Textbook animal breeding

Animal breeding and genetics for BSc students

Author: Kor Oldenbroek en Liesbeth van der Waaij, 2014. Centre for Genetic Resources and Animal Breeding and Genomics Group, Wageningen University and Research Centre, the Netherlands.

Groen Kennisnet

Table of Contents

Chapter 1: Introduction to animal breeding 13

Chapter 1.1 The history of animal breeding: science and application 14

Chapter 1.2 Selection by nature 16

Chapter 1.3 Domestication and animal breeding 17

Chapter 1.4 Domestication continues 20

Chapter 1.5 Origin of animal breeding: a history of science 21

Chapter 1.6 Breeding in the 19-th century 23

Chapter 1.7 Animal breeding in the 20-th century 23

Chapter 1.8 Introduction of DNA in animal breeding 26

Chapter 1.9 Animal breeding: link to societal requirements 27

Chapter 1.10 Organization of breeding activities 29

Chapter 1.11 Relation society and breeding 30

Chapter 1.12 Results of animal breeding 33

Chapter 1.13 Negative effects of animal breeding 36

Chapter 1.14 Key issues in animal breeding 39

Chapter 2: Basics of animal breeding 40

Chapter 2.1 Set up of a breeding program 41

Chapter 2.2 DNA as carrier 43

Chapter 2.3 Structure and composition of chromosomes 45

Chapter 2.4 The transfer from chromosomes and genes from parent to offspring 46

Chapter 2.5 The expression of genes, of their alleles, in the phenotype 46

Chapter 2.6 Meiosis creates differences among offspring of an individual 49

Chapter 2.7 Relatives share similar DNA, they have a relationship 50

Chapter 2.8 Key issues in basics of animal breeding 51

Chapter 3: Reasons to keep animals determine the breeding goal 52

Chapter 3.1 Challenges for animal breeding 54

Chapter 3.2 Breeding goals depend on the production system 54

Chapter 3.3 The choice of the appropriate breed 55

Chapter 3.4 The breeding goal 56

Chapter 3.5 Breeding goals are directed towards the future and require tenacity 56

Chapter 3.6 Breeding goals consists of several traits 57

Chapter 3.7 Measurement of breeding goal traits 58

Chapter 3.8 The breeding goal determines which traits should be recorded. 58

Chapter 3.9 Weighing the different traits of the breeding goal 60

Chapter 3.10 Aspects of sustainability and economics in breeding goals 62

Chapter 3.11 Key issues in reasons to keep animals determine the breeding goal 63

Chapter 3.12 Example: Pig breeding goals 64

Chapter 3.13 Chapter Example: Breeding goals for horses 66

Chapter 3.14 Example: Dutch milk production index 68

Chapter 3.15 Example: Breeding methods of top breeders in dogs 70

Chapter 3.16 Example: Production objectives for village poultry in Ethiopia 71

Chapter 4: Collecting information for breeding decisions 72

Chapter 4.1 The value of a pedigree in animal breeding 73

Chapter 4.2 A unique identification system for animals is essential 76

Chapter 4.3 Collecting phenotypes, monogenic and polygenic traits 76

Chapter 4.4 Mean, variation, standard deviation and coefficient of variation 77

Chapter 4.5 Normal distribution of measurements 78

Chapter 4.6 Covariance and correlation 79

Chapter 4.7 Regression 80

Chapter 4.8 Measurement errors 81

Chapter 4.9 Frequency of measurements 82

Chapter 4.10 Measurements on the animal or on relatives? **82**

Chapter 4.11 Value of indicator traits **83**

Chapter 4.12 Value of information of relatives **83**

Chapter 4.13 The possibilities of DNA analysis **85**

Chapter 4.14 DNA markers **86**
Chapter 4.14.1 Parentage control 86
Chapter 4.14.2 Marker-assisted and genomic selection 89

Chapter 4.15 Key issues in collecting information for breeding decisions **91**

Chapter 5: Genetic models 92

Chapter 5.1 Phenotype and environment during life history **94**
Chapter 5.1.1 Events before birth 94
Chapter 5.1.2 Events after birth 95
Chapter 5.1.3 Events after weaning 95
Chapter 5.1.4 Events after maturity 95
Chapter 5.1.5 Events after reproductive phase 96

Chapter 5.2 Phenotype in a model **96**

Chapter 5.3 Monogenic genetic variation **97**

Chapter 5.4 Polygenic genetic variation **98**

Chapter 5.5 Variance components **100**

Chapter 5.6 Simplify the genetic model **101**

Chapter 5.7 Next generation: transmission model **102**

Chapter 5.8 Heritability **104**

Chapter 5.9 Simply estimating the heritability: parent-offspring regression **106**

Chapter 5.10 Misconceptions about the heritability **107**

Chapter 5.11 Non-genetic influences: The variance due to a Common Environment **109**
Chapter 5.11.1 Importance of common environment 110
Chapter 5.11.2 Examples of common environmental effects 110
Chapter 5.11.3 Special case of a common environmental effect: the maternal effect 111
112
Chapter 5.11.4 A special common environmental effect: (indirect) social genetic effect 112

Chapter 5.12 Key issues of genetic models **114**

Chapter 6: Genetic diversity and inbreeding 114

Chapter 6.1: What is genetic diversity? 116

Chapter 6.2: Forces that influence genetic diversity 117
Chapter 6.2.1: Loss of genetic diversity: genetic drift 118
Chapter 6.2.2: Loss of genetic diversity: selection 120
122
Chapter 6.2.3: Diversity and migration 123
Chapter 6.2.4: Increase in genetic diversity: mutation 125

Chapter 6.3: Change in diversity: inbreeding 125

Chapter 6.4: Causes of inbreeding 126

Chapter 6.5: Inevitable inbreeding 126

Chapter 6.6: Why is genetic diversity important? 127

Chapter 6.7: Toolbox:relationships 128
Chapter 6.7.1: Additive genetic relationship 129
Chapter 6.7.2: Calculation of additive relationships 130

Chapter 6.8: Additive genetic relationship using genomic information 131
Chapter 6.9: Realised additive relationship 133

Chapter 6.10: Inbreeding coefficient and relationship 133

Chapter 6.11: Additive genetic relationship when the common ancestor is inbred 134

Chapter 6.12: Inbreeding at population level: the rate of inbreeding 136

Chapter 6.13: Rate of inbreeding and active population size 138

Chapter 6.14: Predict the rate of inbreeding 140

Chapter 6.15: What rate of inbreeding is acceptable? 142

Chapter 6.16: Key issues on genetic diversity and inbreeding 142

Chapter 7: Inheritance of monogenic traits 143

Chapter 7.1 Calculation of allele frequencies 145

Chapter 7.2 Hardy and Weinberg equilibrium 145

Chapter 7.3 Random effects in matings 147

Chapter 7.4 Breeding aspects of genes with large (positive) effects 148
Chapter 7.4.1 Colour genes in ruminants 149
Chapter 7.4.2 Colour genes in pigs 151

Chapter 7.4.3 Colour genes in horses 151
Chapter 7.4.4 Coat colours in dogs 152
Chapter 7.4.5 Colour genes in poultry 153

Chapter 7.5 Breeding aspects of monogenic traits with negative effects 153

Chapter 7.6 Testing parents for monogenic traits when no genetic markers are available 155

Chapter 7.7 The value of genetic markers in testing parents for genetic defects 155

Chapter 7.8 Elimination of recessive alleles for genetic defects with genetic markers 157

Chapter 7.9 Key issues in inheritance of monogenic traits 158

Chapter 8: Ranking the animals 159

Chapter 8.1: Ranking the animals: an overview of methods 161
Chapter 8.1.1: Mass selection 161
Chapter 8.1.2: Animal Model 162
Chapter 8.1.3: Genomic selection 162

Chapter 8.2: In more detail: breeding value estimation 163
Chapter 8.2.1: The basic statistics 165

Chapter 8.3: Optimising the phenotypic information 166

Chapter 8.4: Accuracy of the breeding value: the basic concept 167

Chapter 8.5: Breeding value estimation: Mass selection 168
Chapter 8.5.1: Special case: repeated observations on a single animal 169

Chapter 8.6: Breeding value estimation: the Animal Model 170

Chapter 8.7: Breeding value estimation: the basic situation 171
Chapter 8.7.1: Breeding value estimation: the rabbit example on mass selection 171
Chapter 8.7.2: Breeding value estimation: a sheep example with information from parents 172

Chapter 8.8: Other types of information sources 172

Chapter 8.9: Examples of estimating breeding values 174

Chapter 8.10: Best Linear Unbiased Prediction 175

Chapter 8.11: Accuracy of estimated breeding values 177
Chapter 8.11.1: Effect of additional information on the accuracy 178

Chapter 8.12: Summary of breeding value estimation 179

Chapter 8.13: Effects of number of offspring on the accuracy 180

Chapter 8.14: Go for highest EBV or highest accuracy? 181

Chapter 8.14.1: Outweighing EBV and accuracy in a dairy cattle example 182

Chapter 8.15: Genomic selection 183
Chapter 8.15.1: Principle of genomic selection 184
Chapter 8.15.2: Composition of the reference population 185
Chapter 8.15.3: Accuracy of genomic selection 186
Chapter 8.15.4: The size of the reference population 187

Chapter 8.16: Key issues on ranking the animals 188

Chapter 9: Predicting response to selection 189

Chapter 9.1: Response to selection: an overview 190

Chapter 9.2: Breeding is about predicting the future 192

Chapter 9.3: Genetic response: the basic principle 193

Chapter 9.4: Response to mass selection 195

Chapter 9.5: Selected proportion and selection intensity 196
Chapter 9.5.1: Appendix: from selected proportion to selection intensity 198

Chapter 9.6: Selection response: the generalized approach 200
Chapter 9.6.1: An example: the Arabian horse 201
Chapter 9.6.2: an example: rabbit breeding 202

Chapter 9.7: Generation interval 203

Chapter 9.8: Optimising genetic gain 204

Chapter 9.9: Selection paths 206
Chapter 9.9.1: an example: beef cattle breeding 207

Chapter 9.10: More detailed selection paths 208
Chapter 9.10.1: An example: dairy cattle breeding 209

Chapter 9.11: Selection intensity and rate of inbreeding 210
Chapter 9.11.1: Special case: indirect selection 211

Chapter 9.12: Practical issues with predicting response to selection 213

Chapter 9.13: Key issues on predicting response to selection 214

Chapter 10: Selection and mating 215

Chapter 10.1: Selection criteria and mating decisions 217

Chapter 10.2: Compensatory mating 218

Chapter 10.3: Long term genetic contribution **219**
Chapter 10.3.1: Example of genetic contributions 220
Chapter 10.3.2: Relation between genetic contribution and inbreeding 221
Chapter 10.3.3: Example of the effect of a popular ram on inbreeding 221

Chapter 10.4: Breeding limitations **223**

Chapter 10.5: Genetic contributions and occurrence of recessive disorders **224**
Chapter 10.5.1: Heavy use of a sire for the frequency of genetic defects 226

Chapter 10.6: Confirmation of parenthood **226**

Chapter 10.7: Key issues on selection and inbreeding **227**

Chapter 11: Crossbreeding 228

Chapter 11.1 Definitions of a breed **230**

Chapter 11.2 Heterosis **231**

Chapter 11.3 The genetic background of heterosis **232**

Chapter 11.4 Effects of heterosis **233**

Chapter 11.5 Motivation for crossbreeding **234**

Chapter 11.6 The different crossbreeding systems and their applicability **236**
Chapter 11.6.1 Two-way cross (pure-breed cross) 236
Chapter 11.6.2 Three-way cross 237
Chapter 11.6.3 Four-way crosses 237
Chapter 11.6.4 Two-way rotation (crisscross) 238
Chapter 11.6.5 Three-way rotation (crisscross) 239
Chapter 11.6.6 Introgression 239
Chapter 11.6.7 Grading-up 239
Chapter 11.6.8 Creating a synthetic breed 240

Chapter 11.7 Key issues on crossbreeding **241**

Chapter 12: The structure of breeding programs 241

Chapter 12.1: Genetic improvement in a breeding program **243**

Chapter 12.2: Breeding programs are less or more in control **243**

Chapter 12.3: Breeding programs with a flat structure **245**
Chapter 12.3.1: Example of a breeding program with a flat structure: the KWPN program 246

Chapter 12.4: Breeding programs with an (open) nucleus **248**
Chapter 12.4.1: Example: an open nucleus breeding program: the CRV dairy program 248

Chapter 12.5: Breeding programs with a pyramidal structure **254**
Chapter 12.5.1: Example: breeding program with a pyramidal structure: the Topigs program 255

Chapter 12.6: Key issues on breeding programs **259**

Chapter 13: Evaluation of the breeding program 259

Chapter 13.1: How to measure genetic improvement? **261**

Chapter 13.2: Genetic trend **262**

Chapter 13.3: What could be influencing the realised response to selection? **263**

Chapter 13.4: Selection limits **265**

Chapter 13.5: Practical issues influencing response to selection **266**

Chapter 13.6: Genotype by Environment interaction **267**
Chapter 13.6.1: Prerequisites determine the environment 268
Chapter 13.6.2: Consequences of GxE for the breeding program 270
Chapter 13.6.3: Correlated response 272
Chapter 13.6.4: Limiting resources 273
Chapter 13.6.5: Role of the environment 275
Chapter 13.6.6: Evidence for the resource allocation model 276
Chapter 13.6.7: Correlations and genotype by environment interaction 277

Chapter 13.7: Solutions to undesirable correlations **277**

Chapter 13.8: Future expectations: where to go from here? **278**

Chapter 13.9: Balance between progress and genetic diversity **279**

Chapter 13.10: Key issues of the chapter on evaluation of the breeding program **280**

Chapter 14: Maintenance of genetic diversity 280

Chapter 14.1: Genetic diversity **282**
Chapter 14.1.1 Genetic diversity in dogs 282
Chapter 14.1.2 Genetic diversity in farm animals 283
Chapter 14.1.3 Importance of variation between breeds 283
Chapter 14.1.4 Origin of differences between breeds 284
Chapter 14.1.5 Origin of differences within breeds 285

Chapter 14.2: Example: FAO's global plan for farm animal genetic resources **286**
Chapter 14.2.1: An example of the use of a conserved breed for research 289
Chapter 14.2.2: An example of the use of a rare breed 290

Chapter 14.3: The conservation of breeds in The Netherlands **292**
Chapter 14.3.1: Revival of the native Dutch Red and White Friesian cow 293

Chapter 14.4: The use of pedigrees for measuring genetic diversity **296**

Chapter 14.5: The impact of DNA information on measuring genetic diversity **296**
Chapter 14.5.1: Genome-wide patters of diversity 297
Chapter 14.5.1.1: An example of haplotypes: B-blood groups of cattle 299
Chapter 14.5.1.2: An example of introgression: Booroola allele in Texel sheep 299

Chapter 14.6: Monitoring populations **300**
Chapter 14.6.1: Voluntary and constrained inbreeding 301
Chapter 14.6.1.1: An outcross is very effective to reduce inbreeding 301
Chapter 14.6.2: Population size 302
Chapter 14.6.3: The ideal population 303
Chapter 14.6.4: Monitoring the rate of inbreeding 304
Chapter 14.6.5: Relationship between the additive genetic relationship and inbreeding 305

Chapter 14.7: Prevention of inbreeding **306**
Chapter 14.7.1: Expansion of the size of the population 306
Chapter 14.7.2: Restrictions in the use of parents 308
Chapter 14.7.3: Mating schemes to control and manage relationship 308

Chapter 14.8: Key issues on genetic diversity **310**

Preface: prof. Johan van Arendonk, chairholder Animal Breeding and Genetics

This textbook contains teaching material on animal breeding and genetics for BSc students. The text book started as an initiative of the Dutch Universities for Applied (Agricultural) Sciences. The textbook is made available by the Animal Breeding and Genomics Centre (ABGC) of Wageningen UR (University and Research Centre) . It is written by two animal breeding scientists from Wageningen UR: Kor Oldenbroek from the Centre for Genetic Resources the Netherlands and Liesbeth van der Waaij from the Animal Breeding and Genomics Centre. Four BSc teachers contributed to this textbook by a critical review of the draft texts : Aline van Genderen from HAS-Den Bosch, Hans van Tartwijk from Van Hall-Larenstein in Wageningen, Jan van Diepen from CAH-Vilentum in Dronten en Linda Krijgsman from Inholland in Delft. Their contribution is gratefully acknowledged. Financial support for writing this textbook came from the WURKS programme of Wageningen University.

When you have questions about the text, please send an email to: kor.oldenbroek@wur.nl

When you are using parts of this textbook for publication, please read and respect the disclaimer of Wageningen UR: http://www.wageningenur.nl/en/Disclaimer.htm

I case you want to refer to material in this textbook: the proper citation is: Kor Oldenbroek and Liesbeth van der Waaij, 2015. Textbook Animal Breeding and Genetics for BSc students. Centre for Genetic Resources The Netherlands and Animal Breeding and Genomics Centre , 2015. Groen Kennisnet: https://wiki.groenkennisnet.nl/display/TAB/

What is animal breeding?

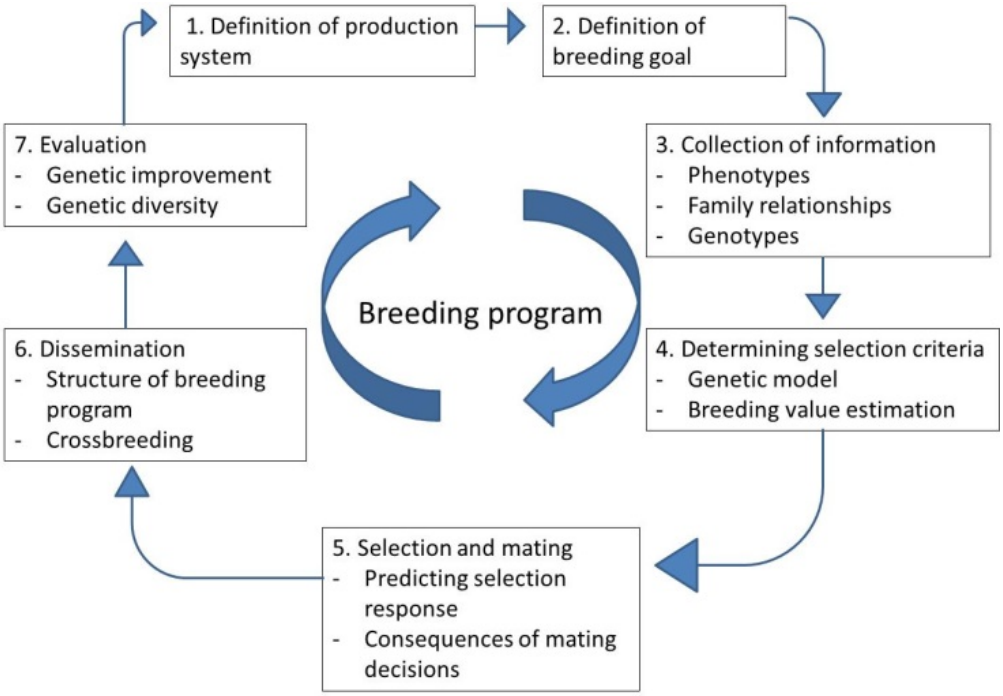
This is a book about animal breeding. But what is animal breeding? Animal breeding is about *selective breeding:* only use males and females for breeding that have passed a certain quality criterion. And with a *predefined goal* in mind: to genetically improve the population in a certain direction. So people make a plan with the *intention to select* the best animals according to a predefined list of requisites (traits), and use those selected animals for breeding the next generation so that the offspring on average will be better than the parents. In other words: selective breeding causes a shift in population average from one generation to the next. Although at first instance you may think that animal breeding involves keeping animals and making sure they reproduce, and it thus would involve optimising reproduction techniques or something along those lines, this is not the case.

Definition

Animal breeding *involves the selective breeding of domestic animals with the intention to improve desirable (and heritable) qualities in the next generation.*

Aim of this book

We start with the basics of animal breeding: the concepts of genetics necessary to understand the genetic processes used in animal breeding. Then, in the following chapters we will 'dive into the deep' and take you on a journey through all the steps that need to be taken in order to develop and run a successful breeding program (see figure 9). As animal breeder you start by defining what you want to improve in your population, collect information on the performance of the animals and their genetic relationships, determine which animals have the best genetic potential, determine what proportion you should use for breeding in order to achieve a certain genetic gain in the next generation, select the animals and mate them, and after producing the offspring evaluate whether what you set out to achieve with your breeding decisions actually happened. Each generation you breed you go through this circle of steps. So each generation you again have the opportunity to adjust these steps to some extent. You should not change the breeding goal every generation, because a single generation will not give you much genetic improvement. Breeding is more about the cumulative success of multiple generations. You can adjust the goal in response to a change in the market. You can also adjust your breeding program in response to an unwanted genetic change in your population. You should do this as soon as you find out as you don't want the cumulative effect of an undesired response to selection. In almost every chapter we will focus on a specific step in the breeding program. We will explain the main goal of that step, introduce the challenges and find out how those could be faced.

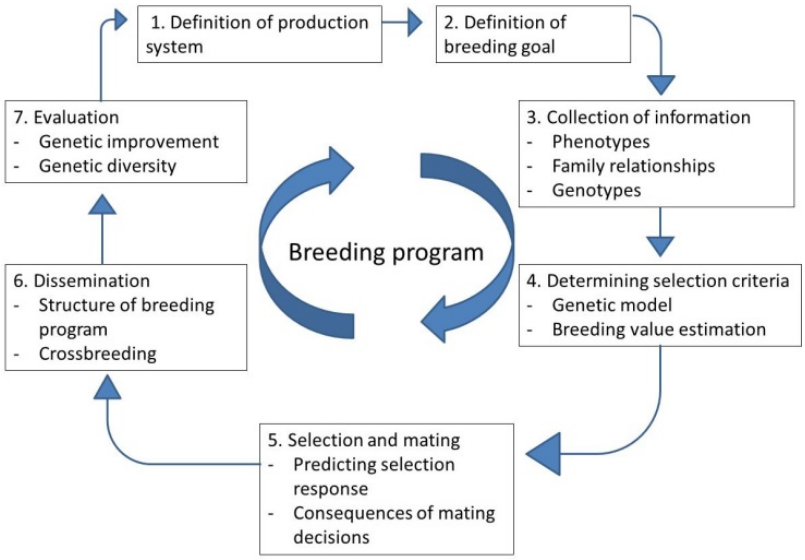


You will notice that some subjects seem to pop up in a number of chapters. That is because they are related to a number of steps and in each step they require a specific attention. That is why they are mentioned in a number of chapters instead of having one chapter just about that subject. The role of genetic relationships, for example, is such a subject. At the end of studying the book you will have gained insight in how a breeding program should be organised, what are

some critical points, and what are consequences of certain breeding decisions. The book is organised such that each chapter starts with a general description of the subject, what is its role in a breeding program, and some points of attention. Then we go a bit deeper and introduce tools (formulas) that provide results to help executing the step in the breeding program accurately. We will use a few formulas to be able to do some basic calculations.

Chapter 1: Introduction to animal breeding

In this first chapter the history of animal breeding is presented. The importance of selection by nature and important aspects of the domestication process will be described. Mankind started to create breeds accompanied with artificial selection 250 years ago. Nowadays, breeding of high productive farm animals, like cattle, pigs and poultry is in the hands of multinational companies which invest a lot of money in state of the art breeding programs. The breeding of sheep, goats, horses and companion animals, e.g. the dog is based on individual breeders collaborating in the setting of a herd book or a breeder's association. Animal breeding is aiming at the improvement of animals by changing their genetic abilities for important traits. These traits are determined by the requirements and wishes from the society which might change over time. Animal breeding is highly influenced by research and developments in population-, quantitative- and molecular genetics. Sometimes, unexpected negative effects of animal breeding are observed that require adequate corrections. A breeding program will be presented here as a circular activity. Each generation, the program starts with formulating the breeding goal and ends with a critical review of the results obtained in the next generation. The evaluation might lead to a reconsideration of the breeding goal for the next round of selection.



Chapter 1.1 The history of animal breeding: science and application

There are 5 very important aspects that should be considered in animal breeding:

1. Most importantly, obviously for selective breeding to be successful it is essential that the trait (e.g. running speed or milk production or coat colour) under selection is heritable.
2. That animals have different genetic backgrounds so that selection is possible.
3. The direction of selection is defined by humans and they decide which animals are allowed to mate and produce members of the next generation.
4. Success of animal breeding can be judged by looking at a shift in population average phenotype from one generation to the next. So animal breeding works at population level, not automatically at individual level.
5. Success of animal breeding can be measured as the cumulative result of multiple generations of selection. Breeding decisions are made with the future in mind.

Definitions

*A **trait** is "a distinguishing phenotypic characteristic, typically belonging to an individual". In practice this means anything you can record or measure on an individual.*
*A **phenotype** is that what you observe or measure on the animal for a certain trait. It can depend both on the genetic background of the animal (provided it is heritable) and external circumstances such as level of nutrition*

Heritable traits

Being able to predict the success of animal breeding relies on one very important factor that we still need to discuss: why does performance in offspring resemble that of the parents? Selective breeding will only be successful in case the trait under selection is *heritable*. Because only a fraction of the animals is selected for breeding, so is allowed to produce offspring, and because the trait is heritable, the performance in the offspring will resemble that of the parents. Therefore only the best parents are used for breeding and the average of the next generation will be better than that of the current. A trait is heritable if the performance for that trait, at least in part, depends on the genetic make-up (DNA) of an animal. Differences in performance between animals can (partly) be explained by genetic differences between animals. More details on what this *heritability* involves will follow later in this book. **The scope of this book and this chapter** In summary, animal breeding relates to intentional selection by humans based on animal performance in a certain environment for predefined and heritable traits. In most practical animal breeding schemes selection will be on more than one trait simultaneously. The animals that are superior in this combination of traits will be selected as breeding animals. In general this combination of traits will consist of traits related to performance (e.g. milk production, number of eggs, growth, sport performance), health, and reproduction. The theory behind selection for a combination of traits easily becomes very complicated. In this book we, therefore, explain the theory behind animal breeding using single trait selection. In the rest of this chapter we will give you a brief history of animal breeding,

starting from scratch (so from domestication). You will see that developments in animal breeding have gone hand in hand with developments in society. Then we will look into the current situation and main challenges. And we will also try to peek into the future: what are the expected developments in society and how will that influence animal breeding decisions? But first we will look back at how it all started: with domestication.

Chapter 1.2 Selection by nature

It sounds as if animal breeding is all in the hands of the humans. Compared to natural populations this indeed is the case, as we decide which animals are allowed offspring and which are not: selective breeding or in other words *artificial selection*. However, as in natural populations there is another force that plays an important role and that is the force of *natural selection*. In natural selection it is not us but the environment that determines survival and reproductive success of animals. So after we have decided which animals we intend as parents, they still need to be able to survive until reproductive age and to be able to reproduce successfully. As you can imagine, natural selection also results in directional change in the population average. Animals *adapt* to their environment and the ones who can do that best will be the most successful in survival and reproduction. In other words: in natural selection the direction of selection is on adaptation to the environment.

Definition

Natural selection *is the process whereby animals that are better adapted to their environment have a higher change to survive and produce more offspring than less adapted animals. The next generation thus, on average, will be more adapted than the current generation.*

Even though animal breeding is defined as intentional selection by humans, you can see that natural selection will also play a role. In some cases natural selection will even work in the opposite direction of selective breeding. In those cases without human intervention the animals with the desired qualities will be less successful in surviving and/or producing offspring. For example, the fact that in many cows there is a negative relationship between high milk production and being able to get pregnant, shows that animals with the desired quality: high milk production, are less likely to produce offspring unless there is extra effort put into it by the farmer. Also, the fact that very high producing cows often have health problems indicates that their chance of producing offspring is reduced compared to their more average producing 'sisters'. Selective breeding often competes with natural selection. We have become so familiar to the fact that some of the best animals in a breed require assistance with some aspects of their survivability and/or reproduction that we think it is normal. Domestic animals are 'created' by humans and to maintain that we accept certain disadvantages. But how far should we go?

For example, some breeds of dogs and beef cattle have been selected such that they are excessively broad in the head and/or shoulders. But being broad shouldered (or big headed) creates birth problems. Without human intervention such as assistance at delivery, or even a caesarean section, both mother and offspring would die. In other words: it is good to keep an eye on unwanted consequences of selective breeding.

Chapter 1.3 Domestication and animal breeding

Now we will discuss the history of the domestic animals and animal breeding. When did domestication start, when and how did selective breeding become organised, what is the role of science in animal breeding, how have animal breeding techniques evolved until now, and what is the role of society and culture in all this? To start with the first question: how and when did it all start?

Definition

Domestication is the process of conversion of wild animals to domestic use.

Domestic animals need to live in (close) association with humans, therefore they have to become tame. They also have to meet the expectations of their owners for the purpose of keeping them. This can be achieved by selective breeding. Expectations of owners will change in time, followed by a change in selective breeding plan. Domestication often has resulted in a type of animals that has become quite different from their wild counterparts. As a result domestication often also involves the development of a dependency on humans so that the animals lose their ability to live in the wild.

Domestication of the dog

The first animal species to be domesticated was the dog. Estimates of when this happened vary a lot, but it was approximately 12,000 years ago. An appealing theory of how this happened is that when people started to settle down and become farmers, they also started to accumulate waste. The tamer than average wolves were brave enough to eat from that waste and thus had a secure source of food. This was an advantage, so natural selection pressure was on being not very afraid of humans. Eventually, a kind of symbiotic relationship developed, where these ancestors of the dog started to perform 'tasks' like warning the humans for approaching danger, helping in hunting, provide warmth, etc. and in return these animals would receive food security. This type of symbiotic relationship is still present in village dog populations in Africa and Asia, and also in some Southern European countries. It is believed that our current domestic dog breeds originate from these village dogs. There is evidence that genetically the village dogs are in between the wolves and the dogs.

Domestication of other species

The symbiotic relationship, such as between human and the ancestors of the dog, most likely is quite unique for dogs. Other types of animals may have been domesticated more forcefully. They were captured and put in an enclosure or tied up, at least during the night, and were only allowed out to graze or scavenge under supervision of a herdsman. Only those animals that were not aggressive, but also not too shy, managed to adapt to these new circumstances. So (mainly natural) selective breeding also in these cases was on temperament. In the table you will find a list of domestic animals with their approximate time and location of domestication. This is approximate, because especially for ancient times it is difficult to make an accurate estimate. But even recent events are not always straightforward. Because when do you call an animal domesticated? And what if it happened in more than one place simultaneously but independently of each other?

Table 1. List of domesticated animals in early times.

Species	Latin	Date	Location
Dog	*Canis lupus familiaris*	>30,000 BC	Eurasia
Sheep	*Ovis orientalis aries*	11000 - 9000 BC	Southwest Asia
Pig	*Sus scrofa domestica*	9000 BC	Near East, China, Germany
Goat	*Capra aegagrus hircus*	8000 BC	Iran
Taurine cattle	*Bos primigenius Taurus*	8000 BC	India, Middle East, North Africa
Zebu cattle	*Bos primigenius indicus*	8000 BC	India
Cat	*Felis catus*	7500 BC	Cyprus and Near East
Chicken	*Gallus gallus domesticus*	6000 BC	India and South East Asia
Llama	*Lama glama*	6000 BC	Peru
Guinea pig	*Cavia porcellus*	5000 BC	Peru
Donkey	*Equus africanus asinus*	5000 BC	Egypt

Domesticated duck	*Anas platyrhynchos domesticus*	4000 BC	China
Water buffalo	*Bubalus bubalis*	4000 BC	India, China
Horse	*Equus ferus caballus*	4000 BC	Eurasian Steppes
Dromedary	*Camelus dromedaries*	4000 BC	Arabia
Honey bee	*Apis*	4000 BC	Multiple places
Silkworm	*Bombyx mori*	3000 BC	China
Reindeer	*Rangifer tarandus*	3000 BC	Russia
Rock pigeon	*Columba livia*	3000 BC	Mediterranean Basin
Goose	*Anser anser domesticus*	3000 BC	Egypt
Bactrian Camel	*Camelus bactrianus*	2500 BC	Central Asia
Yak	*Bos grunniens*	2500 BC	Tibet
Asian elephant	*Elephas maximus*	2000 BC	Indus Valley civilisation
Alpaca	*Vicugna pacos*	1500 BC	Peru
Ferret	*Mustela putorius furo*	1500 BC	Europe
Common carp	*Cyprinus carpio*	Unknown	East Asia
Domesticated turkey	*Meleagris gallopavo*	500 BC	Mexico
Goldfish	*Carassius auratus auratus*	Unknown	China
European rabbit	*Oryctolagus cuniculus*	600	Europe
Japanese Quail	*Coturnix japonica*	1100–1900	Japan

Canary	*Serinus canaria domestica*	1600	Canary Islands, Europe
Fancy rat	*Rattus norvegicus*	1800s	United Kingdom
Fox	*Vulpes vulpes*	1800s	Europe
European Mink	*Mustela lutreola*	1800s	Europe
Cockatiel	*Nymphicus hollandicus*	1870s	Europe
Zebra Finch	*Taeniopygia guttata*	1900s	Australia
Hamster	*Mesocricetus auratus*	1930s	United States
Silver Fox	*Vulpes vulpes*	1950s	Soviet Union
Ball python	*Python regius*	1960s	Africa
Red Deer	*Cervus elaphus*	1970s	New Zealand
Atlantic Salmon	*Salmo salar*	1969	Norway
Atlantic Cod	*Gadus Morhua*	On going	Norway

Chapter 1.4 Domestication continues

Domestication is not only of ancient times. It is still happening today! It often involves species that are used for human consumption or for companion, and that become rare in their natural habitat. To prevent extinction, people try to breed them in captivity. In return benefits are easy access to the animals, and the possibility to optimise the animals through selective breeding to the (expected) demands of the market. And 'market' is a very wide concept: demand for food of animal origin, but also demands of farmers for, for example, dairy cows that can be milked by a robot, demand for dogs that can perform certain tasks, demand for horses with certain temperament, etc. There are some (rare) occasions where new tasks are invented for certain animal species, potentially followed by domestication. A recent example may be that of the use of 'sniffer wasps' for explosives detection. These wasps are trained to smell different types of

explosives and, subsequently, used in places where it is too dangerous for people (or dogs) to go to. Because wasps are small and can fly they can go places where robots can't go. Possibly these wasps in the future will be different from wild wasps. This is due to directional selection on, for example, trainability.

Prerequisites for domestication Domestication is not always successful. Despite many attempts, the zebra, for example, has not been domesticated. Even though it is closely related to the horse and the donkey and you can keep it in an enclosed area where it will survive and reproduce, apart from the exceptional case, it has not been successfully tamed. Several generations in captivity and some selective breeding did not make the zebra genetically tame so that it can be ridden. Why is that? People are not sure, but there is a list of prerequisites for successful domestication that seem to hold. The zebra may not meet one or more of them. The apparent prerequisites are:

1. The animals should be able to adapt to the type of feed they are offered by humans. This may be different (in diversity) from what they were used to in the wild.
2. Animal must be able to survive and reproduce in the relatively closed quarters of captivity. Animals that need a very large territory are not suitable to be domesticated.
3. Animals need to be naturally calm. Very skittish or flighty animals will be hard to prevent escaping.
4. Animals need to be willing to recognise humans as their superior, which means they must have a flexible social hierarchy.

Animal species that do not meet all the above criteria will be very difficult to domesticate. But a fair number of animal species have been domesticated, and the number is still increasing. The early domestication probably was mainly driven by natural selection: the animals that managed best were most successful in producing the next generation. Real selective breeding is of fairly recent origin.

Chapter 1.5 Origin of animal breeding: a history of science

The start in the 18-th century

Until roughly the 1700's animal breeding, as in selective breeding, did not really exist. Of course people mated their animals with animals in the neighbourhood that they liked. There was no *systematic way* of selecting animals for reproduction, based on *predefined characteristics* that did not change from mating to mating, but remained similar in time. In Europe, the origin of animal breeding lays in the United Kingdom. It was Sir Robert Bakewell (1725 – 1795) who introduced keeping accurate records of performance of animals so that objective selection became possible. He used inbreeding (mating of related animals with similar traits) to fix certain

characteristics in animals and he also introduced *progeny testing*: the method of evaluating performance of the first (small) group of progeny and use that information to select the best father of future progeny. He promoted the idea to 'breed the best to the best'. Bakewell developed the New Leicester sheep from the old Lincolnshire breed. The New Leicester had good quality fleece and a good fatty shoulder that was popular at the time. Bakewell also noticed that Longhorn cattle were growing well and used less feed compared to other cattle. So he developed that further in order to grow more meat efficiently. It is amazing he did this without knowing anything about genetics.

Establishment of herdbooks

With time the number of people increased who were using the selective breeding approach introduced by Bakewell. With the growing number of generations of selective breeding, it became increasingly difficult to remember the relationships between the animals, especially further back in the pedigree. This was the reason to start recording pedigree on paper, so that correct information could be reproduced and it could be proven that an animal was of a certain breed. The first herdbook was for the thoroughbred horse and was established in England in 1791. This book did not contain all pedigree, but only those of horses that were winning important races. Following the race horses, the Shorthorn cattle (1822) were next to start a herdbook. In the rest of Europe, herdbooks only started to be established in from 1826 onwards for horses (in France), and from 1855 onwards for cattle (also in France). The first international herdbook was established for the American Berkshire pigs in 1876. The first dog in the Netherlands was registered by the Koningklijke Nederlandsche Jachtvereeninging Nimrod (predecessor of the Raad van Beheer op Kynologisch Gebied in Nederland) in 1874. After the turn of the century animal breeding within herdbook settings became standard.

Creation of breeds

With the establishments of herdbooks, breeds were formed. There is still debate on what is a true definition of the term 'breed'. This is nicely illustrated in dog breeding by the fact that the Fédération Cynologique Internationale (FCI), the international federation of kennelclubs, which are national organisations across herdbooks, recognises 339 separate breeds, while the English Kennelclub recognises 210 breeds, and the American Kennelclub even only 162.

Definitions

*A **breed** is a group of animals of a certain species that through generations of selective breeding has become uniform in performance, appearance, and selection history*
*A **species** is the largest group of animals that are capable of interbreeding and producing fertile offspring*

It is interesting to realise that these herdbooks were established without any knowledge about genetics. Breeders had a feeling about inheritance and that was sufficient to invent this selective breeding.

Chapter 1.6 Breeding in the 19-th century

In 1859, Charles Darwin (1809 – 1882) published his book 'On the origin of species', based on the findings that he collected during his voyage on 'the Beagle'. He discovered the forces of natural selection. He also concluded that the individuals that fit best in their environment have the highest chance to survive and reproduce: they are the fittest. Consequently, different environments result in different directions of selection pressure. He based this on his findings on the Galapagos islands, where finches on one island were different from finches on the next island. His conclusion was that the difference in food source, predators present, etc. between the islands had made the finches develop differently over very many generations. They adapted to their specific environments.

Darwin translated his ideas to domesticated species as well: ***"We cannot suppose that all the breeds were suddenly produced as perfect and as useful as we see now them; indeed, in several cases, we know that this has not been their history. The key is man's power of accumulative selection: nature gives successive variations; man adds them up in certain directions useful to him. In this sense he may be said to make for himself useful breeds"***
C. Darwin. On the Origin of species (1859, p.30)

Still, Darwin did not know about the basic laws of inheritance. It was the monk Gregor Mendel, who in 1865 published the results of his studies of genetic inheritance in garden peas. He showed that genetic material is inherited from both parents, independently of each other. And that each (diploid) individual thus carries 2 copies of the same gene, of which only 1 is passed on to their offspring. Which one is a result of chance (*independent assortment)*. He also showed that these gene copies (alleles) can be dominant (only 1 copy determines the expression of the gene), recessive (2 copies are required for expression), or additive (a copy of both alleles result in an expression that is intermediate to that of having 2 copies of either of the alleles). These findings had no immediate impact on animal breeding and were not recognised as important until 1900.

Chapter 1.7 Animal breeding in the 20-th century

Most of the animal breeding theory we are still using today, was invented in the first half of the 20-th century. The statistician R. A. Fisher (1890 – 1962) showed that the diversity of expression of a trait could depend on the involvement of a large number of so-called Mendelian factors (genes). He published a lot related to statistics and animal breeding, but his keynote paper came out in 1918. Fisher, together with Sewall Wright (1889 – 1988) and J.B.S. Haldane, were the founders of theoretical population genetics. Thomas Hunt Morgan (1866-1945) and co-workers connected the chromosome theory of inheritance to the work by Mendel and created a theory where chromosomes of cells were believed to carry the actual hereditary material.

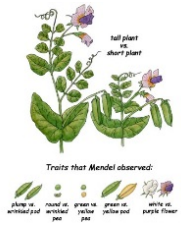
Morgan won the Nobel prize for it in 1933.In the first half of the 20th century Iowa State University in Ames, Iowa, USA was the place to be. It was home to Jay L. Lush (1896 – 1982), who is known as the modern father of animal breeding. He advocated that instead of subjective appearance, animal breeding should be based on a combination of quantitative statistics and genetic information. His book 'Animal Breeding Plans' that was published in 1937 greatly influenced animal breeding around the world. Lanoy Nelson Hazel (1911-1992) was inspired by Lush's book and started working for him, also in Ames. He received his PhD degree in 1941 and in that PhD thesis he developed the selection index theory, a method used for decades to determine what weights should be put on the different traits under selection. In the process of developing this method he also came up with a concept on how to estimate genetic correlations. This is essential for assigning the proper weight to selection traits. Hazel also developed a method using least squares, a statistical technique, for more complicated data with unequal numbers of subclasses as often occur in animal data. Until then, statistical techniques by Hazel were used to optimise weighing the performances for various traits in animals to select those with the most optimal combination. The estimated breeding value (ebv) was only developed later by the statistician C. R. Henderson (1911 – 1989), who was a student of Hazel in Ames. The estimated breeding value made it possible to rank the animals according to their estimated genetic potential (the ebv), which resulted in more accurate selection results and thus a faster genetic improvement across generations. Henderson further improved the accuracy of the estimated breeding value by deriving the best linear unbiased prediction (BLUP) of the ebv in 1950, but the term was only used since 1960. He also suggested to integrate the full pedigree of the population to include genetic relationships between individuals. This way performance of relatives could be included in estimating the breeding value of an individual. The so-called animal model was born. Unfortunately in those days the computer power was too limited to be able to also calculate the breeding values using the animal model. The practical implementation thus had to wait until the later 1980's. Current great minds that have developed a way to incorporate large scale DNA information that has become available in animal model (BLUP) theory to estimate the so-called genomic breeding values are Theo Meuwissen (currently professor in Ås, Norway) and Mike Goddard (currently professor in Melbourne, Australia).

Gallery of persons related to developments in animal breeding

Founder of animal breeding

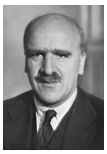

Sir Robert Bakewell

Mendelian inheritance



Georg Mendel

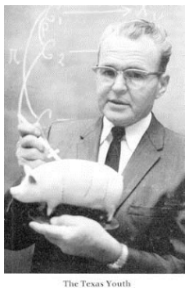
Quantitative genetics

		
R.A. Fisher	S. Wright	J.B.S. Haldane

Father of animal breeding

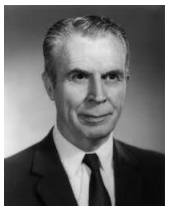
J. L. Lush

Selection index theory



L.N. Hazel

BLUP animal model



C. R. Henderson

Genomic selection


T. Meuwissen, M. Goddard

Chapter 1.8 Introduction of DNA in animal breeding

Until 1953, scientists used statistics and presumed mechanisms to make predictions about inheritance. Nobody knew what exactly was the mechanism behind it. But in 1953 Watson and Crick, using research results from Wilkins and Franklin, discovered the double helix structure of DNA and together they won the Nobel prize for their discovery. Since the discovery of the DNA structure a lot has happened. In the beginning studying the DNA was very labour intensive and thus also very costly. Nowadays robots can perform large scale genotyping e.g. of more than 60,000 genetic markers on thousands of individuals within very limited amounts of time. A

genetic marker can be considered a kind of 'flag' on the genome. Its location and composition ('looks') are known. These genetic markers can, for example, be used to compare animals based on the looks of the various flags.

Main idea behind genomic selection is that the association between the DNA make-up and performance of animals can add to the estimated breeding value, or even replace it. Because you don't have to wait until the phenotype can be measured on the animals anymore as you have the associated DNA information, you can select animals already at very early age and you don't have to wait until they become adult. You can also use this, for example, for traits that are difficult to measure such as disease related traits. The aim would be to prevent diseases in as many animals as possible. It would be highly desirable if you would only need to infect a limited number of animals and measure their response to the infection, link that to their DNA, and use that estimated link to predict the sensitivity of other animals to that disease based on their DNA, without having to infect them. More about genomic selection later in this book. Meuwissen and Goddard (and co-workers) even take it a step further and already work on methods to incorporate full genomic sequences (all DNA in an individual) in estimating genomic breeding values. But large scale availability of full sequences is still something of the future, as that is currently still too expensive. But developments go fast you have to be prepared for the future.

Chapter 1.9 Animal breeding: link to societal requirements

In the Netherlands, and other developed countries, animals breeding, and especially farm animal breeding, has developed to a professional industry with modern technologies, large scale data collection, and analyses. This has resulted in very efficient and effective breeding programs, producing many thousands of genetically improved animals for different parts of the world. However, for this large scale animal breeding a big infrastructure is required, combined with high quality data collection, large computing capacity, and highly educated people to run the breeding program. This level of organisation is not available in all parts of the world (yet). Especially in developing countries the situation is similar to the way it was in Europe before the Industrial Revolution (that started around 1750). In those developing countries animals are kept for multiple purposes: to produce food, labour (traction power), warmth, for their hides and/or wool, their manure is used as fertiliser for the land and also as fuel for the fire, as savings account (sell an animal when needed), and to increase social status (more is better). The surplus of animals or animal products are sold on the market. Also in developing countries efforts are made to improve the productivity of the animals to increase the welfare of people of their, usually poor, owners. For us it is self-evident that animals of a certain breed are uniform in type and performance, that selective breeding usually is well organised and structured, and

that the required infrastructure is present. In many developing countries this is not the case (yet). However, an increasing number of selective breeding programs has been developed in a large range of countries, and many are quite successful. The increasing level of education in those countries is an important factor in that success.

Developments that affect animal breeding

A lot happened in the twentieth century, with its influence on animal breeding. The industrial revolution changed society tremendously. People moved from the farms to the towns to work in the factories, so fewer farmers became available for food production. Increased production per farm was needed. Simultaneously, technical developments went fast. The train was introduced in the end of the 1800's, the car at the beginning of the 1900's, and the plane soon after that. The use of the tractor on farms became more common in the 1950's. Around WWII artificial insemination was introduced in cattle, so that a more offspring could be produced by a single father. With the introduction of storing semen in liquid nitrogen, the possibilities to extensively use a single father in a (very) large area became even bigger. Introduction of these technical developments had its impact on the use of animals. This was especially the case for oxen and horses. Where oxen and horses were the main source of labour to work on the land, the introduction of the tractor made those animals superfluous. Oxen were no longer retained, but slaughtered at younger age. Horses went through hard times, as there was not much use for them. In took until the 1960's for the sport horses to become popular. In the past, the sport was only performed by army officers and rich men. When horse riding became more popular among women, and especially when it became available for more than just the very rich, the number of horses increased again.

Food production important in many species

After WWII it was very clear that food production should have very high priority. The intention was that food of sufficient quality and quantity should be available to everyone at an affordable price. Therefore animals should become more productive. This could be achieved by selective breeding, but also by adjusting the management. Pigs en chickens were kept under controlled circumstances so that the feed they ate was of equal quality and feed intake could be assigned to production, and not to unnecessary things like keeping warm, or fighting infections. Consequently, animals were kept under very efficient and controlled circumstances, so confined and indoors. Farms specialised more into either (only a few types of) crops or animals. In those days it also became easier to transport large amounts of goods over long distances, especially over sea. In the Netherlands this meant it became possible to import large quantities of tropical crops like tapioca and soja. These relatively cheap products were used as replacement for (expensive) grains as raw material for production of concentrates. Where pigs used to be animals that were kept in combination with other types of farming because they could digest a lot of leftovers, the availability of these concentrates made it possible to start specialised pig farms. And the same story holds with respect to poultry farms. Because oxen

were no longer needed, and growing them to slaughter weight was quite expensive, the calves that were not used for replacement in the dairy industry were becoming superfluous. Some farmers specialised in housing these calves and sell them at fairly young age and with that a new branch was born: growing veal calves.

Chapter 1.10 Organization of breeding activities

Start of the organization of breeding activities

Pigs, horses and cattle herdbooks were organised at regional level. Owners of potential breeding males brought those animals to shows where they were judged on looks and owners of female animals could see them to decide which one to breed with. The pig breeders were the first to stop showing their boars in public in the late 1960's, followed by the dairy cattle breeders and their bulls in the 1970's. Main reasons were to prevent infectious diseases to spread, and that production figures had become more important than looks. For export (also of semen) it is important that you can prove that animals have never been in contact with certain pathogens. Stallions are still shown in shows, and also used in riding horse competition. Nowadays all horses entering shows and/or competition are vaccinated. In horses each breed has its own herdbook. There are some exceptions, especially in sport horse breeding, where it is not so much the breed, but more the type of horse that is of interest. The Dutch warmblood (KWPN), for example, has developed from a herdbook registry of originally Dutch horses to a registry of horses bred in the Netherlands. It is aimed at breeding highly successful sport horses. This herdbook is open and market driven, rather than focussing on pure breeding. Use of stallions from other countries is allowed, provided they have been approved of by the KWPN.

Organization of breeding nowadays

In commercial farm animal breeding the situation changed quite drastically. From farmers owning males and females, to artificial insemination (AI) companies owning the males (in cattle) and later also breeding females (in pigs). The number of herdbook owners has decreased dramatically in the past decades from many regional herdbooks to a single national herdbook (in cattle) or international breeding companies (in pigs). At first herdbooks were merged to combine forces, but later also because bigger companies took over smaller ones. Poultry breeders started to specialise completely to breeding of laying hens or broilers and to sell that as a product, rather than eggs or meat. They were developing their own pedigree registration system. In cattle breeding the males are owned by a company, but (most of the) females are privately owned. The breeding company sells semen as main product, not animals. So in a way they sell half of the final product: the calf, the other half, the oocyte (i.e. the cow), is privately owned. The situation is very different for pigs and poultry breeding companies Their final products are animals. This means that if other people get hold of their original animals, they can duplicate the product and not have the costs for development. This is an important reason why pig and poultry breeding companies don't sell purebred animals: then they would give

away their genetics. By selling crossbred animals or semen of crossbred animals, they don't sell the purebreds so no one can reproduce the final product. Commercial companies often keep multiple breeds or lines. Pig and poultry breeding companies keep multiple lines to combine in final products, so they also keep multiple pedigree registrations. But when you buy their products you will not receive specific pedigree information. Pedigree records are only for their own use in selective breeding. In the Netherlands, there are two companies breeding pigs: TOPIGS and Hypor (Hendrix-Genetics). In laying hens (ISA) and Turkey (Hybrid), Hendrix-Genetics is also the owner. Broiler breeding is in the hands of a US company called Cobb, and the division in the Netherlands belongs to Cobb Europe. Globally the number of breeding companies is decreasing, especially related to poultry breeding. In laying hens and in broilers there are only 2 major breeding companies left. In pigs there are a few more, but only 5 larger ones. In cattle there is strong international exchange of semen for many breeds. So in practice, especially in the Holstein Friesian, the population has become one, but with overlapping registries. In summary: farm animal breeding is an increasingly global industry.Farm animal breeding is very different from breeding of horses and companion animals. Especially sport horse breeding studbooks are operating increasingly at an international and competing level. In the Netherlands the largest sport horse breeding studbook is the KWPN. They are very successful internationally in breeding horses for dressage and show jumping. The stallions are approved by the studbook after passing strict selection criteria. The KWPN is an open studbook, which means it does not limit itself to horses that were born and bred in the Netherlands, but it also approves stallions from other studbooks, provided they pass the selection criteria.

Chapter 1.11 Relation society and breeding

So where are we today? Changes that occurred in the field of animal breeding have always been connected to changes in society. It was a matter of a combination of availability of techniques and market demand. So what are current changes that may have a reflection on animal breeding? One big difference in the developed world today compared to 30 years ago is that people are relatively rich and/or food has become relatively cheap. The percentage of average income spend on food in the Netherlands has decreased from 24% in 1980 to only 9.8% in 2010. On average people in European countries spend around 12% of their income on food, whereas people in Russia spend 31%, in India 36%, and in some East African countries even more than 50% (data FAO).

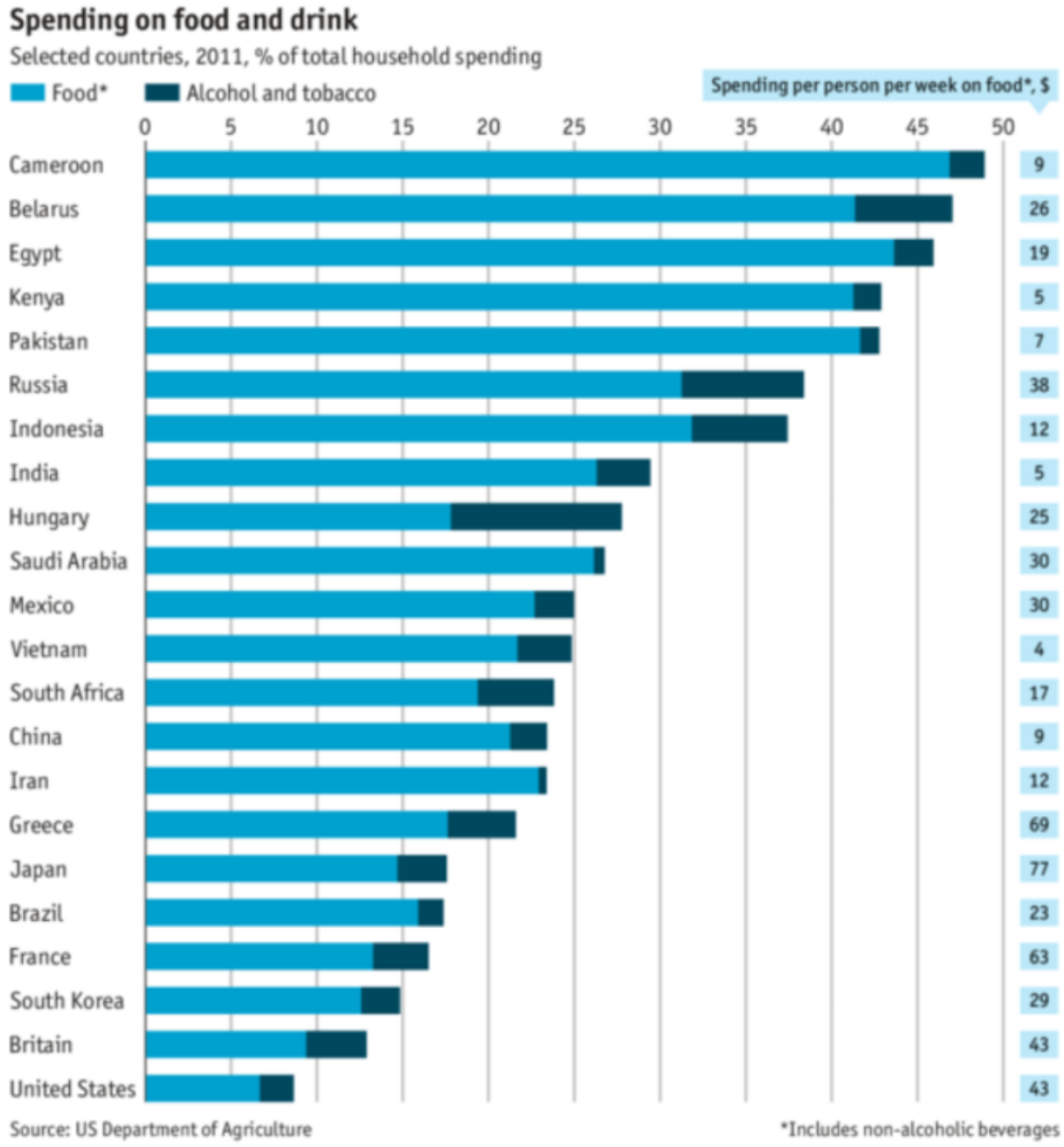


Figure 1: Spending of household income towards food and drink

Cheap food means that you can get more for your money. In the western world there is an increasing concern of people for how their food was produced. It should be healthy and natural and produced locally. In addition animal products should also be produced welfare friendly. In our culture this is considered very normal, but it is a definite sign of wealth that people can afford to care these issues. In poorer areas of the world the main worry is about having enough food of sufficient quality, how it was produced is not a priority.

Challenges for the future

The world population is growing rapidly, especially in urban areas (see Figure). All these people need to be fed. At the moment we are using twice the resources that should be used to ensure that our planet will survive. At the same time about 20% of the food is wasted in developed countries, while there is still considerable food shortage in developing countries. Challenge for the future is to reduce wastage in the developed world, increase food availability in the developing world, and do that with a decreased carbon footprint. Additional challenge lays in the fact that fossil fuel is replaced by biofuel. A lot of crops like wheat or sugar cane are used for biofuel production at the expense of food production.

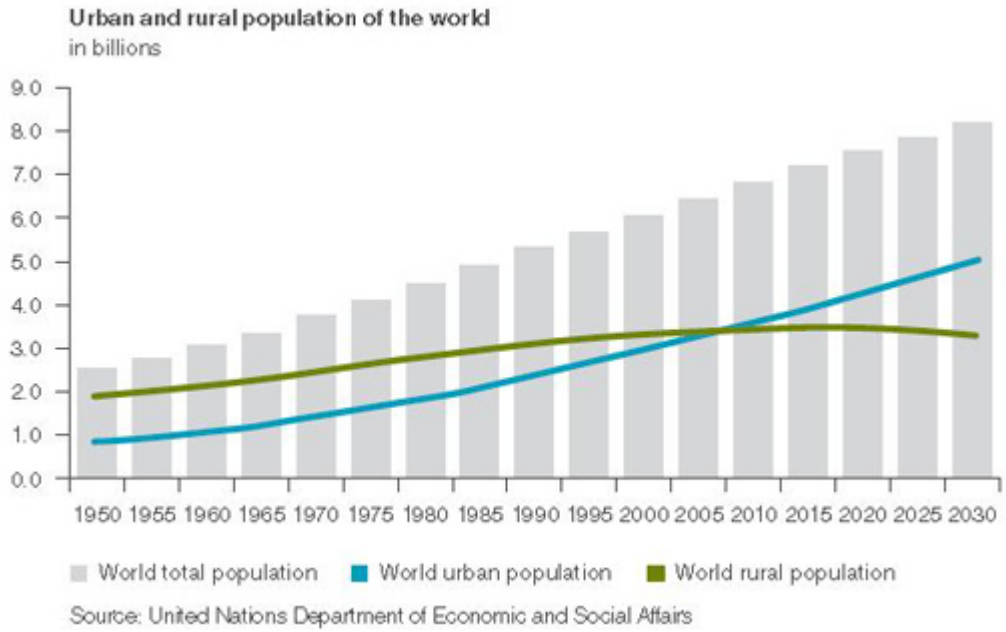


Figure 2: Growth of urban population

Summarising, there are a number of developments that the animal breeding industry need to anticipate to. As farm animal breeding has become a global industry, breeding companies need to develop products that fit the various markets. In the Netherlands there is a growing group of customers who want to spend their money on products with extra attention for environmental and animal friendly production. In other parts of the world the main concern is still to feed the family, and there the emphasis is more on price than on production method. Animal breeding companies supply to both markets. But animal breeding companies also take the moral obligation to breed animals that can produce the products for the specific markets taking into account the carbon footprint that they leave. For example, research is ongoing to see whether pigs and chickens can function on feed that contains waste products from the biofuel industry. And in cattle breeding methods are being developed to reduce methane emission through selective breeding.

Chapter 1.12 Results of animal breeding

Selective animal breeding already has almost 300 years of history. A lot has been achieved since. Obvious results have been achieved in the field of dog breeding. Selective breeding has produced very tall dogs like the Irish Wolfhound (>71 cm), very heavy dogs, like the Boerboel (50-80 kg), very small dogs like the Chihuahua (20 cm), very fast dogs like the Greyhound (17.5 m/sec), and very many more breeds with different looks and purposes. The size of the genetic improvement from generation to generation depends on the technique used to select animals for breeding. Introduction of new selection techniques made it possible to select the best animals for breeding more accurately and efficiently. Especially the introduction of reproduction techniques like artificial insemination (AI), that made it possible to have (very) large numbers of offspring per father, made it possible to select only the very best males for breeding, without decreasing the population in size. Techniques with similar effect on number of offspring per parent are not available for the female reproduction, but also in females techniques like embryo transplantation (ET) or ovum pickup have made it possible to produce much larger number of offspring of excellent females than with normal reproduction techniques in species where normally only one or few offspring per year were feasible.

Results obtained in cattle breeding

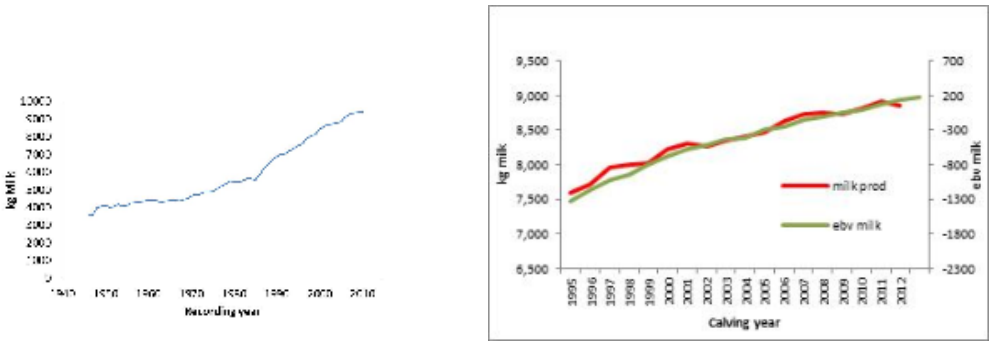


Figure 3. On the left is the phenotypic trend in milk production of Dutch black-and-white dairy cattle in the period 1945 – 2010. On the right is the phenotypic trend (red) compared to the estimated genetic trend (green) for milk production in Dutch black-and-white dairy cattle in the period 1995 – 2013. Ebv = estimated breeding value (source: CRV, the Netherlands).

In Figure 3 on the left you see the increase in milk production in the Netherlands between 1945 and 2000. The increase until 1970 is much less steep than that from 1990 onwards. Reasons for this are many, but important ones are very strong increase in use of AI so that stronger selection in bulls was possible, introduction of more accurate techniques for estimating breeding values, introduction of automatic milking and the free stall instead of the tied stall, and

better quality nutrition. In the graph on the right is the phenotypic trend compared to the genetic trend in the period 1995 – 2013. You see that the increase in phenotypic (= realised) milk production in that period is very similar to the estimated increase in genetic potential for milk production: in both cases approximately 1500 kg. This indicates that systematic improvements in the environment such as automatic milking, loose housing, and diet quality has similar effects on all cows.

Results obtained in poultry breeding

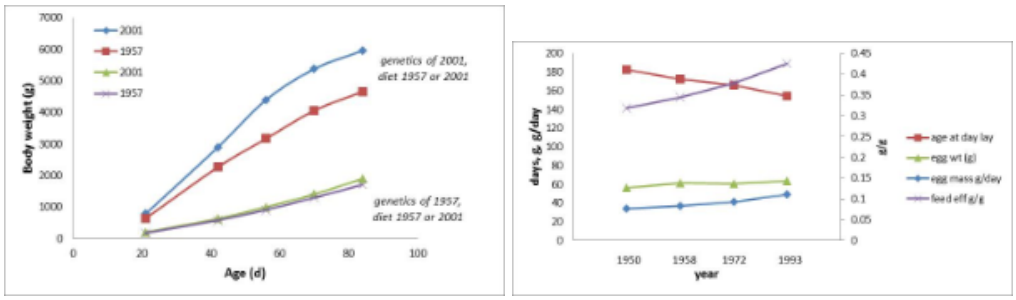


Figure 4. On the left is the effect of improved genetics and nutrition on weight at fixed ages in 2001 compared to 1957 in broilers (after Havenstein at al., 2003). On the right is the effect of selective breeding in laying hens on age at first egg, egg weight and mass, and feed efficiency in the period 1950 to 1993 (after Jones et al, 2001).

In Figure 4 you see examples of what has been achieved in broiler and laying hen breeding since the 1950's. In the figure on the left you see that even though the effect of improved nutrition is present in broilers, selective breeding is the most important reason for the strong increase in body weight at size weeks. It is incredible to realise that selective breeding has increased body weight at 84 days from 1907 g in 1957 to 5958 g in 2001, both on the same diet. The weight has more than tripled in size! In laying hens the effect of selective breeding is not as big, but also here in 43 years of selective breeding the hens start laying 28 days (=15%) earlier, lay 7 g (=12.5%) heavier eggs, lay more eggs, and use about 10% less feed to do so! And that was in 1993, selection has continued since. Body weight of laying hens has remained approximately the same.

Results obtained in horse breeding

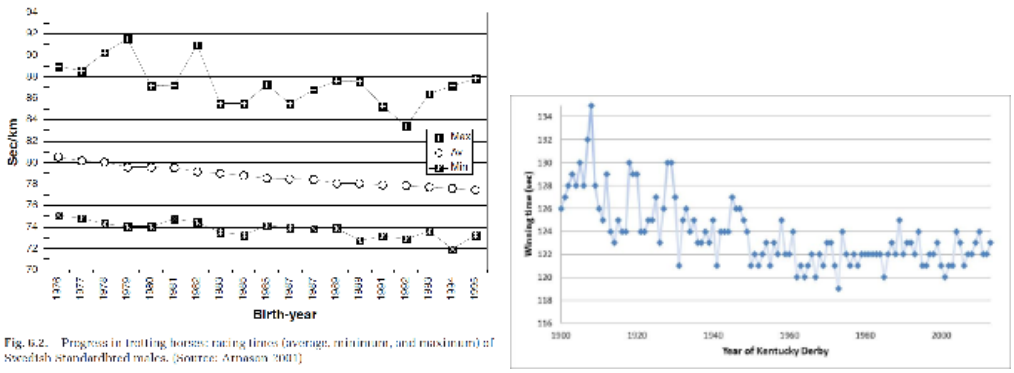


Figure 5. On the left are phenotypic trends in racing speed in Swedish Standardbred male trotters in the period 1976 to 1996 (source: Arnasson, 2001), and on the right are phenotypic trends in winning racing times of Thoroughbred horses during the Kentucky Derby in the period 1900 – 2013 (source: http://www.horsehats.com/KentuckyDerbyWinners.html).

In trotters selective breeding has resulted in a linear decrease in kilometre times (running speed) of about 1 second in 20 years (see Figure 5). There is no evidence that this rate of improvement is going to slow down. In racehorses the story starts also very successful. Selective breeding has caused horses to run faster. However, for the racehorses the success story seems to stop in the early 1950's. The record of the Kentucky Derby, of which the winning times are plotted in figure 5, dates from 1973! Even though selective breeding has continued, and with more advanced techniques, animals have not become any more faster. What happened? That is still a bit unclear, because there still is evidence for genetic variation, so some animals are genetically superior to others, and there is selection for racing qualities. The person who knows how to increase the speed of the racehorses *again is going to be very rich.*

Results obtained in pig breeding In pigs again a similar story can be told. In Figure 6 you see results of 10 years of selective breeding on growth, loin eye (the expensive part of the meat), leanness (backfat thickness), and reproductive performance (number born alive). Also here there is a clear increase in traits generating income (loin muscle and live piglets), and a decrease in money costing factors such as fat and days until slaughter.

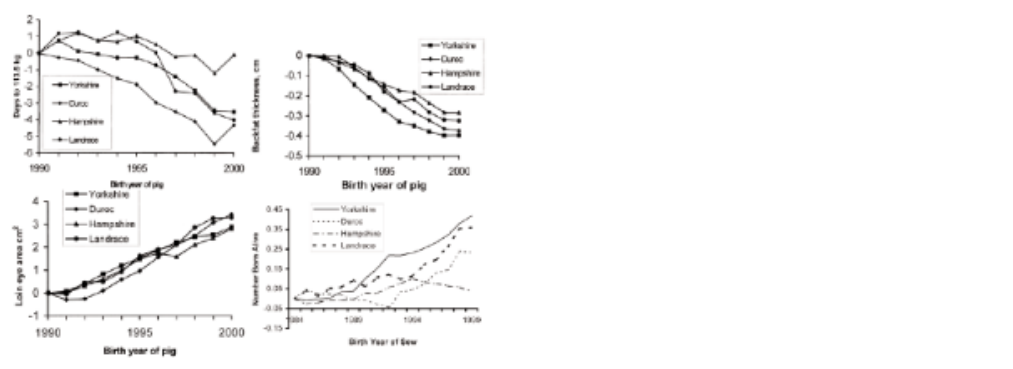


Figure 6. Phenotypic trends for Days to 113.5 kg, backfat thickness, loin eye area, and number born alive in four breeds of pigs in registered herds in the USA in the period 1990 – 2000 (source: Chen et al, 2002, 2003).

Chapter 1.13 Negative effects of animal breeding

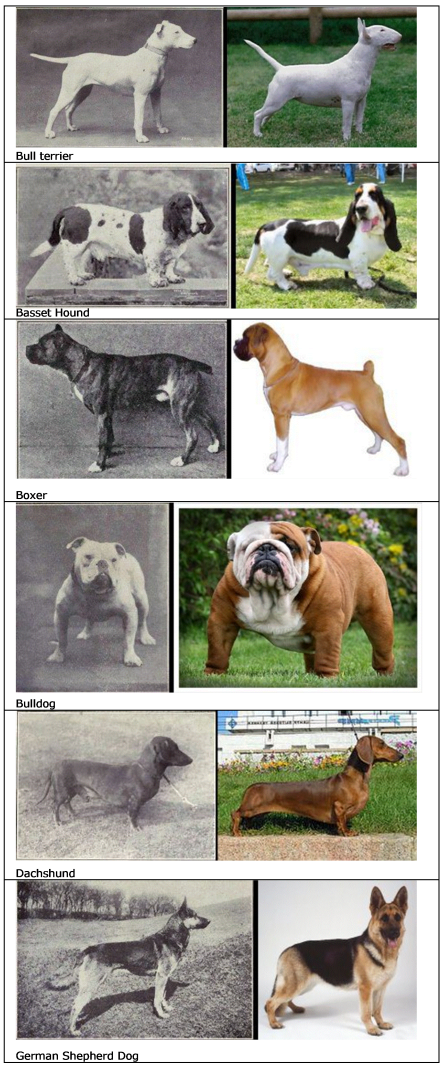
It is not all positive that has come out of animal breeding practices. There are examples where selective breeding has been taken too far. There are also examples where selective breeding has not only improved certain performances, but simultaneously and unintendedly also deteriorated other performances that were not under selection: the so-called negative correlated responses. Both types of negative responses to selective breeding are difficult to predict and are usually only noticed afterwards. This is the case because it takes a while to realise that the negative effects are structural and not coincidence, and that they occur at increasing frequency throughout the population. Even then it sometimes takes stepping back to realise the negative consequences. Changes are going slowly, so you get used to them.

Negative effects in dog breeding

Some clear examples of selection that has gone too far can be found in dog breeding. This is partly because selective breeding in dogs has a long history, but mainly because some dog breeds are selected mainly on looks. And most extreme looks tend to be considered the best, so selection in those breeds has been, and still is, on extreme looks (for some examples see Figure 7). The fact that shape of the skull in some breeds make it difficult for them to eat normal food because of the upper jaw being much shorter than the lower jaw, such as in case of the Boxer or the Bulldog, or breathing, such as in case of all breeds with a short upper jaw resulting in a flat face, or give birth or even mate without medical intervention (e.g. Bulldog), or where there is a risk that the eyes pop out of the socket because the skull is too small for their eyes (e. g. Pekinese, Chihuahua), are clear examples of selection being taken too far. And most of these examples are only related to the skull. Other breed characteristics that are not increasing the dog's wellbeing are, for example, too long ears so that infections are common (e.g. Basset Hound), or long back and neck so that intervertebral disc disease has become common (e.g. Dachshund), or too much skin so that inflammation in between the folds becomes common (e. g. Bulldog), or sloping back so that hip problems are common (e.g. German Shepherd Dog). All examples relate to selective breeding and taking breeds more and more too an extreme, because that is what you win the show with. Looking back only we realise that we have gone too far. And that realisation comes only very slowly because people get used to animals with certain features. They don't consider them abnormal for a very long time. Important is to realise that these effects can be reversed by selecting in the opposite direction.

Figure 7: Examples of representatives of dog breeds in "Dogs of all nations" (Mason, 1915) and in 2012. Respectively the Bull Carrier, Basset Hound, Boxer, Bulldog, Dachshund and German Sheperd Dog.

https://dogbehaviorscience.wordpress.com/2012/09/29/100-years-of-breed-improvement/



Negative effects of breeding in farm animals

It is not just dogs where we have taken selective breeding a step too far. Selection for large offspring has resulted in a high fraction of difficult births, sometimes requiring caesarean sections in the Texel sheep, and even almost as a standard way of delivering in the beef cattle breeds Belgian White-and-Blue cattle and the Dutch Improved Red-and-White. In the Texel sheep selection against difficult births has resulted in a decreasing fraction of birth requiring assistance. In this case the process could be reversed. But in the Belgian White-and- Blue and the Dutch Improved Red-and-White cattle the situation is more problematic and the repair process will take many generations.

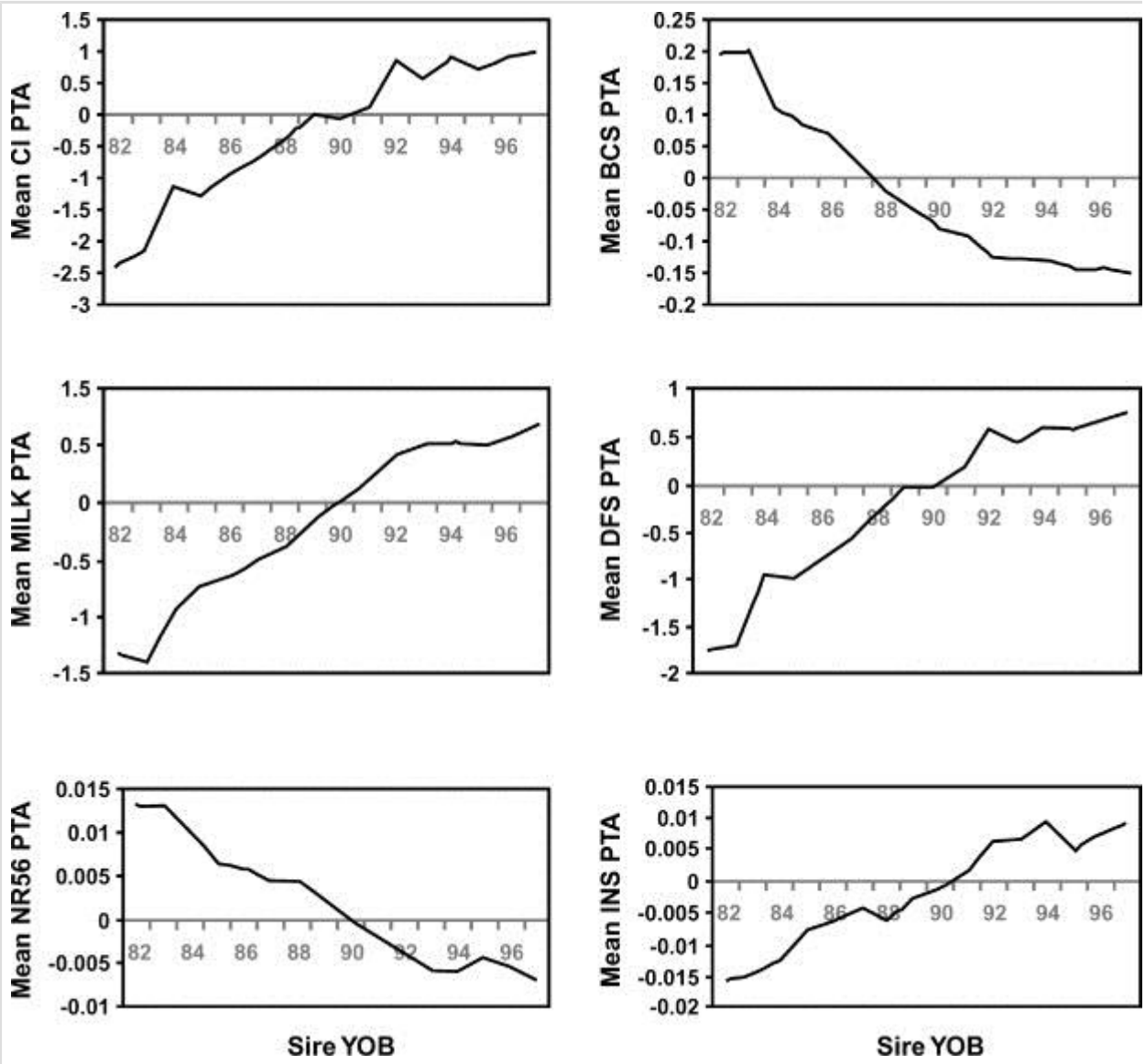


Figure 8. Genetic trends in a number of fertility related traits in dairy cattle in the UK, expressed as the predicted transmitting abilities (PTA) per sire year of birth (YOB), for calving interval (CI), body condition score (BCS), milk, days to first service (DFS), non-return at 56 days (NR56), and number of inseminations (INS). A PTA equals half of the estimated breeding value: the part that is transmitted to the offspring. Source: Wall et al., 2003

Birth problems are not the only unintended negative consequence of selective breeding in farm animals. Remember that the intention was to produce a lot of food, but cheaply so that it would be available for everyone. This has resulted in intensive farming systems, such as in pigs and poultry farming, where the animal products should be produced with as little costs as possible. So fast growth or more eggs with less feed. This has gone very well for many years and breeders really thought that there would be no limits to the genetic improvements as production

increased linearly. Unfortunately, in the 1980's it became more clear that there were also some negative consequences of strong selection for performance traits. For example, broilers started to show metabolic health problems due to the fast growth, laying hens started to have increased bone fractures because they couldn't manage sufficient calcium intake to deposit into the increasing number of eggs, dairy cows and sows started to show reduced fertility during the high production period. This is illustrated in figure 6, where the trends in calving interval, body condition score, milk production, days to first insemination, non-return rate, and mean number of inseminations needed per pregnancy are represented as predicted transmitting abilities (PTA). These PTA are especially used in the UK to indicate what part of the breeding value is transmitted to the offspring. Since those problems became apparent selection pressure has shifted from mainly performance to much more attention to animal health and reproductive performance. This shift has been the trend in all farm animal species. In the figure with the example from dairy cattle this shift in attention has started in the early 1990's as can be seen in the flattening of the slopes.

