively transcribed, provirus.

## References

Allet B (1979) Mu insertion duplicates a 5 base pair sequence at the host inserted site. Cell 16:123–129

Amer CA, Parsons JT, Faras AJ (1981) Direct proof of the 5′ to 3′ transcriptional jump during reverse transcriptase of the avian retrovirus genome by DNA sequencing. J Virol 38:398–402

Bacheler LT, Fan H (1979) Multiple integration sites for Moloney murine leukemia virus in productively infected mouse fibroblasts. J Virol 30:657–667

Baltimore D, Smoler D (1971) Primer requirement and template specificity of the DNA polymerase of RNA tumor viruses. Proc Natl Acad Sci USA 68:1507–1511

Baltimore D, Gilboa E, Rothenberg E, Yoshimura F (1979) Production of a discrete, infectious double-stranded DNA by reverse transcription in virions of Moloney murine leukemia virus. Cold Spring Harbor Symp Quant Biol 43:869–874

Baluda M (1972) Widespread presence, in chickens, of DNA complementary to the RNA genome of avian leukosis viruses. Proc Natl Acad Sci USA 69:576–580

Bandyopadhyay AK, Levy CC (1978) Effect of RNA tumor virus-specific protein p30 on reverse transcriptase. Intraspecies and interspecies interaction between reverse transcriptase and p30. J Biol Chem 253:8285–8290

Banerji J, Rusconi S, Schaffner W (1981) Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell 27:299–308

Beemon KL, Faras A, Haase A, Duesberg P, Masel J (1976) Genomic complexities of murine leukemia and sarcoma, reticulendotheliosis and visna viruses. J Virol 17:525–537

Bender W, Davidson N (1976) Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type 1 oncornavirus RNA molecules. Cell 7:595–607

Benz EW Jr, Dina D (1979) Moloney murine sarcoma virions synthesize full-genome-length double-stranded DNA in vitro. Proc Natl Acad Sci USA 76:3294–3298

Bergmann DG, Souza LM, Baluda MA (1980) Characterization of avian myeloblastosis-associated virus DNA intermediates. J Virol 34:366–372

Berkower AS, Lilly F, Soeiro R (1980) Expression of viral RNA in Friend virus-induced erythroleukemia cells. Cell 19:637–642

Blair DG, McClements WL, Oskarsson MK, Fischinger PJ, Vande Woude GF (1980) Biological activity of cloned Moloney sarcoma virus DNA: terminally redundant sequences may enhance transformation efficiency. Proc Natl Acad Sci USA 77:3504–3508

Boone LR, Skalka A (1980) Two species of full-length cDNA are synthesized in high yield by melittin-treated avian retrovirus particles. Proc Natl Acad Sci USA 77:847–851

Boone LR, Skalka A (1981) Viral DNA synthesized in vitro by avian retrovirus particles permeabilized with melittin. II. Evidence for a strand displacement mechanism in plus-strand synthesis. J Virol 37:117–126

Bosselman RA, Verma IM (1980) Genome organization of retroviruses. V. In vitro-synthesized Moloney murine leukemia viral DNA has long terminal redundancy. J Virol 33:487–493

Calos MP, Miller JH (1980) Transposable elements. Cell 20:579–595

Cameron J, Loh E, Davis R (1979) Evidence for transposition of dispersed repetitive DNA families in yeast. Cell 16:739–751

Canaani E, Aaronson SA (1979) Restriction enzyme analysis of mouse cellular type C viral DNA: emergence of new viral sequences in spontaneous AKRJ lymphomas. Proc Natl Acad Sci USA 76:1677–1681

Canaani E, von der Helm K, Duesberg P (1973) Evidence for 30–40s RNA as a precursor of the 60–70s RNA of Rous sarcoma virus. Proc Natl Acad Sci USA 70:401–405

Chang EH, Maryak JM, Wei CM, Shih TY, Shober R, Cheung HL, Ellis RW, Hager GL, Scolnick EM, Lowy DR (1980) Functional organization of the Harvey murine sarcoma virus genome. J Virol 35:76–92

Chen ISY, Temin HM (1980) Ribonucleotides in unintegrated linear spleen necrosis virus DNA. J Virol 33:1058–1073

Chow LT, Bukhari AI (1977) In: Bukhari AI, Shapiro JA, Adhya SL (eds) DNA insertion elements, plasmids and episomes, Cold Spring Harbor Laboratory, New York, pp 295–306

Coffin JM, Temin HM (1972) Hybridization of Rous sarcoma virus deoxyribonucleic acid polymerase product and ribonucleic acids from chicken and rat cells infected with Rous sarcoma virus. J Virol 9:766–775

Coffin JM, Hageman TC, Maxam AM, Haseltine WA (1978) Structure of the genome of Moloney murine leukemia virus: A terminally redundant sequence. Cell 13:761–773

Cohen JC, Shank PR, Morris VL, Cardiff R, Varmus HE (1979) Integration of the DNA of mouse mammary tumor virus in virus-infected normal and neoplastic tissue of the mouse. Cell 16:333–346

Cohen M, Davidson N, Gilden RV, McAllister RM, Nicolson M, Stephens R (1980) The baboon endogenous virus genome. II. Provirus sequence variations in baboon cell DNA. Nucleic Acids Res 8:4423–4440

Cohen M, Rein A, Stephens R, O'Connell C, Gilden R, Shure M, Nicolson M, McAllister R, Davidson N (1981) Baboon endogenous virus genome: molecular cloning and structural characterization of nondefective viral genomes from DNA of a baboon cell strain. Proc Natl Acad Sci USA 78:5207–5211

Collett MS, Dierks P, Parsons JT, Faras AJ (1978) RNase H hydrolysis of the 5′ terminus of the avian sarcoma virus genome during reverse transcription. Nature 272:181–183
Cooper GM, Okenquist S (1978) Mechanism of transfection of chicken embryo fibroblasts by Rous sarcoma virus DNA. J Virol 28:45–52
Cooper GM, Silverman L (1978) Linkage of the endogenous avian leukosis virus genome of virus-producing chicken cells to inhibitory cellular DNA sequences. Cell 15:573–577
Cooper GM, Temin H (1976) Lack of infectivity of the endogenous avian leukosis virus-related genes in the DNA of uninfected chicken cells. J Virol 17:422–430
Copeland NG, Jenkins NA, Cooper GM (1981) Integration of Rous sarcoma virus DNA during transfection. Cell 23:51–60
Cordell B, Weiss SR, Varmus HE, Bishop JM (1978) At least 104 nucleotides are transposed from the 5′ terminus of the avian sarcoma virus genome to the 5′ termini of smaller viral mRNAs. Cell 15:79–91
Crittenden LB, Smith EJ, Weiss RA, Sarma PS (1974) Host gene control of endogenous avian leukosis virus production. Virology 57:128–138
Darlix JL, Bromley PA, Spahr PF (1977) Extensive in vitro transcription of Rous sarcoma virus RNA by avian myeloblastosis virus DNA polymerase and concurrent activation of the associated RNase H. J Virol 23:659–668
DeLorbe WJ, Luciw PA, Goodman HM, Varmus HE, Bishop JM (1980) Molecular cloning and characterization of avian sarcoma virus circular DNA molecules. J Virol 36:50–61
Dhar R, McClements WL, Enquist LW, Vande Woude GF (1980) Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. Proc Natl Acad Sci USA 77:3937–3941
Dina D, Benz EM Jr (1980) Structure of murine sarcoma virus DNA replicative intermediates synthesized in vitro. J Virol 33:377–389
Donehower LA, Huang AL, Hager GL (1981) Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. J Virol 37:226–238
Donner L, Turek LP, Ruscetti SK, Fedele LA, Sherr CJ (1980) Transformation-defective mutants of feline sarcoma virus which express a product of the viral *src* gene. J Virol 35:129–140
Dunsmuir P, Brorein W, Simon M, Rubin GM (1980) Insertion of the Drosophila transposable element *copia* generates a 5 base pair duplication. Cell 21:575–579
Errede B, Cardillo TS, Sherman F, Duboi E, Deschamps J, Wiame JM (1980) Mating signals control expression of mutations resulting from insertions of transposable repetitive elements adjacent to diverse yeast genes. Cell 22:427–436
Farabaugh PJ, Fink GR (1980) Insertion of the eukaryotic transposable element TY-1 creates a 5 base pair duplication. Nature 286:352–356
Faras AJ, Dibble NA (1975) RNA-directed DNA synthesis by the DNA polymerase of Rous sarcoma virus: structural and functional identification of 4S primer RNA in uninfected cells. Proc Natl Acad Sci USA 72:859–863
Faras AJ, Garapin A, Levinson W, Bishop JM, Goodman DM (1973) Characterization of the low-molecular-weight RNAs associated with the 70s RNA of Rous sarcoma virus. J Virol 12:334–342
Flavell A, Ish-Horowicz D (1981) Extrachromosal circular copies of the eukaryotic transposable element *copia* in cultured Drosophila cells. Nature 292:591–595
Friedrich R, Moelling K (1979) Effect of viral RNase H on the avian sarcoma viral genome during early transcription in vitro. J Virol 31:630–638
Fritsch E, Temin HM (1977a) Formation and structure of infectious DNA of spleen necrosis virus. J Virol 21:119–130
Fritsch E, Temin HM (1977b) Inhibition of viral DNA synthesis in stationary chicken embryo fibroblasts infected with avian retroviruses. J Virol 24:461–469
Fuhrman SA, Van Beveren C, Verma J (1981) Identification of a RNA polymerase II initiation site in the long terminal repeat of Moloney murine leukemia viral DNA. Proc Natl Acad Sci USA 78:5411–5415
Furiuchi Y, Shatkin AJ, Stavneyer E, Bishop JM (1975) Blocked, methylated 5′ terminal sequence in avian sarcoma virus RNA. Nature 257:618–620
Gafner J, Phillipsen P (1980) The yeast transposon TY-1 generates duplications of target DNA on insertion. Nature 286:414–418

Gautsch JW, Elder JH, Schindler J, Jensen FC, Lerner RA (1978) Structural markers on core protein p30 of murine leukemia virus: functional correlation with Fv-1 tropism. Proc Natl Acad Sci USA 75:4170–4174

Gilboa E, Goff S, Shields A, Yoshimura F, Mitra S, Baltimore D (1979a) In vitro synthesis of a 9-kbp terminally redundant DNA carrying the infectivity of Moloney murine leukemia virus. Cell 16:863–874

Gilboa E, Mitra SW, Goff S, Baltimore D (1979b) A detailed model of reverse transcription and tests of crucial aspects. Cell 18:93–100

Gilmer T, Parsons JT (1979) Analysis of cellular integration sites in avian sarcoma virus-infected duck embryo cells. J Virol 32:762–770

Goodman NC, Spiegelman S (1971) Distinguishing reverse transcriptase of an RNA tumor virus from other known DNA polymerases. Proc Natl Acad Sci USA 68:2203–2206

Goubin G, Hill M (1979) Monomer and multimer covalently closed circular forms of Rous sarcoma virus DNA. J Virol 29:799–804

Groudine M, Eisenman R, Weintraub H (1981) Chromatin structure of endogenous retroviral genes and activation of an inhibitor of DNA methylation. Nature 292:311–317

Guntaka RV (1980) Synthesis of circular DNA in avian tumor virus particles. Virology 101:525–528

Guntaka RV, Richards OC, Shank PR, Kung HJ, Davidson N, Fritsch E, Bishop JM, Varmus HE (1976) Covalently closed circular DNA of avian sarcoma virus: purification from nuclei of infected quail tumor cells and measurement by electron microscopy and gel electrophoresis. J Mol Biol 106:337–357

Harbers K, Schnieke A, Stuhlmann H, Jahner D, Jaenisch R (1981) DNA methylation and gene expression: endogenous retroviral genome becomes infectious after molecular cloning. Proc Natl Acad Sci USA 78:7609–7613

Hager GL, Chang EH, Chan HW, Garon CF, Israel MA, Martin MA, Scolnick EM, Lowy DR (1979) Molecular cloning of the Harvey sarcoma virus closed circular DNA intermediates: initial structural and biological characterization. J Virol 31:795–809

Harada F, Sawyer RC, Dahlberg JE (1975) A primer ribonucleic acid for initiation of in vitro Rous sarcoma virus deoxyribonucleic acid synthesis. Nucleotide sequence and amino acid acceptor activity. J Biol Chem 250:3487–3497

Harada F, Peters GG, Dahlberg JE (1979) The primer tRNA for Moloney murine leukemia virus DNA synthesis. Nucleotide sequence and aminoacylation of tRNA. J Biol Chem 254:10979–10985

Haseltine WA, Maxam AM, Gilbert W (1977) Rous sarcoma virus genome is terminally redundant: the 5′ sequence. Proc Natl Acad Sci USA 74:989–993

Haseltine WA, Coffin JM, Hageman TC (1979) Structure of products of the Moloney murine leukemia virus endogenous DNA polymerase reaction. J Virol 30:375–383

Hayward WS (1977) Size and genetic content of viral RNAs in avian oncovirus-infected cells. J Virol 24:47–63

Hayward WS, Braverman SB, Astrin SM (1980) Transcriptional products and DNA structure of endogenous avian proviruses. Cold Spring Harbor Symp Quant Biol 44:1111–1121

Hayward WS, Neel BG, Astrin S (1981) Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290:475–480

Hill M, Hillova J (1972a) Virus recovery in chicken cells tested with Rous sarcoma cell DNA. Nature 237:35–39

Hill M, Hillova J (1972b) Recovery of the temperature-sensitive mutant of Rous sarcoma virus from chicken cells exposed to DNA extracted from hamster cells transformed by the mutant. Virology 49:309–313

Hishinuma F, DeBona PJ, Astrin S, Skalka AM (1981) Nucleotide sequence of acceptor site and termini of integrated avian endogenous provirus ev-1: integration creates a 6 bp repeat of host DNA. Cell 23:155–164

Hopkins N, Schindler J, Hynes R (1977) Six NB-tropic murine leukemia viruses derived from a B-tropic virus of BALB/c have altered p30. J Virol 21:309–318

Hsu TW, Sabran JL, Mark GE, Guntaka RV, Taylor JM (1978) Analysis of unintegrated avian RNA tumor virus double-stranded DNA intermediates. J Virol 28:810–819

Hughes SH (1982) The sequence of the LTR and adjacent segments of the endogenous virus RAV-0. J Virol 43:191–200

Hughes SH, Shank PR, Spector DH, Kung HJ, Bishop JM, Varmus HE, Vogt PK, Breitman ML (1978) Proviruses of avian sarcoma virus are terminally redundant co-extensive with unintegrated linear DNA and integrated at many sites. Cell 15:1397–1410

Hughes SH, Mutschler A, Bishop JM, Varmus HE (1981a) Rous sarcoma virus provirus is flanked by short direct repeats of a cellular DNA sequence present in only one copy prior to integration. Proc Natl Acad Sci USA 78:4299–4303

Hughes SH, Toyoshima K, Bishop JM, Varmus HE (1981b) Organization of the endogenous proviruses of chickens: implications for origin and expression. Virology 108:189–207

Hughes SH, Vogt PK, Stubblefield E, Bishop JM, Varmus HE (1981c) Integration of avian sarcoma virus DNA in chicken cells. Virology 108:208–221

Hurwitz J, Leis JP (1972) RNA-dependent DNA polymerase activity of RNA tumor viruses. I. Directing influence of DNA in the reaction. J Virol 9:116–129

Jaenisch R (1976) Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc Natl Acad Sci USA 73:1260–1264

Jaenisch R, Jahner D, Nobis P, Simon I, Lohler J, Harbers K, Grotkopp D (1981) Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell 24:519–529

Jenkins NA, Cooper GM (1980) Integration, expression, and infectivity of exogenously acquired proviruses of Rous-associated virus 0. J Virol 36:684–691

Jenkins NA, Copeland NG, Taylor BA, Lee BK (1981) Dilute (d) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MULV genome. Nature 293:370–374

Jolicoeur P, Baltimore D (1976) Effect of Fv-1 gene product on proviral DNA formation and integration in cells infected with murine leukemia viruses. Proc Natl Acad Sci USA 73:2236-2240

Jolicoeur P, Rassart E (1980) Effect of Fv-1 gene product on synthesis of linear and supercoiled viral DNA in cells infected with murine leukemia virus. J Virol 33:183–195

Jolicoeur P, Rassart E (1981) Fate of unintegrated viral DNA in Fv-1 permissive and resistant mouse cells infected with murine leukemia virus. J Virol 37:609–619

Ju G, Skalka AM (1980) Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. Cell 22:379–386

Ju G, Boone L, Skalka AM (1980) Isolation and characterization of recombinant DNA clones of avian retroviruses: Size heterogeneity and instability of the direct repeat. J Virol 33:1026–1033

Junghans RP, Duesberg PH, Knight CA (1975) In vitro synthesis of full-length DNA transcripts of Rous sarcoma virus RNA by viral DNA polymerase. Proc Natl Acad Sci USA 72:4895–4899

Kahmann K, Kamp D (1979) Nucleotide sequences of the attachment sites of bacteriophage Mu DNA. Nature 280:247–250

Keshet E, O'Rear JJ, Temin HM (1979) DNA of noninfectious and infectious integrated spleen necrosis virus (SNV) is colinear with unintegrated SNV DNA and not grossly abnormal. Cell 16:51–61

King AMO, Wells RD (1976) All intact subunit RNAs from Rous sarcoma virus contain poly(A). J Biol Chem 251:150–152

Kung HJ, Shank PR, Bishop JM, Varmus HE (1980) Identification and characterization of dimeric and trimeric circular forms of avian sarcoma virus-specific DNA. Virology 103:425–433

Kung HJ, Fung YK, Majors JE, Bishop JM, Varmus HE (1981) Synthesis of plus strands of retroviral DNA in cells infected with avian sarcoma virus and mouse mammary tumor virus. J Virol 37: 127–138

Lee F, Mulligan R, Berg P, Ringold G (1982) Glucocorticoids regulate expression of dihydrofolate reductase cDNA in mouse mammary tumor viruses chimeric plasmids. Nature 294:228–232

Leis JP, Hurwitz J (1972) RNA-dependent DNA polymerase activity of RNA tumor viruses. II. Directing influence of RNA in the reaction. J Virol 9:130–142

Lemons RS, O'Brien SJ, Sherr CJ (1977) A new genetic locus, Bevi, on human chromosome 6 which controls the replication of baboon type C virus in human cells. Cell 12:251–262

Lemons RS, Nash, WG, O'Brien SJ, Benveniste RE, Sherr CJ (1978) A gene (Bevi) on human chromosome 6 is an integration site for baboon type c DNA provirus in human cells. Cell 14:995–1005

Levinson WE, Varmus HE, Garapin A, Bishop JM (1972) DNA of Rous sarcoma virus: its nature and significance. Science 175:76–78

Levinson B, Khoury G, Vande Woude G, Gruss P (1982) Activation of SV40 genome by 72-base pair tandem repeats of Moloney sarcoma virus. Nature 295:568–572

Levis R, Dunsmuir P, Rubin GM (1980) Terminal repeats of the Drosophila transposable element *copia*: nucleotide sequence and genomic organization. Cell 21:581–588

MacHattie LA, Jackowski J (1977) In: Bukhari AI, Shapiro JD, Adhyda SC (eds) DNA insertion elements, plasmids and episomes. Cold Spring Harbor Laboratory New York, pp 219–228

Majors J, Varmus HE (1981) Nucleotide sequences at host-proviral junctions for mouse mammary tumor virus. Nature 289:253–258

Malech HL, Wivel NA (1976) Transfer of murine intracisternal A particle phenotype in chloramphenicol-resistant particles. Cell 9:383–391

Martin, GS, Radke K, Hughes S, Quintrell N, Bishop JM, Varmus HE (1979) Mutants of Rous sarcoma virus with extensive deletions of the viral genome. Virology 96:530–546

McClements W, Dhar R, Blair D, Enquist L, Oskarsson M, Vande Woude G (1981) The long terminal repeat of Moloney sarcoma provirus. Cold Spring Harbor Symp Quant Biol 45:699–705

Mellon P, Duesberg PH (1977) Subgenomic, cellular Rous sarcoma virus RNAs contain oligonucleotides from the 3′ half and the 5′ terminus of virion RNA. Nature 270:631–634

Mitra SW, Goff S, Gilboa E, Baltimore D (1979) Synthesis of a 600-nucleotide-long plus-strand DNA by virions of Moloney murine leukemia virus. Proc Natl Acad Sci USA 76:4355–4359

Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP, Chambon P (1981) The SV40 72-base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res 9:6047–6068

Neel BG, Hayward WS, Robinson HL, Fang J, Astrin SM (1981) Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete viral RNAs: oncogenesis by promoter insertion. Cell 23:323–334

Neiman PE (1972) Rous sarcoma virus nucleotide sequences in cellular DNA: measurement by RNA-DNA hybridization. Science 178:750–753

Oppermann H, Bishop JM, Varmus HE, Leventow L (1977) A joint product of the genes *gag* and *pol* of avian sarcoma virus: a possible precursor of reverse transcriptase. Cell 12:993–1005

Oskarsson M, McClements WL, Blair DG, Maizel JV, Vande Woude GF (1980) Properties of a normal mouse cell DNA sequence (sarc) homologous to the *src* sequence of Moloney sarcoma virus. Science 207:1222–1224

Payne GS, Courtneidge SA, Crittenden LB, Fadly AM, Bishop JM, Varmus HE (1981) Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell 23:311–322

Payne GS, Bishop JM, Varmus HE (1982) Multiple arrangements of viral DNA and an activated host oncogene (c-*myc*) in bursal lymphomas. Nature 295:209–213

Peters G, Dahlberg JE (1979) RNA-directed DNA synthesis in Moloney murine leukemia virus: interaction between the primer tRNA and the genome RNA. J Virol 31:398–407

Peters G, Harada F, Dahlberg JE, Panet A, Haseltine WA, Baltimore D (1977) Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. J Virol 21:1031–1041

Quintrell N, Hughes SH, Varmus HE, Bishop JM (1980) Structure of viral DNA and RNA in mammalian cells infected with avian sarcoma virus. J Mol Biol 143:363–393

Ringold GM, Shank PR, Varmus HE, Ring J, Yamamoto KR (1979) Integration and transcription of mouse mammary virus DNA in rat hepatoma cells. Proc Natl Acad Sci USA 76:665–669

Rose J, Haseltine W, Baltimore D (1976) 5′ terminus of Moloney murine leukemia virus 35S RNA is $m^7G^{5'}$ $ppp^{5'}$ GmpCp. J Virol 20:324–329

Rothenberg E, Baltimore D (1977) Increased length of DNA made by virions of murine leukemia virus at limiting magnesium ion concentration. J Virol 21:168–178

Rothenberg E, Smotkin D, Baltimore D, Weinberg RA (1977) In vitro synthesis of infectious DNA of murine leukemia virus. Nature 269:122–126

Rothenberg E, Donoghue DJ, Baltimore D (1978) Analysis of a 5′ leader sequence on murine leukemia virus 21S RNA: heteroduplex mapping with long reverse transcriptase products. Cell 13:435–451

Sabran J, Hsu T, Yeater C, Kaji A, Mason WS, Taylor J (1979) Analysis of integrated avian RNA tumor virus DNA in transformed chicken, duck and quail fibroblasts. J Virol 29:170–178

Sarkar N, Nowinski R, Moore DH (1971) Helical nucleocapsid structure of the oncogenic ribonucleic acid viruses (oncornaviruses). J Virol 8:564–572
Sawyer RC, Dahlberg JE (1973) Small RNAs of Rous sarcoma virus: characterization by two-dimensional polyacrylamide gel electrophoresis and fingerprint analysis. J Virol 12:1226–1237
Sawyer RC, Harada F, Dahlberg JE (1974) Virion-associated RNA primer for Rous sarcoma virus DNA synthesis: isolation from uninfected cells. J Virol 13:1302–1311
Schindler J, Haynes R, Hopkins N (1977) Evidence for recombination between N- and B-tropic murine leukemia viruses: analysis of three virion proteins by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J Virol 23:700–707
Schwartz DE, Zamecnik PC, Weith HL (1977) Rous sarcoma virus genome is terminally redundant: the 3′ sequence. Proc Natl Acad Sci USA 74:994–998
Shank PR, Varmus HE (1978) Virus-specific DNA in the cytoplasm of avian sarcoma virus-infected cells is a precursor to covalently closed circular viral DNA in the nucleus. J Virol 25:104–114
Shank PR, Cohen JC, Varmus HE, Yamamoto KR, Ringold GB (1978a) Mapping of linear and circular forms of mouse mammary tumor virus DNA with restriction endonucleases: Evidence for a large specific deletion occurring at high frequency during circularization. Proc Natl Acad Sci USA 75:2112–2116
Shank PR, Hughes SH, Kung HJ, Majors JE, Quintrell N, Guntaka RV, Bishop JM, Varmus HE (1978b) Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell 15:1383–1395
Shapiro J (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc Natl Acad Sci USA 76:1933–1937
Shapiro S, Strand M, August JT (1976) High molecular weight precursor polypeptides to structural proteins of Rauscher murine leukemia virus. J Mol Biol 107:459–477
Sherr CJ, Fedele LA, Donner L, Turek LP (1979) Restriction endonuclease mapping of unintegrated proviral DNA of Snyder-Theilen feline sarcoma virus: localization of sarcoma-specific sequences. J Virol 32:860–875
Shimotohno K, Temin HM (1980) No apparent nucleotide sequence specificity in cellular DNA juxtaposed to retrovirus proviruses. Proc Natl Acad Sci USA 77:7357–7361
Shimotohno K, Mizutani S, Temin HM (1980) Sequence of retrovirus provirus resembles that of bacterial transposable elements. Nature 285:550–554
Shoemaker C, Goff S, Gilboa E, Paskind M, Mitra SW, Baltimore D (1980) Structure of a cloned circular Moloney murine leukemia virus molecule containing an inverted segment: implications for retrovirus integration. Proc Natl Acad Sci USA 77:3932–3936
Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517
Steffen D, Weinberg RA (1978) The integrated genome of murine leukemia virus. Cell 15:1003–1010
Stoll E, Billeter MA, Palmenberg A, Weissmann (1977) Avian myeloblastosis virus RNA is terminally redundant: implications for the mechanism of retrovirus replication. Cell 12:57–72
Stoltzfus CM, Synder PN (1975) Structure of B77 sarcoma virus RNA: stabilization of RNA after packaging. J Virol 16:1161–1170
Sutcliffe JG, Shinnick TM, Verma IM, Lerner RA (1980) Nucleotide sequence of Moloney leukemia virus: 3′ end reveals details of replication, analogy to bacterial transposons, and an unexpected gene. Proc Natl Acad Sci USA 77:3302–3306
Sveda MM, Soeiro R (1976) Host restriction of Friend leukemia virus: synthesis and integration of the provirus. Proc Natl Acad Sci USA 73:2356–2366
Swanstrom R, Varmus HE, Bishop JM (1981a) The terminal redundancy of the retrovirus genome facilitates chain elongation by reverse transcriptase. J Biol Chem 256:115–1121
Swanstrom R, DeLorbe WJ, Bishop JM, Varmus HE (1981b) Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. Proc Natl Acad Sci USA 78:124–128
Taylor JM, Hsu TW (1980) Reverse transcription of avian sarcoma virus RNA into DNA might involve copying of the tRNA primer. J Virol 33:531–534
Taylor JM, Illmensee R (1975) Site on the RNA of an avian sarcoma virus at which primer is bound. J Virol 16:553–558
Taylor JM, Cordell-Stewart B, Rohde W, Goodman HM, Bishop JM (1975) Reassociation of 4S and 5S RNAs with the genome of avian sarcoma virus. Virology 65:248–259

Temin HM (1982) Function of the retrovirus long terminal repeat. Cell 28:3–5
Tress E, O'Donnell PV, Famulari N, Ellis RW, Fleissner E (1979) Polymorphism of B-tropic leukemia viruses from BALB/c mice: association of a p30 antigen with N- versus B-tropism. J Virol 32:350–355
Van Beveren C, Goddard JG, Berns A, Verma IM (1980) Structure of Moloney murine leukemia viral DNA: nucleotide sequence of the 5′ long terminal repeat and adjacent cellular sequence. Proc Natl Acad Sci USA 77:3307–3311
Varmus HE, Bishop JM, Vogt PK (1973a) Appearance of virus-specific DNA in mammalian cells following transformation by Rous sarcoma virus. J Mol Biol 74:613–626
Varmus HE, Vogt PK, Bishop JM (1973b) Integration of deoxyribonucleic acid specific for Rous sarcoma virus after infection of permissive and nonpermissive hosts. Proc Natl Acad Sci 70:3067–3071
Varmus HE, Guntaka RV, Fan WJW, Heasley S, Bishop JM (1974) Synthesis of viral DNA in the cytoplasm of duck embryo fibroblasts and in enucleated cells after infection by avian sarcoma virus. Proc Natl Acad Sci USA 71:3874–3878
Varmus HE, Guntaka RV, Deng CT, Bishop JM (1975) Synthesis, structure, and function of avian sarcoma virus-specific DNA in permissive and nonpermissive cells. Cold Spring Harbor Symp Quant Biol 39:987–996
Varmus HE, Padgett T, Heasley S, Simon G, Bishop JM (1977) Cellular functions are required for the synthesis and integration of avian sarcoma virus-specific DNA. Cell 11:307–319
Varmus HE, Heasley S, King HJ, Oppermann H, Smith VC, Bishop JM, Shank PR (1978) Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. J Mol Biol 120:55–82
Varmus HE, Shank PR, Hughes SE, Kung H-J, Heasley S, Majors J, Vogt PK, Bishop JM (1979) Synthesis, structure, and integration of the DNA of RNA tumor viruses. Cold Spring Harbor Symp Quant Biol 43:851–864
Varmus HE, Quintrell N, Wyke J (1981a) Revertants of an ASV-transformed rat cell line have lost the complete provirus or sustained mutations in *src*. Virology 108:28–46
Varmus HE, Quintrell N, Ortiz S (1981b) Retroviruses as mutagens: insertion and excision of a nontransforming provirus alter expression of a resident transforming provirus. Cell 25:23–36
Waters LC, Mullin BC, Bailiff EG, Popp RA (1975) tRNAs associated with the 70S RNA of avian myeloblastosis virus. J Virol 16:1608–1614
Weinberg RA (1977) Structure of the intermediates leading to the integrated provirus. Biochim Biophys Acta 473:39–55
Weiss SR, Varmus HE, Bishop JM (1977) The size and genetic composition of virus-specific RNAs in the cytoplasm of cells producing avian sarcoma-leukemia viruses. Cell 12:983–992
Wells RD, Flugel RM, Larson JE, Schendel PF, Sweet RW (1972) Comparison of some reactions catalyzed by deoxyribonucleic acid polymerase from avian myeloblastosis virus, *Escherichia coli*, and *Micrococcus luteus*. Biochemistry 11:621–629
Williamson VM, Young ET, Ciriacy M (1981) Transposable elements associated with constitutive expression of yeast alcohol dehydrogenase II. Cell 23:605–614
Yamamoto T, de Crombrugghe B, Pastan I (1980a) Identification of a functional promoter in the long terminal repeat of Rous sarcoma virus. Cell 22:787–797
Yamamoto T, Jay G, Pastan I (1980b) Unusual features in the nucleotide sequence of a cDNA clone derived from the common region of avian sarcoma virus messenger RNA. Proc Natl Acad Sci USA 77:176–180
Yang WK, Kiggans JO, Yang D-M, Ou C-Y, Tennant RW, Brown A, Bassin RH (1980) Synthesis and circularization of N- and B-tropic retroviral DNA in Fv-1 permissive and restrictive mouse cells. Proc Natl Acad Sci USA 77:2994–2998
Yoshimora FK, Weinberg RA (1979) Restriction endonuclease cleavage of linear and closed circular form. Cell 16:323–332
Yoshinaka Y, Luftig RK (1977) Murine leukemia virus morphogenesis: cleavage of p70 in vitro can be accompanied by a shift from a concentrifically coiled internal strand ("immature") to a collapsed ("mature") form of the virus core. Proc Natl Acad Sci USA 74:3446–3450

# Defective Avian Sarcoma Viruses

JAMES C. NEIL*



## Abbreviations

RSV - Rous sarcoma virus
FSV - Fujinami sarcoma virus
PRCII - PRCII sarcoma virus
PRCIV - PRCIV sarcoma virus
ESV - Esh sarcoma virus
Y73 - Y73 sarcoma virus
UR1 - UR1 sarcoma virus
UR2 - UR2 sarcoma virus
dASV - defective avian sarcoma virus
dALV - defective avian leukemia virus

* Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow, G61 1BD, Scotland

# 1 Introduction

## 1.1 Scope of the Review

This review was intended initially as a reference source for those interested in the origins and first descriptions of the defective avian sarcoma viruses. Quite a few of these viruses have been characterized in the past few years and their varied nomenclature according to source, discoverer, date of isolation or biological properties could result in some confusion among those attempting to follow the literature. Information will be included on the molecular biology of the sarcoma viruses, rather more of which is available than when the review was first conceived, although in this respect the review will inevitably be out of date by the time of publication. If any bias of content is introduced, this will be towards a more detailed coverage of the author's own area of interest, the gene products of the defective sarcoma viruses.

Rous sarcoma virus (RSV) serves as the model for much of this work and will frequently be referred to for comparative purposes, as will the mammalian defective transforming viruses. Recent reviews provide more complete coverage of these topics (1–4a). This review is complemented by a discussion of the avian acute leukaemia viruses, which appears elsewhere in this volume [5]. As will be seen, concentrating primarily on the defective sarcoma viruses and comparing them to RSV can be justified in terms of their biochemical properties as well as their similar biology. The avian sarcoma viruses offer perhaps the best available system to study the molecular basis of oncogenic transformation.

## 1.2 Nomenclature

Throughout the review, the recently characterized avian sarcoma viruses are referred to as defective sarcoma viruses or dASVs, to distinguish them from the well-characterized Rous sarcoma virus. This nomenclature is imperfect, since strains of Rous sarcoma virus have been generated which are replication-defective and since, in future, non-defective sarcoma viruses related to the dASVs might be isolated. Also, viruses classified as defective avian leukaemia viruses (dALVs) can cause sarcomas. Despite these imperfections, it is convenient to use the term dASV to refer to the viruses listed below Rous sarcoma virus in Table 2.

# 2 Isolation of the Defective Avian Sarcoma Viruses

The first representative of the defective avian sarcoma viruses, Fujinami sarcoma virus (FSV), was reported in Japan soon after Peyton Rous's discovery of RSV [6]. The original isolation was apparently made as early as 1909. Surprisingly, it took until 1980 to demonstrate that FSV carries a transformation-specific sequence different from the RSV *src* gene. One reason for this delay is clear. Some stocks of FSV were contaminated with RSV, as shown by the strong hybridization of a cDNA *src* probe with FSV stocks [7]. However, the recently characterized FSV can apparently be traced to the original isolate [8–10].

The Poultry Research Centre (PRC) isolates, PRCII and PRCIV, have also caused some confusion. The two isolates were reported at the same time, and derived from the same laboratory [11]. It remains unclear whether PRCII and PRCIV are genuinely independent isolates, and biochemical evidence that they are not is covered in some detail in later sections. Briefly, our work on PRCII was begun using virus from two sources. The first, from Dr. L.N. Payne, Houghton, England, yielded the strain which retains the name PRCII. The second, obtained from Dr. L. Crittenden, East Lansing, Michigan, yielded the virus referred to as PRCIIp [12], which was subsequently found to be very similar to PRCIV, obtained from Dr. Payne.

Esh sarcoma virus (ESV) was isolated by *Wallbank* and co-workers in 1966 [14] from a spontaneous tumour in a chicken belonging to a Pennsylvania farmer named Esh. The virus was passaged in chickens to provide "enhanced" stocks, which we obtained directly from Dr. Wallbank [15].

Y73 sarcoma virus was isolated from a transplantable tumour which was discovered, as the name implies, in 1973 by a group working in Yamaguchi University [16]. Again the source was a spontaneous tumour in a chicken. However, it was not until the 33rd passage

Table 1. Origins of the defective avain sarcoma viruses

| Virus | Place and year | Original tumour description | Helper virus subgroup | Reference |
|---|---|---|---|---|
| Fujinami sarcoma virus (FSV) | Kyoto, Japan 1909 | Fibrosarcoma forewing, metastases in liver, intestine | C | 6, 8, 9 |
| PRCII | Edinburgh, Scotland 1958 | Myxofibrosarcoma mesentery | A (+ B) | 11 |
| PRCIV | Edinburgh, Scotland 1958 | Myxofibrosarcoma ovary, mesentery | ? | 11 |
| Esh sarcoma virus (ESV) | Pennsylvania 1966 | Soft myxoid tumor thigh, ovary | A | 14 |
| Y73 | Yamaguchi, Japan 1973 | Myxofibrosarcoma site not specified | A | 16 |
| UR1 | New York 1969 | Not specified | A | 18 |
| UR2 | Connecticut 1963 | Myxofibrosarcoma pancreas, intestine; adenomas in abdominal cavity | A | 18 |
| 16L | New York 1981 | Fibrosarcoma liver, kidney, heart, peritoneum | A | 131 |

of tumour cells that sarcoma virus was demonstrated in the extracts. Most of the biochemical characterization of Y73 has been done by *Toyoshima* and co-workers, who obtained Y73 virus directly from the group who isolated it [17].

The new isolates, UR1 and UR2, were reported in 1981 by *Balduzzi* and co-workers of the University of Rochester (hence UR) [18]. The original isolations were made some time earlier but these were apparently not published. UR1 virus came from primary tumour material supplied by Dr. B.W. Calnek (Cornell University, Ithaca, NY), who obtained this tumour in 1969. UR2 came from material provided by Dr. R.E. Luginbuhl (University of Connecticut, Storrs), who had found the tumour in a chicken in 1963 and passaged washes from the tumour in 8-day-old Spafas chickens. Balduzzi and co-workers inoculated their source materials of UR1 and UR2 into chickens, and isolated viruses after co-cultivating the tumour cells with chicken embryo fibroblasts. Recently, *Neel* and co-workers reported the isolation of a new sarcoma virus, 16L, from a fibrosarcoma in a chicken. In contrast to the other isolates discussed here, which resulted from naturally occurring viral infection, this tumour appeared after deliberate inoculation of a chicken with *td* 107A, a *src* deletion mutant of the Schmidt-Ruppin strain of RSV [131]. Biochemical analysis shows that 16L contains a different transforming gene from the parental SR RSV (see Table 1).

## 3 Biological Properties of the Defective Avian Sarcoma Viruses

### 3.1 Pathogenesis

In this section, the types of tumours induced by the defective sarcoma viruses will be discussed comparing, where possible, the original tumour descriptions with more recent findings.

During the first half of the century, a number of studies were published in which the pathology of FSV infection was compared to that of RSV. These studies concluded that the oncogenic spectrum of FSV was like that of RSV, with some differences in tumour cell morphology and in the incidence of tumour regression [19–22]. Recent studies on FSV have supported these conclusions. FSV induces fibrosarcomas in young chickens after a short latent period (7–10 days). The tumours occur at the site of inoculation (wing web), and attempts to induce haematopoietic tumours by intravenous inoculation produced only sarcomas close to the inoculation site [9, 10].

PRCII induces tumours similar to those of FSV and in both cases the spindle cell morphology of the tumour cells is reflected in the in vitro transformed cells. Recent analyses of PRCII have shown it to be much less pathogenic than the original isolate [11, 12, 13]. Inoculated birds show a relatively low incidence of tumours and the regression rate is significant. There is some reason to believe that PRCII has undergone genetic change during passage since, as discussed later in this review, stocks exist which have larger genomic RNA, encode higher molecular weight transformation-specific proteins, and show an oncogenic efficiency reminiscent of the original isolate [13, 30]. Although evidence linking the putative deletion with altered pathogenicity is still circumstantial, current genetic manipulation techniques could be applied to establish the viral parameters involved in pathogenesis. The virus stock with larger genomic RNA, PRCIIp, is similar in pathogenic properties to PRCIV. Both cause a high incidence of fibrosarcomas when in-

oculated into young chicks. The observed tumours are mainly at the site of inoculation [13]. These recent observations echo the original findings of Carr and Campbell on PRCII and PRCIV. They found that metastases were rare with both viruses.

Recent studies with Esh sarcoma virus reveal properties similar to the original descriptions of *Wallbank* and co-workers [14, 15]. Wing web inoculation of ESV into 1-day-old chickens induced slowly growing fibrosarcomas at the inoculation site in 80% of the birds within 20 days [15]. By these criteria ESV is a sarcoma virus of rather low oncogenicity. However, recently developed high-titred stocks of ESV have not been tested in vivo and these might be expected to produce a more rapid and pronounced effect. By selecting the few progressing tumours, *Wallbank* was able to achieve an in vivo increase in virus titre [14].

Y73, on initial isolation, produced tumours after a short latent period (8 days), again fibrosarcomas at the inoculation site (wing web) [16]. Another group confirmed this observation and further found that on intravenous inoculation into 7-day-old chickens, the virus induced sarcomas in various organs but not detectable leukaemia [17].

*Balduzzi* and colleagues have provided the recent information on the pathogenesis of UR1 and UR2 sarcoma viruses [18]. With UR1, wing web inoculation of 4-week-old chickens produced palpable tumours as early as 8 days after inoculation. Two types of tumour were described: soft tumours with a predominance of mucoid edematous components, and hard tumours containing prominent lymphoid nodules. It will be of interest to determine whether these two tumour types reflect heterogeneity in the virus stocks, susceptibility of two different target cell populations or the result of lymphoid infiltration. UR2, in contrast, produced tumours rather later after inoculation (over 2 weeks). The cellular morphology in UR2 tumours was rather more uniform than with UR1, showing only spindle cells with none of the polygonal or round cells seen in UR1 tumours. These characteristics were noted to be essentially similar to findings with the other avian sarcoma viruses [18].

In summary, the defective sarcoma viruses induce fibrosarcomas at intramuscular inoculation sites and some induce sarcomas on internal organs if inoculated intravenously. The viruses each have distinguishing features of tumour cell morphology, incidence of metastases or efficiency of tumour induction. In only one case, however, has an attempt been made to separate the contribution of helper and defective transforming viruses to this spectrum of pathogenesis [13]. None of the in vivo studies have uncovered any capacity to induce tumours other than sarcomas, although it is probably true to say that insufficient work has been done to conclude that these viruses are unable to influence the proliferation of haemopoietic cells.

## 3.2 In vitro Transformation

All of the defective avian sarcoma viruses transform chick embryo fibroblasts in vitro. The isolation of non-producer cells transformed by defective virus without concomitant helper-virus infection has been instrumental in defining the defective viral genomes and in characterizing their replicative defects.

Foci of cells transformed by FSV, PRCII and PRCIV comprise mainly fusiform cells, reminiscent of the spindle cells seen in tumours induced by these viruses [9–13]. These three viruses fall into the same biochemical class (Table 2) and behave somewhat simi-

Table 2. Avian sarcoma viruses, their cell-derived sequence inserts and gene products

| Virus | Class | Cell-derived[a] sequences | Transformation-specific protein |
|---|---|---|---|
| Rous sarcoma virus (RSV)<br>Avian sarcoma virus B77 | I | *src* | pp60$^{src}$ |
| Fujinami sarcoma virus (FSV)<br>Avian sarcoma virus PRCII<br>Avian sarcoma virus PRCIV<br>Avian sarcoma virus UR1<br>Avian sarcoma virus 16L | II | *fps* | P140$^{gag\text{-}fps}$<br>P105$^{gag\text{-}fps}$<br>P170$^{gag\text{-}fps}$<br>P150$^{gag\text{-}fps}$<br>P142$^{gag\text{-}fps}$ |
| Avian sarcoma virus Y73<br>Esh sarcoma virus (ESV) | III | *yes* | P90$^{gag\text{-}yes}$<br>P80$^{gag\text{-}yes}$ |
| Avian sarcoma virus UR2 | IV | *ros* | P68$^{gag\text{-}ros}$ |

[a] Nomenclature for cell-derived inserts is as proposed by *Coffin* et al. (31)

larly in vitro. It is interesting that the foci induced by a fourth member of the group, UR1, are reported to contain mainly globular cells, with spindle-shaped cells only at the periphery of the foci [18]. The most interesting possibility is that modifications of the cell-derived sequences dictate these differences in focus morphology. However, until the transforming viruses are tested in pseudotypes with the same helper virus, helper contribution to focus morphology cannot be discounted.