Chapter 1.14 Key issues in animal breeding

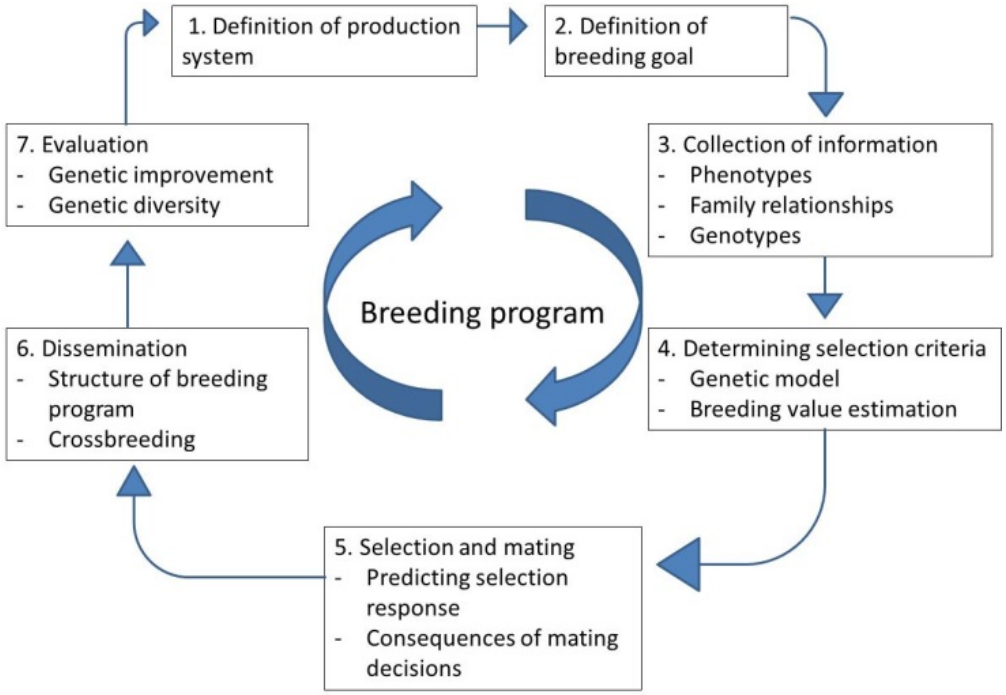
1. In animal breeding people select the animals that will produce the next generation of animals that are on average superior than the present generation of animals.
2. Selection by nature, natural selection, is very important for the adaptation of animals to the circumstances they are kept.
3. An important prerequisite for the success of animal breeding is that traits are heritable, which implies that the ability for the trait is passed on from parents to progeny.
4. Domestication of animals started with the dog. Later on the farm animals were domesticated. Domestication of a species requires specific traits. Within the domesticated species the process is still going on due to new requirements caused by changing circumstances.
5. Selective animal breeding started 250 years ago with the formation of breeds and herd books. The scientific base for animal breeding is developed in the 20-th century. The application of scientific developments in reproduction in the past 50 years made animal breeding more effective. Recently it got a strong impulse from developments in molecular genetics.
6. Breeding activities are directly be influenced by and are related to developments in society: required food production by animals and requirements of mankind for animals as companion and for leisure purposes.
7. Breeding programs with cattle, pigs and poultry yielded sharp increases in the quantity of milk, meat or eggs. Also capabilities of horses improved sharply.

8. Breeding of animals did not only give positive results: in dogs inbreeding and one-sided selection for conformation traits resulted in dogs with more health problems and less welfare. In farm animals selection for yield implies a risk for a deterioration of quality and fitness traits.

Chapter 2: Basics of animal breeding

Animal breeding program

Animal breeding is based on the fact that traits of parents are reflected more or less in their offspring. This is caused by the fact that traits are more or less heritable and that 50 % of the DNA, comprising the heritable ability for the traits of an animal, is passed from a parent to its offspring. In animal breeding potential parents are selected for certain traits and the best ones are indeed used as parents. In this way the next generation will be genetically improved for the desired traits. For the long term the subsequent breeding activities are carried out in a breeding program illustrated in the scheme below:



Chapter 2.1 Set up of a breeding program

Production system

First, the set up starts with a description of the production system (1). In more general terms the analysis of the way we keep the animals and for which purpose. What is relevant in this respect? For a small dog only kept as a companion animal in a comfortable house behaviour and health are important. For heath sheep kept in rough conditions the whole year around fitness traits and grazing behaviour are relevant. For growing broilers in intensive systems with high production costs, daily growth is conclusive.

Breeding goal

Second, the question which traits should be improved in the next generations. What will be the goal(s) for breeding (2)? This question is highly related to the reasons for which we keep the animals. This process deserves a thorough study and a long lasting conclusion as animal breeding is only effective when a breeding goal is consequently maintained for many generations. Examples of breeding goals are defined improvements in traits of production, product quality, health and welfare traits, conformation traits, sport performance, fertility etc.

Collection of information

Third, knowing the breeding goal, relevant information should be collected (3). Relevant in this respect are traits of animals (called phenotypes) that can help to establish the value of an animal with respect to the breeding goal. When jumping performance is a breeding goal trait in horses, data on jumping are collected. When fertility of pigs is in the breeding goal, litter traits are recorded. Other relevant information is the pedigree of the animals. Animal breeding is all about passing genetic abilities from one generation to the next. When you want to trace or influence this process of passing heritable traits, a registration of the parent-offspring relationships, the pedigree of an animal, is crucial. And nowadays, DNA-analysis is possible and practised in animals and can also be used to trace or influence the passing process of genetic abilities for traits.

Breeding value estimation and selection criteria

Fourth, knowing the breeding goal and after recording relevant traits of potential parents, the choice has to be made which animals will indeed be selected as parents (4) and which animals are excluded for reproduction. Based on a genetic model, a statistical model including pedigree information, a breeding value for a trait is estimated. Nowadays, when DNA information of animals is available, it can also be used to estimate breeding values. The estimated breeding value indicates the value of the animal with respect to the breeding goal: the lowest ones will have a negative effect on the breeding goal traits and the highest ones will improve breeding goal traits.

Selection and mating

Fifth, given the estimated breeding values of sires and dams, the actual selection of parents has to take place (5). The parents with a higher than average estimated breeding value will improve the breeding goal traits in the next generation. When for example a group of dairy sires with the highest breeding value for milk yield is selected as sires for the next generations, their

daughters will produce more milk than the present generation of dairy cows. Proper selection of parents will give a positive selection response in the next generations. Selection creates progress in breeding goal traits. After the selection of the parents another choice has to be made: which sire should be mated to which dam? The choice can be made e.g. on base of the pedigree information available or on the traits of the sires and the dams.

Dissemination of genetic gain

Sixth, in many breeding schemes the number of animals from which traits are recorded is rather small in relation to the population of animals used for human purposes. The dissemination of the selection response depends of the structure of the breeding programs. In commercial pigs and poultry programs selection takes place in the top of the breeding program and via a few “multiplying generations” the selection response obtained in the top is disseminated to the animals producing meat or eggs are bred. In cattle breeding artificial reproduction techniques, in particular artificial insemination techniques, give the opportunity to produce high numbers of offspring, disseminating the genes of the superior animals widely. Selection of a small number of animals may have a large impact on the traits of a population. Within the commercial breeding schemes e.g. for poultry and pigs specialized lines are crossed. These lines each are selected for specific traits and crossed in the multiplying phases to obtain crossbred progeny with the breeding goal traits by combining the traits of each of the lines.

Evaluation of results

Seventh, the breeding program should be evaluated regularly (7). The first question is: did we reach what we wanted? Is the new generation of animals better with respect to the breeding goal traits? Do we observe unwanted effects of selection? E.g. we realized a better growth of our meat producing animals, but they have more problems with their legs than their parents. The second question is: what has happened with the relatedness among the animals of the new generation. Are they more related to each other than their parents, due to the fact that we have selected only a few heavily related animals as parents for this generation? Did we decrease the genetic diversity of the population?

Then, the breeding circle starts again with a critical review of changes in the production system. Questions to be answered are: do market requirements change, e.g. for pork of a different quality? Do production circumstances change, e.g. are milk production quota for dairy farms expected to be abolished in the next future?

Chapter 2.2 DNA as carrier

DNA as carrier from genetic information passed to next generation

As a breeder you want to transfer the best genetic material you can get from the present to the next generation of animals. This genetic material is stored in the chromosomes in the nucleus of all cells of the animals that may be selected as parents for the next generation. This transfer process takes place with the transfer of chromosomes containing the genes via the creation of sperm cells and oocytes in the gonads. The combination of a sperm cell and an oocyte into a zygote is the starting point of a new animal with a unique genetic composition. In the transfer of chromosomes from parents to offspring the meiosis plays an important role. It causes that the transfer is determined to some extend by laws, recognized by Mendel (see chapter 1) and to some extend by random processes. The laws connect genetically relatives, e.g. parents and their offspring: the parents each share for 50% the same chromosomes, the same DNA and thus the same genetic value with their offspring. Therefore phenotypic traits of parents, can be found in the phenotypic traits of their offspring, based on the genes they got from their parents. In conclusion: offspring and parents and more in general related animals share a part of their DNA; they have a genetic relationship.

Chromosomes as units of DNA

Body cells of mammals and birds have a nucleus where pairs of chromosomes (units of DNA) are found. Each species has a specific number of chromosomes as is illustrated in the table below:

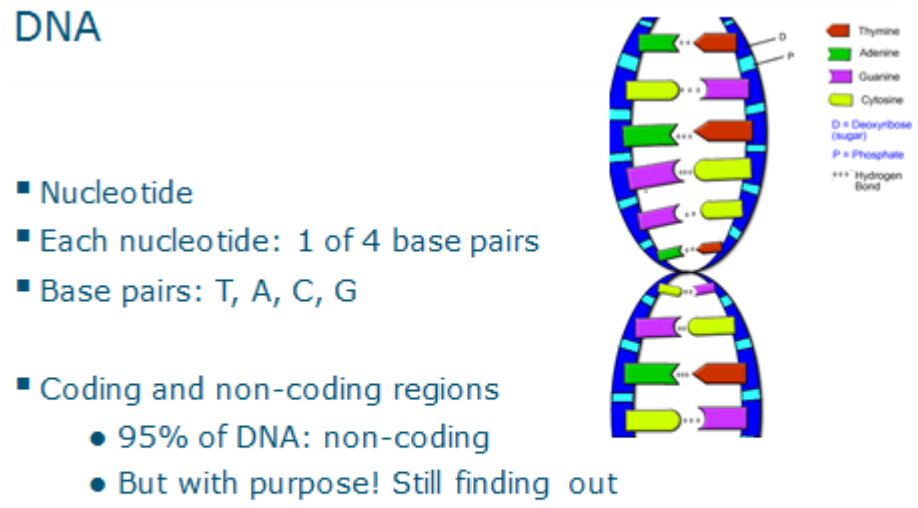
Number of chromosome pairs for different species

Species	Number of chromosome pairs
Man	23
Cattle	30
Horse	32
Pig	19
Sheep	27
Goat	30
Rabbit	22
Chicken	39
Duck	40

Chapter 2.3 Structure and composition of chromosomes

The different number of chromosomes impedes the production of crosses between species. In the formation of a zygote out of a sperm and an oocyte their single chromosomes are regrouped into pairs. With a different number of chromosomes this process fails when sperm cells and oocytes originate from different species.

Chromosomes are double-stranded helices, consisting of two long biopolymers made of simpler units called nucleotides—each nucleotide is composed of a nucleobase (guanine, adenine, thymine, and cytosine), recorded using the letters G, A, T, and C, as well as a backbone made of alternating sugars (deoxyribose) and phosphate groups (related to phosphoric acid), with the nucleobases (G, A, T, C) attached to the sugars (Wikipedia). Within a species and within a breed the nucleotides and nucleobases are situated in a fixed order.



Chapter 2.4 The transfer from chromosomes and genes from parent to offspring

To understand the genetic relationships among related animals it is necessary to know what happens in the creation of sperm cells, oocytes and zygotes as the start of a new animals. In mammals and birds all body cells of an individual are diploid: all chromosomes in the nucleus of a cell are present in twofold grouped in pairs.

Definitions

*A **chromosome** is a discrete block of DNA and is one of the basic structures of the genome. All nuclear DNA is organised into chromosomes with the number varying between animal species. Genes on a chromosome are linked and tend to be inherited together*

***DNA** is Deoxyribonucleic Acid, which is a macromolecule in the form of a double-stranded helix that carries the genetic information in all cells in higher organisms*

*A **gene** is the hereditary unit, a region of DNA on a chromosome containing genetic information that is transcribed into RNA that is translated into a polypeptide chain with a physiological function. A gene can mutate to various forms called alleles*

*An **allele** is a version of the sequence of DNA nucleotides at a locus. Not all individuals carry exactly the same sequence of DNA nucleotides at a locus.This allelic variation is the source of genetic variation*

*A **locus** is a position on a chromosome, for example of a gene. The plural is loci*

To illustrate these definitions: The MC1R-locus (Melanocortin 1 receptor gene) has been identified on chromosome 5 in dogs. From this gene, 2 alleles E and e are known. The original, so-called wild type allele E (non-mutated allele) causes the black color in dogs, the allele e created by mutation causes the loss of gene function and is responsible for bright red or yellow coat color in genotype e/e.

Chapter 2.5 The expression of genes, of their alleles, in the phenotype

In all body cells the chromosomes are present in duplo: one originate from the sire and one from the dam. Therefore all genes are present in duplo. These genes might be identical: the allele originating from the sire is identical to the one from the dam. Then an animal is homozygous for that gene. It implies that progeny of this animal always get that allele from this

parent. The alleles originating from sire and dam might be different. Then an animal is heterozygous for that gene. This implies that progeny of this animal might get one of these two different alleles of this parent.

Definition

Homozygote *is an individual carrying two copies of the same allele at a locus, e.g. ee or EE.*

Heterozygote *is an individual carrying two distinct alleles at a locus, e.g. Ee.*

For a certain gene we can distinguish three different genotypes: e.g. EE, Ee or ee. These combinations of alleles may cause different phenotypes. Suppose E is responsible for the production of the protein eumelanin in skin cells of dogs that gives a black pigmentation of the skin and e is responsible for the production of phaeomelanin in skin cellsof dogs that gives a red pigmentation. It is obvious that EE animals will be black and ee animals will be red. But what will be the color of Ee dogs? It reveals that they are also black!. This phenomenon is called dominance: the expression of the e-allele in heterozygous is not expressed in their phenotype. The allele E is dominant over e or from a point of view from allele e: it is recessive to allele E.

Definitions

A ***dominant allele*** *is an allele that has an effect on the phenotype not only when it is homozygous but also when it is heterozygous. When allele E is dominant over e, EE and Ee have the same phenotypic value.*

A ***recessive*** *allele is an allele that is only has an effect on the phenotype when it is homozygous. Therefore if allele e is recessive, ee yields a different phenotype from Ee and EE, which have the same phenotype.*

In case a gene is involved in the expression of a quantitative trait e.g. body weight of a mature goat the alleles might have a different expression leading to small differences in mature body weight. Take the following examples:

1. GG animals weigh 40 kg, Gg animals 38 kg and gg animals 36 kg. The body weight of the heterozygous animal is exactly the average of the two homozygous animals. The two alleles have an additive effect. We call it co-dominance.
2. GG animals weigh 40 kg, Gg animals 42 kg and gg animals 36 kg. The body weight of the heterozygous animal is higher than the average of the two homozygous animals and even higher than the value of the highest homozygous parent. In this case we call it overdominance.

Definitions

Co-dominance *is a situation in which a heterozygote shows the phenotypic effects of both alleles equally. See 'Additivity'.*

Additivity *is the assumption that each allele influencing a trait does so independently of the other allele present at that locus and all other alleles at all other loci, e.g. if alleles G and g are worth 1 and -1 respectively then additivity assumes GG is worth 2, Gg is worth 0, and gg is worth –2.*

Overdominance *occurs when the heterozygote has a genotypic value more extreme than either parent.*

In addition to the effects of different alleles of a gene at a single locus, alleles of different genes at different loci might influence each other and that might be expressed in the phenotype of the trait they influence. Again we have two possibilities: the effects of different alleles at different loci are additive: the effect is the sum of the effect of the individual alleles. When these effects are non-additive, this is called epistasis.

Definition

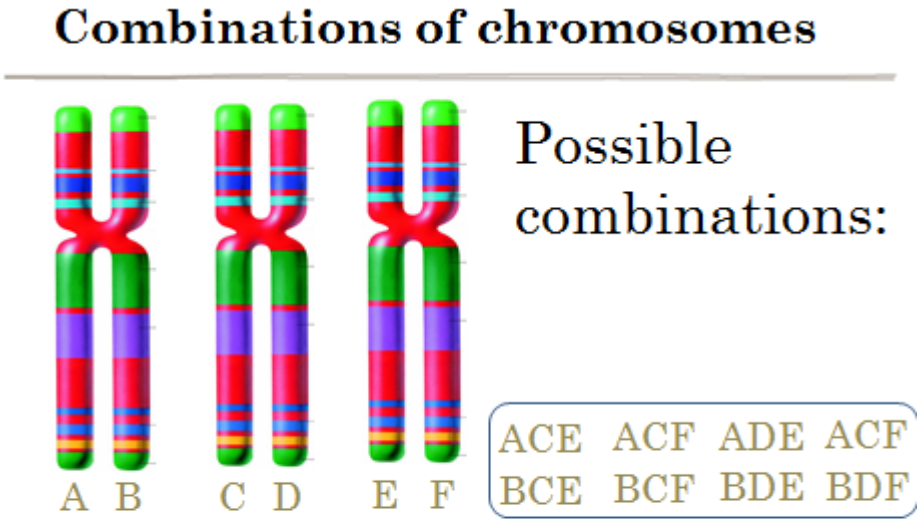
Epistasis *is when loci are non-additive. The genotypic value of a locus on a trait depends upon the genotypes at other loci or a situation in which the differential phenotypic expression of a genotype depends on the genotype at another locus.*

Example of epistatis (Minkema 1966)

Plumage pattern in chicken is amongst other genes, determined by interacting alleles of two different loci: the E-locus and the S-locus. The E-locus has two alleles E and e^{+}. The S-locus: S and s. E is dominant over e^{+} and causes a uniform black plumage. In the $e^{+}e^{+}$ animals the black colour is restricted to restricted parts of the plumage. In the non-black parts of the plumage the SS or Ss animals have at these places the silver colour, while the ss animals have there the golden colour. Thus the golden or silver colour is only expressed in the $e^{+}e^{+}$ animals: black at the E-locus is dominant over golden and silver determined by the alleles at the S-locus.

Chapter 2.6 Meiosis creates differences among offspring of an individual

In sperm cells and oocytes the chromosomes are no longer present in twofold grouped in pairs. In these cells all chromosomes are present in singularity. In the testis and in the ovary the pairs of chromosomes split in singular chromosomes and each chromosome of a pair goes by chance to a sperm cell or an oocyte. This process is called the meiosis and is illustrated below with three chromosome pairs of a male:



Chromosome pair 1 consists of chromosome A and B, pair 2 of C and D and pair 3 of E and F. In the meiosis the pairs the chromosome pairs split and go by chance to a sperm cell. In this way sperm cells are created with 8 (2^3) different combinations of chromosomes: ACE, ACF, BCE, BCF, ADE, ADF, BDE and BDF. When a species has n chromosomes the parents create 2^n different sperm cells or oocytes.

Definitions

Meiosis is the process carried out in the germ cells by which gametes are formed. In diploids this involves the creation of haploid cells (sperm, oocytes) from the diploid progenitor cells

Mendelian sampling *is the random sampling of parental genes caused by segregation and independent assortment of genes during germ cell formation, and by random selection of gametes in the formation of the embryo*

Chapter 2.7 Relatives share similar DNA, they have a relationship

Due to the events in the meiosis a sperm cell and an oocyte contains 50 % of the DNA of the parent (a natural law in the transmission from DNA of parents to offspring) and contains a unique combination of chromosomes of the parent (random process in the transmission of DNA between generations). After the fertilization of an oocyte with a sperm cell the nucleus of the zygote contains again chromosomes in twofold grouped in pairs. This implies that each animal receives half of its chromosomes, half of it genetic value from its sire and half of its dam. Thus the genetic relationship between an animal and each of its parents is 0.5. This is called the additive genetic relationship. But the fact that each sperm cell and each oocyte from a parent contains a unique combination of the chromosomes of the parent causes that the offspring of the same combination of a sire and a dam (full sibs) still show differences in traits. The additive genetic relationship of full brothers and full sisters (full sibs) is 0.5, because on average they share 50 % of the DNA of their parents.

Definition

The ***additive genetic relationship*** *between two animals is the amount of DNA they share due to the fact that they are related*

A few additive genetic relationships are presented below:

Relationship	Percentage of similar DNA they share
Parent-son or daughter	50
Grandparent-grandchild	25
Great-grandparents- great-grandchild	12.5
Full brother(s) – full sister(s)	50

Half-brother(s) – half-sister(s)	25

Thus relatives share similar DNA. The average percentage they share is clear, but without further knowledge of their DNA (genotypes) or their phenotypes, it is not clear which part of the DNA, which alleles they share.

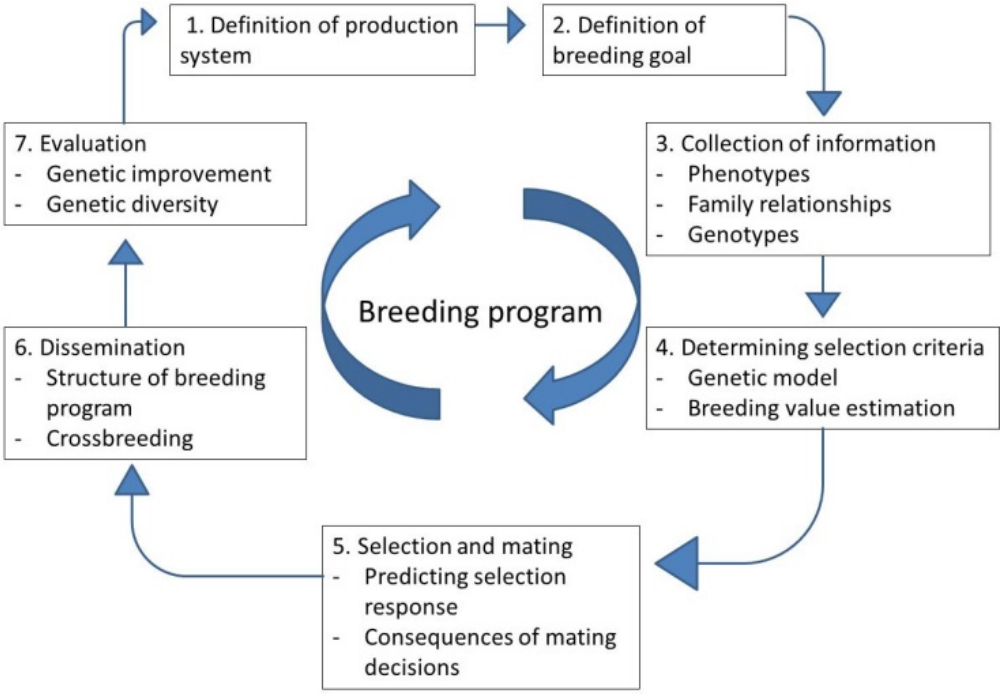
Chapter 2.8 Key issues in basics of animal breeding

1. A breeding program is characterized by a series of subsequent activities: defining the breeding goal, registration of phenotypes, genotypes and pedigrees, estimation of breeding values for selection traits with a genetic model, selection of parents for the next generation based on estimated breeding values, mating of parents and dissemination of genetic superiority towards the production animals and evaluation of the program with respect to the genetic diversity maintained and realized selection response.
2. Body cells of mammals and birds have a nucleus where pairs of chromosomes are found. The different number of chromosomes among species impedes the production of crosses between species. Chromosomes are strings of DNA, Deoxyribo Nucleic Acid, which is a macromolecule in the form of a double-stranded helix that carries the genetic information in all cells of higher organisms.
3. A gene is the hereditary unit, a region of DNA on a chromosome containing genetic information that is transcribed into RNA. This RNA is translated into a polypeptide chain with a physiological function. A gene can mutate to various forms, called alleles.
4. In all body cells the chromosomes are present in duplo: one originate from the sire and one from the dam. Therefore all genes are present in duplo. These genes might be identical: the allele originating from the sire is identical to the one from the dam. Then an animal is homozygous for that gene. The alleles originating from sire and dam might be different. Then an animal is heterozygous for that gene.
5. Alleles can be dominant or recessive (interaction between two alleles at a locus), effects can be additive: co-dominance (the value of the heterozygous lies exactly between the two homozygous genotype) or can be over-dominant (the heterozygote has a genotypic value more extreme than either parent).
6. Alleles of a gene can also interact with alleles of another gene: epistasis.
7. In sperm cells and oocytes the chromosomes are no longer present in twofold grouped in pairs. In the testis and in the ovary the pairs of chromosomes split in singular chromosomes and each chromosome of a pair goes by chance to a sperm cell or an oocyte. This process is called the meiosis and leads to Mendelian sampling: each sperm cell or oocytes contains a unique combination of alleles from its parent.

8. Due to the events in the meiosis a sperm cell and an oocyte contains 50 % of the DNA of the parent. After the fertilization of an oocyte with a sperm cell the nucleus of the zygote contains again chromosomes in twofold grouped in pairs. This implies that each animal receives half of its chromosomes, half of it genetic value from its sire and half of its dam. Thus the genetic relationship between an animal and each of its parents is 0.5. This is called the additive genetic relationship.

Chapter 3: Reasons to keep animals determine the breeding goal

In this chapter we will explain and discuss the setup of a breeding goal (block 2) in relation to the reasons for keeping the animals or the breed (block 1). These reasons can be quite different. This has a great influence on the setup, on the definition of the breeding goal.



The following topics will be presented in this chapter:

- The challenges for animal breeding
- Defining the breeding goal
- Measurement of breeding goal traits
- Weighing the different traits of the breeding goal
- Example: Pig breeding goals are influenced by changes in consumer expectations, in society opinions and in technology developments
- Example: Dutch milk production index in cattle breeding (INET)
- Example: Breeding methods of top breeders in dogs
- Example: Production objectives for village poultry in Ethiopia
- Example: Breeding goals of horse (KWPN)

Chapter 3.1 Challenges for animal breeding

The major challenge for food production in Agriculture is the on-going growth of the human population towards 9 billion in 2050 (United Nations Development Goals, 2005). Livestock systems play an important role in agriculture by producing high quality food. In developing countries animals provide not only meat, milk and eggs, but also fibre, fertiliser for crops, manure for fuel and draught power. In developing countries productivity and fitness traits are the major challenge for animal breeding to facilitate production of food by animals and for food supply. In developed countries with intensive animal production systems, health and welfare traits create a new challenge for animal breeders.

Breeding challenges are influenced by a wide range of factors. They are determined by the needs and priorities of the owners of the animals, the consumers of animal products, the food industry, and increasingly the general public. Finding the right balance between the different demands is a continuous process, and requires anticipation of future conditions and careful planning to establish effective breeding programs.

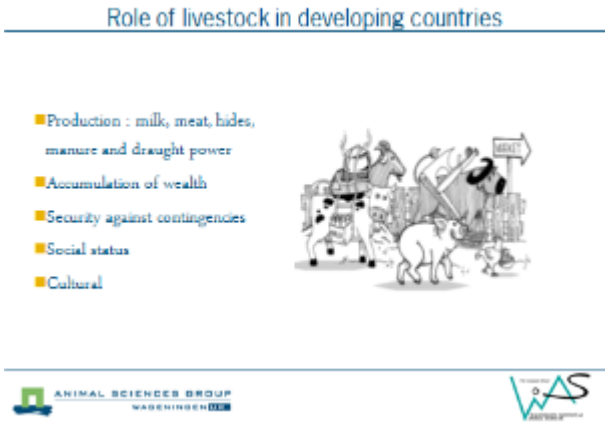
Challenges for animal breeding in small populations

In small populations breeding opportunities for food production are limited. In such populations nearly all animals have to be used as parents for the next generation (at least the females) to get enough offspring. Then, there is no opportunity to select for traits related to food production. In small populations, the main concern is to maintain the population by conserving the genetic diversity and to manage inbreeding. As will be explained later inbreeding causes a lower fitness and increases the incidence of recessive genetic defects. This also implies that in small populations nearly all males and females have to produce offspring. Selection for breeding goal traits is hardly possible in these small populations.

Chapter 3.2 Breeding goals depend on the production system

In addition to food production, animals fulfil a variety of functions for mankind: providing labour, being a companion animal, participation in leisure activities, in cultural events and in nature management. The use of the animals and the wishes of the users determine to a great extend the breeding goal and the breeding programs. In these breeding programs for animal functions other than food production, several traits play an important role in addition to selective breeding for the obvious breeding goal traits: health and welfare of the animals, and adaptation to feed of lower quality and extreme climates to be capable to produce and reproduce. The first slide on the right gives an impression of the role of livestock in developing countries. The second of the socio-cultural value of the Indonesian Madura cattle.

Before defining the breeding goal many questions on the production system have to be answered. For what reason the animals are kept? In which way the products and animals are marketed? What are important aspects of feeding and management? Are the breeders organized? Is there already a breeding program in place? What traits can be recorded? Is artificial reproduction applicable? Hence, these aspects of production systems determine the possibilities for breeding programs and for the selection of breeding goal traits.



Chapter 3.3 The choice of the appropriate breed

The choice of the most appropriate breed to use in a given environment or production system should be the first step when initiating a breeding program and due attention should be given to the adaptive performance of a breed. Numerous examples exist of animals of high productive breeds (e.g. Holstein Friesian dairy cow) imported in tropical countries without any success. The animals are not adapted to the high temperature, they hardly reproduce and heat stress impedes high production levels. In addition many tropical diseases lead to a high mortality. In all food production systems adaptation of the animals to the conditions within the system is important. When ignored, the fitness of the animals will be reduced. Adaptive fitness is characterized by survival, health and reproduction related traits. In the warmer tropical areas,

pathogens and epidemic diseases are widespread, climatic conditions are stressful, and feed and water are scarce. There, locally adapted autochthonous breeds display a far greater level of resistance and adaptation due to their evolutionary roots as compared to imported breeds.

Chapter 3.4 The breeding goal

Setting up a breeding program starts with the definition of a breeding goal and is followed by the design of a scheme that is able to deliver genetic progress in line with this goal. A breeding goal with the relevant traits, collection of performance data, analysis of the data for the identification of superior animals, and the use of superior animals to produce the next generation, are the main components of structured breeding programs.

Definitions

*A **breeding goal** is the specification of the traits to be improved including the emphasis given to each trait. It gives the direction in which we want to improve the population.*

*A **Breeding program** is a program aiming at defined breeding objectives for the production of a next generation of animals. It is the combination of recording selected traits, the estimation of breeding values, the selection of potential parents and a mating programme for the selected parents including appropriate (artificial) reproduction methods. See scheme at the beginning of this chapter.*

A remark: traits with an optimum value

For most of the traits, the objective is a continuous improvement, but for some of the traits the goal is to reach intermediate values. Examples of such traits are egg weight where a market exists for table eggs between 55 and 70 grams. Mature body size is positively related to returns at slaughter but negatively with feed efficiency. The production aim is a high carcass value in combination with low feed cost. Thus in many meat production systems mature body weight of the animals has an optimum.

Chapter 3.5 Breeding goals are directed towards the future and require tenacity

In the ideal situation the breeding goal consists of a single criterion that facilitates the ranking of animals in line with this goal. The breeding goal aims at the future. In practical situations it is often not a single trait but a combination of traits, specified according to their relative importance. The breeding goal usually involves the improvement of multiple traits simultaneously. The breeding goal should be formulated carefully and should hold consequently for many generations to become successful as a breeder. Breeding is a process

where in each new generation small steps towards the breeding goal traits are realized. The success of breeding will become visible as the sum (the accumulation) of all these small steps. Changes in breeding goal traits from generation to generation do not contribute to a visible cumulative success (chapter 1). A trait should be included in the breeding goal based on its economic value and its heritability (see for an example: the Dutch milk production in cattle breeding at the end of this chapter). Breeding goals can be expressed in term of weighing factors for traits based on its economic value or based on the desired genetic improvements for each trait.

Breeding goal traits could be restricted to the wishes of the breeder, to the requests of the producers and processors or even extended to the behaviour of the consumers of products of animal origin or societal wishes. However, the more traits are included in the breeding goal, the less progress in each trait will be obtained per generation.

The outcome of a breeding program is often realized many years after selection decisions are made. This underlines the need to anticipate future demands when defining breeding goals and requires attention towards returns on investments. And most breeding goals are only reached after several generations of selection. This requires tenacity of the breeders involved: frequent changes in breeding goals impedes the generation of progress in breeding programs.

Chapter 3.6 Breeding goals consists of several traits

The breeding goal for food production aims, irrespectively of the species, at: improving gross efficiency (amount of product divided by the amount of feed consumed) en reducing cost price by: 1) improving productivity (higher yield and financial returns, 2) improving feed conversion (less feed per kg of product and less costs) and 3) improving reproduction, health and survival (less replacement animals needed and less costs). Increasing attention to improve welfare and to reduce environmental impact might be additional breeding goal traits.

Nowadays, in commercial breeding programs for dairy cattle, pigs and poultry sophisticated breeding programs with complex breeding goals are in place. For other species breeding programs are less complicated with a limited number of breeding goal traits. E.g. globally in small ruminants kept for meat production (sheep and goat) with less complex breeding programmes growth appears to have the greatest importance in the breeding objective. In commercial pig and poultry breeding special lines with different breeding objectives are developed that are crossed to obtain the final egg or meat producing animal with the optimal combination of objectives selected for in the different lines. Due to a limited number of breeding goal traits in a specialized line, a lot of progress can be made in each line. By crossing the lines, the breeding goals, for which a high level is reached in each line, are combined. His proved to be more profitable than selection for all important breeding goal traits in one line or one breed.

A simplified example of the use of special lines is a three-way cross often applied in in pig breeding: first, sows of a line selected for number of piglets is crossed with a boar of a line selected for growth. Second, crossbred sows are subsequently crossed with boars of a line selected for carcass quality. The result is a lot of piglets born with a good growth and carcass quality. Around 1970 in a sheep breeding trial, ewes from the Finnish Landrace breed (a breed with a high litter size) were mated to a ram of the Ile de France breed (a breed that can be bred irrespective of the season). This resulted in crossbred ewes with a high number of lambs due to three lambing's in two years' time. The sire of the lambs was a ram of the Texel breed famous for growth and slaughter quality.

Chapter 3.7 Measurement of breeding goal traits

Recording breeding goal traits raises a lot of questions:

- What can be measured?
- How often can or should the trait be measured?
- Who or what is measuring the trait?
- What animals can or should be involved?
- At what age?
- How detailed?
- What is the accuracy of the measurement?
- Do systematic effects play a role on the outcome of the measurement?

In this chapter we focus on the traits with respect to the breeding goal. It should be possible to collect them easily at low costs. And it should be possible to measure or to judge them precisely. Last but not least they should be heritable. In chapter 4 we focus on all aspects of measuring traits and on the animals that might be sources of information to reach the breeding goal by selection. Breeding goal traits are heritable traits that can be easily and accurately measured.

Chapter 3.8 The breeding goal determines which traits should be recorded.

Breeding goal traits might be **quantitative**. Milk, meat or egg production, body measurements or performance expressions are examples of quantitative traits. They are measured in units: in kg or simply in numbers: kg of milk, grams of growth and number of eggs.

Breeding goals might be **qualitative** e.g. the quality of a product or an important trait in the breed standard. Product traits, scores for body traits, disease incidences or performance impressions are examples of qualitative traits. They are measured in classes: e.g. a 1 (good), 2 (moderate) or 3 (bad) for meat quality or simply 0 (not present in the animal) or 1 (present).

Some breeding goals traits cannot be measured at the time when it is relevant. E.g. in meat production, meat quality is an important breeding goal trait. However, you cannot measure carcass composition of a young calf, piglet or lamb at the moment you consider to use the animal for breeding. It can be measured only after slaughter and then breeding with that animal is impossible. Indicator traits, obtained by scanning live animals for body composition before you take breeding decisions may help to predict carcass composition.

Breeding goal traits might be complicated consisting of many underlying traits. E.g. in nearly all food producing species reproduction capacity is part of the breeding goal. Reproduction capacity is composed out of male and female reproduction traits. In males sperm quality and insemination results are part of their reproductive capacity. In females age at puberty, interval between litters, number of offspring raised per year are examples of underlying traits. In jumping horses the conformation of the horses and the way they use their legs are very important breeding goal traits. In working dogs, trainability is an important trait in addition to health, behaviour and conformation. The latter three traits are important in all species used for companion purposes.

For some species a few relevant measurements are given in the table below:

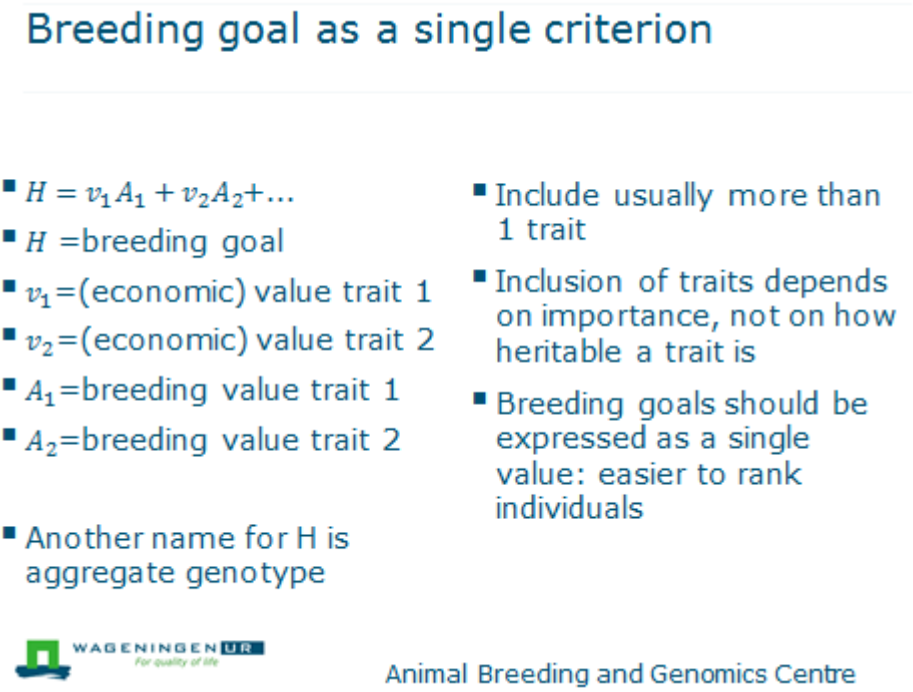
Species	Measurement	Unit	Recorder
Dairy cattle			
	Milk yield	Kg	Controller / farmer / robot
	Fat % in milk	%	Lab
	Pelvic height	Cm	Herdbook inspector
	Udder shape	Score	Herdbook inspector
	Mastitis	Incidence	Farmer/ Veterinarian
Jumping horses	Conformation	Score	Inspector / judge

	Behaviour	Score	Inspector / judge
	Jumping technique	Score	Inspector / judge
	Movement	Score	Inspector / judge
Dogs	Hip dysplasia	Score	Röntgen
	Conformation	Score	Inspector / judge
	Behaviour	Score	Tester
	Genetic defects	Incidence	Veterinarian

Chapter 3.9 Weighing the different traits of the breeding goal

Breeding goals can be simple or complex. In case of commercial breeding programs many traits are recorded and have an influence on the profit of animals produced by these commercial breeding programs. In extensive production conditions or in case of hobby breeding only a few important traits are recorded and simple breeding goals consisting of a few traits are used.

In order to rank the individual selection candidates for the breeding goal traits it is necessary to comprise the values for the respective traits into one single selection criterion. The value of this criterion can be obtained by summing up the breeding value for each trait multiplied by a weighing factor based on the relevance of that trait in the breeding goal. The relevance might be based on the relative economic value of the trait. This principle is outlined in the slide below:



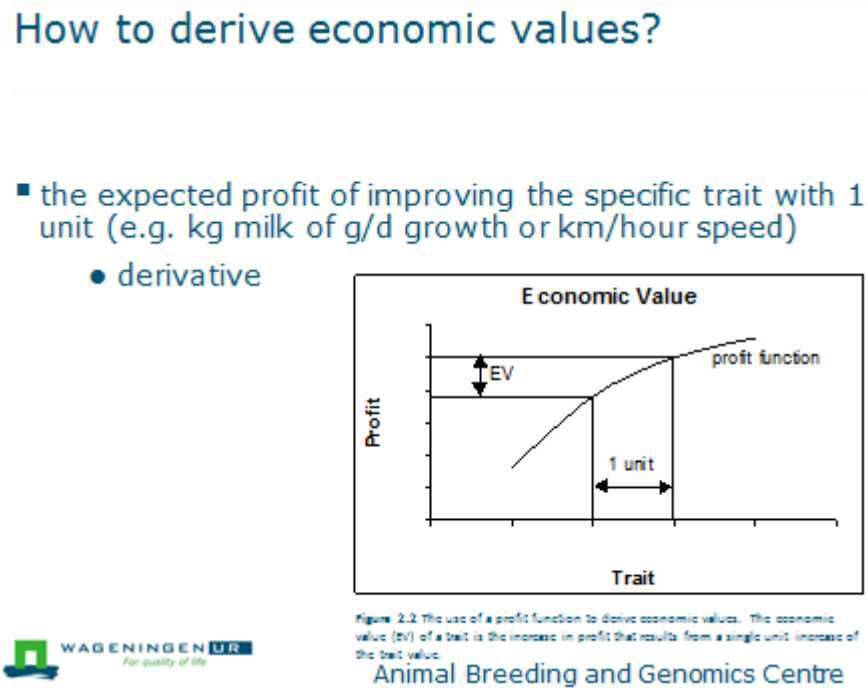
Definition

*The **Breeding value** is the mean genetic value of an individual as a parent for a trait. It is estimated as twice the average superiority of the individual's progeny relative to all other progeny under conditions of random mating*

The economic value of an animal is based on many traits having different effects: production, quality (composition) of the product, disease problems, fertility, and ease of handling and management.

Procedures exist to define breeding goals to weigh market and non-market values. In these procedures animals are seen as an integrated part of a production system (at farm level?). Weighing of traits has been mainly dependent on economic values and frequencies of expression of the genetic gain obtained. The methodology to weigh the traits with respect to resource efficiency and economy is well developed.

A illustrative example of the calculation of the net value of an increase by one unit in a breeding goal trait is given in the slide below:



The composition of traits affects the structure of a breeding program. The breeding goal traits determine from which animals traits should be measured: from parents of the selection candidates, the candidates themselves, their sibs or their progeny? When sib or progeny records are required, they should be bred in sufficient numbers to obtain accurate information for the breeding value of the candidate.

Progeny of selected animals produce at various moments. Therefore it is necessary to relate future income into the present costs of selection.

Chapter 3.10 Aspects of sustainability and economics in breeding goals

Sustainable production systems need long-term and equitable solutions emphasize resource efficiency, profitability, productivity, product quality, environmental soundness, biodiversity, social viability and ethical aspects. Therefore long-term biologically, ecologically, and sociologically sound breeding goals is emphasized.

Improving milk yield in dairy cattle has a favourable impact on greenhouse gas emissions as fewer cows and followers are required to meet the same production level at a farm. Improving the average lifespan of dairy cows with one year has a large favourable impact on greenhouse gas emissions. The expected annual response in milk yield per cow increased when selection index weights were increased from current economic weights to environmental weights.

Defining breeding goals at different levels

Breeding goal traits can be considered at different levels: 1) at individual animal level: what is the effect of the improvement of a trait on the profitability of an animal, 2) at (cross)breeding system level: what is the effect of selection in the grandparents on the profitability of the crossbred grandchild in producing the final product and 3) at farm level: what is the effect on income per farm and 4) at production chain level: what is the effect on production and processing level? On these different levels different effects might be generated. When a beef producer has a contract with a slaughter house to deliver annually a fixed amount of carcass, selection on daily gain will result in heavier carcasses and less carcasses sold by the producer per year. When subsequently he keeps less animals he might be confronted with a surplus of roughage at his farm that he cannot transfer into saleable carcasses. Then the profit from a higher daily gain is lower at farm level. When in a breeding program of dairy cattle selection for milk protein variants is practised, it may lead to a higher cheese yield of milk. If a milk producer is not paid for the protein variants in the milk, the full profit of the selection for milk variants goes to the cheese factory.

Chapter 3.11 Key issues in reasons to keep animals determine the breeding goal

1. Breeding challenges have a large impact on the breeding goal. They are determined by the needs and priorities of the owners of the animals, the consumers of animal products, the food industry, and increasingly the general public. Finding the right balance between the different demands is a continuous process, and requires anticipation at future conditions and careful planning to establish effective breeding programs.
2. A breeding goal is the specification of the traits to be improved including the emphasis given to each trait. It gives the direction in which we want to improve the population. Most breeding goals are only reached after several generations of selection. This requires tenacity of the breeders involved.
3. In small populations nearly all males and females have to produce offspring. Selection for breeding goal traits is hardly possible in these small populations.
4. In breeding programs for animals several traits play an important role in addition to selective breeding for the obvious breeding goal traits: health and welfare of the animals, and adaptation to feed of lower quality and extreme climates to be capable to produce and reproduce.
5. The choice of the most appropriate breed to use in a given environment or production system should be the first step when initiating a breeding program and due attention should be given to the adaptive performance of a breed. When ignored, the fitness of the animals will be reduced. Adaptive fitness is characterized by survival, health and reproduction related traits.

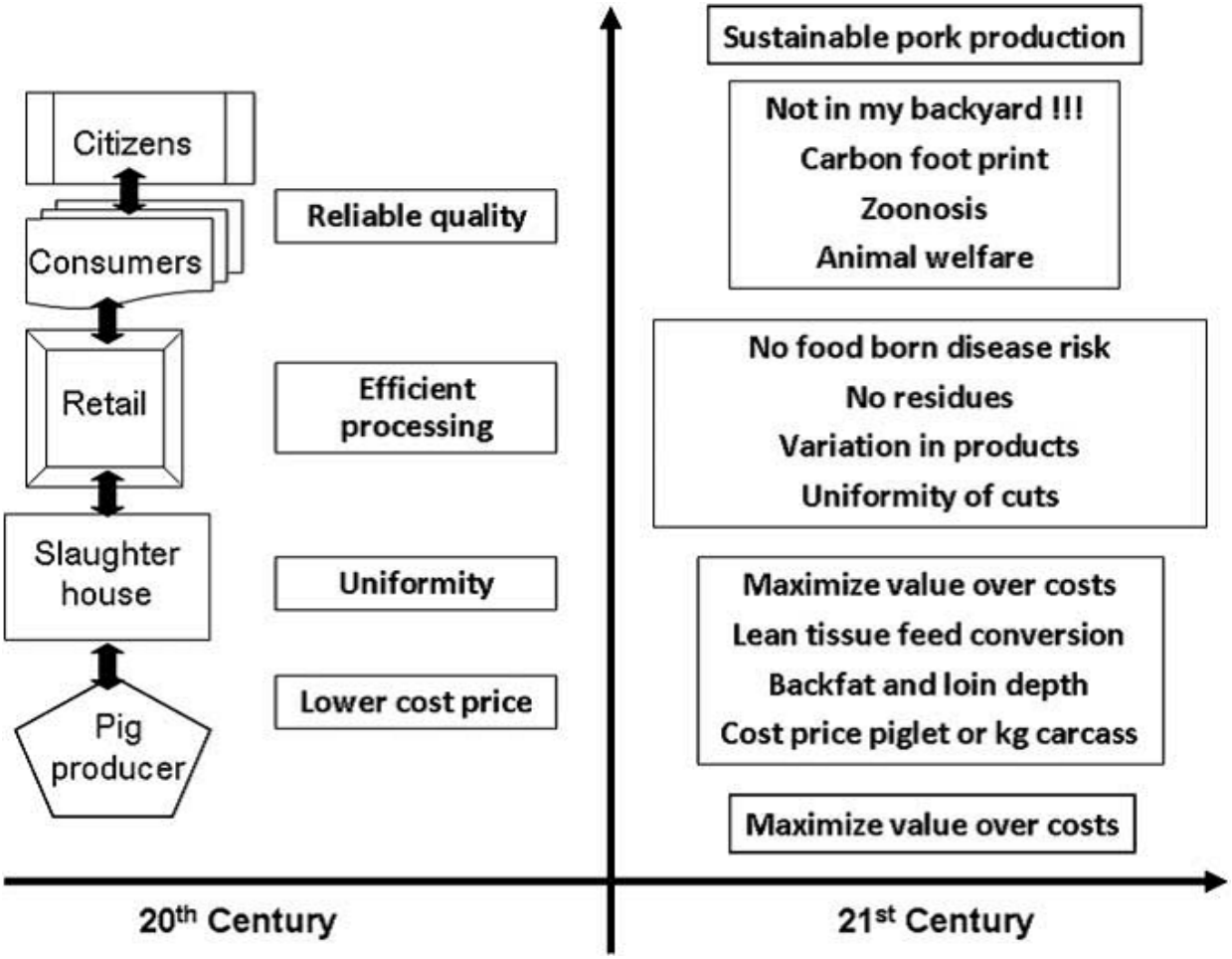
6. Nowadays, in commercial breeding programs for dairy cattle, pigs and poultry sophisticated breeding programs with complex breeding goals are in place. For other species breeding programs are less complicated with a limited number of breeding goal traits. Breeding goal traits are heritable traits that can be easily and accurately measured.
7. To rank the individual selection candidates for the breeding goal traits it is necessary to comprise the values for the respective traits into one single selection criterion. The value of this criterion can be obtained by summing up the breeding value for each trait multiplied by a weighing factor based on the relevance of that trait in the breeding goal. The relevance might be based on the relative economic value of the trait.

Chapter 3.12 Example: Pig breeding goals

Influenced by changes in consumer expectations, in society opinions and in technology developments

Merks, J.W.M. et al, 2012. New phenotypes for new breeding goals in pigs. Animal, 6:4, pp 535-543.

Genetic trends in pigs can be fast but it takes about 3 to 5 years for the changes to actually take place in the production herds and for the consumers to experience the benefits. Present pig breeding programs have breeding goals including traits of interest, the relative significance of these traits and the direction for improvement. The traits of interest depend of the expected market requirements of pork set by consumer's expectations and willingness of the society too accept production methods. Therefore, a good prediction of future trends is necessary. Pork chain development combined with information technology facilitate to sample breeding goal associated phenotypes (traits) in a cost effective manner. As you can see in figure XXX the current interests of farmers, citizens, governments representing the society and food industry requires new "phenotypes": vitality, uniformity, robustness, welfare and health and traits to reduce the carbon food print of pork production, while keeping the production efficiency and product quality.



Improving vitality will result in better survival of piglets during the prenatal period, less piglets born dead, no dead or sick pigs in suckling, nursery and finishing, less sows culled after first parity and lower sickness and mortality in older sows. Improving of uniformity at different levels in the production chain has a positive effect on the management of the animals and on the processing. Uniformity in birth weight of a litter decreases mortality especially of smaller piglets. Uniformity in protein deposition leads to more uniformity in growth and the age at the required slaughter weight and gives a more efficient utilisation of dietary protein. Uniformity in slaughter weight and carcass length increases slaughter plant efficiency. Uniformity in pork chops, meat colour, marbling and drip loss is useful for retail shelves and consumers. Increasing robustness of the animals will improve the ability of pigs to adapt to different stressors they may be face in their life: disease challenges, extremely hot or cold temperatures, low quality feed or changes in housing and management, e.g. by transition from individual to group housing. Reduction of the carbon footprint of pork production can be realised by improvements in digestive efficiency and reductions in maintenance requirements. In pork production males were castrated. Intact males sometimes produce a very bad smell that is smelled after the pork chop is fried. Castration is considered in many countries as a painful interference in the young piglets that should be avoided. But recently selection based on genetic markers opened up selection against boar taint making castration of boars no longer necessary. This is a wish in society

existing already for a long period of time. Genomic selection will be of great help to select for the new “phenotypes” and the new more complex breeding goals. Automation processes for identification of animals and carcasses and recording of their traits in different parts of the production chain will result in very informative databases that provide DNA profiles and phenotypes to be used in selection.

Chapter 3.13 Chapter Example: Breeding goals for horses

Breeding goals of KWPN (source: Selection for performance; KWPN, http://www.kwpn.org/)

Definition of the breeding goal

The goal of the KWPN is the breeding of horses that can perform at Grand Prix level of dressage or jumping. In order to reach this goal, a horse must have a good constitution, a functional and preferably attractive conformation/appearance, correct movement and an agreeable character.

A **good constitution** speaks for itself: training a horse for Grand Prix level is an intensive and long-term project that takes years. A horse that makes its Grand Prix debut before it is eight years old is a rarity. A good constitution, i.e. a healthy physical condition, is very important. The healthier the horse, the greater the chance that it can reach the ultimate Grand Prix level and that it can be maintained for a long period of time.

For the same reason, we also try to align the **conformation**, or the build of a horse, as much as possible with functionality for sport. To even better meet the demand of breeders, owners and riders, the KWPN has had two breeding directions within the Riding Horse type since 2006. Two breeding directions for which the general breeding goal has not changed but expanded. Since 2006, riding horses have been registered and evaluated as either dressage horses or jumpers. The chances of success for a horse that is ‘built’ for his job, is greater than that of a horse whose build works against him. Comparable specific breeding goals also exist for the harness and Gelders breeding directions.

The third factor for success is natural **movement**: does the horse have a correct movement apparatus, one that makes it possible to execute movement with tact, rhythm and balance as well as sufficient power, suppleness and athletic ability.

And finally, the most elusive of all these factors: a **willing character**. The horse after all is not a vehicle, tool or instrument, but an athlete with its own character.

With a view to the intensive and prolonged training path, a teammate that is uncomplicated to handle, easy to ride as well as intelligent and diligent, is probably the most important criterion in the pursuit of the highest sport level attainable.

Breeding directions

Since 2006, the KWPN has distinguished between four breeding directions. Riding horses, which are subdivided into the dressage and jumping disciplines, make up the largest group (85-90%). The other two breeding directions are the harness horse and the Gelders horse. Even though each breeding direction has its own additional objectives, in the first instance all horses fall under the general KWPN breeding goal, which aims at:

breeding a competition horse that can perform at Grand Prix level;

- with a constitution that enables long usefulness;
- with a character that supports the will to perform as well as being friendly towards people;
- with functional conformation and a correct movement mechanism that enables good performance;
- with attractive exterior that is preferably attractive, along with refinement, nobility and quality.

Specialization for Dressage Horse

For the dressage discipline the KWPN also aims for the following breeding goal:

- the dressage horse has a long-lined, generous, correct model with balanced proportions and an attractive appearance.
- the dressage horse moves correctly, is light-footed, in balance with suppleness, carrying power, impulsion, and good self-carriage.
- the dressage horse is easy to handle, easy to ride and intelligent, with a willing and hard-working character.

Specialization for Jumper Horses

For the jumper discipline the KWPN also aims for the following breeding goal.

- the jumper horse has a long-lined, generous, correct model with balanced proportions and an attractive appearance.
- the jumper horse moves correctly, in balance with suppleness, carrying power and impulsion.
- the jumper horse is easy to handle, easy to ride and intelligent, with a willing and hard-working character.

- the jumper horse has courage, jumps with quick reflexes, is careful, has good technique and has a great deal of scope.

Specialization for Harness Horses

For the KWPN harness horse the following applies.

- the harness horse must be able to sustain competition at the highest levels of sport.
- the harness horse moves correctly, in balance with suppleness, carrying power and impulsion.
- the harness horse is easy to handle, easy to drive and intelligent, with a willing and hard-working character.
- the harness horse has proud self-carriage combined with specific trot characteristics: good suspension, a foreleg that moves out well with high knee action and a hindleg that comes under the body with power.

Specialization for Gelders Horses

The breeding goal of the Gelders horse is specifically aimed at:

- a versatile horse that can be used for driving as well as under saddle and that is distinguished by a happy appearance and is reliable with a willing character.
- the Gelders horse moves correctly, in balance with suppleness, carrying power and impulsion.
- in the trot and canter the Gelders horse shows clear knee action in the foreleg and powerful use of the hindleg, good use of the hock and a lot of self-carriage.
- the Gelders horse jumps gladly with good technique and is careful.

Chapter 3.14 Example: Dutch milk production index

In Dutch in cattle breeding the milkproduction index is called INET

Source website Genetic Evaluation Sires (GES): http://www.gesfokwaarden.eu/en/breedingvalues/pdf/E_09_EN.pdf

Introduction

In the world of cattle breeding, selection based on milk production traits plays an important role. The tools used to assist the selection procedure are the cow and bull indexes for milk, fat and protein. The breeding values for the kg of milk, kg of fat and kg of protein are combined to create a single figure: the Dutch production index or Inet. The way in which these breeding values are combined to create the Inet rating is such that selection based on Inet leads to increased profitability in milk production per cow.

The Inet value is calculated according to the following formula:

Inet 2012 = -0.03 * BV kg milk + 2.2 * BV kg fat + 5.0 * BV kg protein

BV stands for breeding value in the formula. The factors –0.03. 2.2 and 5.0 are called the Inet factors. For example: Imagine a bull has the breeding values +1000, +35 and +30 for kg milk, kg fat and kg protein respectively. The Inet of this bull is equal to –0.03* 1000 + 2.2* 34 + 5.0* 30 = 195 euro (to a round figure). The same formula is applied to cows.

Significance of INET

In breeding the central focus is to produce more highly productive and profitable cows through selection. The Inet rating indicates what can be expected from the progeny in extra net yields per lactation if a certain cow is mated with a certain bull. To give an example, we will mate a bull with an Inet of 400 euro with a cow that has an Inet of 200 euro. A calf resulting from this match is expected to have an Inet of 300 euro, which is 100 euro more than its dam. In other words: the calf is expected to yield a net milk production income per lactation of approximately 100 euro more than its dam.

The Inet factors indicate the net yield per kg milk. kg fat and kg protein provided through breeding the production per lactation for milk, fat or protein is increased by one kg. A higher production of one kg of milk per lactation through breeding, without a similar rise in the production of fat and protein, will cost 3 cents. Selective breeding that results in an increased production of one kg of fat will yield €2.20, with €5.00 for one kg of protein.

Calculation model

The economic weighing factors are determined by calculating the difference in farm income if there is a marginal increase in production per cow whereby all the other conditions remain unchanged. The situation (milk price) likely to apply in eight to ten years is taken as the basic assumption in this calculation. The marginal increase in production per cow is the result of the marginal increase of the genetic capacity of the cow for higher production. So, what does an increase in the breeding value of a cow of one kg milk, fat or protein represent at a dairy farm?

Costs for energy and IDP

The calculation model calculates the energy and protein required for milk, fat and protein. To only produce milk or fat, energy is required, producing protein require energy and protein. The feed costs kg milk, fat or protein are calculated as (energy requirement)*price of energy)+ (protein requirement/IDP)*(price of/IDV). Per kg milk, fat and protein is resp. 0.11. 5.9 and 3.0 kFUM (= kVEM) required in energy and for 1 kg of protein 1.56 kIDV= (kVRE) is required.

To calculate the feed costs, a price for medium-priced A-pellets of 18 euro/100 kg is assumed and a price ratio of 6 between kIDV and kFUM : 1. This results in a price of 1 kFUM of € 0.107 and a price of 1 kIDV of € 0.639.

Milk price in the future

In view of the expected trends, the following points have been assumed in the calculation of the Inet factors:

- the milk price is 32 eurocent per kg milk, with 4.2% fat and 3.4 % protein

- the negative land price for 1 kg milk is € -0.015 per kg milk

- the ratio for protein/fat price is 2.25 : 1

- this results in a price for 1 kg fat of € 2.85, and € 6.35 for 1 kg protein.

Results

Based on the energy consumption and the protein demand from feed to produce milk, fat and protein, the feed costs are 0.012. 0.63 and 1.32 euro per kg milk (carrier), kg fat and kg protein.

The yield per kg of milk (carrier), kg fat and kg protein is -0.015. 2.85 and 6.35 euro respectively.

If the costs are subtracted from the yields the net yield is left, taking the feed costs into account:

Inet = -0.027 * BV kg milk + 2.22 * BV kg fat + 5.03 * BV kg protein

Following the weighting factors being rounded up, the Inet in The Netherlands and in Flanders per April 2012 will be as follows:

Inet 2012 = -0.03 * BV kg milk + 2.2 * BV kg fat + 5.0 * BV kg protein

Chapter 3.15 Example: Breeding methods of top breeders in dogs

Breeding methods in dogs

Source: Pekka Hannula and Morjo Nygaerd, 2011. Keys to top breeding ISBN 978-952-67306-5-3. Kirjapaino Jaarli Oy, Turenki, Finland

In Europe, Australia and the United States 22 top breeders of dogs were interviewed for their breeding methods. In their country these breeders were recognized for their long term success in breeding champion dogs at shows: nice healthy dogs with a good behaviour. In their breeding methods they shared the importance of the three breeding goal traits: 1) good health, 2) desired behaviour and 3) a good conformation. At the start of their career they bought dogs according to these criteria and weighed health higher than behaviour and conformation. During their career they insisted on this breeding goal and never bred with dogs that were not healthy or were carrier of a genetic defect. Sometimes they practiced moderate inbreeding (breeding of close relatives) immediately followed by an outcross (breeding of less related dogs). This breeding method and this priority setting of traits in their breeding goal were the key for their success as top breeder.

Chapter 3.16 Example: Production objectives for village poultry in Ethiopia

Breeding poultry in rural areas

Source: Nigussie Dana et al, 2010. Production objectives and trait preferences of village poultry producers in Ethiopia: implications for designing breeding schemes utilizing indigenous chicken genetic resources. Trop. Anim. Health Prod. 42: 1519-1529.

In Ethiopia village poultry systems with indigenous breeds contributes to more than 90 % of the national chicken meat and egg production. This system is characterized by a small flock per household, birds maintained under scavenging conditions in backyards, no supplemental feeding, no separate shelters except for night enclosures in the family house and lack of health care. In a survey the socioeconomic traits of the production system is studied to identify and prioritize the breeding objectives and trait preferences of village producers. These producers prioritize first good adaptability (disease and stress tolerance, flightiness ability to escape predators, scavenging vigour), live weight growth and egg production and second reproduction ability (broodiness, hatchability of eggs) and conformation including size and colour. In addition it revealed that village poultry breeders preferred their local breeds over a reference modern breed for disease ad stress tolerance, escape from predators, the required management level, the scavenging behaviour, the hatchability of eggs and the taste of eggs and the taste of meat. This led to the development of a breeding program with a mass selection breeding scheme for production traits, selection of males and females based on their own production traits, improved the productivity substantially in five generations. Cocks were selected on live weight at 16 weeks and hens at this live weight, their age at first egg and egg production up to 45 weeks. This proves that a breeding scheme does not need to be “sophisticated”. It needs to fit in local conditions (small holders) and add to the value chain.

Breeding program to improve local chicken breed (Horro)

Mass selection based on own performance:

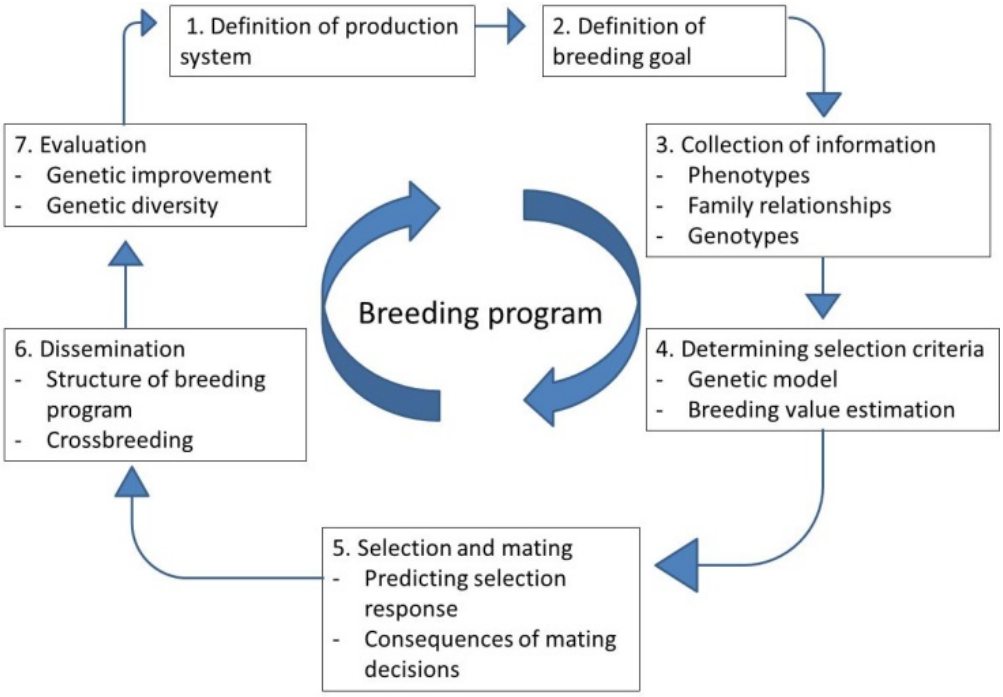
- Growth: based on live weight at 16 wks in both sexes
- Age at first egg in females
- Cumulative Egg number at 45 weeks in females.

Trait	Base	Generation 5
Survival (26 wks)	<50%	97%
Age at first egg (d)	223	150
Body weight 16 wk (g)	550	788
Egg production upto 45 wks	24	65

WAGENINGEN UR
For quality of life

Chapter 4: Collecting information for breeding decisions

Knowing the breeding goal, relevant information should be collected to take breeding decisions. Relevant in this respect are traits of animals (called phenotypes) that can help to establish the value of an animal with respect to the breeding goal. When jumping performance is a breeding goal trait in horses, data on jumping are collected. When fertility of pigs is in the breeding goal, litter traits are recorded. Also, essential information in animal breeding, is the pedigree of the animals. Animal breeding is all about passing genetic abilities from one generation the next. When you want to trace or influence this process of passing on, a registration of the parent-offspring relationships, the pedigree of an animal, is crucial. Nowadays, DNA-analysis is possible and practised in animals and can also be used to trace or influence the passing process of genetic abilities for traits.



Chapter 4.1 The value of a pedigree in animal breeding

Due to the events in the meiosis a sperm cell and an oocyte contains 50 % of the DNA of the animal producing the sperm cell or oocyte (an important fact in the transmission from DNA of parents to offspring). This 50% of the DNA is comprised in an unique combination of chromosomes of the animal (random process in the transmission of DNA between

generations). After the fertilization of an oocyte with a sperm cell the nucleus of the zygote contains again chromosomes in twofold grouped in pairs. This process implies that each animal receives half of its chromosomes (or half of its genetic value) from its sire and half of its dam. Thus the genetic relationship between an animal and each of its parents is 0.5. This is called the additive genetic relationship.

Definition

The additive genetic relationship between two animals is the amount of DNA they share due to the fact that they are related

A few additive genetic relationships are presented below:

Relationship	Percentage of similar DNA they share
Parent-son or daughter	50
Grandparent-grandchild	25
Great-grandparents- great-grandchild	12.5
Full brother(s) – full sister(s)	50
Half-brother(s) – half-sister(s)	25

It is important to stress that each sperm cell and each oocyte from a parent contains a unique combination of the chromosomes of the parent causes that the offspring of the same combination of a sire and a dam (full sibs) still show differences in traits, due to the Mendelian sampling effect. The additive genetic relationship of full brothers and full sisters (full sibs) is 0.5, because on average they share 50 % of the DNA of their parents.

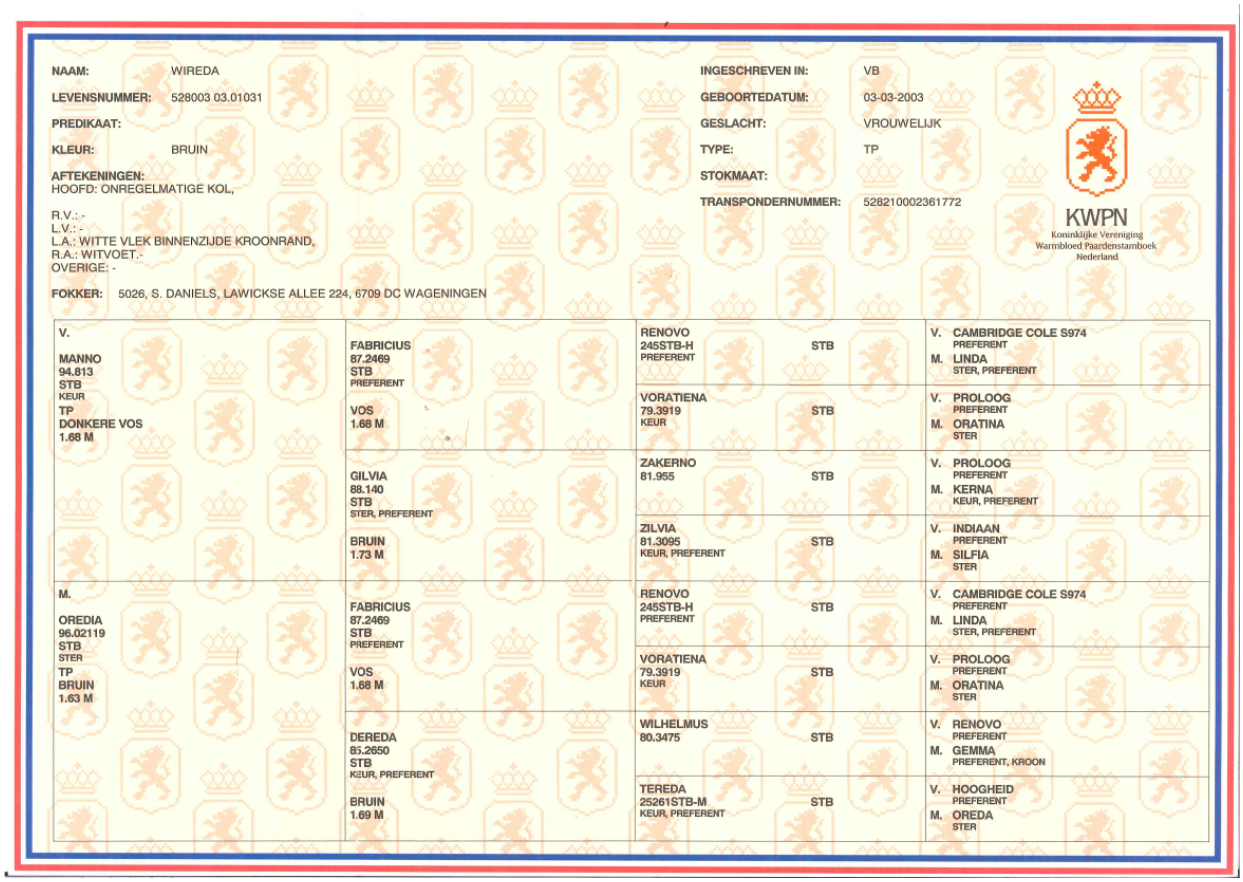
The additive genetic relationship between two related animals is of great importance in animal breeding. Because, for example, the additive genetic relationship between a sire and a daughter is 0.5, they share 50 per cent of their DNA. This means that the traits of a sire have a predictive value for the traits of a daughter. The reverse is also true: the traits of the daughter maybe used to calculate the breeding value of her sire. Of course the heritability of the traits play in this respect a crucial role. For traits with a high heritability the additive genetic relationship between two animals has more impact in breeding than for traits with a low heritability. For instance, height of withers (with a high heritability of 0.6) of a mare is a good predictor of the height of withers of her daughter. However, the success of insemination has a

low heritability of 0.1. This means that the predictive value of the insemination success of the mare for the insemination success of her daughter is low, despite that the additive relationship between the mare and her daughter is 0.5. Therefore, the pedigree of an animal is very informative.

Definition

A pedigree is the set of known parent-offspring relationships in a population, often displayed as a family tree diagram. This can be used to derive the relationships among animals in a population

Below you can find an example of a pedigree of a mare. This is a print of an official studbook as a certificate of the Dutch Studbook for horses (KWPN).



NAAM: WIREDA
LEVENSNUMMER: 528003 03.01031
PREDIKAAT:
KLEUR: BRUIN
AFTEKENINGEN:
HOOFD: ONREGELMATIGE KOL,
R.V.: -
L.V.: -
L.A.: WITTE VLEK BINNENZIJDE KROONRAND,
R.A.: WITVOET.-
OVERIGE: -
FOKKER: 5026, S. DANIELS, LAWICKSE ALLEE 224, 6709 DC WAGENINGEN

INGESCHREVEN IN: VB
GEBOORTEDATUM: 03-03-2003
GESLACHT: VROUWELIJK
TYPE: TP
STOKMAAT:
TRANSPONDERNUMMER: 528210002361772

KWPN
Koninklijke Vereniging Warmbloed Paardenstamboek Nederland

V. MANNO 94.813 STB KEUR TP DONKERE VOS 1.68 M	FABRICIUS 87.2469 STB PREFERENT	RENOVO 245STB-H STB PREFERENT	V. CAMBRIDGE COLE S974 PREFERENT M. LINDA STER, PREFERENT
		VORATIENA 79.3919 STB KEUR	V. PROLOOG PREFERENT M. ORATINA STER
	VOS 1.68 M	ZAKERNO 81.955 STB	V. PROLOOG PREFERENT M. KERNA KEUR, PREFERENT
	GILVIA 88.140 STB STER, PREFERENT	ZILVIA 81.3095 STB KEUR, PREFERENT	V. INDIAAN PREFERENT M. SILFIA STER
	BRUIN 1.73 M		
M. OREDIA 96.02119 STB STER TP BRUIN 1.63 M	FABRICIUS 87.2469 STB PREFERENT	RENOVO 245STB-H STB PREFERENT	V. CAMBRIDGE COLE S974 PREFERENT M. LINDA STER, PREFERENT
		VORATIENA 79.3919 STB KEUR	V. PROLOOG PREFERENT M. ORATINA STER
	VOS 1.68 M	WILHELMUS 80.3475 STB	V. RENOVO PREFERENT M. GEMMA PREFERENT, KROON
	DEREDA 81.2650 STB KEUR, PREFERENT	TEREDA 25261STB-M STB KEUR, PREFERENT	V. HOOGHEID PREFERENT M. OREDA STER
	BRUIN 1.69 M		

Chapter 4.2 A unique identification system for animals is essential

A pedigree only has a predictive value when a unique and reliable identification system is used in a breeding program. At birth each animal should get a unique identification number and its parents should be known without any doubt. In many breeding programmes the pedigree is checked with genetic markers (see example in this chapter). A second requirement is that measurements (phenotypes) on animals (e.g. height of withers, milk production etc.) are combined with the correct identification number. Mistakes in a pedigree and mistakes in the attachments of data to animals in the recording system are disastrous for the predictive value of pedigrees.

The parents of an animal have an additive genetic relationship of 0.5 and the grandparents of 0.25. The shorter the distance between an ancestor and the animal itself, the more valuable are traits of the ancestor in predicting traits of the animal at stake. In the past, herd books started with pedigree registration and pedigree checking to guarantee the buyers of breeding material that the traits of an animal could be derived from the traits of its ancestors.

In addition to the additive genetic relationship with the ancestors more additive genetic relationships between animals can be used in animal breeding. An animal may have full sibs with recorded traits: the additive genetic relationship among full sibs is 0.5. On average, they share 50 per cent of their DNA supplied by their parents. Half sibs share on average 0.25 of their DNA (they have the same dam or the same sire) and this additive genetic relationship of 0.25 with its half sibs might be also valuable for predicting the traits of a (young) half sib. Full or half sibs do have an additive genetic relationship of 0.5 with their common parent and are often used to establish the breeding value of one of these parents. Even offspring in later generations, granddaughters and grandsons with and additive relationship of 0.25, are used to establish the breeding values of their grandparents. In conclusion: it is worthwhile and very informative to extend the pedigree of an animal in a scheme where, in addition to the ancestors, also full and half sibs and offspring are situated. This gives a full picture from all the relatives that might provide information for the estimation of the breeding value of the animal at stake.

Chapter 4.3 Collecting phenotypes, monogenic and polygenic traits

Some traits of animals do not need a lot of knowledge or experience to be recorded. The colour of the animals is a good example: e.g. in a breed of rabbits the animals are black or brown. You may record it in a computer as 0 for black and 1 for brown or 1 for black and 2 for brown. In

genetic terms such traits are monogenic: the expression is determined by the alleles of a single gene. Like colour, some traits are based on a very limited number of genes and that is the cause that only a limited number of classes can be recorded. Many recessive defects in animals are monogenic: the alleles of one gene determine the phenotype: healthy or affected. In statistical terms these are discrete variables, recorded in a limited number of classes. Describing the traits of the breed of rabbits you can calculate that x % of the animals are black and y % of the animals are brown.

Many traits of animals are polygenic, caused by effects of many genes. Many polygenic traits are quantitative and continuous and can be measured in metric units such as kg, l, mm, etc. Some traits are continuous, but measured in categories approximating a linear scale, such as e. g. traits of the conformation of the animals or on performance assessed by judges or inspectors. They score the traits of animals on a scale e.g. from 1-5 or 1-10. Some polygenic traits such as occurrence of disease are on binary scale: sick (e.g. 1) or not (e.g. 0).

Chapter 4.4 Mean, variation, standard deviation and coefficient of variation

In statistical terms continuous variables are described by a mean and measures of variation. To describe the variation, standard deviation, variance and coefficient of variation can be used.

The mean is calculated as follows:

The "mean" of a sample is the sum the sampled values divided by the number of items in the sample:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

For example, the arithmetic mean of five values: 4, 36, 45, 50, 75 is

$$\frac{4 + 36 + 45 + 50 + 75}{5} = \frac{210}{5} = 42.$$

The variance is calculated as follows:

$S^2=(X_i-X)/(N-1)$

The standard deviation is calculated as follows:

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2},$$

For example the standard deviation is the square root from the variance and in this case for the five values: 4, 36, 45, 50, 75 is to be calculated as:

N=5 and the mean for x= 42

X_i	X_i-X	$(X_i-X)^2$
4	-38	1444
36	-6	36
45	3	9
50	8	64
75	33	1089
X_i= 210		$(X_i-X)^2$= 2642

In this case the variance is 2642/4 =660.5 and the standard deviation is 2642/5= 32.5

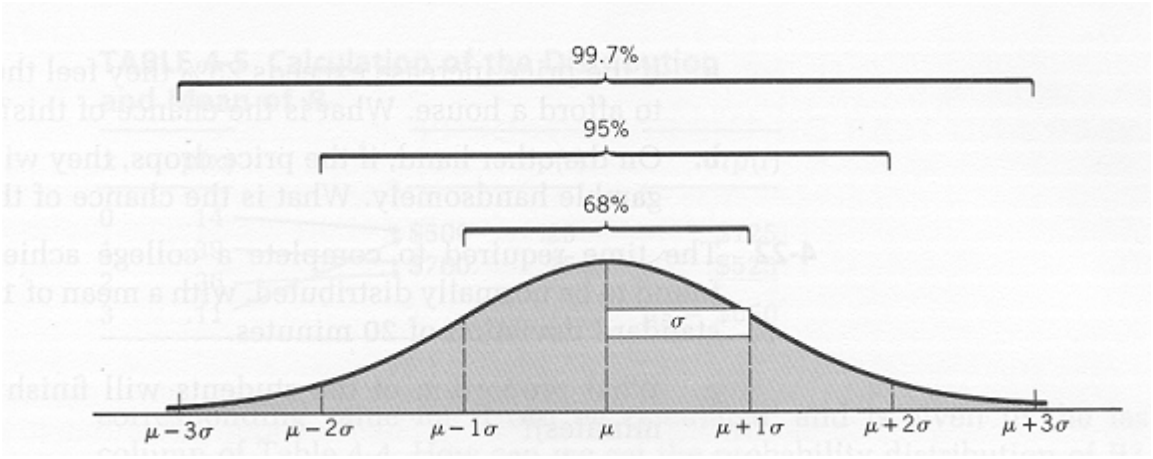
The coefficient of variation is the standard deviation divided by the mean and is calculated as follows:

$$c_v = \frac{\sigma}{\mu}$$

In this case μ is the indication for the mean and the coefficient of variation is: 32.5/42 = 0.77. This means that the size of the standard deviation is 77% of the size of the mean. This implies that you see a lot of differences among animals when the five values above are the value of a trait of five animals.

Chapter 4.5 Normal distribution of measurements

Many traits of animals show a normal distribution. This means that the distribution is symmetric and can be characterized by a mean and variance. Below and above this mean you see the same number of animals. And by the fact that further away from the mean you see a decreasing number of animals with this measured unit. Graphically it can be expressed as:



In such a normal distribution you see that 68 per cent of the animals has a value for this trait within one standard deviation from the mean; 95 per cent within two standard deviations from the mean and 99.7 per cent within three standard deviations from the mean.

Chapter 4.6 Covariance and correlation

Two traits might have a relationship. E.g. when the value for trait one is high you always see that trait two also has a high value (see figure below, the relationship between heart girth and live weight in cows) or just the opposite when trait 1 is high, trait 2 has a low value (see figure below, the relationship between live weight and feed conversion in pigs). The relationship can also be low (see figure below, the low relationship between live weight and sale price in cattle). This might be caused e.g. that these traits are (partly) based on the functions of the same genes. In animal breeding we frequently use the covariance, correlation or regression as a statistical description of such relationships between traits.

In statistical terms **the covariance** is equal to: cov(x,y) = E(xy) – E(x)* E(y)

Where E stands for the expectation, which can be calculated as the summation divided by the number of observations

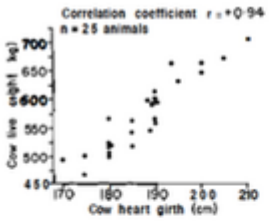
The relation between two traits is in animal breeding mostly described as the correlation between the traits x and y.

In statistical terms **the estimated correlation** is: r(x,y) = cov (x,y) / (var x * var y)

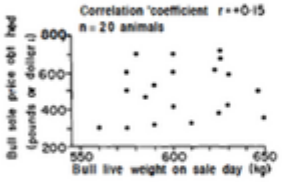
The correlation is usually denoted as r and has a value between -1 and + 1. A plus means that two traits are positively correlated: high values of trait x coincidence in most cases with high value of y (in case r=+1 always). A negative sign means that high values of x coincidence with low values of y.

The scheme at right illustrates in plots relationships (correlations) between two traits in three different cases:

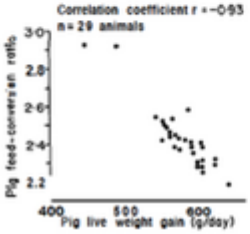
It is very important to understand that the correlation does not indicate cause and consequence or result. A live weight in pigs is not directly the cause of a low feed conversion ratio in pigs (third example in the scheme above) or the reverse. The correlation only indicates that a relationship between the two traits exists. When based e.g. on the function of the same genes, this relationship can be used in breeding.



Heart girth with live weight: r=0.94



Live weight with sale price: r=0.15



Live weight with feed conversion ratio: r= --0.93

Chapter 4.7 Regression

But in case of a rather high correlation the question arises: How much change in y do you see with 1 unit change in x? That question can be answered by the regression coefficient: A measure of relation between x and y, but expressed as proportion the variance of x. It is usually denoted as b.

In statistical terms: $b(x,y) = cov\ (x,y) / (var\ x)^2$

In words b(x.y) = is the change in the value of y when x is one unit higher. It is also possible to calculate the change in x when y is one unit higher.

That regression coefficient can be calculated as: $b(y,x) = cov\ (x,y) / (var\ y)^2$

The regression coefficient can have a positive or a negative sign dependent on the sign of the correlation (covariance) between the traits. The regression coefficient, e.g. from daughter's milk yield on dam's milk yield can be used to estimate the heritability for milk yield.

Chapter 4.8 Measurement errors

The result of a measurement on an animal is called a phenotype. Measurements should always be carried out very carefully and should be critically reviewed. Measurement errors determine how accurate phenotypes can be established.

Definition

*A **phenotype** is the observed value of a trait. It is a consequence of all the enetic and environmental influences and interactions affecting the trait, including errors in measurements*

These measurement errors might have a systematic and/or random character. Systematic errors might be caused by differences between animals for example: diet composition, age at measurement, training etc. Random errors, e.g. measurements errors, may cause a low repeatability for a trait. For instance, when you want to measure the lengths of an animal, and you repeat this ten times you will see a rather large variation in outcomes. Slight movements of the animal are the cause. When you measure pelvic height, you see a rather small variation in outcomes. Two measures of accuracy of measurement are relevant: **repeatability and reproducibility**, both are correlations between measures on the same animal.

Definition

*The **repeatability** is defined as the extent to which measurements on the same object under similar conditions correspond with each other. It indicates how accurate a trait can be established. It is influenced only by measurement errors and time effects*

When the repeatability is low for a trait, the heritability will also be low. That means that is difficult to improve that trait in a breeding program.

Definition

*The **reproducibility** is defined here as the relationship between measurements in different locations and/or by different persons. Besides measurement errors and time effects it is also influenced by systematic effects e.g. of classifiers or technicians*

When the repeatability for a phenotype is high and the reproducibility is low, than efforts need to be taken to standardize the measurement of the phenotype and to train classifiers or technicians. For instance body weight of the piglets is measured at weaning when the piglets are separated from the sow at a fixed time during the day and they are weighed at a fixed time of the day after a fixed time of feed restriction before transportation to the slaughter house.

Then, it is possible to calculate daily gain during fattening without systematic errors. When scoring body traits of horses it is highly recommended to start with a training session for the inspectors and repeat such a training session at regular intervals. Otherwise, two judges might give (systematically) different scores for the same trait of the same animal, leading to a low reproducibility of the trait despite its high repeatability.

Chapter 4.9 Frequency of measurements

The frequency of measuring a phenotype depends on many factors. A milking robot is recording the milk yield of individual cows at each milking. The conformation of animals established by an inspector requires an expensive visit to a farm or and is done only a few times per year. The incidence of a disease is only recorded when a veterinarian is called in by a farmer. Performance of a horse is established at a test-show. Conformation of a dog is recorded at special shows. In general you increase the number of recordings per animal when a trend in a trait may exist. E.g. milk production curves of cows, sheep and goats may differ over the lactation period: top production (kg and days in lactation) and persistency (shape of the curve: flat or sharp decline after the top). To get a reliable estimate of lactation yield (the breeding goal trait) it was figured out that milk recording at least once every six weeks is necessary. At many milk producing farms the first argument for milk recording is to support a proper feeding management of the animals and the second argument is to detect the best animals for milk production.

Chapter 4.10 Measurements on the animal or on relatives?

From which animals phenotypes should be recorded, depends heavily on traits of the trait: growth can be measured in males and females from birth to death; milk production can only be measured in females after the first delivery of progeny; eggs can be counted after the start of the laying period. Carcass traits will be known after slaughter. Disease resistance is only expressed in the presence of pathogens. Longevity can be fully established at the end of life of an animal. Therefore for categories of traits different sources of information are used to get an impression of the genotype of the animal to be considered for selection. These sources of information are informative because they are related to the animal under consideration.

- Parents (pedigree) information: milk production, fertility, longevity
- Sibs or half-sib (brothers and sisters) information: milk production, carcass traits, fertility, longevity or disease challenge
- Progeny: milk production, carcass traits, fertility, health

Recording of traits cannot be done without costs. In dogs and horses, for example, you have to organize shows resulting in costs because inspectors have to be hired. In other species farms have to be visited by inspectors to assess the conformation of animals. For other species costly machines have to be bought to scan live animals or to record automatically important traits in slaughter houses. Low-cost methods facilitate recording traits on a large number of animals, but in case of high costs, e.g. ultrasound scanning of live animals, only animals with a high impact in a breeding program will be scanned.

Phenotypes are recorded by a variety of organizations and this requires additional efforts to combine the phenotypes of an animal before its traits maybe combined to calculate its breeding value for all traits in the breeding goal.

Chapter 4.11 Value of indicator traits

For important traits that are difficult to measure or are expressed later in life or at the end of life indicator traits might be of value. If in horses a good relationship (correlation) exists between the score for the quality of the legs and longevity, the score for the leg quality can be used as a predictor of longevity. In that situation leg quality may be used as a selection criterion in horses to improve the breeding goal trait longevity. In pigs a halothane test was developed in which young pigs were anaesthetized with halothane. This gas gave different reactions in pigs that were a good predictor of their stress susceptibility. Stress around slaughter has a deteriorating effect on pork quality. The outcome of the halothane test is used as an indicator trait be selected for to decrease stress susceptibility and consequently improve pork quality.

Stress susceptible swine Stress resistant swine

Chapter 4.12 Value of information of relatives

Information of ancestors

When an animal is born, or even when an embryo is available, a breeder likes to know the value of this individual for breeding. Is this young animal only capable to perform the intended job, production, leisure activities or company or can I use this animal also as parent for the next generation? Can I use it to improve future generations? The first opportunity is to study the pedigree and to collect all the information of the ancestors in the pedigree. The information of the parents is very valuable as the additive genetic relationship between parents and its progeny is 0.5. Information of grandparents and ancestors in later generations is only valuable when the information of parents is absent or limited. For, alleles that are not passed from a grandparent to a parent cannot be present in the animal at stake. Pedigree information is very useful at the time that the traits or performance of an animal itself cannot be established (yet). You may think of males you want to select for traits only expressed in females and of traits only expressed after puberty (milk and egg production, fertility), after slaughter (carcass traits) or late in life (age related defects, longevity).

Information of the animal itself

As soon as a trait can be measured on the animals itself, the value of the information of its ancestors becomes less. Then, the genetic value of the animals is expressed and it becomes clear which 50 % of the genetic value of the sire and of the dam went to this particular animal. Own information is very valuable when a trait has a high heritability. Hence, the measurement errors or environmental effects for that trait are very limited.

Information of sibs

In some species, poultry and pigs, full sib families exists. In poultry, hens and cocks can be mated and subsequently they may produce hundreds of full sibs. In pigs an average litter comprises 14 piglets being full sibs. The additive genetic relationship among full sibs is 0.5 and this means that full sib data give indeed information on the breeding value of an individual full sib. In case were full sibs are born spread over time, older full sibs might be informative for the breeding value of a younger one. An application of sib selection is seen in pigs were carcass data of a slaughtered full sib is used as information for the carcass traits of a full sib to be selected for breeding. In dogs a full sib trained as seeing-eye dog can give information for a full sib to be selected for a seeing-eye dog breeding program. In most species sires are mated to several dams and in this way half sib groups are created. The additive genetic relationship among half sibs is not that high (0.25). Information of a single half sib has a rather low value for an individual half sib to be selected. Only when a high number of half sibs is available, this information is valuable.

Information of half or full sib groups

In most species sires get relatively large half sib groups. The most pronounced example are dairy bulls used in artificial insemination programs. In the traditional dairy cattle breeding programs young sires produce a first group of daughters. With the first crop of daughters often

consisting of more than 50 daughters each with an additive genetic relationship with the sire of 0.5, it is a very informative half sib group. In some species, pigs, poultry, dogs or fish full sibs are born. The individuals within a litter all have an additive genetic relationship of 0.5 with their dam and with their sire. Full sib groups can be a very valuable source of information.

Combination of information sources

In breeding programs data of animals are sampled and stored continuously in databases. These databases contain traits from ancestors, living breeding animals to be selected, their sibs and their progeny. For living animals, to be selected for breeding, all these data can be combined in statistical methods to estimate their breeding value. The number of generations between the animal to be selected and the animal from which interesting data are stored in the database determines the additive relationship and the usefulness in estimating the breeding value of the animal to be selected. In addition the value of the information depends on the nature of the character (sex limited, when can it be measured in life etc.). And numbers of relatives with data are important: one granddaughter with carcass data is hardly informative for the genetic value of a grandsire, but when there are thousands of granddaughters, as is the case in pork production, these data are very valuable.

Chapter 4.13 The possibilities of DNA analysis

Besides phenotypes, we can do collect DNA information for various purposes. DNA can be found in chromosomes and harbours the genetic differences between animals. These genetic differences in DNA are mainly caused by differences in nucleobases.

Sometimes a nucleobase in a chromosome is substituted by another nucleobase: a point mutation, the base of variation in DNA composition among animals that can be established with molecular genetics methods. An sequence of nucleobases may function as a gene that is responsible for the production of a protein.

A (point)mutation, the substitution of one or a few nucleobases, results in another nucleotide and might result in the production of a new protein or to the absence or malfunctioning of the protein.

Definition

*A **mutation** is an event that creates a change in the DNA sequence on a chromosome of an individual so that the sequence is not the same as that inherited from either sire or dam. In genetics this has most impact when the mutation occurs in germ cells so that it is passed to offspring. Mutational events are caused by irregularities in cellular processes, and when the mutation alters the function of the sequence in which it occurs it may introduce new genetic variation into the population*

Nearly all traits of an animal are determined by many genes. When one of these genes has been subject to mutations in the past and when it has a measurable impact on one of the traits, it is called a Quantitative Trait Locus (QTL).

Definition

A **QTL** is a Quantitative Trait Locus, a discrete, small segment of DNA that has a substantial effect upon a trait. Only a few QTL with large effects have been found. Most complex traits such as body weight and milk production seem to be regulated by many genes, which approaches effectively the assumption of most quantitative genetic theory that traits are affected by an infinite number of genes each with small effect

Chapter 4.14 DNA markers

From a few percentage of the DNA we know its function: it are genes responsible for the production of proteins. From large amounts of DNA, situated between genes, we do not know the function. Nevertheless we are able to establish in the lab differences in the composition of this DNA. Several molecular genetic techniques are used to find genetic markers on the chromosomes.

Definition

A **genetic marker** is a specific and identifiable sequence of DNA

Sometimes a genetic marker is an allele of a gene producing a protein. It is called a functional marker that affects directly the function of the protein. However, in most cases the genetic marker is a piece of DNA from which we do not know the function. It is situated on a chromosome close to a gene and therefore related to one of the alleles of this gene. From the start of the molecular genetic work in animals, animal geneticists have done a lot of work to find genetic markers. Before that start, they only knew that a sire and a dam passed one of their two alleles to their offspring, but they did not known which ones. With genetic markers, it is possible to trace which allele of a sire and which allele of a dam is passed to the offspring.

Chapter 4.14.1 Parentage control

Genetic markers offer some important applications in animal breeding. The first is parentage control. It is based on the fact that a sire and a dam pass one of the two alleles of a genetic marker to its offspring. Thus, from the two alleles you establish in a son or daughter, one should be present in the sire and one in the dam (see example).

Mistakes in pedigrees may be caused by interchanges of parents (or semen) at mating, unnoticed matings, interchange of young animals shortly after birth or administrative mistakes. From experiences in parentage control it is known that 2 – 10 per cent of the animals have a wrong pedigree. In breeding programs with high costs, parentage control is highly recommended, especially when animals are kept in large numbers in a breeding unit, where mistakes easily may pop up.

Example: parentage control with 18 microsatellites in dogs (source: "Het fokken van rashonden", Kor Oldenbroek and Jack Windig, Raad van Beheer op Kynologsich gebied in Nederland) .

Two female dogs named Marjolein and Martha are born on the same day in the same kennel. The kennel owner considered Marjolein to be the daughter of the female dog Lianne and the male dog Boris. In his view Martha has as parents the female dog Lieneke and the male dog Bart. The two males, Borus and Bart, are owned by a neighbouring breeder. Lianne was mated to Boris on the same day that Lieneke was mated to Bart. As usual, before printing the official pedigree, parentage control with 18 micro-satellite markers was carried out to verify the pedigree.

DNA research of all six dogs is presented below with for each microsatellite the number of the two alleles established :

Microsatellite	Marjolein	Lianne	Borus	Martha	Lieneke	Bart
1 AHT 121	102/102	102 /102	97/102	97/102	97/102	102 /102
2 AHT 137	149/151	147 /151	128 /147	147 /149	149 /151	149 /151
3 AHTH 171	219/225	219 /225	212 /233	227 /233	227 /229	219 /219
4 AHTH 260	254/252	254 /246	252 /250	252 /244	244 /244	252 /244
5 AHTK 211	93/93	93/95	91/95	91/93	93/93	93/97
6 AHTK 253	284/288	288 /290	288 /288	288 /288	286 /288	284 /288
7 CXX 279	126/126					

		126 /128	124 /128	124 /128	126 /128	124 /126
8 FH 2054	152/152	152 /164	152 /156	156 /160	152 /160	152 /156
9 FH 2848	230/234	234 /234	230 /230	230 /230	230 /230	230 /234
10 INRA 21	97/101	97/101	95/101	95/101	95/97	95/101
11 INU 005	126/126	126 /126	126 /128	132 /128	132 /126	130 /126
12 INU 030	144/144	144 /150	144 /144	144 /144	144 /150	144 /144
13 INU 055	210/214	210 /218	210 /212	210 /216	212 /216	214 /216
14REN162C04	202/204	200 /202	200 /204	202 /204	200 /202	200 /204
15REN169D01	212/218	212 /212	218 /218	214 /218	214 /218	216 /218
16REN169O18	162/164	162 /162	164 /170	164 /170	164 /168	164 /168
17 247M23	268/268	268 /270	268 /272	268 /268	268 /274	268 /274
18 54P11	226/226	226 /236	226 /232	226 /226	226 /232	226 /234

Marjolein has for microsatillite 1 (AHT 121) allele 102 in twofold. This allele is present in Lianne as well in Borus. For microsatellite 2 (AHT 137) Marjolein has the alleles 149 and 151. Allele 151 can be traced back to her dam Lianne, but her sire Borus (and her dam Lianne!) does not have allele 149! For microsatellite 3 (AHTH 171) Marjolein has the alleles 219 en 225, similar as her dam Lianne, but her sire Borus has the alleles 212 and 233. When you continue to

check in this way all alleles for the 18 microsatellites present in the six animals, you can conclude that Borus cannot be the sire of Marjolein on base of the microsatellites 2, 3, 5, 6, 7, 13. On base of the alleles of Bart for the 18 microsatellites it may be very reasonable that Bart is the sire of Marjolein.

Considering the alleles of Martha and Lieneke you can conclude that it may be very reasonable that they are daughter and dam. But Bart cannot be the sire of Martha based on the alleles for the microsatellites 2, 3, 5, 11, 13 and 16.

A comparison of the alleles of Marjolein with the alleles of Lianne and of Bart indicate that Marjolein may be born out of a mating of Lianne and Bart and that Martha is born out of a mating of Lieneke and Boris. Apparently, something has been going wrong during the matings.

Chapter 4.14.2 Marker-assisted and genomic selection

Marker-assisted selection

The second application of a genetic marker is the tracing of alleles with a favourable effect in marker-assisted selection. Many genetic markers were found in production animals that were closely linked to a QTL with a favourable effect on many traits. Only a few QTL have been found; therefore the use of markers in selection was limited until genomic selection was introduced.

The third application of a genetic marker is the tracing of alleles with an unfavourable effect. First rate examples are monogenic recessive genetic defects that are present in all species. The next table gives an overview of the total number of recorded genetic defects per species, the disorders that are monogenic recessive traits (Mendelian trait), the disorders from which the mutation in DNA is known and for which a genetic marker available and the number of genetic defects that can be used to study human diseases:

Genetic defects: http://omia.angis.org.au/home/

Summary

	dog	cattle	cat	sheep	pig	horse	chicken	goat	rabbit	Japanese quail	golden hamster	Other	TOTAL
Total traits/disorders	580	397	302	214	214	206	190	72	58	41	40	463	2777
Mendelian trait/disorder	223	145	75	88	45	40	114	13	28	31	28	146	976
Mendelian trait/disorder; key mutation known	154	78	40	32	18	29	36	9	7	9	3	58	473
Potential models for human disease	296	142	165	82	77	108	41	28	37	11	14	229	1230

Genetic markers for monogenic recessive traits are very valuable, because they can be used to detect the heterozygote carriers of the allele, heterozygous animals that do not show any symptoms of the genetic defect but do transmit it to 50 per cent of their offspring. The mating of two heterozygous animals gives with a chance of 25 per cent offspring that is showing the symptoms of the genetic defect.

Genomic selection

The fourth application of genetic markers is genomic selection. Genomic selection is a form of marker-assisted selection in which a very large number of genetic markers covering the whole genome are used. In this case all quantitative trait loci (QTL) are closely linked at the chromosomes with at least one marker. The large number of markers is obtained by chips using Single Nucleotide Polymorphims (SNP's), a point mutation of a single nucleotide. The genomic selection is based on the analysis of 10.000 up to 800.000 SNP's. This high number of genetic markers is used as input in a genomic prediction formula that predicts the breeding value of an animal.

In animal breeding, the genetic markers have the highest value for the improvement of traits with a low heritability and for traits that can be established in one sex, late in life or after slaughter.

Definitions

Genomic selection *is selection for a trait of interest with a very large number of genetic markers covering most QTL-loci related to the trait*

An ***SNP*** *is a single nucleotide polymorphism caused by a mutation of a single nucleotide*

A complicating factor is recombination between SNP's and QTL's. This means that the value of animals in the reference population slows down when the number of generations between them and the test population increases (more chance for recombination events). And it implies that it is highly recommended to continue the recording of phenotypic data of future generations.

Whole genome sequencing

Recently whole genome sequencing is introduced. Whole genome sequencing (also known as full genome sequencing, complete genome sequencing, or entire genome sequencing) is a laboratory process that determines the complete DNA sequence of an organism's genome at a single time. This entails sequencing all of an organism's chromosomal DNA as well as DNA contained in the mitochondria . This technique is used in research settings, but practical use in selection may be expected as it opens the possibility to select directly for favourable alleles of QTL's.

Chapter 4.15 Key issues in collecting information for breeding decisions

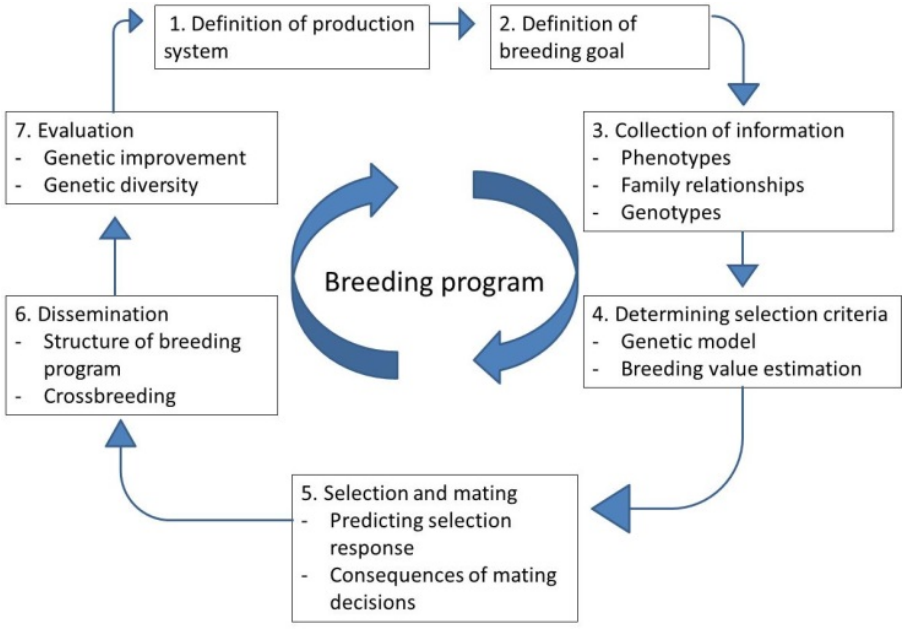
1. Pedigree recording is a backbone of animal breeding because it can be used to establish the additive relationship among animals. This additive relationship is the amount of DNA these animals share because they are related.
2. A pedigree only has a predictive value when a unique and reliable identification system is used in a breeding program.
3. At birth each animal should get a unique identification number and its parents should be known without any doubt. Another requirement is that measurements (phenotypes) on animals (e.g. height of withers, milk production etc.) are combined with the correct identification number.
4. Monogenic traits can be recorded in classes and scored as 0 or 1 (in case of two classes) or 1, 2, 3 , 4.... n, in case of n classes.
5. Polygenic traits can be recorded on a numerical scale. The average value of such traits in a population is presented as the mean and the variation among animals as the standard deviation. The coefficient of variation is the standard deviation divided by the mean.
6. The connection between two polygenic traits can be described by the covariance between the traits, the correlation and the regression.
7. Selection traits should preferably be measured on selection candidates. For important traits that are difficult to measure or are expressed later in life or even at the end of life indicator traits might be of value.

8. Selection traits can also be measured, due to the existence of additive relationships, on ancestors, full sibs or half sibs or on progeny.
9. DNA markers can be used for parentage control, marker assisted selection for positive traits (e.g. product quality traits), marker assisted selection against negative traits (e.g. genetic defects) and genomic selection.

Chapter 5: Genetic models

Even though for some animal species the full genome has been mapped, we still cannot 'see' which in a population of animals would be the genetically superior ones. We cannot read the working of the DNA in full detail yet. So instead we need to make an estimate of the genetic potential of the animal based on its phenotype. How we can do that will be the subject of the chapter about ranking the animals. In this chapter we will look into how we can get an impression of how much of the phenotypic variation in performance we observe in our population is due to actual genetic differences between the animals. For example, is a cow that produces an average of 25 kg per day in her lactation indeed a genetically better cow than a cow that produces 15 kg per day? Is a horse that always has high scores in a dressage test indeed a genetically better horse than the one that scores much lower? And why is the full sister of that excellent hunting dog not also performing at top level with that same trainer? The answers to these question may not all be the same. In this chapter we will find out why.

If we look at the diagram again with the stages involved in the breeding program circle, then we are at the stage number 4. In the previous chapters we have identified the breeding goal and we have collected measurements on phenotypes and genotypes and pedigree of the animals. In this chapter we will define the genetic model that we will use as tool to translate the measurements into a set of criteria that we can use to rank animals for selection.



Chapter 5.1 Phenotype and environment during life history

Generally not all variation in observed phenotypes is a result of differences in genetic makeup between animals. Part of it is determined by variation in what we call *the environment*. Often the environment has a very important influence on the animal's phenotype. An impression of what type of influence the environment has in the phenotype is given in figure 1. The figure represents the life span of an animal, indicated by the green arrow. The vertical bars in the arrow represent important events in life that indicate the start of a new phase in the *life history* of the animals. Above these bars are text boxes that describe the type of event. Below the arrow are text boxes that indicate what kind of environmental influences act on the animal during each stage of its life history.

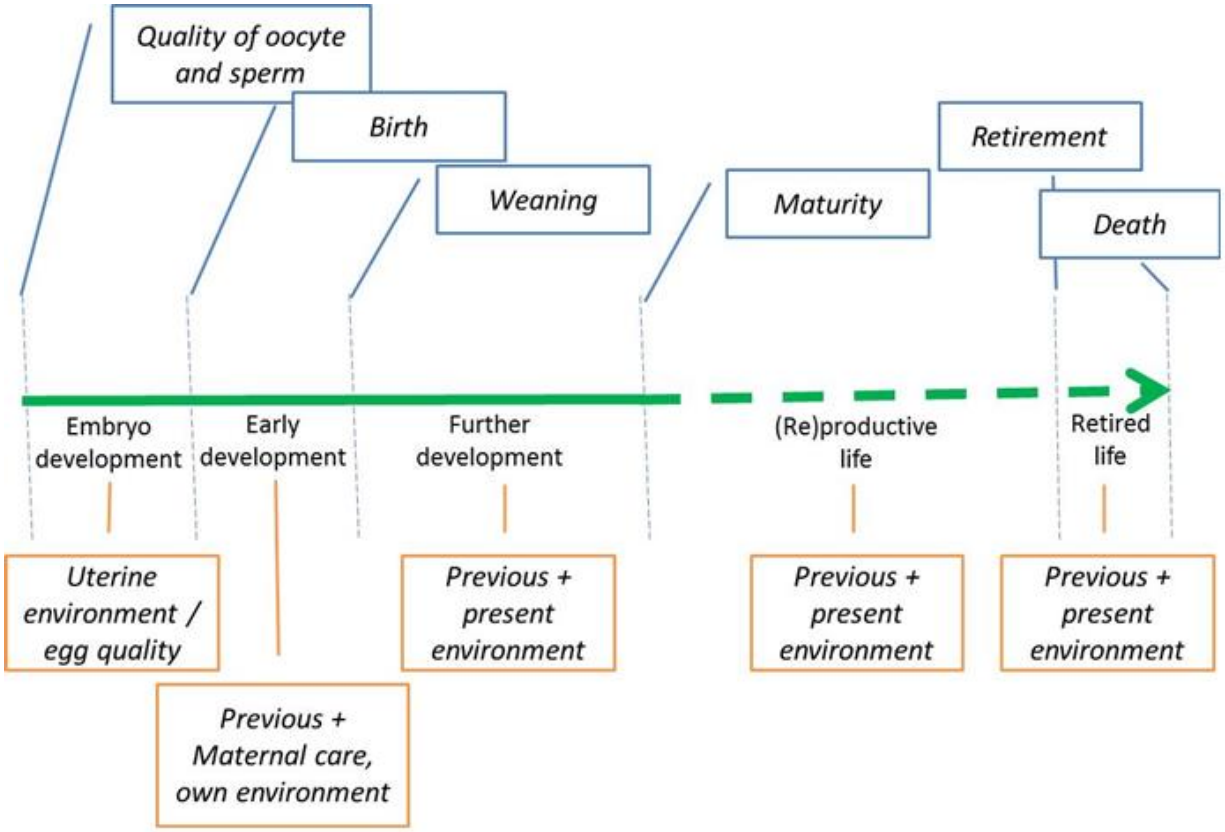


Figure 1. Overview of the lifespan of an animal and important events that indicate changes in environmental influences. In the boxes at the bottom is indicated that the current environment is of influence, but also that in the past. Until weaning the mother determines part of the environment.

Chapter 5.1.1 Events before birth

The life history of an animal starts with the fusion of an oocyte and a spermatozoid: the conception. What happened to these two before conception has an influence on their quality, and thus with the start-up conditions of the animal. The next important phase is that of development through embryonic an foetal phase to the birth of the animal. In mammals this phase occurs in the uterus and the mother has more or less continuous influence, in egg laying species it occurs in the egg and the mother's influence mainly lays in the composition of the yolk and white components. During its time until birth, the animal encounters certain so-called windows of development: periods during which certain parts should develop. These windows often have a fixed time-frame. That means that the development has to take place during that time-frame. After the window closes again, development in that aspect is no longer possible. The development will depend on the genetics of the animal, but also on the environment. If that is not sufficient, the development will be sub-optimal.

Chapter 5.1.2 Events after birth

After birth or hatch, the influence of the mother continues, provided the offspring is allowed to be with the mother. However, the influence of the environment prior to birth is still present. So if there was shortage of something in the uterus or egg, then the development will have adjusted accordingly. After birth there may be compensatory growth, but if the organs were not optimally developed then that is a given fact that cannot be corrected anymore. The development continues after birth, also with a number of windows of development. Post-birth development does not only involve development of the physical part of the animal, but also of the emotional part. Many components of the character are developed during the first few weeks after birth. A sub-optimal early environment can have irreversible influence on the emotional development of the animal. Until weaning the environment of the animal is influenced by the maternal care, including food provision, but also by the other members of the environment, for example litter mates.

Chapter 5.1.3 Events after weaning

After weaning, the irreversible part of the influence of the early environment on the development of the animal remains. On top of that, the current environment of the individual animal will have its influence. For example, whether there is food and water and of what quality, what type of housing, how other animals influence its development, etc. The development of the animal continues until it has finished growing. This may be after it has matured in the sense of after the start of its reproductive life. The first oestrus or fertile sperm production, and sometimes the birth of the first offspring, occur before the animal has finished growing. Therefore, the term maturity is a bit misleading in the figure. It often represents the age at first reproduction, but in this case it represents the end of the animal's development.

Chapter 5.1.4 Events after maturity

After maturity there is the rest of the animal's life. It will depend on the environmental quality how that will influence the animal's phenotype. However, the animal has finished development so many of the environmental influences will be reversible. It may start its reproductive life and create its own influences on its offspring. And because it was influenced by its parents itself, you can imagine that those influences may still be of influence on the development of the animal's offspring. In fact, these influences of the parents of the animal of the development of its offspring would be trans generational effects. This is a relatively new area of research and not a lot is known about the importance of trans generational effects. An example is that of food deprivation of the mother of the animal during pregnancy or before laying the egg. This will have influenced the development of the animal. And because the development was influenced, this in turn may influence the development of its offspring. For example, if the growth of the animal was restricted and it remained small as adult, this will have influenced the size of the uterus as well, and thus the environment of its developing offspring.

Chapter 5.1.5 Events after reproductive phase

After the reproductive phase, some domestic animals are allowed to retire. Also with its own environmental influences on top of its past experiences during life. In nature animals usually do not retire but die before they reach that age. Farm animals are usually culled even before the end of the reproductive period. Note that environmental influences early in life may influence the phenotype later in life. However, not ALL early influences have a lasting effect. Some of the influences will be reversible or of insignificant influence.

Chapter 5.2 Phenotype in a model

As we have seen, anything that has happened during the life history of the animal may have an influence on the current phenotype. For example, if you measure the height of an animal, then that is its phenotype. Genetics will play a role in the height because if the genes determine that the animals will stay small, it can eat as much as it likes but it will never grow big: *the genetic makeup sets the boundary*. However, if that same animal does not get fed properly, or gets ill, or lives in a very cold climate, or his mother was ill when she was pregnant of the animal, the animal will not even reach that boundary and stay smaller than an animal of the same genetic makeup that had the best circumstances for growth. These environmental influences are not always easy to pinpoint because they start so early (oocyte and sperm experience environmental influences as well), but also because it is not always clear what is of influence and what not.

Definition

*The **environment** can be defined as anything that influences the animal's performance that is not related to the genetic makeup of the animal, starting at the earliest possible moment in life, even before conception.*

In general you can consider the following basic model:

Definition

Phenotype = Genotype + Environment

Or

P = G + E

These symbols P and G and E are important to remember as they are very commonly used to describe phenotype, genotype, or 'environment'.

Chapter 5.3 Monogenic genetic variation

Genetic differences between animals are a result of differences in their DNA. If a trait is determined by only a single gene, such as for example in the case of having horns in cattle or not, then the phenotype depends on the combination of alleles for that single gene. Having horns only occurs in case of homozygosity for the recessive allele h. Being polled (having no horns) is a dominant trait, so both Hh and HH animals are polled, and only hh cattle have horns. Especially in dairy cattle the calves are de-horned when they are still very young. However, also in the Holstein-Friesian some animals are naturally polled because they carry the H-allele.

Some aspects of coat colour are also determined by only a single gene with two alleles. But that may result in more phenotypic variation than in the example of having horns or not in cattle. In chestnut horse, for example, horses with no dilution factor (DD) are chestnut (brown coat and brown mane and tail). But if they have a single dilution allele (Dd) they are palomino (lighter manes and tail, slightly lighter coat), and with a double dilution allele (dd) they are cremello (almost white in coat and mane and tail, and also lighter eyes). So this dilution gene results in 3 different phenotypes, whereas the polled gene only produced 2.

It may seem that monogenic traits are never influenced by the environment, so P = G. This is not always the case. An example of a gene that is influenced by the environment is the gene that causes phenylketonuria or PKU in humans. It is a rare recessively inherited metabolic

disorder. Only if two carriers have a child, that child can have PKU. In the Netherlands 1 in 18,000 children has it, but all are tested for it as newborn baby in the blood collected with the heel pinch. In people with PKU the enzym Phenylalaninehydroxylases is not present or not functioning, so that it cannot break down the amino acid Phenylalanine. This is thus accumulating in the blood and the spinal fluid, which causes damage to nerve cells, eventually resulting in brain damage. Patients with untreated PKU are usually mentally retarded with behaviour problems, and often suffer from skin diseases. The treatment is very basic: a life-long diet with very little protein and no intake of aspartaam as that contains phenylalanine. Patients do get amino-acid additives to prevent deficiencies. Thus, the expression of this monogenic traits can be influenced by the environment, by the diet in this example.

Chapter 5.4 Polygenic genetic variation

Many traits are determined by more than a single gene. Figure 2 shows an example of three genes that together determine skin colour. You can see that with three genes you can already form 64 different genotypes! But these 64 genotypes do not result in 64, but only 6 different phenotypes due to epistasis: The expression of the genes depend on the combination of alleles and these 64 genotypes result in only 6 different grades of expression. If you would plot the frequency at which these genotypes occur per phenotype then you get this bell-shape curve. The intermediate phenotype occurs most frequent and both extremes the least frequent. The more genes are involved in expression of a trait, the more the frequency plot will resemble a smooth bell-shape. This bell-shape is a very common shape when you look at frequency distributions of traits with discrete values.

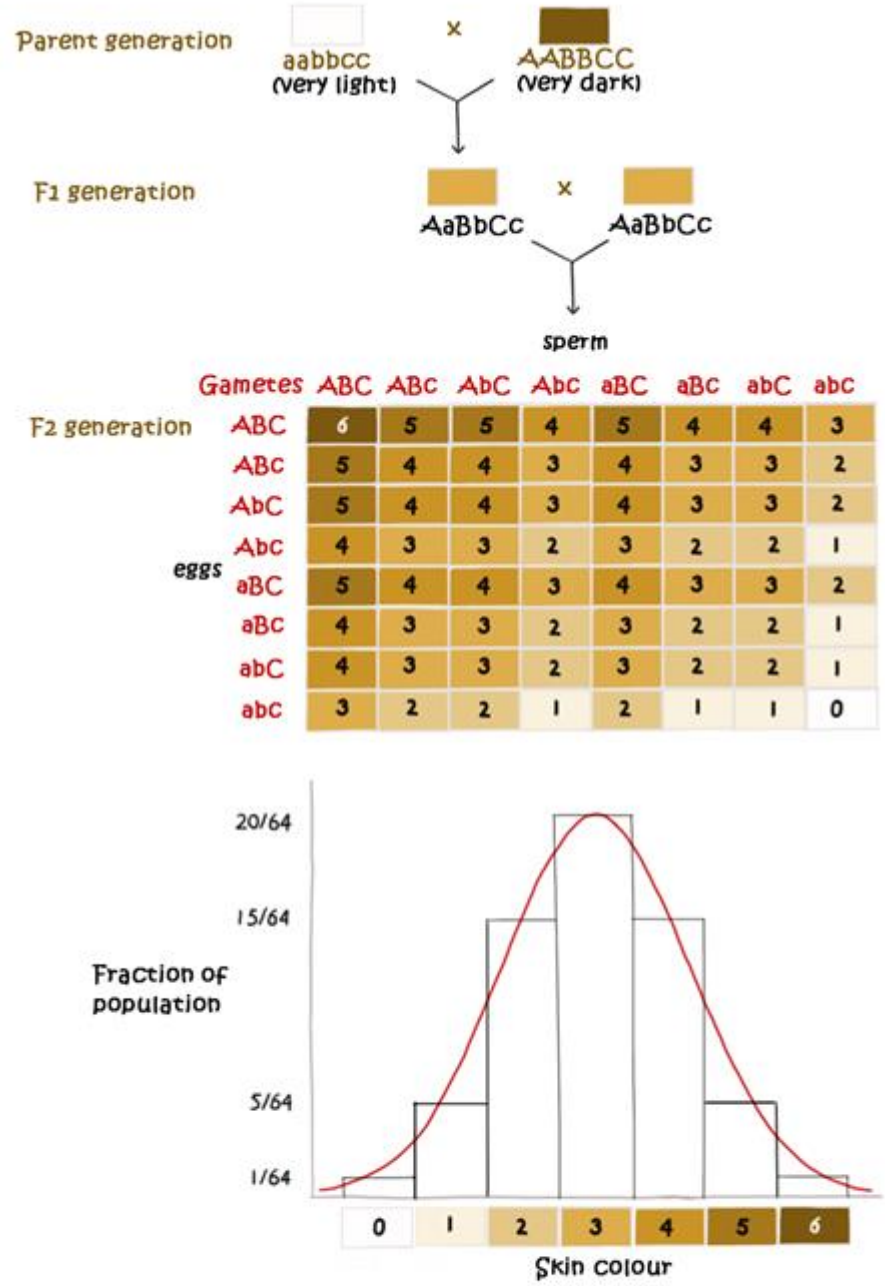


Figure 2. Example of the variation that can be obtained for a trait with only three genes involved: 64 genotypes are possible, resulting in 7 different phenotypes.

In animal breeding a general assumption is that traits are determined by an infinite number of genes, each with a very small effect, so that the bell-shape is very smooth. It is also assumed that this infinite number of gene effects are all additive. The model underlying this assumption is called the *infinitesimal model*.

Definition

*The **infinitesimal model** assumes that all traits are determined by an infinite number of genes, each with an infinitely small effect. This assumption results in a smooth bell-shaped distribution that can be described by the **Normal Distribution**. This distribution comes with a number of rules that animal breeding theory is built on.*

Recent research has shown that, even though of course the number of genes involved in expression of a trait is not infinite, the general assumption that many genes with small effect are involved very often is true. This is very convenient because this bell-shape fits the *Normal Distribution.* The normal distribution comes with a set of statistical rules that make it easier to make predictions. And that is what we want to do in animal breeding: predict genetic potential of animals and predict how the next generation will improve if we decide to use a certain proportion of the animals as parents. More on that in the chapter about ranking the animals.

Chapter 5.5 Variance components

The variation in a population can be quantified using statistics related to the Normal Distribution and is expressed as a so-called variance component. In symbols this generally is indicated with a ^{2}. So the phenotypic variance is called ^2_P, the genetic variance is called ^2_G, and the environmental variance is called ^2_E. Our model of P = G + E is also applicable to variance components:

$$^2_P = {}^2_G + {}^2_E + 2\text{cov}_{G,E}$$

$$= {}^2_G + {}^2_E$$

The covariance between G and E is assumed to be 0. In other words: there is no dependency of the genotype on the environment, or vice versa. The genotype does not change if the environment changes. This assumption generally is justified as we usually only consider a single type of environment when estimating variance components. In the chapter about evaluation of the breeding program we will see that this is not always the case. But for now we will just follow the general assumption that there is no mutual dependency between genotype and environment.

Definition

*Variation in a population is expressed in a **variance component**. The symbol for a variance component is ^{2} and the subscript indicates what type of variance component it is: P, G, or E .*

For estimating these variance components, we make use of the fact that if a trait is heritable it would mean that brothers and sisters perform more alike than unrelated individuals. So we combine the phenotypic information on the animals with their genetic relationships (i.e. the pedigree), and then the only component we have no real information about is the environment. Of course we can identify certain components of the environment, like housing and nutrition. But because the influence of the environment already starts at conception, we cannot identify all components of the environment. And some components we are not even aware of, like the potential influence of the weather three weeks ago on the performance today. We can estimate ^2_E by subtracting ^2_G from ^2_P. So E = P – G. Because this is not a very accurate way of estimating the variance due to environmental influences, this variance component is called *error variance*, rather than environmental variance.

Definition

*The ^2_E is called the **error variance**. This includes the variance caused by environmental influences, but also by some other effects.*

Chapter 5.6 Simplify the genetic model

The G in our model P = G + E is quite complex as it has a number of underlying components. This can be modelled as:

Genotype = additive effect + dominance effect + epistatic effect

Or G = A + D + I

To start at the back: the *epistatic effects* indicate that there are genes interacting with each other. This is, for example, the case if one gene needs the product of another gene to come to expression, resulting in so-called gene-pathways. Expression of one gene thus depends on the allele combination in another gene. The *dominance effects* indicate that expression of the gene itself depends on the allele combination in that gene. Two recessive genes will result in a different expression from one recessive and one dominant allele. The *additive effects* indicate the effect of the gene without the dominance and epistatic effects. So irrespective of the allele combinations of the gene itself or of other genes. What remains are effects that you can add up.

⚠ Definitions

The genetic component consist of three underlying effects:

1. *The **epistatic effect**: interaction between genes*
2. *The **dominance effect**: interaction between alleles of the same gene*

3. *The **additive effect**: everything that is left over after correcting for the interacting effects*

In variance component terms the genetic variance can thus be written as

$^2_G = {}^2_A + {}^2_D + {}^2_I$

To be precise, this equation should be extended by "$+ 2cov_{A,D} + 2cov_{A,I} + 2cov_{D,I}$", however these co-variances are zero by definition and are, therefore, left out of the equation.

Chapter 5.7 Next generation: transmission model

The dominance and epistatic effects depend on allele combinations. They are broken in the gamete production and established again in the offspring, but it is unpredictable how. The additive effects, however, are predictable because they do not depend on specific combination of alleles.

To be able to predict the additive genetic effects, we need to develop another model that describes the transmission of the genetic potential from both parents to their offspring. To illustrate that, in figure 3 you see a family of rabbits. The parents each have two different copies of each gene, but they pass only 1 on to their offspring and you don't know which one. So for each gene there are two alleles per parents, and four different combinations of those alleles possible in the offspring.

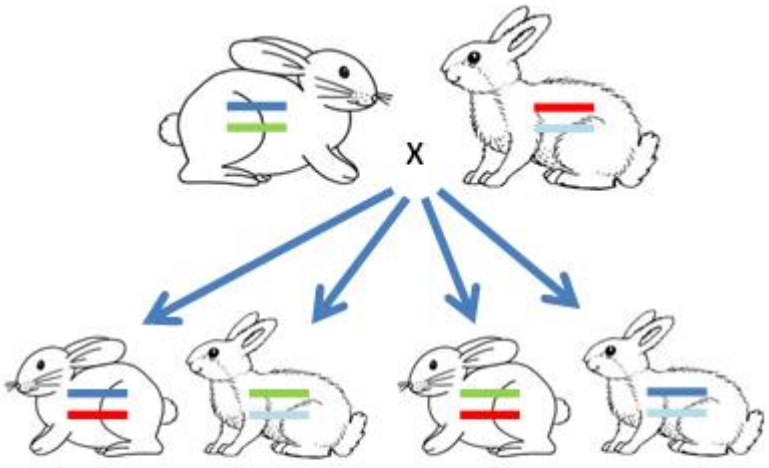


Figure 3. Family of rabbits illustrating that offspring receive half of their genes from each of their parents. Which half is a matter of chance,

Fact is that each animal receives half of its genes from its father (sire) and half from its mother (dam). But you can see from the figure that with this formula you cannot predict what will be the offspring performance as you will need to know WHICH half of the genes are passed on to the offspring. And that is a chance factor, also called the Mendelian Sampling term. So a part you know: half from sire and half from dam, and the other part is the Mendelian Sampling term (MS).

Definition

*The **Mendelian Sampling term** indicates the chance factor in distributing half the genetic material from each parent to their offspring*

Remember that in breeding we are only interested in the additive genetic effect (A), as half of that is inherited to the offspring. This is also called the true breeding value of an animal.

Definition

*The **true breeding value (A)** of an animal is its additive genetic component, half is which is inherited by the offspring*

In model terms you can write the breeding value of an offspring as

$A_{offspring} = ½\ A_{sire} + ½\ A_{dam} + MS$

If you would estimate variance components of all breeding values in a certain generation, then that is the same as the additive genetic variance. So the variance of A can be written as

$^2_A = Var\ (A) = var\ (½\ A_{sire}) + var\ (½\ A_{dam}) + var\ (MS)$

$= ½^2\ var\ (A_{sire}) + ½^2\ var(A_{dam}) + var\ (MS)$

$= ¼\ var(A_{sire}) + ¼\ var(A_{dam}) + var\ (MS)$

Under the infinitesimal model we assume that selection has no influence of the size of the genetic variance from one generation to the next. Therefore we assume that $var(A_{sire}) = var(A_{dam}) = var(A)$. This means that var(MS) must be equal to ½ var(A). This is quite a large component! That explains why they say that breeding is genetic gambling… Fortunately there are tools to decrease the chance factor in breeding. More about that in the chapter about ranking the animals.

Chapter 5.8 Heritability

Because in animal breeding we only make use of the prediction of A, and not of G, we should simplify the model of P = G + E to P = A + E. Note that this last E is larger than before because as we cannot estimate them, E also contains the D and I components. It now becomes more obvious why we call ^2_E the error variance: it contains more than only the effect of the environment.

Footnote: animal breeders tend to be a bit sloppy in the way they use terms. If they talk about P = G + E, they mean P = A + E, unless they specifically tell otherwise. Also, if they mention ^2_G then they mean ^2_A, unless they specifically tell otherwise.

The additive genetic effect is the part of the genetic component that is passed on from both parents to their offspring. In other words: the additive genetic effect is heritable. To indicate how heritable a certain trait is, a parameter is defined that indicates what proportion of the variation you observe (the phenotypic variance) is determined by (additive) genetic differences between animals (the additive genetic variance). This parameter is called the *heritability* and is indicated with the symbol h^2.

Definition

The heritability (h^2) indicates what proportion of the total phenotypic variation is due to genetic variation among individuals. In formula: $h^2 = {}^2_A / {}^2_P$

Boundaries between 0 and 1!

It is possible to estimate the heritability for a trait in a population if phenotypes and genetic relationships (pedigree) is available. A h^2 of 0.3 indicates that 30% of the variation you observe in your phenotypes is due to additive genetic differences between the animals. If ALL phenotypic differences are due to genetic differences, then the h^2 will be 1.0. Larger than 1.0 by definition is not possible. Similarly, if the differences between animals are NOT determined by their genetics, then the h^2 = 0.0. Smaller than 0.0 by definition is not possible.

Restrictions to estimates of the heritability

The estimated heritability is always specific for a trait, but also for a particular population in a particular environment. This has two important reasons. First, the influence of the environment will, of course, depend on the environment. Second, as we have seen in the example about genetic variation in human hair colour, genetic variation for a trait may vary between populations.

Definition

*A **heritability** always is estimated for a specific population in a specific environment because it reflects the genetic variation for a trait in that specific population relative to the phenotypic population*

If phenotypes are recorded in more than one environment, but for the same population, then there may be a third reason for difference in heritability. It could very well be that prerequisites for performance with respect to the trait under consideration may vary between environments. As a consequence, different genotypes may be superior in each of the environments under consideration. For example, if you consider the global Holstein-Friesian cattle population as a single population, then you will compare milk production levels from the Netherlands to those in Bangladesh. It is easy to realise that that may not be fair. It requires different qualities to be a top producer in the Netherlands from those required in Bangladesh. So the genetic variation will be different because partly different genes are required for being a good producer. The environmental variation will also be different because the circumstances are so different. Therefore, you should always estimate the heritability for the trait under selection in your specific population and in a specific environment. However, if somebody else already estimated a heritability for a population very similar to yours that was kept in an environment very similar to yours, then it is fairly safe to assume that the heritabilities will be similar too.

Table 1. Examples of heritabilities for a number of traits in some populations and species.

Animal Species and trait	heritability	Animal Species and trait	heritability
Dairy cattle		**Laying hens**	
Milk yield (kg)	0.36	Age at first egg	0.51
Body condition score	0.22	Egg production (egg/d)	0.22
Somatic cell score	0.15	Egg weight	0.60
Horses		**Sheep**	
Free movement	0.34	Clean fleece weight	0.47
Rideability	0.29	Fibre diameter	0.45
Osteochondrosis	0.23	Daily gain 30 to 90 days	0.52

Pigs		**Dogs**	
Daily gain (g/d)	0.25	Temperament	0.20
Litter size	0.15	Hip dysplasia	0.34
Feed conversion ratio	0.35	Litter size	0.30
Fish (Salmon, trout)			
Survival	0.05		
Body length	0.10		
Body weight	0.20		

Chapter 5.9 Simply estimating the heritability: parent-offspring regression

There are methods to accurately estimate the variance components, corrected for a number of systematic effects. These variance components can be used to calculate the heritability. Potential issue is that for accurate estimates of the variance components a rather large number of records (animals with observations and pedigree) are required. If you only have a limited number of observations, or you don't have good pedigree on the animals, there is a 'quick and dirty' way to get an impression of the size of the heritability: the parent-offspring regression. Parents pass half of their genes on to the offspring. If the trait you are considering is determined by genetics only, you would expect a regression coefficient of 1 if you would plot the average performance of both parents (also called the mid-parent) on the x-axe against the performance of the offspring on the y-axe. If the trait is influenced by the environment to some extent, but also by genetics, then you expect a regression coefficient smaller than 1, but larger than 0. This regression coefficient is an indicator of how much parents and offspring are alike. And the assumption is that the only factor that makes them alike is their common genetic background. In other words: the regression coefficient reflects the heritability. In some situations you will not have observations on both parents, but only on one parent. For example in case of a trait that only comes to expression in either males or females. In that case the regression coefficient does not reflect the complete heritability, but only half of it.

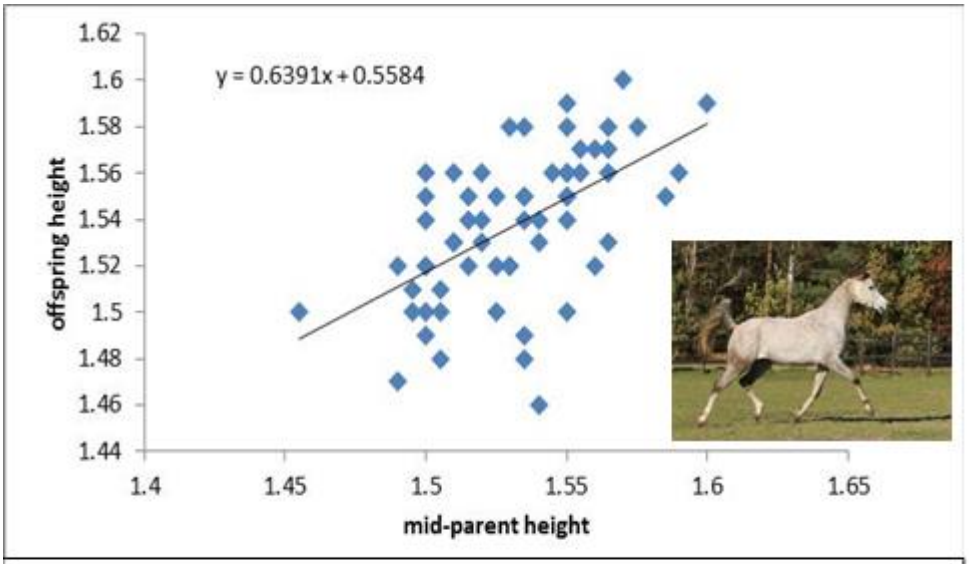


Figure 4. Height at withers of offspring plotted against that in their mid-parents. The regression between the two indicates how heritable the trait (height at withers) in this population is.

In Figure 4 you see an example of regression of the height at the withers of a number of Arabian horses on the average of their parents. The estimated regression coefficient, so the heritability, is 0.64. The estimated intercept of 0.56 suggests that the parents are systematically larger than the offspring. That could be a sign of a change in environment between both generations. This may be the case if the data was collected on one stud farm. However, it could also be a matter of inaccuracy of measuring the height at withers. It has no value in explaining the results.

It is important to keep in mind that this is not an accurate way of determining the heritability. If some families, for example, were kept in an optimal environment and others were kept in a poor environment, this will have influenced the results and will have an increasing effect on the regression coefficient, and thus the heritability. Similarly, if parents are kept in environments of very different quality to that of their offspring, there will be less relation between performance in parents and offspring and the regression will be low. Fortunately there are statistical techniques to take these systematic environmental influences into account.

Chapter 5.10 Misconceptions about the heritability

There are a number of misconceptions about the heritability. We will discuss a few below.

Misconception 1." A heritability of 0.40 indicates that 40% of the trait is determined by genetics"

This is a very common misconception and especially originates from a misunderstanding of the definition. A heritability of 0.40 indicates that 40% of all the phenotypic variation for that trait is due to variation in genotypes for that trait. This has a very different meaning from the definition that in each animal 40% of the expression of the trait is due to genes and the rest due to other influences.

Misconception 2. "A low heritability means that traits are not determined by genes"

A heritability that is larger than 0 always indicates that genes have an effect on the expression of the phenotype. The heritability is determined by the proportion of genetic variance relative to the phenotypic variance. A low heritability thus can indicate that the genetic variance is low. For example, number of fingers on a hand is very much genetically determined, but because by far most people have five fingers on each hand, the genetic variance is very low.

Misconception 3. " A low heritability means that genetic differences are small"

A low heritability not automatically indicates that the genetic variance is small. It can also mean that the error variance is large. And this can be caused by large influence of the environment, but also by inaccurate phenotype recording. For example: resistance to a certain infection will depend on the genetic potential to withstand that infection. The problem is how to measure that potential. If you go into the field and measure once whether a sheep is infected with nematodes, for example, then you will find the ones that are infected at that time. But in the other sheep you cannot distinguish between the ones that either have not been infected yet, already recovered, or are resistant to nematode infections. In other words: there is a lot of inaccuracy in your observations. Because you cannot assign the correct phenotype to each animal, this will result in a relatively large error variance, and thus in a low heritability. If you would improve the recording of nematode infection, for example by going into the field more often and/or improve the measurement methodology, then you obtain more accurate recording of the potential of the sheep to withstand nematode infection, and thus a more accurate estimate of genetic and environmental variance for this trait. The heritability can still remain low if there is not much genetic variation present, but at least it no longer is due to inaccurate phenotypes.

Misconception 4. " A heritability is a fixed value"

The heritability reflects the relative weight of the genetic variance component in the phenotypic variance in a specific population and based on observations that were taken on a specific moment in time. The size of the heritability depends on genetic variance in a population, but also on the influence of the environment and on the accuracy of the observations (see misconception 3). The genetic variance in one population may be (somewhat) different from that in another population. Especially if that other population is of a different breed. But also within a population the heritability can change in time. For example if the new set of phenotypic observations were collected using a more accurate recording method. Or if the housing system has changed since the last recording, so that the influence of the environment has changed. It, therefore, is wise to re-estimate the heritability at regular basis.

To summarise: The heritability indicates what proportion of the phenotypic variance is determined by the additive genetic variance, for a specific population in a specific environment. The specific population dictates the additive genetic variance, the specific environment influences the size of the environmental variance, as is the accuracy of recording of the phenotype so that differences between animals are revealed.

Chapter 5.11 Non-genetic influences: The variance due to a Common Environment

The environment that an animal has experienced in its life in general is very difficult to monitor in detail. But there are components that the animal has shared with others during their development that may have influenced them all in a similar way. The size of that influence can be estimated because we can compare individuals that shared that same common environment to others that had another environment in common. An example of a common environment is the mother, shared by animals in the same litter (think about pigs, dogs, sheep, rabbits, mice, etc). Those animals shared the same intrauterine environment, the same milk composition, approximately the same levels of milk production and maternal care. This shared early environment will have shaped these animals in a similar way. But also animals that were not born in a litter can have a common environment. For example, chicks that were hatched at the same time in the same incubator (or under the same hen), young animals in general that share their first indoor housing (cage or pen). Outdoor housing is much more variable and therefore much less of similar influence on all that shared the environment.

A common environment can, of course, also exist in the adult life of animals. However, in animal breeding we no longer call that a ‘common environment’. A common environment refers to the environment during the development of an animal and has irreversible consequences. If the common environment was abundant, the development of the animals sharing that environment will be according to their potential. However, if animals develop in a restricted environment, then their development will not be according to their potential and the consequences of this suboptimal development are irreversible. A restricted environment during adult life, however, mostly has reversible consequences.

Definition

*A **common environment** is an environment that is shared with others during the development of an animal and that, therefore, is expected to have the same influence on the devlopment of all animals that share a commen environment. The quality of an environment may have irreversible consequences if experienced during devolpment*

Chapter 5.11.1 Importance of common environment

Why would you be interested in the common environmental variance? Most important reason is that its size provides insight in the influence of the common environment on the variation in phenotypes observed. This common environment does not have to be shared at the moment of the recording of the phenotype. For example, age at maturation (first oestrus cycle) in female animals may be influenced by the common environment (e.g. shared litter) months or even years earlier. If the common environment was of good quality this may result in an earlier age at maturation.

The benefit of knowing that these joint shared environmental experiences exist is that being able to quantify the variation in effects of those experiences allows to estimate the heritability more accurately. This is because it is difficult to disentangle the effect of the common environment from the genetic component as closely related animals experience the same common environment. Taking the common environment into account when estimating the variance components helps to 'clean' the genetic variance from actual environmental influences that related animals have in common. And it also provides insight in the size of the influence of the early environment on the phenotype.

The phenotypic variance, taking into account the common environmental effect, can be written as:

$$^2_p = {}^2_G + {}^2_c + {}^2_E$$

We can define a common environmental factor, representing the proportion of common environmental variance relative to the total phenotypic variance. This is indicated with a c^2, analogue to the h^2 for the heritability.

Chapter 5.11.2 Examples of common environmental effects

In Table 2 you see an example of common environmental effects on a number of traits in sows of two different breeds. This common environmental effect represents the effect of being raised in the same litter until weaning. You see that the effect is largest on the leg score of the animals. Possibly this can have something to do with the milk composition of the sow that may have affected the growth and development of the bones. But this would be speculation. What the table also shows is the size of the effect of taking this common environment into account or

not when estimating the heritability for the traits. As explained earlier, this is because of the environmental effect of sharing the same litter. But it is also because it is difficult to disentangle the environmental effect of sharing that same litter from the fact that the animals in the litter were related. This makes it difficult to have accurate estimates of the variances due to the additive genetic and common environmental effects.

Table 2. Examples of heritabilities without (h^2) and with (h^{2*}) taking the common environmental effect (c^2) into account for two different breeds of pigs.

	h^2	h^{2*}	c^2
Landrace			
Leg score	0.06	**0.04**	0.10
Survival to 3rd parity	0.07	**0.05**	0.05
Survival to 5th parity	0.07	**0.05**	0.05
Length of productive life	0.09	**0.07**	0.05
Large white			
Daily gain (g/d)	0.09	**0.06**	0.11
Litter size	0.06	**0.05**	0.05
Feed conversion ratio	0.07	**0.05**	0.05
Body weight	0.08	**0.06**	0.06

Chapter 5.11.3 Special case of a common environmental effect: the maternal effect

The maternal effect can be a special case of the common environment. It is the effect of the environment that is defined by the mother. It starts already prior to birth, and continues for as long as the mother has an influence on the development of her offspring. In case of multiple offspring simultaneously, such as in a litter, the maternal effect can be an important part of the common environmental effect. But also the development of animals that are born as singletons are influenced by their mother. If the mother has multiple offspring, but not at the same time, it

is possible to estimate the effect of the specific maternal environment that is shared by all offspring. For example the size of the uterus, or the temperament of the mother, resulting in a specific type of maternal care.

A complicating factor is that the maternal effect has an environmental component, but also a genetic one! It will depend on the genetics of the mother what uterine environment she can create for the developing offspring. And also how wide her birth canal is, or how much milk she can produce and of what quality. So the maternal effect in fact is an environmental effect for the offspring, but depending on the genetics of the mother.

Definition

*The **maternal effect** is defined as the effect of the environment created by the mother on the evelopment of her offspring. The maternal effect is partly determined by the genetics of the mother.*

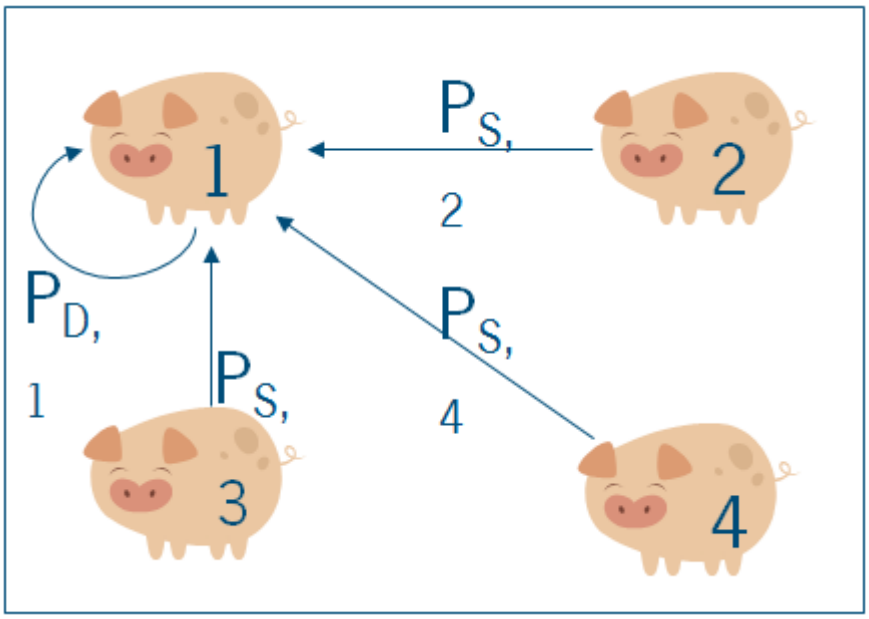
Note that this maternal effect is part of the breeding goal in a number of animal species. After all, having good mothering abilities can be an important component in the breeding program. Maternal effects are included in the breeding goals of, for example, dairy cattle (ease of birth), beef cattle and sheep (ease of birth, mothering ability), pigs and rabbits (the mothers are a very important component in number and quality of offspring).

Chapter 5.11.4 A special common environmental effect: (indirect) social genetic effect

The influence of genetics on the phenotype is even more complicated than you think. Until now we only talked about the genetic potential of an individual itself. With the maternal effect it became clear that it is not only the genetic potential of an animal itself that determines the development, the maternal genetic component also plays a role. However, if you think about it, it is not only your mother who has influenced your development. It were also, for example, your brothers and sisters, and the children in your school, some of whom were your friends and

some may have bullied you. In other words: many people in your surrounding have had an influence on who you are today. Part of this influence was due to what those other people had experienced themselves. But part of it also was because of the genetics of those people. It should now be clear why we also call this indirect effect the social genetic effect.

The phenotype of an animal is influenced by others. The others are part of the environment of the animal, where, just as with the maternal effect, the 'environment' has a genetic component: the genetics of the other animals. In other words: the phenotype of each animal is influenced by a direct genetic effect (its own genes) and its own environment, but also indirectly by phenotypic effects of animals in the surrounding of the animal. Just like the maternal effect, the social phenotype has a genetic and an environmental component. This is illustrated in the figure.



1.

It represents a pen with 4 pigs, with the outline of the direct and social effects on pig 1. Pig 1 has a phenotype due to its own genetics and own environment (P_D), but is also influenced by the social phenotypes (P_S) of its pen mates pig 2, 3, and 4. You can imagine that if these pen mates of pig 1 are quiet and friendly, it will perform much better than when the pen mates are bullying pig 1 around or preventing it from reaching the feed. Social effects may play a role anytime animals are kept in social structures such as pigs in a pen, but also chickens in a cage, horses or cows in a herd, sheep in a flock, etc.

Definition

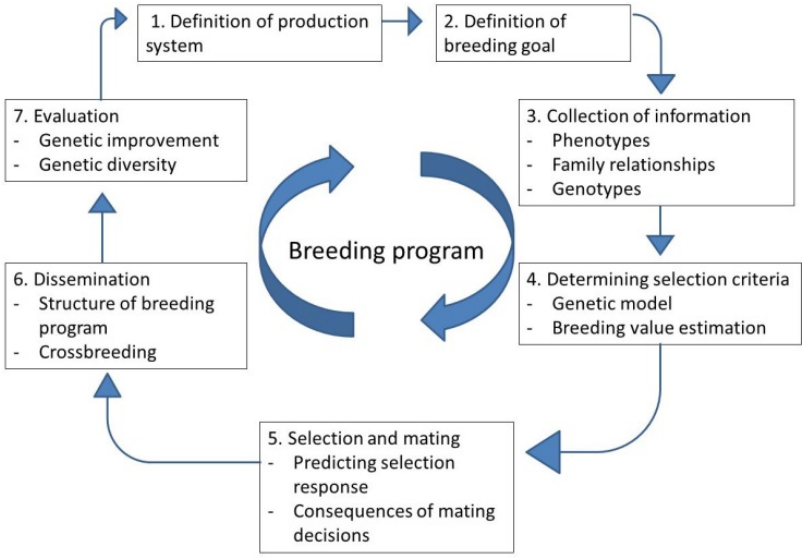
*The **indirect or social effect** describes the effect that the phenotypes of others have on the performance of an animal. Just like with the maternal effect, the social effects are phenotypes that consist of combinations of genetics and environments of the other animals.*

Chapter 5.12 Key issues of genetic models

1. The phenotype is determined by the genotype and the environment.
2. The environment consists of any influence that occurred between conception and the moment of phenotype recording.
3. In breeding we are only interested in the additive genetic effect because they are transmitted to the offspring.
4. The variation in phenotype can be expressed as phenotypic variance.
5. The phenotypic variance consists of additive genetic variance and error variance.
6. The error variance consists of variance due to environmental effects, but it is also a waste bin containing dominance and epistatic effects, errors in the phenotypic measurements, etc.
7. The breeding value of the offspring consists of half the breeding value of the sire and half of the dam.
8. The Mendelian sampling term indicates the part of the additive genetic component in the offspring that cannot be predicted: which half of the sire was transmitted to the offspring and which half of the dam?
9. The heritability indicates what proportion of the phenotypic variance is due to additive genetic variance in the population. It is indicated with $h^{2.}$
10. The common environmental variance is the variance due to the fact that animals shared a common environment during (part of) their development. For example they were raised in the same litter or pen. The proportion of the phenotypic variance that is due to common environment is indicated with c^2.
11. The maternal effect is the effect of the environment created by the mother on the development of the offspring. Part of this maternal effect can be due to genetics of the mother.
12. An indirect genetic or social genetic effect is the effect that others have on the performance of an individual.

Chapter 6: Genetic diversity and inbreeding

Defining a breeding goal and recording of phenotypes and pedigree are important aspects of setting up a breeding program. Accurate pedigree registration is essential for estimating breeding values, as has been indicated in the chapter about collecting information. The pedigree registration can also be used for monitoring the genetic relationship between animals. Knowing the relationhip between animals is very useful for managing genetic diversity in a population. Genetic diversity is a measure of genetic differences between animals in a population (i.e. genetic variation). To make sure that the breeding program remains viable in the future it is essential to monitor and maintain genetic diversity. Genetic diversity allows for selection of superior animals for breeding. If there is no genetic diversity, so if all animals would be genetically the same, selection will not result in an improvement in the next generation. In that case it is useless to set up a breeding program. Genetic diversity also has a clear link to inbreeding. Inbreeding results from mating of related individuals, and has a negative effect on health and reproduction.



In this chapter we are still collecting information (step 3), and we will look more into detail at the role of family relationships in genetic diversity. The chapter will be divided into two parts: first an introduction of the theory, and second a toolbox that can be used for evaluation of genetic diversity and for decision making related to selection and mating. Some of the applications of the tools will be subject of later chapters. To introduce the theory of genetic diversity we will take a top-down appraoch: first consider genetic diversity between populations, then move to within population, and finally look at genetic diversity within an individual. Then we will look at different mechanisms that influence genetic diversity, and discuss their role in animal breeding. We will look at inbreeding and its consequences. The toolbox in the second part of the chapter

will include tools to determine the genetic relationship between animals based on their pedigree, to determine the inbreeding coefficient of an individual animal, and to consider (influences on) the level and rate of inbreeding at population level. In the following chapters you will see that these tools are relevant in many of the steps in the breeding program.

Chapter 6.1: What is genetic diversity?

Diversity is another word for variation: the presence of differences among whatever you consider. Related to genetics, most obvious is the genetic diversity between populations. Different breeds, for example, have specific genetically determined characteristics. Think about differences in size, colour, but also in purposes such as beef versus to dairy cattle, or hunting versus guarding dogs. Genetic diversity also exists within a population, and is related to the genetic differences between animals in that population. It is possible, but very rare, that there is no genetic variation in a population. This occurs in populations that are fully inbred: animals are genetically completely identical to each other. But like said, this is a very rare situation that may occur in genetic lines of laboratory animals that are especially created for that purpose. The purpose of those populations is to provide animals that are as genetically equal as possible so that genetic differences are not a cause of variation in, for example, testing new medicines. A population of clones would be even better from the point of view of having genetically equal animals.

Definition

clone (animal) *is an individual that is genetically identical to another or a group of individuals that are genetically identical to each other*

Such population would have no genetic variation at all. However, in the Netherlands the use of clones is prohibited.

Definition

Genetic diversity *represents the presence of genetic differences between animals within species, both between and within populations*

The number of alleles that are present in a population is a measure of genetic diversity. The more alleles are present, the larger the genetic diversity is. The frequency at which these alleles occur in the population also have an influence on the size of genetic diversity. The more equal the allele frequencies are, the larger the diversity is. This principle is illustrated in Figure 1 for a gene with two alleles.

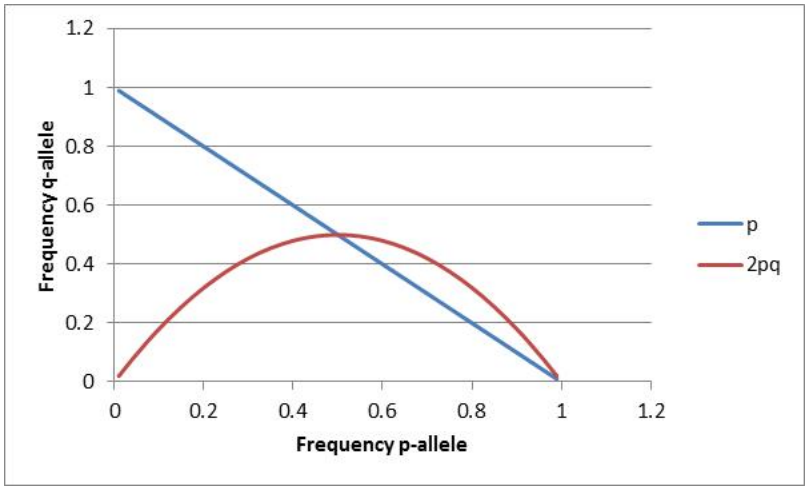


Figure 1. The relation between frequencies of two alleles (blue, straight line), and the consequence for the heterozygosity in the population (curved line). Maximum heterozygosity at p = q = 0.5.