*Lee* and colleagues report that cells transformed by FSV in mass culture, when maintained in culture for long periods, revert to normal cell morphology and can be shown to release only very low titres of FSV [9]. We have noted a similar phenomenon with PRCII, but the effect seems particularly marked with FSV, presenting a technical problem. Among possible explanations, FSV may be rather cytotoxic and the residual cells which grow out may be resistant to FSV transformation by virtue of helper virus infection. In this respect, it is interesting that some strains of Abelson murine leukaemia virus appear to have cytotoxic activity which can be lost with concomitant change in the defective genome [23].

ESV and Y73 foci are both reported to consist of either refractile spindle-shaped or round cells [15, 16] but clear differences between the foci are detectable on close examination, even when pseudotypes with the same helper (RAV-1) are used (J. Ghysdael, personal communication). UR2 foci on chicken cells are reported to be quite large and distinct, contain tight bundles of cells, and the foci often assume a triangular shape after a period of development. In this respect, UR2 foci are claimed to resemble those of fusiform morphology induced by variants of the Bryan high-titre strain of RSV [18].

From the descriptions of the foci induced in vitro, and tumours induced in vivo, it might be anticipated that the avian sarcoma viruses will overlap extensively in target cell specificity. However, no detailed analysis of these target cells has been reported to substantiate this assumption. Also, no comparative analysis of a variety of measureable parameters of transformation has been attempted. Only one study, on temperature-sensitive mutants of FSV, discusses these transformation parameters in any detail [24]. Clearly, there are large gaps in our knowledge of the newly characterized sarcoma viruses, in an

area which should provide a link with the biochemical study of sarcoma virus transformation.

## 4 Classification of the Avian Sarcoma Viruses

All of the defective avian sarcoma viruses appear to be recombinants that have acquired cellular genetic sequences at the expense of most of their replicative gene capacity. On the basis of relatedness of these cell-derived sequences, the known avian sarcoma viruses fall into four classes (Table 2). Relatedness has been measured by DNA-RNA hybridization, using cDNA specific for the cell-derived sequences and probing RNA from virions [25] or infected cells [26] to search for homology. Other criteria which have been used to establish this classification are tryptic peptide mapping of transformation-specific proteins [27–29] and RNase $T_1$ oligonucleotide fingerprinting of genomic RNA [8, 9, 10, 30].

The nomenclature presented here for the cell-derived sequences follows that suggested by a recently published proposal [31]. To discuss the viruses of one group or class, it is helpful to have some brief way of referring to them. Before the invention of the new nomenclature, we used class I, II or III for this purpose and I have included this in Table 2. Alternatively, these could be referred to as v-*src*-, v-*fps*- or v-*yes*-containing viruses, based on the proposed nomenclature. These abbreviations appear to be gaining general acceptance. Under this convention, the viral genomic cell-derived sequence is referred to as v-*onc* (*onc* as a general abbreviation for the three-letter code of the proposed system) and the cellular equivalent sequence as c-*onc*. Where no prefix is used, it is generally the viral sequence which is being discussed. The gene products are designated according to their molecular weight and superscripts refer to the sequence coding for the protein. Polyproteins are distinguished by the use of a capital P before the molecular weight. Thus, the Fujinami virus polyprotein is referred to as $P140^{gag\text{-}fps}$, while the RSV transforming protein is designated $pp60^{src}$, the additional p denoting a phosphoprotein.

An intriguing observation is the homology between the *fps* sequences of FSV and PRCII, and the analogous *fes* sequences of feline sarcoma viruses of the Snyder-Theilen and Gardner-Arnstein strains. This was demonstrated initially by hybridization analysis using the cDNA *fps* probe [25] and confirmed by tryptic peptide mapping of the viral proteins [28] and immunological studies with antisera to the *fes* protein sequences which cross-reacted with the FSV and PRCII *gag-fps* proteins [32]. Thus, it appears that retroviruses from two completely unrelated groups have picked up either the equivalent or a related sequence from the genomes of two distantly related species. Recently, *Scolnick* and colleagues have reported a similar example. Harvey and Balb murine sarcoma viruses have acquired homologous transforming genes from the mouse and rat genomes respectively [33]. The pool of cellular genes which can be acquired to generate a transforming retrovirus may thus be rather small and we may have already identified most of these.

## 5 Genome Structure

The primary objectives of studying the genomes of acute transforming retroviruses are to identify and locate the genetic sequences responsible for oncogenicity, to discover their origin, and to understand how a transforming virus can arise from a replication-compe-

tent non-transforming helper virus. All of the recently characterized sarcoma viruses are defective for replication. Non-producer transformed cells show none of the characteristic replicative gene products, and none have been reported to produce virus particles of any sort. In general, the genome structure of the defective sarcoma viruses has been found to be analogous to that of the defective avian leukaemia virus, MC29 [34]. The methods of analysis of the genomic RNA have included mapping of RNase $T_1$ oligonucleotides and heteroduplexing with cDNA to helper virus sequences. In the first method [35], poly($A^+$)-containing fragments of virion RNA are separated by size and, from the distribution of the large oligonucleotides in various RNA size class, their distance from the 3′ end and hence their 5′-3′ order can be roughly deduced. In addition, related sequences often show identical oligonucleotides, or oligonucleotides with similar base compositions [35]. By heteroduplexing, the location and extent of deletions in the genome can be measured, as can the location and size of the acquired cellular sequences.

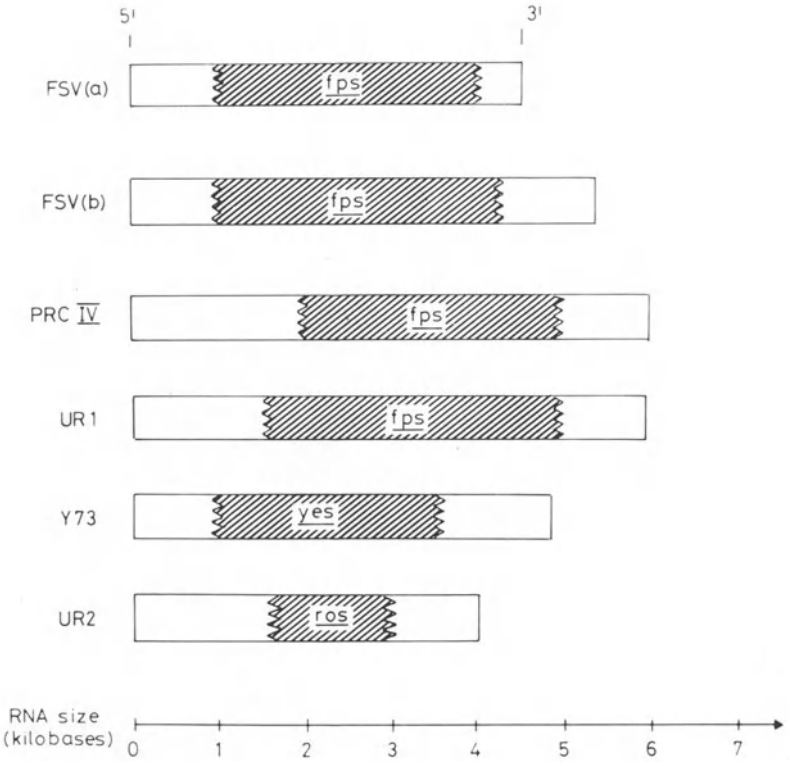



Fig. 1. Genetic structure of the defective avian sarcoma viruses. The genome sizes were estimated by gel electrophoresis or sucrose density gradient centrifugation. The positions and size of the cell-derived sequences were estimated by RNase $T_1$ oligonucleotide fingerprinting of poly($A^+$) size-selected fractions of virion RNA. The *hatched areas* represent the unique sequences, flanked by sequences related to the 5′ and 3′ ends of replication-competent helper virus. The *wavy lines* indicate the large degree of uncertainty of the position of junctions of viral and cell-derived sequences [8, 9, 10, 30, 37, 56, 59]. The two different FSV structures are from analysis of independently derived stocks [9, 10]

Both approaches have been employed to create the scheme summarized in Fig. 1. In each case the viruses appear to have retained the extreme 5′ and 3′ ends of the genome, including *gag* and possibly *env* coding sequences. This now-familiar structure seems by far the most common result of defective transforming virus generation. Other examples can be cited as exceptions to this rule; Rous sarcoma virus appears to be the only case in which a cellular transforming gene has been gained in addition to the total complement of replicative gene sequences. This process is perhaps constrained by the size of the cellular gene rescued and the capacity of the virions to package the larger RNA.

It appears that a range of recombinational events can lead to the insertion and expression of cellular genetic sequences in retroviral genomes. However, in the majority of viruses that have been analysed the 3′ end of the *gag* gene has been lost, generating a hybrid protein with residual *gag* sequences fused to the cell-coded product. This is the case for all the recently characterized defective avian sarcoma viruses.

The v-*fps* sequences of the different isolates (FSV, PRCII, PRCIV, UR1, 16L) are highly related and this is revealed by both RNase $T_1$ fingerprinting and hybridization [25, 30, 39, 56]. *Wang* and co-workers report that out of seven UR1-specific oligonucleotides, three are found in FSV and two others, while not identical, show similar base compositions to two of those of FSV [56].

Whether any virus contains the entire c-*fps* coding sequence is as yet unknown, but the various isolates do contain varying extents of *fps*. *Wong* and colleagues report that their PRCIV cDNA *fps* probe hybridized 78% to FSV and 50% to PRCII [30], while *Wang* and co-workers using an FSV *fps* probe found almost 100% hybridization to UR1 and 60% to PRCII [56]. Since Southern blot analysis suggests that the c-*fps* sequence is present as a single copy in normal chicken DNA [39], the rescue of different related *fps* loci is probably not the reason for this. Rather, the differences appear to be due to the size of the *fps* inserts in the viral genomes. This is further supported by peptide mapping data which show that the non-*gag* peptides of FSV P140$^{gag\text{-}fps}$ and PRCII P105$^{gag\text{-}fps}$ represent overlapping subsets of the non-*gag* peptides of PRCIV P170$^{gag\text{-}fps}$ [78]. In *fps* content, the deduced order would be PRCIV>>FSV>PRCII, with the position of UR1 indeterminate, but containing at least as much of *fps* as FSV. However, strains of FSV differ in *fps* content, and some may contain as much of the *fps* sequence as PRCIV (B. Mathey-Prevot and H. Hanafusa, personal communication). The detected c-*fps* protein (NCP98) has an apparent molecular weight of 98 000 on SDS-polyacrylamide gels [40], which is very close to the value predicted (96 000) from the size of the 2.6-kb *fps* insert in molecularly cloned FSV [39]. The minimal predicted size of the non-*gag* portion of PRCIIp P170$^{gag\text{-}fps}$ is slightly greater than this value at 110 000, though this molecular weight estimate may be influenced by protein phosphorylation. *Adkins* and co-workers suggest that tyrosine phosphorylation may alter the apparent molecular weight of P105$^{gag\text{-}fps}$ by as much as 5000 [127], and the *gag-fps* polyproteins are phosphorylated in vivo on tyrosine residues, in contrast to the NCP detected by *Mathey-Prevot* and colleagues [40]. In any case, the higher molecular weight regions of SDS-polyacrylamide may give rather unreliable molecular weight estimates. This question of whether any virus contains the entire c-*fps* coding sequence will have to be settled by sequence comparison with the cloned c-*fps* gene.

Less information is available for the *yes*-and *ros*-containing viruses. However, cDNA *yes* probe prepared from ESV hybridized 80% to Y73 (R. Cohen, personal communication). Peptide mapping shows the methionine tryptic peptides of ESV P80$^{gag\text{-}yes}$ and Y73

$P90^{gag\text{-}yes}$ to be virtually identical, and the analysis had to be extended to the cysteine-containing tryptic peptides to reveal significant differences [29]. This suggests that the ESV and Y73 *yes* sequences are highly related.

Models have been constructed to explain the generation of acute transforming retroviruses from non-defective helper viruses by recombination with cellular c-*onc* sequences. For a detailed consideration of these models, the recent review by *Graf* and *Stehelin* should be of interest [47]. Recombination at the mRNA level would readily explain the absence of non-coding intervening DNA sequences assumed to be present in the cellular equivalents of the retroviral transforming genes. The c-*src* sequence, for instance, contains at least seven such intervening sequences [132]. However, if a two-step model is proposed, the first step could well be at the DNA level, resulting in the fusion of a deleted provirus to the potential oncogene. The resulting hybrid RNA transcript would lack a viral 3′ end necessary for replication, but may acquire this by non-homologous recombination. This is essentially the model proposed by *Goldfarb* and *Weinberg*, based on their observations on Harvey sarcoma virus [133]. Theoretically, a transforming virus might arise which contains only those viral sequences necessary for replication and expression of the cellular sequences, but no 5′ viral coding sequences. However, most of the defective transforming viruses which have been analysed so far retain at least part of the *gag* coding sequences. It seems possible that *gag* coding sequences may play an important role in replication or gene expression. Alternatively, recombination which cleanly deletes the *gag* sequences but retains other essential functions may be feasible but rare in vivo. This question could be approached by in vitro manipulation of molecularly cloned defective transforming viruses.

## 6 Cellular Equivalents of Avian Sarcoma Virus Transforming Genes

Retroviruses with the capacity to induce cellular transformation directly in vitro and rapid tumours in vivo appear to depend for this ability on a cell-derived sequence. This cellular "gene" is characteristically a low copy number sequence which is expressed at a rather low level in uninfected cells and at a considerably higher level when incorporated in a retrovirus genome and expressed under viral control. These sequences show high evolutionary stability and are thus regarded as likely candidates for essential cellular genes involved in "maintenance" functions such as regulation of cell growth and division. This topic has been reviewed recently [48].

Some considerable knowledge has been acquired on the structure and expression of c-*src*, but only very recently has information become available for c-*fps*, c-*yes* and c-*ros*. The prototype c-*src* has been defined as a low copy number gene, the haploid genome containing one or a few copies, and the c-*src* sequence shows a high degree of evolutionary stability [7]. Thus c-*src* behaves essentially as a cellular gene with a function sufficiently vital to select for its complete retention [132]. The v-*src* and c-*src* sequences appear to be highly similar, and may be functionally identical, with variations in v-*src* occurring by more rapid evolution in the viral genome [46, 49, 50]. However, subtle differences which define v-*src*, but not c-*src*, as an oncogene cannot yet be ruled out. An experiment in which c-*src* is coupled to an efficient transcriptional promoter, and shown to transform cells after DNA transfection, would define c-*src* as having oncogenic potential. Experiments of this sort have been reported for c-*mos* and c-*ros* the cellular homologues of the transforming genes of Moloney and Harvey murine sarcoma viruses [34, 36]. However,

the fact that a c-*onc* can transform cells after transcriptional activation does not rule out that in the generation of v-*onc*, c-*onc* has been altered in a way which renders it more efficient in transformation. In particular, one could speculate on the effects of fusing a cellular protein to the *gag* gene and thereby altering its intracellular distribution as well as its level of expression.

The c-*fps*, c-*yes* and c-*ros* sequences are, like c-*src*, present in normal chicken cells at low copy number and expressed at a low level [25, 26, 39]. The four cellular sequences are similarly conserved through evolution, showing hybridization to mammalian cell DNA, though less strongly than to avian DNA.

*Shibuya* and co-workers have examined the expression of c-*src*, c-*fps*, c-*yes* and c-*ros* in different tissues [39]. The c-*yes* sequence shows by far the highest level of transcription, particularly in kidney, while c-*fps* is expressed in bone marrow and strongly repressed in most other tissues. The significance of these findings is not yet clear. The expression of the c-*onc* genes bears no obvious relation to the spectrum of pathogenicity of the equivalent v-*onc*-containing viruses.

The product of c-*src* has been characterized. Like $pp60^{v\text{-}src}$, it is a phosphoprotein of around 60 000 molecular weight with tyrosine-specific protein kinase activity [51–55]. Recently, *Hanafusa* and colleagues have characterized a cellular *fps* protein of 98 000 molecular weight (NCP98). Antisera to this protein were derived from rats with regressing tumors after inoculation of FSV-transformed rat cells [40]. The antisera had relatively weak reactivity with *gag* proteins but strongly recognized $P140^{gag\text{-}fps}$. Like its viral counterpart, NCP98 has an associated tyrosine-specific protein kinase activity capable of phosphorylating NCP98 itself and exogenous acceptor proteins.

# 7 Avian Sarcoma Virus Gene Products

## 7.1 Composition of the Defective Avian Sarcoma Virus Polyproteins

The defective avian sarcoma viruses containing v-*fps*, v-*yes* and v-*ros* sequences each encode a polyprotein that consists of an *N*-terminal portion of *gag*-related sequences and a *C*-terminal portion encoded by the acquired cellular sequences. In this respect, the viruses resemble the defective avian acute leukaemia viruses AEV and MC29 [57, 58] and the rationale for the study of the dASV polyproteins owes much to earlier studies with the dALVs.

The sarcoma virus polyproteins were readily identified, since they carry *gag* antigenic determinants, obviating the requirement for antiserum specific for the cell-derived sequences which has proved difficult to obtain in most cases. Antisera to individual *gag* proteins or simply to disrupted virions have proved suitable to precipitate the various *gag-onc* proteins. The description of the polyproteins as "transformation-specific" is justified by the demonstration of a given *gag-onc* protein in all cells transformed by that virus, regardless of whether helper virus is also present. The isolation of helper-free (non-producer) transformed cells is thus a crucial step. In some cases, further evidence has come from in vitro translation studies which show that the defective transforming viral genome encodes the polyprotein [9, 59–61].

Avian sarcoma viruses and their gene products are listed in Table 2. The composition of the *gag-onc* polyproteins has been established primarily by tryptic peptide mapping.

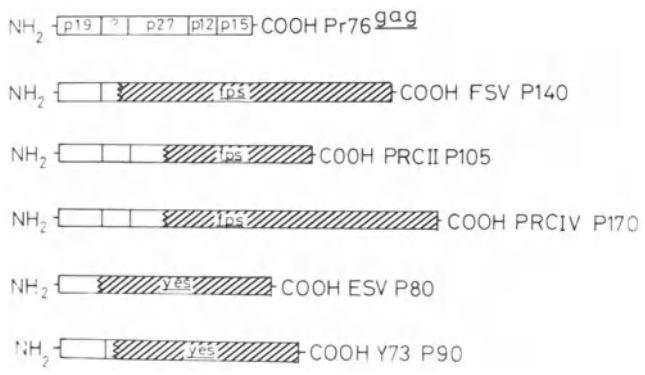



Fig. 2. Composition of the defective avian sarcoma virus polyproteins. The sizes of the polyproteins were estimated by SDS-polyacrylamide gel electrophoresis and the extent of *gag* and non-*gag* sequences estimated by tryptic peptide mapping studies [15, 27, 28, 29, 63, 67, 78] and cleavage with virion protease p15 [67, 78]. The *hatched areas* represent non-*gag* sequences, presumably encoded by the cell-derived sequences. The positions of the *gag*-non-*gag* junctions are uncertain and are therefore represented by *jagged lines*. The structure of $Pr76^{gag}$ is based on the published *gag* gene order [62], with an additional p10 coding region (?) inferred from DNA and protein sequencing studies [D. Schwarz and E. Hunter, personal communication). Virion protease p15 is capable of producing mature p27, p12 and p15 from $Pr76^{gag}$ in vitro [130]

Thus, peptides known to derive from individual *gag* proteins have been identified and the extent of the residual *gag* sequences roughly estimated. This effort has been facilitated by the recent information on the order in which the individual *gag* proteins appear on the precursor, $Pr76^{gag}$ [62]. The polyproteins also contained peptides that could not be matched with the known *gag, pol* or *env* peptides [15, 63]. These are the *fps* or *yes* peptides which show a high degree of homology when viruses of the same class are compared. The available information is summarized in the structural models of the dASV polyproteins shown in Fig. 2.

An early concern about this work was that the polyproteins might have resulted from translation of viral or cellular sequences in an aberrant reading frame, yielding a product of no functional significance for transformation. However, if this were the case, one would not expect to see conservation of this reading frame in every virus of the same class, or the conservation of a strong target site for a protein kinase. Both of these observations provide convincing evidence that the cellular sequences translated in the viral polyproteins play an important role in the transformation process.

In the case of the dALV avian erythroblastosis virus, it has been noted that another transformation-specific protein is encoded by the transforming virus, in addition to the *gag*-related polyprotein $P75^{gag\text{-}erb}$ [42–45]. The possibility remains that an analogous product might be encoded by any of the dASVs, although in a number of cases in vitro translation studies have been performed which failed to reveal any products in addition to the polyproteins [9, 59–61]. With AEV it appears that two transformation-specific proteins are translated from contiguous and non-overlapping sequences [45]. Other viruses might use more than one reading frame to encode multiple proteins. Thus, it would be a premature claim that a polyprotein spanning the entire genomic coding region must be the sole gene product of the transforming virus.

## 7.2 Structural Similarities Among the Avian Sarcoma Virus v-*onc* Proteins

All of the avian sarcoma virus v-*onc* proteins are phosphorylated in vivo on both serine and tyrosine residues. One study with Y73 virus reported that P90$^{gag\text{-}yes}$ was phosphorylated only on serine in vivo [17] but this observation is not borne out by work in other laboratories [29, 64, 65]. The discrepancy may result from relative heat lability of the tyrosyl phosphate bond or perhaps some feature of the labelling conditions used by Yoshida and co-workers.

Surprisingly, tryptic peptides containing the tyrosine phosphorylation sites of pp60$^{src}$ and the *gag-yes* polyproteins (P80, P90) proved to be indistinguishable [64, 65]. Furthermore, the tyrosine phosphorylation site of the *gag-fps* polyproteins shares some features [64, 66, 67], the most prominent of which are the acidic amino acid residues (glu) proximal to the target tyrosine. Projecting from these albeit limited observations, the association of acidic residues and target tyrosine might be a general feature of tyrosine phosphorylation sites. As a precedent, cyclic-AMP-dependent protein kinases consistently recognize a configuration of basic amino acids adjacent to the target serine [68]. Interestingly, a casein kinase has been identified which apparently requires acidic residues on the *N*-terminal side of the target serine. This enzyme is known as casein kinase 2 or glycogen synthase kinase 5 [133, *M. Pinna*, personal communication). However, there have been few opportunities to analyse tyrosine phosphorylation sites. Of the limited numbers of viral or cellular phosphotyrosine-containing proteins, most have not yielded enough labelled material for analysis. Of the examples available, polyoma middle T antigen can be phosphorylated on tyrosine in vitro and it appears that the major phosphorylation site contains the sequence $(glu)_6$-tyr [69]. Since the *fps* and *fes* sequences show homology [25], another obvious comparison to make is with the FeSV *gag-fes* polyproteins which are phosphorylated on tyrosine both in vitro and in vivo. Intriguingly, DNA sequence studies indicate a potential P85$^{gag\text{-}fes}$ tyrosine phosphorylation which contains the sequence shown in

Table 3. Primary sequence around tyrosine phosphorylation sites

| Sequence | Protein | Reference |
|---|---|---|
| -leu-ile-glu-asn-asp-glu-tyr- | pp60$^{src}$<br>P80$^{gag\text{-}yes}$[a]<br>P90$^{gag\text{-}yes}$[a] | 64, 66, 103, 124 |
| -gln-glu-glu-asp-gly-val-tyr- | P105$^{gag\text{-}fps}$<br>P170$^{gag\text{-}fps}$<br>P140$^{gag\text{-}fps}$ | 66, 77, 102 |
| -glu-glu-glu-glu-glu-glu-tyr- | Polyoma middle T | 69 |
| -glu-glu-ala-asp-gly-val-tyr- | P85$^{gag\text{-}fes}$ | 101 |

All of the phosphotyrosine-containing peptides listed above, apart from that of polyoma middle T, begin at a tryptic cleavage site. Thus, in all these in vivo tyrosine phosphorylation sites, the target is seven residues from a tryptic cleavage site.

[a] The sequence around the major tyrosine phosphorylation sites of the *gag-yes* polyproteins is probably identical to that in pp60$^{src}$, based on the comigration of the tryptic peptides and a series of further cleavage products (64, 65, 66, 67) as well as DNA sequence data (103)

Table 3. This sequence contains glutamic acid residues five and six residues from the tyrosine and, like pp60$^{src}$, aspartic acid three residues away.

Among the cellular proteins which have been analysed is a 36K phosphoprotein which was identified initially by *Radke* and *Martin* [70] as a novel phosphoprotein in RSV-transformed cells and shown to act in vitro as a target for the pp60$^{src}$ kinase [71]. By staphylococcal V8 protease cleavage analysis, the closest glutamic acid appears to be ten residues away from the target tyrosine [66; *T. Gilmore*, unpublished). However, this approach to sequence analysis is restricted by the specificity of V8 protease. Some glutamyl bonds may be resistant to cleavage, and if aspartic acid is substituted, V8 resistance will also be encountered [72]. Another cellular protein, 50K, which associates with pp60$^{src}$ in immune complexes [73, 75] and has recently been found to associate with P105$^{gag\text{-}fps}$ in a similar fashion [123, 125], is also phosphorylated on tyrosine in ASV-transformed cells. In this case glutamic acid, located by V8 cleavage, occurs three residues away from the target tyrosine (T. Gilmore, unpublished).

Clearly, this question requires further investigation. Direct sequencing of DNA or protein will supersede the preliminary protease cleavage results. Further understanding of target site specificity may also be gleaned from studies with artificially synthesized peptides. *Wong* and *Goldberg* [41] have found that a synthetic peptide constructed using the pp60$^{src}$ phosphorylation site sequence (Table 3) inhibits the pp60$^{src}$ kinase. Also, antisera raised to this peptide coupled to a carrier protein cross-react with Y73 P90$^{gag\text{-}yes}$, as might be expected from the observed homology in this region [64].

It is instructive to reflect on the apparent conflicts generated by different methods applied to the measurement of relatedness between the *src*, *fps* and *yes* sequences.

The homology of the tyrosine phosphorylation sites of the v-*src*, v-*yes* and v-*fps* products was in sharp contrast to their lack of cross-hybridization using nucleic acid probes [25, 26]. However, when the overall structural features of the molecules were compared, such as the location of phosphorylation and protease cleavage sites, further similarities could be discerned. Topological maps of pp60$^{src}$, P105$^{gag\text{-}fps}$ and P80$^{gag\text{-}yes}$ are presented in Fig. 3 for comparison. The maps were constructed by orienting partial V8 digestion fragments. For pp60$^{src}$, in vitro translation in the presence of *N*-formyl[$^{35}$S] met tRNA allowed identification of the *N*-terminal V8 fragment [77], while the *gag-onc* polyproteins were dissected by cleavage with V8 protease after p15 or carboxypeptidase digestion [67]. Virion protease p15 cleaves the *gag* precursor Pr76 and can remove part or all of the *gag* sequences from the dASV polyproteins [78]. The positions of the tyrosine and serine phosphorylation sites in pp60$^{src}$ were deduced from protease cleavage [77, 79] and sequencing data [80, 124].

Some striking similarities emerged. The tyrosine phosphorylation sites are in the *C*-terminal portions of the proteins, and in the case of the *src* and *yes* products, are on rather V8-resistant fragments of similar size. A further similarity between *src* and *yes* is in the strong serine phosphorylation site in the *N*-terminal portion of the v-*onc* product. The *gag-fps* polyproteins are also phosphorylated on serine residues and some of the sites are probably in the non-*gag* sequence [128]. However, these have not been located on PRCII P105 and hence are not shown in Fig. 3.

Recent DNA sequence studies on molecularly cloned Y73, Fujinami and feline sarcoma viruses confirm the relationships suggested by the protein structural studies. The *src* and *yes* sequences code for proteins of largely similar primary structure. The failure too see relatedness by hybridisation was apparently due to frequent differences

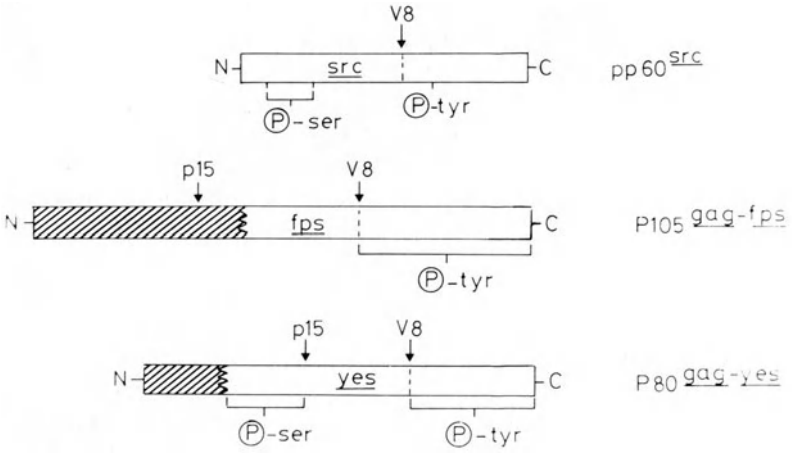



Fig. 3. Structural similarities of avian sarcoma virus gene products. Cleavage sites represented are for virion protease p15 [78] and with staphylococcal protease V8 [67, 77]. The position of the tyrosine phosphorylation site of pp60$^{src}$ has been determined by DNA and protein sequencing [80, 124]. The position of the serine phosphorylation is as yet uncertain. The strong serine phosphorylation site of P80$^{gag\text{-}yes}$ is in an *N*-terminal p15 fragment [78], and is most likely located in the *yes*-coded sequences. The *hatched areas* at the *N*-terminus of the polyprotein represent *gag* gene sequences. The exact locations of the *gag*-non-*gag* junctions are not certain

in third base positions [103]. A slightly less close but nevertheless striking homology was seen between *src* and *fps* [102] and it now seems dear that *fps* and *fes* are the avian and feline representatives of the same conserved cellular gene [101, 102].

In view of the similarities at the level of gross topology (Fig. 3) and primary sequence (Table 3), it would not be surprising if the viral transforming proteins were functionally similar. The *C*-terminal phosphorylation sites of the *fps, yes* and *src* products reflects apparent conservation of the 3′ portions of their coding sequences. Protein kinase function of pp60$^{src}$ appears to reside in this region [79, 82]. It is of particular interest in terms of gene evolution that the relatedness of *src*, *yes*, *fps* and *mos* occurs in limited areas of the coding sequences.

## 7.3 Protein Kinases and the Avian Sarcoma Virus v-*onc* Proteins

This area of work was initiated by the discovery of protein kinase activity associated with pp60$^{src}$ which catalysed the transfer of phosphate to immunoglobulin heavy chain [83, 84], and studies with mutants which implicated the activity in transformation [84–87]. The enzyme activity was later shown to have a novel specificity for tyrosine residues [73, 87]. While protein purification results from different laboratories agree that purified pp60$^{src}$ retains tyrosine-specific kinase activity using a variety of substrates, they differ on the question of whether pp60$^{src}$ is capable of autophosphorylation [88, 89, 126, 129]. This is an important question, since it may be expected that this protein phosphorylation modulates activity. Protein phosphorylation appears to play a major role in regulating the enzymes of central metabolic pathways and in mediating the effects of a variety of hormones [68].

At this stage, a description of the basic in vitro assay system will be helpful. Usually,

the v-*onc* proteins (pp60$^{src}$ or the *gag-onc* polyproteins) are immune precipitated with RSV tumour-bearing rabbit (TBR) sera or anti-*gag* sera. The immune complexes are collected by virtue of their affinity for heat-fixed protein-A-bearing staphylococci, and to a small volume of this suspension are added appropriate cations ($Mg^{2+}$, $Mn^{2+}$) and [$\gamma$-$^{32p}$]-labelled ATP. After a short incubation, the reaction is terminated and the products are analysed on SDS-polyacrylamide gels. This is an outline of the basic method which proved successful for Erikson and colleagues and has been adopted by those investigating the dASVs. Exogenous acceptor proteins can also be introduced to the immune complex suspension before ATP is added.

The protein kinase activity of pp60$^{src}$ was the first clear indication of a gene function which could potentially explain the many manifestations of the transformed state. It was therefore of great interest that other viruses, including the dASVs, encoded phosphoproteins with associated protein kinase activity. The activities detected were in these cases directed towards the polyproteins themselves, but phosphorylation of immunoglobulin heavy chain and endogenous acceptor proteins was also reported [15, 17, 38, 90–92]. Since no purification studies on the *gag-onc* polyproteins have yet been published, it is not possible to say whether the protein kinase activities are intrinsic to the polyproteins or catalysed by associated enzymes. Indeed, some circumstantial evidence suggests that at least the polyprotein phosphorylation is catalysed by another protein. Phosphorylation of FSV P140$^{gag\text{-}fps}$ was reported to be abolished by a pre-incubation with normal serum [93] and we have found that detergents or heat rapidly reduce the associated activity with a number of dASVs [78, 92]. However, inactivation of an intrinsic kinase activity cannot be ruled out as an explanation. The in vitro polyprotein phosphorylating activity labels the major in vivo tyrosine phosphorylation site, suggesting that the in vitro reaction is catalysed by the same enzyme that acts in vivo and that the substrate specificity in this respect is unaltered by the artificial conditions [15, 29, 90–92]. This property, shared by all the dASVs, contrasts with Abelson murine leukaemia virus P120$^{gag\text{-}abl}$, where the in vitro kinase activity labels a number of sites, none of which appears to be an in vivo phosphorylation site [125].

In the light of their studies on temperature-sensitive mutants of FSV, *Lee* and colleagues [94] consider two models to explain their data. The first model proposes that P140$^{gag\text{-}fps}$ is a protein kinase capable of autophosphorylation, while the second model assumes that P140 is not a kinase, but merely a kinase substrate capable of indirectly influencing cellular phosphorylation. They fit their data best to the second model. However, they do not consider a plausible third model which fits their data equally well. P140$^{gag\text{-}fps}$ might have an intrinsic protein kinase activity but may in addition act as a substrate for a distinct kinase. Both features may be required for transformation. This explanation demands consideration particularly in view of evidence that pp60$^{src}$ acts in this fashion [89].

Immune complexes containing the dASV polyproteins also show protein kinase activity for exogenous acceptor proteins, such as casein [90]. Immunoglobulin heavy chain can be phosphorylated on tyrosine whether it is added as an exogenous acceptor or used to precipitate the gag-*onc* proteins by virtue of its anti-*gag* activity. There are some problems with this area of work, however. We and others have experienced some difficulty in absorbing the anti-*gag* activity of the TBR sera with virion proteins and thereby blocking the in vitro phosphorylation. This has not been satisfactorily explained, and we have some preliminary data suggesting that in the case of ESV P80$^{gag\text{-}yes}$ the immuno-

Table 4. Some properties of protein kinase activities associated with avian sarcoma virus gene products. Modified from Feldman et al. (38)

| Gene product | pp60$^{src}$ | P140$^{gag\text{-}fps}$ P105$^{gag\text{-}fps}$ | P90$^{gag\text{-}yes}$ P80$^{gag\text{-}yes}$ | P68$^{gag\text{-}ros}$ |
|---|---|---|---|---|
| cAMP dependence | Independent | Independent | Independent | Independent |
| Phosphate donor | ATP, GTP | ATP | ATP, GTP | ATP |
| Phosphate acceptor | Tyrosine | Tyrosine | Tyrosine | Tyrosine |
| Substrates for immune complex phosphotransfer[a] | IgG, pp60$^{src}$ | IgG, polyprotein, casein | IgG, polyprotein, casein | IgG, polyprotein, casein |
| Cation requirement[b] | Mg, Mn | Mn > Mg | Mn > Mg | Mn > Mg |
| pH optimum | 7.0–8.0 | 6.5–7.5 | 6.5–8.5 | 6.5–8.5 |

[a] Substrates phosphorylated depend on the serum used in the immune complex reaction; [b] The results of Ghysdael and co-workers differ on some aspects of cation dependence (see text)

globulin phosphorylating activity may be due to the presence of pp60$^{c\text{-}src}$ in the complexes (J. Ghysdael, unpublished). This problem may be simplified by molecular cloning of the dASV genomes and expression of the *gag-onc* proteins in prokaryotic cells, where other mammalian cell enzymes can be excluded. This has recently been achieved for pp60$^{src}$ and we can await with interest an answer to the question of pp60$^{src}$ autophosphorylation [95, 129]. It is clear that all of the detected *gag-onc* phosphorylation reactions will have to be re-examined using purified material.

Table 4 summarizes some of the properties of the pp60$^{src}$ and *gag-onc* protein-associated kinases. This table is based on that published by *Feldman* and co-workers [38]. One discrepancy between these and our data is in the question of cation dependence of the in vitro reaction. If the kinetics of the reactions are taken into account, pp60$^{src}$ phosphorylation does show a preference for $Mn^{2+}$ over $Mg^{2+}$. Furthermore, phosphorylation of immunoglobulin heavy chain in P105$^{gag\text{-}fps}$ precipitates shows similar cation dependence to the pp60$^{src}$-Ig reaction, a quite different pattern from the P105$^{gag\text{-}fps}$ phosphorylation reaction (J. Ghysdael, unpublished). Thus, cation dependence should be determined separately for each of the reactions listed in Table 4, as should optimum pH and other properties.

## 8 Tyrosine Phosphorylation and Avian Sarcoma Virus Transformation

Work with the avian and feline sarcoma viruses and Abelson murine leukaemia virus shows that transforming viruses that have acquired five different cellular sequences influence cellular tyrosine phosphorylation. While it must be acknowledged that a still greater number of transforming viruses have not been shown to affect tyrosine phosphorylation, it is still a compelling possibility that a number may act by a similar mechanism. This possibility remains intact despite the recent unfortunate events at Cornell University, where evidence for the sought-after common final pathway was first expounded in detail and then retracted with no clear indication of what residue of fact remained [96–99].

Cell transformation studies with RSV, of much longer standing, have pointed the way for the dASVs. As discussed in the previous section, pp60$^{src}$ is a tyrosine-specific

thermolabile, as did classical transformation parameters such as colony formation in soft agar, increased hexose uptake, and production of plasminogen activator. The incidence of *ts* behaviour in the virus fraction surviving after mutagenization (1%) was assessed, and 11 out of 116 colonies proved to be *ts* for transformation.

*Lee* and co-workers reported isolation of a temperature-resistant (*tr*) variant of the L5 isolate of FSV from a chicken tumour. They also isolated less leaky *ts* mutants after 5-azacytidine mutagenization. In agreement with the other groups, they found that thermosensitivity of transformation correlated with thermolabile phosphorylation of $P140^{gag\text{-}fps}$ [94].

PRCII temperature-sensitive mutants have also been isolated from stocks mutagenized with 5-azacytidine [115]. One of the three mutants characterized, *ts* LA46, is unusual in that intracellular genomic RNA levels drop dramatically at the restrictive temperature, along with yields of infectious virus [116].

## 9.2 Deletion Mutants

For the defective sarcoma viruses, no deletion in the cell-derived sequences has yet been shown unequivocally to affect transformation. However, an abundance of variants suggests that mutants of this sort will not be difficult to obtain. The *fps*-containing viruses show great variation in size of genome and gene products. In this respect they resemble Abelson murine leukaemia virus, of which a larger number of variants have been isolated, with widely varying genome composition and transformation properties [123–125]. *Hanafusa* and colleagues isolated their temperature-sensitive mutant from a wild-type virus which had already undergone change in that it encoded, instead of P140, a $P130^{gag\text{-}fps}$. Whether this change was due to a deletion in the genome or to a substitution leading to premature termination has not been analysed, and no effect on transformation was noted [24]. It is also clear that the FSV strain studied by Hanafusa and colleagues differs from that analysed by Duesberg's group. Apart from differences in thermosensitivity of transformation, there are clear differences in oligonucleotide composition [9, 10] and in polyprotein composition, since p27 peptides were detected by *Pawson* [128] but not by us in studies on $P140^{gag\text{-}fps}$ of FSV obtained from *Hanafusa* [27, 78]. PRCII *ts* LA47 shows a very slightly smaller $P105^{gag\text{-}fps}$ protein on SDS-polyacrylamide gels (A. Hirano, personal communication) although again this change cannot yet be related to the *ts* behaviour or to any specific genomic alteration.

Finally, PRCII might be considered a deletion mutant of PRCIIp for the following reasons. First, PRCII has a smaller polyprotein and a smaller genomic RNA than that of PRCIIp [13, 60]. Second, the methionine-containing tryptic peptides of PRCII P105 are a subset of those of PRCIIp $P170^{gag\text{-}fps}$, apart from one, which might be generated by the putative deletion [78]. Third, V8 cleavage analysis suggests that the *gag*-non-*gag* sequence junction may be identical in PRCII $P105^{gag\text{-}fps}$ and PRCIIp $P170^{gag\text{-}fps}$ [67]. Furthermore, PRCIIp and PRCIV are closely related, if not identical, in their cell-derived *fps* sequences. RNase $T_1$ oligonucleotide fingerprinting reveals very few differences and these map in the helper-related regions of the genomes (T.C. Wong and M. Lai, personal communication). The $P170^{gag\text{-}fps}$ proteins of PRCIIp and PRCIV co-migrate on SDS-polyacrylamide gels and have identical tryptic peptide maps [13]. Thus, it appears most likely that Carr and Campbell isolated only one virus and that subsequent passage of one

Table 5. Tyrosine phosphoproteins in cells transformed by avian sarcoma viruses

| Protein | Molecular weight ($X10^3$) | Increased tyrosine phosphorylation in cells transformed by viruses containing: | | | Reference |
|---|---|---|---|---|---|
| | | *src* | *fps* | *yes* | |
| 36K | 34–39 | + | + | + | 70, 91, 107, 108 |
| 50K | 50 | + | + | + | 73, 75, 121, 123 |
| Vinculin | 130 | + | − | + | 104 |
| k, l, m, n, o, q[a] | 46, 46, 43, 43, 39, 28 | + | + | + | 105, 108 |

[a] Protein p described by Cooper and Hunter is probably the 36K protein (105)

Identifying substrates seems a straightforward task when compared to the problem of establishing a causal relationship between phosphorylation of given substrates and phenotypic features of transformation. *Weber* and colleagues, using RSV mutants, have attempted to find correlations suggesting such causal relationships. This approach has so far revealed an association between 36K phosphorylation and production of plasminogen activator [109]. Perhaps application of the same approach with *fps*-, *yes*- and *ros*-containing viruses will be fruitful.

A further question in this area is how sarcoma virus transformation and the action of growth hormones, such as EGF, might be related. Such a link was first suggested by work with murine and feline sarcoma viruses, when it was found that transformation by these viruses blocked EGF binding [110]. More recently, it has emerged that EGF receptor itself is associated with a protein kinase activity [111] which phosphorylates the receptor itself and is specific for tyrosine [112, 113]. Notably, EGF stimulates tyrosine phosphorylation of the 36K protein in mouse (A431) cells which express an unusually high level of EGF receptors, and produces a temporary increase in the relative level of intracellular phosphotyrosine [114]. It is thus possible that sarcoma virus transformation and EGF-stimulated growth share some common functional aspects, and furthermore that the c-*src* and other c-*onc* products may form part of a pathway in which cell growth and division are regulated by protein phosphorylation. Very recently, it has been reported that platelet-derived growth factor (PDGF) has a similar effect to EGF on cellular tyrosine phosphorylation [76].