If the frequency of the q-allele is 1, the frequency of the p-allele is 0, and vice versa. A high frequency of one allele always coincides with a low frequency of the other allele. The frequency of heterozygotes, calculated as 2pq, depends on both frequencies. The maximum allele frequency is achieved when both alleles have a frequency that is as high as possible, given the other frequency. And that is when both alleles are at equal frequency. For genes with more alleles the principle is the same: maximum heterozygosity at equal allele frequencies. Genetic diversity depends on the presence of a large number of alleles, but also on the frequency of those alleles in the population. Within an animal you can define genetic diversity as whether an animal is homozygous or heterozygous for a certain gene or parts of the genome.

Chapter 6.2: Forces that influence genetic diversity

There are a number of forces that influence genetic diversity. Some can be influenced by us, others occur by coincidence. At population level there are forces that increase genetic diversity, and forces that decrease it. Mutations are events that create new alleles. When it happens in germ cells it has an increasing effect on genetic diversity. Also migration may have an increasing effect, but only immigration (new animals moving into the population). Emigration (animals moving out of the population) usually has a decreasing effect on genetic diversity, especially when the population size is small. Also selection has a decreasing effect: only animals with a specific genetic make-up are allowed to breed, at the expense of others. This will have an effect on allele frequencies, away from the equal frequencies. A final force with a decreasing effect on genetic diversity is the force of coincidence (genetic drift), and related to that is inbreeding. Genetic drift cannot directly be influenced by our selection decisions. More explanation on genetic drift is in the next paragraph.

Forces that influence genetic diversity	Direction of change in genetic diversity
Genetic Drift and inbreeding	-
Selection	-
Migration	- or +
Mutation	+

Chapter 6.2.1: Loss of genetic diversity: genetic drift

Alleles can be lost from the population by coincidence. One reason for allele loss can be that not all animals mate and produce offspring, irrespective of selection decisions. Because animals that we selected for breeding may not all manage to produce offspring. Some may die unexpectedly or some are just never mated (e.g. in dogs and horses where owners of superior animals are not always interested to breed with them). Consequence of this not producing offspring despite being selection candidates does influence the allele frequencies in the offspring generation and alleles that were present at low frequency may be lost.

Mendelian sampling effect on genetic drift

Another reason for allele loss by coincidence has a clear relationship with Mendelian sampling. Even though animals are selected for breeding and mated and they do produce offspring, it is still uncertain which alleles are passed on to the offspring. And in case of multiple offspring: in what proportion they are passed on. Especially in smaller populations allele frequencies are sensitive to this type of genetic drift. Purely by chance allele frequencies change from generation to generation. The smaller the population, the larger the fluctuation in allele frequencies due to genetic drift. Even if all animals participate in breeding, alleles with low frequency run the risk of extinction just because they are not passed on to the offspring by chance. Strangely enough, this can also involve alleles with a positive effect that are under selection! Especially if these alleles have a dominant effect so that heterozygous animals also express the preferred phenotype. At low frequencies most animals that carry the desired allele will be heterozygous. Purely by chance they can all pass on the undesirable allele to their offspring. It sounds strange, but it really happens! Effects of genetic drift can have influences on allele frequencies that are larger than that of selection. Effect of genetic drift is especially important in small populations, where genotypes of individual animals have an influence on the allele frequencies, but it plays a role in populations of all sizes.

Example of genetic drift

In figure 2 is an example of how genetic drift could work. It shows how in a small duck population an allele could be lost without any directional selection. Purely by coincidence not all animals managed to reproduce, and also purely by coincidence, the carriers of the red allele not always passed it on to their offspring. Within four generations the red allele was lost from the population. Of course this is an example. The blue allele could also have been lost, or the frequencies may have just fluctuated a bit. The principle that allele frequencies change and homozygosity can increase purely by coincidence is very realistic and is called genetic drift.

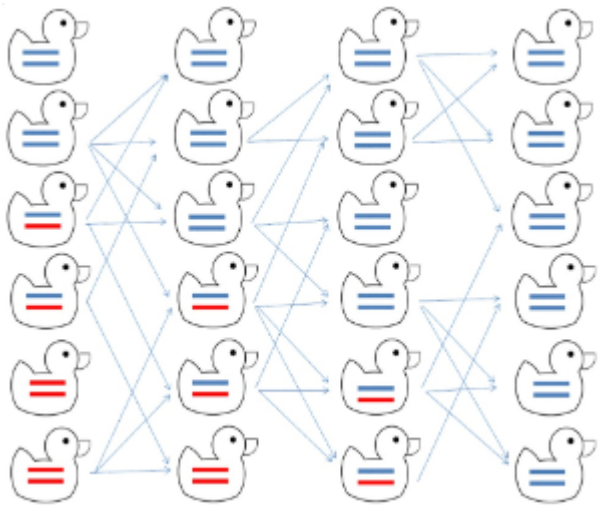


Figure 2. An example of genetic drift in four generations (columns) of a duck population. In generation 1 the red and blue alleles occur at equal frequencies. Duck 1 and 5 do not reproduce (e.g. they may have died earlier or did not find a mate). Duck 2 had the most offspring. Duck 2 could only pass allele blue on to the offspring, ducks 3 and 4 could pass on both colours, and duck 6 could only pass on the red allele. However, of the heterozygous ducks only duck 4 passed on the red allele and only once. In generation 2, the frequency of the blue allele has increased to 8/12 = 2/3. Again not all ducks managed to reproduce, and in generation 3 the frequency of the red allele decreased to 2/12 = 1/6. The red allele, by coincidence, was not passed on to generation 4. The population became homozygous blue.

Consequences of genetic drift

Genetic drift causes changes in allele frequencies, resulting in increase in one frequency at the expense of another. Because of that, it is more likely that animals become homozygous, especially for the most frequent allele. So a loss in genetic diversity at population level has consequences for genetic diversity at individual level. Animals become more alike. Even though they are not closely related through their pedigree, they become more closely related genetically. Genetic drift thus increases relatedness between animals and leads to fixation of alleles in a population.

Thus:

1. *The change in allele frequency by coincidence is called genetic drift.*

2. Coincidence relates to the variation in Mendelian sampling of which allele is passed on to the offspring, and in survival and reproductive success of animals.

3. The consequences of genetic drift on the allele frequencies in the next generation can be substantial, especially in smaller populations.

4. Genetic drift increases relatedness between animals.

Chapter 6.2.2: Loss of genetic diversity: selection

Selection favours some alleles over others. That is the intention of animal breeding! Obviously this has consequences for the allele frequencies in the next generation. Unlike genetic drift, selection has a *systematic* and directional force on the change in allele frequencies. The frequency of the favourable allele increases at the expense of the less favourable allele. Consequently, more animals will become homozygous for the favourable allele and genetic diversity decreases.

Exception to this rule are traits where heterozygous animals are favoured. Selection in that case has an increasing effect on genetic diversity. This, for example, is the case with sickle cell anaemia in humans. Sickle cell anaemia is a hereditary disease where red blood cells obtain a sickle shape and lose their flexibility, resulting in various life threatening complications. The disease is caused by a mutation in the haemoglobin gene. People that are homozygous usually die of the complications at fairly early age. In the sub-saharan region the frequency of the allele causing sickle cell anaemia is higher than elsewhere. Important reason for this is that people who are heterozygous for this gene are less affected by malaria infections. Malaria is an important cause of death in that part of the world and carrying a single allele thus has a clear selective advantage. Another example where the heterozygote is favoured is with the Dutch cattle breed the Witrik. The special type of colour marking only occur when they are heterozygous for the gene involved (see figure 3).

Figure 3: A heterozygous witrik cow with her homozygous witrik calf

Natural selection

Natural selection not only occurs in natural populations. It represents forces of selection that are not decided on by us. Natural selection acts on alleles that contribute to survival and reproductive success, also called *fitness*. For example, animals that have reduced potential to survive until adult age have a lower fitness (natural selective advantage) than animals that are very fit and healthy and will survive until late age. Also, animals with reduced reproductive capacity will have a lower fitness compared to very fertile animals. In domestic animal species the circumstances in which the animals live generally are quite controlled. Resistance to, for example, food shortage is less essential than it can be in natural populations. Still natural selection also works on domestic animal species. The environment the animals are kept in will require some degree of adaptation. For example, animals that are kept indoors need to be able to withstand the lack of sunlight, whereas animals that are kept outdoors will need to be able to withstand the variation in climate and a potentially larger infection pressure. Not being able to cope with the situation results in reduced fitness. If the selected animals require and receive assistance to get pregnant, for example, this will work against the mechanisms of natural selection.

Thus:

> *Both artificial and natural selection favour some alleles over others, resulting in an increase in homozygosity, and thus a decrease in genetic diversity.*
>
> *Exception is where selection favours heterozygous animals, where selection maintains or increases genetic diversity.*

Selection may lead to bottlenecks

Very strong natural selection, for example as a consequence of an outbreak of a very infectious and lethal disease, will result in a severe decrease in population size in combination with a change in allele frequency. Only the animals with some degree of resistance will have survived the infection, together with the lucky few that did not get infected. Those are the animals that need to build up the population again.Consequently, the allele frequency in future generations will depend on the allele frequency in that generation just after the *bottleneck*: the large decrease in population size. Some alleles that were carried by animals that were very susceptible to the disease will have been dramatically reduced in frequency or lost completely. An infamous example of a strong bottleneck is the outbreak of rinderpest in Africa in 1890. It swept through the entire continent and killed 80 to 90% of the indigenous cattle, buffalo, eland, giraffe, wildebeest, kudu, and antilopes (Mack, 1970). Because cattle and goats were killed, it had large social consequences and approximately one third of the human population of Ethiopia and two third of the Masaai of Kenya and Tanzania died of starvation. Rinderpest still causes problems at about 10 year intervals, but in 1890 veterinary support was very limited and the spread of the disease could not be stopped.

Figure 4. Cattle that died of rinderpest in Africa in 1890

Bottlenecks in domestic animal species can also occur because the specific breed lost its original purpose, but has gained a new purpose before the breed died out. Example of this is the Friesian horse that originally was used for work on the farm and lost its purpose with the introduction of the tractor. The breed was severely reduced in size severely before it started to become popular as a breed for leisure sports: both for harness driving and under saddle. Nowadays it is the second largest pure bred horse breed in the Netherlands (the Shetland ponies are the largest breed).

There are other examples of breeds that have lost their popularity, such some of the old Dutch cattle breeds. You can say they are still in the bottleneck because they have reduced in size dramatically, but there are no signs of recovery yet. However, others, such as the Brandrode and the Friesian Red, are showing signs of population growth.

Thus:

A bottleneck in a population refers to a severe decrease in size, followed by a recovery in size.

A bottleneck often has a large influence on the allele frequencies in the population, and thus on the genetic diversity.

Illustration: the role of bottle necks in dog breed formation

The formation of dog breeds is a nice example of what bottlenecks do with genetic variation. First there was the wolf with a lot of genetic variation. At some point in time the dog was formed from the wolf. How that happened exactly will always remain a bit of a mystery. One reasonable sounding theory is that in the times when people started to settle down an became farmers, they also started to accumulate waste. For some wolves that was an easier and more secure source of food that to go hunting. But you had to be very brave to come so close to humans to steal the waste. So only few wolfs managed, and they probably all had the genes for being brave and not so aggressive. That was bottleneck number one. These wolves slowly changed

appearance to dogs. This type of dogs you can still see in many parts of the world. They are the dogs that tend to live in the streets and don't belong to anyone, but still may live at somebody's yard. In return for their guarding, the owner of the yard will throw some food for them. But if you would ask that person he will tell you the dog does not belong to him. Our modern dog breeds are created from these village dogs. First people started to adopt the dogs and select for dogs that were good at hunting or guarding or helping with the sheep or cattle. Then people started to prefer certain looks. Slowly, populations of dogs started to look different from each other. Still there were no rules or regulations and you were free to breed with whatever dog you preferred. Around 1900 the first studbooks were created. Suddenly you were no longer allowed to breed outside the studbook if you wanted to call your dog a purebred. This is a second bottleneck, because most breeds were founded with only relatively few dogs. In modern dog breeds, as in other populations, mutations will increase genetic diversity. However, especially in small populations, inevitable inbreeding due to genetic drift, will decrease it.

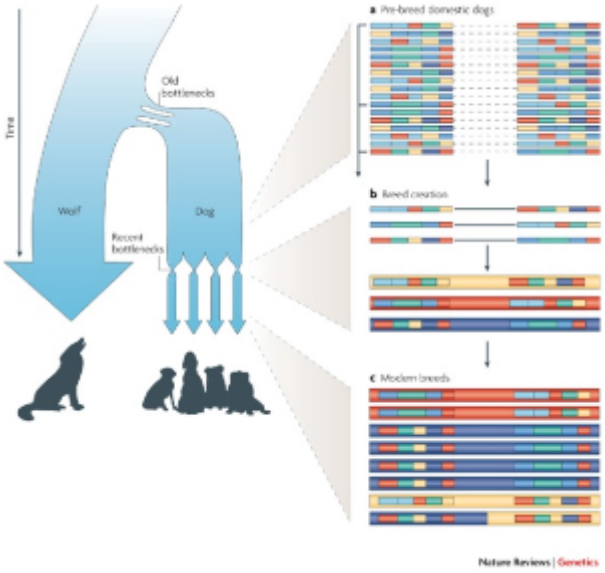


Figure 5. Genetic variation in modern dog breeds is smaller than in de wolf because of some strong bottlenecks that occurred. First when the dogs was formed, and second when our modern breeds were created.

Chapter 6.2.3: Diversity and migration

Migration is the process of leaving one population to join another. When an animal of another population (read breed) joins the population, this may involve introduction of new alleles. In that case the introduction of the new animal results in an increase in genetic diversity. The more

different the population of origin was, the larger the chance of introduction of new alleles, and thus increase in genetic diversity. The opposite of immigration is emigration: an animal leaving the population. This in general has a negligible effect on the genetic diversity, unless the population is either very small or the animal is carrying unique or very rare alleles.

Thus:

Immigration can increase genetic diversity in a population

Diversity and breed origin

Breed formation can roughly be divided into two classes: separation from a main population or created through cross breeding. Many breeds originate from a larger population. The new to be created breed usually already stood out from the main population because of some special feature. Often selection has already been performed for that special feature for some generations before separation from the main population becomes a fact. If the starting population of the new breed is of sufficient size, the genetic diversity in those separated breeds generally is not much lower than in the original population. Such new breed formation is a continuous process and new breeds are also being developed today. For example, the Dutch Kennel Club has special criteria that a group of dogs needs to live up to before it is acknowledged as a new breed. For example, they not only need to look similar, they also need to 'breed pure': the variation in offspring looks needs to be within the developed breed standard.

Instead of separation of from a larger population, breeds can also be developed on purpose by specific cross breeding. In commercial pig and poultry breeding the pure lines consist of multiple backgrounds and form so called hybrids: combinations of breeds that are crossed and the result mated again until the population has become a breed itself. In farm animals this generally is not called a breed, but a line or hybrid or synthetic. Because it was based on combination of breeds (populations), rather than individuals, the genetic diversity still is fairly large.

Creating a synthetic also occurs outside farm animal breeding. However, especially in case these hybrids were created by an enthusiastic breeder with a strong ideas, they tend to be based on only very few founder animals. These few animals are crossed and the resulting offspring are mated again, so that the genetic diversity of those breeds is very limited. The Saarloos Wolfhond is an example of such breed with very small number of founders.

Thus:

There are two main ways to form a breed

- *Separation from a main population*
- *Cross breeding, or creation of a hybrid*

Chapter 6.2.4: Increase in genetic diversity: mutation

A mutation is a change in DNA, and thus creates a new allele and increases genetic diversity. The frequency at which this occurs is small, but differs between species. In humans the mutation rate is estimated at 10^{-5} per gene per meiosis (generation). Mutations tend to occur especially on specific regions on the genome: mutation hotspots. A lot of mutations are deleterious. A dominant mutation often results in mortality (abortion) of the embryo. Recessive mutations hide in heterozygous state and spread through the population. They only come to expression in homozygous state. Some mutations are harmless, and some are positive. Positive mutations will be subject of selection and thus may increase in frequency more efficiently. Not all mutations result in a change in function of the gene, these are called silent mutations. Many of the SNP (single nucleotide polymorphism) are such silent mutations and are used as genetic markers.

Thus:

Mutation will increase genetic diversity

Chapter 6.3: Change in diversity: inbreeding

Inbreeding is the result of mating two related individuals. Related individuals genetically are more alike than non-related individuals because they share alleles. They share alleles because they have an ancestor in common. This common ancestor passed on the same alleles to multiple offspring, who passed them on again, so that eventually they ended up in both related animals. Mating those animals creates the chance that both of them pass the same alleles on to their offspring, resulting in homozygosity in the offspring. The level of inbreeding in an animal depends on the level of relatedness between its parents, and thus the probability that both will pass the same allele on to their offspring.

The level of inbreeding in an animal can be expressed in the inbreeding coefficient. The inbreeding coefficient indicates the *probability* that an individual receives the *same allele* from both his parents, because his *parents are related*. Inbreeding is the result of mating related individuals. The inbreeding coefficient takes values between 0 (0% or not inbred) and 1 (100% or fully inbred). Important to keep in mind is that inbreeding increases homozygosity (and decreases genetic diversity).

Definition:

Inbreeding level or inbreeding coefficient indicates the probability that an animal receives the same allele from both parents because they are related.

Chapter 6.4: Causes of inbreeding

There are two causes of inbreeding: inbreeding due to genetic drift and inbreeding due to non-random mating. Or in other words: inbreeding due to coincidence and inbreeding on purpose, inevitable and evitable inbreeding.

1. *Genetic drift* causes a loss in genetic diversity due to loss of alleles, which leads to an increase in homozygosity and this is also called inevitable inbreeding. Imagine that at some moment in history a mutation occurred. You may assume that this mutation only occurred in a single animal, as it is unlikely that exactly the same mutation also occurred in another animal. Animals that carry the allele today, therefore, must be related because they have that original animal as common ancestor. This is the case for all mutations, even if the animal in which the mutation occurred lived a VERY long time ago. Given the definition that inbreeding is the result of mating related individuals, animals that are homozygous for the allele must be inbred. Homozygosity in a population is an indication of size of allele frequencies. If all animals are homozygous, the other allele(s) is/are lost from the population. Inbreeding due to genetic drift results in a *permanent loss* of genetic diversity because alleles are lost for ever.
2. *Non-random mating* can cause inbreeding, but this is evitable. Mating closely related animals on purpose, like brother and sister or father and daughter matings, results in an increased probability that the offspring of the mating will receive the same allele from both parents. This results in increased homozygosity, and thus in inbreeding. However, this is a *temporary loss* of genetic diversity because if you would stop mating closely related animals on purpose, but use random mating instead, this cause of inbreeding would disappear.

Chapter 6.5: Inevitable inbreeding

Inbreeding due to genetic drift cannot be fully avoided because drift always occurs in a population. To get a feeling for why this is, consider the fact that everybody has two parents, which makes four grandparents, 16 great-grandparents, etc. The number of ancestral parents *n* generations ago thus becomes 2^n. This figure becomes very large in only a limited number of generations. In other words, your parents must be related, so you are inbred. Now it is easier to understand that drift occurs in all populations, but especially in smaller ones. The larger the population the smaller the chance that related individuals will mate by coincidence. Genetic diversity is at its largest when all animals would be heterozygous. Increased homozygosity means reduction in genetic diversity. Mating related individuals increases homozygosity, and thus decreases genetic diversity. Mating related individuals because you can’t avoid it results in

loss of alleles because of genetic drift. Mating related individuals on purpose also creates homozygous animals not necessarily results in loss of alleles because families are no longer mixed. Alleles do get fixed, but different alleles may get fixed in different families. At population level this does not have an influence on allele frequencies.

Chapter 6.6: Why is genetic diversity important?

It has now become clear what genetic diversity involves, and what its relation is with inbreeding. But why would we care about genetic diversity? There are three main reasons for that.

1. An important reason is that genetic diversity allows for flexibility in a population. If circumstances change, different genotypes may be more suitable and selection pressure will change. If the alleles required to adapt to the new circumstances are no longer present, or only at very low frequency, than adaptation of the population will be very difficult. Potentially with devastating consequences.
2. Inbreeding (increase in homozygosity) causes *inbreeding depression*. Inbred animals tend to be less healthy, live shorter, and have reduced reproductive capacity.
3. Related to that: reduced genetic diversity results in increased homozygosity, also of alleles that have deleterious effects. More inbred animals means more animals that suffer from monogenic recessive disorders.

Inbreeding depression

Inbreeding results in increased homozygosity, so also in more homozygous recessives. The negative result of that can be expressed as the performance of animals that are inbred, compared to animals that are not inbred, or as the change in phenotype given 1% increase in inbreeding level. For example, a study on Dutch Shetlandpony stallions showed a clear effect of inbreeding level on sperm quality features. Within the live fraction, inbreeding had a decreasing effect on the percentage normal sperm, and an increasing effect on the percentage with abnormal head (van Eldik et al., 2006. Theriogenology 65:1159-1170). There have been no studies on the effect of sperm quality on the fertilisation results, but it is likely there will be a certain minimum quality required for successful breeding.

Another example is related to inbreeding level in Holstein Friesian dairy cattle. Results in table 2 show the potential effect of a mating between a grandsire and his granddaughter. This is quite an extreme level of inbreeding, but not uncommon. The results show a negative effect of inbreeding on reproduction and milk production related traits. Inbred animals are older when they first calve, have shorter lactations, have longer periods between subsequent calvings, and produce less milk.

Table 2. Effect of 12.5% inbreeding (e.g. result of grandfather-granddaughter mating) on a number of traits in Holstein Friesian dairy cattle. (From Smith et al., 1998 J Dairy Sci 81:2729–2737)

Trait	**Losses from 12.5% inbreeding**
Lifetime days in milk (days)	-129
Age at first calving (days)	+5
First calving interval (days)	+3.3
First lactation total milk (kg)	-464
First lactation total fat (kg)	-15
First lactation total protein (kg)	-15

Chapter 6.7: Toolbox:relationships

Inbreeding is a result of mating related animals. If we know the relationship between animals, we would be able to predict and, to some extent, control inbreeding level in the next generation. If we know the pedigree of the animals it is possible to calculate the level of relatedness between animals, and thus the level of inbreeding of an individual animal. In the following part of the chapter we will look into more detail at how to perform these calculations.

Two animals are related when they have one (or more) ancestors in common. For example, you are related to your cousin because you have the same grandparents, they are your common ancestors. Because you have common ancestors, you and your cousin will have part of your alleles in common. The key-issue is that related individuals have alleles in common.

Thus:

Related individuals have similar alleles

In figure 6 you see two simple pedigrees. In pedigree 1, animals A and B are parents of animals C and D. In other words: A and B are common ancestors of C and D, who thus are full brother and sister. Animals C and D are mated and have offspring E. Because C and D are related, E is inbred. In pedigree 2, animals F and G are parents of H and I. Animals H and I are

mated to non-related animals and have offspring J and K. Those are mated and have offspring L. Animals J and K are related because they have ancestors F and G in common. So animal L is inbred, but less than animal E, because J and K are less related than C and D.

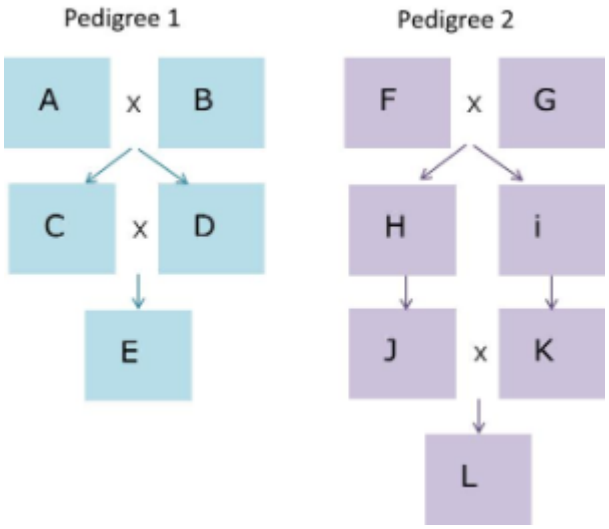


Figure 6. Examples of two simple pedigrees.

The more generations to the common ancestor, the less related two animals are. The less two mated animals are related, the smaller the probability that the same allele is passed on to their offspring, so the less the offspring is inbred.

Thus:

Inbreeding level decreases with decreased relatedness between parents

Chapter 6.7.1: Additive genetic relationship

The additive genetic relationship reflects what proportion of their DNA (alleles) two animals share because they have common ancestor(s). Additive genetic relationships can be calculated from the pedigree. Parents pass half of their alleles on to their offspring so the proportion of alleles that parents and offspring have in common is ½. In other words: the additive genetic relationship between a parent and his/her offspring is ½. The offspring received half of their alleles from the father and half from the mother, so their genome is a mixture of genes originating from both parents. When those offspring have offspring themselves, they again pass on half of the alleles to their offspring. What half of the alleles are passed on to the offspring is a random process (Mendelian sampling). Therefore, the proportion of alleles that a grandparent and his/her grand-offspring have in common is ½ (alleles passed on to their offspring) times ½ (alleles passed on from their offspring to their grandoffspring) is ¼.

Definition

The additive genetic relationship is an estimate of the proportion of alleles that two individuals have in common because they have one or more common ancestor(s)

Chapter 6.7.2: Calculation of additive relationships

There is one important calculation rule that you have to keep in mind when working with probabilities: if this AND that are both supposed to happen you should multiply the probabilities. Think about the situation where the same allele is passed on to the offspring AND to the grand-offspring. If this OR that is supposed to happen you should add the probabilities up. Think about the situation where allele 1 OR allele 2 of a gene is passed on to the offspring. It will become more clear with an example.

The additive genetic relationship (indicated with an 'a') between two individuals depends on the number of common ancestors and on the number of generations to each common ancestor. We will go through the steps to calculate the additive genetic relationship between two animals. Consider pedigree 2 in figure 6.

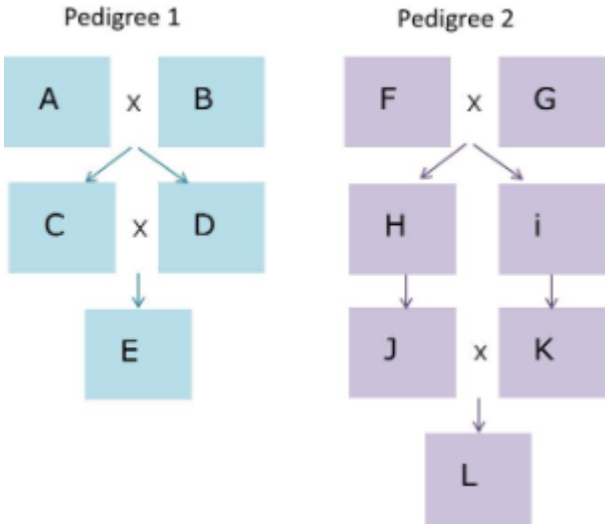


Figure 6: Example of two pedigrees

Question: What is the additive genetic relationship between animals J and K?

Answer in four steps:

Step 1: find the common ancestors.

The common ancestors of J and K are F and G.

Step 2: how many generations (meiosis) are there to each of the common ancestors?

Ancestor 1: F. The number of generations from J to F is 2, and from K to F is also 2.

Ancestor 2: G. The number of generations from J to G is 2, and from K to G is also 2.

Step 3: calculate the additive genetic relationship between the animals.

Through common ancestor 1: The probability that J and K have alleles in common that originate from common ancestor F is equal to the probability that the same alleles are passed on from F to H and from H to J and from F to I and from I to K. So we need to multiply these probabilities that are each equal to ½, resulting in ½ * ½ * ½ * ½ = $½^4$ = 0.0625

These same can be done for common ancestor 2: the probability that J and K have alleles in common that originate from common ancestor G is also equal to $½^4$ = 0.0625 .

These two probabilities can be added up because the animals are related because they share alleles from common ancestor 1 and/or from common ancestor 2. Both probabilities are independent of each other. The additive genetic relationship between J and K thus becomes 0.0625 + 0.0625 = 0.125 or $a_{J,K}$ = 0.125.

These steps for determining the additive genetic relationship can be described in a formula as:

$$a_{X,Y} = \sum_{i=1}^{m} \left(\frac{1}{2}\right)^{(n_i+p_i)}$$

Where X and Y are the animals that we want to know the additive genetic relationship of, m is the number of common ancestors, and for each common ancestor n is the number of generations from animal X to the common ancestor, and p the number generations between animal Y and the common ancestor. You see that per common ancestor the probabilities of sharing alleles are multiplied across generations, because all need to happen, and the probabilities are accumulated across common ancestors, because they are independent of each other.

Chapter 6.8: Additive genetic relationship using genomic information

The additive genetic relationship is estimated using the pedigree. But how accurate is that? We know that parents pass exactly half of their genetics on to their offspring. However, we also know that two full brothers may have received different alleles from the same parents. ON AVERAGE they share half of their genes. However, it could be half, it could be a bit more and it could be a bit less. If we only consider a single gene, two full brothers may even share no alleles at all due to Mendelian sampling. This is illustrated in Figure 7, where you see a pedigree of mice. The mother has alleles A and B, and the father has alleles C and D. Each of the four offspring received one allele from the father and one from the mother. In this example,

each of the mice received a different combination of these two alleles. On average you would expect that these mice share half of their genes as each of them got half from the same parents. But if you compare two of the four mice, they may share one allele or they share none. So instead of a = ½ these mice would have a = 0, at least for this specific gene.

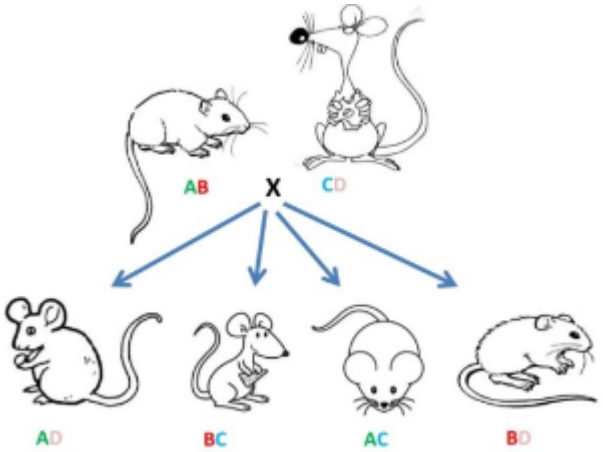


Figure 7. Example of differences between full sibs due to Mendelian sampling in a pedigree of mice.

Of animals course have not only a single, but very many genes. So on average, across all the genes, two full brothers share half of their genes. However, as illustrated in Figure 8, there is some variation around the average, some share a bit more than half and some a bit less. Likewise, half sibs on average share ¼ of their genes. Also here some share a bit more and some a bit less. The variation is half that in full sibs, because half sibs share only one parent instead of two.

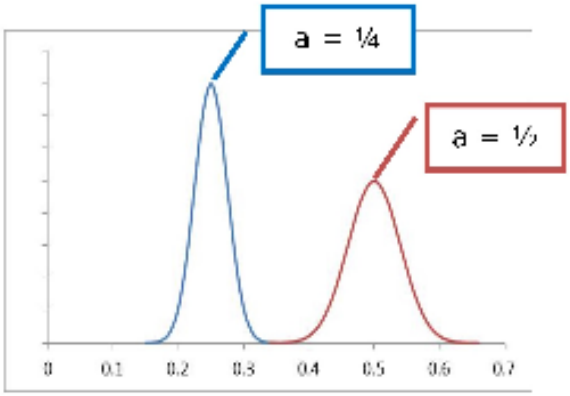


Figure 8. The distribution of true additive genetic relationships around one estimated based on pedigree of ¼ (half sib) or ½ (full sib).

Chapter 6.9: Realised additive relationship

In practice, the additive genetic relationships can be estimated using pedigree, or using genomic information. If genomic information is used (for example SNP markers), you can imagine that the more markers are genotyped per animal, the more accurate you can estimate how much of the genome two animals have in common. This is called the realised additive genetic relationship. In the future, when complete genome sequences are available for the animals, the exact additive genetic relationship can be determined. At this moment this is not the case yet because of financial limitations. In practical animal breeding the additive genetic relationship in most cases is estimated using pedigree relationships. In some cases, like with of genomic selection, the animals are genotyped for a large number of genetic markers, which allows for the more accurate estimate of the additive genetic relationship using the genomic information.

Chapter 6.10: Inbreeding coefficient and relationship

An animal is only inbred if its parents are related. The inbreeding level indicates the probability that an animal receives the same allele from both parents because they are related. In other words: it indicates the probability that an animal becomes homozygous for an allele that both parents share because they have a common ancestor. The inbreeding level of an individual animal is also called the inbreeding coefficient of that animal and can be calculated as:

$F_{animal} = ½ * a_{between\ parents}$

This simple formula indicates that it is easy to calculate the inbreeding coefficient of all animals in a population, as long as you know the additive genetic relationship between their parents. For example, the additive genetic relationship between a full brother and sister is 0.5. If they would be mated and have offspring, those offspring will be inbred. Their inbreeding coefficient would be ½ * 0.5 = 0.25. It means that for each locus the offspring will have a probability of 25% to be homozygous because its parents received the same alleles from their common ancestor. The more generations ago this common ancestor lived, the less the parents are related, so the lower the inbreeding coefficient.

Thus:

Important: An animal is inbred if, and only if, its parents are related!

$F_{animal} = ½ * a_{between\ parents}$

INTERMEZZO: Why is $F_{animal} = ½ * a_{between\ parents}$?

The inbreeding coefficient of an animal indicates the probability that the animal becomes homozygous because it inherits the same allele from both father and mother. For the animal to become homozygous, both parents need to have the same allele in the first place (= $a_{\text{between parents}}$). And then both parents need to pass it on to their offspring. That would result in

$F_{\text{animal}} = a_{\text{between parents}} * ½ * ½$.

This would be correct in haploid organisms. However, animals are diploid: they each have two alleles per locus. So parents have two chances of sharing an allele. Therefore, the probability that their offspring becomes homozygous, expressed as the inbreeding coefficient, becomes

$F_{\text{animal}} = 2 * a_{\text{between parents}} * ½ * ½ = ½ * a_{\text{between parents}}$

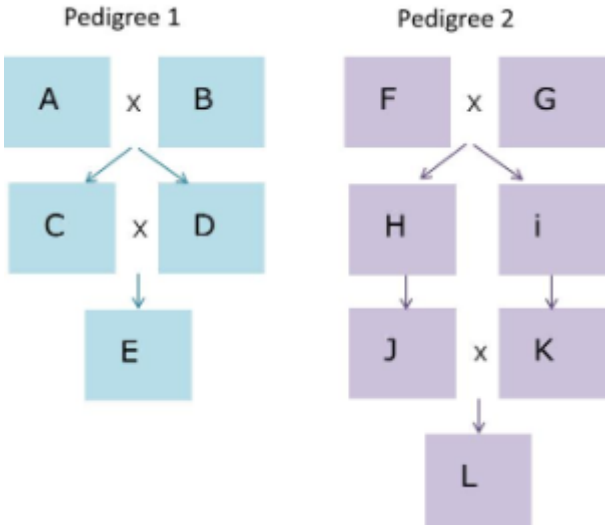
Chapter 6.11: Additive genetic relationship when the common ancestor is inbred

Animals that are inbred are homozygous for more loci (genes) than animals that are not inbred. Consequently, the chance that they pass the same allele on to two of their offspring is larger than in non-inbred animals. The more inbred an animal is, the larger the probability that it is homozygous, thus the larger the probability that it passes the same allele on to two of its offspring. An inbred common ancestor thus results in a higher additive genetic relationship between two animals. How much higher? Proportionally to how much more likely the common ancestor is passing the same allele on to the offspring, which is equal to the level of inbreeding.

Below you see the same formula for the additive genetic relationship between X and Y as before, but now including the inbreeding level (F) of the common ancestor. F indicates how much more likely it is that the same allele is passed on to two offspring.

$$a_{X,Y} = \sum_{i=1}^{m} \left(\frac{1}{2}\right)^{(n_i+p_i)} (1 + F_{w_i})$$

If we go back to our example with pedigree 2 we have no information on the inbreeding of the animals F and G. The additive relationship between J and K is 0.125 in case F and G are not inbred. The question is: how would it make a difference for the additive genetic relationship between J and K if G was inbred? Let's assume that the inbreeding coefficient of G = 0.23, so $F_G = 0.23$*. That implies that the probability that H and I receive the same allele from G is 23% larger than before. So the probability that the same allele is passed on to J and K is 23% larger than before. Before it was* $½^4 * 1 = 0.0625$ *and now it becomes* $½^4 * (1+0.23) = 0.0769$*. If F was not inbred, then the additive genetic relationship between J and K becomes* $½^4 + ½^4 * 1.23 = 0.0625 + 0.0769 = 0.139$*. Animals J and K are more related because common ancestor G was inbred.*



Importance of the number of generations for inbreeding calculations

Conclusions about whether an animal is inbred should always be made with a reference to the number of generations of pedigree that have been taken into account. For illustration, on the right you see the pedigree of Cirius. He is an Arabian horse, bred from Polish lines. Going back three generations in his pedigree you would conclude that Cirius is not inbred: his parents Eternal and Ciarka have no common ancestors.

However, if you would include two more generations, you see that Eternal and Ciarka are related through three common ancestors: Probat, Banat, and Palas. If we would calculate the additive genetic relationship between Eternal and Ciarka, we need to consider the contributions of Probat, Banat, and Palas. Based on this pedigree we would conclude that none of the common ancestors are inbred. Let's first look at the contribution of Probat to the additive genetic relationship between Eternal and Ciarka. Probate lived three generations earlier than Eternal, and four earlier than Ciarka. Eternal thus is more likely to share alleles with Probat than Ciarka is. The probability that both Eternal and Ciarka share alleles from Probat is $½^3 * ½^4 = ½^7 = 0.0078125$. Palas lived four generations earlier than both Eternal and Ciarka, resulting in a probability that they share alleles from Palas of $½^{(4+4)} = 0.00390625$. Banat also lived four generations earlier than both Eternal and Ciarka, so also resulting in a probability that they share alleles from Banat of 0.00390625. Combining these results for the three common ancestors results in an overall probability that Eternal and Ciarka share alleles from common ancestors of 0.0078125 + 0.0039065 + 0.00390625 = 0.015625 = 1.5625%. This is a low additive genetic relationship between Eternal and Ciarka, resulting in a very low inbreeding coefficient of 0.78% (or 0.00078125) for Cirius. Note that the more generations you would include, the higher the additive genetic relationship between Eternal and Ciarka will be, and the higher the inbreeding coefficient of Cirius.

Thus:

Additive genetic relationships and inbreeding coefficients are only informative given the number of generations of pedigree that are considered.

The advice is to use at least five generations of pedigree.

Chapter 6.12: Inbreeding at population level: the rate of inbreeding

In a population the level of inbreeding can be considered as the average inbreeding coefficient across all animals in the population at a certain point in time. As we have seen before, all animals in a population are related, even if it is only very slightly. Consequently, if we compare the average inbreeding coefficient across generations, it will always increase. This increase is called the *rate of inbreeding* or F.

The speed of increase will depend on the relatedness between the animals in the population. The more related the animals in a population are, the more their offspring will be inbred, and the larger the rate of inbreeding. The size of the rate of inbreeding provides an indication of

- the risk of inbreeding depression

- the decrease in genetic diversity (and thus room for adaptive capacity of the animals in response to a change in the environment).

The rate of inbreeding per generation can be calculated in retrospect from the average inbreeding in the current generation in comparison to that in the previous, relative to what remains to 100% inbred:

$F = (F_{t+1} – F_t)/(1 – F_t)$

For example, if the average inbreeding level in generation 5 is 3.5%, and in generation 6 it is 3.9%, then the rate of inbreeding is (0.039 – 0.035)/(1-0.035) = 0.0041 = 0.41%. If you want to do the calculations in % then you subtract F_t from 100 instead of 1.

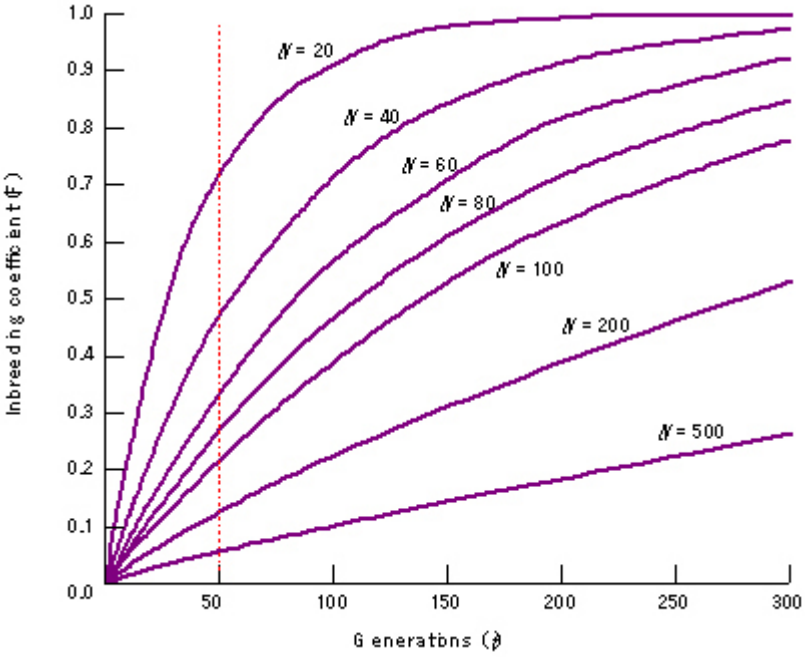


Figure 9. the relation between the non-linear change in inbreeding level across generations in a population and the population size. Reproduced, by permission, from McDonald, B. A. 2004. Reproductive/Mating systems. In: Population Genetics of Plant Pathogens. The Plant Health Instructor. doi:10.1094/PHI-A-2004-0524-01.

If only a single generation is considered then it is a good approximation to just consider the difference between both generations. However, for evaluations across multiple generations it is more accurate to divide by how much is left to complete inbreeding. This is because the increase in inbreeding level across generations is not linear. The maximum inbreeding level is 1 (fully inbred) and further increase is not possible in vertebrate animal populations. An increase in inbreeding level indicates an increase in the probability that an animal becomes homozygous for a locus on the genome. You can imagine that the more inbred an animal becomes, the smaller the probability that the remaining loci become homozygous in the offspring due to mating with a related animal. These related animals are inbred themselves, and thus homozygous for a proportion of their loci. The offspring will also be homozygous for those loci, but that is not an increase in homozygosity, because both parents already were homozygous. Of course the offspring is still very inbred, the level of inbreeding is increasing every generation until all animals are fully inbred. But the speed at which this occurs, relative to full inbreeding, is decreasing when the average inbreeding levels become high. In figure 89you see the relation between the average inbreeding level in a population across generations and population size, assuming random mating amongst the individuals. Obviously, the level of inbreeding is increasing the fastest in the smallest population. The dotted red line indicates that for these population sizes it is ok to assume a linear increase in inbreeding level for the first five

generations *since the population was founded!* In reality the populations usually already exist for many generations, so the initial value of the inbreeding coefficient in the first generation you consider will not be 0. Keeping that in mind, it is always wise, in all populations, to express the rate of inbreeding relative to how much is still left to full inbreeding.

Definition

The **rate of inbreeding** expresses the increase in average inbreeding level in a population from one generation to the next.

Because the increase in inbreeding is non-linear, the rate of inbreeding is expressed relative to how much the population is away from full inbreeding.

F or example, if the average inbreeding level in a population is 0.23, and it was 0.21 in the previous generation, then the rate of inbreeding is (0.23 – 0.21) / (1 – 0.21) = 0.0253. This is more than 0.23 – 0.21 = 0.02, illustrating that not taking the non-linear increase in inbreeding level into account may result in an underestimation of the situation. inbreeding.

Chapter 6.13: Rate of inbreeding and active population size

The rate of inbreeding depends on the population size. It is important to realise that this refers to the *active breeding population*, and not the total population size. To be able to predict the rate of inbreeding to the next generation, you need to know how many males and how many females participate in breeding. In figure 10 you see the relationship between population size and rate of inbreeding in a balanced breeding population, which means equal numbers of breeding males and females. The rate of inbreeding increases very rapidly with decreasing population size, when population size drops below 50. So when 25 males and 25 females are used as parents for the next generation, and they are mated at random.

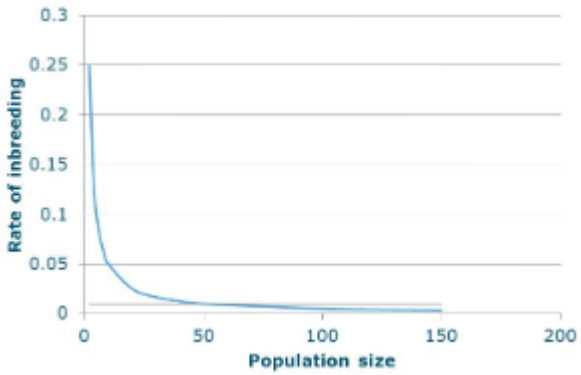


Figure 10. The relation between population size and rate of inbreeding for a balanced situation: equal numbers of breeding males and females.

The fact that the animals are mated at random is quite essential here. Because non-random mating can influence the rate of inbreeding. After all, an animal is inbred if, and only if, its parents are related. So to avoid inbreeding, you can try to mate unrelated animals. This works, but is only a temporary solution. Eventually all animals are related again and mating related animals can't be avoided anymore. The rate of inbreeding will become the same as when random mating would be applied. The consequences of inbreeding are postponed, not avoided. This is illustrated in figure 11b.

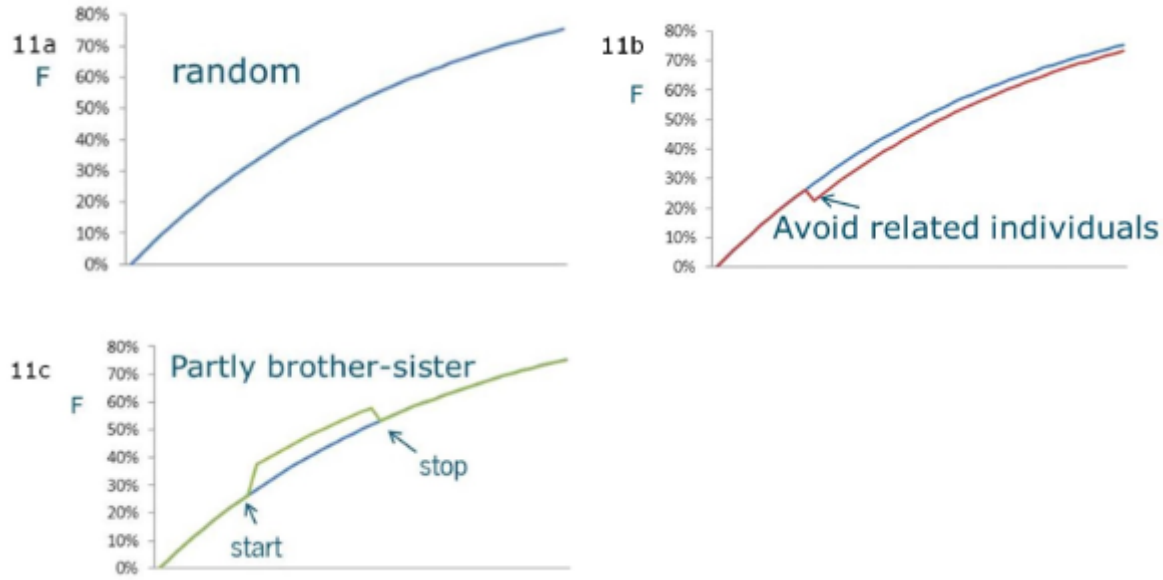


Figure 11. Figure 11a illustrates the increase in inbreeding across generations when random mating is applied. In figure 11b the effect of avoiding mating related animals is indicated. The inbreeding level drops immediately, but slowly approaches the situation under random mating again. In figure 11c the increase in inbreeding due to mating of related animals on purpose is illustrated. This effect can be reversed by returning to random mating again.

Mating systems and inbreeding

It is also possible to perform inbreeding on purpose. For example, by mating a father to his daughter and then also to his granddaughter. This is called *line breeding*. Some breeders perform line breeding because they want to fix the positive alleles of a certain superior male. In theory this is not a bad idea. Inbreeding increases homozygosity, so that also involves the positive alleles. Unfortunately there are two main reasons to avoid line breeding. First of all with line breeding animals become inbred, and thus homozygous, for many and eventually most of the alleles from that superior ancestor animal. However, not all of his alleles are equally desirable. He will be carrier of a number of recessive disorders that may become homozygous in the inbred descendants. The superior animal does not suffer from these disorders because

he is heterozygous, however his inbred, and thus homozygous, descendants will be. Second, line breeding (as any inbreeding) has a serious decreasing effect on the genetic variation in the population if applied by most breeders. This may have consequences for the ability of the population to adapt to a changing environment in the future. The increased inbreeding level because of mating decisions can be reversed by dropping the mating restrictions and introducing random mating. This is illustrated in figure 11c.

The examples in figure 11 show that the rate of inbreeding in the long run depends on the average additive genetic relationship between the animals in the population. You can try to avoid inbreeding or increase inbreeding by adjusting the mating strategies. However, in the end the rate of inbreeding will be determined by the average genetic relatedness between the animals.

Thus:

> *The rate of inbreeding is determined by the average relatedness between the animals in the population.*
>
> *It can be temporarily decreased by avoiding mating of related animals, or increased by mating related animals on purpose.*

Chapter 6.14: Predict the rate of inbreeding

Thus far we have evaluated the rate of inbreeding in retrospect. However, as it provides an indication of the expected increase in inbreeding depression, it would be nice to be able to predict the rate of inbreeding in future generations. Predicting the exact future is not possible, but you can make an approximation. A simple formula to get some idea about the effect of selection decisions with respect to the number of breeding animals is:

$$\Delta F = \frac{1}{8N_m} + \frac{1}{8N_f}$$

In other words, if you know the number of males and females that are used for breeding, you can predict what the rate of inbreeding will be. Of course the exact rate of inbreeding will depend on the genetic relationships between the animals, and that is not taken into account in this formula. And as we have seen in figures 5 and 6, the rate of inbreeding depends on your population size, more than on mating strategy. This formula will provide you with an approximation, assuming that these numbers represent your population size, and assuming absence of selection and no extremely small or large family sizes, relative to the sizes of the other families.

For example, let's assume a population of 3000 animals. But only 20 males and 300 females participate in breeding. Each female gets 10 offspring. What is the rate of inbreeding in this population?

*Answer: even though the population is 3000 animals large, only 320 of them participate in breeding: 20 males and 300 females. Filling that into the formula results in a predicted rate of inbreeding of 1/(8*20) + 1/(8*300) = 0.0067 = 0.67%.*

So 20 males and 300 females, that is 320 breeding animals, results in a rate of inbreeding of 0.67%. Would it matter how these 320 animals were divided across males and females? Try it yourself. What if you would use 160 males and 160 females? And what if you would use 2 males and 318 females? You will find out that the more skewed the proportion of breeding males to females is, the higher the rate of inbreeding.

How about population size? Would that matter? What if you would use only one male and one female for breeding? And how would the rate on inbreeding change if you would increase that number to 10 males and 10 females? Or 100 males and 100 females? You will find out that in very small breeding populations the rate of inbreeding can't be controlled by using equal numbers of males and females for breeding.

So far we have assumed that family size, so number of offspring in males and females, is equal for all families. In real life this is not the case. The rate of inbreeding is most influenced by the largest family, because they will have the largest proportion in the next generation(s). We also have assumed that the population size remains constant across generations. In real life this may not always be the case. Populations may decrease in size because of, for example, decrease in popularity or a disease outbreak. They may increase in size because of, for example, an increase in popularity, or a smaller mortality than anticipated.

Thus:

The rate of inbreeding depends on a combination of

- *the proportion of breeding males to females*
- *the number of breeding males and females*
- *variation in family size*
- *fluctuation in population size*

Example: rate of inbreeding in the Holstein Friesian

Even though the population of Holstein Friesian cattle is very large, so is the average additive genetic relationship among those animals. Because of the use of AI, the number of offspring per sire often is very large. And some sires are used much more for breeding than others, resulting in very unequal family size. Even though Holstein Friesians around the world are related to each other, there is some degree of subpopulations in the different countries. In

Denmark, for example, the rate of inbreeding is estimated at 1%, in Ireland at 0.7%, and in the USA at 1.3%. These are very high values, considering the fact that millions of cows are used for breeding, and hundreds of bulls are available. However, only a limited number are used for breeding, and on a very large number of cows. This is a clear example of the effect of very unequal family size (some bulls used much more than others) on the rate of inbreeding.

Chapter 6.15: What rate of inbreeding is acceptable?

The rate of inbreeding provides an indication of the expected increase in problems due to increased homozygosity. It provides an indication of the risk of getting into trouble. A risk means that it could be better or it could be even worse. As a rule of thumb, the FAO advises to restrict the rate of inbreeding to below 1%, preferably below 0.5%. A rate of inbreeding of 1% indicates that the increase in homozygosity will be 1% per generation. But there is no one who can predict what the consequences will be for the increase in problems due to homozygosity. Because not all genes have equal effect, and also not all alleles within a gene have equal effect.

The FAO advise is related to risk management, not to insurance. A rate of inbreeding above 1% increases the risk that the population in the long run will not survive.

Assuming balanced breeding, a rate of inbreeding of 1% means that at least 25 males and 25 females need to be used for breeding. Likewise, a rate of inbreeding of 0.5% means 50 males and 50 females. If non-balanced breeding is applied, the numbers of breeding males and females need to be chosen such that the rate of inbreeding is restricted to less than 1 (or 0.5)%. Depending on the animal species, this will be a realistic option or not. Consequences of number of breeding animals on the rate of inbreeding is essential to keep in mind when making breeding plans.

Thus:

- *FAO advise to limit the rate of inbreeding to 0.5 – 1%*
- *Essential in population management to use sufficient number of animals for breeding*

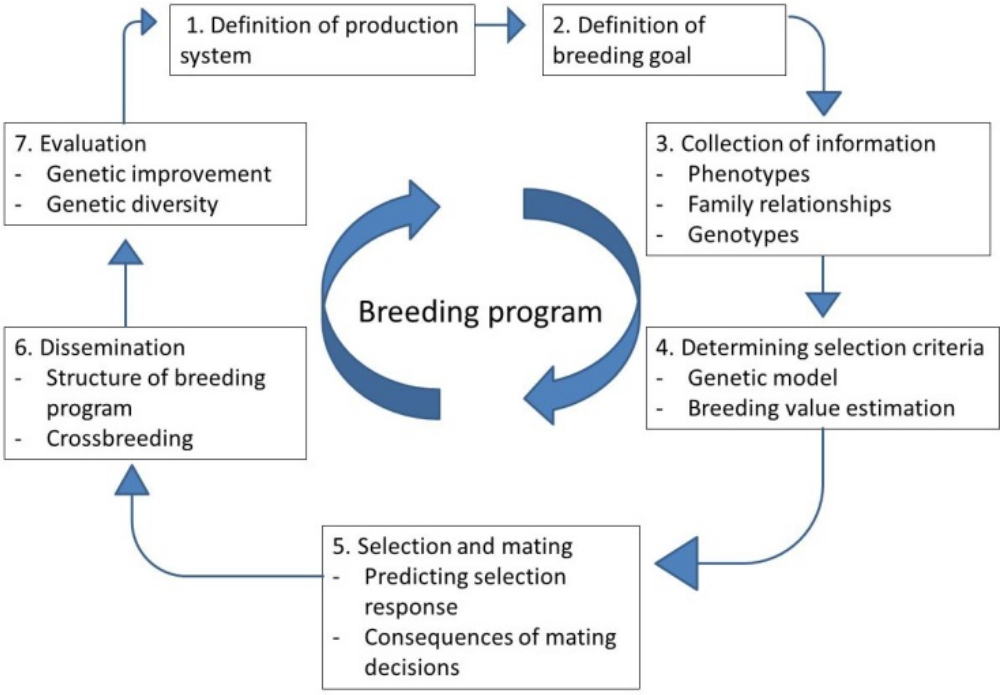
Chapter 6.16: Key issues on genetic diversity and inbreeding

1. Genetic diversity represents the presence of genetic differences within a species between animals, both between and within populations.

2. Genetic diversity is important to maintain flexibility in a population, to prevent inbreeding depression, to prevent increase in frequency of animals suffering from monogenetic recessive disorders.
3. Genetic diversity is influenced by genetic drift and inbreeding, selection, migration, and mutation.
4. Inbreeding indicates the probability that an animal receives the same allele from both parents because they are related.
5. The additive genetic relationship is an estimate of the proportion of alleles that two individuals have in common because they are related. The real additive genetic relationship may differ from the estimated one because of Mendelian sampling.
6. Inbreeding coefficients and additive genetic relationships are only informative given the number of generations of pedigree that are considered (minimal 5).
7. The rate of inbreeding is non-linear and expresses the increase in average inbreeding level in a population from one generation to the next.
8. The rate of inbreeding depends on a combination of the proportion of breeding males to females, the number of breeding males and females, the variation in family size, and the fluctuation in population size.
9. The FAO advises to restrict the rate of inbreeding to 0.5 to 1%

Chapter 7: Inheritance of monogenic traits

As we have learned before, traits of animals can have a monogenic or a polygenic background. We can select animals for polygenic traits based on estimated breeding values as explained in the previous chapter. Examples of monogenic traits are: the colour of the animals, dwarfism, extreme muscularity, malformations or severe health disturbances. The alleles determining the expression of monogenic traits may be dominant, intermediate or recessive. For the alleles of monogenic traits, the allele frequencies can be calculated. Characteristics of monogenic traits will be explained first, then breeding aspects of monogenic traits with desired positive effects will be discussed and finally the way monogenetic with undesired negative effects (genetic defects) can be handled in breeding programs will be outlined. The reason to pay a lot of attention to monogenetic traits is that many are determined by recessive /dominant alleles. Then it is impossible to distinguish heterozygotes from homozygote carriers of the dominant alleles. Then you cannot determine the genetic value of all animals for such monogenic traits. That is a problem, whether we deal with positive or negative aspects of the alleles.



Characteristics of monogenic traits

To repeat briefly: the loci for a monogenic trait may contain identical alleles: the animal is homozygous for this trait. It got an identical allele from the sire and from the dam. A loci also may contain two different alleles: the animal is heterozygous for this trait. It got an allele from the sire and a non-identical from the dam. Homozygous implies that all offspring of an

homozygous animal will get a similar allele from this animal when it is used for breeding and heterozygous implies that offspring gets either one of the two alleles with a chance of 50 %. In heterozygous animals we may have concern with intermediate inheritance (the value for the trait of the heterozygous lies just in between the values of the two homozygous forms) or with dominance / recessiveness. Then the heterozygous animal cannot be distinguished from one of the homozygous forms based on their phenotypes; they are not different.

Chapter 7.1 Calculation of allele frequencies

In a population of animals we can calculate the allele frequencies and the genotype frequencies for a monogenic trait. This is of value when you have an animal with known alleles for a certain monogenic trait and you want to calculate the chance for finding another animal (for mating) with a desired genotype for this trait.

Suppose a monogenic trait has two alleles Z and z. Then animals will have one out of the three genotypes possible: Z/Z, Z/z or z/z. Example: in a population of 630 animals we count 375 animals with the genotype Z/Z, 218 with the genotype Z/z and 37 with the genotype z/z. The frequency of the three genotypes in the population is: 375/630 = 0.595; 218/630 = 0.346 en 37 /630 = 0.059.

The allele frequencies can be calculated as follows: Z/Z animals do have 2 Z alleles; Z/z animals do have 1 Z allele and z/z do have 0 Z alleles. Thus the frequency of the Z allele is: 0.595 + 0.5 * 0.346 = 0.768. The Z/z animals do have 1 z allele and the z/z animals do have two z alleles. Thus the frequency of the z allele is: 0,5 * 0,346 + 0,059 = 0,232. In population genetics the frequency of alleles is noted as the frequency of p and the frequency of q. In this case p = 0.768 and q is equal to 0.232. The sum of p and q is always equal to 1 (in this example: 0,768 + 0,232).

Chapter 7.2 Hardy and Weinberg equilibrium

The frequencies of genotypes and of alleles have a relationship. When the allele frequencies are known, the genotype frequencies can be calculated. This relationship is known in population genetics as *the law of Hardy and Weinberg*. This law is valid in a population, stable (e.g. no migration) over generations. In such a stable population the Hardy and Weinberg equilibrium is at stake.

Definition

Hardy and Weinberg equilibrium *implies that in large populations with random mating among parents, and in the absence of selection, migration, mutation and random drift, the genotype and allele frequencies are constant (do not change from generation to generation) and the genotype frequencies can be calculated from the allele frequencies*

The Hardy and Weinberg equilibrium indicates the stability of a population over generations. The frequency of the genotype Z/Z = p*p = p^2 , because they originate from a combination of sperm cells and oocytes carrying the Z allele each with a frequency of p. The frequency of the genotype z/z = q*q = q^2 , because they originate from a combination of sperm cells and oocytes carrying the z allele each with a frequency of q. The genotype Z/z is created in two ways: by combination of a sperm cell with the Z allele (frequency p) and an oocyte with the z allele (frequency q) and by combination of a sperm cell with the z allele (frequency q) and an oocyte with the Z allele (frequency p). Therefore the frequency of the genotype Z/z = 2*p*q = 2pq.

Table 1: The frequencies of alleles and genotypes when animals are mated in a population in Hardy and Weinberg equilibrium

Oocyte Sperm cell	Allele Z	Frequency p	Allele z	Frequency q
Allele Z	ZZ		Zz	
Frequency P		P^2		pq

Allele z	Zz		zz	
Frequency p		pq		q^2

A llele and genotype frequencies in a population determine the proportion of animals with different phenotypes for monogenic traits. The effect of selection for desired monogenic traits or against undesired monogenic traits depends partly of the underlying allele frequencies. That is the reason to pay attention to allele and genotype frequencies here.

In populations of breeding animals mating of parents is often not ad random and parents are selected for breeding goal traits. Sometimes immigration of animals from other populations is at stake and in small populations random drift may cause shifts in allele frequencies. Therefore, in most breeding populations the presence of the Hardy and Weinberg equilibrium is doubtful. Nevertheless, it is of help to predict genotype frequencies from known allele frequencies for monogenic traits not under selection.

Chapter 7.3 Random effects in matings

When the genotypes of the parents are known, you know what genotypes you may expect in the offspring. These expectations are based on the Mendelian laws, but due to random effects, the real world may deviate from the expectations. E.g. when you mate two parent dogs, one with the genotype Z/Z and the other with Z/z you will get offspring where on average 50 % has the genotype Z/Z and 50 % the genotype Z/z. But in a litter of 4 you even may get 4 puppies with Z/Z and 0 with Z/z. The reason is that by the creation of an embryo each time the chance for the genotype Z/Z is as large as for the Z/z genotype (each 50 %). This is a consequence of the *Mendelian sampling*: each offspring gets 50 % of the genetic package of a parent, but you do not know which half. Then the chance for 4 Z/Z puppies in a litter of 4 is $0.5^4 = 0.0625$. Thus it is still possible to calculate the average expectation for the genotypes of the offspring when you mate two parents with known genotypes, but random effects cause deviations from the average expectation.

Chapter 7.4 Breeding aspects of genes with large (positive) effects

Molecular genetic research increases the number of genes with marked effect (major genes) on the quality of animal's products and affect fertility in several species.

Double muscling gene

For example, in many species (e.g. cattle, sheep, pigs, horses, dogs and humans) the myostatin gene is known. This locus has a recessive allele that in homozygous animals gives rise to the double muscle pattern. A breed where the fixation of this allele is aimed at is the Belgian White Blue cattle breed. The breed is famous for the heavy carcasses, the thick muscles and the high percentage of meat in the carcass. However, a high percentage of the females that are homozygous for the double muscle allele of the myostatin gene cannot give birth in a natural way. Those calves are born through caesarean section, leading to a severe ethical discussion in many countries. In other species the animals homozygous for the myostatin gene give similar problems when they give births and caesarean sections are required.

Milk protein genes

Another example of a gene with a known major effect occurs in dairy cattle. In dairy cattle a number of milk protein genes are known from which the alleles have a different effect on cheese yield. For example, beta-lacto globulin alleles (gene located at chromosome 11)have a marked effect on the efficiency of cheese production. Cows with the BB-allele are the favourite ones for cheese makers. The DGAT1 alleles (gene located at chromosome 14) influence fat percentage in milk and the milk fat composition. The K-allele increases fat and protein percentage and fat yield, while it decreases milk and protein yield. And very importantly, the milk fat composition from cows with K-alleles is different: they produce more fatty acids that are considered to be less favourable for human health.

Meat quality gene

In pigs the halothane gene (located at chromosome 6) is known to influence the stress susceptibility and the meat quality. Research is focused on chromosome 6 in pigs were genes are located that influence the androstenone production in boars. Androstenone is responsible for (an awful) boar taint in the meat of intact boars. Up to now boars are castrated to avoid this boar taint, but selecting against boar taint is a better approach for the welfare of the boars.

Fertility genes

In sheep several genes are described that have an influence on the litter size. An example is that of the Booroola gene present in the Australian Merino has a marked effect: the heterozygous carrier produces one lamb more and the homozygous carrier two lambs more per litter. This allele is now also present in the Dutch Texel breed by crossbreeding merino carrier rams of the allele with Texel ewes and backcrossing the crossbreds with Texel sheep.

Colour genes

In all species a lot of attention is paid to the inheritance of the colour of the coat. The colour of the coat is an important trait in the recognition of breeds. Breed associations often have strict rules for the required colour pattern. In companion animals and animals used for leisure purposes, breeders pay a lot of attention to the inheritance of colour and breed for special phenotypes. In the past, many genes and alleles are described that are involved in the colour inheritance. First, we start to describe a few genes that play a role in coat colour in ruminants and then we mention per species additional particulars.

Chapter 7.4.1 Colour genes in ruminants

In ruminants, as in all mammals, four genes determine the colour of the coat: the Extension, Agouti, Roan and the dilution gene.

The *extension gene* determines the pigment colour of an animal. The dominant allele E is responsible for the production of eumenalin in coat cells of black animals and the recessive allele e for phaeomelanin responsible the red colour in animals. A third allele is responsible for the wild phenotype (red with a light backline).

The *agouti* gene is expressed when at the E locus at least one wild allele is present. It is an example of epistatis: the genotype on the extension locus determines the expression of alleles on the agouti locus. The agouti alleles give a striping pattern of black and red.

A dominant allele at the roan locus is responsible for the presence of white and coloured hairs next to each other. It leads to a grey colour of the animals.

The dilution gene dilutes, when the dominant allele is homozygous, the base colour very much. In case of a black animal, the homozygous animals give light grey patterns. The heterozygous animals have dark grey patterns.

In cattle, sheep and goat white spots may be present caused by several known loci: the spotting, blaze, belting, colour sided and brockling loci. For further information see e.g. : The genetics of cattle, 1999. Editors R. Fries and A. Ruvinsky, and The genetics of sheep, 1997. Editors L.Piper and A. Ruvinsky.

Figure 1: bull with wild type phenotype

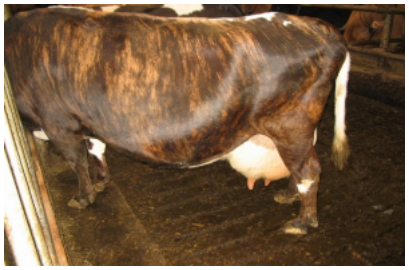


Figure 2: Icelandic cow with agouti pattern

Figure 3: Icelandic cow with the roan pattern

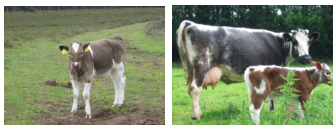


Figure 4: a white grey calf of the Danish Heath Breed imported in The Netherlands (left) and a Dutch “Witrik” (colour sided) cow with the dark grey pattern

Chapter 7.4.2 Colour genes in pigs

In pigs also the extension, agouti, roan and dilution gene determine the base colour. In addition a lot of genes are known, as in ruminants, that give cause to different colour patterns. A dominant allele at the roan locus is responsible for the complete white pigs we have at the moment in many commercial pig breeding programs. For further reading see e.g. The genetics of the pig, 1998. Editors: M.F. Rothschild and A. Ruvinsky.

Chapter 7.4.3 Colour genes in horses

Only a few different genes are responsible for the variation in coat color that we see in horses. As far as we know it concerns eleven different genes. In principle, all color variations can be traced back to two basic colors. Black is the first basic color and chestnut is the second one. The color bay, which is wide spread, in principle is a transformation of the color black. Because of the fact that bay is so abundantly present, it is also often called a basic color. The B of gene B is derived from the word ‘black’. A horse that is homozygous black, therefore is BB. The gene A is responsible for the colour of all bay horses. This A-allele is responsible for the transformation of the pigmentation of the color black (BB) into bay. Horses that have the genetic code BB (black) will express the bay color but their manes, tail and lower legs will remain black. A bay horse not only has the B allele, it also has the A allele. A homozygous bay therefore has the color code: AABB. A black horse does not have the A-allele (note capital A) because the dominant allele A is the allele that causes the transformation of pigmentation. A black horse, therefore must have the homozygous recessive code aa in order to ‘stay’ black. The story of the basic colors is not yet told completely now; for the black color, to come to expression, “permission” is needed from a third gene that is coded with the letter E. The recessive homozygote ee results in the color: chestnut. It all looks a bit complicated but we came simplify things a little. On the European continent (and also in America), the recessive allele b is hardly found. For simplification it is therefore acceptable to not consider the b-gene in color formulas.

This delivers the following color formulas:

Bay 1 = AAEE = homozygous bay	Zwart 1 = aaEE = homozygous black	Zwart 2 = aaEe
Bay 2 = AAEe	Bay 3 = AaEE	Bay 4 = AaEe
Chn 1 = aaee = homozygous chestnut	Chn 2 = Aaee	Chn 3 = AAee

Note that:

- A horse with the genetic code ee, always is a chestnut
- A black horse always has aa and at least one E
- A bay horse has at least one A and at least one E

It will be obvious that if two homozygous parents that have the same color, are mated, their off spring has the same color.

Some examples are:

Bay (AAEE) x black (aaEE) ® bay foal (AaEE)

Bay (AAEE) x chestnut (aaee) ® bay foal (AeEe)

black (aaEE) x chestnut (aaee) ® black foal (aaEe)

On top of the basic colors (B, A and E) some other genes can determine the color of a horse. One of these genes is the *g gene.* The letter G stands for gray. The GG or Gg genotypes overrule all other colors and cause the horse to become grey at later age. These horses are therefore not born as greys. There is a group of genes that dilute the basic colors. The C and C^{cr} are examples of such *dilution* genes. They are responsible for coat colors like Isabel and cremello. Another group of genes causes the colors to not be equally distributed; patches of different colors or white spots can be the result. Overo and Tobiano genes are examples of this group. For further reading see: The genetics of the horse, 2000. A.T. Bowling and A. Ruvinsky

Chapter 7.4.4 Coat colours in dogs

In dogs there is another allele at the extension locus which is dominant over the E allele and that is the E(m)allele, causing a black mask at the head. Further, there is a recessive e allele, when homozygous, causing bright, yellow or red coats. *Brown* is the locus for the dilution of eumelanin and is a recessive trait. When homozygous bb, the eumelanin will dilute to a liver, brown or chocolate colour. Nothing will happen when a B allele is present. The process of *dilution* is recessive epistatic depending on two alleles, D and d. Recessive epistatic means that when the locus is homozygous for the allele d, the dilution is epistatic, overpowering the base colour, causing black to become blue (slate grey). Dilution has only an effect on eumelanin, this differs from cattle and horses. DD and Dd have no effect. The *chinchilla* gene is able to dilute pheomelanin. CC or Cc are needed to form tyrosinase, which is an essential enzyme for the melanin production. When the recessive c is homozygous, the animal is unable to produce pigment and will be an albino. The c(ch), chinchilla allele fade the lighter brown

parts of Agouti and Brindle even more to almost white, but does not affect the black eumelanin. It can also cause the extended white phenotype, causing black animals to have a white phenotype with black nose, feet, eyes and lips. More information: The genetics of the dog, 2012. Editors: E.A.Ostrander and A. Ruvinsky.

Chapter 7.4.5 Colour genes in poultry

In poultry breeding, especially in the work of hobby breeders, a very broad spectrum of plumage colours is produced. The extension locus has a lot of alleles with different expression. It is also known that the main loci for plumage colour interact with each other: epistasis is often at stake. White plumage is preferred in the developed world because the skin of the carcasses is also white. In Asia black carcass skins are preferred. A commercial application of the inheritance of plumage colour is practiced in layer production where cocks are discarded. A sex linked locus for plumage colour gives the opportunity to distinguish at day 1 of live, whether a chick is a male or a female. In the male only one allele of the relevant locus is present and in the female two. Two alleles in the heterozygous (crossbred) female give another plumage pattern than one allele in the (crossbred) male. For further reading see e.g. http://kippenjungle.nl/Overzicht.htm#kippengenen this is a bilingual website (Dutch and English) that also provide knowledge on coat colour inheritance of other species.

Chapter 7.5 Breeding aspects of monogenic traits with negative effects

In animal breeding in all species and in all breeds genetic defects caused by mutations require continuous attention. Due to mutations during the meiosis each individual (even men!) carry mutations with negative effects. When such a mutated allele is dominant, it will have a visible or measurable effect in the carrier and the carrier is not viable or will not be used for breeding. The mutated allele will not be spread in the population: it will be purged. But when a mutated allele is recessive, it will not have an effect for the carrier and it will not be recognized. It will be spread in the population when the carrier is used extensively for breeding. The effect of the mutated allele will be recognized many generations later when, often by accident, two offspring of the carrier will be mated to each other. Then, with a chance of 25 % a homozygous carrier of the mutated allele is born. It might not be viable, malformed or suffer from a serious health problem early or later in life.

In animals a lot of malformations and malfunctions are known. Not all of them have a genetic origin. To determine the genetic background it is highly recommended to record all malformations and malfunctions and to analyse regularly the population for their frequencies. An increase analysed, might show that a certain parent, or ancestor in past generations, gives more than one offspring with a defect. That is the first sign that the deviation is a genetic defect. Per species many genetic defects are known:

Table 2: recorded genetic disorders, genetic defects per species in a global survey

Genetic defects: http://omia.angis.org.au/home/

THE UNIVERSITY OF SYDNEY

Summary

	dog	cattle	cat	sheep	pig	horse	chicken	goat	rabbit	Japanese quail	golden hamster	Other	TOTAL
Total traits/disorders	580	397	302	214	214	206	190	72	58	41	40	463	2777
Mendelian trait/disorder	223	145	75	88	45	40	114	13	28	31	28	146	976
Mendelian trait/disorder; key mutation known	154	78	40	32	18	29	36	9	7	9	3	58	473
Potential models for human disease	296	142	165	82	77	108	41	28	37	11	14	229	1230

The homozygous animals suffering from monogenic genetic defects pop up in populations were in the past a sire got a large number of offspring, much larger than the other sires used at the same time. This fact is illustrated in the figure 5 below:

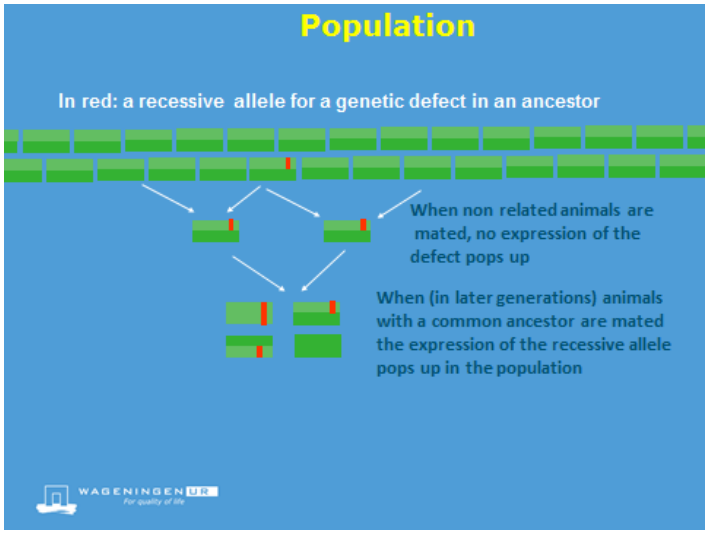


Figure 5: a recessive allele in an ancestor becomes a homozygous sufferer from a genetic defects in a later generation when animals are mated related to this ancestor

From this fact, it can be learned that in animal breeding it is not wise to create a large number of offspring from a selected parent. In well controlled populations this can be easily realized, but in less controlled populations this is very difficult. A directive is that each parent in a generation should produce less than 5 % of the animals in the subsequent population.

Thus, as all individuals carry alleles for genetic defects, it is impossible to discard all these alleles in a population. And once an allele is spread in a population it is possible to reduce its frequency to a low level, but without genetic markers it can never be discarded.

Even with a very low allele frequency (e.g. 0.05) for the genetic defect, you still have (Hardy and Weinberg expectation: 2pq = 2*0.95*0.05 = 0.10) a lot (10% in this example) of carriers in the population.

Chapter 7.6 Testing parents for monogenic traits when no genetic markers are available

In molecular genetics many loci for qualitative (monogenic) traits are detected and alleles are characterized. The latter has resulted in a lot of genetic markers available for testing the presence of alleles for traits from animals that are considered for breeding. But often a genetic marker is not available. Then you have to perform “test” matings.

Suppose you want to mate your brown dog, a bitch, with a black male and you want to know what will be the chance on black puppies. You know that the black allele is dominant over the brown allele. It means that you want to know the chance that a random chosen black male is heterozygous. In the breed the allele frequency of the brown allele is 0.1 and from the black allele 0.9. This means that (assuming that the population is in Hardy Weinberg equilibrium) the fraction of homozygous black animals is 0.9^2 = 0.81 and the fraction of heterozygous animals is 2*0.9*0.1 = 0.18. So the chance for a black male to be heterozygous is 0.18 / (0.81+0.18) = 0.18. Roughly 1 out of 5 black males is heterozygous and will give 50 % black and 50 % brown puppies when mated to your brown bitch.

In the Dutch Groninger Blaarkop cattle breed the frequency of the dominant blaze allele is 0.95. The recessive homozygous allele does result in an unwanted spotted animal. How can we know with an accuracy of 95 % that a sire is homozygous for the blaze allele? Thus with an uncertainty of 5 % you want to get the answer. The best way is to cross the sire with Friesian spotted cows. A homozygous sire will get 100 % offspring with the blaze pattern in the cross with spotted animals. How many test matings have to be performed? Each calf born of a testcross has a chance of 0.5 to be spotted when the sire is heterozygous. With two calves the chance is 0.5*0.5 = 0.25. With 5 calves the chance is 0.5^5 = 0.0325 and lower than 0.05. With less than 0.05 uncertainty you have to perform 5 successful test matings. Thus the number of offspring needed depends on the allele frequencies in the population and the required accuracy or the uncertainty of the test.

Chapter 7.7 The value of genetic markers in testing parents for genetic defects

Nowadays for a lot of monogenic traits, genetic markers are available. Many of these genetic markers are not directly testing the presence of the allele responsible for the monogenic trait. But these markers are establishing the presence of a variable piece of non-functional DNA, an allele at a locus close to the locus of the functional allele. These genetic markers are extremely valuable to detect heterozygous animals, heterozygous for recessive alleles.

But these markers always have the risk that, through recombination effects between the marker and the locus of the monogenic traits in the meiosis, the relation between the marker and the functional allele is disrupted.

Thus, once a genetic marker closely located to the recessive functional allele on a chromosome is detected, the relation between the marker and the allele should be checked regularly in future generations. The closer the marker is situated to the functional allele, the lower the chance for recombination and the higher the accuracy of the genetic marker.

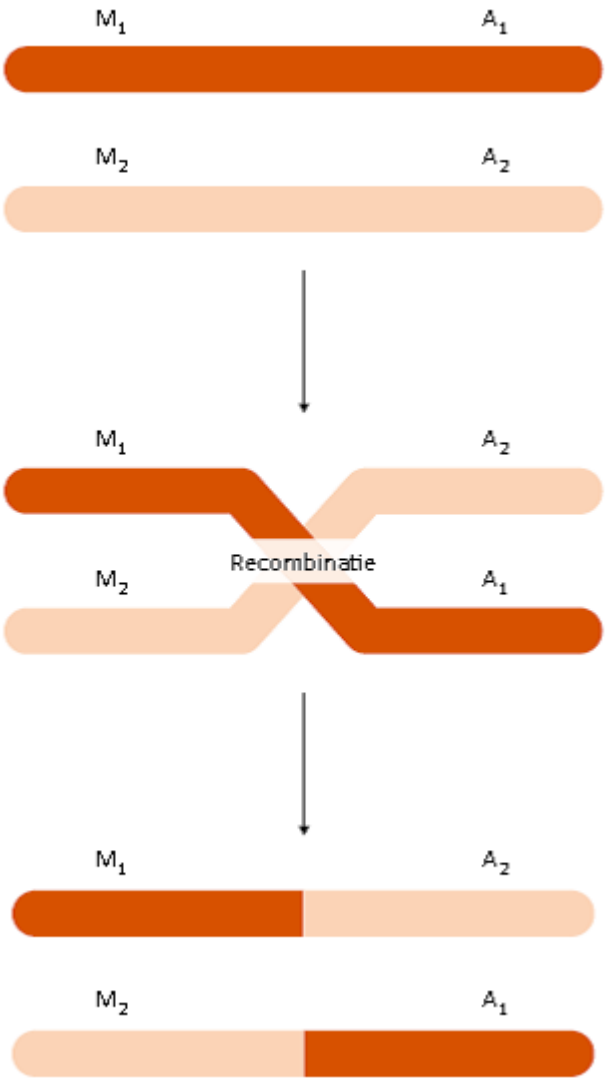


Figure 6: Recombination disrupts the relationship between marker M_1 and allele A_1 and between marker M_2 and allele A_2

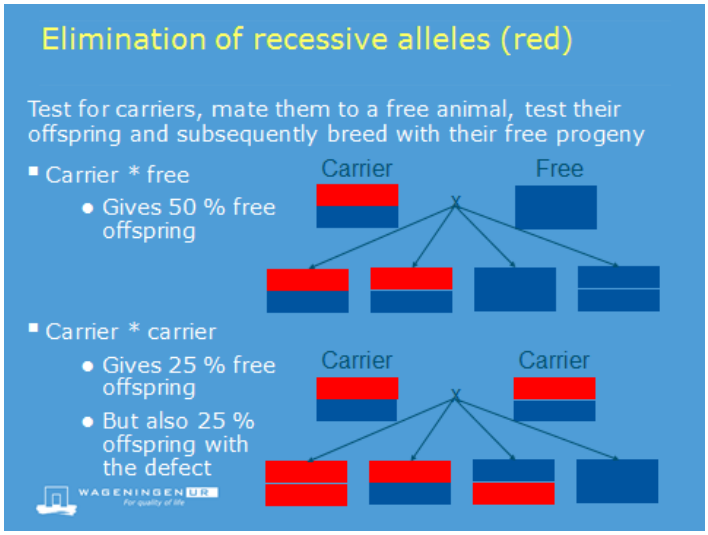
Chapter 7.8 Elimination of recessive alleles for genetic defects with genetic markers

As we have seen before, many genetic defects in animals can be traced back to monogenic recessive traits. A lot of emphasis is given by molecular geneticists to develop genetic markers for such unwanted recessive alleles responsible for genetic defects. The markers give the opportunity to detect the heterozygous carriers of the defect which do not show the defect themselves.

As soon as the first genetic markers became available breed associations started to test animals for these markers and started to eliminate heterozygote animals from breeding. However, it is a hard job to eliminate totally the “bad” allele in the population. It requires testing all the animals and that is often too costly.

Sometimes, e.g. in Dutch sheep breeds tested for scrapie susceptibility, it reveals that elimination of the homozygous and heterozygous sires with the “susceptible” alleles from breeding, would result in limiting the number of rams for breeding drastically. The approved remaining rams with the desired alleles were highly related to each other and would give rise to a huge increase in inbreeding in later generations.

Thus, as in case of scrapie, the allele frequencies for genetic defects might be rather high. Then the best way is to test all the animals considered for breeding with the genetic marker. The use of heterozygous carriers is restricted: they are only mated to free animals. The offspring from such matings consists for 50 % out of carriers that are subsequently excluded from the breeding program. Only the free offspring will be allowed to participate in the breeding program. This way of elimination assures that a broad genetic variation in the population is maintained and the average additive genetic relationship does not increase due to the elimination of the defect.This best strategy to eliminate a recessive allele for a genetic defect in a population with the help of genetic markers is outlined below in figure 7: strategies to eliminate carriers of unwanted recessive alleles by testing and subsequently using only free offspring



Sometimes, when the genetic defect is spread heavily, as was the case for scrapie susceptibility in a few sheep breeds in The Netherlands, the necessity can exist to mate even carriers to carriers. Then 25 % of the offspring is free and can be used for further breeding, but also 25 % of the offspring will be affected by the defect. The latter is problematic when the defect has serious implications for the health and welfare of the affected animal. Then, it should not be applied.

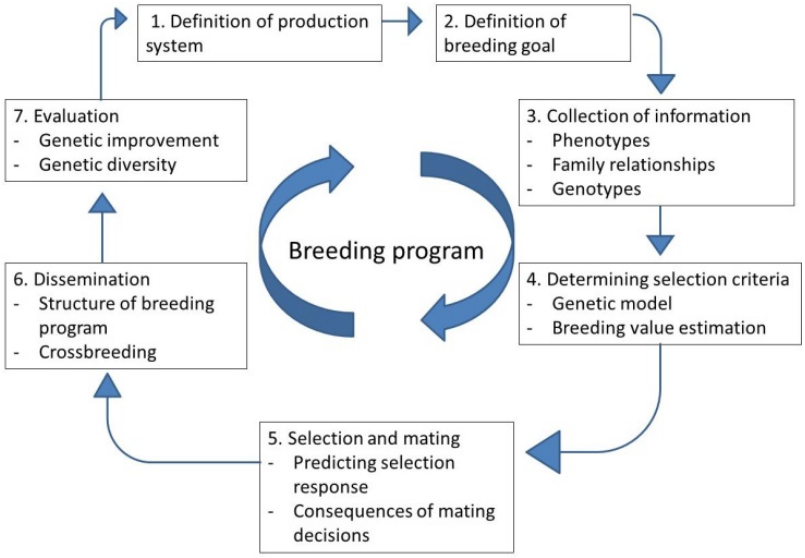
Chapter 7.9 Key issues in inheritance of monogenic traits

1. An animal can be homozygous for a monogenic trait (it got the same allele from its sire and from its dam) or heterozygous (it got a different allele from its sire and from its dam).
2. The alleles determining the expression of monogenic traits may be dominant, intermediate or recessive. The heterozygous genotype has a value equal to one of the homozygous genotypes (dominant), or higher (co-dominant) or equal to the average value of both homozygous genotypes (intermediate).
3. The Hardy and Weinberg equilibrium implies that in large populations with random mating among parents, and in the absence of selection, migration, mutation and random drift, the genotype and allele frequencies are constant (do not change from generation to generation) and the genotype frequencies can be calculated from the allele frequencies.
4. When the genotypes of the parents are known, you know what genotypes you may expect in the offspring. These expectations are based on the Mendelian laws, but due to random effects, the real world may deviate from the expectations.
5. In all species a lot of desired monogenic traits or traits with a large positive effect are known: e.g. colour genes and genes affecting the quality of animal's products.

6. In all species a lot of undesired monogenic traits are known. They are often based on recessive alleles and cause genetic defects when homozygous present in an animal.
7. The allele for a genetic defect will be spread in the population when the carrier is used extensively for breeding and it will pop up when in later generations animals with an additive relationship with the carrier are mated.
8. Genetic markers for a recessive genetic defect are highly valuable in the selection against heterozygous carriers of the allele for the defect. A prerequisite is that the genetic marker is positioned close to the recessive allele, otherwise recombination may disrupt the linkage between the marker and the allele.
9. The best strategy to decrease the allele frequency for the recessive defect is to test the offspring of carriers with the genetic marker and to select for the subsequent generation only animals that do not carry the allele.

Chapter 8: Ranking the animals

In the previous chapter we have seen how we can model the phenotype into genetic and environmental components. That chapter was quite theoretical, but we need that theory to continue with our breeding program. We have collected a lot of information on the performance of the animals, of their pedigree, and of their relatives, and now we want to know which animals are the best to be used in breeding. In other words: how can we rank the animals to be able to select the best? How do we know which have the best genetic potential and thus have the highest value for breeding (best breeding value)? We have observations on the phenotypes of the animals, but the phenotype is not only determined by the genetic potential, but also by environmental influences. In order to rank the animals according to their breeding value we need to find a way to quantify that breeding value, given the information we have: phenotypes and pedigree (i.e. the genetic relationships between animals). In this chapter we will present different techniques to do just that: rank the animals in order of estimated breeding value. We will also discuss which technique is most suitable in which situation as each technique has its own advantages, but also disadvantages. There are two main reasons why 'breeding is gambling': first of all the breeding value needs to be estimated and this may involve inaccuracy of the estimated breeding value. Second, even if you know the breeding value of an animal with 100% accuracy, so you know its true breeding value, you can't predict which half of the genetic potential is passed on to each of the progeny. This factor will always remain a factor of insecurity in breeding decisions because until now it is not possible to know the exact genotypes of a spermatozoid and an oocyte prior to conception.



If we look at the diagram again with the stages involved in the breeding program circle, then we are still at the stage number 4: determining selection criteria. After having developed the genetic model in the previous chapter we will now consider estimating the breeding values of the animals.

Chapter 8.1: Ranking the animals: an overview of methods

The challenge in selective breeding is to find the best animals as parents for the next generation. It is impossible to read the real genetic potential of the animals, but we can make an estimate of it. This estimated genetic potential is also called the *Estimated Breeding Value (EBV)*. Obviously, the estimate will be more accurate if there are more or better indications available of the genetic potential. The EBV is expressed relative to the average animal in a population. So it expresses an estimate of how much better than average an animal is.

Definition

*The **Estimated Breeding Value** (EBV) provides an estimate of the genetic potential of the animal. It is expressed relative to the population average.*

Chapter 8.1.1: Mass selection

The most basic way is to rank the animals according to their phenotype and select the best ones for breeding. This method is also called *Mass Selection*, or Selection on Own Performance. For example, of you want to breed large rabbits, then you rank the animals according to size and only use the largest as parents for the next generation. Will that be a successful method? The answer to that will depend on a number of things. You want to know whether these largest animals indeed were the ones with the best genetic potential. Why were the others smaller? Were they not fed properly when they were young? Or did they have the wrong genes? An answer to these questions lies in the heritability. After all, that is an indication of how much of the variation in phenotypes you observe are caused by genetic variation among the animals. A high heritability indicates that the small rabbits most likely are small because they have lower genetic potential for growth than the larger rabbits. The better the phenotype provides a representation of the genotype, the better you are able to identify the genetically best animals, and thus the better the results of mass selection will be. In addition, an important prerequisite is that own performances are available.

Definition

***Mass selection** is based on ranking the animals on their own performance.*

The success of mass selection is dependent on the heritability of the trait under selection.

Chapter 8.1.2: Animal Model

However, if the heritability is low, then mass selection not necessarily results in selection of the genetically best animals. Also, if the phenotype for some reason is not available for all animals, such as for milk production of males, then mass selection is not sufficient because not all animals have a phenotype. In those cases we can use phenotypes of related animals to estimate the breeding value of the animal without a phenotype. This is possible because, as we have seen in the chapter about genetic relationships, related animals share alleles. And the closer the relationship, the more alleles are shared. This model for estimating breeding values while making use of information on related animals is called the ***Animal Model***. Important prerequisite is that the pedigree recording of the animals is accurate, so that family relationships are known without errors. This method requires quite large numbers of animals to be able to estimate the breeding values accurately. The animals need to be related and/or kept in the same environment to be able to disentangle the genetic and environmental component of the phenotype.

Estimating breeding values with the animal model is very useful in case of missing phenotypes because the genetic relationships with animals that do have a phenotype allow for estimation of breeding values for animals without phenotypes. But even if the own phenotype is available it still has added value to mass selection because it can make use of the additional information on performance of related animals. This gives a more accurate estimated breeding value.

Definition

*The **Animal Model** is a genetic statistical model that combines the information on phenotypes of related animals to achieve a better estimate of the breeding value of an animal.*

Important advantages are:

1. you don't necessarily need a phenotype on each animal to be able to estimate its breeding value.

2. even if you have a phenotype, the extra information on related animals increases the accuracy of the estimated breeding value.

Chapter 8.1.3: Genomic selection

Finally, if you would collect phenotypes only on a select number of animals and you also have detailed genotypes on those animals, for example 60,000 SNP markers, then you can combine that information to estimate the link between the genome and the phenotype. Basic idea is that there are two groups of animals: a select group with detailed phenotypes, also called *the reference population*, and a large group without those phenotypes, also called *the population*.

All animals of both the reference population and the population are genotyped. In the reference population the associations between markers and phenotype are estimated. Then, those associations are combined with the genotypes of the animals in the population to predict their breeding values. This method is called *Genomic Selection.*

Genomic selection is very useful when phenotypes are very difficult or expensive to measure. Think about certain health related traits where you don't want to make the animal ill to be able to measure the phenotype. Or traits that don't require sick animals, but they do require expensive equipment, such as a CT scan. Genomic selection also makes it possible to select animals based on an estimated breeding value before they have reached the age to produce a phenotype themselves. This allows for (very) early selection, and thus can have economic benefits, as well as faster genetic gain because animals can be used as parents earlier. Disadvantage of genomic selection is that the reference population needs to be of sufficient size to be able to estimate accurate associations between genotype and phenotype. It also needs to be updated (=new animals need to be added) at regular basis because the estimated associations between the SNP and the genes determining the phenotype may be lost due to recombination and/or mutation.

Definition

Genomic selection *involves making use of the estimated association between very many SNP's and the phenotype to estimate the breeding value of animals without phenotype, but typed for the SNP's.*

This is particularly useful in case of:

1. phenotypes that are very difficult or expensive to measure

2. you want to estimate the breeding value of very young animals, before they can produce a phenotype

3. sex limited traits

Chapter 8.2: In more detail: breeding value estimation

In the overview you have seen that there are several ways to rank the animals. General aim in animal breeding is to rank the animals as well as possible. A tool for ranking the animals is the estimated breeding value. The more accurate the estimate of the breeding value, the better results of the subsequent breeding can be expected. We will now look into the three methods of estimating breeding values that were described in the overview in more detail.

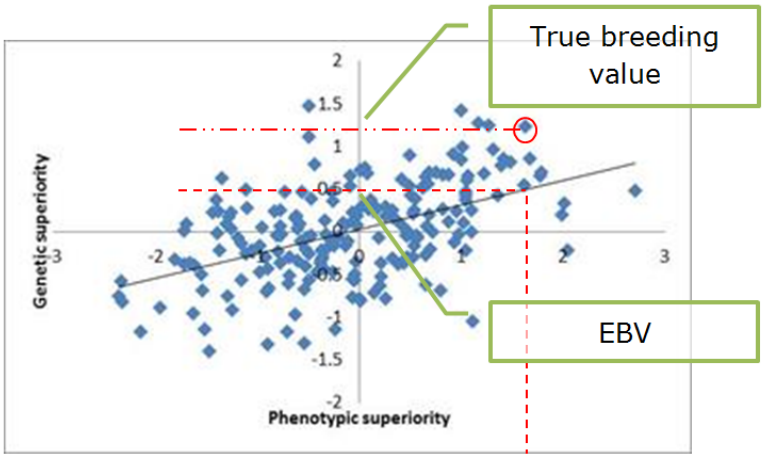


Figure 1: Relationship between the phenotypic superiority and genetic superiority of a group of animals. The regression line indicates the estimated relationship between P and G. This results in the EBV. For some animals this EBV reflects their true breeding value (G) better than for others, indicated by the distance between the data point of the animal and the regression line.

But first we need to know what an estimated breeding value really is. How do we get from information on the phenotypes of the animals and their genetic relationships (pedigree) to an estimate of the breeding value of the animals? In animal breeding we use the principle of regression to achieve this. In figure 1 you can see this principle visualised. If we would plot the true breeding value on the y-axis against the phenotypic superiority on the x-axis, then we can calculate a regression line through the data points. In real life, unfortunately, we cannot create such a plot as we don't know the true breeding values. Instead, we try to find the regression coefficient that, in combination with the phenotypic superiority, would best predict the genetic superiority or true breeding value (TBV). The art of estimating breeding values is based on finding ways to come to the best regression coefficient. This immediately also highlights a critical point in breeding value estimation: it is a linear regression coefficient but animals with the same phenotypic superiority do not always have the same genetic superiority. For some animals, like the animal indicated with a circle in the figure, the TBV is very different from the EBV, whereas for others the EBV would be the perfect estimate of the true breeding value. Part of this difference in how well the EBV resembles the TBV is caused by the fact that the phenotype can be influenced quite substantially by the environment. Therefore, simultaneous to finding the best regression coefficient, it is also important to try to make the phenotypic superiorities fit the regression line as well as possible. In the rest of this chapter we will discuss some options to work on both these problems: predict the best regression coefficient, and make the phenotypic superiority fit the regression line as well as possible.

Definition

*The **true breeding value** (TBV) of an animal represents the genetic potential of that animal: what is the real value of the animal for breeding?*

The perfect EBV would be equal the TBV.

Chapter 8.2.1: The basic statistics

In technical terms it works as follows. To be able to estimate the breeding value we need to re-arrange the formula for calculating the regression coefficient (b):

b = cov(x,y)/var(x)

In the formula b is the regression coefficient, var(x) = variance of the phenotypic superiority, cov (x,y) = relation between the phenotypic superiority and the true breeding value: our estimated breeding values. So we can rearrange our formula to:

Cov(x,y) = b * var(x)

Var(EBV) = b * var(phenotypic superiority)

For individual animals we can translate this to:

EBV = b*phenotypic superiority

The better the phenotypic superiority is predicting the true breeding value of the animals, the more the cov(x,y) will resemble cov(y,y), which equals var(y) = variance of the true breeding values, but at the same time cov (x,x), the variance of the phenotypic observations. In other words: the more the regression coefficient will approach 1.

As a final step: in animal breeding we are interested in identifying the genetically superior animals. That is easier if we would express their EBV relative to an average animal. Any positive EBV would indicate a better than average animal. This is easier than just giving the unadjusted EBV. For example, if you would know that your animal has an EBV of 25, that is nice, but not very informative if you don't know how the other animals score. It is much more informative if you would know that the average animal scores 23, so your animal scores +2. For convenience we, therefore, express the EBV relative to the average.

The phenotypic superiority can be calculated as $(P- P_{gemiddeld})$: the phenotypic information on the animal – the population average. Consequently, the true breeding value is also expressed as genetic superiority: $(A-A_{gemiddeld})$, and the EBV is an estimate of that.

The formula then becomes:

$EBV = b * (P- P_{gemiddeld})$

This formula is the most basic formula to estimate the breeding value of an animal: it combines the phenotypic superiority of the animal and the regression coefficient of the genetic superiority on the phenotypic superiority.

Thus:

For estimating the breeding values of animals we try to find the best regression coefficient and the most informative phenotypic information so that our EBV comes as close to the TBV as possible.

EBV are expressed relative to the average animal to simplify identification of genetically superior animals.

Chapter 8.3: Optimising the phenotypic information

Phenotypes of animals may be systematically influenced by, for example, the (management on the) farm the animals are kept in, or whether they were born in summer or winter, or whether they are male or female, etc. For a fair comparison of animals based on their phenotypes it is important to be aware of these systematic influences and to take them into account when defining the phenotypic superiority of an animal. For example, if males on average are 5 kg heavier than females then correction for effect would involve subtracting 5 kg from the weight of each male, so that males and females can be compared directly on their weight again.

The effect of correcting phenotypic superiority for systematic effects is that the resulting 'cleaned' phenotype will better resemble the genetic superiority. The clean phenotypes thus allow for a better prediction of the regression coefficient. This is illustrated in figure 2. In the figure on the left the cloud of data points indicates the 'raw' data: the data without correction for any systematic effects. When fitting a regression through these data, the regression coefficient would be 0.3. The figure on the right indicates the situation after 'cleaning'. The regression coefficient has increased, indicating that the phenotypic information has become a better predictor of the true breeding value.

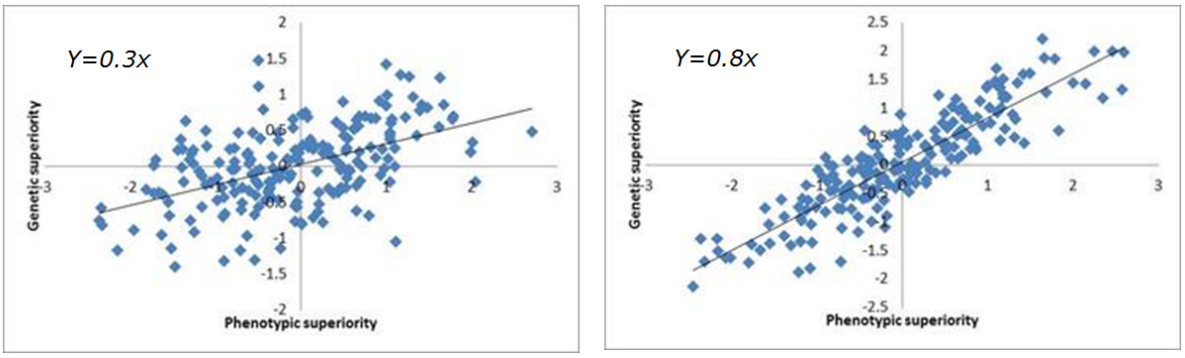


Figure 2: In the figure on the left is the phenotypic superiority uncorrected for systematic influences. In the figure on the right is the data corrected for systematic effects, resulting in better resemblance to the genetic superiority, and thus a higher regression coefficient.

Thus:

The phenotypic superiority can be improved by cleaning the data from systematic environmental effects

Chapter 8.4: Accuracy of the breeding value: the basic concept

Also in case of a high regression coefficient there are still some animals that have an EBV that is higher or lower than the TBV. If we would be able to estimate the breeding value with 100% accuracy, the EBV and the TBV would be the same value. If we would plot the TBV against the EBV then all data points would be perfectly in line. The less the data points are in line, the less certain you are that the EBV indeed is representing the true breeding value: the estimations are not accurate. A measure for data points being in line, and thus the accuracy of the breeding value estimation, is the correlation. If the correlation between estimated and true breeding values is 1, then you have managed to create the perfect estimates. The further away from 1 (i. e. the more they form a cloud), the less accurate the estimated breeding values are.

This is illustrated in figure 3. On the left you see a cloud of data points: some EBV resemble the true breeding value, but some estimates are also way off the true breeding value. The correlation between the EBV and TBV in this figure is 0.76, the EBV not resembles the TBV for all animals. For example, there are two animals with an EBV of 4, whereas their true breeding values are different: 3 and 5. In real life we cannot produce a graph like in this figure because we do not know the true breeding value. But what we can do is estimate the accuracy of the estimated breeding value: the correlation between the phenotypic information and the true breeding value. So how much is the EBV in line with the true breeding value.

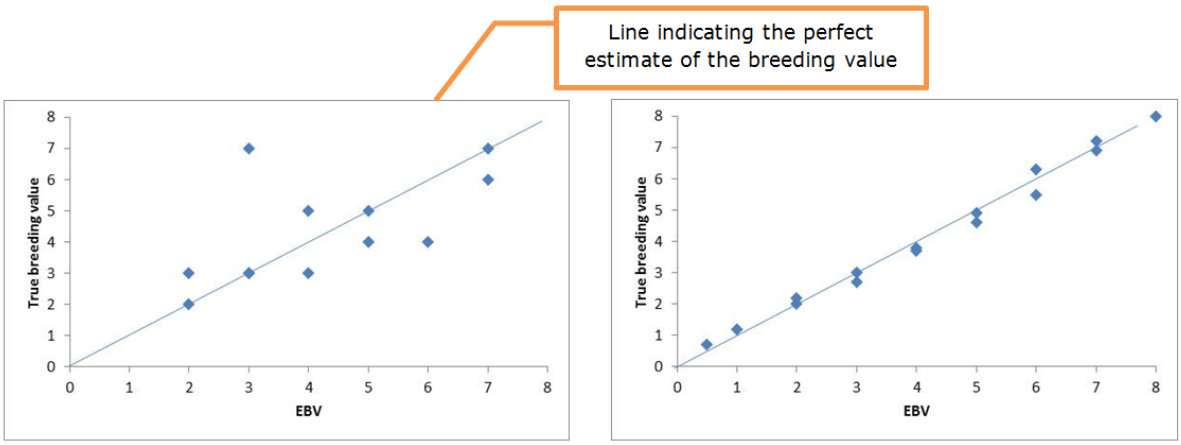


Figure 3: plots of TBV on EBV with the perfect regression line if EBV = TBV. On the left is an example of inaccurate EBV, indicated by the cloud of data points with correlation between EBV and TBV of 0.76. On the right the EBV are estimated much more accurately and are almost in line with the TBV, resulting in a correlation between EBV and TBV of 0.98.

Thus:

The accuracy of the breeding value estimation represents the correlation between the EBV and the true genetic superiority, and has value between 0 (inaccurate) and 1 (100% accurate).

Chapter 8.5: Breeding value estimation: Mass selection

The most basic estimated breeding value is ranking the animals based on their own performance. This is also called *Mass Selection.*

If we would plot the phenotype (true breeding value + environmental influence) on the x-axis, and the genotype (true breeding value) on the y-axis, then the regression coefficient between the two is:

the heritability (cov(x,y) = cov(P,G) = cov(G+E,G) = cov(G,G)+cov(G,E) = var(G) + 0, so b = cov (X,Y)/var(X) = var(G)/var(P) = h^2).

So if we know the phenotypic superiority and we know the heritability, we are ready to estimate the breeding values!

$EBV_{mass\ selection} = h^{2} * (P\text{-}P_{mean})$

Where the P represents the own performance, and the the average own performance in the population. For example, if we consider rabbits and the trait under selection is body weight at 3 months of age, we need to start recording body weight at that age. The average rabbit in our population weighs 2.0kg. If we have a rabbit of 2.3 kg, then its phenotypic superiority would be 2.3 – 2.0 = 0.3 kg. Let's assume that the heritability for this trait in this population is 0.2. Then the EBV for body weight for this rabbit is 0.2*0.3 = 0.06 kg. Note that the unity of the EBV is equal to the unity of the phenotype, in this case kg body weight.

We continue the example with the rabbits. It appeared that even though the rabbits were supposed to have been weighed at 3 months of age, this was not always the case. Due to holidays and weekends, some animals were weighed a bit younger and some a bit older. You can imagine that this has influenced the heritability because there was variation due to age that will have increased the error variance, and thus decreased the heritability. And remember, the closer to 1 the heritability is, the better the phenotype represents the underlying genetic

superiority. The management was changed and now all rabbits were weighed at exactly 3 months of age and the heritability increased from 0.3 to 0.4. Consequently, we are better able to predict the underlying genetic superiority of the animal and the EBV increased to 0.4*0.3 = 0.12 kg.

Thus:

The EBV using own performance can be estimated as: $EBV_{mass\ selection} = h^2 * (P\text{-}P_{mean})$

Chapter 8.5.1: Special case: repeated observations on a single animal

For some traits, in time, more than one record of own performance will be collected per animal. For example, you can have a record on litter size of the first litter, but when the second litter is born this is additional information on the performance of the dam, and so is the birth of the third, and maybe even fourth litter. Because litter size is heritable you would expect that multiple litters within a dam are more alike than litters between various dams. A single record includes genetic and environmental influences. In the second litter the genetics of course is the same, but the environment may be somewhat different. And again, a third litter has the same genetics but maybe somewhat different environment. So the more litters you have, the better you should be able to estimate the genetic potential of the dam. How much better is indicated with a correlation between the subsequent records: the repeatability. The more the subsequent records are alike, the higher the correlation (max = 1).

Availability of more records on an animal allows for a better indication of the phenotypic superiority. You can imagine that the records of an animal in part are influenced by environment that is specific for that record, the so-called temporary environment, but also by environment that is similar across records, the so-called permanent environment. The temporary environment is different every time, so the effect is larger in one records than in the other. By taking the average of the repeated records, the phenotype is corrected for the effect of the temporary environment, and the phenotype based on this average performance thus represents a more accurate representation of the phenotypic superiority as indicated in figure 2.

The better we can express the phenotypic superiority and the higher the repeatability, the better we should be able to estimate the breeding value. Indeed this is the case. Repeated records allow for a better estimate of the regression coefficient. In case of a single record the regression coefficient is h^2, but if there are multiple records it becomes:

$$b_{mass\ selection,\ multiple\ records} = nh^2 / 1+ r\ (n-1)$$

where n is the number of repeated records, and r is the correlation between subsequent records: the repeatability. If the repeatability is 0.5 and we have 2 records, than the regression coefficient increases from h^2 to $2h^2 / 1.5 = 1.33\ h^2$. The value of repeated observations depend on the repeatability and on the number of records that are available. The lower the

repeatability, the more repeated observations are influenced by different environmental influences, and the more added value it has to collect multiple records and re-estimate the breeding value every time a new own performance record becomes available.

Thus:

> *Repeatability is the correlation between subsequent records: the more they are alike, the higher the repeatability (max =1)*
>
> *Repeated observations on own performance adds to the estimation of the regression coefficient. The lower the repeatability, the higher the added value of repeated observations.*

Chapter 8.6: Breeding value estimation: the Animal Model

Even though mass selection is a straight forward way of ranking animals as breeding candidates, it is not always the most accurate way. For example, if all we have is own performance, how would we be able to select animals without own performance? How can we select dairy cattle bulls for improved milk production? Or how can we select animals for meat quality? Meat quality can only be measured after slaughter and then those animals cannot be used for breeding anymore. Fortunately there is a solution for that: we use phenotypic information collected on related animals and use that for estimating breeding values for animals without phenotypes. The success will depend on how much their genetics will resemble that of the animal without own performance: their additive genetic relationship.

Of course for estimating the breeding value of an animal the additive genetic relationship with animals with phenotypes needs to be quite substantial, so that the additional phenotypic information has added value. For example, information on a full brother, who has on average half of its genome the same as your animal (so a=0.5), has more added value than information on a far cousin with additive genetic relationship of only 0.0625. And information on parents or offspring has more added value than information of full brothers and sisters, even though all have additive genetic value of 0.5. That is because parents pass exactly half of their genome on to their offspring. So their additive genetic relationship really is 0.5, whereas full brothers and sisters *on average* share half of their genome. The Mendelian sampling (the uncertainty of which half of the genetic potential of a parent is passed on to each offspring) determines whether this indeed is the same half or not. In other words: this half is not as certain as the half that parents and offspring share.

The method of combining the additive genetic relationships between animals with the phenotypic information of some of those animals to estimate all of their breeding values is called the *Animal Model*. The Animal Model is not only useful in case of missing phenotypic observations, but also to increase the quality of phenotypic information so that the breeding value can be estimated more accurately. So how does the animal model work?

Thus:

The Animal Model represents a method that makes use of phenotypic information of relatives to estimate the breeding value of an animal

Chapter 8.7: Breeding value estimation: the basic situation

For estimating breeding values two things are still required: phenotypic information and the regression coefficient of the true breeding value on the phenotypic superiority. Parent-offspring regression, which can be used to estimate the heritability as we have seen in the chapter on genetic models, in a way is estimating the breeding value of the offspring based on performance of the parents. In case of information on a single parent the regression coefficient equals half the heritability, where the half represents the additive genetic relationship between the single parent and the offspring. The EBV of the offspring thus can be calculated as:

$EBV_{offspring} = b_{single\ parent\text{-}offspring} * (P_{single\ parent} - P_{mean})$

$= ½\ h^2 * (P_{single\ parent} - P_{mean})$

In case information on both parents is available the regression of the phenotype of the offspring on the mid-parent equals the heritability, the additive genetic relationship with both parents is 2*0.5 = 1, so the EBV becomes:

$EBV_{offspring} = h^2 * (P_{mid\text{-}parent} - P_{mean})$

Chapter 8.7.1: Breeding value estimation: the rabbit example on mass selection

Let's go back to our example of the rabbits that we used in the section on mass selection. The average rabbit in our population weighs 2.0kg. If we have a rabbit of 2.3 kg, then its phenotypic superiority would be 2.3 – 2.0 = 0.3 kg. The heritability for weight in this population is 0.2. The EBV for body weight for this rabbit is 0.2*0.3 = 0.06 kg. Now we want to estimate the breeding value of the young offspring of this rabbit. They do not have an own performance yet, so we want to use the phenotypic observation of this parent. The EBV of the offspring thus becomes

$EBV_{\text{-}offspring} = ½ * 0.2 *(2.3 - 2.0) = 0.03$ kg.

This is lower than if the offspring itself would have had a phenotype. Do you see why? This is because to estimate the breeding value for the offspring the phenotypic observation of the single parent was used as information source and the additive genetic relationship between the

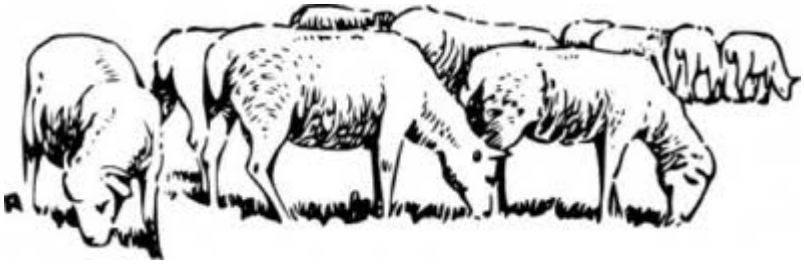
parent and its offspring is 0.5. Important assumption in a case like this with unknown other parents is that these other rabbits that were mated to this single parent were of average quality, so had EBV of 0, and thus did not contribute to the EBV of the offspring. That is why in this example the EBV of the offspring is half that of the parent.

Chapter 8.7.2: Breeding value estimation: a sheep example with information from parents

Another example, now for body weight in sheep and phenotypic information on both parents. The ram weighs 80 kg and the ewe weighs 70 kg. The average sheep of this breed weighs 65 kg. The heritability for body weight is 0.45. Estimate the EBV for unborn offspring. Let's take this step by step. First calculate the mid-parent value: the average weight of both parents is 75 kg. The phenotypic superiority becomes 75 – 65 = 10 kg. If we combine that information the EBV for the lamb becomes:

EBV_{lamb} = 0.45 *(75 – 65) = 4.5 kg.

This means that the lambs are expected to be 4.5 kg heavier than average, and thus weigh 65 + 4.5 = 69.5 kg.



Chapter 8.8: Other types of information sources

In general you can say that if you have information on a single related animal than the regression coefficient is equal to the additive genetic relationship times the heritability:

$b = a*h^2$

In case of parents or grandparents the additive genetic relationship is multiplied by the number of animals you have information on: up to 2 parents or up to 4 grandparents. In other words: the maximum regression coefficient becomes h^2 again.

It becomes a bit more complicated if there is information on a group of related animals that are not (grand)parents, for example a group of half sibs. The animal and its sibs may not only share a genetic component, but potentially also the effect of a common environment (often noted as c). This makes it more difficult to disentangle the effect of genetics and common environment, and thus has a negative influence on the EBV. In table 1 is a short list of formulas for the regression coefficients for a number of different information sources. You don't need to know all

these formulas by heart, but it is good to realise why some have a c^2 in the formula and others don't. In the formulas n indicates the number of records. So in case of progeny testing for each animal you estimate the breeding value for, it indicates the number of progeny that you have observations on, and in case of sib selection it indicates the number of sibs you have information on.

In table 1 you see that the regression coefficient in some cases contain a ½ or a ¼, for example the regression coefficient using full sib information has a ½ in it, and that for half sib information has ¼ in it. These are the additive genetic relationships between the information source and the animal you are estimating the breeding value for. Likewise, the additive genetic relationship with a single parent is ½ and so is that with offspring.

Now you know how to use information on relatives of an animal to adjust the regression coefficient. Of course it is possible to combine various information sources and have even better estimates of the regression coefficient. But for this course it is sufficient if you recognise the type of information source that is available and know how to use that information to estimate the breeding value.

Table 1. Formulas for regression coefficients to estimate breeding values using various information sources.

Information source	b-value (regression coefficient)
Own performance	h^2
Grandparents (mean of 4 records)	h^2
Single parent	$½\ h^2$
Parents (mean of 2 records)	h^2
Sib selection (n full sibs)	$\frac{\frac{1}{2}n\ h^2}{1+(n-1)(\frac{1}{2}h^2+c_{FS}^2)}$
Sib selection (n half sibs)	$\frac{\frac{1}{4}n\ h^2}{1+(n-1)(\frac{1}{4}h^2+c_{HS}^2)}$
Progeny test (n half sibs)	$\frac{\frac{1}{2}n\ h^2}{1+\frac{1}{4}(n-1)h^2}$

Thus:

The regression coefficient to estimate the breeding value using other sources than own performance depends on the additive genetic relationship with the animal, the heritability, the number of information sources, and the size of the shared common environmental effect

Chapter 8.9: Examples of estimating breeding values

First back to estimating breeding values. Remember: for estimating breeding values we needed the regression coefficient, but also the phenotypic superiority. So how do we obtain that if we have information on more than just a single animal? Fortunately that is simple: just take the average. For example, if you want to estimate the breeding value for a sire with 20 offspring based on the offspring performance, then you take the average of the performance of the offspring and relate that to the population average. If the offspring average is 50 and the population average is 40, than

$(P_{offspring} - \bar{P}) = 50 - 40 = 10$

Next step is to combine the regression coefficient and the phenotypic superiority so that we can estimate the breeding value. Remember the basic principle:

$EBV = b * (P - \bar{P})$

There are always three steps you need to take to estimate the breeding value of an animal:

1. determine the phenotypic superiority of your information source
2. determine the regression coefficient
3. combine the previous two to estimate the breeding value

Below you will find some examples on how to apply this in practice.

Examples:

1.What is the EBV for a stallion with excellent parents?

The heritability for rideability in riding horses is 0.29. The sire of this stallion scored 9.5 for rideability, and the dam scored 9.0. The population average is 7.0

Step 1: the phenotypic superiority equals the parent average, which is (9.5+9.0)/2 – 7.0 = 2.25

Step 2: the regression coefficient for mid-parent information is $h^2 = 0.29$

Step 3: the EBV = $b * (P - \bar{P}) = h^2 * (P - \bar{P})$ = 0.29 * 2.25 = 0.65

2. What is the EBV for milk production of a dairy bull with 100 daughters (half-sisters)?

The heritability for milk production is 0.3. The daughters produce on average 10,000 kg, and the population average is 9.500 kg.

Step 1: the phenotype superiority = 10,000 – 9,500 = 500 kg.

Step 2: the regression coefficient (see formula for offspring information in table 1)

b= $\frac{\frac{1}{2}n\,h^2}{1+\frac{1}{4}(n-1)h^2}$ = (½ * 100 * 0.3)/(1+ ¼*(100-1)*0.3) = 15 / 8.425 = 1.78.

Step 3: the EBV for milk production of this bull is 1.78 * 500 = 890 kg.

Note: the maximum regression coefficient of a single parent (usually sire) on offspring is 2 because the sire passes half its genome on to the offspring. Turning that around, and assuming that the sire is mated to average dams, if you have information on the superiority of the offspring than that of the sire is that of the offspring times 2.

Thus:

> *The maximum regression coefficient when using offspring information is 2, and not 1*

3 What is the EBV for average daily gain while growing from 25 to 100 kg of a pig with information on 20 full sibs, but no own performance?

The heritability for slaughter weight is 0.4, the population average is 875 g/d, and that of the 20 full sibs is 900 g/d. The common environmental effect for full sibs (c^2) = 0.45

Step 1: the phenotypic superiority = 900 – 875 = 25 g/d

Step 2: the regression coefficient = (½* 20 * 0.4)/(1+(20-1)*(½* 0.4 + 0.45)) = 4/13.35 = 0.30

Step 3: the EBV for average daily gain from 25 to 100 kg for this pig is 25 * 0.3 = 7.5 g/d.

Note: the regression coefficient is lower than the heritability. Reason is that full sibs perform more alike because they have shared a common environment. Therefore, a smaller proportion of the phenotypic superiority can be assigned to shared genetics than without shared common environment. This is taken into account through the c^2 when determining the regression coefficient for estimating the breeding value.

Thus:

> *The presence of a common environmental effect has a reducing effect on the estimated breeding value*

Chapter 8.10: Best Linear Unbiased Prediction

The Animal model is about utilising alternative information sources through their additive genetic relationship with the animal you want to estimate the breeding value for. Estimating breeding values is about optimising the estimation of the regression coefficient, but also about optimising the phenotypic information. There is a method that combines these two important

factors: it simultaneously corrects the phenotypes for systematic effects, and it estimates breeding values while making use of the additive genetic relationships between the animals. The result is an unbiased estimate of the breeding value. This method is the Best Linear Unbiased Prediction, or in short: BLUP.

It is a method that makes use of matrix algebra. We will not go into details here, but we will try to give the main idea. In formula it would look like this:

$Y = Xb + Za + e$

The Y is the phenotypic information, the Xb corrects the phenotypic superiorities for the systematic effects, and the Za links the phenotypic superiorities to the additive genetic relationships to estimate the EBV. The e indicates the error variance. In a way BLUP does follow the simple model P = E + G, but also provides estimates of G and E.

For example, if animals on one farm are fed much better than on another farm, then ranking animals based on their weight would benefit the animals from the farm with the better nutrition. However, genetically the animals on both farms may be similar. Without taking this systematic influence of farm of origin into account it is likely that the top ranking animals would mainly originate from the farm with the better feed. To be able to compare the performance of the animals more on their genetic potential it is important to take this farm effect into account, and this is what BLUP does (if you provide the information about on which farm each animal was housed). The principle of BLUP is to determine the average weight of the animals on each farm and subtract the difference from the animals on the farm with the highest weight. So if animals on farm 1 weigh 100 kg, and on farm 2 they weigh 120 kg, then you ‘punish’ the animals of farm 2 by subtracting 20 kg from their weight.

Critical issue in correcting for systematic effects is that it only works well if genotypes are sufficiently spread across systematic environmental influences. So the animals on both farms need to be related, for example because the same fathers were used, or because the fathers used on each farm were brothers. If the animals on both farms are unrelated, then part of the reason of the difference in weight may be a difference in genetic potential. And that is what you want to estimate so you don’t want to lose that by correcting the weight. Artificial insemination allows for genetic links between farms because the same sires are used in many farms. In farm animal species where natural mating is common practice, such as in beef cattle and sheep, it often is not possible to estimate systematic farm effects accurately because lack of exchange of animals between farms results in poor genetic links between farms. In species where the sires are brought to their mates on various locations, as can be the case in horse or dog breeding, genetic links will not be a limiting factor. Provided the sires are used often enough.

Thus:

With BLUP it is possible to estimate breeding values using information on relatives and correcting phenotypes for systematic influences.

Critical point is that sufficient genetic links between environments are required to estimate systematic effects of those environments (e.g. farms).

Chapter 8.11: Accuracy of estimated breeding values

The accuracy of an estimated breeding value indicates how well the estimate represents the true breeding value. In other words: it represents the correlation between the estimated and the true breeding value. Unlike a normal correlation, because of some underlying assumptions, this correlation cannot become negative. It can have values between 0 (totally inaccurate estimate) to 1 (estimated breeding value is the perfect predictor of the true breeding value). The accuracy is indicated by the symbol r_{IH}. The r represents the fact that it is a correlation, the I represents the estimated breeding value, and the H represents the true breeding value.

Thus:

The accuracy of an EBV gives an indication of how likely it is that the EBV is correctly estimated. It thus is an indication of the value of the EBV as selection criterion.

The accuracy (r_{IH}) is a correlation, but with values between 0 and 1.

Table 2. The regression coefficient and the accuracy of estimated breeding values for a number of different information sources.

Information source	b-value	Accuracy (r_{IH})
Own performance	h^2	$\sqrt{h^2}$
Grandparents (mean of 4 records)	h^2	$\sqrt{\frac{h^2}{4}}$
Parents (mean of 2 records)	h^2	$\sqrt{\frac{h^2}{2}}$
Sib selection (n full sibs)	$\frac{\frac{1}{2}n\,h^2}{1+(n-1)(\frac{1}{2}h^2+c_{FS}^2)}$	$\sqrt{\frac{\frac{1}{4}n\,h^2}{1+(n-1)(\frac{1}{2}h^2+c_{FS}^2)}}$
Sib selection (n half sibs)	$\frac{\frac{1}{4}n\,h^2}{1+(n-1)(\frac{1}{4}h^2+c_{HS}^2)}$	$\sqrt{\frac{\frac{1}{16}n\,h^2}{1+(n-1)(\frac{1}{4}h^2+c_{HS}^2)}}$
Progeny test (n half sibs)	$\frac{\frac{1}{2}n\,h^2}{1+\frac{1}{4}(n-1)h^2}$	$\sqrt{\frac{\frac{1}{4}n\,h^2}{1+\frac{1}{4}(n-1)h^2}}$

In table 2 are the accuracies of selection for the same information sources as in table 1. From the table it becomes clear that with only information on the parents, or even grand-parents, the accuracy of the EBV can never be as large as what can be achieved with own performance. In the absence of a common environmental effect (c^2), the maximum accuracy that can be

achieved can be determined by assuming a very large n. If n becomes very large, then the maximum r_{IH} that can be achieved with full sib information is equal to $\sqrt{\frac{\frac{1}{4}}{\frac{1}{2}}}$ which equals 0.707. In other words, for any trait with heritability larger than 0.5, selection on own performance gives a higher accuracy than selection on information collected on an infinitely large number of full sibs. Likewise, the maximum accuracy that can be achieved with half sib information, in the absence of a common environmental effect, is $\sqrt{\frac{\frac{1}{16}}{\frac{1}{4}}}$ which equals 0.5. So for any trait with heritability larger than 0.25, selection on own performance gives a higher accuracy than that based on an infinitely large number of half sibs. When there is a common environmental effect, these maximum achievable accuracies for full or half sib selection become smaller. From the formula it becomes clear that a common environmental effect decreases the accuracy of selection.

Thus:

Own performance results in a higher accuracy than full sib information for heritabilities larger than 0.5, and higher than half sib information for heritabilities larger than 0.25.

In the presence of common environmental effects among HS or FS, these heritabilities are lower.

Chapter 8.11.1: Effect of additional information on the accuracy

The more information is available related to the genetics of an animal, the more accurate the EBV. Information on offspring is highly valuable because they share the true half of the genes with their parent. But for young animals that is no realistic source of information yet. In the absence of information of full or half sibs, their parents are the only source of information. They share the true half of their genes with their offspring, and thus should be highly valuable as information source. However, there is a complicating factor with information of parents, because due to Mendelian sampling you do not know WHICH half if the genes are passed on. This is different from (multiple) offspring as information source, because there you do know that half of the genes are passed on. Under the assumption of an average other parent of the offspring the genetic potential (TBV) of the animal can be estimated quite accurately. The fact that each offspring receives one half of the genetics of the animal, but each may receive a slightly different half, makes it possible to quantify the Mendelian sampling component and get an accurate estimate of the EBV. Still the accuracy will depend on the number of offspring and the heritability.

Thus:

Information on offspring is more valuable than that of sibs because the offspring receives half of the genes from the animal. If sufficient number of offspring are available then the Mendelian sampling effect can be quantified and the EBV of the animal can be estimated very accurately.

Chapter 8.12: Summary of breeding value estimation

In summary: Own performance is valuable for trait with a high heritability. Phenotypes collected on half sibs are more valuable than those on full sibs. Even more valuable are phenotypes collected on half sib progeny: no common environment with the animal, generally larger potential numbers available, and offspring receive the true half of the genetic potential of the animal. This all allows for the most accurately estimated breeding value.

Now you know how to calculate the accuracy of selection. But why would you? Because the higher the accuracy of the EBV, the lower the risk of ranking the animals in the wrong order, and thus the lower the risk of selecting the wrong animals for breeding. Figure 4 illustrates the meaning of accuracy. In the figure are three normal distributions, each representing an EBV of 50, but with various levels of accuracy. The most accurate EBV has a 95% confidence interval between 45 and 55. In other words: the best estimate of the breeding value is 50, but with some level of inaccuracy around that estimate. However, it is 95% certain that the true breeding value lays between 45 and 55. The least accurate EBV also indicates that the best estimate is 50, but the 95% confidence interval lays between 35 and 65. That is much larger than that of the most accurate estimate. It means that the risk of ranking the animals in the wrong way is larger.

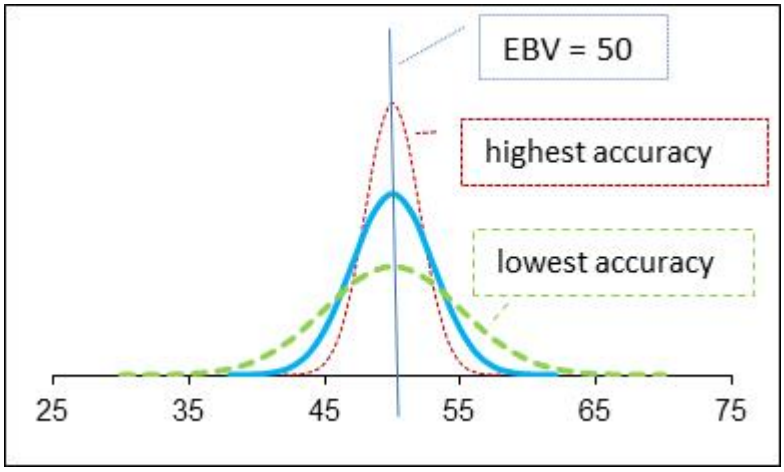


Figure 4. Distributions around the same EBV, but estimated with of high, intermediate, or low accuracy.

This difference in accuracy may be a result of a different heritability. The higher the heritability, the more accurate the EBV. But it can also be a result of suboptimal quality of information source, for example only a limited number of offspring, that could be used to determine the regression coefficient. In that case, in a next round of breeding value estimation it is likely that more information has become available. More information usually means that the regression coefficient can be better determined, and thus that the breeding value can be estimated more accurately. *This can have consequences for the size of the best estimate, and thus for the ranking of the animals!*! This is exactly why EBV, especially of the young animals with little information available, may change with new runs of breeding value estimation. If they do, it is uncertain in which direction they will change: up or down is equally likely.

Thus:

> *The lower the accuracy of an EBV, the larger the risk of a change in EBV when new information sources (e.g. offspring) become available.*

Chapter 8.13: Effects of number of offspring on the accuracy

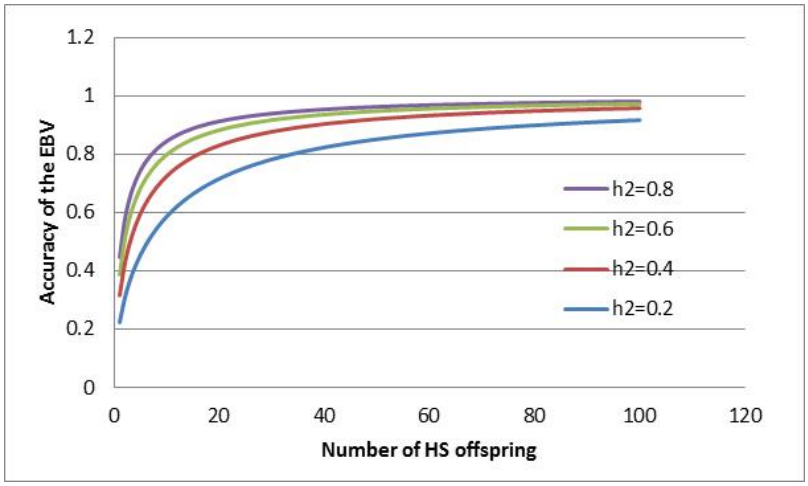


Figure 5. Relation between number of (HS) offspring as information source for estimating breeding values and the accuracy of the EBV for 4 different heritabilities.

In figure 5 you see the relation between the number of half sib offspring with observations for estimating the breeding value of their parent, and the accuracy of the EBV for that parent. The lines are in order of the heritabilities. For a trait with high heritability (0.8) 10 offspring is sufficient for an accuracy of 0.85, whereas you would need 48 offspring to achieve that for a trait with heritability of 0.2. The more information is available, in this example many offspring, the higher the accuracy of the EBV. Even with low heritability, the accuracy eventually will approach 1. Though for low heritabilities you would need very (often unrealistically) large

numbers of offspring: with a heritability of 0.2 and 100 offspring the accuracy is 'only' 91.7%, and with 200 offspring it only increases to 95.6%. A general rule is that the higher the accuracy of the EBV is, the less likely it is to change if additional information (more offspring) becomes available. Traits related to fertility often have a low heritability. For those traits with mass selection the accuracy of the EBV will be low. However, if very large numbers of offspring can be produced, such as in dairy cattle (bulls), pigs (boars), or poultry (hens and cocks), the EBV eventually can be estimated very accurately.

Thus:

Even for traits with low heritability the accuracy can increase to 1 if information on sufficiently large number of offspring is available.

Chapter 8.14: Go for highest EBV or highest accuracy?

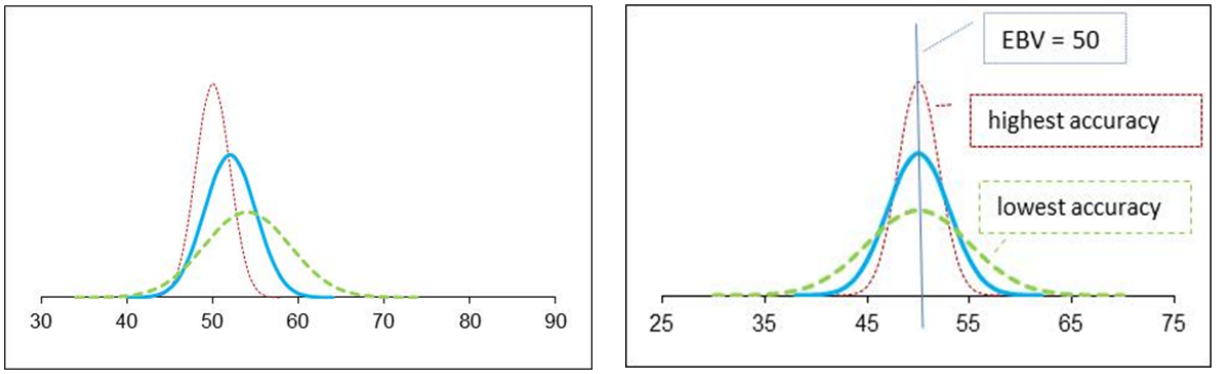


Figure 6. The effect of a difference in accuracy for the same EBV (right) or for a difference in EBV (left).

Now you know how to estimate a breeding value and you have learned about the influence of availability of information sources (related animals with phenotypic observations) on the accuracy of the EBV. Which is more important? Should you go for the animal with the highest EBV or with the highest accuracy? To answer that quickly: go for the highest EBV. To understand why that is the best thing to do, it is important to realise the meaning of both the EBV and the accuracy. As we have seen previously, the EBV is an estimate of the breeding value of the animal. It is an estimate, which means that it may be correct, but it also may be wrong. But it is important to realise it is the BEST estimate of the breeding value. It is the most likely breeding value, given the information that we have on the animal and its relatives. The accuracy of the estimated breeding value indicates how far we may be wrong. This is illustrated in figure 5. In the top you see the same figure that was used previously to indicate the meaning of the accuracy. All estimates are of the same EBV, but with different accuracies. In that case

you would select the animal with the highest accuracy as all have the same EBV. This is to reduce the risk of making the wrong choice. In the lower figure you see the same three accuracies, but now the EBV are different. The highest EBV has the lowest accuracy. But still it is estimated to be the best animal. There is a risk that the true breeding value is lower, but there is also an equal chance that the true breeding value is even higher than the EBV! In other words: even though there is insecurity about the value of the estimate, the EBV is the best estimate and it indicates the most likely value. An animal can have a very low EBV with very high accuracy. Of course a combination of high EBV and high accuracy would be optimal. It depends on the risk you would like to take what level of accuracy you are willing to accept.

Thus:

The EBV provides the BEST estimate of the breeding value of an animal. The accuracy indicates the risk of a difference between EBV and TBV, where the TBV may be higher or lower with equal probability.

Chapter 8.14.1: Outweighing EBV and accuracy in a dairy cattle example

The theory may become more clear with an example. In dairy cattle breeding, for example, a young bull only has a breeding value for milk production based on that of his parents. He may have a very high EBV, the accuracy of that estimate is low. His father may have an EBV with high accuracy of 90% (larger group of daughters with performance records), his mother only has her own performance and maybe some information of relatives, so her accuracy will be around 35%. The accuracy of the breeding value of the young bull will be 0.25*90 + 0.25*35 = 31.25%. Important reason why it is so much lower than the average of this parents is because of the Mendelian sampling: you know he inherited half of his genetics from his father and half from his mother. But you don't know which half. He may have inherited the best halves of each, resulting in a better than expected young bull. But he may also have inherited the worst half of both parents, resulting in very disappointing performance of his daughters. Because you will have to wait for his daughters to produce milk before you get an impression of how good his genetics really are. And you need many daughters before you can be 90% sure (accuracy of 90%). Despite this insecurity, the young bull with the highest EBV still is expected to be the best bull. This example was with dairy cattle but of course exactly the same story could be told for any other type of animal. The main message is that even though you know that the parents are good (high accuracy of high EBV), still the offspring may perform different than expected because of the effect of the Mendelian sampling component.

Thus:

The accuracy of the EBV of young offspring is not equal to the average of that of their parents because of the relatively large influence of Mendelian sampling: breeding remains gambling until you have insufficient information to estimate the EBV accurately.

Chapter 8.15: Genomic selection

From the example of the EBV of the young bull it became clear that the accuracy of his EBV will remain low until there are phenotypic observations on the daughters of the young bull. That takes a lot of time. It would be very interesting if there is a way to increase the accuracy of the EBV already at younger age, without having to wait for daughters to be born. It would also be very interesting if there is a way to estimate breeding values for traits that are difficult or expensive to measure, such as some health related traits or meat quality, without having to infect animals, or make detailed x-rays, or slaughter them. And since a few years there is a method that can do just that: genomic selection.

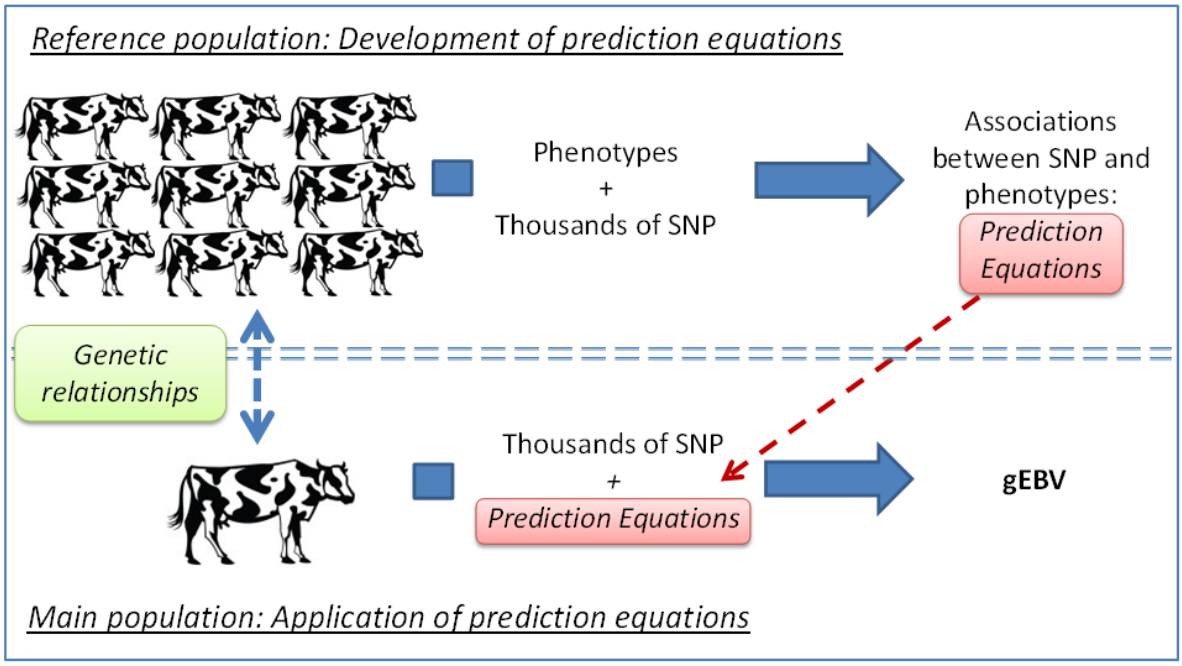


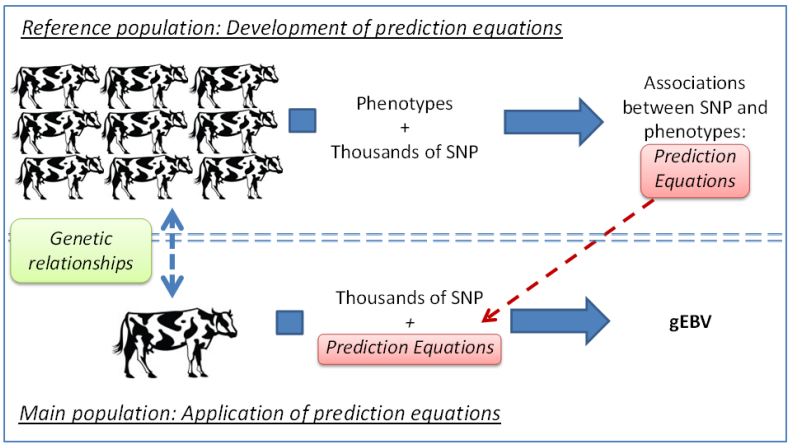
Figure 7. Schematic overview of the logistics behind genomic selection. The reference population provides the information to estimate associations between phenotype and SNP genotypes. These associations are then translated into prediction equations that are used to estimate genomic breeding values for animals without phenotype but with SNP genotypes outside the reference population.

With genomic selection it is possible to estimate an animal's breeding value quite accurately without the need for own performance or performance of large number of offspring. Genomic selection is based on estimation of detailed associations between a very dense set of genetic markers (SNP) and phenotypes on a select group of animals. These associations can then be used to predict the so-called genomic breeding values (gEBV) for related animals that have

been genotyped for a large set of SNP, but that do not have 'traditional' information for accurate EBV's like own performance or a large number of offspring with phenotypes. With genomic selection the DNA of the animal thus provides information for estimating the breeding value, without having to collect phenotypes on the animal itself or on its close relatives.

Chapter 8.15.1: Principle of genomic selection

In figure 7 the general principle of genomic selection is illustrated. First a lot of information needs to be collected on a select group of animals: the *reference population*. All animals in this reference population are genotyped for a very large number of SNP that are nicely spread across the entire genome. How many, is still under debate, but at least multiple thousand (e.g. 60,000). Genotyping for more SNP is more expensive, but will also result in more accurately estimated associations between the SNP and the phenotypes (i.e. the SNP effects). It is also still under debate what would be the optimum number of animals in the reference population. A larger population obviously is more expensive as the phenotyping and detailed genotyping of these animals is costly. But a larger population also allows for more accurate estimates of the SNP effects. Like many aspects in animal breeding the choice of size of the reference population and number of SNP will be a matter of cost-benefit analysis.



Given the phenotypes and the genotypes of the reference population, the associations between genotype and phenotype will be estimated for each of the genetic markers. Subsequently, the estimated effects are combined into so-called *prediction equations*. These are just summations (the estimated SNP effects are additive!) of the effect of the first SNP + that of the second SNP + … + that of the last SNP, so that the end result is the sum of all estimated SNP effects. Because each SNP has 2 alleles, for each SNP there are 3 possible genotypes. The prediction equations are established such that for each SNP the effects of all genotypes that are present in the reference population are estimated. This is one of the reasons why you need a large reference population: to estimate all these SNP effects accurately, each genotype need to be

represented by a sufficiently large number of animals. Now we have a set of equations with estimated SNP effects. The breeding value of animals outside the reference population can now be estimated by applying these equations to their SNP genotypes. These breeding values based on genomic information only are called genomic breeding values, or gEBV.

Thus:

> *Genomic selection is based on estimation of detailed associations between a very dense set of genetic markers (SNP) and phenotypes on a select group of animals: the reference population.*
>
> *The resulting prediction equations are then applied to SNP genotypes of the rest of the population to estimate their genomic breeding value (gEBV), without the need f additional phenotypes.*

Chapter 8.15.2: Composition of the reference population

Apart from sufficient size of the reference population it is also important that the reference population is related to the main population to assure that the estimated associations between SNP and the phenotypes also exist in the main population. The smaller the genetic relationship between reference and main population, the more associations will be different in the main population due to recombination between SNP and genes that determine the effect on the phenotype. Also with little genetic relationships still many of the association will exist because the animals in both populations are of the same breed, but the closer the relationship the better the estimates of the associations.

Thus:

> *The genetic relationship between reference population and main population is of influence on the accuracy of genomic selection. Associations between SNP and phenotype may be lost across generations.*

The need for genetic relationships between reference and main population is the reason for a restricted working life of the reference population. The accuracy of the estimated associations between phenotype and SNP reduces across generations. Main reason is recombination between the genes that are affecting the phenotype and the SNP. The larger the number of SNP that are used in genotyping the reference population, the longer its shelf life. But the linkage between SNP and genes will reduce across generations. The only solution is to update the reference population. It is still unclear what would be the best strategy for increasing the working life of a reference population. Should the population have a very large number of animals from the start? Or is it better to start smaller but add new animals every generation? And how many animals? It has become clear that even though the reference population is very large, it remains essential to add new animals at regular basis to update the estimated SNP associations.

Thus:

Reference populations need to be updated at regular basis to maintain associations between SNP and phenotypes

Chapter 8.15.3: Accuracy of genomic selection

The accuracy of an estimated breeding value depends on 3 factors: the heritability (h^2) of the trait, the number of animals (N) in the reference population, and a parameter called q. The q is a population and trait specific parameter that combines information on the length of the genome with the level of inbreeding for that trait in that population. It is an estimate of the number of independent chromosomal segments. The SNP that are closely together on a chromosome will inherit together without recombination between them. The further apart two SNP are, the more likely it is that there will be a recombination event between them. The higher the inbreeding level, the higher the level of homozygosity on the genome, and thus the less recombination will result in a change in combination of SNP alleles between two SNP. An independent segment is a measure for the likelihood that there will be recombination resulting in different allele combinations. The longer the genome, the more independent chromosomal segments there will be. This is a complicated story. For now it is sufficient to remember that q is specific for a population and may have different values for different traits. Details go beyond the scope of this course. In a formula the accuracy of genomic selection can be expressed as:

$$r_{IH} = \sqrt{\frac{Nh^2}{Nh^2+q}}$$

Thus:

The accuracy of genomic selection depends on the heritability, the number of animals in the reference population, and a population parameter q that reflects the relation between the size of the genome and level of inbreeding

Chapter 8.15.4: The size of the reference population

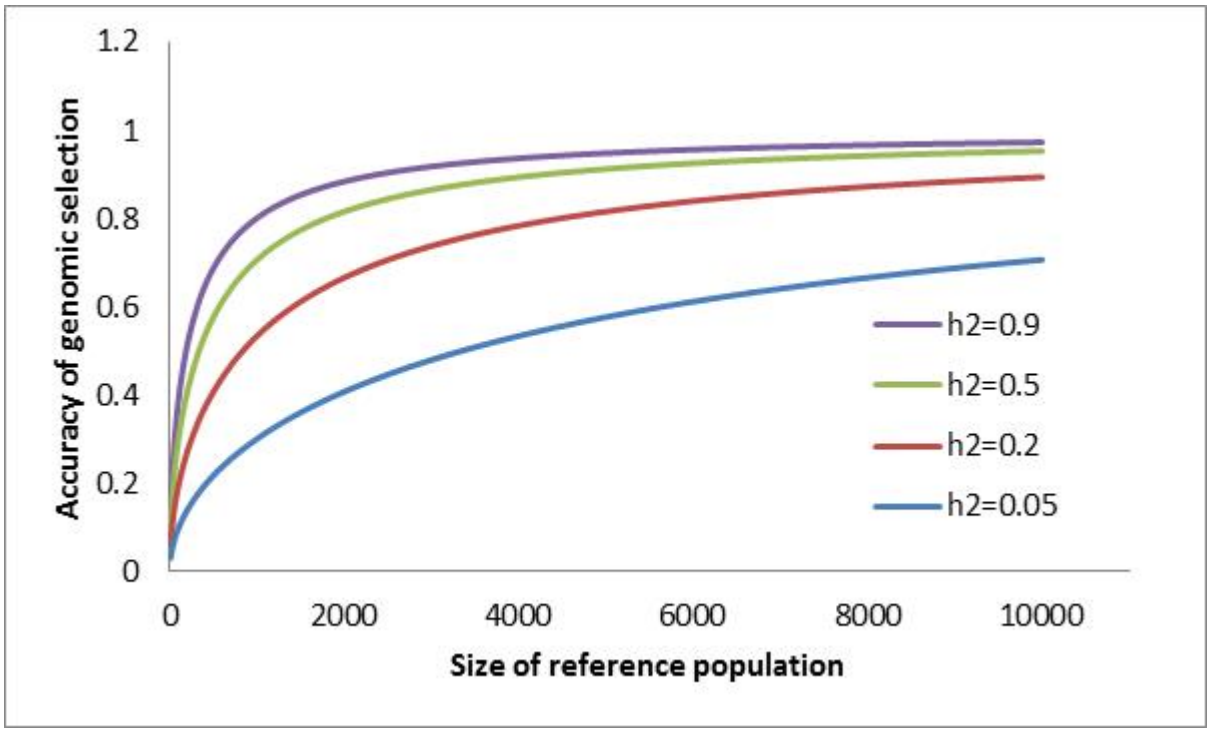


Figure 8. Relation between number of animals in the reference population and accuracy of genomic selection for traits with different heritabilities.

In figure 8 you see the increase in accuracy with increasing size of the reference population for 4 different heritabilities, assuming a q of 500, assuming that the gEBV are estimated using only genomic information. The upper line represents the trait with the highest heritability (0.9), and the lowest that with the lowest heritability (0.05). To achieve the same level of accuracy a decrease in the size of the heritability requires a large increase in the size of the reference population. For example, to achieve an accuracy of 0.6, 5630 are required for traits with heritability of 0.05, whereas only 320 are required for traits with heritability of 0.90. This illustrates that, even though genomic selection is a very nice tool, it is not feasible for small populations, and especially not for traits with a low heritability. As a potential solution, populations (studbooks) could combine forces in composing a reference population, so that they can share the costs, but also the benefits. This is current practice in dairy cattle breeding, where a number of international breeding organisations share a reference population. There is no example that reference populations of different breeds are combined. Theoretical this seems only to be effcetive when the density of SNP's is very high.

Thus:

The size of the reference population can be a imiting factor for achieving accurate gEBV. Solution can be to combine forces across breeding associations.

Apart from being able to have an accurately estimated breeding value already at very young age, genomic selection is especially useful for selection on traits that are expensive to measure accurately because only a relatively limited number of phenotypes are sufficient to improve the EBV of many animals. Even though genomic selection does allow for selection without phenotypic information of the animal itself or on close relative, the accuracy of selection still is also determined by the accuracy of the recording of the phenotype. Especially in the reference population the recording of the phenotypes should be done as accurately as possible because those phenotypes are used for selection of the rest of the population through their associations with the SNP. Inaccurate phenotypes result in suboptimal estimations of association between SNP and phenotypes, and thus in suboptimal estimated gEBV. Remember, the effect of inaccurate observations is directly reflected in the size of the heritability, and that effect is illustrated in figure 8.

It is possible to combine genomic and conventional breeding value estimation. Phenotypic information on the animal or its relatives will add to the accuracy of the gEBV. Details on how that would add go beyond the scope of this course.

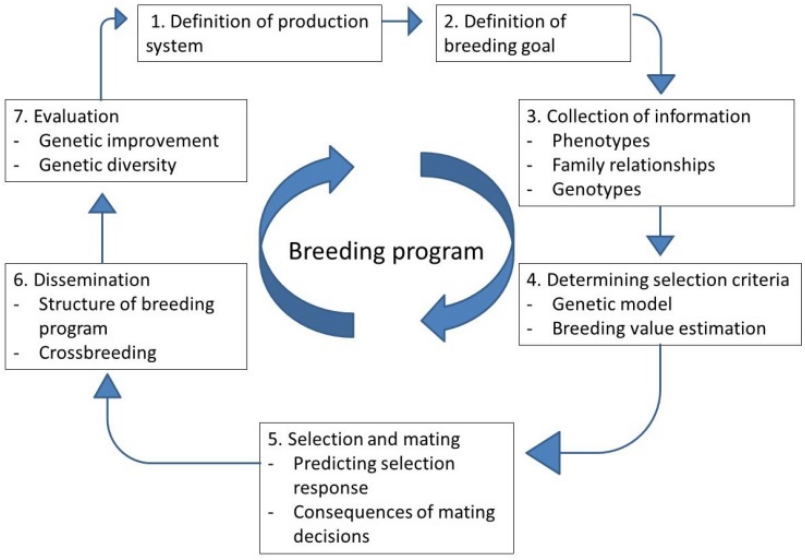
Chapter 8.16: Key issues on ranking the animals

1. Estimated breeding values provide estimates of the genetic potential of animals and indicate their potential value as parents.
2. Estimated breeding values are based on regressing genotypic information on the phenotype.
    a. Better quality of phenotypic information and better quality of the regression coefficient both improve the estimated breeding value.
    b. Accuracy of the estimated breeding value indicates how likely it is that the estimation breeding value resembles the true breeding value.
3. Mass selection involves ranking the animals based on their own performance.
    a. Repeated observations add to the accuracy of mass selection.
4. The animal model allows for incorporation of information on relatives in estimating breeding values.
    a. Often BLUP is used for the estimations. It combines information on relatives and allows for cleaning of the phenotypic data from systematic effects.
    b. Accuracy of information on relatives depends on the heritability and the number of relatives. In case of sib information, the presence of a common environmental effect may reduce the accuracy of estimated breeding values.