## 9 Mutants of the Defective Avian Sarcoma Viruses

### 9.1 Temperature-Sensitive Mutants

The first demonstration of thermolabile dASV transformation was with FSV. The L5 strain of FSV studied by *Martin* and co-workers fortuitously turned out to be temperature-sensitive for transformation. This *ts* behaviour correlated with thermolabile tyrosine phosphorylation of $P140^{gag\text{-}fps}$ and of cellular proteins including 36K [91].

*Hanafusa* and colleagues have also studied a series of *ts* mutants of FSV isolated after mutagenization with 5-azacytidine [24]. Again, polyprotein phosphorylation proved

thermolabile, as did classical transformation parameters such as colony formation in soft agar, increased hexose uptake, and production of plasminogen activator. The incidence of *ts* behaviour in the virus fraction surviving after mutagenization (1%) was assessed, and 11 out of 116 colonies proved to be *ts* for transformation.

*Lee* and co-workers reported isolation of a temperature-resistant (*tr*) variant of the L5 isolate of FSV from a chicken tumour. They also isolated less leaky *ts* mutants after 5-azacytidine mutagenization. In agreement with the other groups, they found that thermosensitivity of transformation correlated with thermolabile phosphorylation of $P140^{gag\text{-}fps}$ [94].

PRCII temperature-sensitive mutants have also been isolated from stocks mutagenized with 5-azacytidine [115]. One of the three mutants characterized, *ts* LA46, is unusual in that intracellular genomic RNA levels drop dramatically at the restrictive temperature, along with yields of infectious virus [116].

## 9.2 Deletion Mutants

For the defective sarcoma viruses, no deletion in the cell-derived sequences has yet been shown unequivocally to affect transformation. However, an abundance of variants suggests that mutants of this sort will not be difficult to obtain. The *fps*-containing viruses show great variation in size of genome and gene products. In this respect they resemble Abelson murine leukaemia virus, of which a larger number of variants have been isolated, with widely varying genome composition and transformation properties [123–125]. *Hanafusa* and colleagues isolated their temperature-sensitive mutant from a wild-type virus which had already undergone change in that it encoded, instead of P140, a $P130^{gag\text{-}fps}$. Whether this change was due to a deletion in the genome or to a substitution leading to premature termination has not been analysed, and no effect on transformation was noted [24]. It is also clear that the FSV strain studied by Hanafusa and colleagues differs from that analysed by Duesberg's group. Apart from differences in thermosensitivity of transformation, there are clear differences in oligonucleotide composition [9, 10] and in polyprotein composition, since p27 peptides were detected by *Pawson* [128] but not by us in studies on $P140^{gag\text{-}fps}$ of FSV obtained from *Hanafusa* [27, 78]. PRCII *ts* LA47 shows a very slightly smaller $P105^{gag\text{-}fps}$ protein on SDS-polyacrylamide gels (A. Hirano, personal communication) although again this change cannot yet be related to the *ts* behaviour or to any specific genomic alteration.

Finally, PRCII might be considered a deletion mutant of PRCIIp for the following reasons. First, PRCII has a smaller polyprotein and a smaller genomic RNA than that of PRCIIp [13, 60]. Second, the methionine-containing tryptic peptides of PRCII P105 are a subset of those of PRCIIp $P170^{gag\text{-}fps}$, apart from one, which might be generated by the putative deletion [78]. Third, V8 cleavage analysis suggests that the *gag*-non-*gag* sequence junction may be identical in PRCII $P105^{gag\text{-}fps}$ and PRCIIp $P170^{gag\text{-}fps}$ [67]. Furthermore, PRCIIp and PRCIV are closely related, if not identical, in their cell-derived *fps* sequences. RNase $T_1$ oligonucleotide fingerprinting reveals very few differences and these map in the helper-related regions of the genomes (T.C. Wong and M. Lai, personal communication). The $P170^{gag\text{-}fps}$ proteins of PRCIIp and PRCIV co-migrate on SDS-polyacrylamide gels and have identical tryptic peptide maps [13]. Thus, it appears most likely that Carr and Campbell isolated only one virus and that subsequent passage of one

stock (PRCIIp) has led to the generation of a deletion mutant, PRCII, with a smaller genome and a smaller polyprotein. Since two tumours were being handled in the same laboratory at the same time, it would seem not unlikely that virus from one tumour could accidentally contaminate the other. The alternative view, that two such highly similar viruses were isolated simultaneously, would seem much less attractive. DNA sequence information will probably be required to settle this issue beyond doubt. If PRCII is indeed a deletion mutant of PRCIIp, their differences in pathogenicity and in vitro transforming properties [13] suggest that a thorough comparison of their effects on cellular phosphorylation may be of value.

## 10 Prospects

Despite the recent acceleration in the rate of identification of cell-derived transforming genes acquired by retroviruses, it must be recognized that the same sequences are being isolated repeatedly. The *fps* sequence, for instance, has been isolated on at least four occasions, and if the related *fes* represents the feline equivalent of the *fps* gene, then on six occasions. It seems, therefore, that to continue to attempt new isolations will be an exercise with diminishing returns. New transforming genes may be identified, but this should become less likely after each new gene is isolated. It is possible, of course, that only a subset of cellular transforming genes can be acquired by retroviruses. A comparison of the retroviral sequence inserts with transforming genes identified by other approaches, such as transfection with tumor DNAs [117–120], should be instructive.

Sequence comparisons of *src, fps, yes* and *ros* should be of considerable interest. Evidence presented in this review suggests that at least *src, fps* and *yes* may be structurally and functionally related, and very recent sequencing studies on the Y73 *yes* and FSV *fps* sequences bear out this suggestion, showing a remarkable sequence overlap with *src* despite the previously observed lack of hybridization [102, 103]. Thus, the c-*src*, c-*fps* and c-*yes* sequences may well represent a family of genes, related in function and derived from a common ancestor. In this respect, one might ask whether the cellular homologues are genetically linked and/or co-ordinately controlled in expression or function.

Although the avian sarcoma viruses with *src, fps, yes* and *ros* sequences all influence cellular tyrosine phosphorylation, and $pp60^{src}$ has been found to have protein kinase activity [88, 89], it has not yet been shown that the *gag-fps, gag-yes* and *gag-ros* polyproteins have intrinsic protein kinase activity. Protein purification or molecular cloning and expression in prokaryotic cells represent two possible approaches to the problem. The latter has recently been applied to $pp60^{src}$ [95, 129].

The study of avian sarcoma viruses, in particular the identification of newly phosphorylated proteins in transformed cells, has already provided some insight into the molecular basis of transformation. This area promises still greater rewards, particularly if conditional transformation mutants of *fps, yes* and *ros* are added to the available tools for study.

*Acknowledgments.* During the early stages of writing this review, I was supported by a Lievre Fellowship from the American Cancer Society. Among the numerous people who offered helpful advice and criticism of the manuscript, I would particularly like to thank Drs. J. DeLamarter, M. Hayman, M. Koury, P. Vogt, N. Wilkie and J. Wyke. I am also grateful to Drs. H. Hanafusa, J. Ghysdael, B. Neel and B. Sefton for discussion and contribution of their unpublished results.

## References


1. Vogt PK (1977) In: Comprehensive Virology 9, pp 341–430
2. Bishop JM (1978) Ann Rev Biochem 47:35–88
3. Bishop JM, Varmus HE (1982) In: RNA Tumor Viruses, Cold Spring Harbor, pp 999–1108
4. Anderson P (1980) Cancer Res 33:109–172

4a. Rosenberg N, Baltimore D In: Klein G (ed) Oncology. Raven, New York, pp 187–205

5. Hayman MJ, Enrietto P (1982) Curr Top Microb Immunol
6. Fujinami A, Inamoto K (1914) Z Krebsforsch 14:94–119
7. Stehelin D, Guntaka RV, Varmus HE, Bishop JM (1976) J Mol Biol 101:349–365
8. Hanafusa H, Wang L-H, Hanafusa T, Anderson SM, Karess RF, Hayward WS (1980) In: Fields B, Jaenisch R, Fox CF (eds) Animal Virus Genetics (ICN-UCLA) Academic Press
9. Lee W-H, Bister K, Pawson A, Robins T, Moscovici C, Duesberg PH (1980) Proc Natl Acad Sci USA 77:2018–2022
10. Hanafusa T, Wang L-H, Anderson SM, Karess RE, Hayward WS, Hanafusa H (1980) Proc Natl Acad Sci USA 77:3009–3013
11. Carr JG, Campbell JG (1958) Br J Cancer 12:631–635
12. Breitman ML, Neil JC, Moscovici C, Vogt PK (1981) Virology 108:1–12
13. Breitman ML, Hirano A, Wong TC, Vogt PK (1981) Virology 114:451–462
14. Wallbank AM, Sperling FG, Hubben K, Stubbs EL (1966) Nature 209:1265
15. Ghysdael J, Neil JC, Wallbank AM, Vogt PK (1981) Virology 111:386–400
16. Itohara S, Hirata K, Inove M, Hatsuoka M, Sato A (1978) Gan 69:825–830
17. Kawai S, Yoshida M, Segawa K, Sugiyama R, Ishizaki R, Toyoshima K (1980) Proc Natl Acad Sci USA 77:6199–6203
18. Balduzzi PC, Notter MFD, Morgan HR, Shibuya M (1981) J Virol 40:268–275
19. Fujinami A, Hatano S (1929) Gan 23:67–75
20. Gye WE (1931) Br J Exp Pathol 12:93–97
21. Duran-Reynals F (1940) Yale J Biol Med 13:77–85
22. Smida J, Thurzo V, Smidova V (1968) Neoplasma 15:329
23. Ziegler SF, Whitlock CA, Goff SA, Gifford A, Witte ON (1981) Cell 27:477–486
24. Hanafusa T, Mathey-Prevot B, Feldman RA, Hanufusa H (1981) J Virol 38:347–355
25. Shibuya M, Hanafusa T, Hanafusa H, Stephenson JR (1980) Proc Natl Acad Sci USA 77: 6536–6540
26. Yoshida M, Kawai S, Yoyoshima K (1980) Nature 287:653–654
27. Neil JC, DeLamarter JF, Vogt PK (1981) Proc Natl Acad Sci USA 78:1906–1910
28. Beemon K (1981) Cell 24:145–153
29. Ghysdael J, Neil JC, Vogt PK (1981) Proc Natl Acad Sci USA 71:2611–2615
30. Wong TC, Hirano A, Lai MMC, Vogt PK (1982) Virology 117:156–164
31. Coffin JM, Varmus HE, Bishop JM, Essex M, Hardy WD Jr, Martin GS, Rosenberg NE, Scolnick EM, Weinberg RA, Vogt PK (1982) J Virol 40:953–957
32. Barbacid M, Breitman ML, Lauver AV, Long LK, Vogt PK (1981) Virology 110:411–419
33. Andersen PR, Devare SG, Tronick SR, Ellis RW, Aaronson SA, Scolnick EM (1981) Cell 26:129–134
34. Duesberg PH, Bister K, Vogt PK (1977) Proc Natl Acad Sci USA 74:4320–4324
35. Wang L-H, Duesberg PH, Beemon K, Vogt PK (1975) J Virol 16:1051–1070
36. Oskarrson M, Maizel JV, Vande Woude GF (1980) Science 207:1222–1224
37. Wang L-H, Hanafusa H, Notter MFD, Balduzzi PC (1982) J Virol 41:833–841
38. Feldman RA, Wang L-H, Hanafusa H, Balduzzi PC (1982) J Virol 42:228–236
39. Shibuya M, Hanafusa H, Balduzzi PC (1982) J Virol 42:1007–1016
40. Mathey-Prevot B, Hanafusa H (1982) Cell 28:897–906
41. Wong T-W, Goldberg AR (1981) Proc Natl Acad Sci USA 78:7412–7416
42. Lai MMC, Neil JC, Vogt PK (1980) Virology 100:475–483
43. Yoshida M, Toyoshima K (1980) Virology 100:484–487
44. Pawson T, Martin GS (1980) J Virol 34:280–284
45. Sheiness D, Vennstrom B, Bishop JM (1981) Cell 23:291–300
46. Wang L-H, Snyder P, Hanafusa T, Hanafusa H (1980) J Virol 35:52–64
47. Graf T, Stehelin D (to be published) Reviews on cancer. Biochim Biophys Acta

48. Bishop JM (1981) Cell 23:5-6
49. Vigne R, Neil JC, Breitman ML, Moscovici C, Vogt PK (1980) Ann N Y Acad Sci 354:384-397
50. Sefton BM, Hunter T, Beemon K (1980) Proc Natl Acad Sci USA 77:2059-2063
51. Collett MS, Brugge JS, Erikson RL (1978) Cell 15:1363-1369
52. Oppermann H, Levinson AD, Varmus HE, Levintow L, Bishop JM (1979) Proc Natl Acad Sci USA 76:1804-1808
53. Karess RE, Hayward WS, Hanafusa H (1979) Proc Natl Acad Sci USA 76:3154-3158
54. Rohrschneider LR, Eisenman RN, Leitch CR (1979) Proc Natl Acad Sci USA 76:4479-4483
55. Collett MS, Erikson E, Purchio AF, Brugge JS, Erikson RL (1979) Proc Natl Acad Sci USA 76:3159-3163
56. Wang L-H, Feldman R, Shibuya M, Hanafusa H, Notter MFD, Balduzzi PC (1981) J Virol 40:258-267
57. Hayman MJ, Royer-Pokara B, Graf T (1979) Virology 92:31-45
58. Kitchener G, Hayman MJ (1980) Proc Natl Acad Sci USA 77:1637-1641
59. Yoshida M, Kawai S, Yoyoshima K (1981) J Virol 38:430-437
60. DeLamarter JF, Neil JC, Ghysdael J, Vogt PK (1981) Virology 112:757-761
61. Adkins B, Hunter T, Beemon K (1982) J Virol, in press
62. Shealy DJ, Mosser AG, Rueckert RR (1980) J Virol 34:431-437
63. Neil JC, Breitman ML, Vogt PK (1981) Virology 108:93-110
64. Neil JC, Ghysdael J, Vogt PK, Smart JE (1981) Nature 291:675-677
65. Patschinsky T, Sefton BM (1981) J Virol 39:104-114
66. Patschinsky T, Hunter T, Esch FS, Cooper JA, Sefton BM (1982) Proc Natl Acad Sci USA 79:973-977
67. Neil JC, Ghysdael J, Smart JE, Vogt PK (1982) Virology 121:274-287
68. Krebs EG, Beavo JA (1979) Annu Rev Biochem 487:923-959
69. Schaffhausen B, Benjamin TL (1981) J Virol 40:184-196
70. Radke K, Martin GS (1979) Proc Natl Acad Sci USA 76:5212-5216
71. Erikson E, Erikson RL (1980) Cell 21:829-836
72. Houmard J, Drapeau GR (1972) Proc Natl Acad Sci USA 69:3506-3510
73. Hunter T, Sefton BM (1980) Proc Natl Acad Sci USA 77:1311-1315
74. Radke K, Gilmore T, Martin GS (1980) Cell 21:821-828
75. Oppermann H, Levinson W, Bishop JM (1981) Proc Natl Acad Sci USA 78:1067-1071
76. Ek B, Westermark B, Wasteson A, Heldin C-H (1982) Nature 295:419-420
77. Collett MS, Erikson RL (1979) J Virol 29:770
78. Ghysdael J, Neil JC, Vogt PK (1981) Proc Natl Acad Sci USA 9:5847-5851
79. Levinson AD, Courtneidge SA, Bishop JM (1981) Proc Natl Acad Sci USA 78:1624-1628
80. Czernilofsky AP, Levinson AD, Varmus HE, Bishop JM, Tischer E, Goodman HM (1980) Nature 287:198-203
81. Van Beveren C, Gallenshaw JA, Jonas V, Berns AJM, Doolittle RF, Donoghue DJ, Verma IM (1981) Nature 289:258-262
82. Oppermann H, Levinson AD, Varmus HE (1981) Virology 108:47-70
83. Collett MS, Erikson RL (1978) Proc Natl Acad Sci USA 75:2021-2024
84. Levinson AD, Oppermann H, Levintow L, Varmus HE, Bishop JM (1978) Cell 15:561-572
85. Sefton BM, Hunter T, Beemon K (1980) J Virol 33:220-229
86. Rubsamen H, Friis RR, Bauer H (1979) Proc Natl Acad Sci USA 76:967-971
87. Collett MS, Purchio AF, Erikson RL (1980) Nature 285:167-169
88. Erikson RL, Collett MS, Erikson E, Purchio AF, Brugge JS (1979) Cold Spring Harbor Symp Quant Biol XLIV:907-917
89. Levinson AD, Oppermann H, Varmus HE, Bishop JM (1980) J Biol Chem 255:11973-11980
90. Feldman R, Hanafusa T, Hanafusa H Cell 22:757-765
91. Pawson T, Guyden J, Kung T-H, Radke K, Gilmore T, Martin GS (1980) Cell 22:767-775
92. Neil JC, Ghysdael J, Vogt PK (1981) Virology 109:223-228
93. Bister K, Lee W-H, Duesberg PH (1980) J Virol 36:617-621
94. Lee WH, Bister K, Moscovici C, Duesberg PH (1981) J Virol 38:1064-1076
95. Gilmer TM, Erikson RL (1981) Nature 294:771-773
96. Spector M, O'Neal S, Racker E (1980) J Biol Chem 255:8370-8373

97. Spector M, O'Neal S, Racker E (1981) J Biol Chem 256:4219–4227
98. Spector M, Pepinsky RB, Vogt VM, Racker E (1981) Cell 25:9–21
99. Vogt VM, Pepinsky RB, Racker E (1981) Cell 25:827
100. Sefton BM, Hunter T, Beemon K, Eckhart W (1980) Cell 20:807–816
101. Mampe A, Loprevotte I, Galibert F, Fedele LA, Sherr CJ (1982) Cell 30:775–786
102. Shibuya M, Hanafusa H (1982) Cell 30:787–795
103. Kitamura N, Kitamura A, Toyoshima K, Hirayama Y, Yoshida M (1982) Nature 297: 205–208
104. Sefton BM, Hunter T, Ball EH, Singer SJ (1981) Cell 24:165–174
105. Cooper JA, Hunter T (1981) Mol Cell Biol 1:165–178
106. Burr JG, Dreyfuss G, Penman S, Buchanan JM (1980) Proc Natl Acad Sci USA 77:3484–3488
107. Erikson E, Cook R, Miller GJ, Erikson RL (1981) Mol Cell Biol 1:43–50
108. Cooper JA, Hunter T (1981) Mol Cell Biol 1:394–407
109. Nakamura KD, Weber MJ (1982) Mol Cell Biol 2:147–153
110. Todaro G, DeLarco JE, Cohen S (1976) Nature 264:26–31
111. Carpenter G, King L, Cohen S (1978) Nature 276:409–410
112. Ushiro H, Cohen S (1980) J Biol Chem 255:8363–8365
113. Hunter T, Cooper JA (1981) Cell 24:741–752
114. Cooper JA, Hunter T (1981) J Cell Biol 91:878–883
115. Hirano A, Vogt PK (1981) Virology 109:193–197
116. Hirano A, Wong TC, Vogt PK (1982) Virology 116:646–649
117. Shih C, Shilo B-Z, Goldfarb MP, Dannenberg A, Weinberg RA (1979) Proc Natl Acad Sci USA 76:5714–5718
118. Cooper GM, Okenguist S, Silverman L (1980) Nature 284:418–421
119. Cooper GM, Neiman PE (1980) Nature 292:857–858
120. Shilo B, Weinberg RA (1980) Nature 289:607–609
121. Brugge JS, Erikson E, Erikson RL (1981) Cell 25:363–372
122. Rohrschneider L (1980) Proc Natl Acad Sci USA 77:3514–3518
123. Brugge JS, Darrow D (1981) Nature 295:250–253
124. Smart JE, Oppermann H, Czernilofsky AP, Purchio AF, Erikson RL, Bishop JM (1981) Proc Natl Acad Sci USA 78:6013–6017
125. Adkins B, Hunter T, Sefton BM (1982) J Virol, in press
126. Purchio AF (1982) J Virol 41:1–7
127. Adkins B, Hunter T (1982) Mol Cell Biol 2:890–896
128. Pawson T, Kung T-S, Martin GS (1981) J Virol 40:665–672
129. McGrath JP, Levinson AD (1982) Nature 295:423–425
130. Vogt VM, Wight A, Eisenman R (1979) Virology 98:154–167
131. Neel BG, Wang L-H, Mathey-Prevot B, Hanafusa T, Hanafusa H, Hayward WS (1982) Submitted
132. Parker RC, Varmus HE, Bishop JM (1981) Proc Natl Acad Sci USA 78:5842–5846
133. Cohen P, Yellowlees D, Aitken A, Donella-Deana A, Hennings BA, Parker PJ (to be published) Eur J Biochem
134. Chang EH, Furth ME, Scolnick EM, Lowy DR (1982) Nature 297:479–483

# Murine Leukemia Viruses with Recombinant *env* Genes: A Discussion of Their Role in Leukemogenesis

NANCY G. FAMULARI*



* Department of Viral Oncology, Memorial Sloan Kettering Cancer Center, 1275 York Ave, New York, NY 10021, USA

# 1 Introduction

In 1951 Ludwik Gross reported experiments which demonstrated the viral etiology of certain murine leukemias. His work identified the presence of an infectious agent in cell-free extracts of AKR thymomas which could transmit leukemia to C3Hf/Bi mice (*Gross* 1951). Subsequent electron microscopic analysis demonstrated the presence of virus particles with the characteristic type c morphology of murine leukemia virus in both preparations of AKR thymomas and of leukemias induced in C3Hf/Bi mice by Gross' protocol (*Dmochowski* and *Grey* 1957; *Bernhard* 1960). In the ensuing 30 years it has become obvious that the murine leukemia viruses (MuLV) are a polymorphic family of infectious agents, heterogeneous in their ability to induce disease in mice and even in their ability to infect and replicate in mouse cells (for review see *Stephenson* 1980). Host range classes of MuLV have been defined that are ecotropic, i.e., infectious for mouse cells, and xenotropic, i.e., infectious for cells of heterologous species but not for mouse cells. In addition, it has become clear that MuLV are maintained in the murine population predominantly via vertical transmission of endogenous viral loci carried by the mouse genome, rather than by horizontal transmission of infectious virus. These viral loci consist of complete and partial viral genomes, the expression of which can result in production of infectious virus or in synthesis of viral gene products in the absence of virus replication. A great deal of effort has been directed toward characterization of the repertoire of MuLV that exists in the mouse, and toward identification of the factors which influence the complex virus-host interaction which governs their expression in vivo (for review see *Lilly* and *Pincus* 1973; *Old* and *Stockert* 1977; *Jolicoeur* 1979). One outcome of these investigations has been the identification by *Hartley* and co-workers (1977) of the particular class of leukemogenic MuLV which Gross assayed in his early experiments. This class of MuLV, the dualtropic, mink cell focus-forming (MCF) viruses, appears to arise de novo in the individual mouse by recombination involving viral *env* gene sequences. The following review is a discussion of what is known currently about MuLV with recombinant *env* genes: (a) the association of this class of MuLV with leukemias of viral etiology, (b) the characteristics of genome structure and viral phenotypes which define this class of MuLV, and (c) the aspects of genomic structure and virus expression which appear to be associated with their ability to induce leukemogenesis.

# 2 Expression of MuLV with Recombinant *env* Genes in Leukemias of Viral Etiology

## 2.1 Leukemias Arising Spontaneously in High-Leukemia-Incidence Mouse Strains

Characterization of ecotropic virus expression in mice of the high-leukemia-incidence strains laid the groundwork for the identification of MuLV with recombinant *env* genes. Insight into the virus-host interaction which results in the generation of this novel class of MuLV came from studies involving two mouse strains in particular: (a) the AKR strain developed by selective inbreeding for high incidence of thymic lymphoma (*Furth* et al. 1933; *Lynch* 1954), and (b) the BALB/Mo substrain constructed experimentally by the

introduction of Moloney-MuLV (M-MuLV) into the germ line of the low leukemia incidence strain, BALB/c (*Jaenisch* 1976).

### 2.1.1 Correlation Between Ecotropic Virus Replication in Young Animals and Leukemia Development in Adults

It has been demonstrated that the inheritance of genes encoding ecotropic MuLV is a prerequisite for development of spontaneous leukemia in adult mice (*Meier* et al. 1973; *Rowe* 1973). However, the simple inheritance of ecotropic virus loci does not result in the high incidence of disease. Mice of the AKR strain are characterized by expression of infectious, ecotropic MuLV encoded by the viral genes they carry (*Rowe* and *Pincus* 1972). *Lilly* and co-workers (1975) demonstrated that both early virus activation and infection of host tissue are required for the development of spontaneous leukemias in this mouse strain. Their study utilized the *Fv-1* restriction of MuLV replication segregating in (BALB/c × AKR) F1 × AKR backcross progeny. A strong correlation was observed between high titer of infectious, $XC^+$, N-ecotropic virus in tail extracts of 6–8 week old animals and leukemia development in these mice as adults. The suppressive influence on leukemogenesis of *Fv-1* restriction of virus replication has been confirmed in several additional studies involving genetic crosses of AKR mice to strains carrying *Fv-1* alleles restrictive for replication of N-ecotropic virus (*Nowinski* et al. 1979; *Lee* and *Ihle* 1979; *Mayer* et al. 1978) and by the low leukemia incidence experienced by AKR mice congenic for the $Fv\text{-}1^b$ allele (E. Stockert personal communication).

Early virus expression and infection of host tissue also occurs in BALB/Mo mice (*Jaenisch* 1976, 1979). Although the parental BALB/c strain does carry endogenous ecotropic virus genes, MuLV encoded by these loci is expressed only late in life (*Hartley* et al. 1969; *Peters* et al. 1972). In contrast, M-MuLV encoded by the *Mov-1* locus of BALB/Mo mice is activated soon after birth resulting in early viremia. De novo infection of host tissue due to expression of M-MuLV has been demonstrated at the molecular level in these animals. By 4 weeks of age an increase in M-MuLV-specific DNA to approximately two copies per haploid mouse genome is detectable in thymus and spleen, but not in liver, brain, or kidney (*Jaenisch* 1979). In addition, unintegrated viral DNA is detectable in thymus tissue, a finding indicative of the occurrence of acute infection (*Jahner* et al. 1980). The development of viremia by these mice can be prevented by passive immunization of newborns with virus-neutralizing antibody. Treated mice do not exhibit an increase in M-MuLV copy member in spleen or thymus, and experience prolonged latent periods and reduced incidence of disease (*Nobis* and *Jaenisch* 1980). Passive immunization with anti-MuLV gp70 antibody is also effective in prolonging onset and lowering disease incidence in AKR mice when animals are treated immediately after birth. Viremia is eliminated, and individual mice develop an antiviral immune response (*Huebner* et al. 1976; *Schafer* et al. 1977; *Schwarz* et al. 1979). It has been postulated that intervention in the leukemogenic process by passive immunization is at the level of immune surveillance directed at preleukemic cells present early in life (*Schwarz* et al. 1981). However, $F(ab)_2$ fragments, generally ineffective in antibody-mediated cytotoxicity, are fully competent in preventing viremia in BALB/Mo mice (*Nobis* and *Jaenisch* 1980). In both BALB/Mo and AKR mice, intervention in the development of lymphoma by immune therapy is effective only when treatment is initiated immediately after birth. It appears likely that the efficacy of passive immunization in prevention of disease deve-

lopment results from interference with virus spread, an interpretation consistent with the inhibition of leukemogenesis by *Fv-1* restriction of virus spread.

### 2.1.2 Emergence of MuLV with Recombinant *env* Genes in the Target Tissue for Transformation

It is clear that the expression of endogenous ecotropic MuLV plays a critical role in leukemogenesis in AKR and BALB/Mo mice. However, several lines of evidence indicate that direct action of endogenous ecotropic viruses in transformation is highly unlikely. The complexity of the role of endogenous ecotropic virus in leukemogenesis was first suggested by studies of *Kaplan* and *Yaffe* (1960; *Kaplan* 1967) and *Nishizuka* and *Nakakuki* (1968). These experiments examined normal thymus of AKR mice for expression of the pathogenic agent which *Gross* (1951) had identified in thymoma extracts of this strain. Cell-free extracts of thymus tissue prepared from 1-month-, 3-month-, and 5-month-old animals were assayed in AKR mice for acceleration of leukemogenesis, i.e., disease onset prior to 180 days of age (*Rudali* et al. 1956). Leukemogenic virus was detectable in thymic extracts from 3-month-old animals and increased in potency in extracts from 5-month-old animals. However, no activity was assayable in thymic extracts prepared from 1-month-old animals or from spleen extracts from animals of 1, 3, or 5 months of age. *Nishizuka* and *Nakakuki* (1968) attributed this age-dependent appearance of leukemogenic virus to an effect of virus concentration that is overcome with age. However, subsequent measurements of ecotropic virus titer in tissues of AKR mice showed that the level of ecotropic virus present in thymus extracts remains fairly constant throughout life and is generally higher in spleen than in thymus (*Rowe* and *Pincus* 1972; *Kawashima* et al. 1976b; *Nowinski* and *Doyle* 1977). In addition, direct assay of the leukemogenicity of in vitro isolates of the endogenous AKR ecotropic virus in AKR and C3Hf/Bi mice demonstrates that these ecotropic viruses cannot induce disease via exogenous infection of newborn mice (*Nowinski* and *Hays* 1978; *Cloyd* et al. 1980; *Pedersen* et al. 1981; *O'Donnell* et al. 1981; *Famulari* and *O'Donnell*, unpublished data).

A clue to understanding events occurring in the virus population of AKR thymus was provided by the analysis of MuLV antigen expression and virus production by thymocytes as a function of age (*Kawashima* et al. 1976a, b). An increase in the cell surface expression of MuLV *env* and *gag* gene-related antigens on thymocytes is detectable at 5–6 months of age when compared to antigen expression on thymocytes of 2-month-old animals (*Kawashima* et al. 1976a). Serological analysis of the gp70 molecules responsible for the increase in *env* gene expression shows that they are distinguishable from those of endogenous ecotropic virus (*Stockert* et al. 1979; *Tung* and *Fleissner* 1980; *Obata* et al. 1981). Comparison of virus expression in the thymus of 2- and 6-month-old animals shows little change in ecotropic virus titer, but does reveal an increase in titer of MuLV with xenotropic host range. This increase does not occur in spleen (*Kawashima* et al. 1976b). Examination of the population of replicating virus with xenotropic host range in 6-month thymus resulted in the identification of MuLV with recombinant *env* genes in this tissue (*Hartley* et al. 1977). It is postulated that this virus class is generated in the thymus of AKR mice at 5–6 months of age by recombination between replicating ecotropic MuLV and endogenous viral sequences which share characteristics with xenotropic MuLV. The recombinational event invariably involves the *env* gene, and frequently results in a virus with dual ecotropic-xenotropic host range and the ability to induce

cytopathic foci (MCF) on mink CCL64 cells (*Hartley* et al. 1977). Such isolates are often referred to as dualtropic, or dualtropic MCF MuLV. However, acquisition of these two phenotypes is not always associated with the recombinant *env* gene, a circumstance which will be discussed later in the review. It is expression of the *env* and *gag* gene products of this class of MuLV which appears to be responsible for the increase in MuLV antigens detected on thymocytes of AKR mice at 5–6 months of age (*Stockert* et al. 1979; *Tung* and *Fleissner* 1980; *Obata* et al. 1981).

The pathogenicity of dualtropic MuLV further distinguishes this virus class from that of ecotropic and xenotropic MuLV. Cloned isolates of dualtropic MuLV recovered from extracts of both preleukemic and leukemic thymus have the ability to induce leukemias in newborn and young adult AKR and C3Hf/Bi mice, whereas the endogenous ecotropic and xenotropic viruses of AKR mice do not (*Nowinski* and *Hays* 1978; *Cloyd* et al. 1980; *O'Donnell* et al. 1981; *Pedersen* et al. 1981; Famulari and O'Donnell, unpublished results). Thus, dualtropic MuLV represent a class of virus with the characteristics which *Gross* (1951), *Kaplan* and *Yaffe* (1960), and *Nishizuka* and *Nakakuki* (1968) identified in their analysis of extracts of AKR thymus; specifically, age-dependent, thymus-specific expression, and the ability to transmit leukemia via exogenous infection of newborn or young adult animals. An apparent relationship between the requirement for expression of endogenous ecotropic MuLV in leukemia development and the appearance of dualtropic virus exists in the postulated mechanism of generation of the dualtropic class.

Expression of MuLV with recombinant *env* genes has also been associated with spontaneous leukemias of several other mouse strains, including those of BALB/Mo strain (*Vogt* 1979; *Bosselman* et al. 1979; *van der Putten* et al. 1981). Two additional examples are the lymphomas of the inbred strain HRS/J (*Green* et al. 1980) and the recombinant inbred strain BXH-2 (*Bedigian* et al. 1981). In both strains ecotropic virus expression is detectable early in life. Individuals of the HRS/J strain develop thymic lymphomas, a phenomenon correlated with homozygosity at the locus of a mutant autosomal gene (*hr*) (*Hiai* et al. 1977). Mice of the BXH-2 recombinant inbred strain experience a high incidence of lymphoma reportedly of non-T cell origin. MuLV with recombinant *env* genes has been isolated from lymphomatous tissue of both strains (*Green* et al. 1980; *Bedigian* et al. 1981).

## 2.2 Presence of MuLV with Recombinant *env* Genes in Leukemias Induced by Exogenous Infection with Ecotropic MuLV

Induction of leukemia by a variety of ecotropic viruses has been described (*Jaenisch* et al. 1975; *Buchhagen* et al. 1976; *Gross* and *Dreyfuss* 1978; *Jolicoeur* et al. 1978; *Rapp* and *Todaro* 1978; *Rosenberg* and *Baltimore* 1978; *Troxler* and *Scolnick* 1978; *Armstrong* et al. 1980; *MacDonald* et al. 1980a; *Oliff* et al. 1980; *Reddy* et al. 1980; *Hoffman* et a. 1981; *Ishimoto* et al. 1981; *Pedersen* et al. 1981). In contrast to endogenous ecotropic MuLV (*Jolicoeur* et al. 1978; *Nowinski* and *Hays* 1978; *Cloyd* et al. 1980; *O'Donnell* et al. 1981), leukemogenic ecotropic viruses can induce disease via exogenous infection of newborn mice. This type of ecotropic MuLV has been isolated from such diverse material as tumor tissue with an extensive and complex in vivo passage history as in the case of Friend, Moloney, and Rauscher viruses (for review see *Gross* 1970), primary leukemias or cultures of primary leukemias as in the case of Gross ecotropic virus (*Hartley* et al. 1969), and

AKR SL viruses (*Nowinski* and *Hays* 1978; *Pedersen* et al. 1981), and even normal tissue of aged animals as in the case of B-tropic isolates from BALB/c mice (*Hartley* et al. 1969). As might be predicted from their varied origins, many characteristics, including the types of neoplasm they induce, distinguish the leukemogenic ecotropic isolates from one another. The phenotypes governing the ability of ecotropic MuLV to induce leukemia via exogenous infection have not been identified; however, recent studies indicate that a common feature of the leukemogenic process is the induction of dualtropic MCF viruses. The following sections will discuss these data. Since it is not possible to assess leukemogenic potential or mechanism of disease induction by ecotropic MuLV if virus mixtures are potentially involved, only studies utilizing cloned, ecotropic virus isolates will be considered.

### 2.2.1 Leukemia Induction by Friend Ecotropic Virus

In the case of Friend virus, both biologically and molecularly cloned ecotropic isolates have been shown to induce rapid disease in newborn mice of a number of strains, including BALB/c and NIH Swiss (*Troxler* and *Scolnick* 1978; *MacDonald* et al. 1980a; *Oliff* et al. 1980; *Ishimoto* et al. 1981; *Ruscetti* et al. 1981a). Friend virus stocks can contain a complex of viruses comprised of an ecotropic component, F-MuLV, that is fully replication competent, and a defective, helper-dependent component, SFFV. Individually both the ecotropic and the defective component have the capacity to induce a lethal, erythroproliferative syndrome. The ecotropic component induces a disease that shares features in common with that induced by the Friend virus complex containing the anemia (FV-A) strain of SFFV. Characterized by anemia, splenomegaly, and elevated numbers of nucleated cells in peripheral blood, the F-MuLV-induced disease is fatal in 100% of animals with a survival time 40–60 days postinjection (*Oliff* et al. 1981).

Characterization of RNA transcripts in spleens of F-MuLV-inoculated animals detects the expression of both F-MuLV- and SFFV-related sequences as early as 14 days after inoculation, prior to development of splenomegaly (*Troxler* and *Scolnick* 1978). The cDNA probe used to distinguish SFFV-related transcription from F-MuLV transcription cross-hybridizes with replicating xenotropic viruses and with AKR and Moloney MCF viruses (*Troxler* et al. 1977b). When the virus repertoire present in spleens of F-MuLV-inoculated animals was analyzed, *env* gene recombinant viruses (Fr MCF viruses) related to F-MuLV were isolated (*Troxler* et al. 1978b). In addition, the expression of the Friend MCF virus *env* gene precursor was detected in passaged F-MuLV-induced leukemias, whereas expression of the SFFV *env* gene product, gp52, was not detected (*Oliff* et al. 1981). Like the F-MuLV/SFFV complex, Fr MCF isolates induce the Friend erythroproliferative disease in adult, as well as newborn, NIH Swiss mice (*Ruscetti* et al. 1981a). The ability of Fr MCF virus to induce disease in adult animals contrasts the recombinant with F-MuLV which is reported to be pathogenic only in newborns (*Troxler* and *Scolnick* 1978; *MacDonald* et al. 1980a).

### 2.2.2 Leukemia Induction by Gross Ecotropic Virus

The induction of thymic leukemias by Gross ecotropic virus (G-MuLV) also involves the generation of dualtropic MuLV (*Famulari* et al. 1982). G-MuLV was isolated by Dr. Janet Hartley from Gross passage A virus preparations, which are extracts of leukemias main-

tained by serial, cell-free mouse passage in the C3Hf/Bi strain (*Gross* 1957). Gross passage A virus is a complex comprised of Gross ecotropic virus and *env* gene recombinant MuLV (*Famulari* et al. 1982) which induces thymomas in 90%–100% of inoculated C3Hf/Bi mice within a 3-month latent period (*Gross* 1957). In contrast, the ecotropic virus component alone induces disease in only 20% of inoculated animals with a 7–8 month latent period (*Buchhagen* et al. 1976; *Gross* and *Dreyfuss* 1978).

Analysis of virus gene expression in primary leukemias induced by G-MuLV demonstrates the synthesis of two primary *env* gene products (PrENV proteins) in these cells. Peptide mapping identifies them as being that of the input ecotropic virus and that of a dualtropic virus. Replication-competent dualtropic virus has been isolated from these leukemias and is being tested for pathogenicity. It is interesting to note that, as is the case with F-MuLV, Gross ecotropic virus is leukemogenic in newborn but not adult mice. The Gross passage A complex containing dualtropic virus is equally pathogenic in adult and newborn mice (*Famulari* et al. 1982).

### 2.2.3 Leukemia Induction by Moloney and Rauscher Ecotropic MuLV

The induction of thymic lymphoma by Moloney MuLV is known to involve the generation of MuLV with recombinant *env* genes. Dualtropic virus has been identified in associated with Moloney MSV stocks, with spontaneous leukemias arising in BALB/Mo mice, and with leukemias induced by inoculation of newborn animals with M-MuLV (*Fischinger* et al. 1975; *Vogt* 1979; *Bosselman* et al. 1979; *van der Putten* et al. 1981). *Vogt* (1979) isolated dualtropic MCF viruses from cultured cell lines derived from spontaneous leukemias arising in BALB/Mo mice, and demonstrated that all cells of the cultured line produce both ecotropic and dualtropic virus. Analysis of M-MuLV-related genomes integrated in tumors of BALB/Mo mice, and BALB/c and 129 mice inoculated with M-MuLV, has identified the presence of proviral sequences which are recombinant between M-MuLV and endogenous viral information, in addition to the authentic M-MuLV genome, in all tumors examined (*van der Putten* et al. 1981).

Stocks of the Rauscher virus (R-MuLV) complex have been shown to contain an R-MuLV-related dualtropic virus which will induce erythroleukemia in NIH Swiss mice (*Van Griensven* and *Vogt* 1980). However, analysis of leukemias induced by cloned R-MuLV for the expression of dualtropic MuLV has not been reported.

### 2.2.4 Leukemia Induction by B-ecotropic MuLV

B-ecotropic virus isolates derived from normal tissue of BALB/c mice induce lymphoma upon inoculation into newborn BALB/c mice (*Jolicoeur* et al. 1978). Again, analysis of the virus repertoire present in these leukemias has not been reported. However, virus expression by lymphoreticular tumors induced by B-ecotropic MuLV isolated from (BALB/c X A) F1 mice has been characterized (*Armstrong* et al. 1977, 1980). The leukemogenic B-ecotropic virus was cloned out of mouse-passaged extracts of a lymphoreticular tumor triggered by protracted graft versus host reaction. Injection of this virus into newborn BALB/c mice induces lymphoreticular tumors which express both B-ecotropic and B-tropic MCF virus.

### 2.2.5 Leukemia Induction by Ecotropic MuLV of Wild Mouse Origin

Nonthymic lymphomas induced by Cas-Br-M, an ecotropic MuLV isolated from a wild mouse, have been shown to express MCF virus-specific antigens and to produce virus which encodes these antigens (*Hoffman* et al. 1981). Expression of an MCF virus type-specific antigen and MCF virus is found in spleen but not thymus of Cas-Br-M virus-inoculated animals, and appears in prelymphomatous, as well as lymphomatous, tissue. A neurological disorder characterized by hind limb paralysis occurs prior to lymphoma. However, it is unclear whether MCF virus expression is associated with the neurological manifestations.

### 2.2.6 Leukemogenic Ecotropic MuLV of AKR Origin

N-ecotropic viruses (AKR SL viruses) which can induce leukemias in newborn AKR and C3Hf/Bi mice have been isolated from a series of cultured, spontaneous AKR leukemias (*Nowinski* and *Hays* 1978; *Pedersen* et al. 1981). Characterization by RNase $T_1$ oligonucleotide fingerprinting and mapping has demonstrated that the genomes of two such isolates share structural similarities with that of G-MuLV and are distinguishable from that of Akv virus in the 3′ coding region of p15(E) and in the U3 region of the LTR (*Buchhagen* et al. 1980; *Pedersen* et al. 1981). Recently we have demonstrated that thymomas induced in C3Hf/Bi and NFS mice by a molecular clone of one of these ecotropic isolates, SL3-3 (*Lenz* et al. 1982) express the PrENV protein of a dualtropic MuLV (Famulari and O'Donnell, unpublished data). Ecotropic MuLV which appear to encode recombinant *env* genes have also been isolated from these AKR cultured leukemias (*Pedersen* et al. 1981, 1982; see sect. 3.1. for discussion).

### 2.2.7 Properties of Leukemogenic Ecotropic MuLV Which Might Contribute to Their Pathogenicity

It is obvious from the studies described in the preceding sections that the generation of dualtropic MuLV is a characteristic of the leukemogenic process of many, if not all, pathogenic ecotropic MuLV. In fact, the events which occur during disease induction by exogenous infection of newborn mice with ecotropic MuLV appear to be completely analogous to those of spontaneous leukemogenesis in high-incidence strains. It is not clear why all ecotropic MuLV cannot initiate this process upon inoculation into newborn mice. Why, for example, ecotropic MuLV encoded by the *Akv* loci of AKR mice do not induce disease upon exogenous infection of C3Hf/Bi mice, but appear to initiate the process of leukemogenesis in AKR mice or NFS mice congenic for these loci is not known. A possible explanation is that the leukemogenic ecotropic viruses replicate more efficiently than do nonleukemogenic ecotropic MuLV when injected into newborn mice. AKR ecotropic virus is activated prenatally in that strain, a fact that might be advantageous to virus spread. It is also possible that leukemogenic ecotropic viruses are characterized by a function which enhances their ability to generate the recombinant virus, i.e., have an increased ability to recombine with endogenous xenotropic virus-related sequences. If this is the case, one might conjecture that the Akv viruses are not the proximal, ecotropic agent in the generation of recombinant viruses in AKR mice, but may be altered to become recombinagenic prior to participation in the formation of MCF virus.

The leukemogenic ecotropic viruses which have been isolated from some cultured AKR spontaneous leukemias might be an example of such an alteration in the endogenous AKR ecotropic MuLV (*Nowinski* and *Hays* 1978; *Pedersen* et al. 1981).

Experiments involving the construction and analysis of in vitro recombinants have been carried out to identify characteristics of the Friend ecotropic virus genome associated with the ability to initiate leukemogenesis. Recombinants have been generated between F-MuLV and the wild mouse, amphotropic MuLV 4070-A (A. Oliff, personal communication). The smallest segment of the F-MuLV genome which has been shown to confer pathogenicity on an in vitro recombinant maps from the start of the amino terminal coding region of gp70 through approximately half of the p15(E) coding region. The *gag* and *pol* genes and sequences spanning from the carboxy terminal coding region of p15(E) through the LTR of the recombinant carrying this segment of the F-MuLV *env* gene are of amphotropic virus origin. Mice inoculated with this recombinant develop Friend erythroproliferative disease, albeit at a lower incidence and with a longer latency than that which results from injection of F-MuLV itself. Expression of Fr MCF virus is invariably associated with these leukemias. Inoculation of mice with amphotropic MuLV does not induce the erythroproliferative syndrome. The mechanism by which these F-MuLV sequences confer pathogenicity on the in vitro recombinant virus is not understood. Although the amphotropic MuLV appears to replicate well in spleen, it is possible that the F-MuLV gp70 is required to confer target cell specificity on the in vitro recombinant. On the other hand, sequence homology between the F-MuLV *env* gene and the endogenous *env* gene sequences which are incorporated into the dualtropic virus genome may facilitate the recombinational event.

## 3 Characterization of MuLV with Recombinant *env* Genes

### 3.1 Phenotypes Which Define *env* Gene Recombinant MuLV

It is evident, as more MuLV with recombinant *env* genes are isolated and characterized, that the phenotypes thus far used to define this virus class, i.e., MCF induction, dualtropism, or polytropism, do not really encompass the range of recombinants that exist in nature. Given the degree of heterogeneity which potentially may result from the recombinant nature of the *env* gene, it may be difficult to identify a marker which will characterize all isolates. The phenotypes initially used to distinguish this virus class are dual ecoxenotropism and mink-cell focus (MCF) induction (*Fischinger* et al. 1975; *Hartley* et al. 1977). MCF induction is an extremely useful marker for the isolation of recombinant MuLV since it is unique to this virus class, and this phenotype has been used extensively to identify recombinant MuLV in virus mixtures. [The mink cell foci are not the characteristic transformed foci of murine sarcoma viruses, but rather localized areas containing pycnotic cells (*Hartley* et al. 1977)]. However, not all recombinant MuLV are MCF inducing or sufficiently dualtropic to be isolated directly in mink cells. For example, there are MCF-negative isolates among a series of recombinants derived from AKR thymus (*O'Donnell* et al. 1981). Several recombinant viruses isolated from Gross passage A extracts are only weakly infectious for mink cells and MCF-negative (*Famulari* et al. 1982). In addition, isolates have been described which induce a cytopathic effect in both mouse and mink cells (*Vogt* 1979; *Bedigian* et al. 1981; *Hamada* et al. 1981). Two rather

interesting recombinant MuLV, NIHC16 (*Rapp* and *Todaro* 1978) and SL3-2 (*Pedersen* et al. 1981), and apparently ecotropic in host range. Further, NIHC16 has been reported to be $XC^+$. Both isolates are leukemogenic, and appear to encode recombinant *env* genes. SL3-2 is known to have an extensive substitution of nonecotropic virus sequences in the *env* gene (*Pedersen* et al. 1981); NIHC16 encodes an MCF-specific *env* gene antigen and ecotropic-virus-related *gag* gene antigens (*Devare* et al. 1978). gp70 of NIHC16 lacks a large, complex oligosaccharide of 5100 daltons (Famulari and Kemp, unpublished data) which is characteristically present on gp70 of ecotropic MuLV but not found on gp70 of xenotropic MuLV or MuLV with recombinant *env* genes (*Kemp* et al. 1979, 1980). In addition, the primary *env* gene product of NIHC16 and SL3-2 has the characteristic low molecular weight (Famulari, unpublished data) described for the PrENV proteins of xenotropic and recombinant MuLV [approximately 5000 daltons less than that of ecotropic MuLV (*Famulari* and *Jelalian* 1978; *Ruscetti* et al. 1979; *Famulari* and *English* 1981)].

Recent work has described a distinct interference group defined by a series of MuLV *env* gene recombinants from diverse origins (*Rein* 1982). To date, recombinant MuLV isolated from several mouse strains and exhibiting a broad range of dualtropic, MCF phenotypes have fallen into this interference group. Ecotropic MuLV (M-MuLV and Akv) and amphotropic MuLV do not cross-interfere with recombinant MuLV in the assay used to define this interference group, a finding which implies that ecotropic and recombinant MuLV infect mouse cells via different receptors. The assay, based on focus formation by MSV pseudotypes on NIH/3T3 cells, does not rely on secondary infection of cells to score an infectious event. Because of this, phenotypic masking of viruses does not influence the results. It will be interesting to determine whether this phenotype defines a characteristic of the recombinant MuLV sufficiently general to include recombinants with ecotropic host range, such as NIHC16 and SL3-2. It is presently known that isolates that induce thymic lymphoma (Moloney MCF, AKR MCF 247), as well as isolates which induce erythroleukemia (Friend MCF), are members of this new interference group.

## 3.2 Genomic Structure of MuLV with Recombinant *env* Genes

Comparison of the genomic structure of recombinant MuLV with that of their respective ecotropic virus parents has been useful in defining and characterizing this class of virus. Analysis of a variety of recombinant viruses suggests that their genomes are colinear with those of the other classes of MuLV containing the three viral genes, *gag, pol,* and *env*, flanked by long terminal repeat (LTR) sequences. The genomes of recombinant MuLV have been characterized directly by hybridization studies using specific probes (*Troxler* et al. 1978), by RNase $T_1$ oligonucleotide mapping (*Faller* et al. 1978; *Faller* and *Hopkins* 1978; *Rommelaere* et al. 1978; *Shih* et al. 1978; *Evans* et al. 1980; *Green* et al. 1980; *Lung* et al. 1980, 1983; *Pedersen* et al. 1981, 1982), by heteroduplex mapping (*Chien* et al. 1978; *Donoghue* et al. 1978; *Bosselman* et al. 1979), and by restriction enzyme analysis (*Chattopadhyay* et al. 1981), and indirectly by antigenic and tryptic peptide analysis of viral gene products (*Elder* et al. 1977; *Hartley* et al. 1977; *Devare* et al. 1978; *Fischinger* et al. 1978; *Troxler* et al. 1978; *Cloyd* et al. 1979; *O'Donnell* and *Nowinski* 1980; *O'Donnell* et al. 1980; *Niman* and *Elder* 1981; *Pinter* et al. 1982). These studies have provided an understanding of the architecture of the genome of recombinant MuLV and of the recombinational events

involved in the generation of this class of virus. However, determination of the precise number and location of the crossover events which occur in the formation of a particular isolate requires sequence analysis of the genome and comparison with at least one of the parent viruses. This type of analysis is in progress in several laboratories. The *gag* gene, and often the *pol* gene of the recombinants, appear to be derived from the ecotropic virus parent. In some instances limited differences have been detected between the *gag* gene of recombinant MuLV and their presumed ecotropic virus parent; however, the two virus types are known to share close homology in this region of the genome. It is unclear what the minor sequence changes encountered in the *gag* gene signify in terms of the mechanism of generation of the recombinant genome. However, it is possible that mutations may be acquired, or secondary recombinational events may occur during the process of isolation. Recombinant MuLV can acquire a *pol* gene of ecotropic origin, or one composed, at least in part, of sequences from the nonecotropic virus parent. The *env* gene, however, always contains substitutions of ecotropic virus information derived from the so-called xenotropic virus-related parent.

Comparative analysis of a series of recombinant viruses derived from mice of the AKR strain and from NFS mice congenic for ecotropic virus loci originating from several mouse strains has provided some understanding of the possible configurations of the *env* gene substitutions (*Lung* et al. 1980, 1983; *Chattopadhyay* et al. 1981). Two general configurations have been identified by these studies (see Fig. 1). In both cases, the xenotropic virus-related substitution begins at or near the start of the coding region for the amino terminus of gp70. (Whether the *pol* gene substitutions which have been mapped in some isolates are continuous with those found in the *env* gene has not been established). In one instance (type A in Fig. 1), the substitution of xenotropic virus-related information extends through the coding region for gp70 and p15(E). However, it appears that the ecotropic virus parent contributes the U3 region of the LTR (*Lung* et al. 1983). Recombinants of this type have been isolated from the NFS congenic mice and have been designated class II viruses. These recombinants do not replicate in vivo and are not pathogenic (*Cloyd* et al. 1980; *Lung* et al. 1980; *Rowe* et al. 1980). The second type of recombinant *env* gene (type B in Fig. 1) is composed of xenotropic virus-related information at both the 5′ and 3′ ends of the gene, interrupted by an insert of ecotropic virus

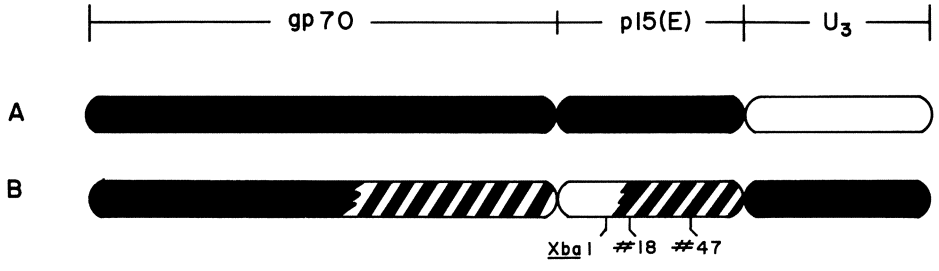



Fig. 1. Schematic representation of the *env* gene of recombinant MuLV isolated from NFS mice congenic for ecotropic virus loci (*A*) and from AKR mice (*B*). *Black area* indicates sequences derived from the xenotropic virus-related parent; *white area* sequences derived from the ecotropic virus-related parent. The *striped areas* can be donated by either parent. The relative location in the p15(E) coding region of the *Xba* 1 site (*Chattopadhyay* et al. 1981) and oligonucleotides 18 and 47 (*Lung* et al. 1983) is determined from DNA sequence analysis of MCF 247 and Akv virus (*Kelly* et al. 1983)

sequences. This insert can begin in the coding region for the carboxy terminus of gp70, or it can begin at or near the amino terminal coding region of p15(E). In either case, the insert continues varying distances downstream through the p15(E) coding region but terminates before the LTR. Thus, for this type of recombinant the amino terminal portion of p15(E) is, in all cases, of ecotropic virus origin and the U3 of the LTR is of xenotropic virus origin (*Lung* et al. 1983; N. Hopkins, personal communication). It appears that the gp70 coding region may be completely substituted by xenotropic virus-related sequences or mosaic with xenotropic virus-related sequences at the 5′ end and ecotropic virus sequences at the 3′ end (*Chattopadhyay* et al. 1981; *Lung* et al. 1983). This type of *env* gene configuration has been identified in a series of recombinants of AKR origin termed class I isolates (*Hartley* et al. 1977; *Cloyd* et al. 1980) and $A^+L^+$ and $A^+L^-$ isolates (*O'Donnell* et al. 1981). In addition, one isolate from an NFS mouse congenic for the *Akv-1* locus encodes this type of *env* gene (*Lung* et al. 1983). These recombinants replicate in thymus, and some, but not all, isolates with this type of genome anatomy are pathogenic.

### 3.3 Defective Viruses with Recombinant *env* Genes: Friend Spleen Focus-Forming Virus (SFFV)

Defective MuLV carrying recombinant *env* gene sequences have been identified in Friend and Rauscher virus preparations (*Steeves* and *Mirand* 1969; *Bentvelzen* et al. 1972; *Steeves* 1975; *Troxler* et al. 1977a, b, c). Although these MuLV require a helper virus to replicate, they encode the information for induction of erythroproliferative disease (*Fieldsteel* et al. 1971; *Steeves* et al. 1971; *Linemeyer* et al. 1980; *MacDonald* et al. 1980b; *Troxler* et al. 1980). The best-studied examples of such isolates are the variants of strains of SFFV found in association with F-MuLV in the Friend virus complex. The SFFV strains were isolated from Friend virus stocks maintained by serial passage in different mouse strains (*Friend* 1957; *Axelrad* and *Steeves* 1964; *Mirand* et al. 1968; *Lilly* and *Steeves* 1973), and were differentiated initially by their ability to induce disease characterized by polycythemia in some cases, and anemia in others (*Mirand* et al. 1968). The polycythemia strain is referred to as FV-P and the anemia strain as FV-A.

Analysis of the genomes of several strains of SFFV has demonstrated that these viruses are composed of a partially deleted *gag* gene which shares sequences in common with the F-MuLV *gag* gene, a deleted *pol* gene, and an *env* gene which shares sequences in common with those of xenotropic and *env* gene recombinant MuLV, as well as with that of F-MuLV (*Troxler* et al. 1977a, b, 1980; *Evans* et al. 1979, 1980; *Bosselman* et al. 1980; *Ruscetti* et al. 1980). Comparison of the genomes of F-MuLV and SFFV by heteroduplex analysis indicates that a deletion exists in the p15(E) coding region of the SFFV *env* gene (*Bosselman* et al. 1980; D. Linemeyer, personal communication). Analysis of SFFV gene products confirms and extends these observations. The *env* gene of all strains of SFFV encodes a glycoprotein of 52 000–55 000 daltons which has been shown to be antigenically and structurally related to gp70 of Friend and Moloney MCF viruses (*Racevskis* and *Koch* 1978; *Dresler* et al. 1979; *Ruscetti* et al. 1979, 1980; *MacDonald* et al. 1980b; *Ruta* and *Kabat* 1980). This protein does not appear to contain p15(E) antigenic determinants (*Schultz* et al. 1980). Despite the lack of p15(E) sequences, gp52 appears to be equivalent to a truncated form of the *env* gene polyprotein (PrENV protein) rather than of gp70. Proteolytic cleavage of gp52 into two components does not occur. However, the oligosac-

charide side chains of this molecule can undergo further processing (*Racevskis* and *Koch* 1978; *Dresler* et al. 1979; *Ruscetti* et al. 1979, 1981b; *Kabat* et al. 1980; *Ruta* and *Kabat* 1980) in a manner which appears to be similar to the processing of the high-mannose oligosaccharide side chains of the PrENV protein into complex side chains during the generation of gp70 and p15(E) (*Witte* and *Wirth* 1979; *Rosner* et al. 1980; *Kemp* et al. 1981). Processing of the carbohydrate side chains of gp52 generates a 60 000–65 000 dalton glycoprotein which is expressed on the surface of SFFV-infected cells (*Ruscetti* et al. 1979, 1981b; *Kabat* et al. 1980; *Ruta* and *Kabat* 1980). The *gag* gene products encoded by SFFV vary from strain to strain, apparently as a result of differences in the deletions sustained in this region of the genome in the generation of the defective virus (*Bernstein* et al. 1977; *Barbacid* et al. 1978; *Evans* et al. 1980; *MacDonald* et al. 1980b; *Ruscetti* et al. 1980; *Ruta* and *Kabat* 1980). Certain FV-P strains have been shown to encode a *gag* gene protein of 45 000 daltons which apparently contains p15, p12, and p30 antigenic determinants. This molecule can be glycosylated and expressed on the cell surface (*MacDonald* et al. 1980b; *Ruscetti* et al. 1980; *Ruta* and *Kabat* 1980).

Thus, the defective SFFV genome has the features of a deleted replication-competent *env* gene recombinant. The events leading to the appearance of this type of MuLV and the exact relationship between the nonecotropic virus sequences of SFFV and the Friend MCF viruses are not known.

## 3.4 The Origin of the Xenotropic Virus-Related Sequences

The nonecotropic virus parent involved in the generation of recombinant MuLV shares sequences in common with, but does not appear identical to, replicating xenotropic MuLV. Comparison of the genomes of xenotropic and recombinant MuLV demonstrates sequence homology between the two virus classes in the nonecotropic virus substitutions found in the recombinants (*Troxler* et al. 1977a, b; *Troxler* and *Scolnick* 1978; *Bosselman* et al. 1979, 1980; *Green* et al. 1980; *Chattopadhyay* et al. 1981; Rommelaere and Hopkins, unpublished data). However, nucleic acid sequences are found in the nonecotropic virus substitution which are not present in any replicating xenotropic MuLV examined thus far (*Green* et al. 1980; *Chattopadhyay* et al. 1981; Rommelaere and Hopkins, unpublished data). gp70 of recombinant viruses has been shown to contain tryptic peptides found in gp70 of xenotropic MuLV (*Elder* et al. 1977; *Fischinger* et al. 1978; *Troxler* et al. 1978), but tryptic peptide analysis and partial protease digest analysis (method of *Cleveland* et al. 1977) of gp70 of several recombinants which appear to be completely substituted in the gp70 coding region show these molecules to be distinct from all xenotropic virus gp70 examined (*Elder* et al. 1977; *Famulari* and *English* 1981; Famulari, unpublished data). Serological analysis of gp70 of recombinant MuLV demonstrates the presence of xenotropic virus type-specific antigens on these molecules (*Hartley* et al. 1977; *Hoffman* et al. 1981; *O'Donnell* et al. 1980). However, several gp70 antigens have been identified which are type-specific for recombinant MuLV, i.e., not shared by known ecotropic or xenotropic virus (*Devare* et al. 1978; *Cloyd* et al. 1979; *Ruscetti* et al. 1979; *Stockert* et al. 1979).

The distribution on normal mouse tissue of one of these antigens, $G_{AKSL2}$, provides a clue to the identity of the xenotropic virus-related parent participating in the generation of at least some recombinant MuLV. $G_{AKSL2}$ is expressed on normal thymocytes and

bone marrow cells of high-leukemia-incidence strains (AKR, PL, C58) early in life, but is not encoded by replicating ecotropic and xenotropic MuLV of these strains (*Stockert* et al. 1979). However, many recombinant MuLV isolated from preleukemic and leukemic thymus of AKR mice and one isolate from a leukemia of an NFS mouse congenic for *Akv-1* have been shown to encode this antigen; recombinant MuLV of PL and C58 origin have not been examined (*O'Donnell* et al. 1980, 1981). It has been suggested that the $G_{AKSL2}$ antigen is carried on a differentiation-linked gene product, related to, but not identical with, gp70 of replicating xenotropic MuLV. This thymocyte gp70 which may carry both the $G_{AKSL2}$ and the xenotropic virus-related antigen $G_{ERLD}$ (*Obata* et al. 1981) could be encoded by an endogenous virus locus analogous to the *Gv-1* locus of 129 mice (*Stockert* et al. 1971), the locus encoding the differentiation-linked gp70 carrying $G_{IX}$ and $G_{ERLD}$ in the 129 strain (*Obata* et al. 1975, 1981; *Tung* et al. 1975). Recombination between replicating ecotropic virus and the endogenous sequences encoding $G_{AKSL2}$ and $G_{ERLD}$ could generate recombinants such as have been isolated from AKR thymus.

Multiple copies of sequences which may encode these endogenous xenotropic virus-related *env* genes have been identified in DNA of AKR embryos (*Chattopadhyay* et al. 1982). The probe used to detect these endogenous sequences is an 0.7-kbp segment of the *env* gene coding region of SFFV defined at the 5′ limit by a *Bam*HI site and at the 3′ limit by an *Eco*RI site (BE probe). Recombinant MuLV have a corresponding sequence in their *env* gene mapping between 6.2 and 6.9 kbp from the 5′ end of the genome. This probe, which cross-hybridizes to xenotropic MuLV, but not to ecotropic MuLV, recognizes specific restriction enzyme fragments present in DNA from AKR embryos. These restriction enzyme fragments are found in the genome of replicating MCF viruses, but not in that of replication-competent xenotropic virus. Thus, the xenotropic virus-related *env* gene sequences which participate in the generation of recombinant MuLV may be present in the mouse genome in the form of complete or partial proviruses, an hypothesis consistent with the appearance of substituted *pol* genes in some *env* gene recombinant viruses (*Chattopadhyay* et al. 1981; *Lung* et al. 1983).

## 3.5 A Repertoire of Xenotropic Virus-Related Sequences Available for Recombination with Ecotropic Virus

Comparison of the substituted regions encoded by recombinants isolated from a particular mouse strain suggests that a repertoire of endogenous MuLV sequences is available for recombination with the ecotropic virus parent (*Chattopadhyay* et al. 1981; *Lung* et al. 1983). Analysis of gp70 which appear to be completely substituted provides a means of characterizing the endogenous sequences involved in recombination because heterogeneity detected among gp70 of this type can be attributed to the endogenous sequences participating in the recombination rather than to the mosaic nature of the gp70 molecule. The recombinants (class II and class I intermediate) isolated from spontaneous neoplasms of NFS mice congenic for ecotropic virus loci from various inbred strains (*Cloyd* et al. 1980) appear to encode gp70 completely substituted by nonecotropic virus sequences (*Lung* et al. 1980, 1983; *Chattopadhyay* et al. 1981; N. Hopkins, personal communication; see Fig. 1, type A). gp70 of four such isolates has been analyzed for expression of several antigens and for in vitro host range (*Cloyd* et al. 1979; *O'Donnell* et al. 1980, 1981).

All but one of these viruses originated from nonthymic leukemia, from sources such as splenic and lymph node reticulum cell sarcoma and mixed hematopoietic neoplasm. One isolate has an uncertain origin, having been derived from a thoracic lymphoma which arose in either thymus or mediastinal lymph node (*Cloyd* et al. 1980). *Akv-1*-C36, the recombinant originating from a thoracic lymphoma of an NFS mouse congenic for the *Akv-1* locus (*Cloyd* et al. 1980), encodes the $G_{AKSL2}$ antigen (*O'Donnell* et al. 1981) as well as the MCFA-3 antigen (*Cloyd* et al. 1979). None of the recombinants of nonthymic origin, *Akv-2*-C34, *C58v-1*-C48, and *C58v-2*-C45, (class II isolates) has acquired the ability to encode $G_{AKSL2}$ antigen (*O'Donnell* et al. 1981); *C58v-1*-C48 encodes the MCF-specific antigen MCFA-3 (*Cloyd* et al. 1979). *Akv-1*-C36 plates with equal titers on mouse SC-1 cells and mink lung cells in vitro. In contrast, *Akv-2*-C34, *C58v-1*-C48, and *C58V-2*-C45 exhibit a preferential tropism for mink cells relative to mouse cells (*O'Donnell* et al. 1981). In addition, the recombinants of nonthymic origin fail to replicate in vivo in spleen or thymus, whereas *Akv-1*-C36 will replicate in thymus, albeit poorly (*Cloyd* et al. 1980).

Thus, the analysis of recombinants arising on the NFS background indicates that a family of sequences is available in this strain for recombination with ecotropic virus. Analysis of four isolates has described the acquisition of three different gp70 coding regions by recombinants arising in this mouse strain. The acquisition of $G_{AKSL2}$ antigen by the *Akv-1*-C36 isolate demonstrates the presence of sequences encoding this antigen in NFS mice, despite the fact that the antigen is not expressed on normal tissue of this strain. It is tempting to speculate that the failure of the nonthymic isolates to acquire this antigen is due to a tissue-specific availability for recombination of sequences which encode $G_{AKSL2}$.

Suggestive evidence exists that a family of xenotropic virus-related sequences can participate in the generation of recombinants in AKR mice as well. Most of the recombinants from this strain which have been characterized for phenotypes of *env* gene products have mosaic gp70. However, one nonpathogenic recombinant which has been analyzed, 26-4 virus (*O'Donnell* et al. 1980, 1981), appears to be completely substituted in the gp70 coding region, and the pathogenic isolate, MCF 13 (*Hartley* et al. 1977), appears to have at most a small insert of ecotropic virus information in the gp70 coding region. gp70 of a third AKR isolate, AKR L5 (1375-2), which appears to be completely substituted has not been analyzed (*Chattopadhyay* et al. 1981; *Lung* et al. 1983; N. Hopkins, personal communication). Comparison of gp70 of 26-4 and MCF 13 viruses demonstrates that the nonpathogenic isolate does not encode $G_{AKSL2}$ and has a preferential tropism for mink cells relative to mouse cells in vitro (*O'Donnell* et al. 1980). The peptide map of gp70 of 26-4 virus is closely related, but not identical, to that of replicating xenotropic MuLV of AKR (*Famulari* and *English* 1981). MCF 13 virus, in contrast, encodes the $G_{AKSL2}$ antigen and plates with equal efficiency on mouse and mink cells in vitro (*O'Donnell* et al. 1981). The peptide map of gp70 of this recombinant is quite distinct from that of replicating AKR xenotropic virus (*Elder* et al. 1977; *Famulari* and *English* 1981). Since the sequences which encode the $G_{AKSL2}$ antigen appear to be derived from the nonecotropic virus parent (*Stockert* et al. 1979), it is likely that 26-4 and MCF 13 virus have different nonecotropic virus parents. It also seems likely that the particular xenotropic virus-related sequences acquired by a recombinant influence the biological properties of that virus. For example, all thymotropic recombinants examined to date encode the $G_{AKSL2}$ antigen, a finding which suggests that this antigen may be a marker for sequences which can confer tissue specificity on a recombinant.

# 4 Characteristics of Recombinant MuLV Associated with Pathogenicity

## 4.1 Aspects of Virus-Host Interaction Required for Pathogenicity by Recombinant MuLV: The Ability to Infect Target Cells

Recombinant MuLV which can be isolated from premalignant and malignant tissue represent a heterogeneous population of viruses with regard to their ability to induce disease. Isolates which cannot be demonstrated to establish an obvious infection in vivo do not induce disease (*Cloyd* et al. 1980; *Rowe* et al. 1980; *O'Donnell* et al. 1981; *Ruscetti* et al. 1981a).

A variety of lines of evidence support the idea that pathogenic recombinants must infect and replicate in the target cells for transformation. In AKR mice it appears that thymocytes become infected with recombinant MuLV at 5–6 months of age prior to onset of disease. This event was observed initially as an increase in the level of expression of MuLV-related antigens, both *gag* and *env* gene encoded, on the surface of thymocytes (*Kawashima* et al. 1976a). An identical high expression of MuLV gene products occurs on the surface of thymocytes of young AKR mice (2 months of age) as a consequence of exogenous infection by certain recombinant MuLV (*O'Donnell* et al. 1981, 1982). When virus expression has reached maximum levels, greater than 90% of thymocytes exhibit high levels of virus gene expression. This phenomenon has been referred to as "antigen amplification" because it represents an eight- to tenfold increase in antigen expression over that detected on uninfected thymocytes of 2-month-old animals (*Kawashima* et al. 1976a, b; *Nowinski* and *Doyle* 1977). PrENV protein and gp70 encoded by recombinant MuLV can be detected biochemically in the cytoplasm and on the surface of these infected thymocytes (*Tung* and *Fleissner* 1980; *Famulari*, unpublished data), and replicating recombinant MuLV can be isolated from these cells (*Hartley* et al. 1977; *O'Donnell* et al. 1981). A similar phenomenon has been observed in NFS mice injected with a wild mouse ecotropic MuLV, Cas-Br-M. In these animals, which develop nonthymic lymphoma, expression of an MCF virus-related antigen carried by gp70 is detected in spleen, but not thymus, several weeks prior to the onset of disease (*Hoffman* et al. 1981).

Integrated viral genomes of recombinant MuLV have been identified in both virus-induced and spontaneous leukemias of BALB/c, BALB/Mo, and AKR mice by the presence of characteristic restriction enzyme fragments in the mouse genome (*Pedersen* et al. 1980; *van der Putten* et al. 1981; *Quint* et al. 1981; *Chattopadhyay* et al. 1982). The expression both of the gene products of recombinant MuLV and of infectious virus, albeit at low titers in some instances, has been documented in spontaneous leukemias of high-incidence strains and in leukemias induced by exogenous infection with either ecotropic or recombinant MuLV (*Hartley* et al. 1977; *Racevskis* and *Koch* 1978; *Dresler* et al. 1979; *Ruscetti* et al. 1979; *Vogt* 1979; *Armstrong* et al. 1980; *Bilello* et al. 1980; *Cloyd* et al. 1980; *Famulari* et al. 1980, 1982; *Green* et al. 1980; *Housman* et al. 1980; *Pedersen* et al. 1980; *Rowe* et al. 1980; *Ruta* and *Kabat* 1980; *Van Griensven* and *Vogt* 1980; *Anand* et al. 1981; *Bedigian* et al. 1981; *Hoffman* et al. 1981; *Ishimoto* et al. 1981; *Lyles* and *McConnell* 1981; *O'Donnell* et al. 1981, 1982; *Oliff* et al. 1981; *Linemeyer* et al. 1982).

Discussion of the requirement for recombinant MuLV to infect the target cell for transformation raises the question of target cell specificity of these viruses. It has been hypothesized that thymotropism is a phenotype which must be acquired by recombinant MuLV which induce thymic leukemias (*Cloyd* et al. 1980; *Rowe* et al. 1980; *O'Donnell* et

al. 1981). Thymotropism is operationally defined in this context on the basis of an isolate's ability to establish infection in thymocytes in vivo either by the criterion of production of input virus or expression of detectable viral gene products by these cells. It is not known where the block to establishment of infection by the so-called nonthymotropic recombinants occurs. Both virus-specific or mouse-strain-specific mechanisms can be envisaged, including acquisition by the recombinant of a highly immunogenic gp70 or one with poor affinity for receptors for infectivity on mouse cells. This is particularly fertile ground for speculation because relatively little is known about host and virus phenotypes which effect replication of recombinant MuLV in vivo. The identity of the receptors which mediate infection by recombinant MuLV is not clear, nor is it known whether such receptors are expressed in a tissue-specific fashion. The description of a distinct interference group for recombinant MuLV in mouse cells suggests that these viruses do not infect thymocytes via the ecotropic virus receptor (*A. Rein* 1982). In fact, although an ecotropic virus receptor has been measured on thymocytes of adult mice (*Fowler* et al. 1977; *DeLarco* et al. 1978), some question exists as to whether de novo infection of thymocytes by ecotropic MuLV can occur in adult mice. [It does appear that ecotropic virus can infect mouse thymocytes or their precursors early in life (*Nobis* and *Jaenisch* 1980; *Tung* et al. 1980)]. In AKR mice the block to infection by ecotropic virus observed in young adult animals may be at the level of virus penetration and be due to interference, since these thymocytes express ecotropic virus gp70 (*Tung* and *Fleissner* 1980). However, a similar block to ecotropic virus infection of thymocytes observed in C3Hf/Bi mice, a strain which is ecotropic-virus-negative until late in life, suggests that other factors may also be involved (*Famulari* et al. 1982). In both strains, thymocytes of adult animals are susceptible to infection by recombinant MuLV (*Cloyd* et al. 1980; *O'Donnell* et al. 1981; *Famulari* et al. 1982).

Evidence also suggest that Friend MCF virus may not use the ecotropic virus receptor to infect target cells in spleen (*Ruscetti* et al. 1981a). Certain mouse strains, including DBA/2, are resistant to induction of erythroproliferative disease by F-MuLV. F-MuLV is not blocked from establishing infection in spleens of these animals, but Friend MCF virus is blocked even when inoculated as a phenotypic mixture with F-MuLV or amphotropic MuLV. It has been suggested that the expression of a recombinant-like gp70 detectable on the cell surface of tissue of DBA/2 mice interferes with infection of this strain by recombinant MuLV. (Such interference would be expected to block disease induction by F-MuLV because recombinant MuLV generated as a result of F-MuLV infection would not be able to spread through the spleen).

In some instances it has been observed that the block in the establishment of in vivo infection by recombinant MuLV can be overcome by inoculation of the recombinant as a phenotypic mixture with ecotropic or amphotropic MuLV (*Vogt* 1979; *Haas* and *Patch* 1980; *Van Griensven* and *Vogt* 1980; *Ruscetti* et al. 1981a). In addition, it has been observed that NFS mice congenic for *Akv-1* or *Akv-2*, the ecotropic virus loci of AKR, and (AKR × NFS) F1 mice are more susceptible to leukemogenesis by MCF 247 virus than are ecotropic-virus-negative NFS mice (*Cloyd* et al. 1981). One function that the ecotropic MuLV is thought to serve in phenotypic mixtures with recombinant MuLV is to protect the recombinant virus from inactivation by serum-neutralizing factor (*Haas* and *Patch* 1980). Whether this is the only function provided by the ecotropic virus is not known. However, data discussed in the preceding paragraphs suggest that the ecotropic virus does not provide recognition of receptors for infectivity. It should be borne in mind that

helper-dependent recombinants, like SFFV, do not appear to infect target cells via novel receptors. The defective viruses seem to rely on receptor recognition provided by helper virus, and are infectious in vivo when pseudotyped by a variety of helpers. Incorporation of gp52-like molecules into virus produced by SFFV/helper-virus-infected cells has not been detected, a finding which suggests that the SFFV *env* gene does not contribute to receptor recognition of these virions (*Racevskis* and *Koch* 1978; *Dresler* et al. 1979).

## 4.2 Aspects of Genomic Structure and Gene Expression of SFFV Which Correlate with Pathogenicity

Although it appears likely that the acquisition of a gp70 appropriate for establishing in vivo infection is prerequisite for a helper-independent recombinant MuLV to be pathogenic, the question remains open as to what function, if any, the protein products of the recombinant *env* gene play in transformation per se. Much effort has been directed toward answering this question by both analysis of genomic structure and gene products. Naturally occuring variants and in vitro recombinants of the helper-dependent SFFV and of a series of thymotropic, helper-independent MuLV have been studied to determine what aspects of the recombinant MuLV genome correlate with pathogenicity. Evidence is emerging from the analysis of both virus systems suggesting that expression of recombinant *env* gene sequences is involved in the disease induction capacity of these MuLV.

### 4.2.1 Nucleic Acid Sequences Required for Induction of Disease by SFFV

Molecular cloning techniques have been used to probe the genome of the FV-P strain of SFFV for sequences which direct the induction of the polycythemia form of the Friend erythroproliferative syndrome (*Linemeyer* et al. 1980, 1981, 1982). Various restriction enzyme fragments of the SFFV genome have been rescued by cotransfection of NIH3T3 cells with infectious, nonpathogenic helper virus DNA. Virus produced by such cotransfected cells can then be assayed in vivo for the ability to induce Friend disease. Since transfection of NIH3T3 cells with helper virus DNA alone does not yield virus preparations with demonstrable pathogenicity, the production of pathogenic virus by cotransfected cells can be attributed to information contributed by the subgenomic fragments of SFFV (*Linemeyer* et al. 1980). The organization of SFFV DNA and helper virus DNA in these virus preparations is not yet known (D. Linemeyer, personal communication). Both restriction enzyme fragments and deletion mutants of SFFV have been assayed by this method. A 1.5-kbp restriction enzyme fragment which maps in the SFFV *env* gene has been identified which encodes sufficient information to induce Friend erythroproliferative disease in mice (*Linemeyer* et al. 1982). This fragment was isolated from a molecular clone of SFFV into which a *Sal* I site had been introduced experimentally, and spans the region from the *BAM* HI site located in the 3′ end of the genome to a *Sal* I site 1.5 kbp downstream. gp52 is encoded by this fragment, which is only slightly larger than the estimated 1.4 kbp of DNA required to encode the unglycosylated form of gp52 [45 000–46 000 daltons (*Dresler* et al. 1979; *Schultz* et al. 1980)] (*Linemeyer* et al. 1982).

In addition, a series of SFFV mutants carrying deletions in the *env* gene have been shown to be inactive in vivo. The deletion sustained by one such mutant involves only *env*

gene sequences and appears to span the junction between the xenotropic virus-related and ecotropic virus information in this gene. Expression of gp52 is not detectable in spleen of mice injected with virus preparations containing this fragment. In contrast, virus preparations containing mutants with deletions in the 5′ coding region (*gag* gene) and LTR have been shown to be pathogenic and to encode gp52 in vivo. Oligonucleotide fingerprinting of the LTR deletion mutants indicates that, at a minimum, the right-hand half of the SFFV LTR has been deleted from these viruses. This analysis of partial SFFV genomes provides the best evidence to date that expression of the SFFV *env* gene protein is necessary and sufficient for induction of the Friend erythroproliferative disease in susceptible mice.

### 4.2.2 Synthesis and Processing of gp52

Analysis of the synthesis and processing of gp52 in both nonproducer fibroblast cell lines and erythroleukemia cell lines has demonstrated that this molecule is an extremely stable component of the cytoplasm of infected cells. Little evidence of chase is observed for greater than 2 h after synthesis (*Racevskis* and *Koch* 1978; *Dresler* et al. 1979; *Ruscetti* et al. 1979; *Lyles* and *McConnell* 1981). A stable association of gp52 with the nuclear fraction of SFFV-infected cells has been reported (*Lyles* and *McConnell* 1981). The kinetics of association of gp52, PrENV of F-MuLV, and the G protein of vesicular stomatitis virus with the nuclear fraction have been compared in studies which indicate the function of the nuclear membrane as part of the rough endoplasmic reticulum. Immediately after synthesis, all three glycoproteins are associated with the nuclear membrane fraction to a comparable extent. However, unlike the G protein and PrENV protein, gp52 appears to be unable to move through the pathway for glycoprotein processing efficiently and remains associated with this particular membrane fraction (D. Lyles, personal communication). These data are consistent with the findings that little gp52 is expressed on the cell surface, and apparently none is exfoliated from cells or incorporated into virus (*Racevskis* and *Koch* 1978; *Dresler* et al. 1979; *Ruscetti* et al. 1979, 1981b; *Kabat* et al. 1980; *Lyles* and *McConnell* 1981). The fact that p15(E) sequences are missing from gp52 may account for the static nature of the interaction of this molecule with intracellular membranes. It has been hypothesized that the accumulation of intracellular gp52 which results from the block in normal processing of this glycoprotein might play a role in pathogenesis by SFFV (*Ruta* and *Kabat* 1980; *Lyles* and *McConnell* 1981).

Clues to a second mechanism by which expression of the SFFV *env* gene may induce the Friend erythroproliferative syndrome have been provided by a series of studies which compare the two SFFV strains, FV-A and FV-P. It has been demonstrated that the particular strain of SFFV present in the Friend virus complex determines whether an infected mouse will develop the anemia or polycythemia form of Friend disease (*MacDonald* et al. 1980b; *Troxler* et al. 1980). Further, it has been demonstrated by tryptic peptide analysis that gp52 encoded by the FV-A and FV-P strains are different (*MacDonald* et al. 1980b). Although it is not known how structural differences in gp52 effect the phenotype of disease induced by SFFV, analysis of the intracellular processing and cell surface expression of *env* gene products of the two strains suggests a possible mechanism (*Ruscetti* et al. 1981b). While the rate of synthesis of the primary *env* gene product, gp52, appears to be similar in FV-A- and FV-P-infected cells in vivo and in vitro, processing of this molecule into higher molecular weight protein species is more efficient in FV-P-infected cells. This

processing, which appears to involve a shift in the oligosaccharide side chains of gp52 from the high-mannose to the complex type, results in a 65 000-dalton species in FV-P-infected cells and a 60 000-dalton species in FV-A-infected cells (*Ruscetti* et al. 1981b; *Ruta* et al. 1982). Only a small fraction of gp52 synthesized in cells infected by either strain of SFFV undergoes this processing, an estimated 5%–10% in FV-P-infected cells (*Ruta* and *Kabat* 1980; *Lyles* and *McConnell* 1981; *Ruscetti* et al. 1981b). However, it is estimated that eight- to tenfold less processing occurs in FV-A-infected cells (*Ruscetti* et al. 1981b). Expression of the 65 000-dalton FV-P *env* gene product, but not the 60 000-dalton FV-A protein, is detectable on the surface of infected fibroblasts and spleen cells (*Ruscetti* et al. 1981b).

This differential cell surface expression is of interest considered in conjunction with the differential erythropoietin requirements for burst formation of FV-A- or FV-P-infected erythroid progenitor cells. Bone marrow and spleen cells infected in vivo with the FV-P strain will produce erythroid bursts in vitro independent of addition of erythropoietin (*Hankins* and *Krantz* 1975; *Horoszewicz* et al. 1975; *Liao* and *Axelrad* 1975); in contrast, cells infected by FV-A in vivo require erythropoietin for formation of bursts in vitro (*Fagg* et al. 1980; *MacDonald* et al. 1980b). Infection of bone marrow cells in vitro by FV-A does result in low production of erythroid bursts in the absence of hormone (*Hankins* and *Troxler* 1980). Cells in these bursts exhibit erythroblast morphology, but they are poorly hemoglobinized. Addition of low levels of erythropoietin to FV-A-infected cells in this system results in a three- to eightfold increase in the number of bursts formed; bursts formed in the presence of the hormone are strongly hemoglobin-positive. While addition of erythropoietin to FV-P-infected bone marrow increases burst formation twofold, all bursts are hemoglobin-positive, irrespective of addition of the hormone. At the concentration of erythropoietin used in these experiments, uninfected bone marrow preparations are incapable of burst formation. These data suggest that sensitivity to the action of erythropoietin by burst-forming cells is effected by SFFV in a virus-strain-specific fashion. Since the FV-P strain directs the expression of more processed gp52 onto the surface of infected cells than does the FV-A strain (*Ruscetti* et al. 1981b), it is tempting to speculate that this molecule facilitates binding of the hormone to the cell surface.

In vivo, the induction of erythroblastosis by SFFV appears to be absolutely dependent on the expression of the protein products of the recombinant *env* gene. The hyperplasia observed in vivo as a consequence of infection may be due to alterations in the normal hormone requirements of erythropoiesis in infected cells. The expanded population of erythroblasts which results in vivo is composed of a repertoire of cell types, some of which are not normal components of the erythroid lineage (*Tambourin* et al. 1979). It appears likely that it is out of this population of proliferating, but not yet transformed, cells that the malignant phenotype emerges. In this hypothetical schema the role of expression of the SFFV *env* gene in the appropriate target cell would be to place this cell population in a proliferative state no longer subject to normal host controls. Secondary events occurring in this cell population would then generate the transformed (i.e., transplantable) phenotype.