5. Genomic selection combines phenotypes and extensive SNP genotypes of a reference population, and uses the resulting associations to estimate breeding values for other animals with SNP genotypes.
    a. Allows for accurate EBV at very young age.
    b. Useful for phenotypes that are difficult or expensive to measure.
    c. The reference population needs to be of sufficiently large size and requires regular updating.

Chapter 9: Predicting response to selection

Now that we have tools to rank the animals according to the best estimate of their genetic potential, we are ready to start selecting the best animals for breeding. Immediately new questions are rising, because how many animals should we select? What are consequences if we select more or fewer animals? And for how long should a breeding animal be in service? Should we differentiate between parents that breed the next generation and eliminate parents that breed new parents? In this chapter tools will be presented to answer these questions. You will also learn that selection decisions often are steered by practical limitations.



If we look at the diagram again with the stages involved in the breeding program circle, then we have moved from the stage of determining selection criteria to the stage of actual selection and mating. In this chapter we will concentrate on the consequence of selection decisions by predicting response to selection. Mating will be subject of the next chapter.

Chapter 9.1: Response to selection: an overview

A number of steps are required to achieve response to selection, and thus to make genetic progress or genetic gain. First of all the animals need to be ranked in order of their predicted genetic potential. How to obtain the most accurate estimates of the genetic potential (the EBV) was subject of the previous chapter. Now that the animals can be ranked, the next step is to select the best for breeding. Success of the selection decisions depend on a number of factors:

1. How heritable is the trait under selection (i.e. the trait in the breeding goal)?

2. How much genetic variation for that trait is there in the population?
3. What is the average accuracy of the EBV, and thus the accuracy of selection?
4. What proportion of the animals will be selected for breeding?
5. In case genetic gain is to be expressed per year, rather than per generation: how long is a generation?

The heritability and the genetic variance are population parameters and cannot be influenced by the breeder. This is assuming that the phenotypes that were collected for estimating the heritability was of good quality, and the pedigree was recorded without errors.

A factor that can be influenced by the breeder is the accuracy of selection. If sufficient number of offspring can be used for estimating the breeding value, then the accuracy will be higher than if only performance of a few sibs is available. However, a downside of waiting for many offspring to be born before selecting the animals is that it takes a very long time before sufficient information is collected.

In Figure 1 is an illustration of a population that is ranked according to their phenotype for a certain trait. Most animals have an average phenotype, few are scoring very low, and few are scoring very high. After ranking you can select the best animals. The size of the proportion of the population that will be selected will depend on the how many animals are required for breeding. The selected proportion is the factor that is easiest to influence. A smaller proportion results in larger genetic response because the selected animals will be more superior than with a larger selected proportion.

However, the selected proportion cannot be unlimitedly small for two main reasons: first of all the intention is to maintain population size. So if few animals are selected, these need to be able to produce sufficient number of offspring to replace the entire generation of animals. Especially in females the number of offspring can be a limiting factor. Second, few animals selected as parents with large number of offspring, results in many animals that are closely related in the next generation, and thus a rate of inbreeding that may exceed the limit of 0.5 to 1 % as advised by the FAO.

Even though the genetic progress PER GENERATION is increased, the genetic progress PER YEAR is not or even decreased. In other words: there is a balance between increasing accuracy of selection and the time required to achieve the information to achieve the largest genetic gain per year.

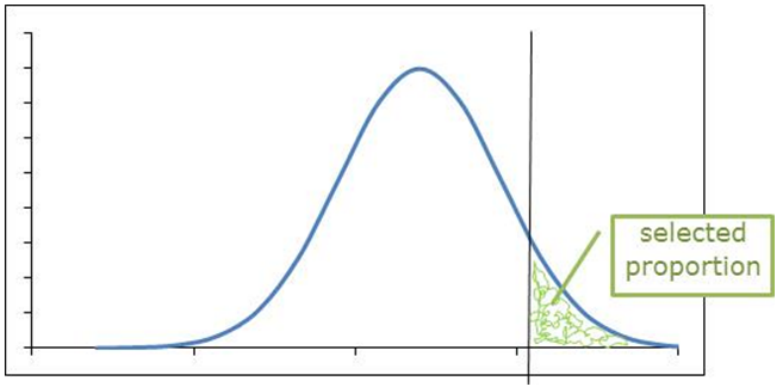


Figure 1. Illustration of a population with the fraction best animals selected. On the y-axe is the frequency of animals with that phenotype, and on the x-axe is the trait under consideration.

Thus:

To optimise the success of a breeding program it is important to balance the relatively short-term decisions: acquire high genetic gain, and the long term maintenance of the population: controlling rate of inbreeding.

Chapter 9.2: Breeding is about predicting the future

Breeding is always aimed at the future. Decisions you make now will influence the future generation(s). The breeding goal that you have defined indicates what you think will be important in the future. You have analysed the market and have an idea about what customers will demand some years from now. Will it be mainly milk or butter or cheese? Will it be mainly pork chops or ham or bacon? Will it be mainly breast meat or legs or full carcasses? Will it be top class sport horses or recreational horses? You also have an idea about how the market will develop in size: will it grow or downsize? Is that a temporary or long term development? Should you increase your breeding population or sell animals off? Will the market be the same at national and at international level? Finally, you have an idea about the expected developments in production systems and regulations. What are new developments related to housing systems, nutrition, etc and how are they expected to influence the performance of your animals? Has the (inter)national government announced new regulations that may limit your current production system? Should you anticipate to these upcoming changes?

Defining and updating your breeding goal is a very important part of the potential success of your breeding program. It is important to keep in mind that decisions you make today will only start to show after the offspring is born and start to perform. And, depending on the species, that may be years from now. Real effect of your decisions will only show after a number of generations. The breeding goal is aimed at the future, and often defined for a period of 10 to 15 years.

Thus:

> *Breeding is about predicting the future. All expected developments in the market and the production circumstances play a role in what the future will look like. All need to be anticipated when defining the breeding goal.*

Given the breeding goal, you will need to decide on the number of animals that you will need to select for breeding, and the number of offspring that these selected parents will produce. The selection strategy of your choice allows you to predict the performance of the offspring. Results may suggest some changes in the strategy. Therefore it is important to predict these results beforehand, so that adjustments can be made if necessary. That is what we will focus on in this chapter: how to predict the performance in the offspring, how can that prediction be improved, and what are consequences of selection decisions.

Chapter 9.3: Genetic response: the basic principle

Breeding is about selecting the best animals for breeding. The success of breeding decisions can be evaluated in the next generation. Animals in the next generation outperform animals in the current generation if the animals in the next generation *on average* are genetically superior to those in the current generation. Why on average? Because even though you have selected the very best animals as parents, their offspring combine genetics from both parents. Some of those combinations will be even better than that of the individual parents: they outperform the parents. Others, however, will have received a combination of genetics from both parents that is of less quality than that in the individual parents. They will perform less than the parents. This variation in performance is the result of the Mendelian sampling: each offspring receives half of the genetics from each parent, but each will have a different combination of chromosomes that have also recombined in the process of gamete production (meiosis). This is an important force of maintaining genetic variation in a population.

Back to predicting genetic response to selection. In figure 2 you see the process of selection and response in a schematic overview. The two normal distributions represent two generations. The top distribution is the generation of the parents. The best animals are selected for breeding. The other animals are not reproducing within the breeding program. The selected parents perform better than the population average. The size of the difference in performance between the parents and the total population (i.e. the superiority of the parents) is called the *Selection differential*, abbreviated as S. Note that ‘performance’ in this case indicates the selection criterion. That could be the phenotype (mass selection) or the EBV. The selected parents produce the next generation (lower distribution). This generation on average will perform better than the previous generation. The difference in average performance between both generations is called the *selection Response*, abbreviated as R. In general, the offspring generation will not perform as well as the average of the parents. Why is that? Because selection was not based on the true genetic potential (TBV) of the animals, but on an estimate

of that (EBV). It depends on the quality of the estimate how big the difference is between the selection differential and the selection response. The better the estimate, the closer the performance of the offspring will be to the performance of the selected parents. However, also the poorer the estimate the more often wrong selection decisions will be made because the genetically best animals have not been recognised as such. This results in less genetic gain. Note that lack of accuracy of the EBV (almost) never results in a selection response that is larger than the selection differential. The only situation in which this may happen is in case of cross breeding (see that chapter for further explanation). With within population selection the aim is to select the very best animals as parents. Any inaccuracies will lead to a result (offspring) that perform less, not better, than expected.

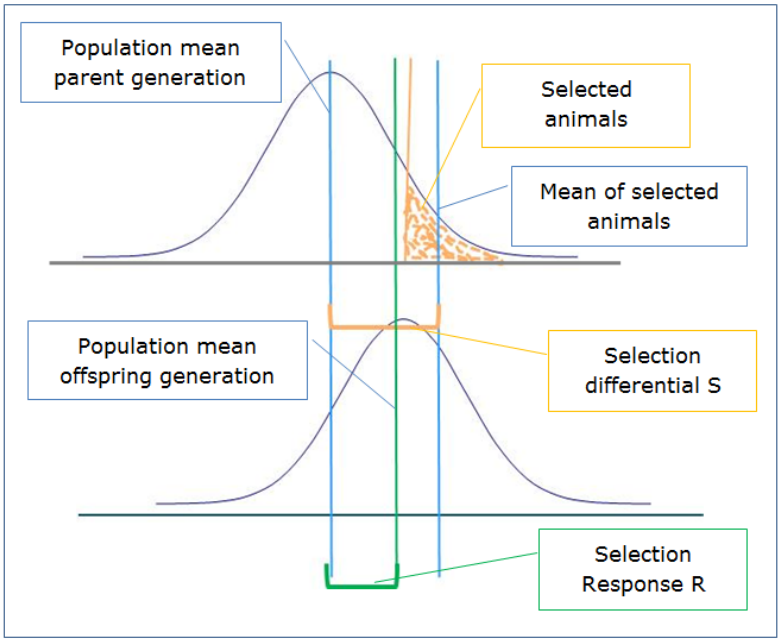


Figure 2. Schematic overview of the principle of selection and response to selection

Thus:

Predicting genetic gain is about predicting the future: how much better will the offspring perform compared to the current generation.

The superiority of the selected parents in comparison to their generation is called the selection differential (S)

The superiority of the offspring in comparison to their parents is called selection response (R)

Chapter 9.4: Response to mass selection

Mass selection is the most elementary type of selection: based on observed phenotypes. The variation among animals is represented by the phenotypic variation. The selection differential S would be the difference in average performance of the population and the selected parents. We are interested in the genetic response, so we need to translate the difference in phenotype to difference in genotype. To achieve this, we can scale the result with the heritability, because the heritability indicates what proportion of the phenotypic variance can be attributed to genetic variance. The result of this scaling would give the expected genetic response to selection in the offspring generation. In formula this looks like:

$$R = (\bar{P}_{selected\ parents} - \bar{P}_{parental\ generation}) * h^2$$

$$= S * h^2$$

Response to selection given a selection strategy causes a shift in genetic potential, also called genetic gain, indicated with . For mass selection response to selection is equal to genetic gain. The equation thus becomes:

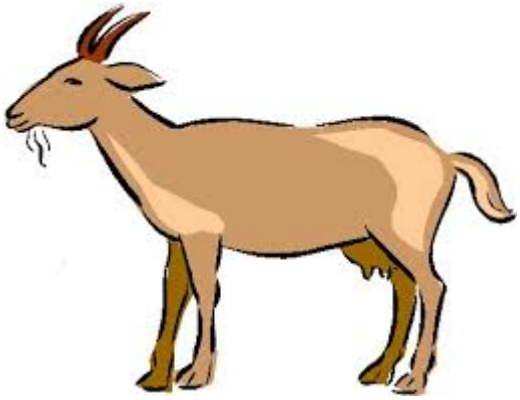
$$\Delta G = (\bar{P}_{selected\ parents} - \bar{P}_{parental\ generation}) * h^2$$

Note that this formula is very similar to that of estimating the EBV using mass selection:

$$EBV_{mass\ selection} = h^2 * (P - \bar{P})$$

In fact, estimating genetic gain is the same as estimating the average EBV of the parents, as that is how much better than the average of the parental generation you would expect the offspring generation to perform.

How does this work in practice? For example, you are managing a goat population and you want to increase their adult body weight. Your average adult goat weighs 50 kg. You have selected a number of males and females that on average weigh 55 kg (ignore for now the fact that males are heavier than females). The heritability of adult body weight in your population of goats is 0.42. You want to know, given your current breeding decision, what your average goat will weigh in the next generation. If we do the calculations: S = 55 – 50 = 5 kg, and your G = 5 * 0.42 = 2.1 kg. So you expect your next generation to be 2.1 kg heavier than the current generation: They will weigh on average 50 + 2.1 = 52.1 kg.



An important assumption in evaluating predicted genetic gain, is that the environmental influences will not change from generation to generation. Following the goat example: the next generation is expected to weigh 52.1 kg, provided that the environmental influences will remain the same. Of course this may not be true, but because we have no accurate idea about the environmental influences in the next generation. Therefore we assume they do not change.

Thus:

Important assumption in evaluating predictions of genetic gain: environmental influences remain constant across generations

Chapter 9.5: Selected proportion and selection intensity

The size of the genetic gain depends on the size of the selection differential (i.e. how much better than average the parents are). This can be influenced by three main factors.

1. First of all: if there is a lot of variation (^2_p) in the population, then it is easier to find animals that perform much better than average, compared to when there is little variation in the population. This is illustrated in the top figure in figure 3.
2. Second, what proportion (p) of the population you are using for breeding. A large proportion used for breeding means that on average parents are not that much better than the population average. The larger the selected proportion, the less superior the parents will be on average. This is illustrated in the lower part of figure 3. A small selected proportion will result in more superior parents than a large selected proportion.
3. Third, the accuracy (r_{IH}) of the selection: how certain are you that you have selected the genetically best animals for breeding?

The selected proportion alone is not a very good representation of how much better than average the parents are. It needs to be evaluated in combination with the size of the variation. A way to do that, is expressing the mean of the selected proportion in units of variation: the standard deviation. As is described in the chapter about introduction to statistics, the normal distribution can be divided into standard deviations according to a fixed pattern, such that 68% of the observations lay between plus and minus one standard deviation around the mean, 95% between plus and minus two standard deviations, and 99.7% between plus and minus three standard deviations. Many phenotypes tend to be normally distributed in a population. A phenotypic value can thus be expressed as being so many standard deviations away from the mean. We can use the selected proportion of animals and use properties of the normal distribution to determine the mean of the animals in that selected proportion, expressed in phenotypic standard deviations: the *selection intensity.*

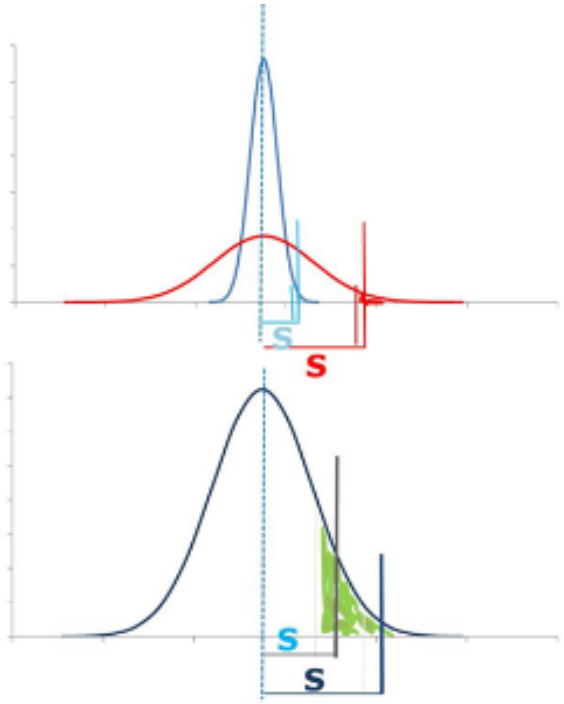


Figure 3. Illustration of the effect of size of variation in performance for the selection criteria (top figure) and selected proportion (lower figure) on the size of the selection differential.

Thus:

> *Genetic gain is determined by 3 main factors: phenotypic variance, accuracy of selection, and selected proportion.*

The selection intensity is abbreviated as *i.* In formula:

$i = S / _P$

so

$S = i * _P$

In summary: the selected proportion, in combination with the phenotypic variance, is enough to predict the average performance of selected parents.

Thus:

> *The selection intensity represents the mean of the selected proportion in phenotypic standard deviations.*

In chapter 9.5.1 you will find a table where you can look up *i* for any given selected proportion. This table is valid for selection on any trait that is normally distributed, so it is not specific for a trait or a population.

Chapter 9.5.1: Appendix: from selected proportion to selection intensity

Table to translate selected proportions to selection intensities. In the table p% is the selected proportion in percentages, and i is the corresponding selection intensity. For selected proportions in between the values mentioned in the table, take the linear approximation of i. For selected proportions larger than 50%, take the i for (1-p) and multiply that i by (1-p)/p. The table is a simplified version from the table published in Falconer and Mackay, 1987.

p%	**i**	**p%**	**i**	**p%**	**i**
0.01	3.960	1.0	2.665	10	1.755
0.02	3.790	1.2	2.603	11	1.709
0.03	3.687	1.4	2.549	12	1.667
0.04	3.613	1.6	2.502	13	1.627
0.05	3.554	1.8	2.459	14	1.590
0.06	3.507	2.0	2.421	15	1.554
0.07	3.464	2.2	2.386	16	1.521
0.08	3.429	2.4	2.353	17	1.489
0.09	3.397	2.6	2.323	18	1.458

0.10	3.367	2.8	2.295	19	1.428
		3.0	2.268	20	1.400
0.12	3.317	3.2	2.243	21	1.372
0.14	3.273	3.4	2.219	22	1.346
0.16	3.234	3.6	2.197	23	1.320
0.18	3.201	3.8	2.175	24	1.295
0.20	3.170	4.0	2.154	25	1.271
0.22	3.142	4.2	2.135	26	1.248
0.24	3.117	4.4	2.116	27	1.225
0.26	3.093	4.6	2.097	28	1.202
0.28	3.070	4.8	2.080	29	1.280
0.30	3.050	5.0	2.063	30	1.259
0.32	3.030			31	1.138
0.34	3.012	5.5	2.203	32	1.118
0.36	2.994	6.0	1.985	33	1.097
0.38	2.978	6.5	1.951	34	1.078
0.40	2.962	7.0	1.918	35	1.058
0.42	2.947	7.5	1.887	36	1.039
0.44	2.932	8.0	1.858	37	1.020
0.46	2.918	8.5	1.831	38	1.002

0.48	2.905	9.0	1.804	39	0.984
0.50	2.892	9.5	1.779	40	0.966
		10.0	1.755	41	0.948
0.55	2.862			42	0.931
0.60	2.834			43	0.913
0.65	2.808			44	0.896
0.70	2.784			45	0.880
0.75	2.761			46	0.863
0.80	2.740			47	0.846
0.85	2.720			48	0.830
0.90	2.701			49	0.814
0.95	2.683			50	0.798

Chapter 9.6: Selection response: the generalized approach

Knowing the phenotypic variance and the selected proportion allows us to determine the superiority of the selected parents (how much better than average the selected parent are) on forehand. This is based on the variance and the selected proportion only, so before the actual animals have been identified. This is very convenient! We can use this information to predict the genetic response to selection, given a certain proportion. This response can then be evaluated, and, if desired, compared to predicted selection results when using a larger or smaller selected proportion. The selection intensity is an important tool in the decision making process.

Apart from the selected proportion and the phenotypic variance, what is missing for predicting response to selection, or genetic gain, is the translation of the phenotype into an estimate of the genetic potential (EBV). We need to know how accurate that estimate is, and how to translate the phenotype to the genetic potential. There is a general formula for predicting genetic gain:

$$\Delta G = i * r_{IH} * \sigma_a$$

Even though it looks different, this formula actually is the same as the one for mass selection that was discussed earlier in this chapter:

$G = i * r_{IH} * {}_a$

$= S/_p * {}_a/_p * {}_a (= S * h^2)$

If we consider the components of the formula for genetic gain, it actually makes sense. The $S/_p$ indicates the genetic superiority of the parents, expressed in $_p$. The $_a/_p$ translates $_p$ into $_a$, so into genetics. The final $_a$ translates the result in units of the trait under selection (e.g. kg milk, or ride ability score). Once again: advantage of working with selected proportions (and thus with selection intensity) is that results can be predicted prior to the actual selection decision. From now on we will only consider $G = i * r_{IH} * {}_a$ as that is generally applicable, and not only for mass selection. Note that the order of the components of the formula does not matter.

Chapter 9.6.1: An example: the Arabian horse

In a population of Arabian race horses the breeding goal is to increase running speed in 2,000 m races for 3 year olds. The average time across this distance at that age in the current population is 117.0 seconds. The plan is to select the best 10% for breeding (for now we ignore the difference between males and females, both in speed and in reproductive capacity). The genetic standard deviation is 3.0 seconds, and the accuracy of selection is 0.24. Predict the genetic gain using these selection decisions. What would be the average 2,000 m time in the next generation?

Answer: Look in the table at the end of the chapter and find that a selected proportion of 10% results in a selection intensity of 1.755. This means that this selected top 10% on average performs 1.755 standard deviations better than the population average. It will depend on the variance for the trait how much better that is in trait units: 1.755 *3.0 = 5.265 seconds. This is a bit optimistic because we were not able to estimate the genetic potential 100% accurately. In fact, it was only 24% accurate (r_{IH} = 0.24). Combining all information results in: G = i * r_{IH} * $_a$ = 1.755 * 0.24 * 0.3 = 1.26 seconds. The example is about improving running speed, so the next generation will be 1.26 sec faster: 117.0 – 1.26 = 115.74 seconds.

Chapter 9.6.2: an example: rabbit breeding

Not only horse people like show jumping, also for rabbits there is a jumping competition. One very fanatic rabbit trainer decided to start selection on jumping capacity (measured in cm of fence size). Males and females both perform equally well in the jumping circuit. As females are capable of producing large number of offspring in a short period of time, selected proportion can be of equal size in males and females. Our breeder is selecting the 20% best jumpers for breeding, based on performance of 10 ? of their offspring. The heritability for show jumping capacity in rabbits is 0.14, with a phenotypic variance of 40. Predict the genetic gain that the breeder will achieve using this selection strategy.

A selected proportion of 20% results in a selection intensity of 1.4. The accuracy of selection can be determined using table 2 of the chapter on ranking the animals:

$$\sqrt{\frac{\frac{1}{4}n\,h^2}{1+\frac{1}{4}(n-1)h^2}}$$

.

If we fill in the heritability and the number of offspring, we come to an accuracy of selection of 0.266. The genetic variance can be determined from the heritability and the phenotypic variance: 0.14*40 = 5.6, so the genetic standard deviation is the square root of 5.6, which is 2.37.

Filling all that information into the formula results in a genetic gain of: G = 1.4 * 0.27 * 2.37 = 0.90 cm

The prediction is that the next generation on average will jump 0.90 cm higher than the current generation.

The breeder is disappointed and wants more genetic gain. Predict how much improvement he can achieve be selecting the top 15% instead of the top 20% for breeding.

The accuracy of selection and the genetic standard deviation remain the same, but the selection intensity increases from 1.4 to 1.554. The predicted response to selection then becomes 0.99 cm.

The breeder is still not completely satisfied because he wants a genetic gain of more than 1 cm and decides to base the selection on the performance of 12 instead of 10 offspring. Predict the selection response for this new situation.

Selection based on performance of more offspring will increase the accuracy of selection. Re-calculating that results in an r_{IH} of 0.30. The predicted selection response now becomes 1.554 *0.30*2.37 = 1.10 cm.

Chapter 9.7: Generation interval

The genetic response to selection predicts how much better the next generation will perform compared to the current generation. Playing around with the components of the formula of genetic gain will give you a feeling for how selection decisions can influence the response to selection.

From the example on show jumping of rabbits you have seen that selected proportion and accuracy of selection are of influence on the predicted response to selection. The problem that we are facing is that the genetic gain that we predicted is expressed per generation. How long is a generation?

Definition

The generation interval *is the average age of the parents at birth of their offspring that in their turn will produce the next generation of breeding animals. The generation interval facilitates to calculate the genetic response per year instead of per generation.*

A genetic gain of 1.10 cm per generation does not give much information. Obviously, if rabbit generations would last 0.3 years, this is much faster genetic gain than if generations last 1 year. To get a good impression of the achieved genetic gain, even if you don't know how long a generation lasts, the genetic gain needs to be expressed in time units. A commonly used time unit is a year. To be able to express the genetic gain per year you need to know how many years a generation lasts. Obviously, the first offspring is born at an earlier age of the parent than the last offspring. Some animals have their first offspring earlier than others, and some animals will have only one offspring, whereas others will have more. How to take all that into account? The definition of the length of the generation interval (abbreviated as L) is to take the average age of animals when the average offspring (number) is born, AFTER selection of the animals as parents. The word 'after' is important here, because it is important to realise that in case of selection based on progeny performance testing, the progeny that are used to base the selection on officially are not part of the generation interval. Figure 4 presents a schematic overview of the concept of generation interval. In the top is the situation where animals are selected either on own performance or on performance of their sibs.

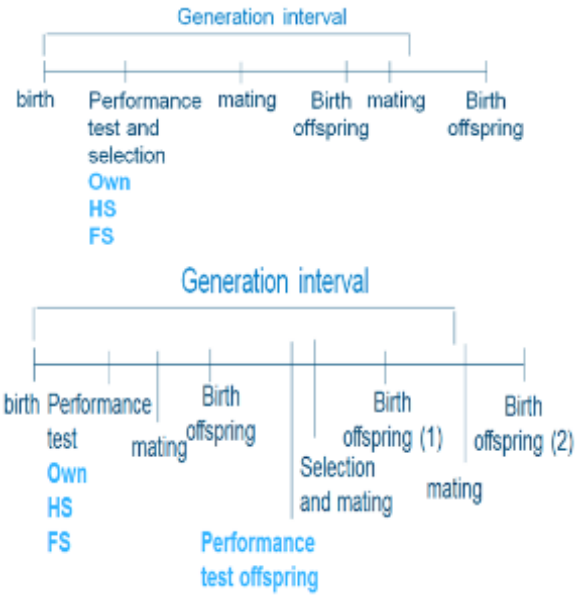


Figure 4. Schematic overview of the principle to the generation interval. In the top figure the situation where animals are selected on their own performance or that of their sibs, and in the bottom figure is the situation where animals are selected based on the performance of their first progeny. Those offspring do not count in determining the generation interval, resulting in an extended generation interval in populations where parents are selected on progeny performance.

Chapter 9.8: Optimising genetic gain

After selection the animals are mated for the first time and offspring will be born. In this example animals on average will have two batches of offspring (single or litters). The length of the generation interval is equal to the age in between the births of both batches. Assumption in this figure is that the number of offspring born in each batch is the same. If not, then the (length of the) generation interval needs to be weighted according to the number of offspring in each batch.

For example, in a sheep breed the ewes will have their first batch of offspring (a single) at 1 year of age, and the second offspring (a single) at 2 years of age. The generation interval in that case is (1*1 + 1*2)/(1+1) = 1.5 years . However, if those same ewes would generally have single in the second batch, but some would have twins, so that the average number of offspring in the second batch would be 1.3 lambs, then the generation interval would become (1*1 + 1.3 *2)/(1+1.3) = 1.56 years.

For animals that are selected on the performance of their first progeny the ‘counting’ only starts from the second batch of offspring onwards. This is presented in the lower part of figure 4. Otherwise the principle is exactly the same as with selection based on own performance or sibs. It is clear that the generation interval will become longer if selection is based on progeny testing.

If we continue the example with the sheep, selection is based on the first offspring. Now the ewes all get the opportunity to produce an extra batch of offspring, so that each ewe will produce 2 batches after being selected as parents. The average age of the third batch of offspring is at 3 years of age, and the ewe on average will have 1.5 lambs. The generation interval will become: (1.3*2 + 1.5*3)/(1.3+1.5) = 2.54 years.

The genetic gain thus far was expressed per generation. Now that we have calculated how many years are in a generation, we can express the genetic gain per year:

$$\Delta G = \frac{R}{L} = \frac{i * r_{IH} * \sigma_a}{L}$$

Note that there is a relation between accuracy of selection and generation interval. The accuracy can be increased by improving the information sources to base the EBV on. Information on performance of (large numbers of) progeny gives the highest accuracy. However, it also takes a long time to collect this information. In other words: the generation interval increases. So the improvement in genetic gain per year because of the increased accuracy may be outbalanced by the increase in generation interval. Also, producing lots of offspring of parents that have not been approved for breeding will cost a lot of money.

If we briefly go back to the example with the jumping rabbits: The breeder was happy when selection was based on performance of 12 offspring. However, he may want to look into the matter in more detail because it will depend on the litter size whether this number can be achieved with a single batch of offspring, or whether multiple batches are required. Multiple batches mean more time and the generation interval in rabbits is low. In such situations it may be a consideration to accept a slightly lower accuracy of selection, but manage more generations of selection in the same time frame. It may result in more genetic gain per time unit in the longer run.

Thus:

Optimising genetic gain will require a balance between increase of the accuracy and increase of the generation interval

Chapter 9.9: Selection paths

Thus far we have not considered difference in selection strategy between males and females. However, in most species there is a difference. There are 3 main reasons for this difference:

1. Important reason for that is the limited reproduction capacity in females, especially in mammalian species. A general assumption in animal breeding is that the population size remains the same across generations. This has consequences for selection strategies, because it means that the selected animals should be capable to produce sufficient number of offspring to maintain population size! Males can produce more offspring than females, especially when AI is available. Selection in males, therefore, often is more stringent than in females. In other words: there may be a difference in selection intensity.
2. Another reason is a difference in information sources to estimate the breeding values in males and females. There are some traits that cannot be measured in both sexes, for example milk production. Consequently, males may be selected based on progeny performance, whereas females are selected on own performance, resulting in a difference in accuracy of selection.
3. Related to that, a third reason is the age at which animals can be selected and at which the average progeny is born. If males are selected based on progeny test results, then they on average will be older than the females, provided those are selected based on own performance, for example. On the other hand, in most species males mature earlier than females. In other words: there may be a difference in age at first progeny, and thus in generation interval.

Consequence of these potential differences in selection in males and females is that the difference in selection paths need to be considered when determining the genetic gain. The way to handle these different selection paths is fairly straight forward: just split the equation in a part for males (m) and a part for females (f):

$$\Delta G\ per\ year = \frac{R_m + R_f}{L_m + L_f} = \frac{i_m r_{IH,m} \sigma_a + i_f r_{IH,f} \sigma_a}{L_m + L_f}$$

The intensity of selection and accuracy of selection and generation interval may be different in males and females. The genetic standard deviation, however, is a population parameter and is, therefore, the same between males and females.

Chapter 9.9.1: an example: beef cattle breeding

Beef cattle breeding is rather small in the Netherlands. In countries like Australia or the USA, or within Europe countries like France or the UK, beef cattle breeding is a much larger business. The cows graze on large areas of land and are not handled at regular basis. Therefore, AI is not a very useful tool in the reproduction. Most farmers buy bulls and let them graze with the cows. The very large farms also breed their own bulls.

Consider a population of beef cattle that is selected for increased growth. The heritability is 0.35 and the phenotypic standard deviation ($_p$) is 0.2 kg / day. The females are selected on their own performance. As the population size is supposed to remain constant and females can produce about three calves in their lives, 2/3 of the females need to be selected to produce sufficient animals for replacement (remember that both male and female calves are born!). A selected proportion of 0.67 results in a selection intensity (i_f) of 0.54. The accuracy of selection for selection on own performance is equal to h^2 so $r_{IH,f} = 0.59$.

*The males are selected based on the performance of 100 progeny, resulting in a r_{IH} of 0.95. Each male is mated to 10 females, resulting in a selected proportion of 0.10 * 0.67 = 0.067. The selection intensity thus is 1.95 (check in the table).*

*Finally, the genetic standard deviation is equal to the square root of h^2 * = 0.35 * 0.2^2 = 0.118. What is the genetic gain in this population?*

Filling all that information into the formula results in a genetic gain per generation of:

= 0.257 (kg/day)

Genetic gain per generation does not provide the insight in the genetic improvement that was intended. To achieve that, the genetic gain per generation needs to be scaled to genetic gain per year. The average age of the females when they produce their average offspring is 4.5, so $L_f = 4.5$. Males are selected after progeny information has become available, resulting in a generation interval of 5 yrs. Predict the genetic gain per year in this population.

Scaling the genetic gain per generation to the generation interval results in a genetic gain per year of:

$$\Delta G \; per \; year = \frac{1.95*0.95*0.118+0.54*0.59*0.118}{5+4.5} = 0.027 \text{ (kg/day) per year}$$

Thus:

Selection intensity and accuracy of selection may differ between males and females. Selection response in each of these selection paths are calculated separately, and afterwards combined into a genetic gain for the entire population.

Chapter 9.10: More detailed selection paths

Selection paths in males and females can be very different for a number of reasons. Especially because males in general can produce (many) more offspring than females, the contribution of males to the next generation can be considerably larger than that of the females. For that reason, in many species more attention is paid to selecting the males as accurately as possible. Females often are allowed to breed without, or with weak, selection criteria. In many animal species, the deviation into different selection paths is taken one step further and discriminate between males that are selected to produce new breeding males, and males that are selected for producing females. The same can be applied to the female side: females to breed new breeding males, and females to breed females. Males that have been selected for breeding are called sires, and females that have been selected for breeding are called dams.

We can thus define four selection paths:

1. Sires to breed sires (**SS**)
    a. This is the most stringent selection path to breed new fathers of the fathers. Only elite sires make it to sire father.
    b. Sires to breed dams (**SD**)
        i. Within the sires this is a less stringent selection path. These sires will be the fathers of the breeding females (the dams).
        ii. Dams to breed sires (**DS**)
            1. This is the most stringent selection path within the dams to breed new sires. Only the elite dams will make it to sire mother.
            2. Dams to breed dams (**DD**)
                a. This is the least stringent selection path. It depends on the studbook whether there are selection criteria for new dams.

Combining these selection paths into a formula:

$$\Delta G \ per \ year = \frac{R_{SS} + R_{SD} + R_{DS} + R_{DD}}{L_{SS} + L_{SD} + L_{DS} + L_{DD}}$$

Thus:

Selection response can be divided into a number of selection paths, the number depending on the number of differences in selection intensity and the accuracy of selection

Chapter 9.10.1: An example: dairy cattle breeding

On dairy farms the cows are handled at daily basis (usually twice a day in the milking parlour). Mating often occurs through use of AI because it allows farmers to concentrate on their cows and not to handle a bull. Introduction of AI in dairy cattle has caused a clear deviation in tasks and ownership: cows tend to be owned by farmers and the breeding bulls by a breeding company. The farmers and the breeding company work together because they need each other' s input in their business. Dairy cattle is a species where all four selection paths can be recognised. On the one hand there are the bulls. Proven bulls can all be used to produce new cows, that is what they are selected for. But of the bulls with the very highest EBV also sons are retained as selection candidates to become breeding bull. Other male calves are sold for the meat industry. The larger part of the cows are used for breeding new cows ('replacement'). The very best of cows are mated to the very best of bulls to produce new bulls: they become bull dam. On the other side of the spectrum there are some cows that are considered of insufficient quality to produce replacement cows. These cows are selected out of the breeding program and are often mated with a beef breed to produce calves of some value. The cows are milked until they are replaced.

Consider a small dairy cattle population of 2000 cows in a time when genomic selection was not an option. Of those cows, 80% are used to produce replacement cows. Sexing of sperm is not available, so 50% male and 50% female calves are born. Of the male calves that are born, 1.5% is selected as breeding bull, and 0.25% as bull sire. Of the female calves born, 3.5% is selected as bull dam. Cows are on average 4 years old when they produce their calves, bull mothers on average are 5.5 when they produce their potential breeding bull sons, breeding bulls on average are 6 when they produce their average calves, and bull sires are on average 8 when they produce their average potential breeding bull sons. Animals are selected on their EBV for milk production. The EBV for cows was based on own performance records, the EBV for breeding bulls on the performance of 10 of their daughters, and the EBV of bull sires was based on the performance of 20 of their daughters. The heritability for milk production is 0.3, and the genetic variance in this population is 122,500 kg.

Calculate the genetic gain per year.

Answer:

This question requires a stepwise approach. There are four different selection paths and for each selection path we need to calculate the *i* and the r_{IH}. As we need to calculate the genetic gain per year, we need to divide by the generation interval. They have already been defined in the text. Let's take this path by path.

First the SS path. The selected proportion is 0.8*0.5*0.0025 = 0.001 (0.1%), which coincides with an *i* of 3.367. How to calculate the r_{IH} is in table 2 of the chapter on ranking the animals. The formula is:

$$\sqrt{\frac{\frac{1}{4}n\,h^2}{1+\frac{1}{4}(n-1)h^2}}$$

If we fill that in (20 daughters, $h^2 = 0.3$) we come to an r_{IH} of 0.619.

In the SD path the selected proportion is 0.8*0.5*0.015 = 0.006 (0.6%), resulting in an i of 2.834. The r_{IH} (10 daughters, h^2=0.3) is 0.448.

In the DS path the selected proportion is 0.8*0.5*0.035 = 0.014 (1.4%), resulting in an i of 2.549. The r_{IH} (own performance) = h^2 = 0.548.

In the DD path the selected proportion is 0.8*0.5 = 0.4 (40%), resulting in an i of 0.966. The r_{IH} is the same as for DS: 0.548.

Now we have all the calculations for the separate selection paths and we need to combine them into an overall genetic gain. The genetic standard deviation is 122,500 = 350 kg.:

$$\Delta G = \frac{3.367*0.619*350+2.834*0.448*350+2.549*0.548*350+0.966*0.548*350}{8+6+5.5+4} = 78.64 \text{ kg}$$

This population is expected to produce 78.64 kg more milk per year on average.

Note that this is not a realistic example as in reality there are all sorts of complicating factors such as that animals are selected based on more than one information source, for example on own performance AND on information on sibs, AND on progeny. The older an animal gets, the more information becomes available and the more accurate the breeding values will become. In addition, we assume that breeding occurs within generation. However, in reality there is overlap between generations. Some animals are used for breeding much longer than others.

Chapter 9.11: Selection intensity and rate of inbreeding

From the previous it has become clear that decrease in selected proportion, so an increase in selection intensity, results in an increase in genetic gain. Fast genetic gain can thus be achieved by selecting only the very few very best animals for breeding. That is simple, why not do that? Apart from the fact that the reproductive capacity will determine the minimum number of animals that need to be selected in order to maintain population size, there is another important issue: inbreeding. Remember from the chapter about genetic relationship and inbreeding that the rate of inbreeding in a population can be predicted by $1/8N_m + 1/8N_f$ and

that a smaller number of parents thus results in a higher rate of inbreeding. Especially in case of unbalanced numbers of males and females. If we use the recommendation of the FAO not exceed the rate of inbreeding of 0.5 to 1% for the population to remain viable, this may have consequences for the selection strategy.

In large populations, for the same rate of inbreeding, larger selection intensities can be applied than in small populations. For example, in a population of 20,000 animals (half male, half female) a selected proportion of 1% would result in 100 animals. Equal selected proportions in males and females would result in a rate of inbreeding of 0.25%. If, however, this population was not 20,000 but only 2000 animals large, then a selected proportion of 1% in males and females would result in a rate of inbreeding of 2.5%, which is too large. Often the selected proportion in males is (much) smaller than in females. If we take the population of 20,000 animals again, a selected proportion of 0.1% in males (select the best 10 males) and use all 10,000 females for breeding, results in a rate of inbreeding of 1.25% (1.25125 to be precise). Despite the fact that 10,010 animals are used for breeding, the rate of inbreeding is still too high for the population to remain viable. And with AI 1000 offspring per sire in most species is not a problem, provided sufficient number of females are available.

Breeding companies earn their income from selling genetic material (usually sperm). So it is in their interest to keep their breeding stock viable. However, there are competing breeding companies who want to provide the same market of genetic material. Breeding companies solve this conflict by trying to make as much genetic progress as possible to keep (or increase) market share, but restrict the rate of inbreeding to 1%.

Thus:

> *Decisions on the intensity of selection depend on the consideration of genetic gain versus rate of inbreeding*

Chapter 9.11.1: Special case: indirect selection

Thus far we have assumed that phenotypic information is available for the trait under selection on at least a large part of the population. If selection is on growth then weights at different ages are easily made available, if selection is on milk production then production records of the females are available, and even if selection is on meat quality then records are available on relatives. However, in some cases phenotypes are not available, for example in case of an infectious disease, and/or phenotypes for traits that are expensive or invasive to collect. In those cases it is possible that a second trait can be used as *indicator* of the trait that you want to select for. Important prerequisite is that the indicator trait is correlated with the trait that you want to improve (i.e. the trait in the breeding goal). Obviously, the higher the correlation, the better. In formula the genetic gain in the breeding goal trait, given selection on the indicator trait, can be predicted as:

$$\Delta G = i * r_{IH\ indicator\ trait} * \sigma_{a\ breeding\ goal\ trait} * r_{indicator\ trait,breeding\ goal\ trait}$$

The selection intensity is the same as with direct selection and depends on the selected proportion. The accuracy of selecting the breeding animals is predicted using the heritability for the indicator trait, as that is what you base your selection on. You are interested in the response to selection in your breeding goal trait, so you want to express the results in units of the breeding goal and thus use the genetic standard deviation of the breeding goal trait. It will depend on the size of the correlation to what extend selection on the indicator trait indeed will result in genetic progress in the breeding goal trait. You, therefore, have to multiply the result with the correlation between indicator and breeding goal trait. The overall accuracy of selection thus both depends on the accuracy of selection on the indicator trait, and the correlation between indicator and breeding goal trait.

Note that it depends on the combination of $r_{IH,\ indicator\ trait}$, determined by the heritability of the indicator trait, and the correlation between breeding goal and indicator trait whether more genetic gain can be achieved with indirect compared to direct selection.

Thus:

An indicator trait provides an indication of the performance for the breeding goal trait, and is useful as replacement of traits that are very difficult or expensive to measure.

Requirements for success are the heritability of the indicator trait and the correlation with the breeding goal trait.

Indirect selection can be a very good solution when traits are difficult or expensive to record.

An example: white line disease in dairy cattle

Results of research on a large number of Dutch dairy cattle herds in 2002-2003 indicated a prevalence of white line disease of 9.6%. The heritability was low, only 0.02% (so the r_{IH} was 0. 02 = 0.14). Important reason for the low heritability was that the farms were visited only once, and any cow that was not diseased was considered healthy. However, some cows may have been only just recovered or not been infected. The genetic variance was 0.078, so $_a$ was 0.28. These are not very promising figures to achieve large genetic gain with. However, the genetic correlation between white line disease and foot angle is 0.64. Steep angled feet are affected more often than flat feet. Foot angle is an easy to measure trait with heritability of 0.18 (so r_{IH} of 0.42). If we would select on observations related to white line disease, then the genetic gain would be *i* * 0.02 * 0.28 = *i* * 0.040. If we would select on the foot angle to decrease the prevalence of white line disease, then the genetic gain would be *i* * 0.18 * 0.28 * 0.64 = *i* * 0.076. This is almost a double response if selection is indirectly on foot angle instead of directly on white line disease prevalence!

Chapter 9.12: Practical issues with predicting response to selection

Thus far we have been discussing the optimal situation in which someone decides on which animals are allowed to breed and which are not. There are two main points where the breeder has influence: the selected proportion and the accuracy of selection. For good predictions of the genetic response it is essential that these selected proportions and accuracy of selection are correct. How realistic is that?

This is fairly realistic in case this someone owns all potential breeding animals, such as in commercial pig and poultry breeding. The selected proportion may depend somewhat on the expected market situation, but at least changes are recorded. The accuracy of selection of the breeding animals also is in the hands of the breeding company. Animals are selected based on performance of other animals (sibs, offspring), and the exact number depend on the number available and may vary a bit between animals. The influence of this small fluctuation on the predicted genetic gain will be very limited. But even in those breeding companies there may be unexpected practical events, such as a disease outbreak, that prevent selecting the intended proportion. However, in general in these species the prediction equations as presented in this chapter are very useful.

In dairy cattle breeding the situation becomes a bit more difficult as the large part of the cows are owned by individual farmers. Each farmer will have his or her own breeding goal, though in general terms these will resemble that of the breeding company, who owns the bulls. The selected proportion in bulls is in the hands of the breeding company, but the subsequent use of those bulls is in the hands of the farmers. There are popular bulls and not so popular bulls. Even though both are selected for breeding, the popular bull will have a larger number of offspring in the next generation than the not so popular bull. The assumption when predicting genetic response to selection is that all selected bulls will have equal chance of ‘spreading their genes’. This obviously is not the case. Depending on which bulls are used more often than expected, this will lead to an over or under prediction of the genetic response. Popular bulls obviously also will have a more accurate EBV than not so popular or young bulls. This difference in accuracy needs to be taken into account when predicting the response to selection. On the females side there may be very little effect of selection in cows because of two reasons: first of all the selected proportion is very large as most cows are used to produce replacements. Second, the farmers may have slightly different selection criteria that result in an overall even smaller effect of selection in the cows. In practice, this selection path can be ignored.

In horses the situation is a bit more complicated than in dairy cattle. The stallions are approved for breeding if they live up to the standards that are defined by the studbook. However, the approved stallions do not necessarily represent a well-defined selected proportion because not all males foals are selection candidates because not all owners are willing to present their colt (young stallion) at the stallion inspection. The selected proportion based on the number of colts

selected in the stallion inspection thus may not represent the true selected proportion. As in dairy cattle, some stallions will be much more popular and, consequently, will have more offspring in the next generation than others. As in dairy cattle, the accuracy of selection will depend on the information that is available and may differ between stallions. That can be taken into account when predicting the response to selection. In most studbooks, as in dairy cattle, all mares are allowed to breed. However, different from dairy cattle is that it are not necessarily only the best mares that are used for breeding. Some owners of a very good mare do not want to breed with her, and some owners of poor mares do want to breed a foal. In mares it seems a fair approximation that the selected proportion is 100%. Accuracy of selection generally will be low because of the generally limited number of foals. Some mares with a popular father potentially have a large number of half sibs.

In dogs the situation is more complicated again. The breeding associations and the ‘Raad van Beheer’ have defined some basic prerequisites for males to be allowed to breed. Females in some breeds are also required to present a health certificate related to some potential breed specific health issues. So far so good. However, in males there is no selected proportion because very few people enjoy having a breeding male, and this is not necessarily related to the quality of the dog. Similarly, few owners of a bitch want to breed a litter, despite the quality of the bitch. Consequently, prediction of genetic response based on selected proportion and accuracy of selection is not feasible in dogs. Exception to this is the breeding of working dogs, where quality of the dog is the selection criteria and selected dogs are used for breeding.

So what to do with these situations where selected proportion and accuracy of selection cannot be defined very well? One solution can be to predict the average genetic potential in the offspring of each mating based on the EBV of the sire and the dam. As we have seen in the chapter about genetic models, knowing the EBV in the parents only gives you some idea about the EBV in the offspring due to the Mendelian sampling: $A_{offspring} = ½ A_{sire} + ½ A_{dam} + MS$. It will depend on the accuracy of the EBV in the parents to what extend the inaccuracy of the prediction of the genetic response will be further increased.

Thus:

In conclusion: predicting genetic response to selection assuming a selected proportion and an accuracy of selection is very useful, but be aware of the (lack of) accuracy of your assumptions!

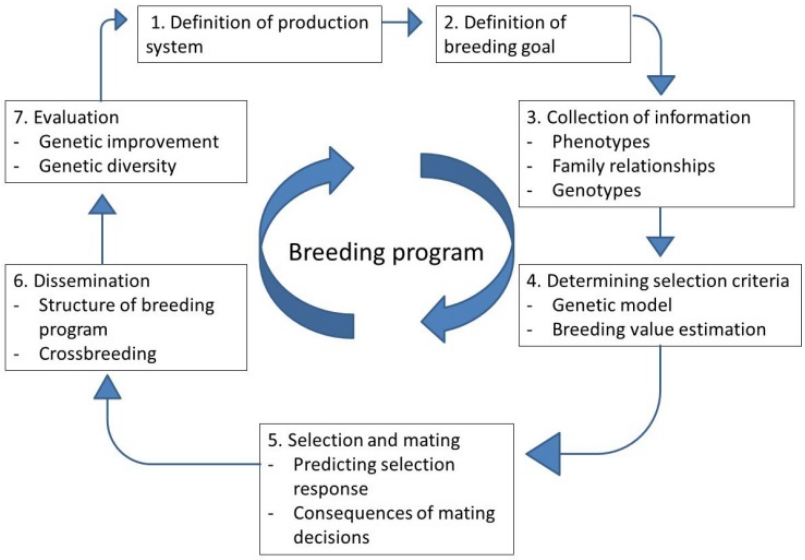
Chapter 9.13: Key issues on predicting response to selection

1. Breeding is about predicting the future.

2. Predicting genetic gain is about predicting future performance, but also about evaluating breeding decisions beforehand.
3. Genetic gain is determined by 3 main factors: phenotypic variance, accuracy of selection, and selected proportion.
4. The generation interval allows to express genetic gain per year instead of per generation.
5. Optimising genetic gain will require a balance between increase in accuracy and decrease of the generation interval.
6. Selection paths allow for difference in selected proportion and accuracy of selection in males and females.
7. Decisions on the intensity of selection depend on the consideration of genetic gain versus rate of inbreeding.
8. Indirect selection using indicator traits can be a very good solution when traits are difficult to measure or expensive to record.
9. Predicting genetic response to selection assuming a selected proportion and an accuracy of selection is very useful, but be aware of the (lack of) accuracy of your assumptions!

Chapter 10: Selection and mating

After estimating the breeding values and predicting the influence of selection decisions on the genetic response to selection, we are ready for action: actual selection and mating of the animals! Like many aspects of animal breeding, mating can have a different effect on individual mating decisions as compared to the overall population. The breeding goal for the overall population may differ from the breeding goal of the individual breeder. However, to achieve genetic improvement at population level, it is essential that selection decisions are made at population level. Individual breeders, subsequently, can apply their selection criteria to select the mate for their animal from the selection candidates that have been identified at population level. Mate choice may depend on a number of aspects such as the intended use of the offspring, the quality (or lack of quality) of the female, the price of the desired mating, or the distance to the mate (in case of natural mating). The aim is to find a suitable mate and produce good offspring, given the limits that you set yourself with respect to mate choice. Mating decisions at the level of individual breeder may have consequences for the rate of inbreeding at population level. Because if your mate choice is the same as the mate choice of many other breeders, than the mate of your choice will have many offspring in the next generation, whereas others may have none. Desired mate choice at individual level and consequences of individual mate choices at population level, therefore, may conflict with each other.



In this chapter we will consider reasons for, and consequences of, mating decisions such as to compensate points of lesser quality in your female, or to achieve certain qualities in your offspring (e.g. colour). We will also discuss potential consequences of intensive use of popular sires on population level, and briefly about reasons for parentage testing.

A note: mating and single gene traits

Mating decisions may also be based on trying to create or avoid homozygosity. In the situation of recessive monogenetic disorders, for example, mating decisions will be aimed at avoiding the risk of homozygous recessive offspring. However, there are also monogenetic traits that have a desired effect. For example having horns or not in cattle, or to create a specific coat colour. More about that in the chapter about monogenetic inheritance.

Chapter 10.1: Selection criteria and mating decisions

In commercial farm animal breeding of pigs and poultry selection of the best animals is followed by more or less random mating. Individual mating decisions are not made because on average, at population level, there is no evidence for additional value. In other words: mating does not result in a directional change in the average of the traits under selection. If any, there may be a decreasing effect on the variation for the trait. But only if all breeders make selection and mating decisions with the same breeding goal in mind. At individual level there may be some additional value of mating decisions, especially related to monogenetic traits.

Reasons for an owner to use his/her female for breeding may differ from the need to produce offspring to initiate milk production, via striving for offspring of the highest quality, to ‘even though she is not well anymore we can always use her for breeding’. Reasons for specific choice of a mate for a female may vary from practical reasons, such as costs and travel distance in case of natural mating, via avoiding certain problems, such as heritable disorders, and compensating shortcomings, to going for the most popular mate. Selection criteria should be defined before selecting the best mate for breeding, but in practice these two processes are often interrelated. But be aware genetic improvement is created by selection and not by mating.

Owners of females make the actual decision to use an approved male for breeding or not. Effectively they are in charge of the actual breeding. Owners of the males only have a ‘product’ on the market. They may need to put some effort into marketing of their product. Quite often marketing is at least as important as quality of the males where it comes to reasons for mate choice by owners of females. A top sire determines the competitive value of a breeding organization.

Thus:

> *Mating decisions have no influence at population level, but may have some effect at the level of individual mating.*

Relation of mate selection with inbreeding

Remember that an animal is inbred if his parents are related: $F_{animal} = ½ * a_{between\ parents}$. Parents are related if they have ancestors in common. The fewer generation away this common ancestor is, the more the parents are related. Offspring of a full brother and sister have an inbreeding coefficient of ½*0.5 = 0.25. The less two parents are related, the less the offspring is inbred. Some breeding organisations have regulations to avoid mating of close relatives. The Dutch Kennel Club (overall organisation of the various dog breeds in the Netherlands and in charge of pedigree registration of all registered purebred dogs), for example, has the regulation that no pedigree will registered for offspring of a bitch that was mated to her grandfather, her father, her brother, her son, or her grandson.

Thus:

> *Mate selection should take the additive genetic relationship between both potential parents into account as that is a direct indication of the inbreeding coefficient of the offspring.*

Chapter 10.2: Compensatory mating

Mate selection may be aimed at compensating specific shortcomings. For example, a mare may have superb gaits, but her legs are not of superior quality. The stallion of choice would need to have perfect legs, preferably with evidence that his offspring also have good legs. His legs would be considered more important that a strong trot, as the mare should bring that into the offspring. Another mare may have perfect legs but could have a somewhat stronger gallop, then the stallion should have an excellent gallop but is allowed to have somewhat lesser leg quality. At least, that is the idea. Specific mate choice for each female so that the offspring is of the best possible quality. A practical advice for breeding organisations would be to first select the parents as that defines the genetic improvement, and then define a compensation mating scheme as potential advisory service.

However, even though this sounds very logical, there is no guarantee for success! It is clear that with compensatory mating breeders will have different breeding goals that they base their mate choice on. Aspects of lesser quality in one female may be the stronger point in another female. Obviously, the additive effects of all mate decisions in a population is not likely to point in a certain direction. In other words, mating decisions will result in extra genetic gain. Also at individual level there are a number of factors that may influence the expected result of mate choice:

1. Mendelian sampling. This introduces a chance factor, even if you know the EBV for father and mother very accurately.

2. Pleiotropic (one gene affects multiple traits) and epistatic effects (gene-gene interactions). It is possible that a trait, for example gait quality in horses, is affected by a gene that interacts with another gene. If one of those genes has the wrong allele in the offspring then the gaits won't be improved.
3. What is the accuracy of the information that the selection decision is based on? For example, does genetic or training make a champion? You should especially ask yourself such questions in the absence of an accurate EBV.

Thus:

Compensatory mating involves finding the best mate for individual females to compensate her shortcomings.

Compensatory mating may have an effect on individual mating results, but has not effect at population level.

Chapter 10.3: Long term genetic contribution

Why do animals have common ancestors? Because their common ancestor apparently was popular enough to have multiple offspring, which, possibly after some generations, resulted in the birth of both parents. The more popular a breeding animal was in the past, the larger the chance that two potential parents have this ancestor in common. The more animals share that common ancestor, the larger the chance that mating two animals will result in an inbred offspring. In other words, there is a relation between the long term genetic contribution of an animal to the population and the rate of inbreeding in the population. The long term genetic contribution is a measure of the level of relatedness between animals in a population because of a shared common ancestor. To illustrate this concept, consider a male that became very famous because he won an important competition. Many breeders decide to use him as mate for their females. In the next generation it becomes clear that he was a champion for good reason because a number of his sons also perform (much) better than average, so they are also used for breeding relatively often. In the next generation, again some sons of these sons are better than average and are used a lot for breeding. You can imagine that within a number of generations, a very large proportion of the animals will have that first champion male as ancestor. If those animals are mated, they are inbred to that champion.

Definition

Long term genetic contribution is a measure of relatedness in the population because of a shared common ancestor.

There is a relation between the long term genetic contribution of an animal to the population and the rate of inbreeding in the population.

If you would consider your own pedigree and that of your neighbour, you will find common ancestors if you go back far enough in time. Eventually, everyone is related. The point is that how close you are related will depend on the number of offspring that your common ancestor(s) produced and the number of generations between the common ancestor and the present individuals. The more offspring, the more 'paths' there are to the common ancestor, and the larger the proportion of genes you share with that common ancestor. The same principle can be applied to animal breeding, though in general the process of shared genes from common ancestors goes a lot faster because the breeding population is relatively small, and introduction of animals from outside the population usually is not practised. After a number of generations, (almost) all animals are related to that common ancestor. After some more generations the contribution of the common ancestor does not change anymore: (almost) all animals have the same proportion of genes of that ancestor in common.

Chapter 10.3.1: Example of genetic contributions

This concept is illustrated in figure 1, where the genetic contributions of 2 sires across 6 generations is shown by the proportion of colour in a box. Each sire was mated to 10 dams who had 10 offspring. So in total there were 100 offspring per generation, of which 50 were male. Out of those 50 males only 5 were selected for breeding. And selection was irrespective of the parents, so good sires have more offspring that were used for breeding than poor sires. The contribution of the original sire to the offspring is indicated by proportion of colour in the boxes. The sire in the picture on the left was a very successful sire. His offspring was very popular and two of his sons were selected and used a lot. Again their sons and grandsons were popular and, consequently, the proportion of the original sire in generation 6 is large. This in contrast to the sire in the figure on the right. He was not very successful and none of his sons were selected for breeding. Only in generations 4 and 5 males were selected in this family. Consequently, the genetic contribution of this original sire in generation 6 is very small. Note that once the genetic contribution of the original sires have stabilised in the population it is no longer possible to change the size of the contribution in the population.

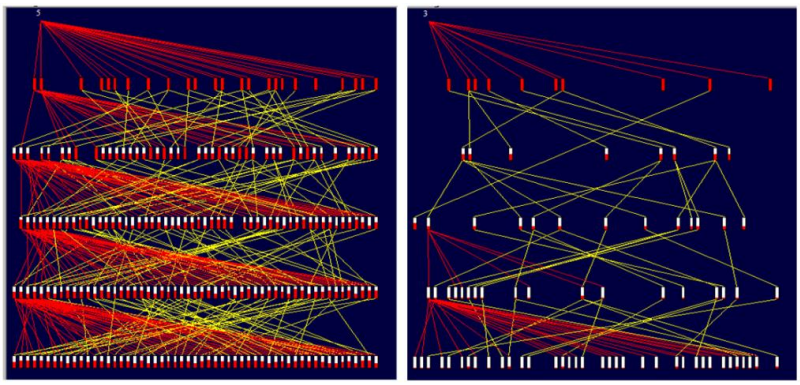


Figure 1. Example of genetic contributions of 2 sires in the same population across 6 generations, males are in red and females in yellow. Initially there were 5 sires mated to 10 dams each that produced 10 offspring each. The picture on the left shows the contribution of a genetically superior and popular sire. In generation 6 in all animals a considerable proportion of their genes originated from this sire. The picture on the right shows the contribution of a much less popular sire. Two of his children never produced any offspring themselves. In generation 6 the contribution of his genes is very small. (the pictures were created using the freeware program GENUP, by Brian Kinghorn)

Chapter 10.3.2: Relation between genetic contribution and inbreeding

In other words, decisions about mating intensity of selected parents made today can have large consequences for future generations. Once the contribution of an animal has spread through the population you can never delete it again. It is possible the take the genetic contribution of each breeding animal into account when predicting the rate of inbreeding. This is an accurate method for predicting the rate of inbreeding, but only if the contributions of the animals that founded the population are considered. When only contributions of more recent animals are considered the relatedness between those animals is ignored. Consequently, in that case the resulting rate of inbreeding will be an underestimation of the true rate of inbreeding.

In formula the influence of the genetic contribution of animals to the rate of inbreeding looks like:

$$\Delta F = \frac{1}{4}\sum c^2$$

Where F is the rate of inbreeding due to the genetic contribution of the animals that are considered, and c^2 is the contribution of an animal to the next contribution, squared. This formula can be used to predict the rate of inbreeding in the future, given mating intensity decisions today!

Thus:

Mating intensity can have irreversible consequences for the rate of inbreeding in future generations. Large genetic contributions of genetically superior animals will spread through the population and remain there as fixed proportions in each animal.

Chapter 10.3.3: Example of the effect of a popular ram on inbreeding

In figure 2 is a more applied (though quite extreme) example of a small sheep population that was initiated with 5 breeding rams. Some rams were more popular than others and there was one particular ram (ram 2) that became champion and was very popular with the breeders. He was used in 45% of the matings, whereas the contribution of the others varied between 10 and

20%. Apparently this popular ram became champion for good reason, and his descendants were also very popular. Within 6 generations, the animals in the population on average shared 60% of his genes. This slowly increased to 66% in generation 25, after which it remained constant: on average all animals shared 66% of their genes with this ram 2. Most likely there were no more animals that were not related to ram 2 in generation 35. Note that the contributions of the other 4 original rams are also still present. Also all animals will have some of their genes in common with these rams. Apparently, all of them had some genetic quality that they passed on to their offspring, that again were selected for breeding etc. However, the contribution of ram 2 is much larger than that of the other rams of that generation, indicating a larger contribution to the rate of inbreeding, and a larger risk of homozygosity for the recessive alleles that ram 2 was carrier of.

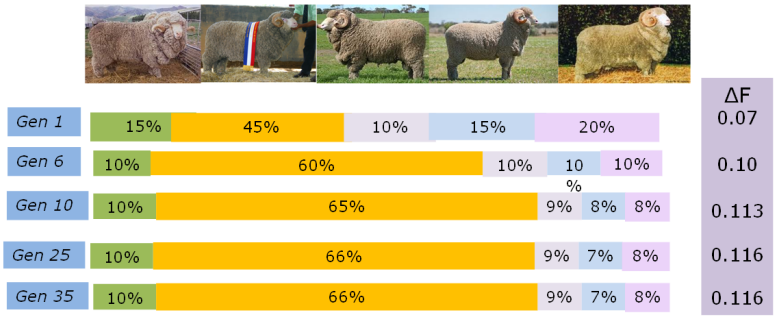


Figure 2. Overview of genetic contributions of breeding rams that were available in generation 1 across a large number of generations, and the consequence of those contributions for the rate of inbreeding in those generations. Note: this is the rate of inbreeding due to the genetic contributions of these rams only, genetic contributions of other animals are not taken into account.

Important message: intensive use of popular individuals may lead to a conflict of interest in the short and in the long run. In the short run everyone wants to be able to use that superior male because everyone wants to create the largest chance to breed a new champion. It also creates considerable income for the owner of the superior male. However, in the long run this may have negative consequences for the population, and thus also for individual breeders. May have, because if the popular sire appeared to be a poor sire based on his offspring, his sons will have a small chance of being selected for breeding, and the contribution of the popular sire thus will relatively remain small or die out.

Thus:

There is a conflict of interest between short term benefit and long term cost of intensive use of genetically superior animals.

Chapter 10.4: Breeding limitations

Some breeding associations have regulations with respect to the mating intensity of the breeding males. Aim of these regulations is to control the genetic contributions of the males to the future generations, and thus the rate of inbreeding. To realize this they strive for equal contributions of each selected male. In general these breeding limitations are not appreciated by the breeders because their interest mainly lays in the short term result: they want to use the superior male for breeding or they want to earn money from selling matings. Those personal interests generally outweigh the long term interest of the entire population. People tend to believe the regulations are good to have, as long as they don't influence their interest. Why should they restrict themselves where others were allowed to breed with that male? Why should they not be allowed to make money of selling breeding to earn back some of the costs they made for getting that approved male? You can imagine that this especially is an issue for species where males are privately owned, such as in dogs, horses, or sheep. In dairy cattle the males are owned by a limited number of AI service stations and they can sell semen from other males instead.

Example: breeding limitations and rate of inbreeding in Friesian horses

The Friesian is a relatively large (i.e. in numbers) Dutch horse breed. However, it has not always been large. Originally it was used as farm horse in Friesland. During the week it was mainly used to pull farm equipment and on Sunday it was a used to show off when it pulled the 'sjees' (carriage) to church. In the 1950's the tractor was gaining popularity and the Friesians, like other farm horses, were out of a job. The result was that the population decimated in size. In the 1980's the Friesian became popular again as horse to be used in leisure sport (both in harness and under saddle). The increased popularity, in combination with the availability of AI, resulted in a rapid increase in population size based on relatively few stallions. The result was a rate of inbreeding in 2000 was 2%. There we signs of inbreeding depression like a reduction in semen quality and mare fertility. There was also an increase in frequency of genetic disorders such as cryptorchidism, placenta retention, dwarfism, and hydrocephalus. The breeding organisation decided that action was required and in 2003 the number of breedings per male is limited to maximal 180 breedings per stallion per year for a maximum of 6 breeding seasons. From the age of 9 years onwards the limitations are removed. These breeding limitations have economic consequences, especially for the stallion owners, but are positive for the future perspectives of the population as a whole. In 2013 the rate of inbreeding was reduced to 0.5%.

Chapter 10.5: Genetic contributions and occurrence of recessive disorders

You can ask yourself why is it risky to use a superior animal excessively? The answer is because of the increasing effect it has on the rate of inbreeding. The rate of inbreeding is an indication of the risk of increased frequency of genetic disorders. It has been estimated that every individual is carrier of approximately 25 recessive disorders, most of them still unknown, and irrespective of animal species. Genetic contributions are directly related to the rate of inbreeding. Large contributions increase the risk of becoming homozygous for these genetic disorders in the future.

To get some insight in how genetic disorders spread through the population we will consider the ‘birth’ of a new disorder: the occurrence of a mutation with negative but recessive effect. It may take quite some time before this new mutation is noticed because at first it only occurs in a few offspring (approximately 50%) that became carrier. In the next generation usually it still will only occur as carrier because often brother-sister matings are prohibited. So it will take another generation for homozygous animals to occur, and only if matings between generations (parent-offspring or uncle-niece type of matings) are allowed. Even then the number of homozygous recessives will be small so, depending on the type of defect, it may still go unnoticed. Only in the fourth generation after the mutation occurred there will be animals that are homozygous because of mating between not very close relatives. It will depend on the severity of the disorder whether it will be recognised as such in the first generations of homozygosity of the disorder. Especially if the disorder does not lead to very severe problems it may go unrecognised for a very long time. By the time it is recognised the allele frequency in the population already may be quite substantial.

In Figure 3 is a numerical example to create some feeling for the chance of timely detection of a mutation with negative effect. Underlying this table are a number of assumptions related to number of offspring (each animal will have 10 offspring) and avoiding close inbreeding (no brother-sister or parent-offspring mating). Given those assumptions, the table shows that a new mutation takes about 4 new generations before there is some chance of being noticed. It will only be noticed if the consequences of the mutation are very negative AND cannot be attributed to something else. For example, if the mutation has a negative effect on embryo survival, it will take much longer before it is realised that the seemingly lack of fertility is due to embryo mortality and not due to other reasons like poor sperm quality. To come back to the numerical example: At some point a negative recessive mutation occurs. The offspring of the animal will partly be carrier: half of the offspring inherited the mutated allele and half to wild type: In generation 1 5 animals are carrier and 5 are not. All other animals in the population are wild type, they are not in the table, but can be used for mating, as happened in generation 1 to create generation 2 (no mating between sibs was allowed). Again 10 offspring per animal, resulting in 25 carriers and 75 wild type. In the next generation again mating was only with wild type resulting in 125 carriers and 825 wild type. Then mating between carriers is allowed, but still not between sibs. So out of the 125 carriers only 100 are allowed to mate. IF ALL OF THESE ANIMALS MATE to each other (so only carrier mates to carrier), then the number of affected animals would be 25, out of 10,000! So given all assumptions, that are fairly realistic although the 10 offspring per animal may be a bit much, after 4 generations only a maximum of 0.25% of the animals are affected. If the effect of the genetic disorder is not very extreme or unusual, it will take many more generations before people realise that the number of affected animals is increasing and perhaps could it be heritable??

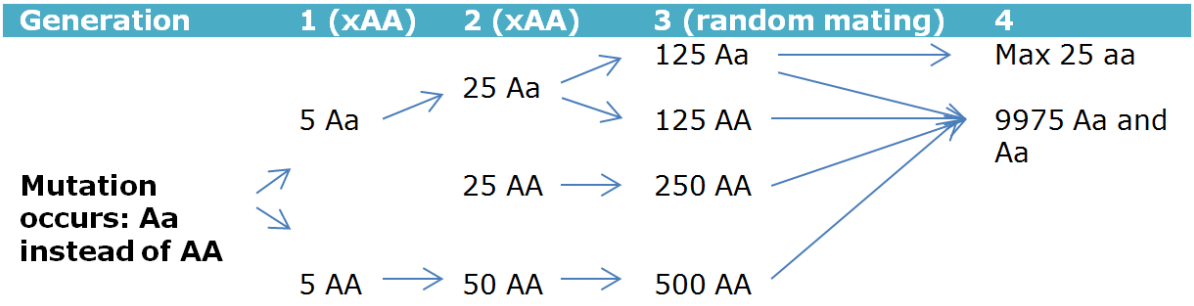


Figure 3. Numerical example of how long a mutation with negative effect can remain unnoticed. Important assumptions: 10 offspring per animal per generation, random mating within generation, but not between sibs (full and half). Note that the population most likely is larger than this because these are only the descendants of a single animal.

This was only a numerical example, but these things really happen! A famous (or rather infamous) example is that of the heritable disorders BLAD (bovine leucocyte adhesion deficiency) and CVM (complex vertebral malformation) in Holstein dairy cattle. A very large genetic contribution of a single bull resulted in two heritable disorders that were spread widely though the Holstein population.

Chapter 10.5.1: Heavy use of a sire for the frequency of genetic defects

In the 1980's of the previous century there was a bull named Carlin M Ivanhoe Bell (nickname: Bell). He was very popular because his daughters were very good milk producers, and he was heavily used as sire for two decades. His sons were also heavily used and today most Holstein cows are inbred to Bell, often through more than one pathway.

Unfortunately, many generations later it appeared that Bell was carrier of two genetic disorders: BLAD and CVM, and because he was used so heavily as sire, the disorders spread throughout the Holstein Friesian dairy cattle population. The disorder BLAD causes immunodeficiency, resulting in recurrent infections. It is traced back to Osborndale Ivanhoe, the grandfather of Bell. The disorder CVM often (88% of the cases) results in abortion within 260 days after insemination. Only 4 to 5% of the foetuses are born alive. It was traced back to Pennstate Ivanhoe Star, the father of Bell. CVM was only discovered in 1999.

So even though the mutations for neither BLAD nor CVM originate from Bell, but possibly from his father and grandfather, the very large genetic contribution of Bell caused the extensive spread of both disorders through the Holstein Friesian dairy cattle population. Neither of the disorders was serious enough to link to a genetic disorder already at small frequencies. So both of them were allowed to spread widely before the link to a heritable disorder was made.

Thus:

Speed of spread of recessive genetic disorders through a population is increased with intensive use of breeding animals.

The disorder will have spread widely before it is recognised as such.

Chapter 10.6: Confirmation of parenthood

In many breeding associations the mating needs to be confirmed by DNA test when the offspring is born. Important reason is to protect the quality of the pedigree recording. There are a number of reasons why pedigree recording may be of poor quality:

1. In some group mating systems (e.g. fish) where multiple males are females are housed together. The exact pedigree of offspring will only be revealed after DNA check.
2. When many offspring are born on (approximately) the same day, tagging of the animals is not always done vary accurately and the tag with pedigree registration for one animal may end up on the wrong animal.
3. In large pasture systems (e.g. in New Zealand) of dairy cattle, where the cows have been inseminated with different bulls in a period of six weeks, it is not always clear which calve belonged to which cow after the daily collecting the calves born without supervision at the pasture. In that case both parents are uncertain and parentage needs reconstruction based on what information is available on the DNA of each of the parents.
4. When matings fail females are mated again. Sometimes to the same male, sometime to another male. As sperm may survive in the female for some time, a DNA test will reveal which male is the father.
5. At the start of a breeding program, when little or no pedigree is available. An extensive DNA test (large SNP set for example) can be used to test how related two animals are with incomplete pedigree.
6. To discourage false matings. Especially in situations where the actual mating is expensive, the owner of the male may be tempted to use an alternative when the number of matings is exceeding the potential of the male. Or when the fertility of the male is insufficient. Especially in the past, before the DNA test became available, some of these practices did occur in horses.

Thus:

> *To keep the pedigree records accurate the parentage of the offspring can be DNA tested. There are a number of situations where DNA testing is the only way to confirm parentage.*

Chapter 10.7: Key issues on selection and inbreeding

1. Mating involves finding a suitable partner among the selected parents and produce offspring.
2. Mating decisions have no influence on genetic improvement at population level, but may have some influence at individual level.
3. Compensatory mating involves finding the best mate to compensate the female's shortcomings.

4. Mate selection should take into account the additive genetic relationship between the potential parents as that reflects directly the inbreeding coefficient in the intended offspring.
5. Unbalanced mating intensity may have irreversible consequences for the rate of inbreeding in future generations.
6. There is a conflict of interest between short term benefit (profit for producer and breeder) and long term cost (inbreeding related problems) of intensive use of genetically superior animals.
7. DNA tests can be used to confirm parentage of offspring.

Chapter 11: Crossbreeding

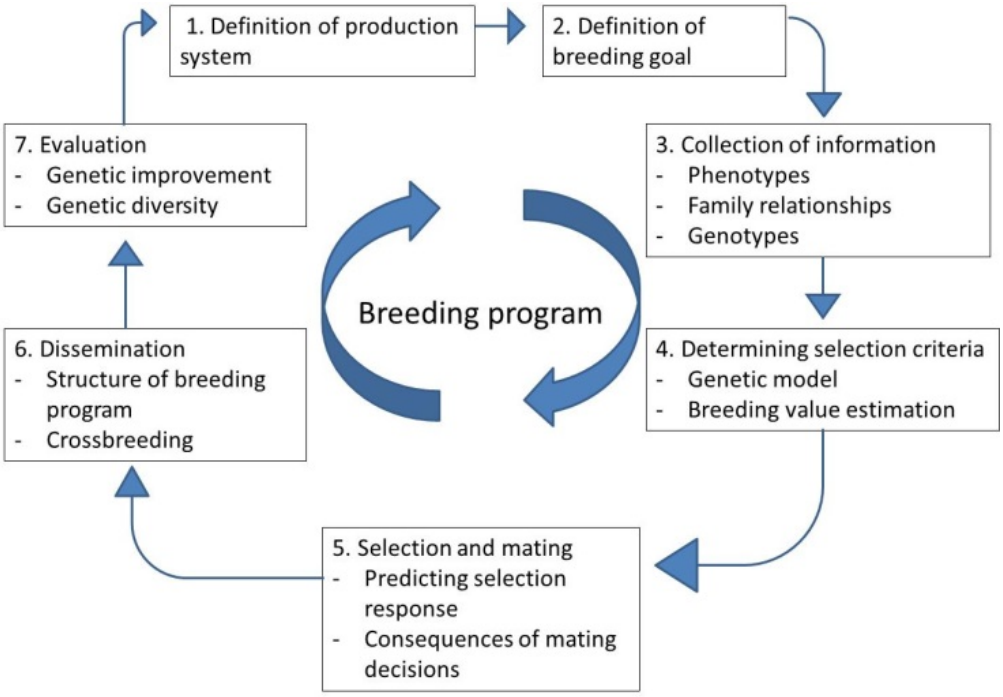
Bakewell (see chapter 1) was not only the founder of practical breeding but also the founder of the first standardized breeds. Before 1850 mankind used landraces in farming and other activities. It were populations very well adapted to the places they were kept. But their characteristics, their phenotypes were highly variable and the characteristics of their progeny were hardly predictable. Bakewell bred the first standardized breeds according to a breeding standard for conformation and a defined breeding goal. In this way selection for a few characteristics in a landrace led to the creation of a (standardized) breed.

Definitions

A breed *is an interbreeding group of animals within a species with some identifiable common appearance, performance, ancestry or selection history. Many definitions are used to define this concept. See 11.1 for more details.*

Crossbreeding *are matings between animals of different breeds or lines*

Crossing of animals of different breeds is systematic mating of animals of different breeds as part of a well-designed breeding program. What are the pro's en cons of crossbreeding in commercial beef, poultry and pig breeding? In this chapter we first explain the theoretical background of crossbreeding and the different crossbreeding systems before we outline the structure breeding programs in the next chapter. Crossing takes place after the selection of parents in different breeds or lines (phase 5 in the figure below) and is structurally embedded in a breeding program (phase 6).



Many standardized breeds in horses and dogs are the result of crossing animals of different breeds (landraces or standardized breeds) in combination with a strong selection among the crossbred animals for characteristics of the breeding standard. Thus, consequent selection for specific characteristics has led to the existence of a wide variety of breeds within the domesticated species. Breeds differ in characteristics and for specific production goals, combining characteristics of different breeds might be required. For that reason sometimes breeds are crossed. E.g. in tropical countries a local cattle breed with a high tick resistance is crossed with an exotic breed with a high production to obtain animals with a moderate production resistant to ticks.

In this chapter we will explain the following topics:

- The genetic background of heterosis
- Motivation for crossbreeding
- The different crossbreeding systems and their applicability

Chapter 11.1 Definitions of a breed

What is a breed?