## 4.3 Thymotropic MuLV with Recombinant *env* Genes: Aspects of Genomic Structure and Gene Expression Which Correlate with Pathogenicity

Comparative analysis of genomic composition has been used to define features of the helper-independent recombinants which correlate with the ability of an isolate to induce thymic leukemia. The studies which will be emphasized here focus on the analysis of a series of thymotropic recombinants of AKR origin. All the isolates under consideration score as thymotropic by the criteria of induction of virus gene expression and production of infectious virus by thymocytes infected in vivo. Some isolates accelerate leukemia development in AKR mice and some do not. The pathogenic viruses examined include class I isolates of *Hartley* et al. (1977) and *Cloyd* et al. (1980) and an $A^+L^+$ isolate of *O'Donnell* et al. (1981). This latter nomenclature refers to the phenotypes of induction of high levels of viral antigen expression in thymocytes, "A," and induction of lymphomas in the AKR acceleration test, "L." The nonpathogenic viruses examined are the $A^+L^-$ isolates of *O'Donnell* et al. (1981). These recombinants infect thymocytes and synthesize high levels of viral gene products in these cells, but do not induce leukemia. The comparison of these two classes of recombinant has taken on two aspects: (a) the characterization of gene products and genomic architecture, and (b) the analysis of gene expression and virus replication in vitro and in vivo. Very limited differences have been detected between the pathogenic and nonpathogenic isolates in the analysis of gene products and genome anatomy.

### 4.3.1 Genomic Structures Correlating with Pathogenicity

Fingerprinting of RNase $T_1$ oligonucleotides and restriction enzyme analysis of proviruses demonstrate that the *gag* gene coding region of these thymotropic isolates is of ecotropic virus origin (*Lung* et al. 1980, 1983; *Chattopadhyay* et al. 1981; N. Hopkins, personal communication), and indeed all the recombinants encode the ecotropic virus type-specific *gag* gene antigen, GCSA (*O'Donnell* and *Stockert* 1976; *O'Donnell* et al. 1981). Two recombinants, MCF 13 (pathogenic) and 28-7 (nonpathogenic), appear to have acquired substitutions of endogenous MuLV sequences in the *pol* gene coding region. The remaining five isolates encode *pol* genes of ecotropic virus origin (*Lung* et al. 1980, 1983; *Chattopadhyay* et al. 1981; N. Hopkins, personal communication).

The *env* gene of these isolates is composed of xenotropic virus-related sequences beginning at the start of the gp70 coding region and continuing through the U3 region of the LTR, interrupted by an insert of ecotropic virus information (see Fig. 1, type B; *Lung* et al. 1983; N. Hopkins, personal communication). Many of these thymotropic recombinants, both pathogenic and nonpathogenic isolates, carry an ecotropic virus insert starting in the coding region of the carboxy terminal portion of gp70, and thus have mosaic gp70. The extent of the ecotropic virus information in gp70 of MCF 69L1 is somewhat less than for the other isolates with mosaic gp70. However, the entire gp70 coding region of AKR-L5 virus appears to be derived from the nonecotropic virus parent; it is unclear whether MCF 13 virus carries a small ecotropic virus insert in the gp70 coding region or not (*Chattopadhyay* et al. 1981; *Lung* et al. 1983). Both of these isolates are pathogenic (*Cloyd* et al. 1980; *O'Donnell* et al. 1981; *Lung* et al. 1983). Thus, it appears that acquisition of a gp70 coding region containing both ecotropic and xenotropic virus sequences is not prerequisite for thymotropism or pathogenicity. [An additional pathogenic recombinant

of C58 origin has been identified which appears to encode a gp70 derived entirely from the xenotropic virus-related parent (*Cloyd* et al. 1980; *Lung* et al. 1983)]. It also seems likely that the phenotype of dualtropism in vitro is acquired from the xenotropic virus-related parent, since AKR-L5 and MCF 13 are both infectious for mouse and mink cells in culture (*Cloyd* et al. 1980; *O'Donnell* et al. 1981). This speculation is consistent with the findings of *Rein* (1982) which indicate that recombinant MuLV and ecotropic MuLV may use different receptors for infection of mouse cells.

Antigenic analysis and peptide mapping of gp70 of recombinants with mosaic gp70 does not discriminate between the pathogenic and nonpathogenic isolates (*O'Donnell* et al. 1980, 1981; *Famulari* and *English* 1981), although analysis of in vitro host range indicates that gp70 of two nonpathogenic isolates, SC37 and SC30, may differ from that of the other isolates. These two viruses exhibit a preferential tropism for mouse cells relative to mink cells in vitro. All of the other thymotropic recombinants are equally infectious for mouse and mink cells (*O'Donnell* et al. 1981). Antigenic and peptide analysis of gp70 of MCF 13 shows it to be unique among the series of isolates (*Elder* et al. 1977; *Famulari* and *English* 1981), but experiments have not been carried out to determine whether the amino terminal portion of MCF 13 gp70 is the same or different from that of the recombinants with mosaic gp70. gp70 of AKR-L5 has not been analyzed.

In light of the findings which indicate that expression of gp52 is required for disease induction by SFFV (*Linemeyer* et al. 1982), it will be important to reexamine the structure of gp70 encoded by these thymotropic recombinants to determine whether common sequences are acquired by the isolates which are pathogenic. If expression of gp70 in target cells is relevant to transformation by these viruses, it is likely that only part of the molecule is specifically involved in the pathogenic function, since viruses carrying mosaic gp70 and completely substituted gp70 are equally pathogenic and appear to induce the same malignancy. Comparison of the xenotropic virus-related sequences acquired in the amino terminal portion of gp70 of pathogenic and nonpathogenic isolates will be of particular interest.

Analysis of RNase $T_1$ oligonucleotides derived from the p15(E) coding region discriminates between the pathogenic and nonpathogenic isolates that are thymotropic (N. Hopkins, personal communication). It appears that the ecotropic virus insert extends further downstream into p15(E) coding sequences in the pathogenic isolates than in the nonpathogenic isolates. Since the sequence of the p15(E) coding region is known for the AKR ecotropic virus, Akv, and the AKR recombinant, MCF 247, it has been possible to locate particular oligonucleotides and restriction sites in the gene, and thus define the boundaries of the ecotropic virus insert (*Kelly* et al. 1983); see Fig. 1, type B for schematic). All of the thymotropic recombinants share an *Xba* 1 site with Akv virus that maps 168 bases downstream from the start of the p15(E) coding region (*Chattopadhyay* et al. 1981, personal communication; N. Hopkins personal communication). All of the pathogenic recombinants share an oligonucleotide (oligo 18) with Akv virus which maps from 227 to 244 bases downstream from the start of p15(E), and which is missing from all of the nonpathogenic recombinants except SC37. The next oligonucleotide (oligo 47), which the recombinants can share with ecotropic virus, maps from 424 through 438 bases downstream from the start of p15(E), and is missing from many of the pathogenic recombinants and from all of the nonpathogenic recombinants, including SC37 (N. Hopkins, personal communication). Thus, it appears that the ecotropic virus insert extends at least 244 bases into the p15(E) coding region of the pathogenic isolates, but may continue con-

siderably further into the gene. In contrast, the ecotropic virus insert in the nonpathogenic isolates appears to be shorter, apparently terminating between 168 and 227 bases into the p15(E) coding region. It will be important to determine whether the differences detected between the pathogenic and nonpathogenic viruses by these methods are reflected in the amino acid sequence of the p15(E) proteins encoded by these isolates.

The nonpathogenic recombinant, SC37, represents an exception to the correlation between the size of the ecotropic virus insert in p15(E) and pathogenicity. The existence of this exception may indicate that the correlation is invalid or it may indicate that the ecotropic virus insert must extend beyond oligo 18 for acquisition of pathogenicity. Because there is a stretch of approximately 200 nucleotides between oligo 18 and oligo 47 in which there are no RNase $T_1$ oligonucleotides which can be mapped, the minimum 3′ boundary of the ecotropic virus insert in the pathogenic isolates is not known precisely. A third explanation exists to reconcile the characteristics of the ecotropic virus insert of SC37 with the hypothesis that the length of these sequences effects pathogenicity. If the region of p15(E) which is contributed by the ecotropic parent is required because it interacts with gp70 to provide an in vivo function, then loss of appropriate sequences in the gp70 coding region might also render an isolate nonpathogenic. In this regard, the gp70 of SC37 virus is known to confer an altered in vitro host range on this isolate (*O'Donnell* et al. 1981). Possibly the gp70 of SC37, while conferring thymotropism on the isolate, lacks the hypothetical phenotype which confers pathogenicity. The concept that interaction between two virus gene products, in this case two *env* gene products, is involved in disease induction is an interesting one. Evidence has been presented which suggests that interrelated functions of the large T and small t proteins of SV40 may be required for induction of the transformed phenotype by this virus (*Kaplan* et al. 1981).

Additional studies (*Lung* et al. 1980, 1983; *Chattopadhyay* et al. 1981) have been carried out comparing the genomes of pathogenic isolates (class I, *Hartley* et al. 1977; *Cloyd* et al. 1980) with those of nonpathogenic isolates which do not replicate in vivo (class II, *Cloyd* et al. 1980). These nonpathogenic isolates which originated from nonthymic lymphoid neoplasms of NFS mice congenic for ecotropic virus loci appear to contain a substitution of xenotropic virus-related information which spans both the gp70 and p15(E) coding regions (Fig. 1, type A). Thus, the ecotropic virus derived *Xba* 1 site (*Chattopadhyay* et al. 1980) and oligo 18 (*Lung* et al. 1983) are not found in these isolates. Oligonucleotide mapping data indicates that the LTR region of these isolates is derived from the ecotropic virus parent (*Lung* et al., 1983). It is not known why these nonpathogenic isolates are unable to establish infection in vivo. These recombinants appear to replicate to high titer in vitro (*Cloyd* et al. 1980; *O'Donnell* et al. 1981), but exhibit a preferential tropism for mink cells relative to mouse SC-1 cells (*O'Donnell* et al. 1981). In addition, they are distinguishable from the thymotropic recombinants by their relative plating efficiency on SC-1 cells versus NFS embryo cells (*Rowe* et al. 1980). It is not known whether these characteristics of in vitro tropism have bearing on in vivo replication. It is interesting to draw a parallel between these NFS isolates and the Friend MCF viruses which also are unable to establish infection in vivo. In the case of Friend MCF virus, inoculation of animals with phenotypic mixtures of the recombinant and ecotropic or amphotropic MuLV overcomes the block to replication and results in induction of disease (*Ruscetti* et al. 1981a). Utilization of a similar approach in the analysis of these recombinants of NFS origin might help to establish whether they have pathogenic potential.

### 4.3.2 Expression of *env* Gene Products of Recombinant MuLV in Thymocytes

The results of experiments comparing the gene expression and replication of the pathogenic and nonpathogenic recombinants in thymocytes have been encouraging, in that no evident defects in the replication of the nonpathogenic viruses have been detected. Intrathymic injection of either type of isolate into weanling AKR mice results in expression of viral *env* and *gag* gene products on the surface of both cortical and medullary thymocytes (*O'Donnell* et al. 1981, personal communication). Approximately sixty percent of thymocytes can be shown to become infected by this criterion, probably as a result of virus spread from cell to cell through the thymus. Expansion of a small population of initially infected cells has not been formally excluded. However, this mechanism seems unlikely because the size distribution of cells in the thymus does not appear to change as a consequence of virus infection, and because the percentage of antigen-positive cells detectable at 30 days postinfection follows a linear dose response to virus dilution. Thus, the kinetics of virus spread through the thymus appear to be proportional to the titer of virus in the inoculum. Although an equivalent percentage of cells remain uninfected in animals inoculated with pathogenic or nonpathogenic isolates, it is not known whether this population differs qualitatively, depending on the pathogenicity of the recombinant. Nor is it known whether thymocytes which appear to remain uninfected represent a specific cell population.

Within 40–60 days after virus infection has spread through the thymocyte population, or 70–90 days after virus inoculation, leukemic blast cells emerge in thymus of mice inoculated with the pathogenic isolates, but not in those inoculated with nonpathogenic recombinants. These blast cells are transplantable into histocompatible recipients, whereas populations of virus-infected thymocytes which do not contain blast cells are not (*McGrath* and *Weissman* 1979; P. O'Donnell, personal communication). An analogous series of events occurs during the spontaneous development of leukemia in AKR mice. The thymocyte population becomes infected by recombinant MuLV at approximately 6 months after birth; the emergence of leukemic cells occurs 40–60 days later. The delay between virus infection of thymocytes and appearance of transformed cells indicates that replication of recombinant MuLV in thymocytes does not transform these cells directly, but appears to put the population at risk for transformation. What this priming event directed by virus might be is not known. However, models proposing mechanisms for virus induction of leukemia must be consistent with this lag between virus infection of thymocytes and transformation.

Preliminary experiments have been conducted in my laboratory and in the laboratory of P. O'Donnell to measure the level of *env* gene expression on the surface of infected thymocytes. Flow microfluorometry has been used to quantitate an anti-gp70 immunofluorescence reaction to these cells. The results of this analysis suggest that thymocytes infected by nonpathogenic isolates express the same level of the protein products of the recombinant *env* gene on their surface as cells infected by the pathogenic isolates. In addition, the rate of synthesis of the primary *env* gene products of the pathogenic and nonpathogenic viruses appears to be the same in thymocytes. It is likely that the immunofluorescence reaction being measured is detecting the level of gp70 expression on thymocytes. Although both gp70 and the primary *env* gene protein (PrENV protein) are found on the surface of thymocytes infected by either pathogenic or nonpathogenic recombinants, gp70 represents the predominant virus protein accessible to

radioiodination. The PrENV protein of *env* gene recombinants is a fairly stable component in leukemic cells; gp70 is difficult to detect as a cleavage product by kinetic measurements of processing of viral proteins (*Famulari* et al. 1980, unpublished data). Comparison of the kinetics of *env* gene processing of the pathogenic and nonpathogenic recombinants in vitro does not distinguish between the two types of isolate. However, a comparative kinetic analysis of *env* gene expression in thymocytes has not yet been carried out. It is known that the PrENV protein of recombinant MuLV interacts with the cellular machinery for glycoprotein processing in a way that distinguishes this virus class from ecotropic and xenotropic MuLV. The PrENV protein of recombinant MuLV is expressed on the surface of infected fibroblasts and leukemia cells, whereas that of ecotropic and xenotropic virus is not (*Famulari* and *Jelalian* 1979; *Famulari* et al. 1980; *Famulari* and *English* 1981). The form of the PrENV protein expressed on the cell surface appears to be transported to the plasma membrane without undergoing processing of its oligosaccharide side chains (*Kemp* et al. 1981). However, the PrENV proteins of both the pathogenic and nonpathogenic recombinants are processed in this fashion (*Famulari* and *English* 1981, unpublished data).

### 4.3.3 Models of Virus-Induced Transformation of Lymphocytes of the T Cell Lineage

Weissman and McGrath have proposed a model of virus-induced leukemogenesis based on their identification of a virus-binding activity carried by leukemic blast cells (for review, see *McGrath* et al. 1980a; *Weissman* and *McGrath* 1982). This virus-binding activity is a specific marker for leukemic cells and is not detectable on nontransformed thymocytes even if they are infected by virus. Originally it was proposed that a small proportion of normal thymocytes would carry a receptor responsible for this virus binding and that cells would be susceptible to infection by leukemogenic viruses via this receptor. The same receptor which mediated infection was proposed to act as a mitogen-binding receptor which would specifically recognize the *env* gene products of the infecting virus. Thus, once infected, thymocytes would be subject to sustained proliferation driven by binding of viral gene products, specifically gp70, produced as a consequence of infection. In light of current knowledge it seems unlikely that the virus-binding activity detectable on leukemic blasts is related to the virus receptor for infection of thymocytes. The thymocyte population is infected by recombinant MuLV well before the appearance of cells which bear the virus-binding activity assayed on leukemic blasts. Indeed, a thymocyte population infected by recombinant MuLV appears to be a hallmark of the preleukemic thymus in the AKR strain. It is possible, however, that the mitogen-binding function proposed for these receptors is involved in virus-induced leukemogenesis. The concept of self-driven proliferation which the hypothesis of receptor-mediated leukemogenesis espouses is an interesting one. In support of this hypothesis, McGrath and Weissman have presented evidence that monoclonal antibodies (anti-Thy-I) which block virus binding to a cultured AKR leukemia are cytostatic for this cell line, whereas antibodies which bind to these cells, but do not block virus binding are not cytostatic (*McGrath* et al. 1980b). Identification of the receptor responsible for virus binding by leukemic blasts and assessment of its potential for recognition of virus as mitogen will be an important step in testing this model.

Ihle and co-workers have proposed a second mechanism for virus-induced leukemogenesis based on their finding that the ability of a mouse strain to mount a cellular

immune response to MuLV correlates with susceptibility to leukemias of viral etiology (for review see, *Ihle* and *Lee* 1982). The antiviral immune response is measured as T cell blastogenesis in preparations of splenic lymphocytes which have been exposed to gp70 in vitro. This response has been observed in preleukemic animals of several mouse strains, including AKR and M-MuLV-injected BALB/c mice, and appears to be dependent on the establishment and maintenance of viremia in the host. Analysis of the blastogenic cell population indicates that two lymphocyte populations are involved in the proliferative response in vitro: (a) an antigen-specific population which manufactures mitogenic factors in response to exposure to MuLV antigens, and (b) an antigen-nonspecific population which is induced to proliferate by these factors (*Lee* and *Ihle* 1981b). Interleukin 2 (IL-2) and interleukin 3 (IL-3) have been identified among the factors produced by stimulated antigen-specific lymphocytes. The percentage of splenic lymphocytes capable of responding to factors elaborated in vitro by antigen-specific, virus-stimulated lymphocytes is increased 10- to 40-fold in preleukemic (viremic) mice as compared to control animals. This increase has been interpreted to be the consequence of a proliferation of factor-responsive cells resulting from increased production of mitogenic factors in vivo. Thus, a chronic cellular immune response to MuLV would result in an increase in production of mitogenic factors and, secondarily, in proliferation of various lymphocyte subpopulations. Ihle and co-workers hypothesize that the leukemic cell emerges from this population of proliferating lymphocytes.

Evidence that this immune response to MuLV is tied to the process of leukemogenesis comes from the analysis of CBA/N mice inoculated with M-MuLV (*Lee* and *Ihle* 1981a). These mice become viremic as a consequence of virus infection. However, they mount no detectable cellular immune response to MuLV and do not develop leukemia. Thus, the CBA/N strain provides an example of dissociation between development of viremia and disease. The resistance of DBA/2 mice to leukemogenesis by F-MuLV is a second example of such dissociation; however, it appears that in DBA/2 mice, a block in the leukemogenic process may occur at the level of resistance to infection by recombinant MuLV (*Ruscetti* et al. 1981a; see Sect. 4.2.2). Ihle and co-workers suggest that recombinant MuLV can be generated and replicate in the CBA/N strain; however, documentation of MuLV expression in these animals will be of importance in assessing the significance of the correlation between failure to develop cellular immunity to MuLV and resistance to leukemogenesis. The value of understanding virus expression in this strain is underscored by the finding that CBA/N mice, like DBA/2 mice, are resistant to induction of erythroproliferative disease by F-MuLV (*Ruscetti* et al. 1981a).

The model proposed by Ihle and co-workers does not consider the role of recombinant MuLV in the induction of leukemias. However, evidence from work of a number of laboratories strongly suggests that the generation of recombinants is a required step in virus-induced leukemogenesis in the mouse and, further, that the transformed cell is infected by this class of MuLV. A possible link between the model of Ihle and co-workers and what is known about the expression of recombinant *env* gene products in leukemia cells might involve an altered sensitivity of virus-infected thymocytes to mitogenic factors. One might hypothesize that infection of thymocytes by pathogenic recombinants would effect the interaction of these cells with mitogenic factors elaborated as a consequence of the cellular immune response to MuLV. The end result might be stimulation of proliferation or differentiation of infected thymocytes.

*Acknowledgments.* I would like to express my gratitude to Sisir Chattopadhyay, Nancy Hopkins, David Kabat, David Linemeyer, Douglas Lyles, Michael McGrath, Paul O'Donnell, Allen Oliff, and Alan Rein for their generosity in sharing their unpublished data with me. My thanks to Allen Oliff, Nancy Hopkins, Peter Besmer, Edward Stavnezer, and Erwin Fleissner for many discussions, and to Allen Oliff for his critical reading for the manuscript. Special thanks to Paul O'Donnell for encouragement and insight and much more.

It is a pleasure to acknowledge the assistance of Susan Jordan and Dawn Cieplensky in preparing this manuscript.

NGF is supported by grants CA-27950 and CA-16599 from the National Cancer Institute.

## References

Anand R, Lilly F, Ruscetti S (1981) Viral protein expression in producer and nonproducer clones of Friend erythroleukemia cell lines. J Virol 37:654–660

Armstrong MYK, Ruddle NH, Lipman MB, Pierce SK, Richards FF (1977) Role of endogenous murine leukemia virus in immunologically triggered lymphoreticular tumors. I. Development and use of oncogenic cell-free preparations serially passage in vivo. J Natl Cancer Inst 58:67–72

Armstrong MYK, Weininger RB, Binder D, Himsel CA, Richards FR (1980) Role of endogenous murine leukemia virus in immunologically triggered lymphoreticular tumors II. Isolation of B-tropic mink cell focus-inducing (MCF) murine leukemia virus. Virology 104:164–173

Axelrad AA, Steeves RA (1964) Assay for Friend leukemia virus: rapid quantitative method based on enumeration of macroscopic spleen foci in mice. Virology 24:513–518

Barbacid M, Troxler DH, Scolnick EM, Aaronson SA (1978) Analysis of translational products of Friend strain of spleen focus-forming virus. J Virol 27:826–830

Bedigian HG, Taylor BA, Meier H (1981) Expression of murine leukemia viruses in the highly lymphomatous BXH-2 recombinant inbred mouse strain. J Virol 39:632–640

Bentvelzen P, Aarssen AM, Brinkhof J (1972) Defectivity of Rauscher murine erythroblastosis virus. Nature 239:122–123

Bernhard W (1960) The detection and study of tumor viruses with the electron microscope. Cancer Res 20:712–727

Bernstein A, Mak TW, Stephenson JR (1977) The Friend virus genome: evidence for the stable association of MuLV sequences and sequences involved in erythroleukemia transformation. Cell 12:287–294

Billelo JA, Colletta G, Warnecke G, Koch G, Frisby D, Pragnell IB, Ostertag W (1980) Analysis of the expression of spleen focus-forming virus, (SFFV)-related RNA and gp55, a Friend and Rauscher virus-specific protein. Virology 107:331–344

Bosselman RA, van Griensven LJLD, Vogt M, Verma IM (1979) Genome organization of retroviruses VI. Heteroduplex analysis of ecotropic and xenotropic sequences of Moloney mink cell focus-inducing virus RNA obtained from either a cloned isolate or a thymoma cell line. J Virol 32:968–978

Bosselman RA, van Griensven LJLD, Vogt M, Verma IM (1980) Genome organization of retroviruses IX. Analysis of the genomes of Friend spleen focus-forming (F-SFFV) and helper murine leukemia viruses by heteroduplex-formation. Virology 102:234–239

Buchhagen DL, Pincus T, Stutman O, Fleissner E (1976) Leukemogenic activity of murine type C viruses after long-term passage in vitro. Int J Cancer 18:835–842

Buchhagen DL, Pedersen FS, Crowther RL, Haseltine WA (1980) Most sequence differences between the genome of the Akv virus and a leukemogenic Gross A virus passaged in vitro are located near the 3′ terminus. Proc Natl Acad Sci USA 77:4359–4363

Chattopadhyay SK, Lander MR, Gupta S, Rands E, Lowy DR (1981) Origin of mink cytopathic focus-forming (MCF) viruses: comparison with ecotropic and xenotropic murine leukemia virus genomes. Virology 113:465–483

Chattopadhyay SK, Cloyd MW, Linemeyer DL, Lander R, Rands E, Lowy DR (1982) Cellular origin and role of mink cell focus-forming viruses in murine thymic lymphomas. Nature 295:25–31

Chien Y-H, Verma IM, Shih TY, Scolnick EM, Davidson N (1978) Heteroduplex analysis of the sequence relations between the RNAs of mink cell focus-inducing and murine leukemia viruses. J Virol 28:352–360

Cleveland D, Fischer SG, Kirschner MW, Laemmli VK (1977) Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J Biol Chem 252:1102–1106

Cloyd MW, Hartley JW, Rowe WP (1979) Cell-surface antigens associated with recombinant mink cell focus-inducing murine leukemia viruses. J Exp Med 149:702–712

Cloyd MW, Hartley JW, Rowe WP (1980) Lymphomagenicity of recombinant mink cell focus-inducing murine leukemia viruses. J Exp Med 151:542–552

Cloyd MW, Hartley JW, Rowe WP (1981) Genetic study of lymphoma induction by AKR mink cell focus-inducing virus in AKR and NFS crosses. J Exp Med 154:450–458

DeLarco JE, Rapp UR, Todaro GJ (1978) Cell surface receptors for ecotropic MuLV: detection and tissue distribution of free receptors in vivo. Int J Cancer 21:356–360

Devare SG, Rapp VR, Todaro GJ, Stephenson JR (1978) Acquisition of oncogenicity by endogenous mouse type C viruses: effects of variations in *env* and *gag* genes. J Virol 28:457–465

Dmochowski L, Grey CE (1957) Subcellular structures of possible viral origin in some mammalian tumors. Ann NY Acad Sci 68:559–615

Donoghue DJ, Rothenberg E, Hopkins N, Baltimore D, Sharp PA (1978) Heteroduplex analysis of the nonhomology region between Moloney MuLV and the dual host range derivative HIX virus. Cell 14:959–970

Dresler S, Ruta M, Murray MJ, Kabat D (1979) Glycoprotein encoded by the Friend spleen focus-forming virus. J Virol 30:564–575

Elder JH, Gautsch JW, Jensen FC, Lerner RA, Hartley JW, Rowe WP (1977) Biochemical evidence that MCF murine leukemia viruses are envelope (*env*) gene recombinants. Proc Natl Acad Sci USA 74:4676–4680

Evans LH, Duesberg PH, Troxler DH, Scolnick EM (1979) Spleen focus-forming Friend virus: identification of genomic RNA and its relationship to helper virus RNA. J Virol 31:133–146

Evans L, Nunn M, Duesberg PH, Troxler D, Scolnick E (1980) RNAs of defective and nondefective components of Friend anemia and polycythemia virus strains identified and compared. Cold Spring Harbor Symp Quant Biol 44:823–835

Fagg B, Vehmeyer K, Ostertag W, Jasmine C, Klein B (1980) In: Rossi G (ed) In vivo and in vitro erythropoiesis: the Friend system. Elsevier, Amsterdam pp 163–172

Faller DV, Hopkins N (1978) $T_1$ oligonucleotide maps of Moloney and HIX murine leukemia viruses. Virology 90:265–273

Faller DV, Rommelaere J, Hopkins N (1978) Large $T_1$ oligonucleotides of Moloney leukemia virus missing in an *env* gene recombinant, HIX, are present on an intracellular 21S Moloney viral RNA species. Proc Natl Acad Sci USA 75:2964–2968

Famulari NG, English KJ (1981) *Env* gene products of AKR dualtropic viruses: examination of peptide maps and cell surface expression. J Virol 40:971–976

Famulari NG, Jelalian K (1979) Cell surface expression of the *env* gene polyprotein of dualtropic MCF murine leukemia virus. J Virol 30:720–728

Famulari NG, Tung J-S, O'Donnell PV, Fleissner E (1980) Murine leukemia virus *env*-gene expression in preleukemia thymocytes and leukemia cells of AKR strain mice. Cold Spring Harbor Symp Quant Biol 44:1281–1287

Famulari NG, Koehne CF, O'Donnell PV (1982) Leukemogenesis by Gross passage A murine leukemia virus: expression of viruses with recombinant *env* genes in transformed cells. Proc Natl Acad Sci USA 79:3872–3876

Fieldsteel AH, Dawson PJ, Kurahara C (1971) In vivo and in vitro recovery of defective Friend virus by various leukemia viruses. Int J Cancer 8:304–309

Fischinger PJ, Nomura S, Bolognesi DP (1975) A novel murine oncornavirus with dual eco- and xenotropic properties. Proc Natl Acad Sci USA 72:5150–5155

Fischinger PJ, Frankel AE, Elder JH, Lerner RA, Ihle JN, Bolognesi DP (1978) Biological, immunological, and biochemical evidence that HIX virus is a recombinant between Moloney leukemia virus and a murine xenotropic C-type virus. Virology 90:241–254

Fowler AK, Twardzik DR, Reed CD, Weislow OS, Hellman A (1977) Binding characteristics of Rauscher leukemia virus envelope glycoprotein gp71 to murine lymphoid cells. J Virol 24: 729–735

Friend C (1957) Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia. J Exp Med 105:302–318
Furth J, Seibold HR, Rathbone RR (1933) Experimental studies on lymphomatosis of mice. Am J Cancer 19:521–604
Green N, Hiai H, Elder JH, Schwartz RS, Khiroya RH, Thomas CY, Tsichlis PN, Coffin JM (1980) Expression of leukemogenic recombinant viruses associated with a recessive gene in HRS/J mice. J Exp Med 152:249–264
Gross L (1951) "Spontaneous" leukemia developing in C3H mice following inoculation in infancy, with AK leukemic extracts or AK embryos. Proc Soc Exp Biol Med 76:27–32
Gross L (1957) Development and serial cell-free passage of a highly potent strain of mouse leukemia virus. Proc Soc Exp Biol Med 94:767–771
Gross L (1970) Oncogenic viruses. 2nd edn. Pergamon, Oxford
Gross L, Dreyfuss Y (1978) Relative loss of oncogenic potency of mouse leukemia virus (Gross) after prolonged propagation in tissue culture. Proc Natl Acad Sci USA 75:3989–3992
Haas M, Patch V (1980) Genomic masking and rescue of dualtropic murine leukemia viruses: role of pseudotype virions in viral lymphomagenesis. J Virol 35:583–591
Hamada K, Yanagihara K, Kamiya K, Seyama T, Yokoro K (1981) Leukemogenicity and cell transformation mechanisms in vitro by Gross murine leukemia virus: analysis of virus subpopulations. J Virol 38:327–335
Hankins WD, Krantz SB (1975) In vitro expression of erythroid differentiation induced by Friend polycythemia virus. Nature 253:731–734
Hankins WD, Troxler D (1980) Polycythemia and anemia-inducing erythroleukemia viruses exhibit differential erythroid transforming effects in vitro. Cell 22:693–699
Hartley JW, Rowe WP, Capps WI, Heubner RS (1969) Isolation of naturally occurring viruses of the murine leukemia virus group in tissue culture. J Virol 3:126–132
Hartley JW, Wolford NK, Old LJ, Rowe WP (1977) A new class of murine leukemia virus associated with development of spontaneous lymphomas. Proc Natl Acad Sci USA 74:789–792
Hiai H, Morrissey R, Khiroya R, Schwartz RS (1977) Selective expression of xenotropic virus in congenic HRS/J (hairless mice). Nature 270:247–249
Hoffman PM, Davidson NF, Ruscetti SK, Chused TM, Morse HC (1981) Wild mouse ecotropic murine leukemia virus infection of inbred mice: dualtropic virus expression precedes the onset of paralysis and lymphoma. J Virol 39:597–602
Horoszewicz JS, Leung SS, Carter WA (1975) Friend leukemia: rapid development of erythropoietin-independent hematopoietic precursors. J Natl Cancer Inst 54:265–267
Housman D, Levenson R, Volloch V, Tsiftsoglou A, Gusella J, Parker D, Kernen J, Mitrani A, Weeks V, Witte O, Besmer P (1980) Control of proliferation and differentiation in cells transformed by Friend virus. Cold Spring Harbor Symp Quant Biol 44:1177–1185
Huebner RJ, Gilden RV, Toni R, Hill PR, Trimmer TW, Fish DC, Sass B (1976) Prevention of spontaneous leukemia in AKR mice by type-specific immunosuppression of endogenous virogenes. Proc Natl Acad Sci USA 73:4633–4635
Ihle JN, Lee JC (1982) Possible immunological mechanisms in C-type viral leukemogenesis in mice. Curr Top Microbiol Immunol 98:85–101
Ishimoto A, Adachi A, Sakai K, Yorifuji T, Tsuruta S (1981) Rapid emergence of mink cell focus-forming (MCF) virus in various mice infected with NB-tropic Friend virus. Virology 113: 644–655
Jaenisch R (1976) Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc Natl Acad Sci USA 73:1260–1264
Jaenisch R (1979) Moloney leukemia virus gene expression and gene amplification in preleukemic and leukemic BALB/Mo mice. Virology 93:80–90
Jaenisch R, Fan H, Croker B (1975) Infection of preimplantation mouse embryos and of newborn mice with leukemia virus: tissue distribution of viral DNA and RNA and leukemogenesis in the adult animal. Proc Natl Acad Sci USA 72:4008–4012
Jahner D, Stuhlmann H, Jaenisch R (1980) Conformation of free and of integrated Moloney leukemia virus proviral DNA in preleukemic and leukemic BALB/Mo mice. Virology 101: 111–123
Jolicoeur P (1979) The *Fv-1* gene of the mouse and its control of murine leukemia virus replication. Curr Top Microbiol Immunol 86:67–122

Jolicoeur P, Rosenberg N, Cotellessa A, Baltimore D (1978) Leukemogenicity of clonal isolates of murine leukemia viruses. J Natl Cancer Inst 60:1473–1476

Kabat D, Ruta M, Murray MJ, Polonoff E (1980) Immunoselection of mutants deficient in cell surface glycoproteins encoded by murine erythroleukemia viruses. Proc Natl Acad Sci USA 77:57–67

Kaplan HS (1967) On the natural history of the murine leukemias: presidential address. Cancer Res 27:1325–1340

Kaplan HS, Yaffe D (1960) Accelerated development of Strain Ak/n mouse lymphomas by cell-free tissue and tumor extracts. Proc Am Assoc Cancer Res 3:332

Kaplan PL, Topp WC, Ozanne B (1981) Simian virus 40 induces the production of a polypeptide transforming factor(s). Virology 108:484–490

Kawashima K, Ikeda H, Stockert E, Takahashi T, Old LJ (1976a) Age-related changes in cell surface antigens of preleukemic AKR thymocytes. J Exp Med 144:193–208

Kawashima K, Ikeda H, Hartley JW, Stockert E, Rowe WP, Old LJ (1976b) Changes in expression of murine leukemia virus antigens and production of xenotropic virus in the late preleukemic period in AKR mice. Proc Natl Acad Sci USA 73:4680–4684

Kelly M, Lung ML, Holland CA, Herr W, Corbin V, Chattopadhyay S, Lowy DR, Hopkins N (1983) A comparison of nucleotide sequences of the p15(E) gene and LTR of Akv and MCF 247 murine C-type viruses. J Virol 45:Jan issue

Kemp MC, Basak S, Compans RW (1979) Glycopeptides of murine leukemia viruses. I. Comparison of two ecotropic viruses. J Virol 31:1–7

Kemp MC, Famulari NG, O'Donnell PV, Compans RW (1980) Glycopeptides of murine leukemia viruses. II. Comparison of xenotropic and dualtropic viruses. J Virol 34:154–161

Kemp MC, Famulari NG, Compans RW (1981) Glycopeptides of murine leukemia viruses. III. Glycosylation of *env* precursor glycoproteins. J Virol 39:463–470

Lee JC, Ihle JN (1979) Mechanisms of C-type viral leukemogenesis I. Correlation of in vitro lymphocyte blastogenesis to viremia and leukemia. J Immunol 123:2354–2358

Lee JC, Ihle JN (1981a) Chronic immune stimulation is required for Moloney leukemia virus-induced lymphomas. Nature 289:407–409

Lee JC, Ihle JN (1981b) Increased responses to lymphokines are correlated with preleukemia in mice inoculated with Moloney leukemia virus. Proc Natl Acad Sci USA 78:7712–7716

Lenz J, Crowther R, Klimenko S, Haseltine W (1982) Molecular cloning of a highly leukemogenic, ecotropic retrovirus from an AKR mouse. J Virol 43:943–951

Liao SK, Axelrad AA (1975) Erythropoietin-independent erythroid colony formation in vitro by hemopoietic cells of mice injected with Friend virus. Int J Cancer 15:467–482

Lilly F, Pincus T (1973) Genetic control of murine viral leukemogenesis. Adv Cancer Res 17: 231–277

Lilly F, Steeves RA (1973) B-tropic Friend virus: A host range pseudotype of spleen focus-forming virus (SFFV). Virology 55:363–370

Lilly F, Duran-Reynals ML, Rowe WP (1975) Correlation of early murine leukemia virus titer and H-2 type with spontaneous leukemia in mice of the BALB/c × AKR cross: a genetic analysis. J Exp Med 141:882–889

Linemeyer DL, Ruscetti SK, Menke JG, Scolnick EM (1980) Recovery of biologically active spleen focus-forming virus from molecularly cloned spleen focus-forming virus pBR322 circular DNA by cotransfection with infectious type C retroviral DNA. J Virol 35:710–721

Linemeyer DL, Ruscetti SK, Scolnick EM, Evans LH, Duesberg PH (1981) Biological activity of the spleen focus-forming virus is encoded by a molecularly cloned subgenomic fragment of spleen focus-forming virus DNA. Proc Natl Acad Sci USA 78:1401–1405

Linemeyer DL, Menke JG, Ruscetti SK, Evans LH, Scolnick EM (1982) Envelope gene sequences which encode the gp52 protein of the spleen focus-forming virus are required for the induction of erythroid cell proliferation. J Virol 43:223–233

Lung ML, Hering C, Hartley JW, Rowe WP, Hopkins N (1980) Analysis of the genomes of mink cell focus-inducing murine type-C viruses: a progress report. Cold Spring Harbor Symp Quant Biol 44:1269–1274

Lung ML, Hartley JW, Rowe WP, Hopkins NH (1983) Large RNase T1-resistant oligonucleotides encoding p15(E) and the U3 region of the LTR distinguish two biological classes of MCF C-type viruses of inbred mice. J Virol 45:Jan issue

Lyles DS, McConnell KA (1981) Subcellular localization of the *env*-related glycoproteins in Friend erythroleukemia cells. J Virol 39:263–272

Lynch CJ (1954) The R.I.L. strain of mice: its relation to the leukemic AK strain and AKR substrains. J Natl Cancer Inst 15:161–176

MacDonald ME, Mak TW, Bernstein A (1980a) Erythroleukemia induction by replication-competent type C viruses cloned from the anemia- and polycythemia-inducing isolates of Friend leukemia virus. J Exp Med 151:1493–1503

MacDonald ME, Reynolds Jr FH, van de Ven WJM, Stephenson JR, Mak TW, Bernstein A (1980b) Anemia- and polycythemia-inducing isolates of Friend spleen focus-forming virus. Biological and molecular evidence for two distinct viral genomes. J Exp Med 151:1477–1492

Mayer A, Duran-Reynals ML, Lilly F (1978) *Fv-1* regulation of lymphoma development and of thymic, ecotropic and xenotropic MuLV expression in mice of the AKR/J × RF/J cross. Cell 15:429–435

McGrath MS, Weissman IL (1979) AKR leukemogenesis: identification and biological significance of thymic lymphoma receptors for AKR retroviruses. Cell 17:65–75

McGrath MS, Pillemer E, Kooishra D, Weissman IL (1980a) The role of MuLV receptors on T-lymphoma cells in lymphoma cell proliferation. Cont Top Mol Immunol 11:157–183

McGrath MS, Pillemer E, Weissman IL (1980b) Murine leukemogenesis: monoclonal antibodies to T-cell determinants arrest T lymphoma cell proliferation. Nature 285:259–261

Meier H, Taylor BA, Cherry M, Huebner RJ (1973) Host-gene control of type-C RNA tumor virus expression and tumorigenesis in inbred mice. Proc Natl Acad Sci USA 70:1450–1455

Mirand EA, Steeves RA, Ayila L, Grace JT Jr (1968) Spleen focus formation by polycythemic strains of Friend leukemia virus. Proc Soc Exp Biol Med 127:900–904

Niman HL, Elder JH (1980) Molecular dissection of Rauscher virus gp70 by using monoclonal antibodies: localization of acquired sequences of related envelope gene recombinants. Proc Natl Acad Sci USA 77:4524–4528

Nishizuka Y, Nakakuki K (1968) Acceleration of leukemogenesis in AKR mice by grafts, cell suspensions, and cell-free centrifugates of thymuses from preleukemic AKR donors. Int J Cancer 3:203–210

Nobis P, Jaenisch R (1980) Passive immunotherapy prevents expression of endogenous Moloney virus and amplification of proviral DNA in BALB/Mo mice. Proc Natl Acad Sci USA 77:3677–3681

Nowinski RC, Doyle T (1977) Cellular changes in the thymuses of preleukemic AKR mice: correlation with changes in the expression of murine leukemia viruses. Cell 12:341–353

Nowinski RC, Hays EF (1978) Oncogenicity of AKR endogenous leukemia viruses. J Virol 27:13–18

Nowinski RC, Brown M, Doyle T, Prentice RL (1979) Genetic and viral factors influencing the development of spontaneous leukemia in AKR mice. Virology 96:186–204

Obata Y, Ikeda H, Stockert E, Boyse EA (1975) Relation of $G_{IX}$ antigen of thymocytes to envelope glycoprotein of murine leukemia virus. J Exp Med 141:188–197

Obata Y, Stockert E, DeLeo AB, O'Donnell PV, Snyder HW Jr, Old LJ (1981) $G_{ERLD}$: a cell surface antigen of the mouse related to xenotropic MuLV defined by naturally occurring antibody and monoclonal antibody. J Exp Med 154:659–675

O'Donnell PV, Nowinski RC (1980) Serological analysis of antigenic determinants on the *env* gene products of AKR dualtropic (MCF) murine leukemia viruses. Virology 107:81–88

O'Donnell PV, Stockert E (1976) Induction of GIX antigen and Gross cell surface antigen after infection by ecotropic and xenotropic murine leukemia viruses in vitro. J Virol 20:545–554

O'Donnell PV, Stockert E, Obata Y, DeLeo AB, Old LJ (1980) Murine leukemia virus-related cell surface antigens as serological markers of AKR ecotropic, xenotropic and dualtropic viruses. Cold Spring Harbor Symp Quant Biol 44:1255–1263

O'Donnell PV, Stockert E, Obata Y, Old LJ (1981) Leukemogenic properties of AKR dualtropic (MCF) viruses: amplification of murine leukemia virus-related antigens on thymocytes and acceleration of leukemia development in AKR mice. Virology 112:548–563

O'Donnell PV, Nowinski RC, Stockert E (1982) Amplified expression of murine leukemia virus (MuLV)-coded antigens on thymocytes and leukemia cells of AKR mice after infection by dualtropic (MCF) MuLV. Virology 119:450–464

Old LJ, Stockert E (1977) Immunogenetics of cell surface antigens of mouse leukemia. Annu Rev Genet 11:127–160

Oliff AI, Hager GL, Chang EH, Scolnick EM, Chan HW, Lowy DR (1980) Transfection of molecularly cloned Friend murine leukemia virus DNA yields a highly leukemogenic helper-independent type C virus. J Virol 33:475–486

Oliff A, Ruscetti S, Douglass EC, Scolnick E (1981) Isolation of transplantable erythroleukemia cells from mice infected with helper-independent Friend murine leukemia virus. Blood 58: 244–254

Pedersen FS, Buchhagen DL, Chen CY, Hays EF, Haseltine WA (1980) Characterization of virus produced by a lymphoma induced by inoculation of AKR MCF 247 virus. J Virol 35:211–218

Pedersen FS, Crowther RL, Tenney DY, Reimold AM, Haseltine WA (1981) Novel leukemogenic retroviruses isolated from cell line derived from spontaneous AKR tumor. Nature 292:167–170

Pedersen FS, Crowther RL, Hays EF, Nowinski RC, Haseltine WA (1982) Structure of retroviral RNA's produced by cell lines derived from spontaneous lymphomas of AKR mice. J Virol 41:18–29

Peters RL, Hartley JW, Spahn GJ, Rabstein LS, Whitmire CE, Turner HC, Huebner RJ (1972) Prevalence of the group-specific (gs) antigen and infectious virus expressions of the murine c-type RNA viruses during the life span of BALB/c Cr mice. Int J Cancer 10:283–289

Pinter A, Honnen WJ, Tung J-S, O'Donnell PV, Hammerling U (1982) Structural domains of endogenous murine leukemia virus gp70s containing specific antigenic determinants defined by monoclonal antibodies. Virology 116:499–516

Quint W, Quax W, van der Putten H, Berns A (1981) Characterization of AKR murine leukemia virus sequences in AKR mouse substrains and structure of integrated recombinant genomes in tumor tissue. J Virol 39:1–10

Racevskis J, Koch G (1977) Viral protein synthesis in Friend erythroleukemia cell lines. J Virol 21:328–337

Racevskis J, Koch G (1978) Synthesis and processing of viral proteins in Friend erythroleukemia cell lines. Virology 87:354–365

Rapp UR, Todaro GJ (1978) Generation of oncogenic type C viruses: rapidly leukemogenic viruses derived from C3H mouse cells in vivo and in vitro. Proc Natl Acad Sci USA 75: 2468–2472

Reddy EP, Dunn CY, Aaronson SA (1980) Different lymphoid cell targets for transformation by replication-competent Moloney and Rauscher mouse leukemia viruses. Cell 19:663–669

Rein A (1982) Interference grouping of murine leukemia viruses: a distinct receptor for the MCF-recombinant viruses on mouse cells. Virology 120:251–257

Rommelaere J, Faller DV, Hopkins N (1978) Characterization and mapping of RNase T1-resistant oligonucleotides derived from the genomes of AKV and MCF murine leukemia viruses. Proc Natl Acad Sci USA 75:495–499

Rosenberg N, Baltimore D (1978) The effect of helper virus in Abelson virus-induced transformation of lymphoid cells. J Exp Med 147:1126–1141

Rosner M, Grinna LS, Robbins PW (1980) Differences in glycosylation patterns of closely related murine leukemia viruses. Proc Natl Acad Sci USA 77:67–71

Rowe WP (1973) Genetic factors in the natural history of murine leukemia virus infection: GHA Clowes Memorial Lecture. Cancer Res 33:3061–3068

Rowe WP, Pincus T (1972) Quantitative studies of naturally occurring murine leukemia virus infection of AKR mice. J Exp Med 135:429–436

Rowe WP, Cloyd MW, Hartley JW (1980) Status of the association of mink cell focus-inducing viruses with leukemogenesis. Cold Spring Harbor Symp Quant Biol 44:1265–1268

Rudali G, Duplan J-F, Latarjet R (1956) Latence des leucoses chez des souris injectees avec un extrait leucemique a-cellulaire AK. CR Soc Biol (Paris) 242:837–839

Ruscetti SK, Linemeyer D, Feild J, Troxler D, Scolnick EM (1979) Characterization of a protein found in cells infected with the spleen focus-forming virus that shares immunological cross-reactivity with the gp70 found in mink cell focus-inducing virus particles. J Virol 30:787–798

Ruscetti S, Troxler D, Linemeyer D, Scolnick EM (1980) Three laboratory strains of spleen focus-forming virus: comparison of their genomes and translation products. J Virol 33:140–151

Ruscetti S, Davis L, Feild J, Oliff A (1981a) Friend murine leukemia virus-induced leukemia is associated with the formation of mink cell focus-inducing viruses and is blocked in mice expressing endogenous mink cell focus-inducing xenotropic viral envelope genes. J Exp Med 154:907–920

Ruscetti SK, Feild JA, Scolnick EM (1981b) Polycythemia- and anemia-inducing strains of spleen focus-forming virus differ in post translational processing of envelope-related glycoproteins. Nature 294:663–665

Ruta M, Kabat D (1980) Plasma membrane glycoproteins encoded by cloned Rauscher and Friend spleen focus-forming viruses. J Virol 35:844–853

Ruta M, Clarke S, Boswell B, Kabat D (1982) Heterogeneous metabolism and subcellular localization of a potentially leukemogenic membrane glycoprotein encoded by Friend erythroleukemia virus. J Biol Chem 257:126–134

Schafer W, Schwarz H, Thiel H-J, Fischinger PJ, Bolognesi DP (1977) Properties of mouse leukemia viruses XIV. Prevention of spontaneous AKR leukemia by treatment with group-specific antibody against the major virus gp71 glycoprotein. Virology 83:207–210

Schultz AM, Ruscetti SK, Scolnick EM, Oroszlan S (1980) The *env*-gene of the spleen focus-forming virus lacks expression of p15(E) determinants. Virology 107:537–542

Schwarz H, Fischinger PJ, Ihle JN, Thiel H-J, Weiland F, Bolognesi DP, Schafer W (1979) Properties of mouse leukemia viruses XVI. Suppression of spontaneous fatal leukemias in AKR mice by treatment with broadly reacting antibody against the viral glycoprotein gp71. Virology 93:159–174

Schwarz H, Ihle JN, Wecher E, Fischinger PJ, Thiel H-J, Bolognesi DP, Schafer W (1981) Properties of mouse leukemia viruses XVII. Factors required for successful treatment of spontaneous AKR leukemia by antibodies against gp71. Virology 111:568–578

Shih TY, Weeks MO, Troxler DH, Coffin JM, Scolnick EM (1978) Mapping host range-specific oligonucleotides within genomes of the ecotropic and mink cell focus-inducing strains of Moloney murine leukemia virus. J Virol 26:71–83

Steeves RA (1975) Spleen focus-forming virus in Friend and Rauscher leukemia virus preparations. J Natl Cancer Inst 54:289–297

Steeves RA, Mirand EA (1969) Separation of members of the Friend virus complex by sucrose gradient centrifugation. Proc Am Assoc Cancer Res 10:86

Steeves RA, Eckner RJ, Mirand EA, Priore RL (1971) Rapid assay of murine leukemia virus helper activity for Friend spleen focus forming virus. J Natl Cancer Inst 46:1219–1228

Stephenson JR (1980) Molecular biology of RNA tumor viruses. Academic Press, New York

Stockert E, Old LJ, Boyse EA (1971) The $G_{IX}$ system: a cell surface allo-antigen associated with murine leukemia virus; implications regarding chromosomal integration of the viral genome. J Exp Med 133:1334–1355

Stockert E, DeLeo AB, O'Donnell PV, Obata Y, Old LJ (1979) $G_{AKSL2}$: a new cell surface antigen of the mouse related to the dualtropic mink cell focus-inducing class of murine leukemia virus detected by naturally occurring antibody. J Exp Med 149:200–215

Tambourin PE, Wendling F, Jasmin C, Smadja-Joffe F (1979) The physiopathology of Friend leukemia. Leuk Res 3:117–129

Troxler DH, Scolnick EM (1978) Rapid leukemia induced by cloned Friend strain of replicating murine type-C virus. Virology 85:17–27

Troxler DH, Boyers JK, Parks WP, Scolnick EM (1977a) Friend strain of spleen focus-forming virus: a recombinant between mouse type C ecotropic viral sequences and sequences related to xenotropic virus. J Virol 22:361–372

Troxler DH, Lowy D, Howk R, Young H, Scolnick EM (1977b) Friend strain of spleen focus-forming virus is a recombinant between ecotropic murine type C virus and the *env* gene region of xenotropic type C virus. Proc Natl Acad Sci USA 74:4671–4675

Troxler DH, Parks WP, Vass WC, Scolnick EM (1977c) Isolation of a fibroblast nonproducer cell line containing the Friend strain of the spleen focus-forming virus. Virology 76:602–615

Troxler DH, Yuan E, Linemeyer D, Ruscetti S, Scolnick EM (1978) Helper-independent mink cell focus-inducing strains of Friend murine type-C virus: potential relationship to the origin of replication-defective spleen focus-forming virus. J Exp Med 148:639–653

Troxler DH, Ruscetti SK, Linemeyer DL, Scolnick EM (1980) Helper-independent and replication-defective erythroblastosis-inducing viruses contained within anemia-inducing Friend virus complex (FV-A). Virology 102:28–45

Tung J-S, Fleissner E (1980) Amplified *env* and *gag* products on AKR cells. Origin from different murine leukemia virus genomes. J Exp Med 151:975–979

Tung J-S, Vitetta ES, Fleissner E, Boyse EA (1975) Biochemical evidence linking the $G_{IX}$ thymocyte surface antigen to the gp69/71 envelope glycoprotein of murine leukemia virus. J Exp Med 141:198–205

Tung J-S, Boyse EA, Shen F-W (1980) Influence of *Fv-1* alleles on cellular expression of gp70. J Exp Med 151:980–983

van der Putten H, Quint W, van Raiij J, Maandag ER, Verma IM, Berns A (1981) M-MuLV-induced leukemogenesis and structure of recombinant proviruses in tumors. Cell 24:729–739

Van Griensven LJ, Vogt M (1980) Rauscher "mink cell focus-inducing" (MCF) virus causes erythroleukemia in mice: its isolation and properties. Virology 101:376–388

Vogt M (1979) Properties of "mink cell focus-inducing" (MCF) virus isolated from spontaneous lymphoma lines of BALB/c mice carrying Moloney leukemia virus as an endogenous virus. Virology 93:226–236

Weissman IL, McGrath MS (1982) Retrovirus lymphomagenesis: relationship of normal immune receptors to malignant cell proliferation. Curr Top Microbiol Immunol 98:103–112

Witte ON, Wirth DF (1979) Structure of murine leukemia virus envelope glycoprotein precursor. J Virol 29:735–743

# Avian Acute Leukemia Viruses

MICHAEL J. HAYMAN*



## 1 Introduction

On the basis of their oncogenic properties, avian retroviruses can be divided into three groups: (a) sarcoma viruses (ASV), which cause solid tumors of the connective tissue after a short latency period; (b) lymphatic leukemia viruses (LLV), which cause a B-cell lymphoma only after a long latency period of approximately 6 months or more; and (c) acute leukemia viruses (ALV), which upon infection of young chickens rapidly bring about the death of the birds from a variety of different neoplasms, usually, but not exclusively, of the hematopoietic system (*Purchase* and *Burmester* 1973; *Hanafusa* 1977; *Graf* and *Beug* 1978). This classification is not absolute, in that there are overlaps in the oncogenic spectra of these viruses. For example, certain LLV have been shown to cause erythroblastosis and occasionally sarcomas (*Purchase* and *Burmester* 1973); similarly, certain ALV can also cause sarcomas (*Rothe-Meyer* and *Engelbreth-Holm* 1933; *Graf* et al. 1977). However, by using this classification the seven field isolates of avian retroviruses listed in Table 1 have been grouped together as acute leukemia viruses. There is another isolate, reticuloendothelosis virus (REV-T), which is, by this definition, an ALV; however, since it is not related to the classical avian leukosis/sarcoma complex as judged by nucleic acid homology, it will not be included here. Interested readers are referred to *Witter* (1978) and *Graf* and *Stehelin* (1982). There are numerous reports in the literature regarding the isolation of other ALV; unfortunately, these strains are no longer available (see *Graf* and

* Viral Leukaemogenesis Laboratory, Imperial Cancer Research Fund, P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, England

*Beug* 1978). Consequently, it is only the viruses listed in Table 1 which will be considered in this review.

The general strategy of infection by ALV is thought to be identical to that of the other avian retroviruses. Since this has been the subject of a number of excellent reviews over the last few years (*Hanafusa* 1977; *Vogt* 1977; *Beemon* 1978; *Bishop* 1978; *Coffin* 1979) and is covered in great detail in a recent book (eds. *Weiss* et al. 1982), readers are referred to these texts for general information regarding the principal features of the genetic and molecular biology of retroviruses. However, the following is a summary of certain key features which are pertinent to an understanding of the ALV.

Retroviruses contain diploid genomes composed of two identical single-stranded RNA subunits, the size which varies from 5 to 10 kilobases depending on the virus in question. The nondefective LLV have a subunit size of approximately 8.5 kilobases and contain three known genes. These are *gag*, which codes for a precursor protein of 76 000 daltons that is subsequently cleaved into the four internal nonglycosylated core proteins of the virion (*Vogt* and *Eisenman* 1973; *Eisenman* et al. 1974; *Vogt* et al. 1975; *Eisenman* and *Vogt* 1978); *pol*, which codes for the virion RNA-dependent DNA polymerase or reverse transcriptase; and *env*, which codes for the viral envelope glycoproteins, gp85 and gp37. The order of these three genes is 5′-*gag-pol-env-poly*A-3′. ASV contain an additional gene which, in the case of Rous sarcoma virus (RSV), is termed *src* and is inserted between the *env* gene and the 3′ end of the RNA. This gene appears to have been acquired by recombination with cellular sequences, since related sequences can be found in all uninfected vertebrate cellular DNA (*Stehelin* et al. 1976a, b). Genetic evidence indicates that the acquisition of this gene confers on the virus the ability to cause sarcomas. The other ASV have also acquired genes related to cellular sequences, but in these cases at the expense of viral replicative functions. They are therefore referred to as defective sarcoma viruses and they are reviewed by *Neil* in this volume. With this as background, I would now like to review what we know about avian ALV, emphasizing information that is pertinent to the mechanism by which ALV transform cells. In doing so my principal objective is to illustrate how, by studying retroviruses in general and ALV in particular, it should be possible to gain insight into the mechanisms of cellular growth control and differentiation.

## 2 Transformation by Avian Acute Leukemia Viruses (ALV)

The seven strains of ALV to be covered are all field isolates which have been identified in different countries over the past 50 years or so (Table 1). Upon infection of young chickens, these viruses rapidly bring about the death of the birds from a variety of different neoplasms, usually, but not exclusively, of the hematopoietic system (*Graf* and *Beug* 1978). The predominant neoplasms that these viruses cause are also listed in Table 1. Experimental manipulation of these viruses has been greatly facilitated, firstly, by the ability of these viruses (with the exception of AMV) to transform fibroblasts in tissue culture (*Langlois* et al. 1969; *Ishizaki* and *Shimizu* 1970; *Graf* 1973; *Graf* et al. 1976) and, secondly, by the establishment for all seven strains of an in vitro assay for the transformation of chicken bone marrow cells (*Baluda* and *Goetz* 1961; *Graf* 1975; *Graf* et al. 1979). These assays allow the isolation of colonies of transformed cells and thus provide pure populations for analysis.

Table 1. Avian acute leukemia virus strains

| Virus strain | Predominant neoplasm(s)[a] | Country and year when isolated |
|---|---|---|
| AEV[b] | Erythroblastosis, sarcomas | Denmark 1931 |
| AMV | Myeloblastosis | USA 1941 |
| E26 | Myeloblastosis, erythroblastosis | Bulgaria 1962 |
| MC29 | Myelocytomatosis, liver and kidney tumors[c] | Bulgaria 1964 |
| CMII | Myelocytomatosis | Federal Republic of Germany 1964 |
| MH-2 | Liver and kidney tumors[c] | England 1927 |
| OK10 | Liver and kidney tumors[c] | Finland 1975 |

[a] Only those neoplasms induced in chickens are listed here, and it should be borne in mind that it is highly likely that the host will have some influence over the disease produced; [b] Recently two other viruses than can cause erythroid leukemia have been isolated, AEV-H (*S. Kawai*, personal communication) and S13. (*S. Benedict, H. Beug, P.K. Vogt,* personal communication), but as these two viruses have yet to be completely characterized, they will not considered here; [c] These tumors have been referred to as either endotheliomas or carcinomas

## 2.1 Transformation of Avian Fibroblasts

All the ALV, except AMV, are able to transform avian fibroblasts, although E26 has so far only been shown to be capable of transforming quail fibroblasts (*Graf* et al. 1979). The best studies on the analysis of the transformed cells are those in which cloned chicken fibroblasts were used (*Royer-Pokora* et al. 1979; *Beug* et al. 1979). The use of cloned fibroblasts is preferable to chicken embryo cells, as the primary embryo cells are a mixture of cell types, making comparisons of the transformed cells unreliable. Indeed, AMV was initially reported to transform fibroblasts, but the transformed cells observed were later shown to be contaminating macrophages in the embryo cultures (*Moscovici* et al. 1969; *Graf* 1973). Studies involving cloned fibroblasts have been performed mainly with RSV, AEV, and MC29. These studies have shown that, both phenotypically and with regard to certain parameters of transformation, it is possible to distinguish the virus-induced transformed cells (*Royer-Pokora* et al. 1978). For example, although cloned chicken cells transformed by all three viruses grow in agar, only MC29-transformed cells display an increased growth rate and only AEV- and RSV-transformed cells exhibit increased hexose uptake and lack fibronectin on their surface (*Royer-Pokora* et al. 1979). The morphology of the transformed cells varied widely. RSV-transformed cells were rounded, AEV-transformed cells exhibited a spindly refractile morphology, and MC29-transformed cells were somewhat small and epithelioid in shape with prominent nuclei. These data suggest that these three different viruses transform fibroblasts by means of different mechanisms, since in each case the expressed parameters of transformation are different.

## 2.2 Transformation of Avian Hematopoietic Cells

The hallmark of ALV is their ability to transform hematopoietic cells in vitro. The in vitro assays usually use bone marrow cells as a source of target cells, although it is possible to

Table 2. Cells transformed in vitro by acute leukemia viruses

| Virus strain | Hematopoietic cells | Fibroblasts |
|---|---|---|
| AEV | Erythroblasts | Yes[a] |
| AMV | Myeloblasts | No |
| E26 | Myeloblasts plus erythroid cells | Yes |
| MC29 | Immature macrophages | Yes[a] |
| CMII | Immature macrophages | Yes |
| OK10 | Immature macrophages | Yes |
| MH-2 | Immature macrophages | Yes |

[a] Also capable of transforming mammalian fibroblasts

use other cells, for example, spleen cells or yolk sac cells (*Moscovici* et al. 1975). Bone marrow is the cell source of choice, since it contains target cells for transformation by all the ALV tested to date, unlike certain of the other cell sources (*Graf* et al. 1980). The in vitro-transformed hematopoietic cells have been characterized with regard to their differentiation phenotype by both biochemical and immunological techniques (*Beug* et al. 1979) and compared with leukemic cells obtained in vivo.

Hematopoietic cells transformed by AEV, both in vivo and in vitro, were found to belong to the erythroid lineage. AMV-transformed cells resembled myeloblasts, and MC29-type viruses resembled macrophages (*Beug* et al. 1979). Cells transformed by E26 either are myeloblasts or cells of the erythroid lineage, depending on the assay conditions (*Radke* et al. 1982), a point that will be considered in greater detail later in this review. The main points that can be derived from these data are that the bone marrow assay seems to provide a faithful in vitro correlate of the in vivo disease pattern, and that the seven viruses can be divided into three distinct groups on the basis of the differentiation phenotype of the cells they transform in vitro (Table 2). On the basis of these data it was postulated that the ALV contain three different types of oncogenes, which were named after the cells they transform (*Beug* et al. 1979; *Graf* et al. 1980). This hypothesis was confirmed by analysis of the viral genomes (see Sect. 3).

## 2.3 Transformation of Mammalian Fibroblasts

Certain strains of RSV can transform mammalian fibroblasts, in addition to avian fibroblasts. Such transformed cells have proved very useful for the isolation of viral mutants and the production of antiserum against the transforming protein of RSV, pp60$^{src}$. They have also provided an excellent system for studying eukaryotic gene expression and virus-cell interactions. These results prompted investigations into the ability of ALV to transform mammalian fibroblasts. AEV and MC29 have both been found capable of transforming rat fibroblasts following direct infection (*Quade* 1979), and NIH-3T3 fibroblasts following transfection (*Copeland* and *Cooper* 1980). The presence of AEV and MC29 in these cells was demonstrated by virus rescue following fusion with chick fibroblasts and, in the case of the transformed rat cells, the rescued virus was shown to be able to transform the appropriate bone marrow cells (*Quade* 1979). In addition, the presence

and expression of the viral genomes in the rats cells could be demonstrated (*Quade* et al. 1981), and reversion of the transformed phenotype in the AEV-transformed cells correlated with reduced expression of the AEV provirus (*Quade* et al. 1981). The transformed rat cells have been characterized with regard to parameters of transformation expressed and have been found to display the same characteristic set as transformed avian fibroblasts, with the notable exception that the MC29-transformed cells cause sarcomas (K. Quade, personal communication). Recently, it has proven possible to raise antisera against the transforming protein of AEV in rats bearing tumors induced by AEV rat cells (*Hayman* et al. 1983). These results will be considered in greater detail in Sect. 3.1. There have been no reports of any ALV-transforming mammalian hematopoietic cells.

## 3 Genetic Structure, Expression, and Protein Products of ALV

Initial attempts to elucidate the structures of the genomes of ALV utilized the techniques of liquid hybridization, heteroduplex mapping, and oligonucleotide mapping (*Roussel* et al. 1979; *Bister* and *Duesberg* 1980; *Lai* et al. 1979; *Hu* et al. 1979; *Stehlin* et al. 1980). These studies were in the main very successful, and the summation of the data obtained is shown in Figs. 1–3. With the recent advent of molecular cloning it has been possible to isolate cloned DNA copies of several members of the ALV, and these are presently being sequenced (*Souza* et al. 1980; *Vennström* et al. 1980, 1981; *Lautenberger* et al. 1981). Since there are interesting features to each of the ALV genomes, I will consider them separately. However, it is appropriate to point out that the single important feature of their genetic structure is that they have all acquired sequences which were originally derived from cellular DNA. It is these new sequences which, in all likelihood, are responsible for the transforming ability of the virus. Hence, they are known as viral oncogenes, or v-*onc* for short, the cellular counterparts being referred to as c-*onc*. The seven isolates can be divided into three groups on the basis of the v-*onc* sequences in their genomes. These three sequences are known as v-*erb* in the case of avian erythroblastosis virus, v-*myb* in the case of AMV and E26, and v-*myc*, which is found in the genomes of MC29, CMII, MH-2, and OK10. Let us now consider these viruses separately with the emphasis on how the viruses transform cells.

### 3.1 Avian Erythroblastosis Virus

The genome structure of AEV is in some ways typical of that of all the ALV and, for that matter, also the defective sarcoma viruses (*Neil,* this volume). The v-*onc* sequence has been acquired in the middle of the genome of a LLV at the expense of viral replicative information, in this case all of the *pol* gene, and most of the *gag* and *env* genes (Fig. 1). The v-*erb* insert is, however, unique in that it actually gives rise to two gene products; therefore, the gene is divided into two separate regions known as *erb A* and *erb B*. The *erb A* gene product is a fusion protein of the partial *gag* gene and the *erb A* region and has a molecular weight of 75 000 daltons (p75) (*Hayman* et al. 1979b). As we will see, the expression of the v-*onc* genes as *gag* fusion proteins is the most common mode of expression used by the ALV. p75 is synthesized from a genomic length messenger RNA of approximately 5.3 kb. The *erb B* gene product is synthesized from a spliced subgenomic mRNA

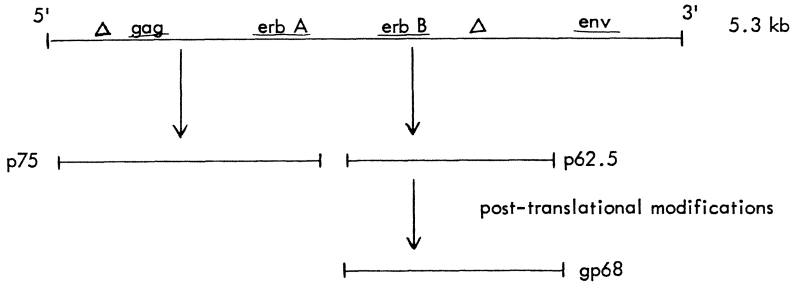



Fig. 1. Genome structure and gene product of avian erythroblastosis virus (not to scale)

of 3.5 kb. When immunoprecipitated with tumor-bearing rat sera, a doublet of 66 000 and 68 000 daltons (p66 and p68) is seen. Pulse chase experiments show that p66 can be chased into the p68 form (*Hayman* et al. 1983). These two proteins are identical by tryptic peptide mapping, and although p68 appears to be phosphorylated whereas p66 is not, it is unknown as yet whether this modification is sufficient to account for the difference in molecular weight. It should be mentioned that the *erb B* gene product was originally described as a 40 000-dalton protein on the basis of in vitro translation studies (*Lai* et al. 1980; *Pawson* and *Martin* 1980; *Yoshida* and *Toyoshima* 1980). However, this protein seems to be the result of an internal initiation event, and the true in vitro translation product is, in fact, a protein of 62 500 daltons (*Privalsky* and *Bishop* 1982; *Hayman* et al. 1983). Additional modifications have been identified which convert p62.5 into the p66 and p68 form and appear to involve glycosylation as judged by endoglycosidase H experiments (*Hayman* et al. 1983). Subcellular fractionation experiments have located p66 and p68 in cell membranes, and immunofluorescence studies on live cells indicate that they are expressed on the cell surface (*Hayman* et al. 1983).

AEV is capable of transforming two cell types, fibroblasts and erythroblasts, and causing two distinct neoplasms, sarcoma and erythroid leukemia. Which of these two proteins is responsible for transformation, or are both necessary? At the present time, there is no clear-cut answer to this question. However, the best clues come from experiments involving viral mutants. Conditional mutants of AEV have been isolated, and there are now five available (*Graf* et al. 1978; *Palmieri* et al. 1982). Transformation studies have shown that these mutants are temperature sensitive for both erythroblast and fibroblast transformation. Since it is highly unlikely these are all double mutants, these data imply that a mutation in one gene affects both types of transformation. This effectively rules out the possibility that there is one gene for fibroblast transformation and one for erythroblast transformation, and suggests that either one gene is responsible for both types of transformation or that both genes are necessary for transformation. Studies with a nonconditional mutant, *td* 359 AEV, which can transform fibroblasts but not erythroblasts and causes sarcomas but not erythroblastosis, would indicate that the two types of transformation can be dissociated. However, it is still possible that the 'dose' of biologically active protein required to transform erythroblasts is higher than that needed to transform fibroblasts. The mutations in *td* 359 could have reduced the biological activity of the transforming protein, such that it is no longer sufficient for erythroid transformation but is still sufficient for fibroblast transformation. Whereas attempts to map the conditional mutants have proved unsuccessful, the mutation in *td* 359 was originally

reported to be in p75 (*Beug* et al. 1980). However, with the availability of antiserum against the *erb B* protein, it has been possible to show that there is also a deletion in the *erb B* gene product of approximately 10 000 daltons. These experiments are, therefore, not very helpful in assigning various transforming abilities within the *erb* locus. The most useful approach seems to be site-specific mutagenesis of cloned AEV DNA, and the results of such experiments are now being reported. Deletions within the *erb A* locus do not appear to remove the fibroblast- or erythroblast-transforming ability of the virus, whereas deletions within the *erb B* do (*Frykberg* et al. 1983).

These data imply that the *erb B* gene product is sufficient for transformation. However, it should be mentioned that the virus with *erb A* deletions is not as efficient at transformation as wild-type virus. Therefore, although *erb B* is sufficient for transformation, an ancillary role for *erb A* is still likely.

Taken together, these data imply that a cell surface membrane glycoprotein (p68) is responsible for the proliferation of erythroid cells by AEV. In this regard, it is interesting to note that two other viruses that transform erythroid cells, namely, S13 virus (S. Benedict and P.K. Vogt, personal communication) and the murine Friend leukemia virus (*Linemeyer* et al. 1982), also appear to do so via glycoproteins. Whether these membrane glycoproteins exert their oncogenic effects as receptors for growth factors, or via some other mechanism, is presently unknown. However, it is obvious that, having identified a membrane glycoprotein which has such a profound effect on cell proliferation, an important step has been taken towards defining erythroid differentiation and growth control in general.

## 3.2 Avian Myeloblastosis Virus and E26

Although avian myeloblastosis virus and E26 are often considered together, since they both contain the oncogene v-*myb*, hybridization studies using a probe derived from AMV (*Roussel* et al. 1979) show that there is only 66% homology between their *myb* genes. Both viruses can transform myeloid cells in vitro and the cells are essentially indistinguishable by a variety of biochemical, immunological, and functional markers (*Beug* et al. 1979). However, the similarities stop there, since E26 is able to transform quail fibroblasts and chick cells of the erythroid lineage (*Sotirov* 1981; *Graf* et al. 1979; *Radke* et al. 1982), whereas AMV is not known to transform any other cell type. Since their genome structures and their modes of expression of v-*myb* are drastically different, I feel it is appropriate to consider them separately.

The genome structure of AMV is depicted in Fig. 2, where it can be seen that the v-*myb* gene has essentially taken the place of the *env* gene of a LLV. The AMV genome has been cloned (*Souza* et al. 1980), and sequence analysis of cloned DNA has revealed that the *pol* gene has also sustained a deletion of approximately 30 bases (*Rushlow* et al. 1982; J.M. Bishop, personal communication). These two independent groups have also been able to show that the 5′ end of the v-*myb* gene appears to start with noncoding sequences that are derived from an intron found in the c-*myb* gene, implying that the recombinational event that gave rise to AMV took place at the level of DNA. Analysis of intracellular RNA has shown that the v-*myb* gene is expressed as a subgenomic 21S mRNA (*Chen* 1980; *Gonda* et al. 1981), and although no protein product has yet been identified, the sequence analysis has identified an open reading frame that predicts a protein

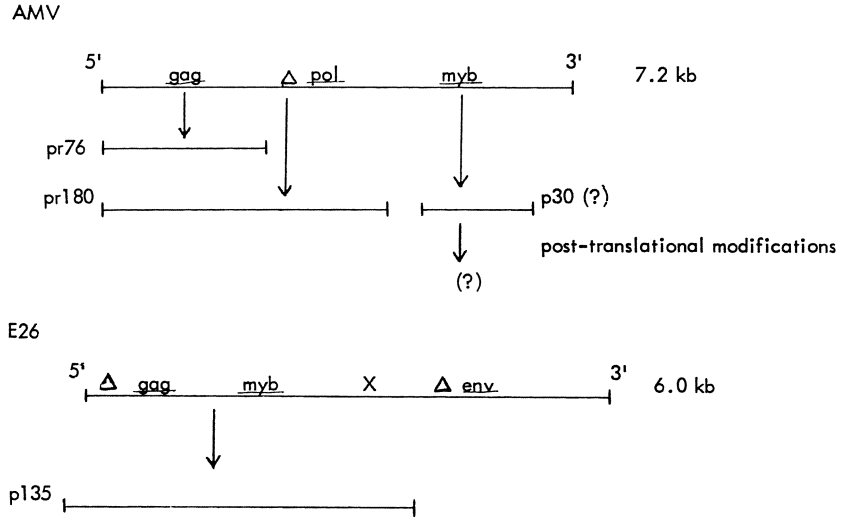



Fig. 2. Genome structure and gene products of avian myeloblastosis virus and E26 (not to scale). (?) indicates the gene product has still not been definitely characterized. X is an unknown sequence present in the E26 genome which is presumed to be also contained in p135

of approximately 30 000 daltons containing no viral structural gene information. The actual size of the in vivo protein could, of course, be larger if there were posttranslational modifications, as was found in the case of the AEV *erb B* protein or if an RNA splicing event expanded the open reading frame.

The genome structure of E26 depicted in Fig. 2 is based on circumstantial evidence gathered from a number of observations (*Roussel* et al. 1979; *Bister* et al. 1982a; *Beug* et al. 1982b) and until a virus DNA copy is cloned the exact structure will remain unknown. However, certain salient features of the genome structure are clear. Firstly, E26 encodes a p135 *gag* fusion protein which must contain *myb* information (*Bister* et al. 1982a; *Beug* et al. 1982b). Whether, in addition, it synthesizes a product from a subgenomic mRNA containing only *myb* sequences is unknown. Also, given that the size of the E26 genome is 5.7 kb (*Bister* et al. 1982a), sequences other than those of *myb*, the partial *gag*, and *env* genes can easily by accommodated. These differences in the genome structure and mode of expression of the *myb* sequences may well be relevant to the differences in the transforming potential of E26 compared to AMV. It is not clear if this difference in the expression of *myb* allows E26 to transform fibroblasts and erythroid cells, or if as yet uncharacterized sequences in the E26 genome are responsible for this phenomenon. Answers await the cloning of the E26 genome. Whatever the result, E26 is an intriguing virus. It either harbors two different v-*oncs* in its genome or, by expressing v-*myb* as a *gag*-related polyprotein it has widened its oncogenic potential. It becomes even more interesting when one considers the intriguing growth properties of the cells it transforms.

One of the major factors that has inhibited the experimental dissection of the E26 genome has been the inability to grow E26-transformed erythroid or myeloid cells easily (M.J. Hayman; T. Graf and H. Beug, frustrating observations). Recently the reason for this has become clear. The E26-transformed cells are still dependent for growth on cellu-

lar growth factors. Either erythropoietin-like factors for the erythroid cells (H. Beug, personal communication) or conditioned medium from concanavalin-A-stimulated spleen cells (*Beug* et al. 1982b) are required for growth. These data raise the intriguing possibility that the E26 oncogene product requires these growth factors to function and, therefore, may itself be a receptor for such factors. Determining if this is indeed the case is one of the most interesting areas of research presently underway, and the answers are eagerly awaited.

## 3.3 Myelocytomatosis Virus Strain, MC29, and Related Viruses

The avian myelocytomatosis virus, MC29, has been shown to contain an oncogene, v-*myc*, which is centrally located in the 5.5-kb genomic RNA, as depicted in Fig. 3. This same sequence, or a highly related one, has also been found in three other independently isolated ALV, namely CMII, MH-2, and OK10. Its location within the genome of these other viruses is also shown in Fig. 3. As MC29 is the best studied of these four viruses, we will consider it first, and then address the interesting features of the other group members.

MC29 transforms fibroblasts and macrophages in culture, and the transformed cells contain only a single mRNA of genomic length (*Sheiness* et al. 1981). The only gene

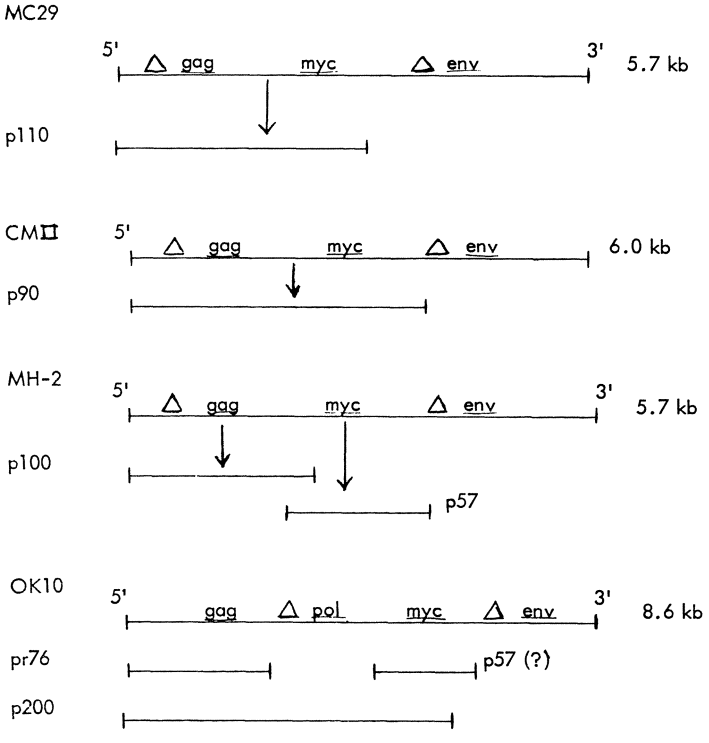



Fig. 3. Genome structure and products of MC29-type viruses (not to scale). (?) indicates the gene product has still not been definitely characterized

product found in such cells is a *gag-myc* polyprotein of 110 000 molecular weight (p110) (*Bister* et al. 1977; *Kitchener* and *Hayman* 1980). This protein, therefore, was considered the candidate to be the transforming protein for both fibroblasts and macrophages. Studies with partially transformation-defective mutants, which are severely hindered in their ability to transform macrophages in culture whilst still able to transform fibroblasts (*Ramsay* et al. 1980), confirmed this. These mutants contained deletions within the *myc* gene, giving rise to smaller proteins which had lost specific phosphorylation sites in the *myc* region (*Ramsay* et al. 1980, 1982b; *Enrietto* and *Hayman* 1982; *Ramsay* and *Hayman* 1982; *Bister* et al. 1982b). Furthermore, analysis of these mutants in vivo showed they no longer caused the characteristic liver and kidney tumors associated with wild-type MC29 (*Hayman* et al. 1982b; *Enrietto* et al. 1983). These data confirm that the *myc* gene is responsible for transformation by MC29. However, they leave unanswered the question of why the mutants still transform fibroblasts in vitro without causing a neoplasm in vivo relating to this property. Various arguments suggesting dose response differences between fibroblasts and macrophages have been proposed (*Ramsay* et al. 1980), but so far they remain untested.

Recently a back mutant of one of the partially transformation-defective viruses, termed MC29-HB1, has been isolated. This back mutant has regained the ability to transform macrophages in culture, and analysis of the viral genome revealed that it had regained the *myc* sequence lost from the original mutant (*Ramsay* et al. 1982a). In addition, it had also acquired a different LTR and had minor changes in the *gag* gene region. These latter changes were presumed to be due to recombination with the helper virus with which it was propagated, whereas the *myc* sequences had presumably been acquired by recombination with the c-*myc* gene. One of the most interesting aspects of the HB1 virus is its pathology, since when it is injected into chickens, it does not cause the disease associated with wild-type MC29. Rather, it causes predominantly lymphomas and occasional myelocytomas (*Hayman* et al. 1982b). The lymphomas do not appear to be the same as those caused by lymphoid leukosis viruses, as they do not involve the bursa and appear as well in bursectomized birds (L.N. Payne, P.J. Enrietto and M.J. Hayman, unpublished observations). This new disease pattern is very interesting since it has recently been shown that the majority of the LLV-induced lymphomas are caused by insertional mutagenesis in the vicinity of the c-*myc* gene (*Neel* et al. 1981; *Hayward* et al. 1981; *Payne* et al. 1981, 1982; for recent review see *Varmus* (1982). It has always been an intriguing fact that the tumor in which elevated c-*myc* expression is observed is not one commonly associated with those caused by v-*myc*. The observation that in HB1 induces a novel spectrum of disease suggests that very subtle differences in the viral genome have very significant effects on the pathology of the virus. It is also worth noting that host factors can affect the disease pattern, as recently shown for MC29 and MH-2, where MC29 was virtually non-pathogenic in quail under conditions where MH-2 was highly pathogenic (*Linial* 1982).

This leads to considerations of how MC29 transforms cells. The most informative data on this come from recent experiments which loate the p110 protein within the cell nucleus (*Abrams* et al. 1982; *Donner* et al. 1982). This would imply some mechanism involving perhaps an interaction with the chromatin. In this regard it is worth noting that in DNA transfection studies on bursal tumors which had elevated c-*myc* expression, although the tumors contain activated transforming genes as evidenced by assay on NIH-3T3 cells, the transforming DNA transfected did not contain c-*myc* or any viral sequences (*Cooper* and *Neiman* 1980, 1981). These authors postulated that the c-*myc* enhancement

was a primary stage in tumorigenesis, with the actual transforming gene being distinct from c-*myc* but being turned on somehow by c-*myc*. The finding that v-*myc* seems to act in the nucleus might lend credence to this hypothesis. Obviously, though, we are still a long way from understanding the underlying mechanism of *myc* – induced oncogenesis.

Of the other viruses in this group, CMII can probably be considered to be very similar to MC29. The v-*myc* sequence is again centrally located within the genome (*Bister* et al. 1979), and a *gag-myc* protein is encoded with a molecular weight of 90 000 (*Hayman* et al. 1979a). As this protein appears to be the only one found in transformed cells, it presumably is the transforming protein. The two other members of this group are, however, more interesting. MH-2 was originally reported to direct the synthesis of a 100 000-molecular weight *gag*-related protein (*Hu* et al. 1978), and it was presumed that this was a *gag-myc* fusion protein. However, recent data suggest the presence of a subgenomic mRNA in MH-2-transformed cells (*Pachl* et al. 1982; *Graf* and *Stehelin* 1982) which could encode a *myc* gene product lacking *gag* sequences; this raises the possibility that p100 is not the transforming protein. p100 may not even contain v-*myc* information, a possibility given some weight by the recent finding that p100 does not contain the v-*myc* phosphopeptides found in CMII p90 and MC29 p110 (*Ramsay* et al. 1982b). This difference in expression of the v-*myc* sequence in MH-2 may have some bearing on the pathology of the virus (*Linial* 1982) as discussed above. The last member of this group, OK10, is somewhat similar to MH-2, in that although it makes a fusion protein, p200 (*Ramsay* and *Hayman* 1980), it also synthesizes a subgenomic mRNA that could encode a v-*myc* protein (*Chiswell* et al. 1981; *Saule* et al. 1982). The genome structure of OK10 is somewhat different from the members of the MC29 group, in that the v-*myc* sequence is inserted between *pol* and *env* (*Ramsay* and *Hayman* 1980; *Bister* et al. 1980). Thus, in addition to the p200 *gag-pol-myc* protein, this virus contains a complete *gag* gene protein (*Ramsay* and *Hayman* 1980). Whether these unique features of the OK10 genome have any effect on its oncogenic potential is unknown, as the required studies, namely, the isolation of various mutants, have yet to be performed. Also, it is worth mentioning that due to the absence of a v-*myc*-specific antiserum no v-*myc* product from the subgenomic mRNA has been found in either MH-2 or OK10-transformed cells; consequently it is not known if this protein is also located in the nucleus.

These different modes of expression of the v-*myc* information within the MC29 group of viruses serve to highlight one of the features of ALV-induced transformation. On the one hand, these viruses were originally grouped together on the basis of in vitro transformation assay similarities which later were shown to correlate with the presence of the v-*myc* sequence. On the other hand, subtle differences in the v-*myc* sequence lead to marked changes in the virus pathology, as evidenced by the partial transformation-defective mutants and HB1. Therefore, there are obviously subtleties involved in the function and expression of the *myc* sequence, both viral and cellular, which have a dramatic effect on transformation. Only when we completely understand how the *myc* gene functions will these subtle differences be revealed.

## 4 Use of ALV to Study Cellular Differentiation

One of the most fascinating features of ALV is their ability to transform only specific cell types; for example, AEV transforms erythroid cells in bone marrow, whereas AMV

transforms myeloblasts. Since this specific transformation is a function of the viral oncogene, this has led to the hypothesis that these oncogenes may play a role in normal differentiation. Whether this will turn out to be the case obviously awaits further experimentation. However, the finding that the cellular oncogenes are expressed in cells other than those which are transformed by the appropriate v-*onc*-containing virus (*Gonda* et al. 1982; *Graf* and *Stehelin* 1982) means that the earlier hypotheses were probably too simplistic. Yet the concept remains an important one and may, in fact, turn out to be true for certain specific oncogenes. The ALV are proving to be invaluable tools for studying not only leukemogenesis, but also cellular differentiation. This has recently been the subject of a specific review, and the interested reader is referred to this for more details (*Beug* et al. 1982a). However, it is worthwhile here to outline the salient features of the system in order to point out what advantages the ALV provide over other systems. For this purpose I will use the AEV erythroid cell transformation system as an example, as this is the best studied.