A simple question but difficult to answer, and the following are published definitions from a variety of groups, each relevant and pertinent to their stakeholders:

i. "*Animals that, through selection and breeding, have come to resemble one another and pass those traits uniformly to their offspring.*" (http://www.ansi.okstate.edu/breeds/, 28/09/2006)

ii. "*A breed is a group of domestic cats (subspecies felis catus) that the governing body of CFA has agreed to recognize as such. A breed must have distinguishing features that set it apart from all other breeds.*" (Cat Fanciers Association, http://www.cfa.org/breeds/breed-definition.html, 28/09/2006)

iii. "*A race or variety of men or other animals (or of plants), perpetuating its special or distinctive characteristics by inheritance.*" (*http://www.biology-online.org/dictionary/Breeds*, 28/09/2006)

iv. "*Race, stock; strain; a line of descendants perpetuating particular hereditary qualities.*" (Oxford English Dictionary, 1959)

v. "*Either a sub-specific group of domestic livestock with definable and identifiable external characteristics that enable it to be separated by visual appraisal from other similarly defined groups within the same species, or a group for which geographical and/or cultural separation from phenotypically separate groups has led to acceptance of its separate identity.*" (FAO World Watch List, 3rd Edition)

vi. "*A breed is a group of domestic animals, termed such by common consent of the breeders, ... a term which arose among breeders of livestock, created one might say, for their own use, and no one is warranted in assigning to this word a scientific definition and in calling the breeders wrong when they deviate from the formulated definition. It is their word and the breeders' common usage is what we must accept as the correct definition.*" ('The Genetics of Populations'; Lush, 1994)

vii. "*A breed is a breed if enough people say it is*." (K. Hammond, personal communication)

Continuing definition (v), FAO argue that breed is very often a cultural term and should be respected as such, a perspective clearly articulated in definition (vi), and succinctly summarised in (vii). This is acknowledged, but the concept of resemblance through common hereditary descent is a useful addition to the definition of a breed.

Source: Chapter 3. What is genetic diversity? John Woolliams and Miguel Toroin Utilization and conservation of farm animal genetic resources, 2007. Editor. Kor Oldenbroek. Wageningen Academic Publishers.

Chapter 11.2 Heterosis

Crossbreeding is not only applied with different breeds, but also with different selection lines. This is fully exploited in commercial pig and poultry breeding. These lines are formed by pure breeding or by crossing different breeds. After the formation of the line the animals in such a selection line are selected for a limited number of breeding goal traits. After generations of selection they excel in these specific breeding goal traits. When such lines are crossed, the crossbreds not only combine the characteristics of each of the lines, but for some characteristics the performance of the crossbreds is higher than the average performance of the parent breeds due to heterosis.

Definition

Heterosis or hybrid vigour (term mostly used in plant breeding) is *the extent to which the performance of a crossbred in one or more traits is better than the average performance of the two parents*

Chapter 11.3 The genetic background of heterosis

To start with a simplified example: suppose (not valid in the real world) that in chicken a single gene with two different alleles A and a determines the amount of eggs produced in a year.

Breed 1 is homozygous (fixed) for A: all animals have the genotype AA.

Breed 2 is homozygous (fixed) for a: all animals have the genotype aa.

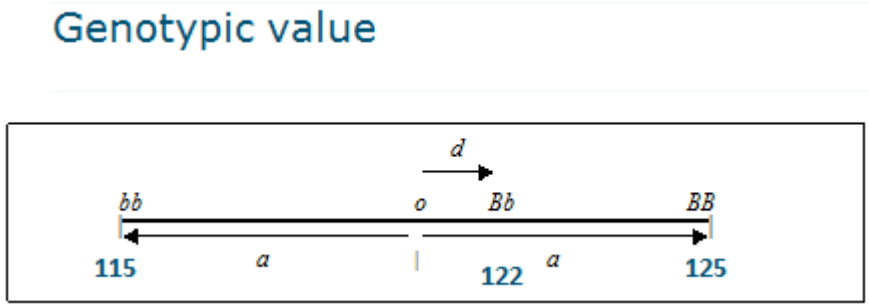
Breed 1 is producing 96 eggs per year and breed 2 94 eggs per year. Cocks of breed 1 are mated to hens of breed 2.

Their offspring with genotype Aa is expected to produce 95 eggs per year, the average of the two parent breeds, but they do produce 100 eggs per year. This is the effect of heterosis: the performance of a crossbred (Aa) is better than the average performance of the two parental breeds (AA and aa). The heterosis is 5 eggs per year of expressed as a percentage 5/95 = 5.2 %. Heterosis is based on the phenomenon of dominance: the genotype Aa has a value that is higher than the average of the genotype AA and aa.

Definition

Dominance *is when the alleles of a locus are non-additive. When a locus shows dominance, the genotypic value of the heterozygote on a trait is not the average of the two homozygotes. An extreme type is overdominance. It occurs when the heterozygote has a genotypic value more extreme than either parent*

Below you will find another example as an illustration of heterosis due to a dominance effect on a single locus. The genotypic value of the homozygote BB = 125. The genotypic value of the homozygote bb = 115. The heterozygote Bb has a genotypic value of 122. The additive effect of B over b = 125-115/2 = 5. The dominance effect of Bb is 122-120 (120 = the average value of BB and bb) = 2.

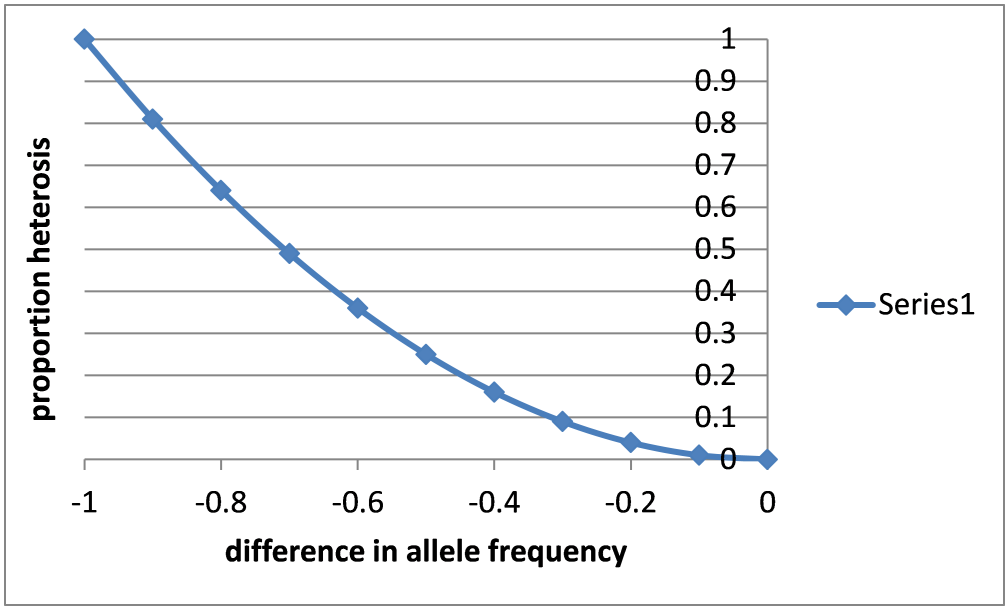


origin: o = (bb+BB)/2 = (115+125)/2 = **120**
additive effect: a = (BB – bb)/2 = (125 - 115)/2 = **5**
dominance effect: d = Bb – o = 122 – 120 = **2**

Chapter 11.4 Effects of heterosis

Heterosis has a positive effect because in the crossbreds many genes are heterozygous that were homozygous in the parent breeds. Alleles with a negative effect are often recessive. In the crossbreds these negative recessive alleles are ruled out. The amount of heterosis to be expected for a specific characteristic in a cross of two breeds depends of the number of loci involved and the differences between the two breeds in the relevant allele frequencies at these loci.

In the figure below this is illustrated when a trait is determined by the two alleles of one gene. The higher the difference in gene frequency, the higher the amount of heterosis. After crossing the two lines for multi genic traits the amount of heterosis is determined by the average difference in allele frequencies for the total number of genes involved and the dominance effects on each locus. When the difference in allele frequencies is -1 (for all genes the breeds are homozygous for different alleles, than the heterosis is equal to 1 (100%).



Chapter 11.5 Motivation for crossbreeding

Heterosis is one of the reasons to apply crossbreeding of breeds or lines. These effects of dominance are observed in all species and across species it can be concluded that estimates for heterosis are higher for characteristics with a low heritability and lower for traits with a high heritability. Heterosis is often substantial for fertility and health characteristics that cannot be easily improved by selective breeding due to the low heritability. Thus, improvement of health and fertility traits is often an important motivation to apply crossbreeding. In New Zealand the cross between Jersey and Holstein cattle has a long history and from 2000 onwards in North America and in Western Europe crossbreeding dairy cattle is increasingly applied. These crosses are applied to improve fertility and health characteristics of high producing dairy cows. These characteristics can hardly be improved in selection programs where milk production is an important breeding goal trait. Danish research show (summarized below) that in a cross of different dairy breeds heterosis can be utilized to improve these characteristics that have a strong relationship with longevity and overall profit.

Heterosis estimates (cattle)

Trait	Heterosis
Production	3%
Fertility	10%
Calving ease (direct)	10-15%
Stillbirth (direct)	5-10%
Calving ease (maternal)	10-15%
Stillbirth (maternal)	5-10%
Longevity	10-15%
Total merit	≥ 10%

- Sorensen et al. (2008)

WAGENINGEN UR

Reference: Crossbreeding in dairy cattle: A Danish perpective. Sorenson, M.K. et al, 2008. Journal of Dairy Science, Vol 91 (11), pp. 4116-4128

The second reason for crossbreeding is to exploit the complementarity of breeds or lines: combination of the characteristics of two breeds or lines is favourable. An example is the cross of sows of a pig breed with a high litter size with a boar of a breed that gives a fast growth up to the slaughter weight. The cross leads to more pigs per litter that grow fast during fattening. This gives a higher profitability than keeping a purebred sow with the same litter size and a moderate growth of the piglets or keeping a sow of a breed with a moderate litter size and a fast growth of her piglets.

The third reason is that crossbreds combine characteristics that cannot easily be improved simultaneously in a single breed. An example is growth of lean meat and meat quality in pigs. These traits are within breeds or lines negatively correlated: animals with a higher growth of lean meat have a lower score for meat quality and the reverse: genes that influence growth of lean meat also influence meat quality, but they have an opposite effect.

The last reason for crossbreeding is the protection of the genetic improvement in the selection lines of commercial companies. They invest a lot in maintaining animals, recording of traits etc. By selling only crossbred animals to farmers, they avoid that competitors can use their purebred parent stock. And selling crossbreds to farmers, generation after generation generates the income for the breeding companies.

Chapter 11.6 The different crossbreeding systems and their applicability

First: where we use the word breed, you can also read selection line. In commercial pig and poultry breeding selection in specialized selection lines is much more common than selection within breeds. Second: in all crossbreeding systems before animals are crossed, they are first selected for the relevant traits. Crossbreeding does not make selection redundant. Third: crossbreeding schemes require a strict implementation by all participants. Hence, crossbreeding is applied for several reasons and they can only be realized when the chosen crossbreeding system is strictly implemented.

In these crossbreeding systems heterosis percentages vary as can be seen in the table below (the crossbreeding systems will be outlined in the subchapters):

Heterosis in different crossing systems with breeds S and T

Type of cross	Heterosis %
F1 (S × T)	100
F2 (S × T) × (S × T)	50
Back cross S × (S × T) or T × (S × T)	50
Second generation of a rotational cross S × (T × (S × T))	75
Third generation of a rotational cross T × (S × (T × (S × T)))	62.5
Rotational cross after many generation	66.6
Second generation of a synthetic line (= F2) (S × T) × (S × T)	50
Third generation of a synthetic line (= F3) (S × T) × (S × T) × (S × T)	50
Synthetic two-breed line after many generations	50
Synthetic three-breed line after many generations	66.6

WAGENINGEN UR
For quality of life

In the pure-breed cross the heterosis in the F_1 is 100 %. In the F_1 population the difference in allele frequencies between the F_1 and one of the parent breeds is half the difference between the two parent breeds. This fact makes that e.g. when the F_1 is mated with one of the parent breeds in a backcross the heterosis in the F_2 is 50 % compared to the original two way cross. The % of heterosis depends of the difference in allele frequencies between the dam and the sire population. As stated before, the heterosis is most pronounced and most valuable for the improvement of traits with a low heritability like health and fertility traits.

Chapter 11.6.1 Two-way cross (pure-breed cross)

Two breeds are crossed and the offspring is used only for production purposes and is not used for breeding. For the crossbreds the notation **F_1** is used. The full effect of heterosis can be exploited in the offspring. It requires the maintenance of the pure breeds, both with a breeding program. This system is widely used in dairy cattle and in sheep. The females not needed to produce purebred female replacements are mated to a breed with a high growth and slaughter quality. In this way the value of the offspring not needed for replacement gets a much higher value compared with purebred offspring.

A * B

F_1 (AB = 50% A, 50% B)

Chapter 11.6.2 Three-way cross

In this cross two way cross females (F_1's) are crossed to a purebred male of a third breed. For their offspring, the second generation in the cross, the notation **F_2** is used. In this system the full effect of heterosis in the crossbred dams (F_1) can be utilized. A special case of a three-way cross is the mating of female crossbreds to one of the sires of the parent breeds. This is known as a **backcross**. In the past in The Netherlands the three-way cross with the Landrace and the Dutch Large White was very popular. The Landrace sows were excellent mothers. The Dutch Large White was excellent in growth and carcass characteristics, but worse in maternal characteristics. The first cross gave excellent sows with a high littersize (due to the Landrace characteristics and the heterosis) and after the backcross a high number of piglets were born with excellent growth and carcasses.

A * B

F_1 (AB) * C

F_2 (ABC = 25% A, 25% B, 50% C)

Chapter 11.6.3 Four-way crosses

In this cross two way cross females (F_1's) are crossed to crossbred males of a third and fourth breed. For their offspring, the second generation in the cross, the notation **F_2** is used. In this system the full effect of heterosis in the crossbred dams (F_1) and males (F_1) can be utilized. Four-way crosses are widely used in commercial poultry breeding programs. All three reasons form the motivation for these four-way crosses in layer and broiler production: exploiting

heterosis, combining ability of breeds and crosses and selection in breeds for traits that cannot be easily improved simultaneously in single breed. In these poultry breeding programs a lot (15-20) of traits are important; many of them are negatively related to each other and/or have a low heritability. First the selection for traits in specialized lines and then combining the traits of the lines by crossbreeding gives the full profit for heterosis.

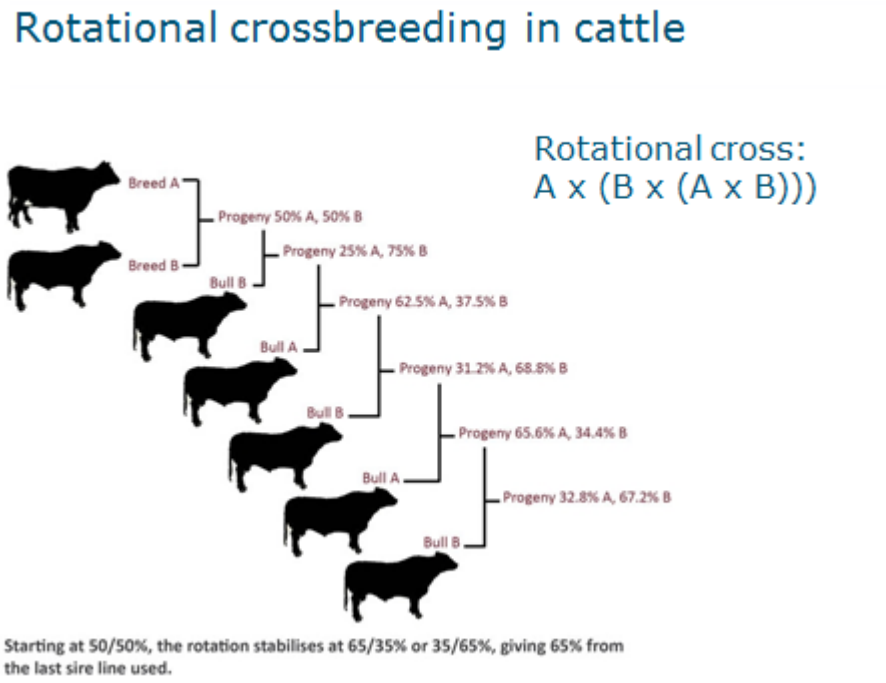
A * B C * D

F_1 (AB) * F_1 (CD)

F_2 (ABCD = 25% A, 25% B, 25% C, 25% D)

Chapter 11.6.4 Two-way rotation (crisscross)

In this cross the start is similar to backcrossing. Females of breed A are crossed to males of breed B. Their female offspring (F_1) are (back) crossed with a male of breed B. Their female offspring (F_2) are crossed with a male of breed A. Their female offspring (F_3) are crossed with a male of breed B etc. Alternately in each generation males of breed A and breed B are used and crossbred females in all generations can be used to produce replacements. In such a cross a part (2/3 compared to a pure-breed cross) of the heterosis effect can still be exploited. The pure breeds still need to be maintained with a breeding program, but this can be done by other breeders. Below the two-way rotation cross in cattle is illustrated:



Chapter 11.6.5 Three-way rotation (crisscross)

In this cross females of breed A are crossed to males of breed B. Their female offspring (F_1: AB) are crossed with a male of breed C. Their female offspring (F_2: 25 % A, 25 % B and 50% C) are crossed then with a male of breed A. Their female offspring (F_3: 62.5 % A, 12.5 % B and 25 % C) are crossed with a male of breed B etc. Alternately in each generation males of breed A, B and C are used and crossbred females in all generations can be used to produce replacements. In such a cross a large part (6/7) of the heterosis effect can still be exploited.

Chapter 11.6.6 Introgression

In this cross males of breed B or crossed with females of breed A to incorporate a characteristic that is present in breed B with a high frequency and that is absent or has a low frequency in breed A. The females of the first cross (F_1: AB) are selected for the wanted characteristic of breed A and mated with males of breed A. This is repeated in the next generations. Breed B is used once and selection in the crossbreds leads to the introgression of the wanted characteristic of breed B. Genetic markers for the wanted characteristic improve the applicability of this method greatly. An example of introgression is the transfer of the Booroola allele in the Texel sheep breed in The Netherlands. It increases litter size and is present in Merino sheep. Crossing once Texel ewes with Merino rams followed by selection of carriers of the Booroola allele in the F_1 resulted in purebred Texel ewes producing twice as many lambs compared to non-carriers of the wanted allele.

A * B

F_1 (AB) * A

F_2 * A etc

Chapter 11.6.7 Grading-up

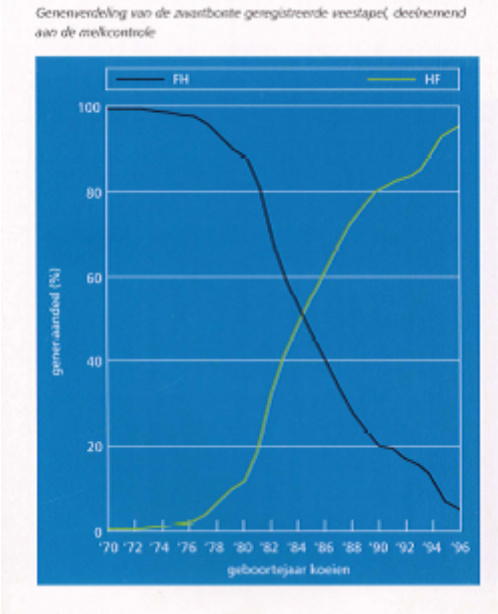
In this cross the method used aims to change a population of animals quickly from one breed to another. Sires of the newly desired breed are continually backcrossed to the females from the previous generation. After three generations the F_3 animals contain already for 87,5 % the genes of the desired breed and after four or five generations the population fully resembles the desired parental breed.

A * B

F_1 (AB) * B

F_2 * B etc

Such a process took place in Western Europe in the 70-ties of the past century when local Black and White cattle populations were graded-up with Holstein Friesian sires from North America creating the present European Holstein Friesian cattle breed.



Chapter 11.6.8 Creating a synthetic breed

This cross starts as follows: two breeds are crossed and males and females of the F_1 generation are reciprocally mated. This is continued in the F_2, F_3, F_4 etc. In this way a new (synthetic) breed is created containing equal parts (50%) of the alleles of the two founder breeds. According to this principle also three or four breeds can be used to create a synthetic breed. Then the reciprocally mating of males and females in the F_2 generation is the real start of the new synthetic breed. The Dutch Flevolander is a recent example of a created new breed. It started with the cross of Finnish Landrace sheep, a breed with a high litter size, with Ile-de-France sheep, a breed with a non-seasonal oestrus pattern. The F_1 ewes produce large litters and give three lambings in two years. The F_1 rams were mated to F_1 ewes, F_2 rams to F_2 ewes, F_3 rams to F_3 ewes etc. In this way a very prolific sheep breed is created.

A * B A * B

F_1 (AB) * F_1 (AB)

F_2 * F_2

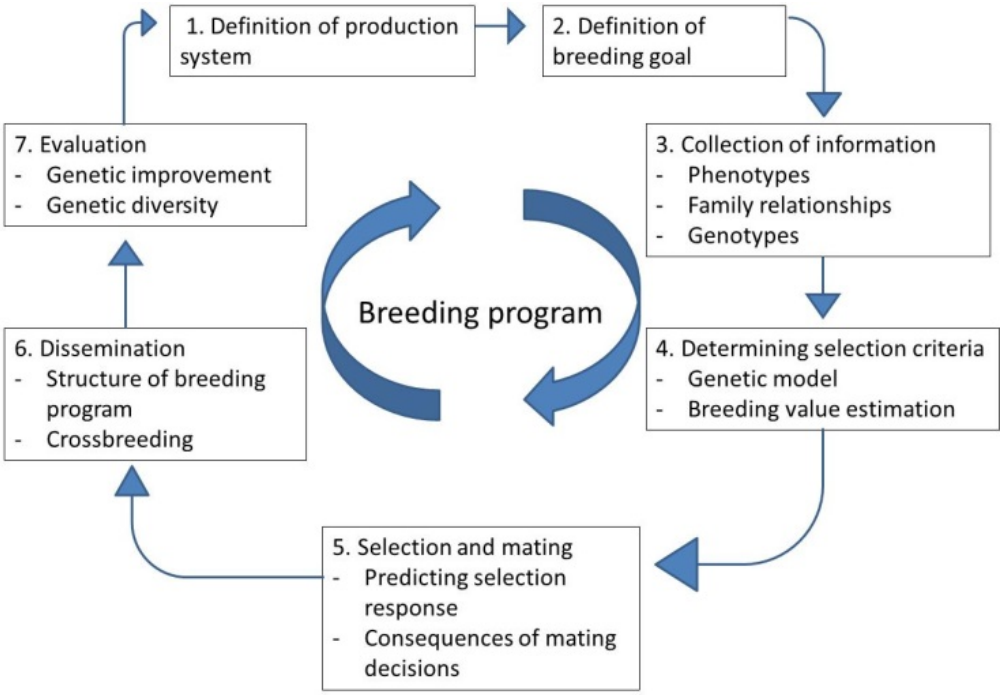
F_3 * F_3 etc

Chapter 11.7 Key issues on crossbreeding

1. Crossbreeding are matings between animals of different breeds or lines.
2. A breed is an interbreeding group of animals within a species with some identifiable common appearance, performance, ancestry or selection history.
3. Selection lines are formed by pure breeding or by crossing different breeds. After the formation of the line the animals in such a selection line are selected for a limited number of breeding goal traits. After generations of selection they excel in these specific breeding goal traits.
4. When breeds or selection lines are crossed, the crossbreds not only combine the characteristics of each of the breeds or lines, but for some characteristics the performance of the crossbreds is higher than the average performance of the parent breeds or parent lines due to heterosis.
5. Heterosis or hybrid vigour is the extent to which the performance of a crossbred in one or more traits is better than the average performance of the two parents.
6. Heterosis has a positive effect because in the crossbreds many genes are heterozygous that were homozygous in the parent breeds. Alleles with a negative effect are often recessive. In the crossbreds these negative recessive alleles are ruled out.
7. Heterosis is one of the reasons to apply crossbreeding of breeds or lines. The second reason for crossbreeding is to exploit the complementarity of breeds or lines: combination of the characteristics of two breeds or lines is favourable. The third reason is that crossbreds combine characteristics that cannot easily be improved simultaneously in a single breed. The last reason for crossbreeding is the protection of the genetic improvement in the selection lines of commercial companies.
8. In all crossbreeding systems before animals are crossed, they are first selected for the relevant traits. Crossbreeding does not make selection redundant. Crossbreeding schemes require a strict implementation by all participants. Hence, crossbreeding is applied for several reasons and they can only be realized when the chosen crossbreeding system is strictly implemented.

Chapter 12: The structure of breeding programs

Breeding programs aim at generating genetic improvement in a population. In breeding programs a permanent selection response is created for the breeding goal traits by collection of information on selection candidates, estimation of breeding values, selection among selection candidates in combination with a mating scheme for these candidates. Thus, the genetic improvement generated by the selection of animals in phase 5 in the breeding program is disseminated in phase 6: the structure of breeding programs.



Definition

*A **breeding program or a breeding scheme** is a program aiming at defined breeding objectives for the production of a next generation of animals. It is the combination of recording selected traits, the estimation of breeding values, the selection of potential parents and a mating program for the selected parents including appropriate (artificial) reproduction methods*

Chapter 12.1: Genetic improvement in a breeding program

As we have seen before, the permanent selection response depends on the selection intensity, the accuracy of the breeding value, the genetic variation and the generation interval. The first three are in the nominator of the formula and the last one in the denominator. Thus a higher the selection intensity, accuracy and genotypic variation and a lower generation interval gives the highest response. The formula is:

$SR = i * r* var_g / GI$

where: SR = selection response; i = selection intensity; r = accuracy; var_g = genotypic variation and GI = generation interval

Within a breeding program the genotypic variation can hardly be influenced, but the other three parameters can be influenced. They are interrelated especially the accuracy of selection and the generation interval. E.g. a choice can be made for selection with a high accuracy, but you have to wait a long time before you have all the information of the selection candidates resulting in a long generation interval. When the choice is for a lower accuracy, you may select younger animals for breeding, resulting in a short generation interval. Thus breeding programs can be optimized with respect to selection intensity, accuracy and generation interval.

As shown before, within a breeding program the selection response can be generated in four different selection paths. The total selection response is the addition of the selection response in these four different paths, with in the following order, a decreasing impact on genetic improvement: sires to breed sires, sires to breed dams, dams to breed sires and dams to breed dams. In the structure of breeding programs these differences in impact are reflected in the structure.

Chapter 12.2: Breeding programs are less or more in control

In a breeding program tenacity, accuracy and discipline of the active breeders is crucial. Crucial are: tenacity with respect to the breeding goal, accuracy in collecting phenotypes, genotypes and pedigree registration and last but not least discipline in selection and mating. It are all important human factors that should be kept under control. But in nearly all species the female animals, and sometimes also the males, of the population under selection are owned by individuals: farmers or citizens. They own the animals and decide whether they want to breed

with their animal(s) and they decide on the mating partner for their animals. And in case of individual ownership, the willingness and help of the owners is greatly required to stick to the breeding goal, to collect information systematically and accurately and to perform the selection and mating process.

In animal species kept for companion or leisure purposes control over the breeding program by the breed associations is very loose. These programs have a flat structure: nearly all females can be selected and in most cases the breed association only has a strong vote in the selection of the males for breeding. Shows where conformation of the animals is leading, determine to a great extend which males are used, often resulting in a few “champion” males very heavily used in the whole population. In these breeding programs only the selection paths sires to breed sires and sires to breed dams are effective in creating a selection response in the whole population. As an example we will outline a breeding program for horses (the Dutch KWPN program).

On the contrary, in pig and poultry production (pork, eggs and broiler meat) commercial breeding programs have full control over all breeding activities. They own a limited number of breeding animals that are part of their selection lines. In these lines the companies determine the breeding goal, perform the collecting of data and the breeding value estimation, and take care for the selection and mating of parents to produce a new generation. The final product is produced by large numbers of animals that are three- or four-way crosses from the selection lines they own. In these breeding programs all the selection paths are effective in creating a selection response in the whole breeding program with a pyramidal structure. Working with different selection lines, that differ in breeding goal traits, the breeding companies have the flexibility to produce different three- or four way crosses. E.g. in pork production markets are different. Requirements for slaughter weight, carcass composition and meat quality are not equal. The globally operating companies have different sire lines for pork production which are used to serve different markets for pork. Controlled by contracts, farmers multiply and cross the selection lines in a pyramidal structure. As an example we will outline a breeding program for pork production (the Topigs program).

In between breeding programs with a flat and loose structure and breeding programs with a pyramidal structure fully in control, there are the breeding programs with an open nucleus. In these programs a part of the population is owned by a limited number of breeders and/or a breeding company. This part is used to select the sires and to select the dams of the sires for the next generation. In these breeding programs only the selection path dams to breed dams is of few importance in creating a selection response in the whole population. As an example we will outline an open nucleus for dairy cattle breeding (the CRV program)

Chapter 12.3: Breeding programs with a flat structure

In many species used by mankind, breeding programs have a simple structure: e.g. dogs, horses and meat producing sheep and goats. In such programs an intense selection of males takes place, because you need a limited number of males to produce the next generation. Some selection in the females is practised, because you need a lot of them as dams to produce the next generation, but this selection is hardly effective. In these species the breeding animals (especially the females) are in the hands of individual owners which take their own decisions on selection and mating. As consequences breeding goals change too often and are not used steadily, recording of traits and pedigree is less complete, selection and mating can hardly be influenced. This results in a low genetic improvement rate over generations.

In these species herdbooks play a prominent role in the breeding program. They do the pedigree recording and set the rules for the characteristics of the males and females to be selected as parents for the next generation. The rules for the males are often very strict and only a limited amount is approved for breeding. Often a lot of emphasis in this approval is given to conformation. For the females the rules are very loose, disapproval of females is seldom practised.

In horse breeding a lot is done nowadays to professionalize the breeding program. A limited number of stallions in the breed is approved for breeding by the studbooks and females with the best conformation, health and performance traits are promoted for breeding. This results indeed in a genetic improvement that can be established in analyses.

In the breeding of dogs, shows play an important role in the selection of males. A limited number of males with the best conformation score at shows is used for breeding, often without any control by the breed association. Nowadays, dog breeding is under discussion in the society due to the negative side effects of a strong selection for conformation and the high relationship among animals in the population leading to inbreeding and the high frequency of genetic defects.

In meat producing sheep and goat breeds the selection of rams and bucks is most effective as you need a limited number of them to produce the next generation. Simple breeding goals for these species are weight and muscularity at a fixed age. With the use of young rams and bucks the generation interval can be kept short and creates genetic improvement. From a farm economics point of view, the ewes should produce litters as long as possible. So, the generation interval on the dam’s side is rather long. In more intense sheep and goat production systems fertile dam breeds are crossed with males from specialised meat breeds, resulting in a lot of lambs with a good growth and slaughter quality. The Dutch Texel breed is famous for these traits and is globally often used as a terminal sire line to produce crossbreds lams from females of local fertile breeds. In milk producing sheep and goat breeds, the selection of the

dams of the males used for breeding is practised. Progeny testing of males is hardly performed due to a low percentage of milk sheep and milk goats in milk recording schemes. Up to now, little genetic improvement is obtained in milk sheep and goats.

Chapter 12.3.1: Example of a breeding program with a flat structure: the KWPN program

The KWPN breeding goal (2014)

Since 2006, the KWPN distinguishes four breeding directions. Riding horses, that are subdivided into the dressage and jumping disciplines, make up the largest group (85-90%). The other two breeding directions are the harness horse and the Gelders horse. Even though each breeding direction has its own additional objectives, all horses fall under the general KWPN breeding goal, which aims at:

- breeding a competition horse that can perform at Grand Prix level;
- with a constitution that enables long usefulness;
- with a character that supports the will to perform as well as being friendly towards people;
- with functional conformation and a correct movement mechanism that enables good performance;
- with attractive exterior that is preferably attractive, along with refinement, nobility and quality.

The KWPN has formulated a separate breeding standard for each breeding direction. These standards are, in fact, descriptions of the ideal dressage-, jumping-, harness- and Gelders horse. The breeding standard helps to evaluate horses objectively and uniformly. It provides framework in which jury members give their evaluations. This reduces the risk of pronounced personal preferences and therefore increases the uniformity, reproducibility and the reliability of the evaluation.

A distinction must be made between mares and stallions.

The selection and use of the best mares is stimulated by the studbook by issuing predicates. Predicates can be seen as ‘quality stamps’ that a mare has. The mare can earn predicates based on her own qualities (performance, conformation and heath) or based on qualities of the offspring (performance, conformation).

For the stallions there is an obligatory route to become a studbook approved stallion.

The stallion route has four steps:

1. Inspections of the stallion: on hard surfaces for correctness of legs and body and on free jumping (jumper horses) or free movement (dressage stallions).

2. Health checks: in the various stages of the stallion selection process, stallions must prove to meet minimum demands for clinical correctness of body and functioning, X-rays, semen quality and roaring.
3. Station performance test: depending upon their age, stallions must prove their sport ability during a central performance test of maximum 70 days
4. When a stallion has offspring, these offspring are inspected and/or followed in sport. Based on the collected information, various breeding values are estimated for each stallion. Based on their breeding values, the stallions are evaluated on the moment when their eldest offspring is 1 year, 3 year, 7 year and 11 years of age.

KWPN collects various types of information on many different moments:

Linear scores of conformation and movement 20 random foals of each new approved stallions

X-rays for ostechondrose 20 random yearlings of each new approved stallion

Linear scores on conformation, movement, free all (mainly) 3 year old mares offered for studbook inspection

Jumping or free movement + valuing marks all young stallions offered for the stallion selection process

Marks obtained in performance tests one day performance tests for Mares (IBOP) and

station tests (EPT) for mares and stallions

Competition results off all registered horses

These data are used to estimate breeding values for all stallions and mares.

Breeding values are estimated for:

All linear scored traits (conformation, movement, free jumping)

Osteochondrosis

Sport dressage

Sport jumping

These breeding values are communicated to breeders, used during selection processes, to evaluate the breeding program and for the best combination of stallion and mare.

Chapter 12.4: Breeding programs with an (open) nucleus

Nucleus programs are characterized by a limited number of female animals with a genetic superiority. In potential, these are the dams to breed sires. They are owned by a breeding organization or a limited number of breeders and called a nucleus (breeding farm(s) or breeding unit(s)). They deliver the next generation of sires to breed sires and sires to breed dams. They are recorded for a large number of traits. The breeding organization takes the decisions on selection and mating in the nucleus and in the contracts. As a consequence breeding goals are used steadily, recording of traits and pedigree is complete, selection and mating in the nucleus is under full control. This results in a high genetic improvement rate over generations. Nucleus programs can be closed, as is the case in commercial pig and poultry breeding. Once the breeding animals for the nucleus are chosen at the start, no animals from outside the nucleus are added to the nucleus population. It is called a closed nucleus breeding program.

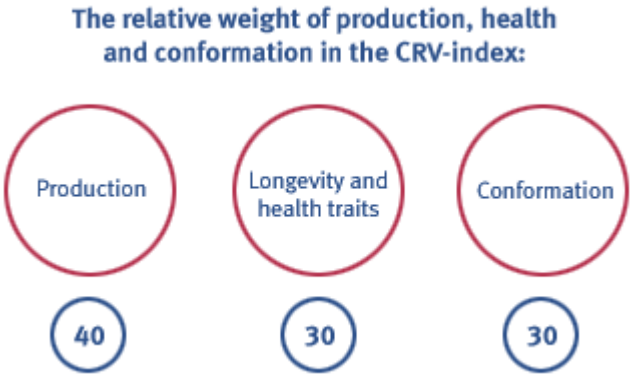
In cattle breeding artificial reproduction techniques, in particular artificial insemination techniques and in vitro fertilization in combination with embryo implantation are well developed and heavily used in the open nucleus. This gives the opportunity to produce high numbers of offspring from superior sires and from dams and disseminating the genes of these superior animals widely in the production population. In the disseminated part of the population (mainly used for production purposes = production population) the offspring of the sires is tested for important traits. When the estimated breeding value of females in the production population is comparable with (or higher than) the breeding value in the nucleus population they can enter the nucleus. In that case they can be bought by the breeding company or contracted by the breeding organization. This is called an open nucleus breeding program.

Open nucleus breeding programs can be considered in species where the present breeding programs have a flat structure: horses, dogs, sheep and goats. Then genetic improvement can be generated in a controlled situation and males from the nucleus can be used widely by individual private breeders. In dogs open nucleus breeding programs are developed for breeding working dogs and for eye-seeing dogs even closed nucleus programs are set up.

Chapter 12.4.1: Example: an open nucleus breeding program: the CRV dairy program

The breeding goal and CRV selection index

For its nucleus CRV uses as a breeding goal (2014): '*a healthy and long lasting dairy cow that will give the optimum contribution to farm profit'.* For the selection of bull dams and sires of sons, for the nucleus animals or potential nucleus animals CRV in The Netherlands uses its own index. In this 'CRV-index' a lot of weight is put upon health traits and longevity as can be seen below:



Each of these three categories is a sub-index in which respectively a number of production traits, longevity and health traits and a number of conformation traits are combined in one single figure.

The traits for these three categories originate from several sources. An important source is the milk recording service. For management reasons farmers like to know individual milk production data (kg and content). Conformation data are gathered by inspectors visiting farms with regular intervals. Longevity and health traits are composed out of milk recording and conformation data.

For the bulls selected to disseminate the genetic improvement into the production population another selection index is used: the NVI. This NVI is the total net merit index that is used in the Netherlands and Flanders for ranking bulls with the aim of putting those bulls at the top that are able to produce daughters that come closest to the national breeding goal. The NVI is derived from a formula that takes into consideration three different components: Production, Health and Conformation. Below the underlying traits are given and their contribution to the NVI:

The selection response obtained in the milk production population

In the table below the genetic gain (in breeding values) in the production population from selection with NVI after 1 generation is given (e.g. with selection with the NVI formula the next generation of animals will have a breeding value for milk that is 272 kg higher than the breeding value for milk of the present generation of animals :

Trait	Gain	Value

Milk	272	Kg
Fat	13	Kg
Protein	8.7	Kg
Longevity	200	Days
Udder health	2.3	Pnt
Udder	1.8	Pnt
Feet & Legs	2.2	Pnt
Calving interval	0.8	Pnt
Interval from first to last insemination	1.0	Pnt
Calving ease (paternal)	1.4	Pnt
Maternal calving process	1.1	Pnt
Vitality (paternal)	0.7	Pnt
Maternal vitality	0.9	Pnt

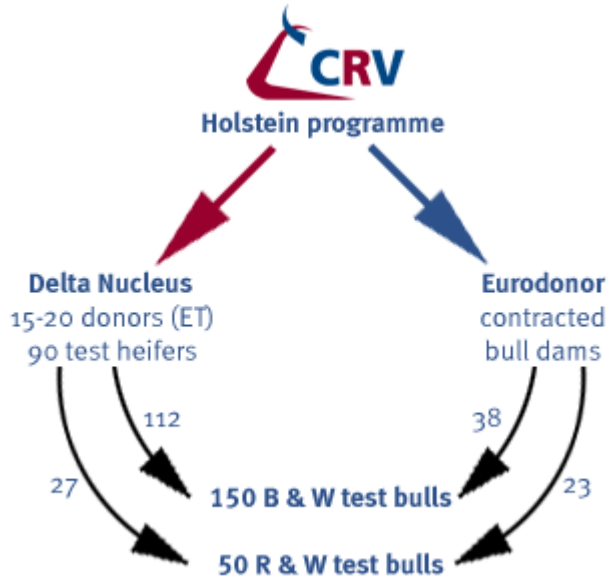
The structure of the CRV breeding program

In parallel CRV performs a breeding program for Black and White (B&W) and for Red and White (R&W) Holstein Friesians. The selected females originate from the nucleus with animals CRV owns (Delta donor) and from farmers in the production populations worldwide (Euro donor). Within Euro donor selected females are contracted with the essence that the selected female is inseminated with a sire of the choice by CRV and a calf born out of this insemination is first offered for sale to CRV. This Euro donor scheme makes the CRV nucleus an “open” nucleus. The offspring of selected Euro females at production farms with the highest CRV index are added to the offspring of the Delta females owned by CRV. In this group CRV selects the sires and dams to breed the next generation of potential breeding animals. Within the Delta scheme In Vitro Fertilization (IVF) is applied. From one year old heifers oocytes are collected and fertilized in a lab with semen of a selected sire. In this way full and half sib progeny (males

and females) of a donor is born out of foster dams. The best males and females, the ones with the highest predictive value for the CRV index are selected. Males are selected to use as sires to use in the nucleus and in the production population and females for the Delta nucleus. When enough embryos become available from a heifer by IVF, the heifer is inseminated and transported to one of the testing farms of CRV. Here, the heifers are raised up to their first calving and subsequently tested under uniform conditions for milk production and conformation in their first lactation. The test results (own performance) are used to calculate a realized breeding value. Annually 100 Delta heifers and 150 heifers from the Euro donor are tested at the farms under contract of CRV. The very best heifers are used again as donor in their first lactation and a part of the embryos obtained are also sold to production farms.

In the scheme below (2014) from the 150 B&W test bulls only 38 were from donors at production farms. For the 50 R&W test bulls 23 were from such donors. In the Delta nucleus an intense selection is applied in the one year old heifer donors: 15-20 are dam of 90 test heifers and of 112 test bulls. They are selected on their predictive breeding value, nowadays a genomic breeding value (see later).

Annually CRV produces 5700 embryos within the Delta program and 3000 embryos within the Euro donor program. And in addition a few embryos are bought on the North American market. The selection in the male calves born is intense: 1 out of 15 is used as a young sire in the production population.



Genomic selection in the CRV breeding program for dairy cattle (personal communication Marieke de Weerd, November 2013)

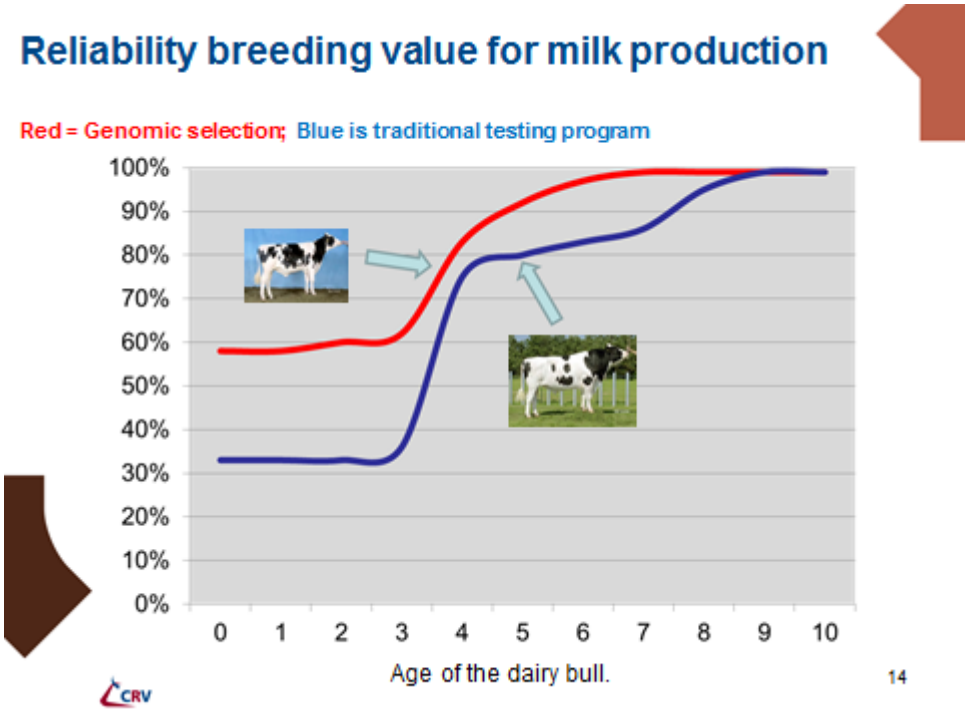
For commercial breeding companies the attractiveness of genomic selection consists of two factors. First, it is possible to calculate differences in breeding value between full sibs before they have their own phenotypic records or records from their progeny. The SNP analysis makes clear which genes of the parents are transmitted to each full sib. Second, young animals

that should have been progeny tested before they obtain an accurate breeding value, can be used intensively directly after puberty. The accuracy of the genomic breeding value comes close to the accuracy obtained after progeny testing. In this way the generation interval in breeding programs with genomic selection can be very short. This speeds up the genetic gain of the breeding programs.

In the past CRV applied progeny testing to breed the bulls that were intensively used by the dairy farmers. On average dairy farmers bought approximately 25 per cent of semen from young unproven bulls. CRV sold in a short term 1000 doses of semen from these young bulls. This resulted four years later in at least 50 daughters of such an unproven bull with a completed first lactation. Then, the selection of the best bulls took place and the best proven bulls were used heavily by the dairy farmers (75 per cent of the inseminations performed at dairy farms).

In this traditional testing scheme the selection of young bull dams and young bull sires takes place in year 0, the birth of the young proven bulls is in year 1, their semen is collected and used in year 2, their calves are born in year 3, their heifers start to produce milk in year 5 and finish their first lactation in year 6. Then, the proven bulls are selected on the first lactation data of their daughters. After year 6, they are intensively used and their progeny becomes productive in year 10. In this traditional testing scheme it takes 10 years before the selection of the young dams and sires leads to an increase in profit for the dairy farmers .

Nowadays (2014) genomic selection is used and leads to a rise in the use of young bulls by the dairy farmers. This is due to the fact that the accuracy of their genomic breeding value approaches the accuracy of the breeding value of the proven sires based on the phenotypic data of their daughters. See, figure below where you should realize that the reliability is the $(\text{accuracy})^2$.



Within CRV breeding program genomic selection of the young bull dams decreases also their age at selection. Nowadays, 75 percent of the young bull dams are one year old heifers and 25 per cent are first calved heifers. In the traditional scheme young bulls dams had at least a full lactation at the moment of selection. An increasing percentage of young bulls with a genomic breeding value is used as sire of the new generation of young bulls. In the traditional scheme only proven bulls were sire of the young bulls to be tested. All these facts sharply decrease the generation interval in the breeding program of CRV and this accelerates the genetic improvement at least with a factor 2.

Another effect of genomic selection are the relatively small costs of genomic testing. Therefore, annually 2600 young bull calves are tested for their SNP variants. Then, based on their genomic breeding value, one out of 15 is selected to put into the breeding program. This sharp selection, even within a group of full sibs, is very attractive. In the breeding program the number of young bulls selected for a high breeding value increases. The bulls that have to wait four years for the milk production data of their progeny, decreases sharply. This has a very favourable effect on the costs of the genomic selection breeding program in comparison with the traditional scheme with a large amount of “waiting” bulls, waiting for years before their daughters finished their first lactation.

CRV is co-operating with a lot of breeding companies in other European countries to create a large reference population. Nowadays (2014) the reference population consists of approximately 30.000 progeny tested bulls that were also screened for their SNP profile. This high number is the reason why the accuracy of the genomic breeding value of young unproven bulls comes very close to the accuracy of progeny tested bulls.

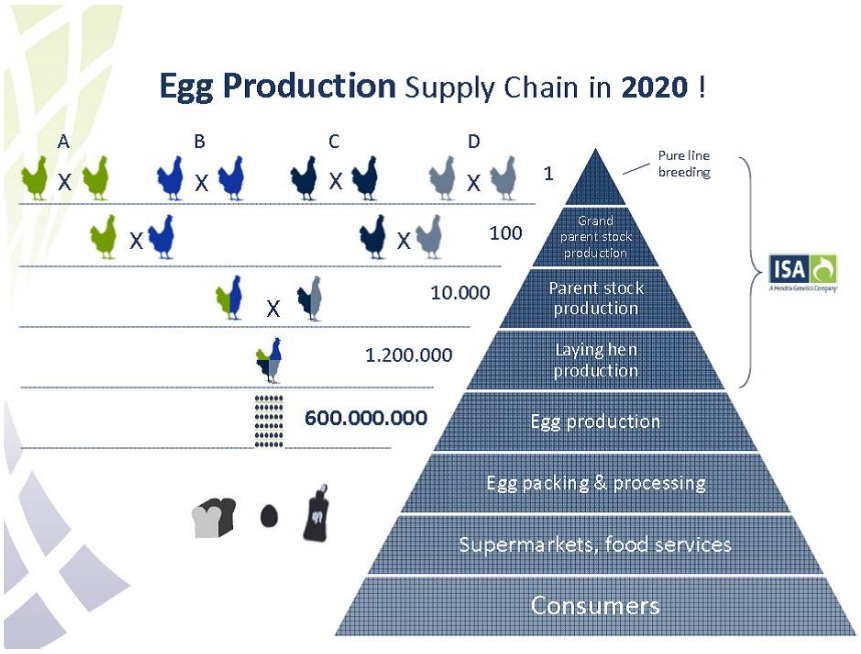
Value of genomic selection within the dairy production farms

At dairy farms a large proportion of the dairy cows has to be used as dam for the next generation of dairy cows. This low selection intensity can be increased in two ways: 1) increase the longevity and with that the number of calvings per cow and 2) by the use of sexed semen, because then the change to obtain a replacement heifer calf is 90 % instead of 50%. In a situation with a higher intensity due to a low replacement rate and the use of sexed semen, genomic selection in young born heifer calves is calculated to be profitable for the dairy farmer.

Chapter 12.5: Breeding programs with a pyramidal structure

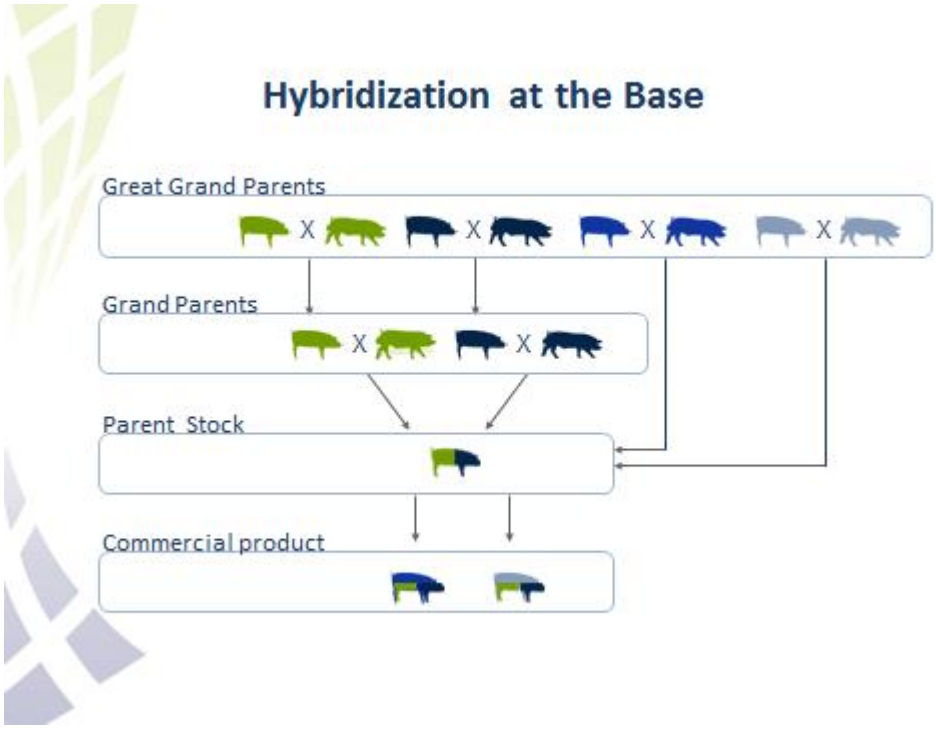
In many breeding programs trait recording is expensive and therefore the number of animals from which traits are recorded is rather small in relation to the whole population of animals. Then, the genetic improvement is realized with a limited number of animals in which the selection is practiced. Subsequently, selected animals are used to disseminate the genetic improvement realized in the recorded population into the whole population. The selection of a limited number of animals, the multiplication in the next generation with a much larger number of animals and the production of the "production" animals in very large numbers in the final generation, leads to a pyramidal structure of such a breeding and production program.

The dissemination of the selection response depends of the structure of the breeding programs. In commercial pigs and poultry programs, selection takes place in the top of the breeding program. Via a few "multiplying generations" the selection response obtained in the top is disseminated into the animals producing meat or eggs. see the figure below where the structure of the Hendrix-Genetics (ISA) breeding program is outlined:



Within the commercial breeding schemes, e.g. for poultry and pigs, the selection response is realized in specialized lines. In commercial poultry (broiler) breeding programs usually a four-way cross is applied. The two female lines are selected for fertility and egg quality and the male lines for growth traits. Crossbreeding results in a high number of healthy chicks due to the full exploitation of heterosis. In the broiler scheme above, the selection takes place in the pure lines in a limited number of Great Grandparents. When the selected Great Grandparents are multiplied in sufficient numbers, they are crossed. The generation interval in the poultry selection lines is very low, less than a year to accelerate the selection response. The A*B line cross gives the A*B F_1 progeny that act as Grandparent and the C*D line cross the C*D F_1 Grandparent. The pure line and the F_1 animals are owned by the breeding company to protect the characteristics of their lines and the realized genetic improvement in these lines. They have the possibility, by maintaining different selection lines to create animals that can produce for different markets and offers the opportunity for a quick reaction at market changes.

In commercial pig breeding programs usually a three-way cross is applied. See e.g. the Hypor program in 2014:



Chapter 12.5.1: Example: breeding program with a pyramidal structure: the Topigs program

Basis: three way-crosses

Topigs piglets used for pork production are three-way crosses: they are the progeny of crossbred sows and boars of a sire line. Topigs works with two sow lines Topigs 20 and Topigs 40 that are crossed to produce the F_1 sows that are subsequently crossed to a boar of one of

the four sire lines: Tempo, Talent, Top Pi and Tybor. The Topigs 20 line is a selection line that was created from a cross between Landrace and Large White lines. The Topigs 40 line is a selection line that was created from a cross between Large White lines. The two sow lines slightly differ in the breeding goal traits. When they are crossed the maternal traits of the F_1 sow (Topigs sow) show heterosis which is an important reason to obtain an excellent dam giving birth to one F_2 piglet extra compared to the average of Topigs 20 and Topigs 40. It results in a robust sow that is able to nurse and wean all the piglets born in the cross with one of the sire lines. The breeding goal traits of the dam lines and the sire lines (the two dam lines and the four sire lines differ among themselves in the weights given to the different breeding goal traits).

Structure and performance testing in the Topigs lines

Health status is very important for a breeding company. Transport of a live animals or semen / embryos between farms and between countries may never spread diseases. Therefore within the breeding structure the process started of converting to an international SPF breeding structure. The goal is to set up 3 SPF genetic nucleus farms each with minimally 250 GGPS (Great Grandparents) per dam line, and a minimum of 600 present sows. Artificial insemination is used more intense because semen gives less health threats than alive boars, it can be used for the worldwide exchange of genetic material to connect breeding lines worldwide and on-farm AI is used to lower generation intervals. A performance test is applied in the dam lines.

A dam line consists of more than 2000 great grandparent (GGPS) and grandparent sows (GPS) tested for fertility and maternal traits. They produce over 7500 gilts that are performance tested for growth traits and carcass composition. Some 1200 out of these 7500 gilts replace the GGPS and GPS sows. Annually 40 GGP boars are selected with a high intensity out of the 2500 boars performance tested per year. For the boars feed intake is an additional trait in the performance test. Per line 20000 crossbred parent sows also provide additional information on fertility and stayability for the selection of the GP and GGP sows. A performance test is also applied in the sire lines.

Explanation of and additional information to the breeding scheme for the sire lines: each sire line consists of 500 GGP sows which are replaced each year (short generation interval). They produce 3500 gilts that are performance tested for growth traits and carcass composition. Of them 500 are selected for replacement. Annually 40 GGP nucleus boars are selected with a very high intensity out of 3500 boars performance tested for growth traits, carcass composition and feed intake. From each nucleus boar from 50-100 crossbred progeny (F_2) the growth performance is known and 25 purebred progeny of these boars the carcass is dissected and meat quality is assessed.

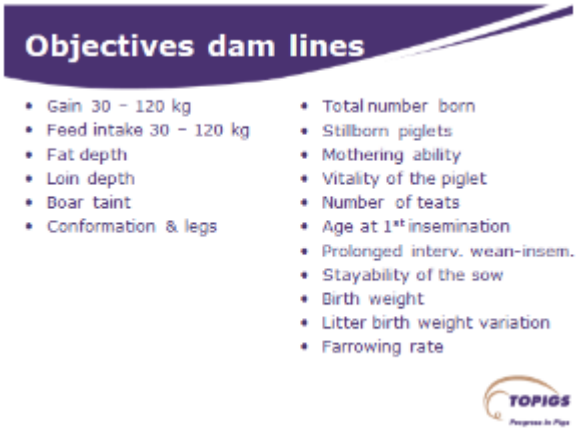
Breeding value estimation

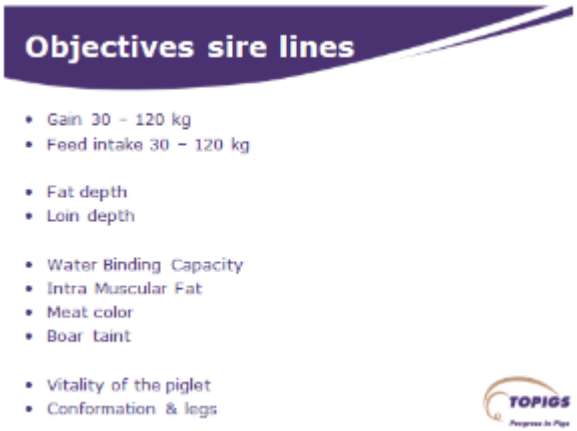
Topigs operates a very large database for the different lines. Weekly the data from more than 300000 sows is updated with data from more than 750 breeding farms in more than 25 countries. The data are collected from all levels of the breeding and production pyramid

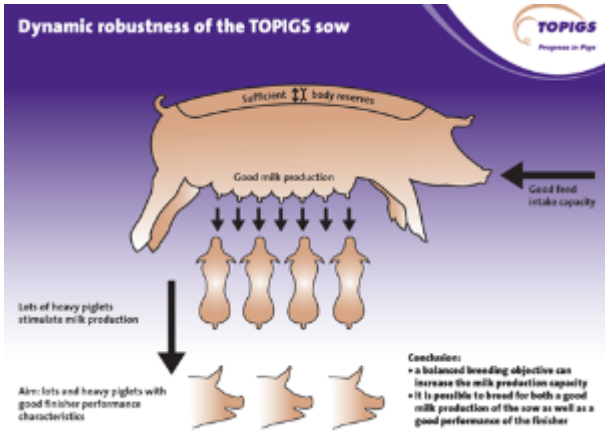
simultaneously. Weekly breeding values are renewed. The breeding values are calculated in a multi trait animal model with a mixed-BLUP procedure. For each animal the Topigs index value is calculated for ranking and selection. Topigs started with genomic selection especially for fertility and carcass traits.

Combined purebred and crossbred performance

In the whole breeding pyramid pedigree registration is carefully done. This facilitates to collect data from all their offspring of the GGP in three generations upto the finishers. As in subsequent generations the number of progeny is multiplied, a wealth of data per GGP is available within a few years. All this data is stored and used for breeding value estimation. Relevant data are survival of sows, gain under practical conditions, slaughter data, especially on carcass composition and meat quality. To get a better prediction for these traits and for fertilty and survival traits, genomic selection is developed and can be used to select GGP animals with a higher accuracy at an earlier phase in the breeding piramyd.







GN Performance testing system *Dam lines*

- > 2000 **(G)GP sows**
- Fertility, stayability, teats, piglet weighing protocol
- > 7500 gilts performance tested/year
- Life gain, test gain, US loin & fat
- 60% replacement (> 100% in the GGPS)

- 40 **GGP boars** selected/year
- > 2500 boars performance tested/year
- Life gain, test gain, US loin & fat, feed intake (FCR)

- Additional sow performance data commercial farms (CCPS):
- Fertility, stayability
- >= 20,000 sows per line

TOPIGS

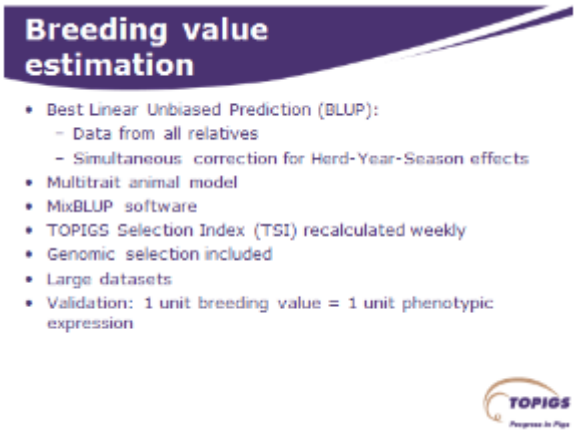
GN Performance testing system *Sire lines*

- 2x 250 **GGP sows**
- 100% replacement
- 3500 gilts performance tested/year
- Life gain, test gain, US loin & fat

- 40 **GGP boars** selected/year
- 3500 boars performance tested/year: till av. 120 kg
- Life gain, test gain, US loin & fat, feed intake (FCR)

- Per nucleus boar:
  - 50–100 crossbred finisher offspring performance (CCPS)
  - 25 dissected pure line animals (kg loin & meat quality)

TOPIGS

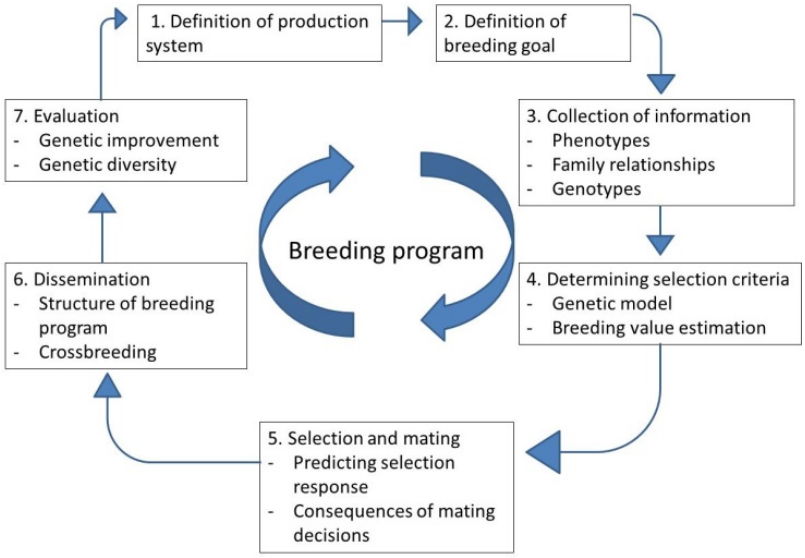


Chapter 12.6: Key issues on breeding programs

1. A breeding program or a breeding scheme is a program aiming at defined breeding objectives for the production of a next generation of animals. It is the combination of recording selection traits, the estimation of breeding values, the selection of potential parents and a mating program for the selected parents including appropriate (artificial) reproduction methods.
2. In a breeding program tenacity, accuracy and discipline of the active breeders is crucial. Crucial are: tenacity with respect to the breeding goal, accuracy in collecting phenotypes, genotypes and pedigree registration and last but not least discipline in selection and mating. It are all important human factors that should be kept under control.
3. In animal species kept for companion or leisure purposes control over the breeding program by the breed associations is very loose. These programs have a flat structure: nearly all females can be selected and in most cases the breed association only has a strong vote in the selection of the males for breeding.
4. In pig and poultry production (pork, eggs and broiler meat) commercial breeding programs have full control over all breeding activities. They own a limited number of breeding animals that are part of their selection lines. In these lines the companies determine the breeding goal, perform the collecting of data and the breeding value estimation, and take care for the selection and mating of parents to produce a new generation.
5. The breeding programs with an open nucleus are in between breeding programs with a flat and loose structure and breeding programs with a pyramidal structure fully in control. In these open nucleus programs a part of the population is owned by a limited number of breeders and/or a breeding company. This part is used to select the sires and to select the dams of the sires for the next generation.

Chapter 13: Evaluation of the breeding program

After the design and implementation of the breeding program it is essential to evaluate the result: how big is the genetic response to selection? If all went according to plan, this will resemble your predicted genetic response. However, in some situations the predicted and realised genetic response will be different. In that case it is essential to find out and to make adjustments where needed. We have reached step 7 in the breeding program cycle: evaluation of the breeding program. In this chapter we will discuss a number of issues that need to be considered so that the success can be maintained or improved on in the next generation. Important subject for evaluation is of course to see whether response that has been achieved is approximately the same as was predicted. If not, you need to find out reasons why this may be the case. Is the quality of pedigree and phenotype recording without error? Have animals that were selected for breeding indeed be used for breeding? Is the selection environment a good representation of the environment that the offspring need to perform in? In other words: have we selected the right animals? Could it be that the population has reached a selection limit? Another point of evaluation is what other changes have occurred that were not anticipated. Has selection to improve the traits in the breeding goal also caused (undesired) response in other traits? Evaluation of a breeding program not only involves evaluating the breeding practice. It will also involve looking ahead. Has there been any change in legislation or market situation, or are changes expected in the near future? What is your competitor doing? What should you do to keep/expand your market share? You see that even though the breeding program is up and running, it still requires effort to keep it that way.



Apart from genetic progress that has been achieved, breeding programs need to be evaluated on another important aspect: to what extend has the genetic diversity been maintained? Is the rate of inbreeding under control? If not, what could be changed to improve that? This evaluation of the genetic diversity is very important in evaluating the breeding program. Without genetic diversity there is no future for the breeding program, and a decrease in genetic diversity results in an increase in inbreeding depression and the frequency of genetic disorders in the population. The evaluation of the breeding program from a genetic diversity point of view will be discussed in a separate chapter (Chapter 14).

Chapter 13.1: How to measure genetic improvement?

Genetic improvement indicates how much better animals are genetically in the current generation as compared to the previous. To be able to determine the genetic improvement, it is required to know the genetic potential of the animals. The actual genetic potential cannot be measured. But if the breeding values of the animals can be estimated with very high accuracy, using information on many offspring for example, these EBV can be used as good approximation of the true genetic potential. Apart from high accuracy, the EBV also need to be estimated without bias of systematic effects. For example, animals that are fed with a high quality feed may perform systematically better to animals that are fed simple feed. If EBV are not corrected for this effect, the EBV of animals that were fed the high quality feed will be biased upwards. Their performance was better, but not because of their genetics. To get the best estimation of the EBV they need to be estimated using BLUP, which can take these systematic effects of feeding, housing, season, or other environmental influences into account (see chapter on ranking the animals).

Summarising, it is possible to get a good approximation of the genetic improvement across generations when EBV are estimated with high accuracy and using BLUP. The realised genetic improvement or realised genetic response can be determined by taking the difference between the average EBV in both generations.

Realised response to selection = average $EBV_{\text{generation t+1}}$ – average $EBV_{\text{generation t}}$

This simple formula will give the best possible approximation of the realised genetic response. Remember that the predicted genetic response is calculated as

$$\Delta G = \frac{i * r_{IH} * \sigma_a}{L}$$

The difference between the realised and predicted response is what requires evaluation. The smaller the difference, the better the realisation resembles the prediction, and the less evaluation is required. However, if the difference is substantial, it is essential to find out what causes it.

Thus:

> *Realised genetic improvement can be determined by taking the difference in average EBV, for example between generations*

Chapter 13.2: Genetic trend

To get an impression of the realised genetic response in the longer term, it is possible to evaluate *genetic trends* across generations. A genetic trend is a compilation of average EBV per generation, and indicates the direction of change across generations. It is often visualised in a graph and is useful to check whether there are unexpected deviations from linearity, e.g. due to a selection limit. Figure 1 shows the genetic trend for milk production in Dutch black-and-white dairy cattle from 1995 to 2013, expressed per birth year, rather than per generation. Reason for this is that in dairy cattle generations are not discrete, but overlap because some cows get older than others. By expressing a genetic trend per birth year this overlap in generations is overcome. It also provides more insight in what has happened during a defined time span. The EBV are estimated relative to the average EBV in 2009. In technical terms: 2009 was the reference year. Animals with higher genetic potential than the average in 2009 have EBV larger than 0, and animals with genetic potential lower than the average of 2009 have negative EBV. In the figure are the annual average EBV's. The fact that there is an increase in EBV indicates a positive genetic trend. Animals are genetically improved across birth years, so selection has been successful. In the same figure you also see the phenotypic trend in milk production (the green line that starts just below the genetic trend in 1995). The fact that both trends have approximately the same slope indicates that environmental factors such as management, feeding, housing have been supporting, and not limiting the expression of the genetic improvement.

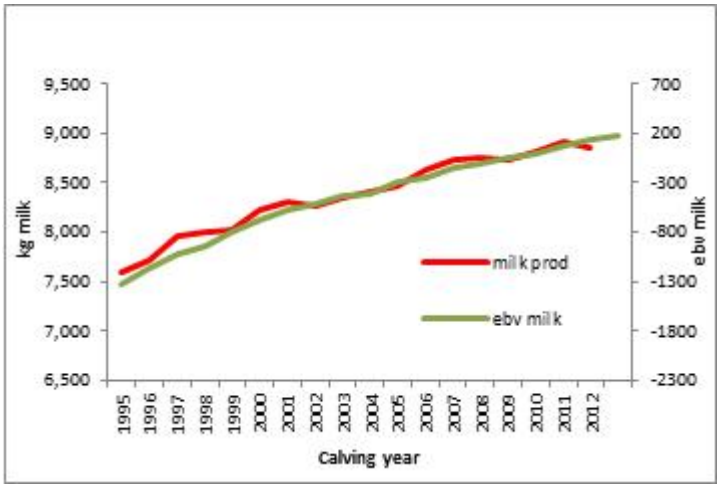


Figure 1. The increase in milk production in Dutch dairy cattle between 1995 and 2013, and the increase in EBV since 1995. Even though the scale is expressed differently, both increase approximately 1500 kg.

Definition

*A **genetic trend** represents realised genetic response across a period of time (e.g. years or generations)*

Chapter 13.3: What could be influencing the realised response to selection?

The predicted genetic response to selection is used to make a plan, to design the breeding scheme. Afterwards, genetic gain is recorded, but it often is not compared to the predictions. That is a shame, because the difference between predicted and realised selection response provides an indication of the realised success of the breeding program. It is, therefore, very wise to monitor because it provides insight in the success of the breeding program and in aspects that do not go as you initially planned them.

Assumptions are not realised

Why would there be a difference between estimated and realised genetic response? Consider the formula for predicting response to selection and evaluate each of the components. Were their real values the same as used in the prediction?

Selection intensity

To start with, the realised selection intensity may have been lower than used for predicting the response. For example, some of the animals that were selected for some reason were not able to participate in breeding. This means less superior animals will have taken their place, which

will have decreased the realised genetic response to selection. Or some of the selected animals were used as parent very much more than others. This will affect the selection response.

Accuracy of selection

The next component is the accuracy of selection. This is influenced by the heritability and the information sources, e.g. own performance versus sib testing. The information sources are influenced quite easily. If in the prediction equation it was assumed that the EBV of all animals would be based on 5 offspring each, and in reality some animals had fewer offspring, this will reduce the accuracy. Similarly, it could also be that instead of 5, some animals had 8 offspring. This will improve the accuracy of their EBV, and thus increase the probability of selecting the genetically best animals for breeding. The heritability cannot be influenced easily. As we have seen in the chapter about genetic models, one potential way to increase a heritability is by improving the measuring method for obtaining the phenotype. The heritability may also change due to a change in additive genetic variation.

Additive genetic standard deviation

The next component in the formula is the additive genetic standard deviation, which is the square root of the additive genetic variation. The additive genetic variance is estimated by combining the phenotypic information and the additive genetic relationships between the animals. See the chapter on genetic models. It makes use of the fact that related animals are more alike than unrelated animals. However, if these pedigree relationships are not recorded accurately, the related animals (on paper) no longer perform that much more alike. Fewer of the similarities between the animals can be assigned to genetic relationships. Pedigree errors thus reduce the size of the estimated additive genetic variance.

Even if the pedigree recording was correct and the estimated additive genetic variance is as accurate as possible, still the estimate may change somewhat across generations. As we have seen in the chapter about relationships and inbreeding, there are some forces that will have an influence on the additive genetic variation, even though the changes will not be large from generation to generation. In the longer term it does make a difference. Therefore, it is important to re-estimate the additive genetic variation at regular basis. Potential reasons for change are that selection increases the frequency of the desired alleles. Genetic drift, however, may cause the alleles that were under selection to decrease, rather than increase, in frequency. Mutations may create new variation, that cannot be predicted at forehand.

Generation interval

The last component in the prediction equation is the generation interval. This only matters if you chose to express the response per year, rather than per generation. Determining the generation interval can be quite tricky, as we have seen in the chapter on genetic response to selection. It hardly ever is the same for all families, so you have to assume an average. In real life, the generation interval may be longer or shorter than anticipated, causing the realised genetic response per year to be different from the predicted one.

Thus:

When failing to realise predicted genetic response to selection: are all assumptions with respect to the components of the prediction equation met?

Chapter 13.4: Selection limits

A reason for not meeting the expected response to selection may be that the population is reaching a selection limit. A selection limit indicates that the population has reached the point where further change is no longer feasible. This can be due to the fact that there is no more genetic variation, but there are other reasons.

Limit due to opposing natural selection

In Figure 2 you see an example of a population that apparently has reached a selection limit. The high line is still responding to selection and is growing bigger and bigger. The low line, on the other hand, has linearly decreased in size for approximately 25 generations, but then further decrease was no longer feasible. Even though selection was on the lightest chickens every generation, the next generation was not getting any lighter anymore. It is not clear why this is the case. Selection results are always expressed at phenotypic level. It could be that the genetically smallest birds showed the same phenotype as the genetically larger birds, so directional selection was no longer possible. In that case this selection limit represents a physiological limit, rather than limited genetic variation. Which of the two was the case could be tested by selecting the light birds upwards again. If that is still possible then the genetic variation is still present. Another reason for the reached selection limit could be that the smallest birds were no longer capable of reproduction. That would be a typical example of natural selection working in the opposite direction of artificial selection. Selection limits due to opposing natural selection usually are hard to undo. In some cases an improvement of the environment may take away the natural selection limit.

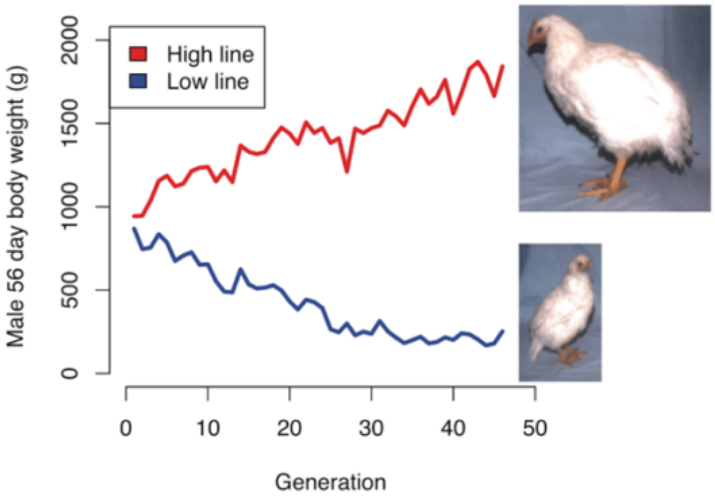


Figure 2. Phenotypic averages across generations of divergent selection in white leghorns. From Johanssen et al., 2010 Genome wide effect of long-term divergent selection DOI: 10.1371 /journal.pgen.1001188

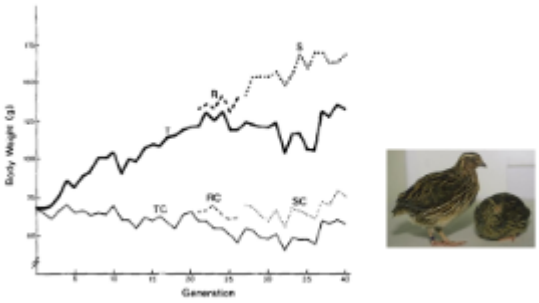


Figure 3. Selection for increased body weight at 4 weeks of age under different nutritional environments in quail. The T-line reached a selection limit around generation 25. This limit disappeared after improving the diet (R and S-line). After Marks, 1996. Poultry Science 75: 1198-1203.

Limiting environment to express genetic potential

For example, in order to show growth potential, animals require sufficient nutrient intake. If those nutrients are unavailable, they can't show their genetic potential and only grow as the feed intake allows them. This is shown in a long term selection experiment in quail, where at some point a selection plateau occurred (see figure 3, the solid line). A selection plateau seems to be the same as a selection limit, only it can be elevated by changing the environment. This is not possible with a real selection limit. The selection plateau in the quail was elevated by improving the feed quality. With this experiment it was clearly shown that limits to selection in some cases actually may be plateaus because they are caused by limits in the environment, rather than by genetics.

There are three main reasons for the occurrence of a selection limit:

- loss of genetic diversity (irreversible)

- opposing natural selection through reduced fertility or even mortality (usually irreversible)

- limiting environment to express potential (often reversible)

Chapter 13.5: Practical issues influencing response to selection

The realised genetic trend is not always the same as the predicted genetic trend. In practice it generally is not higher than expected, but it can be lower. If it is lower than expected, as breeding organisation it is essential to find out why this is the case. We have already discussed

the potential issues with difference between predicted and realised selection response. If these issues remain across generations, this will have an influence on the genetic trend. But there are a few additional reasons that we will discuss.