One of the major problems is studying the regulation of differentiating erythroid cells is obtaining enough cells from the early stages to analyze. The ability to transform cells in vitro with AEV and isolate large numbers of clonal populations of erythroblasts circumvents this problem. These cells represent a cell population that normally constitute less than 1% of bone marrow cells; by a variety of different parameters the cells appear to be identical to the normal cells (*Beug* et al. 1979). Furthermore, virus mutants are available which are temperature-sensitive for transformation (*Graf* et al. 1978; *Palmieri* et al. 1982), and it has been shown that when the cells are grown at 41.5 °C, the virus is inactivated and the cells differentiated (*Graf* et al. 1978). The important features of this differentiation are: (a) it is of a clonally pure population of cells; (b) the induction is by growth at 41.5 °C, the normal body temperature of a chicken, and not following the addition of nonphysiological chemicals; and (c) the differentiation is still under the control of normal cellular regulatory mechanisms, such as "erythropoietin-like" molecules (*Beug* et al. 1982a). Therefore, this system provides an excellent model to study the regulatory processes involved in erythropoiesis, and any key observations that are made can then be confirmed in normal erythroid differentiation. This system has already been exploited to produce monoclonal antibodies against cell surface molecules (*Hayman* et al. 1982a), to identify membrane glycoproteins that change on differentiation (*Savin* and *Beug* 1981), to study the regulation of globin gene expression (*Weintraub* et al. 1982), and I am sure that many more exciting features remain to be discovered.

## 5 Summary and Concluding Remarks

As stated at the outset, my intention in this review was to consider how studies on the ALV will help us understand normal growth control and even certain aspects of differentiation. They represent a useful model system for studies of growth control, since they use ostensibly cellular genes, c-*oncs,* under viral control of expression as v-*oncs,* to bring about cell transformation. Whether the cellular genes have in some way been mutated to make them oncogenic or whether it is the abundance of the virus-controlled gene product that leads to transformation is a problem that still needs clarification; the answer may indeed be different for different oncogenes. Nevertheless, an understanding of how

the various ALV v-oncogenes function will be a large step forward towards understanding many of the mechanisms of normal growth control.

In considering where within the cell interactions might take place that would be important for regulating the growth and differentiation of that cell, the major sites would be the plasma membrane, the cytoplasm, and the nucleus. Interactions with the external environment via receptors in the plasma membrane would provide important signals that would be transported through the cytoplasm to give rise eventually to signals in the nucleus that result in the regulation of gene expression. Examination of the localization of the different ALV oncogene products shows that they provide examples of regulating growth control from at least two of these sites. The MC29 v-*myc* gene product appears to function directly within the nucleus, where it may be controlling the expression of other genes. The AEV *erb B* gene product is a membrane glycoprotein, where it may be acting as a receptor molecule which interacts with normal cellular growth factors and therefore passes signals to the cell which lead to proliferation. The E26 v-*myb* gene product gives rise to transformed cells which are still dependent on normal growth factors and may indeed interact with them. Thus, the avian ALV have identified for us at least three genes that can play central roles in growth control, and each one provides a different example of an important regulatory molecule. An understanding of how just one of these important molecules functions will be a large step forward in understanding growth control in general.

*Acknowledgments.* I would like to thank Dr. Paula Enrietto for her invaluable help in writing this review, and Mrs. Joyce Newton for her patience and typing the review.

## References

Abrams HD, Rohrschneider LR, Eisenman RN (1982) Nuclear localization of the putative transforming protein of avian myelocytomatosis virus. Cell 29:427–439

Baluda MA, Goetz IE (1961) Morphological conversion of cell cultures by avian myeloblastosis virus. Virology 15:185–199

Beemon KL (1978) Oligonucleotide fingerprinting with RNA tumor virus RNA. Current Top Microbiol Immunol 79:73–110

Beug H, v Kirchbach A, Döderlein G, Conscience J-F, Graf T (1979) Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. Cell 18:375–390

Beug H, Kitchener G, Doederlein G, Graf T, Hayman MJ (1980) Mutant of avian erythroblastosis virus defective for erythroblast transformation: Deletion in the *erb* portion of p75 AEV suggests function of the protein in leukemogenesis. Proc Natl Acad Sci USA 77:6683–6686

Beug H, Hayman MJ, Graf T (1982a) Leukaemia as a disease of differentiation: Retroviruses causing acute leukaemias in chickens. Cancer Surveys 1:205–230

Beug H, Hayman MJ, Graf T (1982b) Myeloblasts transformed by the avian acute leukemia virus E26 are hormone-dependent for growth and for the expression of a putative *gag-myb*-containing protein, p135 E26. EMBO Journal 1:1069–1073

Bishop JM (1978) Retroviruses. Annu Rev Biochem 47:35–88

Bister K, Duesberg PH (1980) Genetic structure of avian acute leukemia viruses. Cold Spring Harbor Symp Quant Biol 44:801–822

Bister K, Hayman MJ, Vogt PK (1977) Defectiveness of avian myelocytomatosis virus MC29: Isolation of long-term nonproducer cultures and analysis of virus-specific polypeptide synthesis. Virology 82:431–448

Bister K, Loliger HC, Duesberg PH (1979) Oligonucleotide map and protein of CM11: Detection of conserved and non-conserved elements in avian acute leukemia viruses. J Virol 32:208–219

Bister K, Ramsay G, Hayman MJ, Duesberg PH (1980) OK10, an avian acute leukemia virus of the MC29 subgroup with a unique genetic structure. Proc Natl Acad Sci USA 77:7142–7146

Bister K, Nunn M, Moscovici C, Perbal B, Baluda M, Duesberg P (1982a) Acute leukemia viruses E26 and avian myeloblastosis virus have related transformation-specific RNA sequences but different genetic structures, gene products and oncogenic properties. Proc Natl Acad Sci USA 79:3677–3681

Bister K, Ramsay GM, Hayman MJ (1982b) Deletions within the transformation-specific RNA sequences of acute leukemia virus MC29 give rise to partially transformation-defective mutants. J Virol 41:754–766

Chen JA (1980) Expression of endogenous avian myeloblastosis virus information in different chicken cells. J Virol 36:162–170

Chiswell DJ, Ramsay G, Hayman MJ (1981) Two virus-specific RNA species are present in cells transformed by defective leukemia virus OK10. J Virol 40:301–304

Coffin JM (1979) Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. J Gen Virol 42:1–26

Cooper GM, Neiman PE (1980) Transforming genes of neoplasms induced by avian lymphoid leukosis viruses. Nature 287:656–658

Cooper GM, Neiman PE (1981) Two distinct candidate transforming genes of lymphoid leukosis virus-induced neoplasms. Nature 292:857–859

Copeland NG, Cooper GM (1980) Transfection by DNAs of avian erythroblastosis virus and avian myelocytomatosis virus MC29. J Virol 33:1199–1202

Donner P, Greiser WI, Moelling K (1982) Nuclear localization and DNA binding of the transforming gene product of avian myelocytomatosis virus. Nature 296:262–266

Eisenman RN, Vogt VM (1978) The biosynthesis of oncovirus proteins. Biochim Biophys Acta 473:187–239

Eisenman R, Vogt VM, Diggelmann H (1974) The synthesis of avian RNA tumor virus structural proteins. Cold Spring Harbor Symp Quant Biol 39:1067–1075

Enrietto PJ, Hayman MJ (1982) Restriction enzyme analysis of partially transformation defective mutants of the acute leukemia virus MC29. J Virol 44:711–715

Enrietto PJ, Hayman MJ, Ramsay GM, Wyke JA, Payne LN (1983) Altered pathogenicity of avian myelocytomatosis (MC29) viruses with mutations in the v-*myc* gene. Virology 124: 164–173

Frykberg L, Palmieri S, Beug H, Graf T, Hayman MJ, Vennström B (1983) Transforming capacities of avian erythroblastosis virus mutants deleted in the *erb A* or *erb B* oncogenes. Cell 32:227–238

Gonda TJ, Sheiness DK, Fanshier L, Bishop JM, Moscovici C, Moscovici MG (1981) The genome and the intracellular RNAs of avian myeloblastosis virus. Cell 23:279–291

Gonda TJ, Sheiness DK, Bishop JM (1982) Transcripts from the cellular homology oncogenes: Distribution among chicken tissues. Mol Cell Biol 2:617–624

Graf T (1973) Two types of target cells for transformation with avian myelocytomatosis virus. Virology 54:398–413

Graf T (1975) In vitro transformation of chicken bone-marrow cells with avian erythroblastosis virus. Z Naturforsch C30:847

Graf T, Beug H (1978) Avian leukemia viruses. Interaction with their target cells in vivo and in vitro. Biochim Biophys Acta 516:269–299

Graf T, Stehelin D (1982) Avian leukemia viruses oncogenes and genome structure. Biochim Biophys Acta 651:245–271

Graf T, Royer-Pokora B, Schubert GE, Beug H (1976) Evidence for the multiple oncogenic potential of cloned leukemia virus: in vitro and in vivo studies with avian erythroblastosis virus. Virology 71:423–433

Graf T, Fink D, Beug H, Royer-Pokora B (1977) Oncornavirus-induced sarcoma formation observed by rapid development of lethal leukemia. Cancer Res 37:59–63

Graf T, Ade N, Beug H (1978) Temperature-sensitive mutant of avian erythroblastosis virus suggests a block of differentiation as mechanism of leukaemogenesis. Nature 257:496–501

Graf T, Oker-Blom N, Todorov TG, Beug H (1979) Transforming capacities and defectiveness of avian leukemia viruses OK10 and E26. Virology 99:431–436

Graf T, Beug H, von Kirchbach A, Hayman MJ (1980) Three new types of viral oncogenes in defective leukemia viruses II Biological, genetical and immunochemical evidence. Cold Spring Harbor Symp Quant Biol 44:1225–1234

Hanafusa H (1977) Cell transformation by RNA tumor viruses. In: Fraenkel-Conrat WRR (ed) Comprehensive, Virology. Plenum, New York, pp 401–488

Hayman MJ, Kitchener G, Graf T (1979a) Cells transformed by avian myelocytomatosis virus strain CM11 contain a 90K *gag*-related protein. Virology 98:191–199

Hayman MJ, Royer-Pokora B, Graf T (1979b) Defectiveness of avian erythroblastosis virus: Synthesis of a 75K *gag*-related protein. Virology 92:31–45

Hayman MJ, Beug H, Savin K (1982a) Changes in the expression of membrane antigens during the differentiation of chicken erythroblasts. J Cell Biochem 18:351–362

Hayman MJ, Enrietto PJ, Ramsay GM, Bister K, Graf T, Payne LN (1982b) Studies on the molecular basis of the oncogenic potential of the avian myelocytomatosis virus MC29. In: Fox CF (ed) Tumor viruses and differentiation, CETUS-UCLA symposium on cellular and molecular biology. Liss, New York

Hayman MJ, Ramsay GM, Savin K, Kitchener G, Graf T, Beug H (1983) Characterisation of the avian erythroblastosis virus *erb B* gene product as a membrane glycoprotein. Cell (in press)

Hayward WS, Neel BG, Astrin SM (1981) Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290:475–480

Hu SSF, Moscovici C, Vogt PK (1978) The defectiveness of Mill Hill 2, a carcinoma-inducing avian oncovirus. Virology 89:162–178

Hu SSF, Lai MMC, Vogt PK (1979) The genome of avian myelocytomatosis virus MC29 analyzed by heteroduplex mapping. Proc Natl Acad Sci USA 76:1265–1268

Ishizaki R, Shimizu T (1970) Heterogeneity of strain R avian erythroblastosis virus. Cancer Res 30:2827–2831

Kitchener G, Hayman MJ (1980) Comparative tryptic peptide mapping studies suggests a role in cell transformation for the *gag*-related proteins of avian erythroblastosis virus and the avian myelocytomatosis virus strains CM11 and MC29. Proc Natl Acad Sci USA 77:1637–1641

Lai MMC, Hu SSF, Vogt PK (1979) Avian erythroblastosis virus: transformation-specific sequences from a contiguous segment of 3.25 kb located in the middle of the 6-kb genome. Virology 97:366–377

Lai MMC, Neil JC, Vogt PK (1980) Cell-free translation of avian erythroblastosis virus RNA yields two specific and distinct proteins with molecular weights of 75,000 and 40,000. Virology 100: 475–483

Langlois AJ, Fritz RB, Heine U, Beard D, Bolognesi DP, Beard JW (1969) Response of bone marrow to MC29 avian leukosis virus in vitro. Cancer Res 29:2056–2074

Lautenberger JA, Schulz RA, Garon CF, Tsichlis PN, Papas TS (1981) Molecular cloning of avian myelocytomatosis virus (MC29) transforming sequences. Proc Natl Acad Sci USA 78: 1518–1522

Linemeyer DL, Menke JG, Ruscetti SK, Evans LH, Scolnick EM (1982) Envelope gene sequences which encode the gp52 protein of spleen focus-forming virus are required for the induction of erythroid cell proliferation. J Virol 43:223–233

Linial M (1982) Two retroviruses with similar transforming genes exhibit differences in transforming potential. Virology 119:382–391

Moscovici C, Moscovici MG, Zanetti M (1969) Transformation of chick fibroblast cultures with avian myeloblastosis virus. J Cell Physiol 73:105–108

Moscovici C, Gazzolo L. Moscovici MG (1975) Focus assay and defectiveness of avian myeloblastosis virus. Virology 68:173–181

Neel BG, Hayward WS, Robinson HL, Fang J, Astrin SM (1981) Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell 23:323–334

Pachl C, Linial M, Eisenman R, Biegalke B (1982) Analysis of *myc* gene products in ALV-transformed lymphoid cells and in MH-2 transformed cells. J Cell Biochem 56:p233

Palmieri S, Beug H, Graf T (1982) Isolation and characterization of four new temperature-sensitive mutants of avian erythroblastosis virus (AEV). Virology 123:296–312

Pawson T, Martin GS (1980) Cell-free translation of avian erythroblastosis virus RNA. J Virol 34:280–284

Payne GS, Courtneidge SA, Crittenden LB, Fadly AM, Bishop JM, Varmus HE (1981) Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell 23:311–322

Payne GS, Bishop JM, Varmus HE (1982) Multiple arrangements of viral DNA and an activated host oncogene (c-*myc*) in bursal lymphomas. Nature 295:209–213

Privalsky ML, Bishop JM (1982) Proteins specified by avian erythroblastosis virus: coding region localization and identification of a previously undetected *erb B* polypeptide. Proc Natl Acad Sci USA 79:3958–3962

Purchase HG, Burmester BR (1973) Leukosis sarcoma group. In: Hofstad MS, Calnek BW, Helmboldt CF, Raa WM, Yoder HW Jr (eds) Disease of poultry, 6th edn. Iowa State University Press Ames, pp 502–567

Quade K (1979) Transformation of mammalian cells by avian myelocytomatosis virus and avian erythroblastosis virus. Virology 98:451–465

Quade K, Saule S, Stehelin D, Kitchener G, Hayman MJ (1981) Revertants of rat cells transformed by avian erythroblastosis virus. Virology 115:322–333

Radke KE, Beug H, Graf T (1982) Transformation of both erythroid and myeloid cells by E26, an avian erythroleukemic virus which contains the *myb* gene. Cell 31:643–653

Ramsay G, Hayman MJ (1980) Analysis of cells transformed by defective leukemia virus OK10: Production of non-infectious particles and synthesis of $Pr76^{gag}$ and an additional 200,000-dalton protein. Virology 106:71–81

Ramsay GM, Hayman MJ (1982) Isolation and biochemical characterization of partially transformation-defective mutants of avian myelocytomatosis virus MC29. J Virol 41:745–753

Ramsay G, Graf T, Hayman MJ (1980) Mutants of avian myelocytomatosis virus MC29 with smaller *gag* gene-related proteins have an altered transforming ability. Nature 288:170–172

Ramsay GM, Enrietto PJ, Graf T, Hayman MJ (1982a) Recovery of *myc* specific sequences by a partially transformation-defective mutant of avian myelocytomatosis virus MC29, correlates with the restoration of transforming ability. Proc Natl Acad Sci USA 79:6885–6889

Ramsay GM, Hayman MJ, Bister K (1982b) Phosphorylation of specific sites in the *gag-myc* polyproteins encoded by MC29 type viruses correlates with their transforming ability. EMBO Journal 1:1111–1116

Rothe-Meyer A, Engelbreth-Holm J (1933) Experimentelle Studien über die Beziehungen zwischen Hühnerleukose und Sarkom an Hand eines Stammes von übertragbarer Leukose-Sarkom Kombination. Acta Pathol Microbiol Scand 10:380–427

Roussel M, Saule S, Lagrou C, Rommens C, Beug H, Graf T, Stehelin D (1979) Three new types of viral oncogene of cellular origin specific for hematopoietic cell transformation. Nature 281:452–455

Royer-Pokora B, Beug H, Clavier M, Winkhardt HJ, Friis RR, Graf T (1978) Transformation parameters in chicken fibroblasts transformed by AEV and MC29 avian leukemia viruses. Cell 13:751–760

Royer-Pokora B, Grieser S, Beug H, Graf T (1979) Mutant of avian erythroblastosis virus (AEV) with restricted target cell specificity. Nature 282:750–752

Rushlow KE, Lautenberger JA, Papas TS, Baluda MA, Perbal B, Chirikjian JG, Reddy EP (1982) Nucleotide sequence of the transforming gene of avian myeloblastosis virus. Science 216: 1421–1423

Saule S, Sergeant A, Torpier G, Raes MB, Pfeifer S, Stehelin D (1982) Subgenomic mRNA in OK10 defective leukemia virus-transformed cells. J Virol 42:71–82

Savin KW, Beug H (1981) Cell-surface glycoprotein synthesis during differentiation of chicken erythroblasts transformed by temperature-sensitive avian erythroblastosis virus. Cell Differ 10:163–170

Sheiness D, Vennström B, Bishop JM (1981) Virus-specific RNAs in cells infected by avian myelocytomatosis virus and avian erythroblastosis virus: modes of oncogene expression. Cell 23:291–300

Sotirov N (1981) Histone H5 in the immature blood cells of chickens with leukosis induced by avian leukosis virus strain E26. J Natl Cancer Inst 66:1143–1147

Souza LM, Komaromy MC, Baluda MA (1980) Identification of a proviral genome associated with avian myeloblastic leukemia. Proc Natl Acad Sci USA 77:3004–3008

Stehelin D, Guntaka RV, Vamus HE, Bishop JM (1976a) Purification of DNA complementary to nucleotide sequence required for neoplastic transformation of fibroblasts by avian sarcoma viruses. J Mol Biol 101:349–365

Stehelin D, Vamus HE, Bishop JM, Vogt PK (1976b) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 260:170–173

Stehelin D, Saule S, Roussel M, Lagrou C, Rommens C (1980) Three new types of transforming genes in acute defective avian leukemia viruses. Specific viral nucleotide sequences correlating with distinct phenotypes of virus-transformed cells. Cold Spring Harbor Symp Quant Biol 44:1215–1223

Varmus HE (1982) Recent evidence for oncogenesis by insertion mutagenesis and gene activation. Cancer Surveys 1:309–321

Vennström B, Fanshier L, Pessano S, Bishop JM (1980) Molecular cloning of the avian erythroblastosis virus genome and recovery of oncogenic virus by transfection of chicken cells. J Virol 36:575–585

Vennström BC, Moscovici C, Goodman HM, Bishop JM (1981) Molecular cloning of the avian myelocytomatosis virus genome and recovery of infectious virus by transfection of chicken cells. J Virol 39:625–631

Vogt PK (1977) Genetics of RNA tumor viruses. Comprehensive Virology 9:341–430

Vogt VM, Eisenman R (1973) Identification of a large polypeptide precursor of avian oncornavirus proteins. Proc Natl Acad Sci USA 70:1734–1738

Vogt VM, Eisenman R, Diggelmann H (1975) Generation of avian myeloblastosis virus structural proteins by proteolytic cleavage of a precursor polypeptide. J Mol Biol 96:471–493

Weintraub H, Beug H, Groudine M, Graf T (1982) Temperature-sensitive changes in the structure of globin chromatin in lines of red cell precursors transformed by *ts* AEV. Cell 28:931–940

Weiss R, Teich N, Varmus HE, Coffin J (eds) (1982) Molecular biology of tumor viruses. Part III, RNA tumor viruses. Cold Spring Harbor Press, New York

Witter RL (1978) In: Hofstad MS, Calneck BW, Helmbodt CF, Reid WM, Yoder Jr JW (eds) Reticuloendothelosis Virus. Diseases of poultry. Iowa State University Press, Ames, pp 480–486

Yoshida M, Toyoshima K (1980) In vitro translation of avian erythroblastosis virus RNA: identification of two major polypeptides. Virology 100:484–488

# Molecular and Cellular Biology of Abelson Virus Transformation

OWEN N. WITTE*



## 1 Overview

A tremendous increase in our knowledge of the variety of retroviral transforming genes and their complexity of structure and function has been attained within the last decade. Rapidly transforming, replication-defective retroviruses have gained a prominent position as systems to exploit in the study of the interactions of specific growth control genes and their host cells. A large number of independent isolates of such viruses from diverse vertebrate species (see *Coffin* et al. 1980) have been documented. The phenomenon of a specific host cell gene being recombined with part of the sequences of a replication-competent retrovirus and, hence, generating a new viral form and function has been demonstrated at the biological and molecular levels for a number of cases.

Three large and interrelated areas dominate much of this field. Certainly, the range and variety of transforming genes that can be documented, catalogued, and compared to the existing pool of retroviral forms is a major problem whose solution will take the concerted effort of many laboratories. Studies of the details of the mechanisms by which

* Department of Microbiology and Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

the activity of the various retroviral genes influence the host cell's transformed phenotype are still in their infancy and will be vigorously pursued. Finally, a number of this group of retroviruses, including the Abelson virus, have found special utility because of their ability to transform cell types that occur in well-studied developmental pathways. Thus, the ability of Abelson virus to transform cells of the B-lymphocyte pathway provides useful reagents (continuous transformed cell lines) for the molecular and cellular developmental immunologist. Conversely, detailed knowledge of B-lymphocyte development can be exploited to refine our attempts to study the interaction of Abelson virus with its host target cells both in vitro and in vivo. These latter points have generated much of the intensity of experimental effort in the Abelson virus system.

## 2 Isolation and Biological Studies

The isolation of Abelson murine leukemia virus (A-MuLV) by *Abelson* and *Rabstein* (1970a, b) occurred unexpectedly during studies of the leukemogenic activity of the Moloney murine leukemia virus (M-MuLV). The particular experimental protocol used and the characteristics of A-MuLV-induced disease hold several key lessons.

Moloney virus is a member of a group of replication-competent murine retroviruses that, after in vivo inoculation of susceptiable mice, give rise to thymic lymphomas with a latent period of 3–6 months. These tumors eventually spread to peripheral lymph nodes and spleen (*Moloney* 1960, 1962). Even prior to the discovery of the viral etiology of thymus-derived leukemias in certain inbred mouse strains (*Gross* 1951), it had been demonstrated that thymectomy could significantly reduce the incidence of spontaneous leukemias in such strains (*Furth* 1946; *Kaplan* 1950).

Several groups had demonstrated that viral-induced thymic leukemias were reduced following inoculation if the thymus was physically removed. In such animals tumors with distinctive pathologies occasionally appeared (*Moloney* 1962; *Gross* 1960), including myeloid and lymphosarcoma-type leukemias.

*Abelson* and *Rabstein* (1970a, b) used a similar protocol, inoculating M-MuLV into a large number of BALB/c Cr mice which had been treated with prednisolone (a potent corticosteroid) since birth. This drug treatment essentially performs a chemical thymectomy, since the thymus is particularly sensitive to the lympholytic effect of steroids (*Branceni* and *Arnason* 1966). Of the original group of over 150 mice, only ten developed tumors; of the four animals examined in any detail, the cell-free extracts prepared from only one mouse showed some unusual (non-M-MuLV) properties. This single tumor extract was the source of all preparations of Abelson murine leukemia virus. The original tumor was classified as a lymphosarcoma, and involved both cervical and inguinal lymph node groups and lymphoid tissue along the vertebral column. In addition, prominent infiltration of the marrow cavity of the calvaria with invasion into the meninges had occurred. These pathological findings were seen repeatedly 25–40 days following injection of cell-free tumor extract into susceptible mice, whether or not they had been treated with steroids. In neither the original tumor nor in the secondary recipients was there leukemic involvement of the thymus.

This lack of thymic involvement was further supported by the observation that congenitally athymic nude mice were good hosts for A-MuLV-induced lymphomas (*Raschke* et al. 1975). This distinctive pathological spectrum of nonthymic lymphoma, of short

latent period, with early involvement of both bone marrow and peripheral lymphoid tissues, provided a unique experimental system.

A close relationship between the derived A-MuLV and the inoculated M-MuLV was suggested from neutralization and dilution experiments of *Potter* et al. (1973). Antisera specific for M-MuLV neutralized A-MuLV activity, and highly diluted preparations containing A-MuLV induced classic M-MuLV-like disease. This strongly suggested that A-MuLV contained M-MuLV-like envelope antigens, and that the virus preparations used were mixtures of both A-MuLV and M-MuLV.

Although very young mice of most inbred strains were susceptible to A-MuLV-induced disease (*Abelson* and *Rabstein* 1970b), adult mice of most strains were resistant. *Risser* et al. (1978a) have identified two unlinked genes (called Av-1 and Av-2) that regulate susceptibility to A-MuLV disease. Adult BALB/c mice are the best host for A-MuLV disease; they carry the dominant sensitivity alleles at both loci. Neither allele appears to regulate the M-MuLV component.

Since the early in vivo-passaged stocks were never cloned, the interpretations of some of the original work on the pathological characteristics and species-specific reaction to A-MuLV infection should be read with caution (*Abelson* and *Rabstein* 1970b; *Rabstein* et al. 1971; *Siegler* et al. 1972).

A major improvement in the dissection of the A-MuLV/M-MuLV complex came from the work of *Scher* and *Siegler* (1975), who demonstrated that A-MuLV could directly transform established lines of mouse fibroblasts, like NIH3T3 cells. In particular, they could demonstrate that A-MuLV transformants in some cases were negative for release of virus (transformed nonproducers), but could release infectious A-MuLV after superinfection with M-MuLV or a related virus. This pattern of a replication-incompetent, rapidly transforming virus that can be rescued and packaged in *trans* is certainly not unique to A-MuLV. This represents a form of transduction of new genetic information, as will be detailed in the following sections. One of the transformed nonproducer cell lines they isolated, called ANN-1 (for Abelson NIH nonproducer number 1), became the source of clonally purified A-MuLV for a number of laboratories.

## 3 In vitro Transformation of Lymphoid Target Cells

A critical advance in our ability to examine the interactions of A-MuLV during lymphoid transformation came with the development of suitable in vitro transformation assays. Two groups showed that in vitro infection of adult hematopoietic tissues with A-MuLV, followed by short-term culture and subsequent implantation in vivo, could produce Abelson lymphomas (*Sklar* et al. 1975; *Raschke* et al. 1975). This clearly demonstrated that A-MuLV could initiate transformation in vitro; it was not a particularly useful procedure, however, because the in vivo passage was not quantitative. Some risk of virus spread to host tissues was always a relative drawback, and the complex selective environment of the whole animal made interpretation of the tumors and their phenotypes difficult.

The first in vitro system was developed by *Rosenberg* et al. (1975). They used fetal liver as a source of rapidly dividing hematopoietic cells. In a liquid culture system using fetal calf serum and a thiol reagent (2-mercaptoethanol), they were able to establish continuous cell lines which could induce tumors in syngeneic animals. A more direct and

quantitative assay, employing direct plating of A-MuLV-infected young adult bone marrow in semisolid agarose media, was reported the following year (*Rosenberg* and *Baltimore* 1976).

This agarose transformation system has proven to be of exceptional value in the comparison of different A-MuLV stocks of wild-type and mutant strains, and as a source of clonal cell lines used in developmental and transformation phenotype studies (described below). Within 10–12 days postinfection with A-MuLV/M-MuLV, macroscopic colonies (2–4 mm) can be counted and, if desired, removed and expanded into continuous cell lines in liquid culture. Generally lymphoblastoid in morphology with a rapid doubling time (12–15 hours), these cell lines are easily frozen, can be grown to high cell density ($3–5 \times 10^6$/ml), and maintain a normal diploid chromosome complement (*Rosenberg* and *Baltimore* 1976; *Klein* et al. 1980). Although many of the cell lines derived in this manner do not produce virus, all examined to date (a total of over 500 in several laboratories) contain the A-MuLV genome and express its transforming gene product (*Rosenberg* and *Baltimore* 1976; 1978; *Witte* et al. 1978; N. Rosenberg and O. Witte, unpublished). This strong correlation suggests a direct transforming effect on a susceptible target cell for A-MuLV. Those cell lines which do not produce A-MuLV occasionally show no evidence of the helper virus used in the infection. These cell lines can be easily superinfected with M-MuLV or another helper virus; they then produce A-MuLV at normal titer. A major fraction of the cell lines appear to contain a defective helper virus which is capable of releasing a low number of particles into the supernatant fluids (*Shields* et al. 1979). The precise molecular defects in the helper are not clear, but the released particles contain very small amounts of genomic RNA and surface glycoproteins. Generally, these defective particle producers can be superinfected with helper virus to rescue the A-MuLV genomic RNA to a higher titer.

A most surprising, and as yet unexplained, observation has been made about the role of the helper virus component in the Abelson lymphoid transformation system, both in vitro and in vitro (*Rosenberg* and *Baltimore* 1978; *Scher* 1978). Only those helper viruses which are themselves oncogenic (like the thymoma-inducing M-MuLV) serve as good helpers for lymphoid transformation. Nononcogenic helper viruses can serve to replicate A-MuLV and transform fibroblast lines, but they are very inefficient in the lymphoid transformation process. The molecular mechanism for this difference is not known, but it may involve a difference in the rate and duration of replication of each helper virus in lymphoid cells. Superinfection of a lymphoid A-MuLV-transformed nonproducer line with an inefficient helper leads to a transient, low-level release of A-MuLV over a few days (*Witte* et al. 1978).

## 4 Biochemical Characterization of the A-MuLV Genome and Gene Product

A close relationship between A-MuLV and M-MuLV was indicated both by the history of isolation (*Abelson* and *Rabstein* 1970a), and by the neutralization studies of *Potter* et al. (1973). The replication defect of A-MuLV, coupled with the helper activity of M-MuLV, suggested that A-MuLV genomic RNA contained sufficient sequences to be recognized, reverse transcribed, and packaged in *trans* by M-MuLV. The first demonstration of such sequences came from the study of *Parks* et al. (1976). They used an A-MuLV-trans-

formed nonproducer cell line as a source of viral RNA, and a representative small fragment cDNA probe preparation of M-MuLV sequences to demonstrate by liquid hybridization that A-MuLV retained about 25% of the sequences of M-MuLV. The precise arrangement of retained M-MuLV sequences was demonstrated by *Shields* et al. (1979). First, they were able to demonstrate by agarose-gel electrophoresis of virion RNA that A-MuLV/M-MuLV stocks contained a new RNA species of 5.6 kilobases, in addition to the 8.8-kilobase M-MuLV form. Using a full length minus strand cDNA probe from M-MuLV, they formed heteroduplexes between the A-MuLV/M-MuLV RNA and the M-MuLV DNA. They then analyzed them by electron microscopy, and with gel electrophoresis of fragments following digestion with single-strand-specific nuclease S1. This combination of techniques clearly demonstrated that A-MuLV had retained sequences closely homologous to M-MuLV at its 5′ end (about 1320 bases), and at its 3′ end (about 730 bases). A large central region of 3500 bases had no homology to M-MuLV, and was presumed to have derived from a normal mouse cellular sequence via a recombination event (Fig. 1). Many other rapidly transforming retroviruses have now been shown to have arisen by such a recombination event with distinct cellular genes (reviewed by *Bishop* 1980).

Prior to such studies demonstrating that a large region of new genetic information was contained in the A-MuLV RNA, an alternative hypothesis, in which A-MuLV represented a more subtly mutated form of M-MuLV, was entertained. Cell lines containing A-MuLV were examined for the expression of proteins serologically related to those expressed by the parental M-MuLV strain. M-MuLV, like all replication-competent retroviruses, expresses its genetic information in three major precursor polyprotein groups, *gag, gag-pol,* and *env* (*Bishop* 1978; *Shinnick* et al. 1981). The initial translation product of the M-MuLV *gag* gene, called $Pr65^{gag}$, contains sequences encoding the major virion structural proteins p15, p12, p30, and p10, in that order, from amino- to carboxyl-terminus. Using monospecific antisera for these four *gag* gene products, reverse transcriptase (the *pol* gene), and gp70 (the *env* gene), metabolically labeled cellular lysates from a large number of A-MuLV-transformed nonproducer and producer fibroblast and lymphoid cell lines were examined by direct immunoprecipitation and gel electrophoresis (*Witte* et al. 1978). A new protein of 120 000 molecular weight (called P120) was

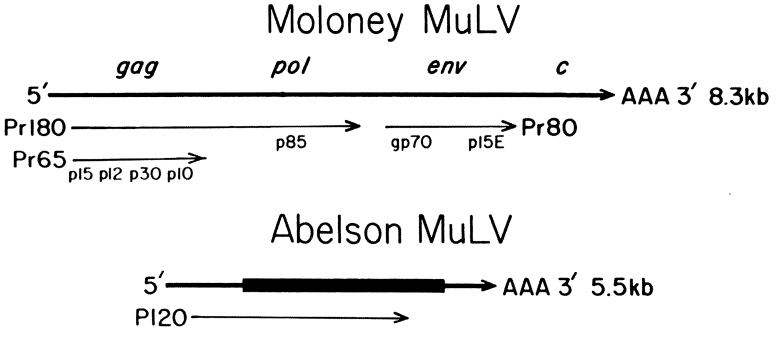



Fig. 1. RNA genome structures of Moloney and Abelson virus. The homologous 5′ and 3′ regions are shown in the *medium dark line.* The central *abl*-specific portion of A-MuLV (about 3500 bases) is the *heavy dark line.* The approximate map position of the proteins encoded from each virus is shown in *thin line*

seen in all A-MuLV-transformed cell lines examined. P120 could be precipitated by antisera specific for p15, p12, and some anti-p30 sera, but not by antipolymerase or antienvelope reagents. In addition, A-MuLV-containing virion RNA preparations could be translated in vitro in a reticulocyte lysate system to generate a P120-size molecule with identical serological reactivities (*Witte* et al. 1978).

Using a competition radioimmunoassay system and column size fractionations of cellular lysates, *Reynolds* et al. (1978a) independently observed a high-molecular-weight *gag*-related polyprotein in A-MuLV-transformed cells. They subsequently demonstrated in vitro translation of P120 by ocyte injection and immunoprecipitation (*Reynolds* et al. 1978b).

Both sets of studies, in conjunction with the RNA structural work of *Shields* et al. (1979), strongly suggested that the P120 protein was a direct product of the A-MuLV genome which was a fusion between a portion of the *gag* gene of M-MuLV containing p15, p12, and a portion of p30, and the newly acquired, centrally located sequences specific to A-MuLV (called v-*abl*).

When molecular clones of A-MuLV representing proviral circles (*Goff* et al. 1980) and cellular integrated forms (*Srinivasan* et al. 1981) became available, the relationship between M-MuLV and A-MuLV could be compared more directly. The junction between M-MuLV *gag* gene sequences and the v-*abl* sequences occurs in phase, early in the reading frame for the p30 portion of the *gag* gene when restriction map analysis and direct DNA sequence for M-MuLV (*Shinnick* et al. 1981) is compared to A-MuLV (discussed in *Goff* and *Baltimore* 1982).

The unique information expressed in the P120 protein from the v-*abl* sequences was identified serologically by preparation of mouse syngeneic tumor regressor serum that could immunoprecipitate P120 even in the presence of excess *gag*-related proteins (*Witte* et al. 1979). Almost all Abelson-induced tumors progress and kill the host in syngeneic challenges, but some tumors derived from the C57L strain would be rejected after subcutaneous inoculation. After a lengthy immunization protocol, some mice will produce antibodies of sufficient titer to use in immunoprecipitation and immunofluorescent studies (*Witte* et al. 1979; N. Rosenberg and O. Witte unpublished).

## 5 Abelson Protein has an Integral Tyrosine Kinase Activity

The key observation of *Collett* and *Erikson* (1978) that Rous sarcoma virus ($pp60^{src}$) could phosphorylate the heavy chain of immunoglobulin in an immunoprecipitate kinase reaction gave the first hint to a potential function for any retroviral transforming gene product. Although, this type of nonphysiological reaction is far removed from describing the mechanism of a transformation pathway, it has provided a major clue and established a standard against which to compare other viral transforming genes and cellular enzymes.

Immunoprecipitates of A-MuLV P120 also contain a kinase activity, but its properties are distinct from those of $pp60^{src}$ (*Witte* et al. 1979c, 1980a; *van de Ven* et al. 1980). Either anti-*gag* gene sera or anti-Abelson tumor sera was used to precipitate P120. When washed immunoprecipitates were incubated with $^{32}$P-ATP and $Mn^{++}$ (or, less efficiently, $Mg^{++}$), the P120 protein itself became phosphorylated, but the heavy chain of immunoglobulin did not. The reaction was extremely fast, and not augmented by the addition of cyclic

nucleotides. In addition, partially purified soluble preparations of P120 also apparently autophosphorylated in an in vitro kinase reaction (*Witte* et al. 1980a).

Most striking was the observation that the end product of this phosphorylation reaction was phosphotyrosine (*Witte* et al. 1980a). This phosphoamino acid had not previously been detected in any biological system. At the same time, work by *Eckhart* et al. (1979) on polyoma middle T antigen, and by *Hunter* and *Sefton* (1980) on pp60$^{src}$ also demonstrated phosphotyrosine as the end product of in vitro kinase reactions. The list of other retroviral transforming gene kinases, as well as cellular kinases that utilize tyrosine as a substrate, has grown exponentially since that time (for review see *Bishop* 1980).

For each retroviral kinase, it is necessary to show that the kinase activity is an integral part of the viral gene product, and not a contaminating cellular activity that is using the viral protein as a substrate. Also, the relationship of the kinase activity to the transforming activity needs to be established. For the A-MuLV protein a number of types of evidence can be used. First, a variety of partial and complete transformation-defective mutants of A-MuLV have been isolated from several laboratories (detailed in a following section). In each case, the relative transforming potency and in vitro kinase activity have been well correlated (*Rosenberg* et al. 1980; *Reynolds* et al. 1980; *Witte* et al. 1980b). Second, strain variants whose proteins can be electrophoretically distinguished were used to demonstrate a *trans* labeling reaction between two A-MuLV proteins in vitro (*Witte* et al. 1979c) and recently in vivo (*Ponticelli* et al. 1982).

A more direct demonstration of intrinsic kinase activity for A-MuLV proteins comes from analysis of labeling of synthetic peptide substrates that resemble the in vivo phosphotyrosine sites of some transforming proteins (*Patschinsky* et al. 1982; *Hunter* 1982) by preparations of wild-type and mutant strains (J. Konopka and O. Witte, unpublished). The extent of labeling of a synthetic peptide containing a tyrosine acceptor site by a large number of mutant and wild-type A-MuLV strains paralleled the biological potency of each strain. By comparing the reaction efficiency on this secondary substrate, one eliminates the problem of the putative enzyme and substrate being one and the same molecule. Perhaps the most direct demonstration comes from the analysis of Abelson protein expressed in *E. coli* (unpublished observations of C. Queen, J. Wang, and D. Baltimore, discussed in *Goff* and *Baltimore* 1982). Even in this circumstance, the *abl* protein can phosphorylate on tyrosine residues.

Very little is known about the precise nature of the in vivo phosphorylations carried out by P120 in fibroblasts or lymphoid cells, but *Sefton* et al. (1981) have shown a large increase in the total cellular level of phosphotyrosine following A-MuLV transformation. This increase is much larger for the fibroblast lines than the lymphoid ones examined. Some information is available from the work of *Cooper* and *Hunter* (1981) on potential cellular substrates. In fibroblasts transformed by A-MuLV and several other retroviruses, a few common phosphotyrosine-containing proteins can be detected by two-dimensional gel electrophoresis. Little is known about any of these proteins and their role in transformation. One, called 36K for its apparent size in kilodaltons, is not detected in some lymphoid transformants, and its role in this process is probably unimportant (J. Cooper, personal communication).

## 6 Viral Mutants

The relative simplicity of the A-MuLV genome and gene expression made the search for viral mutants by conventional approaches a difficult task. Since only a single protein with one defined function (transformation) was known, it was not possible to isolate mutants in transformation by selection for any linked marker using standard virological approaches. Almost all work on the molecular biology of A-MuLV had been carried out with virus stocks and cell lines derived from a transformed nonproducer fibroblast line called ANN-1 isolated by *Scher* and *Siegler* (1975). When virus and cell lines were prepared from earlier in vivo passaged stocks, a new A-MuLV strain, which encoded a protein of 160 000 daltons (P160), was found (Fig. 2). Serological analysis with anti-*gag* and anti-*abl* reagents showed that the P160 and P120 strains were closely related (*Rosenberg* and *Witte* 1980; *Rosenberg* et al. 1980). A comparison of the genome structures by Southern blot technique showed that a single major deletion of 800 base pairs in the *abl* region of P120 was the only major difference between these two strains (*Goff* et al. 1981). This has been confirmed by heteroduplex analysis between cloned DNA copies of each strain (S. Latt, S. Goff, J. Wang, and D. Baltimore, unpublished observations, discussed in *Goff* and *Baltimore* 1982).

These data would suggest that an in phase deletion must have occurred to generate P120. Both proteins would have very similar amino- and carboxyl-terminal portions. This

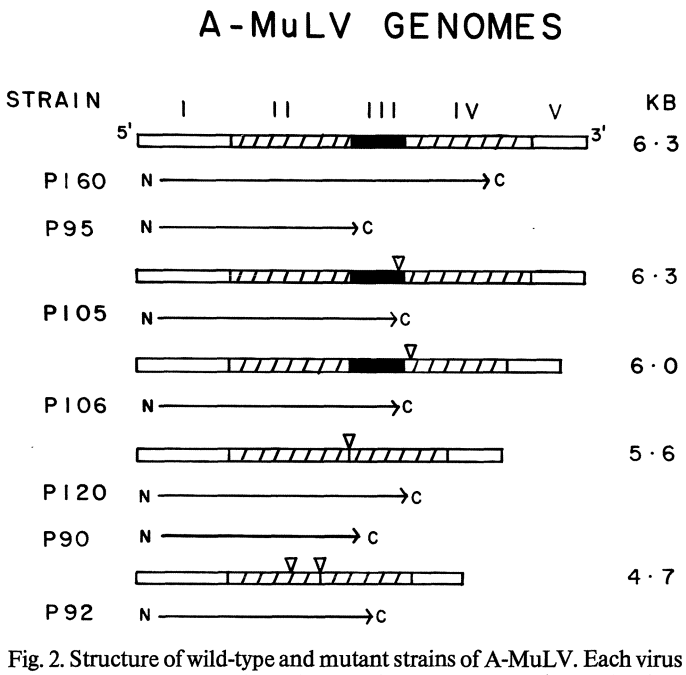



Fig. 2. Structure of wild-type and mutant strains of A-MuLV. Each virus strain is represented as the RNA form in a 5′ to 3′ orientation. Regions I and V are the retained M-MuLV sequences that are unchanged in all cases shown. Region II contains the kinase-essential sequences which are deleted in the P92*td* strain. Region III is found in the P160 strain, but deleted in P120 without markedly affecting transformation efficiency. Region IV encodes the carboxy terminal portion of the A-MuLV protein, and is required for expression of the lethal phenotype. From *Ziegler* et al. (1981)

has been confirmed by peptide mapping using partial fragmentation (*Witte* et al. 1981) and complete tryptic digestion (*Grunwald* et al. 1982; *Ponticelli* et al. 1982). Although some other minor differences (probably point mutations or very small deletion substitutions) may be present, the P160 and P120 strains are closely related.

Since both P120 and P160 strains are wild-type for kinase activity, and transformation of fibroblasts and lymphoid targets in vitro and in vivo, one must assume that the deleted sequences lost in P120 are not essential for transformation, as measured by these assays (*Rosenberg* and *Witte* 1980; *Rosenberg* et al. 1980). A number of different mutants isolated from both P160 and P120 strains have further confirmed this concept of nonessential sequences in the *abl* region. Spontaneous mutants isolated by nonselective cloning from P120 stocks have P120-sized genomes but carboxyl-terminal-truncated proteins that are still able to efficiently transform fibroblasts in vitro. These mutants, called P90 and P100, which presumably have premature stop codons secondary to point mutations or small deletions, show a reduced kinase activity and proportional decrease in lymphoid transformation efficiency both in vitro and in vivo (*Rosenberg* and *Witte* 1980; *Rosenberg* et al. 1980; *Goff* et al. 1981). These partially transformation-defective mutants have proved useful in establishing a correlation between the kinase activity and the transformation behavior. A number of other carboxyl-terminal-truncation mutants have been prepared by site-directed mutagenesis of cloned P160 and P120 genomes (*Watanabe* and *Witte* 1982; R. Prywes and D. Baltimore, unpublished observations, personal communication) with very similar results to the spontaneous mutants like P90. One mutant isolated from P160, called P87, has lost expression of over half of the *abl* sequences, yet still transforms fibroblasts (*Watanabe* and *Witte* 1982).

The correlation between amount of kinase activity and biological activity was further substantiated by the isolation of transformation-defective mutants which express serologically cross-reactive, but functionless, A-MuLV proteins (*Witte* et al. 1980b; *Reynolds* et al. 1980; *Ponticelli* et al. 1982). The molecular defect of one mutant, called P92*td*, has been determined by Southern blot analysis and peptide mapping (*Goff* et al. 1981; *Ponticelli* et al. 1982). This mutant was derived from P120 by an in phase deletion of 700 base pairs from the 5′ region of the *abl* sequences. This locates essential sequences for kinase and transformation in this portion of the *abl* segment.

An alternative mapping of sequences required for transformation comes from the work of *Srinivasan* et al. (1981), who used restriction enzyme inactivation patterns of direct DNA-mediated fibroblast transformation to define essential *abl* sequences. This analysis is limited by the available restriction sites, but the results are in good agreement with mutants recovered as virus and mapped by other means.

Most mutants described above were isolated by mass screenings of clonal lines by immunological techniques. In order to select mutant strains of A-MuLV, it was necessary to define a secondary function, other than transformation, carried by the A-MuLV protein. An unexpected variation in the stability of A-MuLV infection on different mouse fibroblast lines has provided such a secondary function (*Ziegler* et al. 1981). We observed that when some sublines of BALB3T3 were infected with wild-type strains of A-MuLV and cultured on tissue-culture-grade dishes, the cells would transform but over a period of a few weeks, the nontransformed population would overgrow. In fact, a direct lethal effect of the A-MuLV gene product could be demonstrated. Cells which lost expression of wild-type A-MuLV had a selective advantage. Transformation-defective mutants could stably infect such cell lines.

Wild-type A-MuLV stocks were used to infect such sensitive indicator cell lines, and a continued selection pressure for transformed cells (growth on bacterial-grade petri dishes which do not allow substratum adherance) was applied. This selected cell population contained a very high frequency of A-MuLV mutants. This class of mutants, called lethal minus, had lost expression of most carboxyl-terminal *abl* sequences, leading to truncated proteins which were still kinase- and transformation-positive at near wild-type levels (*Ziegler* et al. 1981).

Using site-directed mutagenesis on cloned DNA copies of the P160 A-MuLV, further separation of the sequences essential for kinase activity and transformation, and those additionally required for the lethal effect on fibroblasts, have been mapped (*Watanabe* and *Witte* 1982). Although the function required for the lethal phenotype is not understood, the genetic data from mutants and coinfections imply that it must work synergistically in *cis* with the kinase function to exert its effect.

An independent demonstration of the lethal capability of A-MuLV is seen in direct transfection of fibroblasts with cloned viral DNA. The efficiency of transfection of a selectable metabolic marker could be reduced by cotransformation with A-MuLV DNA. Presumably, many cells that took up multiple copies of each DNA succumbed to the lethal effect, and were never recovered (*Goff* et al. 1982). It is interesting that many of the recovered transformants from wild-type A-MuLV DNA transfections express truncated proteins that resemble the lethal-minus variants described above (*Watanabe* and *Witte* 1982; R. Prywes and D. Baltimore, personal communication).

Since the A-MuLV protein is a fusion between a portion of the *gag* gene sequences of M-MuLV and the cellular *abl* gene, it is possible that the transforming properties of P120 are partly due to the amino-terminal *gag* gene portion. One could imagine a number of conformational or cell localization properties regulated in some manner by the retained M-MuLV piece. To test this directly, R. Prywes and D. Baltimore (personal communication) have prepared site-directed A-MuLV mutants that lack almost all of the *gag* gene sequences. Such mutants are clearly able to transform fibroblasts. Their lymphoid-transforming capability in vivo and in vitro is not yet clear. Interestingly, these mutants lack detectable autokinase activity as measured in the immunoprecipitate assay, but they are able to label synthetic peptide substrates in vitro (J. Konopka and R. Prywes, unpublished observations).

An additional complication regarding the nature of sequences regulating transformation efficiency comes from the work of N. Rosenberg (personal communication) on the in vivo selection of mutants of higher oncogenic potential. When highly susceptible mice are inoculated with A-MuLV P90 strain virus with M-MuLV helper, the incidence and rate of tumor appearance is very low compared to wild-type (*Rosenberg* et al. 1980). In those animals who get tumors, a very large proportion contain a new mutant A-MuLV virus which makes an 85 000-dalton protein, P85. This appears to be a further carboxyl-terminal truncation from wild-type. Although P85 stocks are as inefficient as P90 in lymphoid transformation in vitro, P85 stocks are highly efficient on in vivo inoculation. The molecular mechanism of this difference is unclear, but the role of host environment and defenses will need to be sorted out.

A final note is the absence of useful temperature-sensitive transformation mutants in the A-MuLV system. A number of laboratories have been unsuccessful using standard biological approaches and have now turned to in vitro mutagenesis on cloned DNA.

# 7 Posttranslational Modification of the A-MuLV Protein

## 7.1 Phosphorylation

Because of the in vitro and probable in vivo kinase activity of the A-MuLV protein, the details of phosphorylation have been examined. When P120 labeled by the in vitro autokinase reaction is examined by partial proteolysis- and large-fragment mapping, and complete tryptic digestion and two-dimensional peptide mapping techniques (*Witte* et al. 1981), a number of tyrosine phosphorylation acceptor sites are seen. All of these sites cluster in the amino-terminal 40 000-dalton region, which includes the *gag* gene sequences and some *abl* sequences. The sites utilized in the in vivo phosphorylations are different, and are mainly found in the *abl* region (*Witte* et al. 1981). *Sefton* et al. (1981) have analyzed each site of phosphorylation for specific phosphorylated amino acids, and found two unique tyrosine acceptor sites and a number of serine and threonine sites. A series of P120 and P160 derived deletion/truncation mutant strains were examined by the same two-dimensional mapping technique (*Ponticelli* et al. 1982), and all transformation-competent strains showed the same tyrosine acceptor sites, but could differ slightly in their serine/threonine sites. From the known structure of the apoprotein truncation of several mutants, these tyrosine acceptor sites fall near the sequences encoded by the kinase-essential region. Two transformation defective (*td*) mutant strains were examined and found to lack both phosphotyrosine peptides. However, when an electrophoretically distinguishable transformation-competent strain was coinfected with the mutant, the *td*-encoded protein did contain the same two phosphotyrosine-containing peptides as all competent strains. This strongly suggested that one A-MuLV protein might function in vivo to *trans*-kinase another. Coinfection with several other transforming retroviruses does not show this *trans* labeling (*Ponticelli* et al. 1982; E. Scolnick, F. Ponticelli and O. Witte, unpublished). This in vivo *trans*-kinase is similar to the in vitro *trans*-kinase described previously by *Witte* et al. (1979c) in the immunoprecipitate system. Both systems suggest a dimer or oligomer association between A-MuLV protein molecules may be occurring in vivo as well as in detergent extracts. We suspect that the *gag* gene sequences that normally mediate core structure formation during retrovirus budding play a role in this association.

## 7.2 Glycosylation

Carbohydrate modification was not found using several criteria when P120 or mutant proteins derived from this strain were examined. P120 was insensitive to endoglycosidase H or tunicamycin blockade, and no incorporation of radioactive sugar precursors into this protein could be detected (*Witte* et al. 1979a).

The 5′ region of the A-MuLV genome, including a part of the *gag* gene coding sequences, is very closely related to its M-MuLV parent (*Shields* et al. 1979; *Goff* et al. 1980; *Srinivasan* et al. 1981). The M-MuLV *gag* gene is primarily expressed as a nonglycosylated form ($Pr65^{gag}$) which initiates with the coding sequence of core structural protein, p15 (*Shinnick* et al. 1981). A second *gag*-containing translation product initiates at an upstream site which has not been precisely determined despite knowledge of the complete nucleotide sequence. This upstream translation start site generates an apoprotein

form with a "leader" piece amino-terminal, and in phase with the normal $Pr65^{gag}$ reading frame results in a protein of about 70 000–75 000 daltons. This apoprotein form is modified by carbohydrate additions of the asparagine-linked high-mannose-type which finally leads to complex sugar processing, cell surface deposition, and some proteolytic variants (see *Edwards* and *Fan* 1981; *Schultz* et al. 1981 for general discussion and references).

Although P120 did not express this carbohydrate-modified form utilizing an alternative start site, the P160 strain and mutants derived from it do express both protein forms (*Ziegler* et al. 1981). P160 and mutants derived from it which have lost expression of carboxyl-terminal *abl* sequences all express two protein forms separable on low-percentage acrylamide gels. The lower molecular weight component of each doublet appears to utilize the P15 start site analogous to $Pr65^{gag}$. The higher molecular weight form is carbohydrate modified, shifts after endoglycosidase H digestion or tunicamycin blockade, and contains extra protein sequences at its amino-terminus (F. Ponticelli, C. Whitlock, S. Ziegler, and O. Witte, unpublished observations). It appears analogous to the form of the M-MuLV *gag* gene which becomes glycosylated. Since both P120 and P160 are wild-type for transformation, this carbohydrate-modified form probably has no major role in oncogenicity but does provide a useful marker to distinguish these strains. A possible explanation for this difference between P160 and P120 was found in comparing the DNA sequence upstream from the P15 ATG start site in both strains. The P160 sequence corresponds to the open reading frame of M-MuLV, while P120 has a single base mutation which would terminate any proteins started upstream prematurely in the "leader" piece. The P15 ATG start would not be affected (D. Robertson and O. Witte, unpublished observations).

## 7.3 Subcellular Localization

Any evaluation of subcellular localization should recognize that the hybrid nature of the A-MuLV protein containing both *gag* and *abl* sequences, potential alternative start sites, phosphorylation, carbohydrate modifications, and major differences in *abl* sequence between P160 and P120 and their mutants will complicate the picture.

Using simple fractionation schemes of differential centrifugation following homogenization of A-MuLV-transformed fibroblasts or lymphoid cells, P120 was largely recovered in the particulate-membrane fraction (*Witte* et al. 1979a; *Rotter* et al. 1981). A large proportion of this membrane-associated form can be removed by high salt wash, indicating a more peripheral attachment (F. Ponticelli and O. Witte, unpublished). Some of this pool may be related to P120 which remains bound to cytoskeletal structures in detergent permeabalized cells (*Boss* et al. 1981).

A portion of the *abl* region of P120 and P160 appears to cross the cell membrane and become exposed on the cell membrane for serological detection by immunofluorescence using anti-Abelson tumor sera (*Witte* et al. 1979a). In our original study, none of the anti-*gag* sera we tried were able to the stain surface of A-MuLV transformed cells strongly. Recently L. Schiff-Maker and N. Rosenberg (personal communication) have prepared anti-*gag* protein-specific hybridoma antibodies which bind to the surface of P120 and P160 A-MuLV-transformed nonproducer cell lines. Determining the precise orientation and extent of sequences exposed outside the membrane has been hindered

by lack of labeling of A-MuLV proteins with standard surface iodination procedures (*Witte* et al. 1979a).

A very recent development that may be related to this issue is the demonstration that the A-MuLV protein can occur as a lipoprotein-modified form by covalent attachment of fatty acid derivatives (B. Sefton, personal communication). Whether this alteration or other modifications regulates the final subcellular localization remains to be tested.

## 8 The Cellular *abl* Gene and Protein

The earlier studies of RNA (*Shields* et al. 1979) and protein structure for A-MuLV (*Witte* et al. 1978; *Reynolds* et al. 1978a) suggested that new sequences from the mouse genome were recombined with M-MuLV. The formal demonstration of this required molecular clones of A-MuLV and preparations of subclones to use as viral-Abelson-specific (v-*abl*) probes (*Goff* et al. 1980). Using these v-*abl* probes, it was demonstrated that mice contained a unique copy gene (called c-*abl* for cellular form) that cross-hybridized. These sequences are highly conserved throughout evolution, being detected at high stringency in all vertebrates and at lower stringency even in Drosophila (*Goff* et al. 1980; *Dale* and *Ozanne* 1981; *Shilo* and *Weinberg* 1981).

A detailed analysis of these c-*abl* sequences isolated in cloned DNA form from the mouse has been carried out by S. Goff, J. Wang, and D. Baltimore (unpublished observations discussed in *Goff* and *Baltimore* 1982). These c-*abl* sequences are spread out over a 25–30-kilobase region. Their studies have utilized a number of v-*abl* probes in Southern blot analysis, heteroduplex mapping, and some DNA sequencing. The overall picture that has emerged suggests that the sequences of v-*abl* represent a portion of the exon sequences of the c-*abl* gene.

These v-*abl* and c-*abl* probes have been used to detect two large RNA species (5.3 and 6.5 kilobases) of low copy number in lymphoid and fibroblastic tissues and cell lines that presumably represent the transcription products of the c-*abl* gene (*J. Wang,* unpublished observations). The large size of either of these mRNA species is consistent with the single identified c-*abl* gene product, called NCP150 (for normal cell protein 150 000 molecular weight), identified by *Witte* et al. (1979b) in lymphoid tissues by direct immunoprecipitation using anti-Abelson-tumor serum. Very small amounts of NCP150 can be detected in fibroblast cell lines if $^{32}$P-orthophosphate-labeling is utilized (*F. Ponticelli* and *O. Witte,* unpublished observations).

More direct evidence that NCP150 is a product of c-*abl* sequences has come from comparison of two-dimensional tryptic peptide maps of the cellular and viral phosphoproteins *(Ponticelli* et al. 1982). Although a number of homologous peptides for the c-*abl* and v-*abl* proteins were seen, the c-*abl* gene product did not contain detectable phosphotyrosine. This suggests that a qualitative difference between these two proteins in functional tyrosine kinase activity may be present. To date, NCP150 has failed to show any kinase activity (O. Witte, unpublished observations); however, the assays utilized are so artificial that negative results are difficult to interpret.

Recently, a number of laboratories have utilized specific v-*abl* probes to examine the expression of c-*abl* mRNA in tumors and normal tissues. No specific correlation of *abl* expression and tumor phenotype was apparent, although many tumors do contain *abl* RNA (*Westin* et al. 1982). When fetal mouse tissues at different stages of gestation are

examined, an increase in *abl* expression is seen around day 10 (*Muller* et al. 1982). Since a variety of cultured cell types and tissues including adult testis express c-*abl*, it seems likely that its expression may be related more to active cell growth than to any specific cell lineage.

## 9 Variation in Hematolymphoid Transformation by A-MuLV

A large amount of data has been accumulated to define the most common cell phenotype of A-MuLV transformants as pre-B lymphoid cells (*Pratt* et al. 1977; *Siden* et al. 1979; *Boss* et al. 1979). Using in vitro transformation of adult bone marrow or fetal liver to generate a large number of continuous clonal cell lines, *Alt* et al. (1981) showed evidence of rearrangements in the mu heavy chain gene, but infrequent rearrangements at either light chain locus. A similar differentiation phenotype is recovered from in vivo transformants in most cases. A small number of A-MuLV transformants produce both mu heavy chain and kappa light chain. Many of these cell lines express the early lymphoid marker terminal deoxynucleotidyl transferase (*Silverstone* et al. 1980, Fig. 3).

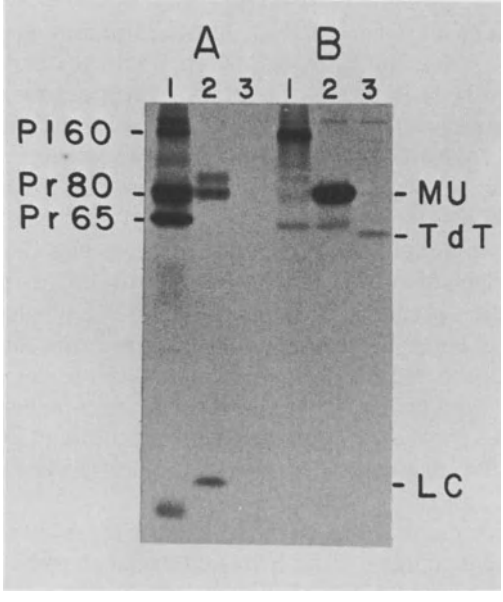



Fig. 3. Lymphoid differentiation and viral gene markers expressed in A-MuLV transformants. Two A-MuLV transformed BALB/c clones lymphoid transformants were labeled with $^{3}$H-leucine and extracted for immunoprecipitation-gel electrophoresis analysis with a polyvalent goat anti-Moloney virion serum (*lane 1*), a rabbit anti-mouse immunoglobulin serum (*lane 2*), or a rabbit anti-terminal-deoxynucleotidyl transferase (TdT) serum (*lane 3*). One cell line (*Panel A*) exhibits a B-cell phenotype with expression of both membrane and secreted mu immunoglobulin heavy chains and kappa light chain, and no detectable expression of TdT. Both A-MuLV P160 and helper virus proteins Pr80 and Pr65 are seen. Another clone (*Panel B*) expresses the more common pre-B cell phenotype with intracellular mu chains, no light chain, and positive TdT reactivity. In this line very low helper virus protein levels are seen with good expression of the A-MuLV P160 gene product. Details of the experimental techniques are found in *Whitlock* and *Witte* (1981) and prior references

How the A-MuLV/M-MuLV complex determines this preference is not clear. Since the cell surface receptor specificity is determined by the M-MuLV *env* gene components which can mediate infection of a wide variety of cell types, this seems unlikely to determine the tumor phenotype. Alternative explanations would center on specific interactions of the A-MuLV gene product with different cell types. Since NCP150 appears in a variety of cell types, unless P120 has a unique function, it is hard to imagine a unique specificity of the tyrosine kinase activity determining selection of target cells. Although difficult to quantitate, the growth rate, passage efficiency, and available numbers of different target cell types could mediate the efficiency of recovery of different phenotypes.

A number of unusual phenotypes for A-MuLV transformants have been reported when the transformation is carried out under special conditions. By explanting in vivo derived tumors and cloning, *Raschke* et al. (1978) were able to derive cell lines with macrophage characteristics. The relationship of this phenotype to those in the B cell lineage is not clear. Recently, *Boyd* and *Schrader* (1982) have reported that one A-MuLV-transformed pre-B type cell line could be induced to acquire macrophagelike characteristics.

If the initial conditions of A-MuLV infection are varied it is possible to derive unique phenotypes. Using a pristane priming system followed by A-MuLV in vivo infection, *Potter* et al. (1973) were able to isolate cell lines with plasmacytoma characteristics. Most striking is the recent report of *Cook* (1982) that direct intrathymic injection of A-MuLV/M-MuLV could induce thymic tumors of T cell origin. In earlier work thymocytes were not able to be transformed in vivo or in vitro (*Abelson* and *Rabstein* 1970b; *Rosenberg* and *Baltimore* 1976). An intriguing further expansion of the range of cell types that A-MuLV can stimulate to grow was observed by *Waneck* and *Rosenberg* (1981). When very early fetal liver cells were infected with A-MuLV and plated in agarose media, a portion of the colonies obtained were red in color when observed macroscopically. Further analysis clearly demonstrated that these colonies contained cells at various stages of erthyropoiesis, and that A-MuLV was required for the effect.

Thus, the growth of cells of the T, B, macrophage, and erythroid lineages, as well as cultured fibroblasts, is altered by A-MuLV. It is not possible from such tumor phenotyping studies to firmly establish the nature of the target cell in the fetal liver or bone marrow that was initially transformed. It is possible that a target cell at a much earlier stage of development is able to undergo further limited differentiation before fixing at a particular stage in one cell lineage.

One indirect approach has been to compare the target cell frequencies in normal and mutant strains of mice carrying defects in T cell, late B cell, or hematopoietic microenvironment. In all cases examined, the number of A-MuLV targets was unchanged (N. Rosenberg and V. Sato, unpublished, discussed in *Baltimore* et al. 1979).

An alternative approach has been to search for antisera or monoclonal antibodies with specificity for A-MuLV target cells that could be used to enrich such populations. *Shinefeld* et al. (1980) have prepared monoclonal antibodies in rats immunized with mouse brain tissue. The antibodies have the ability to recognize surface antigens on A-MuLV target cells in bone marrow. The antibodies clearly recognize other cell types as well, but are a promising start to this type of purification.

Any physical purification of a specific cell type from bone marrow is a difficult job, since the complexity of the mixture is so great. We have recently begun a long-range

project to biologically purify nontransformed cell types in the B cell lineage by selective culture techniques. These highly enriched or pure clonal lines could then be used as target populations for A-MuLV transformation. Using a modification of the long-term marrow culture technique developed by *Dexter* et al. (1977), we have established long-term cultures of cells and their precursors from murine bone marrow (*Whitlock* and *Witte* 1982a). These cultures continuously produce pre-B and B cells with random rearrangements of their heavy and light chain genes over many months. These cultured cells serve as excellent targets for A-MuLV transformation.

Recently, we have been able to derive clonal lines of normal pre-B and B cells from such cultures. Both cell types serve as targets for A-MuLV as evidenced by agar growth assays (C. Whitlock, S. Ziegler, J. Stafford, and O. Witte, unpublished observations).

## 10 Cellular Changes During A-MuLV Transformation

Following A-MuLV transformation of fibroblastic cells, there is a loss of functional receptors for the peptide hormone epidermal growth factor (EGF, *Blomberg* et al. 1980). Recently *Twardzik* et al. (1982) have reported that A-MuLV-transformed fibroblasts release a low-molecular-weight peptide called TGF, for tumor growth factor. TGF competes with EGF for binding. The specific biological relevance for transformation of lymphoid cells remains to be determined.

Another cell change induced by A-MuLV is the appearance of new cell surface antigens. In two different systems, animals with regressing syngeneic A-MuLV tumors produce antibodies reactive with nonviral cell surface antigens. In one system (*Rotter* et al. 1980, 1981), a protein of 50 000 daltons is expressed at increased level and antigenic accessibility compared to normal cells. This protein, or a close analogue, is found on the surface of tumors induced by a variety of biological or physical agents. A second system has identified an antigen present on some normal hematopoietic cells and in high amounts on A-MuLV-transformed lymphoid tumors (*Risser* et al. 1978b). Although, initially thought to be a viral gene product, this antigen is clearly encoded by cellular information (*Grunwald* et al. 1982). No functional significance for either cell surface antigen is known.

A complication to consider when trying to define such cellular changes is that A-MuLV lymphoid transformation does not appear to be a synchronous, all-or-none effect. It is possible to show that most A-MuLV-infected lymphoid target cells are poorly oncogenic early after infection. They require weeks to months to progress to a highly oncogenic state, as monitored by in vivo and in vitro growth assays (*Whitlock* and *Witte* 1981) This progression pattern is seen even when clonal cell lines are examined (*Whitlock* et al. 1983). The nature of the secondary cellular changes that presumably occur as a part of the transformation process is not known. In an extreme case, one reported tumor has lost all expression of A-MuLV after long-term passage in vivo (*Grunwald* et al. 1982).

## 11 Future Studies

The utilization of molecular genetic techniques should rapidly expand our knowledge of certain aspects of A-MuLV biology. Details of protein structure, genetic regions, and relationships to other viruses should be forthcoming from direct sequence and other

structural studies now in progress. A number of expression vector systems should provide an ample supply of both v-*abl* and c-*abl* gene products to analyze and use in deduction of intracellular function.

The expanded host range of cell types able to be transformed by A-MuLV will be useful for studying developmental aspects of hematopoiesis and perhaps other organ systems. The perplexing problems of viral transformation specificity and efficiency will require some new approaches that can quantitate the effects of A-MuLV at early and late times after infection on defined populations of target cells.

Finally, although the role of the phosphotyrosine kinase function in fibroblast and hematopoietic transformation is well established, there is no direct clue as to its mode of action. The demonstration of the cellular target molecule(s) for the kinase would be a first step in seeing the path of its action.

## References

Abelson HT, Rabstein LS (1970a) Influence of prednisolone on Moloney leukemogenic virus in BALB/c mice. Cancer Res 30:2208–2212

Abelson HT, Rabstein LS (1970b) Lymphosarcoma: virus-induced thymic-independent disease in mice. Cancer Res 30:2213–2222

Alt F, Rosenberg N, Lewis S, Thomas E, Baltimore D (1981) Organization and reorganization of immunoglobulin genes in A-MuLV-transformed cells: rearrangement of heavy but not light chain genes. Cell 27:381–390

Baltimore D, Rosenberg N, Witte ON (1979) Transformation of immature lymphoid cells by Abelson murine leukemia virus. Immunol Rev 48:3–22

Bishop JM (1978) Retroviruses. Annu Rev Biochem 47:35–88

Bishop JM (1980) Enemies within: the genesis of retrovirus oncogenes. Cell 23:5–6

Blomberg J, Reynolds FH Jr, van de Ven WJM, Stephenson JR (1980) Abelson murine leukemia virus transformation involves loss of epidermal growth factor binding sites. Nature 286: 504–507

Boss M, Greaves M, Teich N (1979) Abelson virus-transformed haematopoietic cell lines with pre-B cell characteristics. Nature 278:551–553

Boss MA, Dreyfuss G, Baltimore D (1981) Localization of the Abelson murine leukemia virus protein in a detergent insoluble subcellular matrix: architecture of the protein. J Virol 40: 472–481

Boyd AW, Schrader JW (1982) Derivation of macrophage-like lines from the pre-B lymphoma ABLS 8.1 using 5-azacytidine. Nature 297:691–693

Branceni D, Arnason BG (1966) Thymic involution and recovery of immune responsiveness and immunoglobulins after neonatal prednisolone in rats. Immunology 10:35–44

Coffin JM, Varmus HE, Bishop JM et al. (1981) Proposal for naming host cell-derived inserts in retrovirus genomes. J Virol 40:953–957

Collett MS, Erikson RL (1978) Protein kinase activity associated with the avian sarcoma virus *src* gene product. Proc Natl Acad Sci USA 75:2021–2024

Cook WD (1982) Rapid thymomas induced by Abelson murine leukemia virus. Proc Natl Acad Sci USA 79:2917–2921

Cooper JA, Hunter T (1981) Four different classes of retroviruses induce phosphorylation of tyrosines present in similar cellular proteins. Mol Cell Biol 1:394–407

Dale B, Ozanne B (1981) Characterization of mouse cellular DNA homologous to Abelson murine leukemia virus-specific sequences. Mol Cell Biol 1:731–742

Dexter TM, Allen TD, Lajtha LG (1977) Conditions controlling the in vitro proliferation of hemopoietic stem cells in vitro. J Cell Physiol 91:335–340

Eckhart W, Hutchinson MA, Hunter T (1979) An activity phosphorylating tyrosine in Polyoma T antigen immunoprecipitates. Cell 18:925–933

Edwards SA, Fan H (1981) Immunoselection and characterization of Moloney-murine-leukemia-virus-infected cells lines deficient in surface *gag* antigen expression. Virology 113:95–108
Furth J (1946) Prolongation of life with prevention of leukemia by thymectomy in mice. J Gerontol 1:46–54
Goff SP, Baltimore D (1982) The cellular oncogene of the Abelson murine leukemia virus genome. In: Klein G (ed) Viral oncology. Raven, New York
Goff SP, Gilboa E, Witte ON, Baltimore D (1980) Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: Studies with cloned viral DNA. Cell 22:777–786
Goff SP, Witte ON, Gilboa E, Rosenberg N, Baltimore D (1981) Genome structure of Abelson murine leukemia virus variants: proviruses in fibroblasts and lymphoid cells. J Virol 38: 460–468
Goff SP, Tabin CJ, Wang J, Weinberg R, Baltimore D (1982) Transfection of fibroblasts by cloned Abelson murine leukemia virus DNA and recovery of transmissible virus by recombination with helper virus. J Virol 41:271–285
Gross L (1951) Pathologic properties and vertical transmission of the mouse leukemia agent. Proc Soc Exp Biol Med 78:342–349
Gross L (1960) Development of myeloid (chloro-) leukemia in thymectomized C3H mice following inoculation of lymphatic leukemia virus. Proc Soc Exp Biol Med 103:509–514
Grunwald DJ, Dale B, Dudley J, Lamph W, Sugden B, Ozanne B, Risser R (1982) Loss of viral gene expression and retention of tumorigenicity by Abelson lymphoma cells. J Virol 43:92–103
Hunter T (1982) Synthetic peptide substrates for a tyrosine protein kinase. J Biol Chem 257: 4843–4848
Hunter T, Sefton B (1980) Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc Natl Acad Sci USA 77:1311–1315
Kaplan HS (1950) Influence of thymectomy, splenectomy and gonadectomy on incidence of radiation-induced lymphoid tumors in strain C57 black mice. J Natl Cancer Inst 11:83–90
Klein G, Ohno S, Rosenberg N, Weiner F, Spira J, Baltimore D (1980) Cytogenetic studies on Abelson-virus-induced mouse leukemias. Int J Cancer 25:805–811
Moloney JB (1960) Biological studies on a lymphoid leukemia virus extracted from sarcoma S37. I. Origin and introductory investigations. J Natl Cancer Inst 24:933–995
Moloney JB (1962) The murine leukemias. Fed Proc 21:19–31
Muller R, Slamon DJ, Tremblay JM, Cline MJ, Verma IM (1982) Differential expression of cellular oncogenes during pre- and postnatal development of the mouse. Nature
Patschinsky T, Hunter T, Esch FS, Cooper JA, Sefton BM (1982) Analysis of the sequence of amino acids surrounding sites of tyrosine phosphorylation. Proc Natl Acad Sci USA 79: 973–977
Parks WP, Hawk RS, Anisowicz A, Scolnick EM (1976) Deletion mapping of Moloney type C virus: polypeptide and nucleic acid expression in different transforming virus isolates. J Virol 18:491–503
Ponticelli AS, Whitlock CA, Rosenberg N, Witte ON (1982) In vivo tyrosine phosphorylations of the Abelson virus transforming protein are absent in its normal cellular homolog. Cell
Potter M, Sklar MD, Rowe WP (1973) Rapid viral induction of plasmacytomas in pristane-primed BALB/c mice. Science 182:592–594
Pratt DM, Strominger J, Parkman R, Kaplan D, Schwaber J, Rosenberg N, Scher CD (1977) Abelson-virus-transformed lymphocytes: null cells that modulate H-2. Cell 12:683–690
Rabstein LS, Gazdar AF, Chopra HC, Abelson HT (1971) Early morphological changes associated with infection by a murine non-thymic lymphatic tumor virus. J Natl Cancer Inst 46:481–491
Raschke WC, Ralph P, Watson J, Sklar M, Coon H (1975) Oncogenic transformation of murine lymphoid cells by in vitro infection with Abelson leukemia virus. J Natl Cancer Inst 54: 1249–1253
Raschke WC, Baird S, Ralph P, Nakoinz I (1978) Functional macrophage cell lines transformed by Abelson leukemia virus. Cell 15:261–267
Reynolds FH Jr, Sacks TL, Deobaghar DN, Stephanson JR (1978a) Cells non-productively transformed by Abelson murine leukemia virus express a high-molecular-weight polyprotein containing structural and non-structural components. Proc Natl Acad Sci USA 75:3974–3978

Reynolds RK, van de Ven WJM, Stephenson JR (1978b) Translation of type C viral RNAs in *Xenopus laevis* oocytes: evidence that the 120 000-molecular-weight polyprotein expressed in Abelson-leukemia-virus-transformed cells is virus coded. J Virol 28:665–670

Reynolds FH, van de Ven WJM, Stephenson JR (1980) Abelson murine leukemia virus transformation-defective mutants with impaired P120-associated protein kinase activity. J Virol 36:374–386

Risser R, Potter M, Rowe WP (1978a) Abelson-virus-induced lymphomagenesis in mice. J Exp Med 148:714–726

Risser R, Stockert E, Old LJ (1978b) Abelson antigen: a viral tumor antigen that is also a differentiation antigen of BALB/c mice. Proc Natl Acad Sci USA 75:3918–3922

Rosenberg N, Baltimore D (1976) A quantitative assay for transformation of bone marrow cells by Abelson murine leukemia virus. J Exp Med 143:1453–1463

Rosenberg N, Baltimore D (1978) The effect of helper virus on Abelson-virus-induced transformation of lymphoid cells. J Exp Med 147:1126–1141

Rosenberg N, Baltimore D (1980) Abelson virus. In: Klein G (ed) Viral oncology. Raven, New York, pp 187–203

Rosenberg N, Witte ON (1980) Abelson murine leukemia virus mutants with alterations in the virus-specific P120 molecule. J Virol 33:340–348

Rosenberg N, Baltimore D, Scher CD (1975) In vitro transformation of lymphoid cells by Abelson murine leukemia virus. Proc Natl Acad Sci USA 72:1932–1936

Rosenberg N, Clark DR, Witte ON (1980) Abelson murine leukemia virus mutants deficient in kinase activity and lymphoid cell transformation. J Virol 36:766–774

Rotter V, Witte ON, Coffman R, Baltimore D (1980) Abelson-murine-leukemia-virus-induced tumors elicit antibodies against a host cell protein P50. J Virol 36:547–555

Rotter V, Boss MA, Baltimore D (1981) Increased concentration of an apparently identical cellular protein in cells transformed by either Abelson murine leukemia virus or other transforming agents. J Virol 38:336–346

Scher CD (1978) Effect of pseudotype on Abelson-virus- and Kirsten-sarcoma-virus-induced leukemia. J Exp Med 147:1044–1053

Scher CD, Siegler R (1975) Direct transformation of 3T3 cells by Abelson murine leukemia virus. Nature 253:729–731

Schultz AM, Lockhart SM, Rabin EM, Oroszlan S (1981) Structure of glycosylated and unglycosylated *gag* polyproteins of Rauscher murine leukemia virus: carbohydrate attachment sites. J Virol 28:581–592

Sefton BM, Hunter T, Raschke WC (1981) Evidence that the Abelson virus protein functions in vivo as a protein kinase that phosphorylates tyrosine. Proc Natl Acad Sci USA 78:1552–1556

Shields A, Otto G, Goff S, Baltimore D (1979) Structure of the Abelson murine leukemia virus genome. Cell 18:955–962

Shilo B-Z, Weinberg RA (1981) DNA sequences homologous to vertebrate oncogenes are conserved in *Drosophila melanogaster.* Proc Natl Acad Sci USA 78:6789–6792

Shinefeld LA, Sato VL, Rosenberg NE (1980) Monoclonal rat anti-mouse-brain antibody detects Abelson murine leukemia virus target cells in mouse bone marrow. Cell 20:11–17

Shinnick TM, Lerner RA, Sutcliffe JG (1981) Nucleotide sequence of Moloney murine leukemia virus. Nature 293:543–548

Siden EJ, Baltimore D, Clark D, Rosenberg N (1979) Immunoglobulin synthesis by lymphoid cells transformed in vitro by Abelson murine leukemia virus. Cell 16:389–396

Siegler R, Rich MA (1966) Pathogenesis of murine leukemia. Natl Cancer Inst Monogr 22:525–597

Siegler R, Zajdel S, Lane I (1972) Pathogenesis of Abelson-virus-induced murine leukemia. J Natl Cancer Inst 48:189–218

Silverstone A, Sun L, Witte ON, Baltimore D (1980) Biosynthesis of murine terminal deoxynucleotidyl transferase. J Biol Chem 255:791–796

Sklar MD, Shevach EM, Green I, Potter M (1975) Transplantation and preliminary characterization of lymphocyte surface markers of Abelson-virus-induced lymphomas. Nature 253:550–552

Srinivasan A, Premkumar-Reddy E, Aaronson SA (1981) Abelson murine leukemia virus: molecular cloning of infectious integrated proviral DNA. Proc Natl Acad Sci USA 78:2077–2081

Twardzik DR, Todaro GJ, Marguardt H, Reynolds FH Jr, Stephenson JR (1982) Transformation induced by Abelson murine leukemia virus involves production of a polypeptide growth factor. Science 216:894–896

van de Ven WJM, Reynolds FH Jr, Stephenson JR (1980) The nonstructural components of polyproteins encoded by replication-defective mammalian transforming retroviruses are phosphorylated and have associated protein kinase activity. Virology 101:185–197

Waneck G, Rosenberg N (1981) Abelson leukemia virus induces lymphoid and erythroid colonies in infected fetal cell cultures. Cell 26:79–89

Watanabe SM, Witte ON (1983) Site directed deletions of Abelson murine leukemia virus define transformation-essential and lethal sequences. J Virol (in press)

Westin EH, Wong-Staal F, Gelman EP et al. (1982) Expression of cellular homologies of retroviral *onc* genes in human hematopoietic cells. Proc Natl Acad Sci USA 79:2490–2494

Whitlock CA, Witte ON (1981) Abelson-virus-infected cells can exhibit restricted in vitro growth and low oncogenic potential. J Virol 40:577–584

Whitlock CA, Witte ON (1982a) Long-term culture of B-lymphocytes and their precursors from murine bone marrow. Proc Natl Acad Sci USA 79:3608–3612

Whitlock CA, Ziegler SF, Witte ON (1983) Progression of the transformed phenotype in clonal lines of Abelson-virus-infected lymphocytes. Cell Mol Biol (in press)

Witte ON, Wirth D (1979) Structure of the murine leukemia virus envelope glycoprotein precursor. J Virol 29:735–743

Witte ON, Rosenberg N, Paskind M, Shields A, Baltimore D (1978) Identification of an Abelson-murine-leukemia-virus-encoded protein present in transformed fibroblast and lymphoid cells. Proc Natl Acad Sci USA 75:2488–2492

Witte ON, Sun L, Rosenberg N, Baltimore D (1979) A transacting protein kinase identified in cells transformed by Abelson murine leukemia virus. Cold Spring Harbor Symposia on Quantitative Biology 44:855–857

Witte ON, Rosenberg N, Baltimore D (1979a) Preparation of syngeneic tumor regressor serum reactive with the unique determinants of the Abelson-MuLV-encoded P120 protein at the cell surface. J Virol 31:776–784

Witte ON, Rosenberg N, Baltimore D (1979b) Identification of a normal cellular protein cross-reactive to the major Abelson murine leukemia virus gene product. Nature 281:396–398

Witte ON, Dasgupta A, Baltimore D (1980a) Abelson murine leukemia veins protein is phosphorylated in vitro to form phosphotyrosine. Nature 283:826–831

Witte ON, Goff S, Rosenberg N, Baltimore D (1980b) A transformation-defective mutant of Abelson murine leukemia virus lacks protein kinase activity. Proc Natl Acad Sci USA 77: 4993–4997

Witte ON, Ponticelli A, Gifford A, Baltimore D, Rosenberg N, Elder J (1981) Phosphorylation of the Abelson murine leukemia virus transforming protein. J Virol 39:870–878

Ziegler SF, Whitlock CA, Goff SP, Gifford A, Witte ON (1981) Lethal effect of the Abelson murine leukemia virus transforming gene product. Cell 27:477–486

# Current Topics in Microbiology and Immunology

Editors: M. Cooper, W. Henle, P. H. Hofschneider, H. Koprowski, F. Melchers, R. Rott, H. G. Schweiger, P. K. Vogt, R. Zinkernagel

## Volume 97

1982. 28 figures. IV, 204 pages.
ISBN 3-540-11118-2



## Volume 98

## Retrovirus Genes in Lymphocyte Function and Growth

Editors: E. Wecker, I. Horak

1982. 8 figures. VIII, 142 pages.
ISBN 3-540-11225-1



## Volume 100

## T Cell Hybridomas

A Workshop at the Basel Institute for Immunology
Organized and edited by H. V. Boehmer, W. Haas, G. Köhler, F. Melchers, J. Zeuthen
With the collaboration of S. Buser-Boyd

1982. 52 figures. XI, 262 pages.
ISBN 3-540-11535-8

## Volume 101

## Tumorviruses, Neoplastic Transformation and Differentiation

Editors: T. Graf, R. Jaenisch

1982. 27 figures. VIII, 198 pages
ISBN 3-540-11665-6

## Volume 102

1983. 40 figures. VIII, 180 pages
ISBN 3-540-12133-1



Springer-Verlag
Berlin
Heidelberg
New York
Tokyo

# Virology Monographs

Editor: C. Hallauer

Volume 17
D. R. Strayer, D. H. Gillespie, H. L. Orlowitz

## The Nature and Organization of Retroviral Genes in Animal Cells

1980. 50 figures, 3 tables. V, 117 pages
ISBN 3-211-81563-5

The book discusses the origin, evolution, and oncogenic basis of retroviruses. A brief introduction outlines the general structure and genetic information of retroviruses, after which their genetic origin in virogenes is discussed.
On the basis of data published here for the first time, the authors discuss the organization of virogenes, probing their function in evolution, development, and cell physiology, and weighing their possible role in cell-to-cell communication.
Interspecies transfer of retroviruses has created new Mendelian genes in a species; the authors discuss the problems of homozygosity and multiplicity of these new virogenes, some of which, recombinant with host DNA, are thought to form an oncogenic host-virus hybrid.
This monograph proposes that newly evolved repeated DNA sequences of cell chromosomes serve as integration sites for recombination with viral chromosomes and that the details of the recombination event specify the consequences of retroviral infection in terms of oncogenicity. The origins of cancer in humans are discussed from this vantage point.

Volume 18
S. Dales, B. G. T. Pogo

## Biology of Poxviruses

1981. 27 figures. VIII, 109 pages
ISBN 3-211-81643-7

The book contains an up-to-date, critical analysis of the biology of poxviruses which is accessible both to specialists and to readers with general interest in microbiology. Salient historical discoveries into the nature of poxviruses and the diseases they cause is related to contemporary investigations and discoveries. Emphasis is placed on unique molecular biological and biochemical features associated with developmental cycles of these DNA cytoplasmic viruses. The sequence and regulation of the genome replication and transcription, as they pertain to virion-associated and induced enzymatic activities, are considered in considerable detail. The programme of development involving distinctive phases of assembly is evaluated in the context of temporally coordinated synthesis, processing and modification of > 100 polypeptides constituting the virion. Phenomena related to the disease process are considered, among them genetic variability, cytopathology, transmission of infectivity and immunological responses. Attention is drawn to potential natural reservoirs of smallpox-like agents among the orthopoxviruses.

Springer-Verlag Berlin Heidelberg New York Tokyo