An obvious reason is that the animals that were selected and used as breeding animals were not as good as expected. This may have been caused by the fact that some animals were scoring systematically better than others because of a treatment that was not included as systematic effect in the breeding value estimation (the details go beyond the scope of this course). This better scoring has been assigned to their genetic potential. Therefore, all these animals have an overestimated EBV, which placed them higher in the ranking then they should have been based on their genetic potential. When it has become clear what is the systematic effect causing this overestimation of some of the EBV, this can be fixed quite easily.

A change in the breeding goal in between the time of prediction and realisation most likely will result in a difference between the realised and the predicted selection response. The predictions were made using the old breeding goal, but the animals are selected already using the new, slightly changed breeding goal. Obviously this will result in a difference between predicted and realised genetic response to selection. Related to this, even though the breeding goal remains the same, a change in phenotype recording may have similar consequences. For example, introduction of new and improved equipment will improve the accuracy of the measurements, which may have an increasing effect on the heritability (see chapter on genetic models), which in turn has an increasing effect on the accuracy of selection, and thus on the realised response to selection. Solution to both of these issues is to adjust the predicted response to the new situation and compare the newly predicted to the realised selection response.

Thus:

> *Deviations from expected genetic trend may be due to a change in breeding goal or a change in phenotype recording*

Chapter 13.6: Genotype by Environment interaction

A special case is when selection is based on performance of animals who live in a different environment from the one that their offspring is going to live in. This means there is a potential risk of selecting animals that may be best in the parental environment, but not in the offspring environment. This risk is negligible when both environments are similar. However, it may become an issue if these environments are obviously different. For example, a dairy cattle bull may have been selected based on the top performance of his daughters in the Netherlands. He is also used as sire on cows in Spain. However, his daughters can’t stand the heat very well and perform less than expected because milk production in Spain requires tolerating high temperatures, whereas this is much less the case in the Netherlands. Another bull is also used

as sire in Spain. He has a lower EBV in the Netherlands because his daughters were not the very best. However, he is very popular in Spain because there his daughters are doing really well. In other words, to be able to express the same trait: milk production, the genetic potential of a cow in the Spain needs to be slightly different from that of a cow in the Netherlands. It requires the capacity to tolerate high temperatures. The best bull in Spain thus is not the best bull in the Netherlands as both environments require slightly different genotypes. This re-ranking of animals according to the specific combination of genotype and environment is called *Genotype by Environment Interaction or GxE.* The 'genotype' can refer to individual animals, such as these two dairy bulls, but it can also represent population averages, instead of the individual animals.

Thus:

> *Genotype by environment interaction (GxE) = when the difference in performance of two genotypes depends on the environment in which the performance is measured.*
>
> *GxE can refer to a change in size of the difference in performance, or to a change in ranking of animals, in different environments*

Chapter 13.6.1: Prerequisites determine the environment

The environment can be considered as a set of prerequisites. An animal performs at its best if it possesses all prerequisites for that environment. If some are missing, the performance is reduced. Each environment has its own set of prerequisites. However, some environments are so much alike that animals can use the same set of prerequisites. The more different two environments are, the more important it becomes to have the specific prerequisites to manage in either of the environments. Sometimes, the prerequisites are even adverse: if you have the one, you can't have the other. For example, if you have a thick fur to tolerate cold, you can't stand the heat. These adverse prerequisites are called *trade*-offs. An animal can be capable of handling one prerequisite, for example digesting low quality feed, but that requires a type of digestion physiology which makes it impossible to grow very fast when the feed quality is good. You can be good at one thing, but it automatically means you are not so good at something else. Often meeting a prerequisite in one environment only causes a small negative effect on performance in the other environment. The opposite can be much more of a problem: not being resistant to an infection is quite problematic in an environment with that infection.

The size of genotype by environment interaction

To get insight in the size of GxE you can plot the performance of different genotypes across multiple environments in a graph. On the x-axe of such a graph is the environmental gradient, for example temperature, or protein content in the diet, or some other component in the environment of the animals. On the y-axe is the performance of the animals in the environments. The resulting plot is called a *reaction norm*. The slope of the reaction norm will tell you how sensitive the genotype is. A horizontal slope indicates absence of sensitivity to the

environments that are considered. A slope indicates that performance in one environment is better than in the other. Parallel reaction norms indicate that the two genotypes are equally sensitive to changes in the environment across the environmental range. However, if one reaction norm is more steep than the other, this is an indication that one genotype is more sensitive to a change in the environment than the other and only then we call that a genotype by environment interaction. In the extreme situation the reaction norms even cross each other. That means that in the one environment the one population would be best, whereas in the other environment the other population would be best.

In figure 4 is an example of both type of reaction norms to indicate GxE. In the top figure the reaction norms are non-parallel, but not crossing, indicating that one population remains superior to the other in both environments. In the bottom figure the reaction norms are crossing, indicating that the genetic superiority changes with the environment change. The reaction norms in these figures are straight lines because only two environments are considered. They can become non-linear when more environments are included, and thus multiple performances are compared. Non-horizontal reaction norms indicate that some breeds are more capable of dealing with less optimised environments than others.

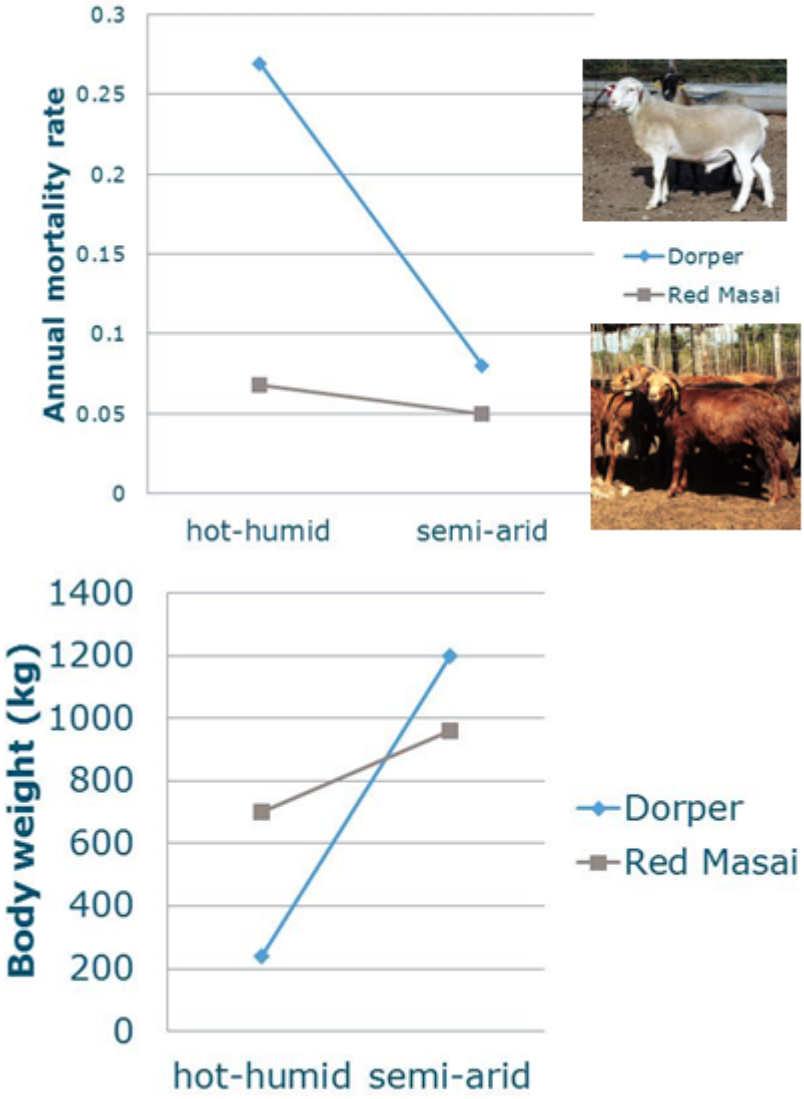


Figure 4. Two examples of GxE in the sheep breeds Dorper and Red Masai. In the top graph the non-parallel reaction norms are an indication of the higher sensitivity of the Dorper to semi-arid conditions than the Red Masai, expressed in mortality rate. But the mortality rate of the Dorper remains higher in both environments. In the bottom graph the GxE is stronger and indicating that the Dorper is heaviest in hot and humid conditions, whereas the Red Masai is the heaviest in the semi-arid conditions.

Thus:

A reaction norm represents performance of genotypes (i.e. animals or populations) in a range of environments.

Non-parallel reaction norms of genotypes indicate the presence of genotype by environment interaction

Chapter 13.6.2: Consequences of GxE for the breeding program

Genotype by environment interaction occurs in any situation where one population is more sensitive to changes than others, indicated by a steeper slope in the reaction norm. What are consequences for the breeding program? When do you need to take into account that the offspring is supposed to perform in a different type of environment than what the parents are selected in?

An answer to these questions lays in the genetic correlation between performance in both environments. If the correlation is low, or even negative, selection based on performance in one environment may result in poor performance of the offspring in the other environment. For example, if you would select the best performing animals in an environment that is optimised in all possible ways (perfect housing, feeding, health care, etc), you may get the animals with the best genetic potential for your trait if interest IN THAT ENVIRONMENT. However, if you would use those animals as parents for offspring in a more average environment, they may perform poorly because they miss the prerequisites that it takes to perform in that average environment. The genetic correlation between performance in two environments is a measure of how applicable the genotype in one environment is to perform in the other environment. In other words, that genetic correlation provides an indication of whether the same breeding program can be used for both environments or not. Obviously, performance testing of sibs or offspring in the environment that the offspring is supposed to perform in is very valuable information to optimise the selection decisions.

If the same breeding program is supposed to serve a too large range of environments, it may become wiser to split the breeding program in two. That decision will depend on the outcome of a number of questions. What is the genetic progress now and how will that improve when the breeding program is split? Very important in relation to that is your competitive position with respect to others that operate in the same market. As a breeding company, you can save money by maintaining a single breeding program, but you can lose much more by losing genetic gain, and thus market share, to your competitor. Obviously this is the case in farm animal breeding, but it also is the case in riding horse breeding! The KWPN, for example, has decided to split its breeding program in two: one for dressage and one for show-jumping. The selection criteria for both specialisations are different. Stallions for the dressage specialisation are no longer tested on their show jumping skills but they should have excellent gates and show real potential in the performance test. Stallions for the show jumping specialisation are not punished for having less superior gates, but they should show real potential for show jumping. The idea was that this would allow more genetic progress through specialisation. There were costs involved, but the idea was that the benefits would outweigh the costs and market share could be further increased. Because the split of the breeding program into two has been relatively recent, the results have not been evaluated yet. But preliminary results suggest that from genetic progress point of view the split has been a success.

A rule of thumb for running a breeding program is that if the genetic correlation between performance in two environments drops below 0.6, so if different genetics is required to perform well in either environment, it becomes worthwhile to split the breeding program into two: one for each environment. A correlation higher than 0.6 indicates that even though the selection of the parents may be sub-optimal, it still outweighs the costs of running two separate breeding programs. Costs involve not only the financial consequences, but the costs with respect to loss in genetic response to selection if the population is reduced in size, and with respect to maintaining genetic diversity (of the two smaller populations).

Thus:

> *Two environments require separate breeding programs if the correlation between performance in both environments is smaller than 0.6*

Chapter 13.6.3: Correlated response

As we have seen in the chapter about response to selection, sometimes it is possible to use performance for one trait as indication of performance for another, potentially more difficult or expensive to measure trait. This type of trait is called an indicator trait and because of its correlation with the trait in the breeding goal, selection based on this trait automatically improves the trait in the breeding goal. The stronger the correlation, the higher the response to selection in the breeding goal trait. This is an example of making use of existing correlations with a clear advantage for the breeding program. The correlation is used as tool for selection.

Linkage

Correlations exist for a number of reasons. One reason is that the genes that are involved in the correlated trait are located closely to the ones that are involved in the trait under selection. Therefore, recombination events between them are rare, and the allele of the correlated trait often is inherited together with the positive allele of the trait under selection. In technical terms this is called that they are in *linkage disequilibrium*: allele combinations of both genes are inherited together. For example, if a gene that is involved in adult size is closely located to a gene involved in hip dysplasia in dogs, the alleles for size and hip dysplasia are inherited together (see figure 5). If the allele for large size is close to the allele for presence of hip dysplasia, and the allele for small size is close to the allele for absence of hip dysplasia, then there is a negative correlation between hip dysplasia and size. This is caused by the fact that these genes almost always inherit together and the combination of alleles is undesirable in a breed that is selected for large size.

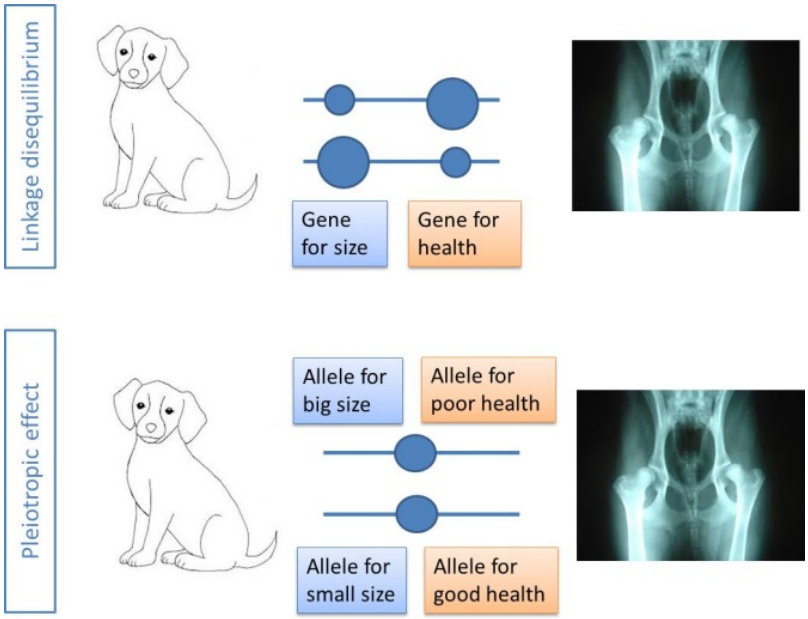


Figure 5. Two causes of genetic correlations. Linkage disequilibrium, where genes affecting different traits do not inherit independently, and the pleiotropic effect, where one gene affects multiple traits.

Pleiotropic effects

A genetic correlation can also exist because the gene affecting the trait is influencing another trait as well. This is called a *pleiotropic effect* of the gene. For example, if a gene that affects adult size in dogs is also affecting the risk of developing hip dysplasia size in dogs, and the allele that results in large dogs is also resulting in higher risk, then selection for adult size will result in poorer hip quality (see figure 5).

Chapter 13.6.4: Limiting resources

A third reason for the existence of an, often negative, correlation between the trait under selection and environmental sensitivity has not much to do with locations on the genome or with multiple functions of individual genes. It has more to do with the fact that the animal has ‘to make choices’ about what to spend its resources on. First of all, these are not conscious choices. We call them choices because resources spent on one process or trait cannot be spend on something else. The mechanism behind the ‘decision making’ what to spend the resources on is not very clear yet, but most likely consist of a combination of genetics, and factors such as life stage of the animal, condition, health, and a number of other factors. Whether the resources that are required for the various processes are available will depend on the availability, and quality of the available resources, but also on the intake capacity of the animal. Think about feed intake capacity in dairy cattle. Some cows cannot ingest enough feed to maintain body condition. They spend all resources on milk production, but have restricted feed intake capacity so they have to use some of their resource reserves (body condition) as well.

Thus:

Genetic correlations can exist for several reasons:

- *Linkage disequilibrium*
- *Pleiotropic effect*
- *Conflicting resource allocation*

Animals have some (but not complete!) flexibility in shifting resources to various processes. It seems that some animals are better capable of doing that than others, and there are indications that this is heritable. For sure there is a heritable component to feed intake capacity. So if we would compose a very simple model, we can say that resource intake needs to be divided across anything related to survival on the one hand, and reproduction on the other. This is illustrated in figure 6A. We now put the animal in a more demanding environment. In order to survive it will need to obtain more resource for survival. If possible it will increase its feed intake. However, if that already was at its maximum, it will have to shift resources that were allocated to reproduction to survival (figure 6B). In this environment there is a negative correlation between survival and reproduction. Better survival is achieved at the expense of reproduction.

In domestic animals, the selection criteria can be considered a part of the ‘survival’ . After all, if they are not good enough, they are not selected and are no longer part of the breeding population. From breeding perspective they are ‘dead’. Dairy cattle, for example, is selected for high milk production. High producing cows often have more fertility problems than lower producing cows. We can use the resource allocation model to get some insight in why that may be. The cows cannot spend their resources on both survival (e.g. milk production) and reproduction at the same time. The cows that spend most on milk were at selective advantage (had higher survival potential). They could achieve this by increasing feed intake, but also by shifting resources away from reproduction, resulting in a negative correlation between milk production and reproduction. This is what may have happened.

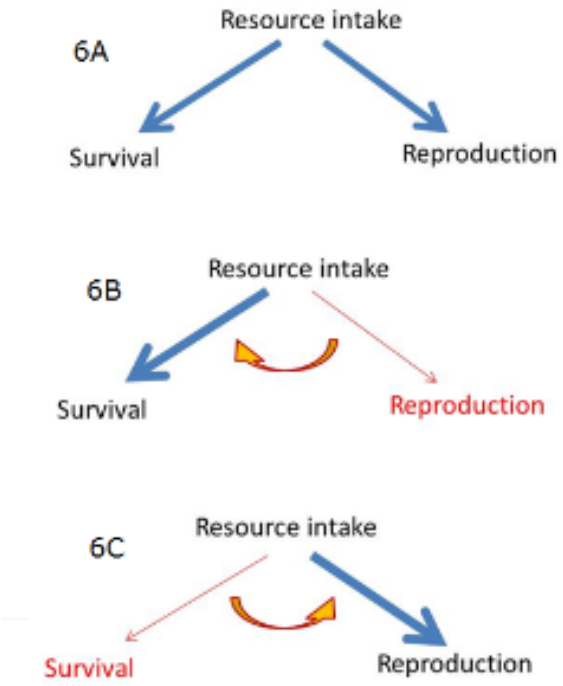


Figure 6. The simple model of resource allocation. 6A shows the basic model: resources need to be divided across traits related to survival and reproduction. 6B represents a situation where more resources are required to survive. In order to achieve that, resources need to be taken away from reproduction. This may lead to reduced reproductive performance. 6c shows the opposite situation: resources are required for reproduction, and are taken away from survival to achieve the required amount. This may lead to a conflict between survival and reproduction.

Chapter 13.6.5: Role of the environment

The environment can also be improved so that less resources need to be allocated to survival and more can be spend on reproduction, resulting in a shift towards reproduction (figure 6C). Animals that have the highest proportion of their resources allocated to reproduction will have the largest share in the next generation. So selection pressure automatically will be on the large proportion of resources allocated to reproduction. You can imagine that after a number of generations the animals that allocate more to survival have been reduced in number. In this environment that is not a problem, because only limited resources are required for survival. However, if you put those animals in a poorer quality environment, they will not manage anymore. A negative correlation has developed between reproduction and survival. This is another type of negative correlation than with the milk production and reproduction, even though the reasons for development are similar. In this case the animals are very reproductive, but need a good environment for survival. In the previous case the animals had very good ‘survivability’ (i.e. milk production), but at the expense of reproduction.

In our modern farm animal species both types of negative correlations occur. We have selected the animals for very high performance and, simultaneously, we tried to optimise their environment so that they could show their potential. By doing so, we have created animals that perform very well under optimised conditions. However, they also have become more sensitive to a decrease in environmental quality compared to what they were selected for. For example, broilers have become quite sensitive to fluctuations in ambient temperature. The range at which they feel comfortable is much decreased as compared to that of, for example, laying hens. Of course there is more than just resources involved, but the principle of not having to deal with a demanding environment so that all effort can be put in the selection criterion (growth) is applicable.

Chapter 13.6.6: Evidence for the resource allocation model

Even though this is a simple model, the directions of selection response do occur in reality. To illustrate that, in figure 7 you see the results of a selection experiment in mice. A population of mice was divided in two and kept on two different diets with respect to protein content for 6 generations. Then they were put in the opposite environment and their growth performance was measured. Both sub-populations performed best in their own environment. However, the group that was selected on the high protein diet suffered much more in the other environment than the group that was selected on the low protein diet. The results are presented as reaction norms in figure 7. The fact that they are non-parallel, and even crossing, shows that genotype by environment interaction can be created in only a limited number of generations. Populations, or breeds, adjust to their environment.

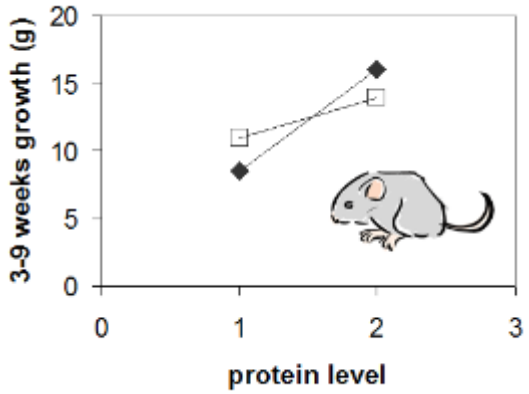


Figure 7. A population of mice was divided into two and bred on two diets with different protein level. In generation 7 their performance was recorded on both diets. Each population performed best on their own diet.

Chapter 13.6.7: Correlations and genotype by environment interaction

One last point, related to genotype by environment interaction, is that correlations between traits may differ between environments. They may even change sign. For example, the age at maturity (when animals are able to reproduce) is positively correlated to growth in a good quality environment. Animals that grow large mature later than animals that remain smaller. In a poor environment, however, this may be reversed. Even though there is no re-ranking in size, there is re-ranking in who matures first. The reason could be that the poorer growers in the good environment reached adult size earlier and could start to invest in reproduction earlier. In the poor environment, however, they may have struggled to grow to adult size for some reason, and started reproduction later. This is an extreme example, where the sign of the correlations even change. But it is good to realise that correlations between traits may change across environments. Good to keep in mind when using indicator traits for selection, for example. Results in one environment cannot automatically be translated to other environments.

Summarising, there are different reasons for the presence of genetic correlations. Some may develop as consequence of selection strategy, whereas others are based on the genome. It is important to be aware of the presence of genetic correlations, because it implies that selection on one trait will have consequences for other traits. We make use of that with indicator traits, but what to do with undesirable correlations?

Chapter 13.7: Solutions to undesirable correlations

Sometimes two traits are correlated in an unfavourable way. For example, milk production in cows (but also in other species) is negatively correlated with fertility. High producing animals often have more difficulties in becoming pregnant again. However, there are animals that combine high production with good fertility. If those are selected for breeding, milk production can be improved without decreasing fertility. The fact that two traits are undesirably correlated does not automatically mean that they cannot both be improved in the population. Unless the correlation between both traits is 1 or -1, there will always be some animals that have the desired genotype for both traits. Obviously, the genetic gain for each of the traits will be lower than if their correlation would have been desirable. Because some of the best animals for that trait will not be very good in the other and those should not be selected. So the selection intensity decreases.

As we have seen in the chapter on ranking the animals, selection on multiple traits simultaneously can be applied for traits that are both in the breeding goal. In chapter 2 on defining the production system and the breeding goal, you have learned how traits can be weighed into an index (single value). The weights can be the economic value of the trait: how much profit can be made from 1 trait unit genetic improvement? But sometimes economy is not the best way to weigh the traits under selection and desired gains would be a better option. For

example, if changes in the market or legislation are expected in the future, then selection weights can be defined based on how fast traits need to change to meet these expectations. Similarly, sometimes societal pressure is large to change the result of the breeding program. For example, broilers should grow less fast, or certain dog breeds should have be able to breath freely instead of having certain looks, or calves should be born without caesarean section. Even though from direct economic perspective selection weight should be small, the society demands otherwise. Apart from serving the consumers, so your market, taking the societal demand serious is also wise to maintain a good reputation. And a good reputation is very important to keep and increase market share.

Thus:

> *Even if a genetic correlation between two traits is undesirable, selection for both traits is still possible. Though genetic response will be affected by a lower selection intensity*

Chapter 13.8: Future expectations: where to go from here?

Breeding is about predicting the future. Changes in your breeding program made today will only show in a few generations from now. Of course it is not possible to predict the future in detail, but it is possible to foresee some general changes. For example, will there be changes in legislation in the near future that will affect your product? Changes in legislation with respect to housing, for example, will influence the type of animal that you will need to produce. Animals that have always been selected for being housed alone that suddenly need to live in groups most likely will not be very successful. Likewise, if export legislation is going to be changed, the breeding company needs to be aware of that. If de-horning, or tail docking, or beak trimming is no longer allowed, the breeding companies need to have selected animals that will manage well without those interventions.

Legislation

Notification of a change in legislation usually is made well in advance. However, changes in the market are much less predictable. What is the economic situation of your clients going to be like in the future? And will the market keep on demanding chicken breast or will it shift towards full carcasses, for example? Will dressage remain as popular as it is today or will the new equine star rider be a show jumper or an eventer, resulting in an increased demand for those horses? Predicting the market changes is a profession on its own.

Related to the market demand is the market share. If your breeding program is successful your market may expand. Are you prepared for the new demands of that market? For example, if you start to sell to other parts of the world, are the animals that are required there the same as

are required here? Is there a genotype by environment interaction to be expected? If so, how big is it? Do you need to create a new population for that purpose or can you manage with your current population? Is it possible to select your animals here or do you need to select your animals close to where your new market is so that they can adapt to the new environment?

Market developments

And what is your market position in relation to your competitor? Are you operating in the same market or can you go for a different market? What are your strong points and what are the strong points of the product of the competitor? It is important to have a good impression of the market as a whole: what is the demand, who are players on the market, what are their plans, what is the quality of your product, etc.

Science and technology

A final component that you have to keep an eye on are new developments in technology. New technology may make things possible that were not possible before. But, at least as important, if your competitor is using the technology and you are not, you may lose your advantage and then market share and then go out of business. It, therefore, is important to incorporate new technology related to, for example, measuring complex phenotypes or estimating breeding values using genomic information, even before it is completely clear what the benefits are. Because if you don't use it and your competitor does, and it appears to be a beneficial technology, you are too late. You will have lost the lead you may have kept if you would have used the technology from the start. In practice this is what you see happening a lot. New technologies are adopted before it is completely clear how beneficial they will be, because the competitors are using it too. To avoid being too late, it is important to jump in in time.

Thus:

> *The future with respect to limits (legislation) and opportunities (market expansion) can be predicted to some extent. It is important to do so.*

Chapter 13.9: Balance between progress and genetic diversity

As we have seen in the chapter on evaluating genetic diversity, there are a number of conflicts of interest between maintaining genetic diversity and achieving genetic response. The most obvious is the selection intensity. A higher selection intensity means selecting relatively few animals for breeding. Few animals for breeding results in higher rate of inbreeding, and thus loss of genetic diversity. Related to that: using the best animal(s) more intensely than other

animals may result in larger genetic gain, but is disastrous for maintaining genetic diversity in the future. Every breeding program needs to be aware of both sides of the coin. There is no ready-to-go solution for all, specific situations will require specific ways to balance genetic progress and genetic diversity.

Chapter 13.10: Key issues of the chapter on evaluation of the breeding program

1. Realised genetic improvement can be determined by taking the difference in average EBV

2. There are three main reasons for the occurrence of a selection limit:

a. Loss of genetic diversity (irreversible)

b. Opposing natural selection through reduced fertility or mortality (usually irreversible)

c. Limiting environment to express potential (often reversible)

3. A genetic trend represents realised genetic response across a period of time (e.g. year or generations)

4. Deviations from expected genetic trend may be due to a change in the breeding goal or in phenotype recording

5. Genotype by environment interaction = when ranking of animals based on their EBV differs between environments

6. A reaction norm represents performance of genotypes in a range of environments

7. Two environments require separate breeding programs if the correlation between performance in both environments is smaller than 0.6

8. Genetic correlations can exist for several reasons:

a. Pleiotropic effect

b. Linkage disequilibrium

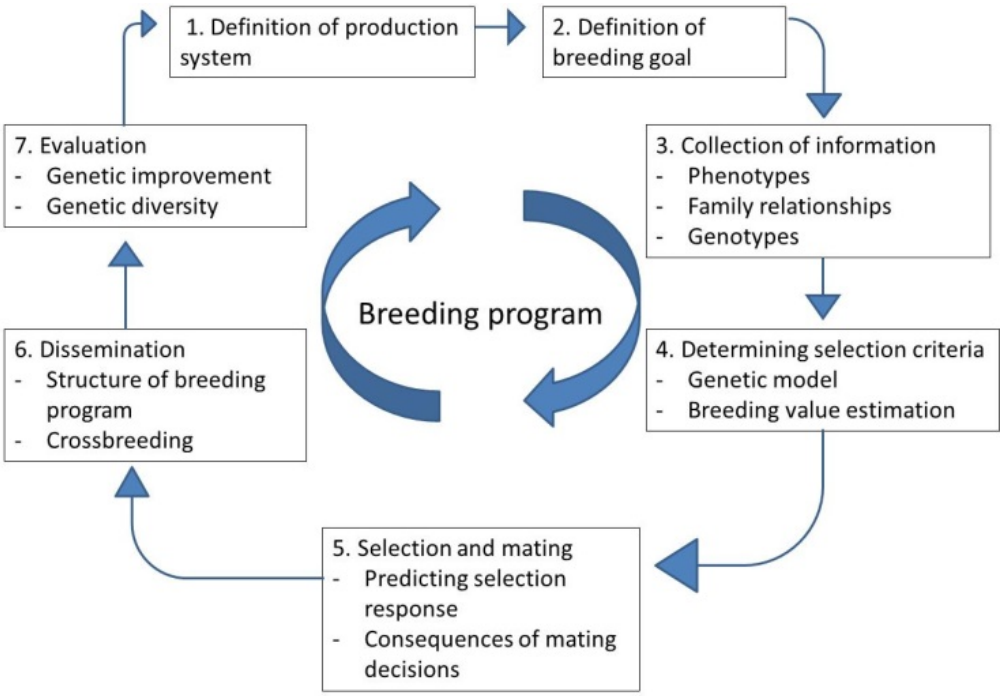
c. Conflicting resource allocation

9. Even if a genetic correlation between two traits is undesirable, selection for both traits is still possible. Though genetic response will be affected by a lower selection intensity

10. The future with respect to limits (legislation) and opportunities (market expansion) can be predicted to some extent. It is important to do so.

Chapter 14: Maintenance of genetic diversity

In previous chapters we have learned that breeding programs are set up to create genetic improvement. The genetic variance of the traits to be improved is of crucial importance. This aspect is fully recognized in the previous chapter where we evaluated the realized genetic improvement. In addition to the improvement obtained, the genetic relationship among animals in a population is important. When it increases as a consequence of the selection of parents, in the future strongly related animals have to be mated. Then, inbreeding effects become relevant: inbreeding depression and the occurrence of recessive genetic defects. Genetic variation of traits and variation in the composition of pedigrees is relevant in the evaluation of breeding programs. It are aspects of the genetic diversity that should be continuously evaluated when running a breeding program. The genetic variance of traits in a breeding program or a production program (pyramidal structure in pig and poultry breeding) is not limited to the genetic variance of the breed at stake. It can extended to the genetic variance in the species that can be utilized in a breeding program with crossbreeding. Thus attention for genetic diversity is not restricted to genetic diversity within a breed but comprises the genetic diversity within and between breeds. Therefore the conservation of breeds is important as will be explained too in this chapter. In succession we will discuss: what is genetic diversity, how can we measure it, what is the value of conserving breeds, the importance of the relationship among animals in avoiding inbreeding within a breed and how can we prevent excessive increases in relationship in a breeding program.



In writing this chapter two books were frequently used: “Utilization and conservation of farm animal genetic resources” (editor Kor Oldenbroek; Wageningen Academic Publishers, 2007) and “Het fokken van rashonden” (Kor Oldenbroek en Jack Windig, Raad van Beheer op Kynologisch gebied in Nederland, 2012; in Dutch).

Chapter 14.1: Genetic diversity

Worldwide mankind has domesticated more than 30 animal species for agricultural purposes (14 of these species are responsible for more than 90 % of the food produced by animals). In addition other animal species are domesticated for leisure purposes or for a variety of other services for mankind (hobby, guarding, nature management, hunting etc.). Within these animal species a lot of variation is observed. The animals of a species differ more or less in a lot of traits: they show diversity in nearly all traits. A diversity that has a genetic origin.

Within species we recognize landraces: within a landrace animals resemble each other but among individuals you still may observe diversity in a lot of phenotypic traits. Out of the landraces mankind created standardized breeds (and later on, out of these, special selection lines). In the standardized breeds animals resemble each other more than in landraces. They are more uniform; but still among individuals of standardized breeds diversity can be observed. To conclude: within animal populations (species or landraces or breeds or selection lines) diversity exists that has a genetic origin. The origin is determined by the fact that animals differ in their DNA composition: within a species more than in a landrace, within a landrace more than in a standardized breed and within a standardized breed more than in a selection line. A broad definition of genetic diversity is:

Definition

Genetic diversity *is the set of differences between species, breeds within species, and individuals within breeds expressed as a consequence of differences in their DNA*

Chapter 14.1.1 Genetic diversity in dogs

The dog is eminently illustrative for the concept of genetic diversity. Mankind has domesticated wolves. After the domestication people kept “village dogs” belonging to landraces developed by natural selection and adapted to the local environments where they were kept. In this way, among the landraces, differences in traits developed based on differences in DNA. In the past 200-150 years standardized breeds were developed that differ greatly in phenotype: conformation (e.g. for body weight from 1 to 100 kg), hair coat and coat colour and in behaviour. Worldwide hundreds of standardized breeds were developed. These developments were facilitated by the fact that in the dog’s species a few genes with multiple alleles are responsible for these conformation and appearance traits. Thus dogs share similar DNA

because they all have the wolf as ancestor and belong to the dog's species, but differ in alleles for certain genes that creates the diversity among breeds (and to a lesser extend within breeds).

Chapter 14.1.2 Genetic diversity in farm animals

In farm animals, the variation between breeds in performance is important in the start of a breeding program or farm activities. What will be the best breed given the production circumstances and which breed fits the best to our defined breeding goal? In the past 50 years the observation of substantial breed differences, especially in cattle, pigs and poultry, has led to a sharp selection of breeds, that were still developed. Many breeds and selection lines were set aside in this process. Specialization took place in animal production and only a few breeds were considered to be the best in milk, beef, pork, egg or broiler production. This global concentration on a few breeds is strengthened by the application of modern technology in breeding programs requiring high investments. Such investments are only profitable when the genetic improvement realized in breeding programs can be largely disseminated: when a lot of breeding material can be sold and the genetic improvement becomes profitable in a lot of offspring. The concentration worldwide on a limited number of breeds leads to an increasing number of breeds being considered unprofitable and consequently leads to risk of extinction.

Chapter 14.1.3 Importance of variation between breeds

The key question in these developments is: what proportion of the genetic variation lies between breeds? If this parameter is typically small then it may be expected that the variation within breeds can be utilized in selection programs to overcome weaknesses in traits when they arise. If the between breed variation is large then it may be not at all realistic to expect an adequate selection response within breeds. This is an important reason to be alert on the chance of extinction of breeds.

Example: At the moment a lot of breeders of Holstein Friesian cattle start crossbreeding with Swiss, German, French and Scandinavian dual purpose breeds to improve the health and fitness traits of their cattle. They experience a deterioration of these traits and do not expect that they can stop this process when they continue pure breeding with Holstein Friesians. The French and Scandinavian breeds do have better health and fitness traits and in stable (rotation) crossbreeding systems these traits get an additional boost from heterosis.

When the required livestock production has to adapt quickly to new challenges, then the variation between breeds might be of great help. Therefore conservation of breeds with a diversity of traits is a rational and important strategic response to the uncertainties in production circumstances and market developments today.

Broadly it appears that variation between breeds accounts for approximately half of the total genetic variation within a species. In formula form:

$^2_S = 0.5\ ^2_B + 0.5\ ^2_W$, where

^2_S = genetic variance within a species

^2_B = genetic variance between breeds

^2_W = genetic variance within breeds

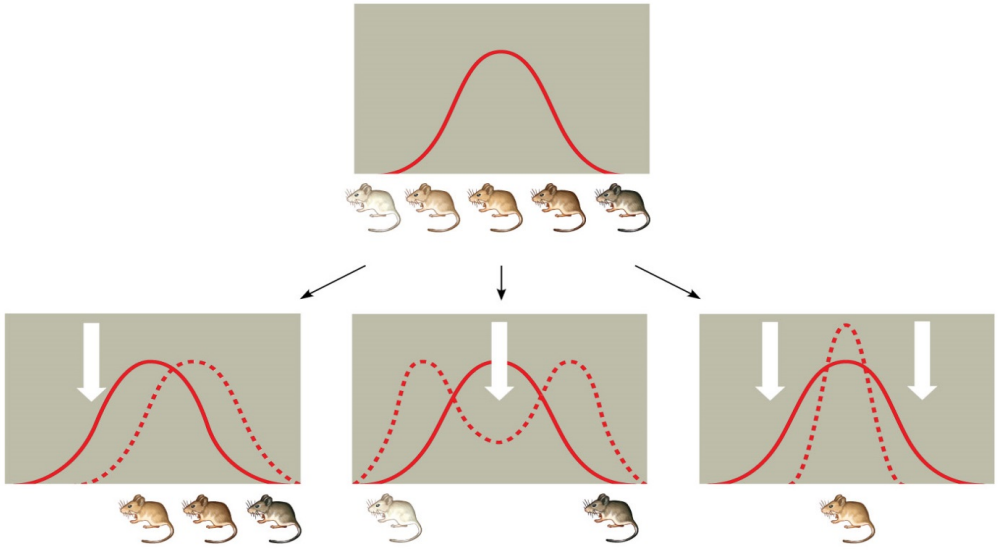
Chapter 14.1.4 Origin of differences between breeds

The differences between breeds have been developed through a combination of four evolutionary forces: genetic drift, migration, selection and mutation.

Genetic drift is a term for the random fluctuations of allele frequencies due to random sampling processes involved when genes are passed from parent to offspring, and is one of the phenomena linked to inbreeding. It plays a larger role in smaller populations. Over time genetic drift will lead to increasing genetic differences between two breeds drawn from the same population and then maintained in isolation.

The *migration* of individuals moving from one breed to another, acts against inbreeding, since it lessens the genetic differences that exist between the breeds, and increases the variation within the recipient breed.

If *selection* occurs, carriers of favourable alleles have a selective advantage in the next generation. Selection may lead to convergence or divergence between breeds depending on the selection goals used in each breed. In livestock, selection can be both artificial and natural; for example, natural selection will have played an important role in improving adaptive fitness for particular breeds kept over many generations in environments with specific challenges e.g. periodic droughts. The selective advantage of favourable alleles is illustrated by the figure below:



When coloured mice have a selective advantage the alleles for less coloured mice disappear in future generations, when brown mice have a disadvantage their brown alleles disappear and white and grey mice survive, and when brown mice have advantages the alleles for white and grey disappear.
In general, *mutation* in the genome increases the genetic differentiation between breeds and creates genetic diversity. However, mutation occurs with a low frequency and, in the absence of selection, the influence of mutation becomes measurable only over a relatively large number of generations. However at some point in the past, mutation has been responsible for creating the polymorphisms that lie at the heart of all genetic diversity.

Chapter 14.1.5 Origin of differences within breeds

Within breeds genetic drift, migration, selection and mutation are also relevant actors. In addition to these evolutionary forces the way the breed is created is essential for the genetic variation within the breed established today.

E.g. in dogs (standardized) breeds were created in crossing often only a few animals from a limited number of breeds. Their offspring was selected according to a strict breeding standard. A dog breed is often based on a limited number of founder animals and this is the cause that often a limited genetic variation within dog breeds is established.

E.g. in pigs and cattle, breeds were developed from landrace breeds by culling the animals that did not fit into the breeding standard (not the right colour or the wrong conformation) and by promoting the use of males that highly fitted the breeding standard.

Genetic drift can be avoided when breeds are kept in large numbers.

Migration of animals has often a positive effect on the size of the genetic variation within a breed. In practical terms: when studbook regulations allow that animals from outside the breeds can be used (under certain regulations) in that breed, the genetic variation will be enlarged. Therefore it is highly recommended to work with ”open” studbooks instead of closing the breed for animals from outside.

Selection of animals as parents for the next generation might also have a rather high negative impact on the genetic variation within a breed when the selection is very intense. Then only a few parents are selected that determine the genetic variation in the next generation.

Mutation is at short term of minor importance for the genetic variation within breeds. The mutation rate is estimated to be that low, that in a short period of time the number of animals within the breeds is too small to have the opportunity that a mutation takes place.

Chapter 14.2: Example: FAO's global plan for farm animal genetic resources

In the sixties of the past century, scientific and farmer communities draw attention to the high rate of erosion of animal genetic resources. In Europe, farmers were leaving the rural areas where much breed diversity was present and many local breeds were replaced by a few highly promoted and intensively selected breeds. These intensively selected breeds were also exported to developing countries outside Europe and replaced breeds which were well adapted to circumstances and management systems deviating sharply from those inEurope. In 1992 FAO launched a special action program for the Global Management of Farm Animal Genetic Resources and in 2007 this was replaced by a "Global Plan of Action" after publishing the state of the animal genetic resources. In the FAO terminology animal genetics resources refers to the number of breeds within a species. The situation of the breeds at global level is illustrated in the figure at the right:

A lot of attention is given by FAO to define the risk status of breeds. That is not simply a matter of numbers of animals. Of course that is a main criterion, but there is more: are females only used for pure breeding or are they more or less used for crossbreeding? A decisive factor is the reproductive capacity of the breed: does a female produce hundreds of offspring per year as is the case in commercial poultry breeding or does a female produce on average one replacement in ten years as is the case in horses? FAO uses the criteria: not at risk, vulnerable, endangered and critical to qualify the risk status of a breed. In the figure below it can be seen how this works out for breeds of species with a high and low reproductive capacity:

Based on the risk status of a breed different management strategies are appropriate to use the breed and to safeguard its conservation. FAO has developed a nice flowchart to find the appropriate strategy for a breed dependent on its risk status. For the breeds at risk first the value of a breed is considered: e.g. for the relationship with other breeds (is it a unique breed?), the presence of special adaptive traits, the value of its use in the society and the cultural historic value. After this consideration you may conclude that the breed is worthwhile to conserve and a conservation program will be put in place. This might be an *in vivo* or an *in vitro* conservation program.

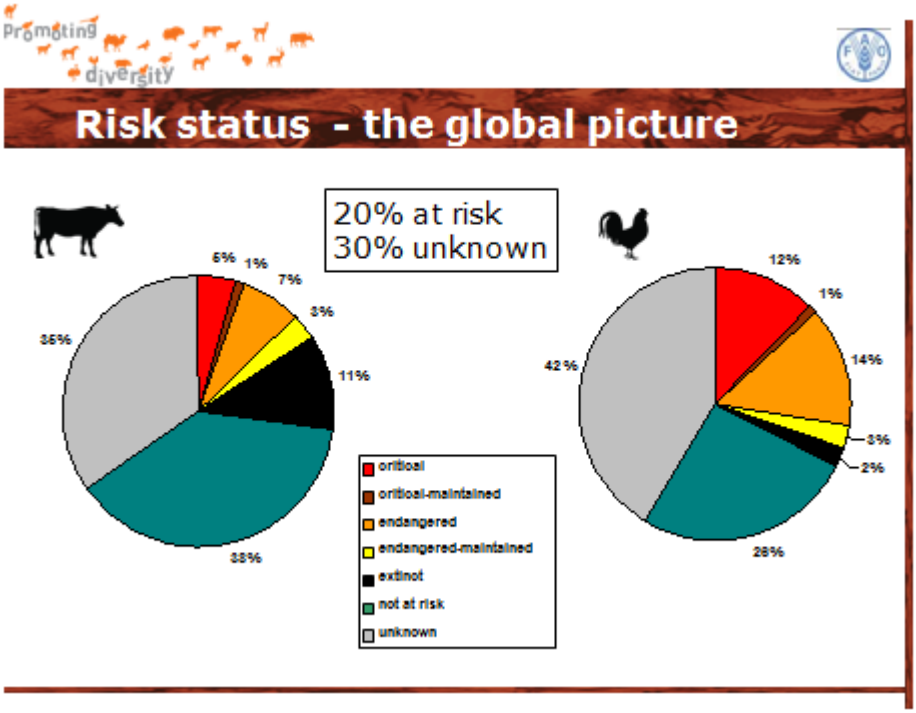
Definitions

*In **vivo conservation** is conservation through maintenance of live populations kept under normal farm conditions and/or inside of the area in which they evolved or are normally found.*

*In **vitro (cryo) conservation** is the storage of gametes of embryos in liquid nitrogen*

For the breeds not at risk or potentially at risk genetic improvement is still possible in breeding programs. Of course, the possibilities are limited for breeds potentially at risk, due to a low number of animals that can be used for breeding. For these breeds, conservation programs might be developed. These are breeding programs where minimising the relatedness among parents has the first priority, instead of maximising genetic improvement. As will be explained in detail later in conservation programs a relatively high number of sires and dams have to be selected as parents for the next generation. In conservation programs long generation intervals are stimulated and sires with semen conserved in a gene bank may be used when they appear to have a very low number of offspring in the population alive today.

The *in vivo* conservation of a breed requires a well-designed breeding program that accounts for the small number of animals and that is strictly followed by the breeders and evaluated frequently. The main objective for *in vivo* conservation is to facilitate the use of the breeds in the rural area for 1) nature management, 2) production of regional products with a high additive value and 3) the maintenance of cultural historic activities.



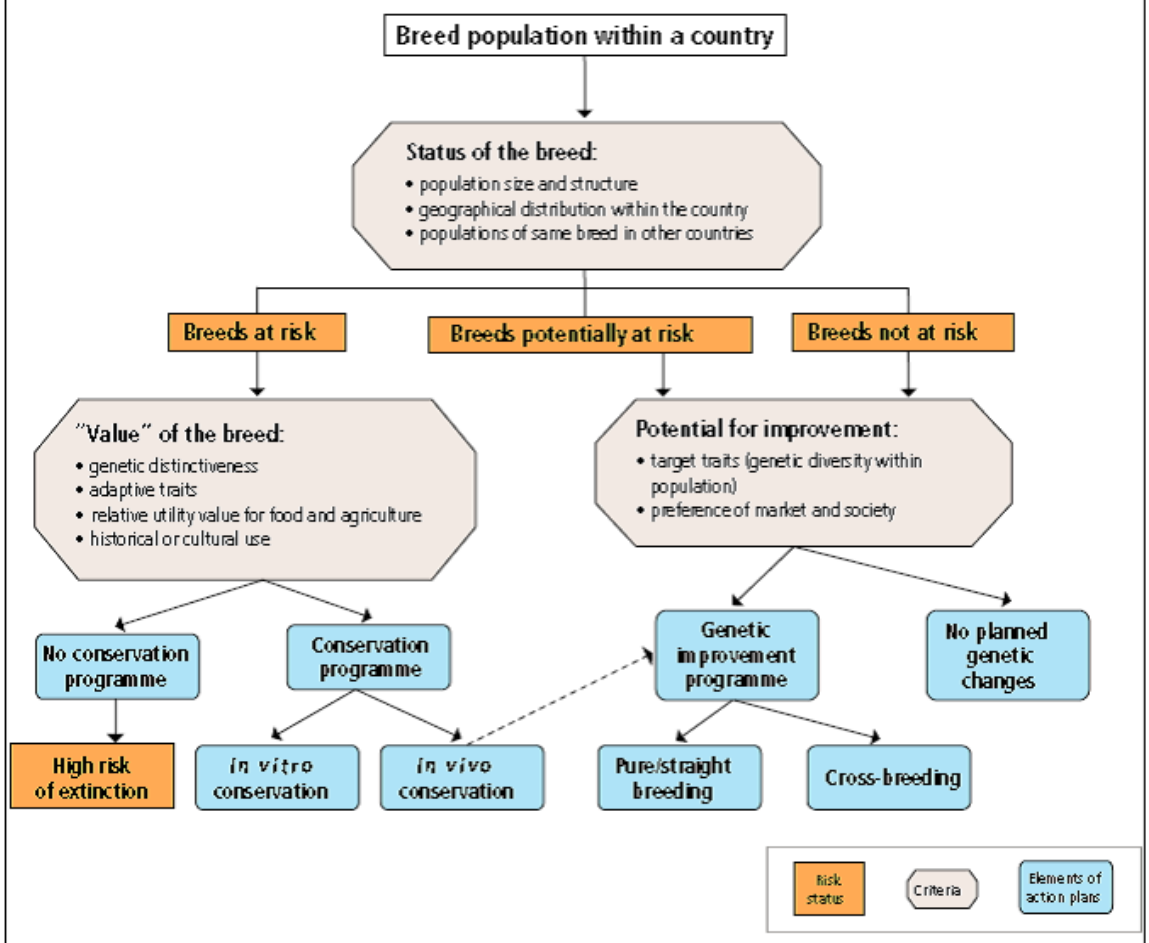
Risk categories according to species' reproductive capacity

Reproductive capacity	Males (n)	Breeding females (n) ≤100	101 - 300	301 - 1 000	1 001 - 2 000	2 001 - 3 000	3 001 - 6 000	>6 000
High*	≤5							
	6 - 20							
	21 - 35							
	>35							
Low**	≤5							
	6 - 20							
	21 - 35							
	<35							

= critical, = endangered, = vulnerable and = not at risk.

*High reproductive capacity species = pigs, rabbits, guinea pigs, dogs and all poultry species.
**Low reproduction capacity species = horses, donkeys, cattle, yaks, buffaloes, deer, sheep, goats and camelids.

FIGURE 1.
Flow chart for national management of animal genetic resources



Note: Originally published in FAO. 2007. *The State of the World's Animal Genetic Resources for Food and Agriculture*, edited by B. Rischkowsky & D. Pilling. Rome (available at http://www.fao.org/docrep/010/a1250e/a1250e00.htm).

Chapter 14.2.1: An example of the use of a conserved breed for research

Colour sidedness: an example of genetic diversity conserved for research

Colour sidedness is a dominantly inherited phenotype of cattle characterized by pigmented areas on the flanks, snout and ear tips. It is also referred to as "lineback" or "witrik" (which means white back), as colour-sided animals typically have a white band along their spines. In several countries, animals are specifically bred for this colour pattern, and thus the trait is conserved. Colour sidedness has been documented at least since the European Middle Ages and is presently segregating in several cattle breeds around the world, including Belgian Blue, some Nordic breeds, Dutch Witrik, American Randall Lineback and Brown Swiss. By genotyping animals from several colour-sided breeds and comparing the data to those from a breed lacking this trait, scientists in Belgium were able to determine that colour sidedness in cattle is caused by segments of the genome that have been duplicated and exchanged between chromosomes 6 and 29 (Durkin *et al.*, 2012).

This study marked the first example of a phenotype determined by duplicated genes found on separate chromosomes. The maintenance of several cattle breeds with the colour pattern facilitated the detection of this genetic mechanism, previously unknown in mammals.

Chapter 14.2.2: An example of the use of a rare breed

Doubling the price of Drenthe Heath lambs in the Netherlands

The Drenthe Heath sheep arrived in the northeastern part of the Netherlands 6 000 years ago. They were kept and survived on this region's infertile sandy heathlands. Through adaptation and natural selection the Drenthe Heath sheep became a rather small animal with sturdy legs and low fleshiness. As a result, the carcass weight and the meat-to-bone ratio are low relative to standard meat-sheep breeds. It is the only Dutch sheep breed with horns. Nowadays, the flocks are primarily used for nature management. They are guided by shepherds, a sight that tourists visiting the area find very appealing. Approximately 2 000 ewes are registered in the Drenthe Heath Sheep Herdbook. Recently, the owners of three flocks have started to market their lambs as *Drènts Heidelaom,* an organic product. The lambs are produced in a well-defined market chain. This has doubled the price the shepherds get for lambs, relative to the anonymous lamb market.

The production chain was set up as follows. First, the organic management of the flock and the organic growing of the lambs until slaughter were organized. These management practices are controlled and verified by Skal*, the Netherland's official certification and inspection body for organic production. Second, a small local abattoir was contracted to slaughter the lambs in the most humane manner possible. Third, arrangements were made for the carcasses to be transported and sold to a specialized butcher producing organic lamb chops, ham of lamb and lamb sausages. These products are sold by the butcher at organic farmers' markets in cities in the western part of the Netherlands. Fourth, together with the Foundation for Conservation of the Drenthe Heath Sheep, the Slow Food organization in the Netherlands was consulted. Because of the special natural management and nutrition of the sheep and lambs, Drenthe Heath lamb has a special "wild" taste. Because of this and the cultural–historic significance of the sheep and the product, Drenthe Heath lamb was recognized by the Slow Food organization as part of the Ark of Taste (a catalogue of heritage foods that are often at risk of extinction). Fifth, arrangements were made for collaboration among flocks and this resulted in a "Presidium"** of the Slow Food organization: Drenthe Heath Lamb or in the language of the region, *Drènts Heidelaom.*

---

* http://www.skal.nl/english/tabid/103/language/nl-nl/default.aspx

** A "Presidium" is a small project to support groups that champion the production and marketing of an artisan food that addresses economic, environmental, cultural and/or social objectives that are considered favourable by the Slow Food organization.

Chapter 14.3: The conservation of breeds in The Netherlands

In vivo

In The Netherlands, the Dutch Rare Breed Foundation (in Dutch: Stichting Zeldzame Huisdierrassen: SZH) stimulates the *in vivo* conservation of the Dutch Native Breeds. Over 70 breed organizations exist and are connected to the SZH. Native is defined as present and bred in the country for more than 6 generations plus 40 years. The number of breeds in the large (farm) animals is rather limited: 7 cattle breeds, 2 pig breeds, 4 horse breeds, 8 sheep breeds, 3 goat breeds, 9 dog breeds and 7 rabbit breeds. There are numerous native Dutch breeds of chicken and other birds (e.g. goose and pigeon).

The SZH provides services to the breeding organizations and breeders in three different ways: 1) monitoring breeds and development and evaluation of breeding programs, 2) creation of awareness and of education materials and 3) stimulating the use of rare breeds in nature management and in the production of regional products for niche markets. The main genetic problem in the rare breeds is the relatedness among the animals in the small populations and the threat for inbreeding. In close collaboration with the Centre for Genetic Resources The Netherlands (CGN) breed organizations get advise how to minimise the increase in relatedness and inbreeding in these small rare breeds. In close collaboration SZH and CGN conserve semen from males from the rare breeds and in a few breeds, the breeding program is supported by the use of “gene bank” semen.

In vitro

In The Netherlands, the Centre for Genetic Resources The Netherlands (CGN) is responsible for the *in vitro* conservation of the genetic diversity comprised in the breeds of farm animals currently present in The Netherlands. It are the widely used breeds or selection lines from breeding companies and the rare native breeds. The CGN has a gene bank in which primarily semen is stored in liquid nitrogen. In cattle it are samples of bulls from breeds widely used (25 doses of each bull entering the breeding program) or from rare breeds (400 doses of selected /available bulls, collected at farms were they are used for natural service). In pigs, once every 10 years a snapshot is taken from the boars at AI stations, boars of different selection lines or from the two rare native breeds. In poultry and birds semen from males from rare breeds is stored. In dogs a start has made to freeze semen from males of rare breeds. The next slide gives an overview of the gene bank collection in 2013:

In vivo and *in vitro* conservation is complementary: when you do both, you guarantee flexibility of breeding programs by safeguarding rare breeds and have the possibility to use them for present activities. Gene banks can have an important role in the support of small rare breeds and can help to make these populations viable again. This is illustrated by the example of the Dutch Red and White Friesian cow:

Chapter 14.3.1: Revival of the native Dutch Red and White Friesian cow

Around 1800 the cattle population in theprovinceof Friesland consisted mainly out of Red Pied cattle. Many red ancestors were imported from Denmark and Germany after outbreaks of “rinderpest”. Since 1879 the Friesian Cattle Herd book registered Red and White animals. Later

on, pushed by export markets, black and white became a more popular color than red and white. For Black and White breeders it was a shame to obtain a red and white calf out of black and white parents. A sire with a strong influence in the Red and White population was named “Foundling” and was abandoned by its Black and White breeder.

In 1970 only 50 farmers with 2500 animals joined the Association of Red and White Friesian cattle breeders. Then, after a short period of specialization and intensification in dairy production (1970-1990) and the import of Holstein Friesians from theUSAandCanada, in 1993 only 21 purebred animals were left: 17 females and 4 males. A group of concerned owners started the Foundation for native Red and White Friesian Cattle and this foundation contacted the Gene Bank for Animals, which was just founded.

A breeding program was jointly developed. Semen from “old” sires in the gene bank was used to breeding females under a contract. Males subsequently born were raised by the breeder, who got a subsidy from the gene bank. Semen of these young males was collected and conserved and could be used in new contracts. In this way, the breed increased in numbers: in 2004 256 living females and 12 living males were registered. From 43 males 11780 doses of semen are stored by the gene bank and are available for AI. Only a few females are still used for milk production and their milk is used for cheese production. The majority of females is kept by hobbyists as suckler cow.

Chapter 14.4: The use of pedigrees for measuring genetic diversity

Theoretically, we can estimate the degree of diversity between breeds by simple, often costly, experiments in which animals of a large number of different breeds are kept together in the same environment. Providing: 1) the numbers of animals per breed are sufficiently large, so that the errors in estimating the breed mean are negligible compared to the scale of differences between breeds, and 2) the breeds are a fully representative sample of the breeds available, the variance $_B^2$ can be derived from the breed means. Which environment for testing and how different the answers would be in other environments are important research questions that have global implications for agriculture, the environment and conservation. Given this uncertainty, it is important that testing environments are directly relevant to the intended environment for implementation. In the seventies of the past century a lot of breed comparisons were done globally in dairy and beef cattle and in pigs. However, most of the experiments were carried out with a limited number of breeds in different environments. This fact made the calculation of between breed variance less reliable.

Quantifying the amount of genetic variation in a trait within a breed is also difficult and involves associating known genetic similarities between individuals with similarities in phenotypes. A major source of the reliable information on the relatedness is the pedigree, i.e. a record of sire and dam for each individual, accumulated over generations. In the absence of detailed information on DNA from individual animals, which will continue to be the case for most populations for some time into the future, there is a need to identify these relationships through observing and recording the pedigrees of animals, at least in sufficient depth to identify sires and dams that are responsible for these relationships among individual animals. The deeper the pedigree, the better the true relatedness can be calculated. Hence, in each generation the number of parents increases exponentially ($2^{n)\cdot}$ It is generally accepted that a complete pedigree in five generations is required to establish relatedness. As parameter for the quality of pedigree information in a population, the *pedigree completeness* is used. It can be calculated as the percentage of complete pedigrees e.g. in five or six generations.

Chapter 14.5: The impact of DNA information on measuring genetic diversity

The last decade has seen the cost of genotype information reduced enormously, making such information much more affordable for science and for commercial applications. This is opening up new opportunities for evaluating genetic diversity. A number of different marker types have been used in scientific studies and their popularity has changed with advances in technology.

Informative DNA markers can therefore help the measurement of diversity in two ways. The first way is to overcome the problem that in some species it may be impossible or very costly to observe pedigree directly, e.g. in many fish species; and by genotyping a small number of polymorph markers (say 10 to 20 micro satellites), chosen to be informative, on all offspring and all possible parents, then it is possible to identify the sires and dams of almost all the offspring. The second involves the extensive genotyping (say 50.000 SNP's) across all chromosomes of the genome in order to estimate the actual proportion of DNA shared by sibs or other relatives more precisely than simply using the expectation for sharing DNA between relatives that is provided by the pedigree.

However, the availability of DNA allows us to measure diversity in different ways, since we can obtain the nucleotide sequence of individuals in specific areas of the genome and identify the alleles that are present in a population at each position and the genotypes of each individual.

Options for addressing genetic diversity with this information include the following:

1. *Examining the diversity in allele frequency*, by defining an allele frequency for an individual as 0, ½ or 1 depending on whether it carries 0, 1 or 2 copies of the allele. This trait can be treated as it was a continuous trait, and diversity measured, both between and within breeds. Note that in this approach the breed mean is the estimate of the allele frequency for the breed. As an example, if two breeds are fixed for different alleles then no diversity will be observed within breeds and all the diversity will lay between breeds.

2. The breed means for the frequencies of several alleles, usually from unlinked loci, are combined by some pre-defined function to measure what is called *a genetic distance between the breeds*. There are several such distance measures , that will not be discussed here.

3. Instead of using gene frequency, *the frequency of heterozygotes* may be measured. A heterozygote has two different alleles at a locus, and is function of the allele frequencies and the mating of relatives and survival rates. The justification for this is that in the absence of diversity there will be no heterozygotes in the population. Reliable comparisons between breeds must be made on very dense sets of markers, which may be possible in a lot of species, where DNA chips can contain in excess of 50,000 markers.

4. A further simple but limited measure of diversity is *counting the number of different alleles* appearing in the population for a set of loci, with the more alleles the more diverse. Counting the number of alleles in each breed and the number shared with each other breed offers an opportunity of examining differences between breeds. A variation on this is to count the number of 'private alleles', where *a 'private allele' is defined as an allele found in one breed but in no other*.

The counting approach to measuring diversity appears less valuable than measuring the allele frequencies themselves. Nevertheless observations on private alleles can be very useful in other ways, for example in studies where the conservation of breeds is at stake and in traceability schemes (is this meat indeed produced by a breed with this rare allele?).

Chapter 14.5.1: Genome-wide patters of diversity

An important assumption in studies with markers using ”non-functional” DNA (not responsible for the production of a protein) is that these markers are neutral i.e. are not associated with alleles for traits under selection. The issue over the neutrality of the markers is important since it is assumed that these markers change in frequency only by genetic drift, rather than by drift *and* selection. The neutrality of a locus may differ between breeds since: (i) one breed may have important alleles that are not present in another; and (ii) different livestock breeds will be subject to different selection criteria, with these largely dominated by the selection objectives of the breeders concerned.

The genome is organised into chromosomes and this introduces the phenomenon of linkage of alleles on different loci (when no recombination between these loci occur during meiosis, these alleles are transferred as a fixed combination of alleles from one generation to the next). One consequence of linkage is that alleles that are on the same chromosome and close to a new favourable mutation will be tend to increase in frequency alongside the mutation in a process termed *‘hitch-hiking’*.

It is very likely that the alleles very closely linked to the mutation will also become fixed in the population. Therefore this region of the chromosome, very close to the locus under selection, will display very low diversity in the neighbouring loci within a breed. An examination of allelic diversity throughout the genome may therefore show patterns of regions of high diversity and by regions of relatively low diversity. This pattern of diversity *within* the genome is called a *signature of selection* or *selection footprint* and may indicate loci important for domestication, or for the characteristics of particular breeds, or simply highly-conserved regions for the species as a whole, whether wild or domesticated. Effective searching for selection footprints is only just beginning in livestock species with the availability of genome-wide markers such as SNPs.

Definitions

Hitch-hiking *is the change in the frequency of an allele due to selection on a closely linked locus with a positive allele*

Signature of selection or selection footprint *is the pattern of reduced diversity neighbouring on a chromosome to a gene that has been strongly selected for or against within a population*

More generally, the expansion in DNA information will allow the diversity of allelic combinations at loci distributed throughout the genome to be studied. This type of diversity within breeds will depend not only on the allele frequencies but also on the extent of linkage of alleles disequilibrium (LD) that is observed. This LD, the transfer of a fixed combination of alleles from

one generation to the next, may arise from the breed history of population size and population management over time. E.g. in the past such that combination of alleles was present in the limited number of animals used for breeding in one generation (the intensive use of a popular sire or a genetic bottleneck) or were introduced in the breed by introgression.

Definitions

Linkage *is the phenomenon by which alleles at loci that are close together on a chromosome and which have been inherited together from one parent of an individual tend to be passed on together to an individual's offspring. The closer the loci are on a chromosome the stronger is this phenomenon. When the loci are on different chromosomes then this tendency is completely absent*

Linkage disequilibrium *is a fixed combination of alleles in haplotypes. Over time recombination events between loci will remove this combination, more quickly the further away the loci are from each other*

Bottleneck *is a period when the number of parents used to reproduce the breed was particularly small. In such a period the genetic drift is high due to a marked reduction in the size of the population*

Introgression *is transfer of an allele or set of alleles from one breed to another. This is achieved by the crossing of a number of parents from the donor breed to the recipient breed, followed by systematic backcrossing to the recipient breed, with parents chosen to be carriers of the desired alleles. Markers can be used to detect these carriers*

Chapter 14.5.1.1: An example of haplotypes: B-blood groups of cattle

Examples that combinations of linked alleles, a haplotype, was present in a few animals intensively used, can be found in the literature on cattle blood groups. The blood group B in cattle is determined by linked loci on chromosome 12. Each of the 20 different loci is responsible for the production (or the absence) of an antigenic factor, a protein, that can be established in the lab. Over 300 B blood groups are determined in cattle that differ in the combination of antigenic factors. It was found that B blood groups of intensively used sires and their son gave a sharp increase of the frequency of that blood group in the population. The alleles for the different antigenic factors are situated as follows on chromosome 12.

__Q__Y_2__G__D'__G'__G"__F'---'F'$_1$__BKP'__I_1__J'__K'__I_2__O_1__O_3A__I"__I'__

E.g. when a intensively used sire with the haplotype BO_1Y_2D' is intensively used, the frequency of this combination of linked alleles, this haplotype responsible for this B blood group increases in the population.

Chapter 14.5.1.2: An example of introgression: Booroola allele in Texel sheep

An example of the introgression of linked alleles was found in the Texel breed, where Merino sheep carrying the Booroola allele were used for introgression. The Booroola allele increases litter size: when heterozygous present in a ewe + 1 lam and when homozygous + 2 lams. In an experiment Texel ewes were mated to Merino rams carrying the allele. The F_1's were mated (backcrossed) to animals of the Texel breed. The intention was to incorporate only the Booroola alleles while maintaining as much as possible all other alleles of the Texel breed. This process of introgression was accelerated with the use of a genetic marker for the Booroola allele. In the experimental phase the presence of the Booroola allele revealed to give rise to a marked increase in the percentage of lams born dead. In subsequent generations this unwanted effect of more dead lams disappeared slowly. The explanation was that in the Merino sheep used, very close to the Booroola gene a gene was linked on the chromosome that gave rise to the higher percentage of lams born dead. With selection for a high litter size in combination with selection against dead lams, the linkage between the alleles in the population disappeared slowly.

Chapter 14.6: Monitoring populations

Breeding programs should not only be evaluated for realized genetic improvement but also for the amount of inbreeding. Inbreeding takes place when related animals, that do have an additive genetic relationship different from zero, are mated. Their offspring is inbred. And as we have seen, the inbreeding coefficient is equal to half the additive genetic relationship of the parents. Inbreeding may lead to the expression of monogenic recessive defects and to inbreeding depression. Inbreeding gives homozygosity at many loci and then favourable dominance effects disappear. In this respect inbreeding has the opposite effect of crossbreeding. There we have seen that crossbreeding leads to heterosis, especially in health and fittness traits. Signs of inbreeding depression is also higher for these traits. Examples of inbreeding in production traits are shown below:

Inteeltdepressie in de landbouw

Species	trait	Afname door 10% inteelt
koeien	Melkgift	3.2%
schapen	Vachtgewicht	5.5%
	Lichaamsgewicht	3.7%
varkens	Aantal biggen	3.1%
	Lichaamsgewicht	4.3%
muizen	Aantal jongen	7.2%
	Lichaamsgewicht	0.6%
mais	Planthoogte	2.1%
	Zaadopbrengst	5.6%

Chapter 14.6.1: Voluntary and constrained inbreeding

Inbreeding can take place in two different ways in a population: 1) breeders mate intentionally a sire and dam that are more related to each other than on average in population. This is called voluntary inbreeding. 2) breeders have to mate sires and dams that are related due to the fact that all animals in the population are related to each other. This is called constrained inbreeding and will be the main type of inbreeding we will discuss when monitoring populations. Constrained inbreeding is caused by the limited size of closed breeding populations as is the case in many breeds.

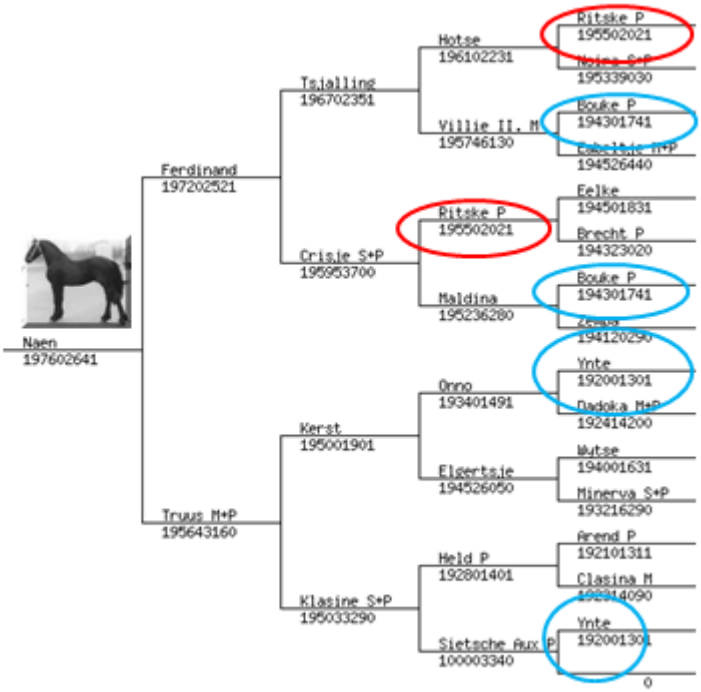
In each generation further in the pedigree the number of ancestors increases exponently. E.g. in generation 10 an animal has 2^{10} = 1024 ancestors. In most breeds in the period the ancestors of generation 10 lived, less than 1024 animals were used for breeding. Thus further in de pedigree the same animals pop up in the pedigree of the sire and the dam: they are related and therefore their offspring becomes inbred. This emphasizes that the deeper the pedigree, the better the relationship between sire and dam can be established. As (constrained) inbreeding is indeed an issue in dog breeding, most examples we will present, will originate from dog breeds.

Chapter 14.6.1.1: An outcross is very effective to reduce inbreeding

Inbreeding takes place when related animals, a sire and a dam, are mated. Then their offspring is inbred. When the sire is inbred and the dam is inbred, but they share no common ancestors, they are not related, thus their offspring is not inbred. In other words inbreeding is not heritable. When a male from another breed is used in a population that male has no common ancestors

with any of the females in the breed at stake. Even when this breed has a high average coefficient of inbreeding all the offspring of the “foreign” sire has an inbreeding coefficient of zero. An outcross is very effective to reduce inbreeding and to reduce the inbreeding problems. The outcross can be illustrated with the pedigree of an individual animal where the sire and dam are not related.

In the pedigree of Naen below, the sire Ferdinand is inbred: his parents Tsjalling and Crisje share Ritske P and Bouke P as common ancestors. The dam Truus is inbred: her parents Kerst and Klasine share Ynte as common ancestor. But, based on the 5 generations in this pedigree the son Naen of Ferdinand and Truus is not inbred because Ferdinand and Truus are not related: they do not share a common ancestor. Practical breeders call such a mating within a breed often an “outcross”.



Chapter 14.6.2: Population size

The population to monitor consists of all animals that can be mated: the maximum size of the breeding population. Looking backwards in the history of a breed its population size varies.

Therefore the first parameter to monitor is *population size*. A large population has the advantage of less chances for random drift and constrained inbreeding. In well managed commercial breeding populations the population size is determined before the breeding programs starts and is maintained in all generations later on. But in less controlled breeds, e.g. in horses or dogs, it depends of several factors how population size develops. In dogs the popularity of the breed might change; in horses low prices for foals may lead to less matings and in the long term to a decrease in population size.

The *number of offspring born per year* is the second parameter to monitor population size. Over years it presents the stability of a breed: is there an increase (favourable) or a decrease (unfavourable)? But in animal breeding we know that not all animals born will be used for breeding the next generation. They are not selected for breeding or the owner does not want to breed with the animal. The latter is often seen in species not kept for commercial purposes: dogs and horses.

Chapter 14.6.3: The ideal population

In the ideal population inbreeding is at a very low level, mutations with adverse effects do not pop up and the random loss of alleles with a low frequency does not happen. Therefore the ideal population is large and many males and females are used for breeding. This decreases the forced mating of related animals, carriers of rare alleles produce indeed progeny and carriers of mutation with adverse effects have not be used to produce progeny.

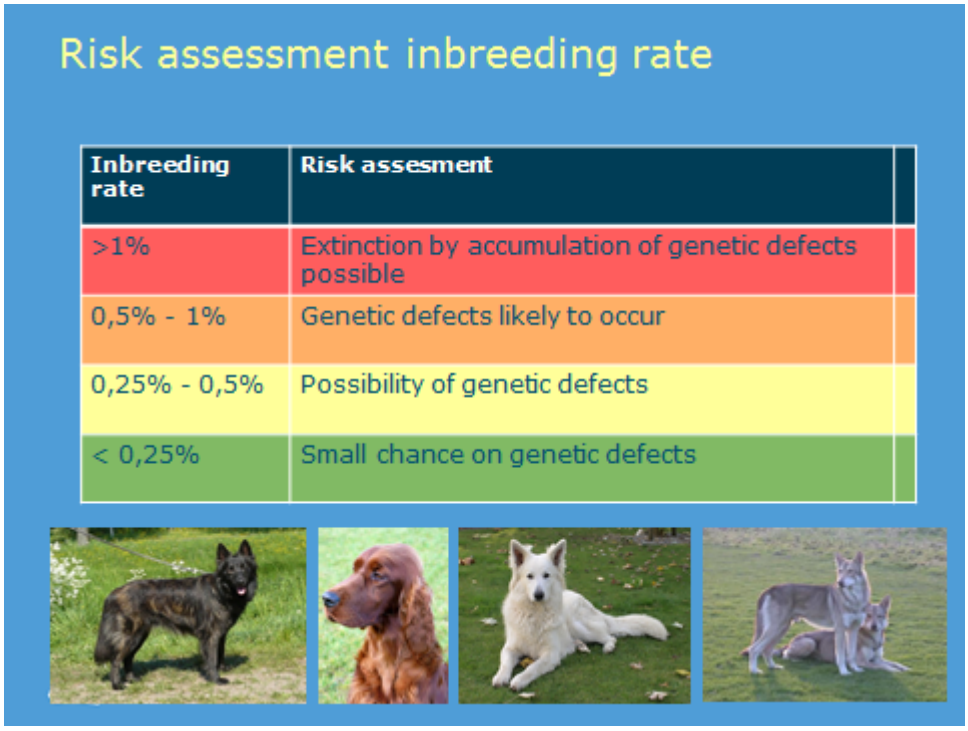
The ideal population is large: scientific literature pleads for the use of more than 100 animals as parents for the next generation. This facilitates natural selection to cull mutations with adverse effects, it avoids the random loss of rare alleles and it maintains a large genetic variation in the population.

In addition to the size of the population, the structure of the population contributes to the ideal population. The structure depends among others of the contribution of the parents to the number of offspring. When this contribution is proportional (evenly spread) then the average genetic relationship does not increase more than necessary. When a few sires dominate as parent of offspring then the relationship among animals increases sharply in the next generations and consequently inbreeding will increase sharply. See also the explanation of genetic contributions in the chapter “mating”.

The number sires and dams and the variation in the number of their offspring determine the genetic composition of the next generation of their offspring. The variation in the number of their offspring is of great importance. In well controlled breeding programs it is tried to keep this variance as small as possible: it is tried to get an equal number of progeny selected individual. But in a lot of species dams give birth to multiple offspring and then variation in litter size is always present. In less controlled breeding programs a lot of variation is often found in the number of offspring per sire. Popularity of sires is responsible for that: show champions are widely and often unlimitedly used as sires. The number of animals that contribute in different extend to the next generation is crucial in the management of a population when you aim at reducing the rate of inbreeding.

Chapter 14.6.4: Monitoring the rate of inbreeding

To avoid problems with inbreeding, the occurrence of recessive genetic defects and inbreeding depression, the rate of inbreeding F should be less than 0,5 % per generation. Internationally it is agreed that 0,5 % inbreeding is the maximum acceptable value. The higher the estimated inbreeding rate per generation, the higher the chance for inbreeding problems as is illustrated below:



The inbreeding rate F can simply be calculated per year by calculating the inbreeding coefficient of all animals born in subsequent years and to calculate the difference in inbreeding coefficients between two subsequent years.

Inbreeding rate F per generation can be calculated first by calculating the inbreeding coefficient of all animals born in a year. Second the generation interval should be calculated, being the period of time taken to renew the population of parents.

Definition

The generation interval for male and female parents is the average age of the parent when its replacement is born. The generation interval for the population is then the average of these two values since males and females each contribute half the genes to renewing the population

The inbreeding coefficients are averaged over generation 1 and the subsequent generation 2. The difference is divided by the average generation interval:

$F = (F_2 - F_1) / GI,$

where F_2 and F_1 are the average inbreeding coefficients and GI is the generation interval.

A short generation interval accelerates not only the genetic improvement, but also the inbreeding rate per year. In small populations where genetic improvement has not the highest priority, a long generation interval is recommended. Then, e.g. more time is available to monitor the results of mating schemes and to realize still matings that were intended, but did up to now not result in offspring.

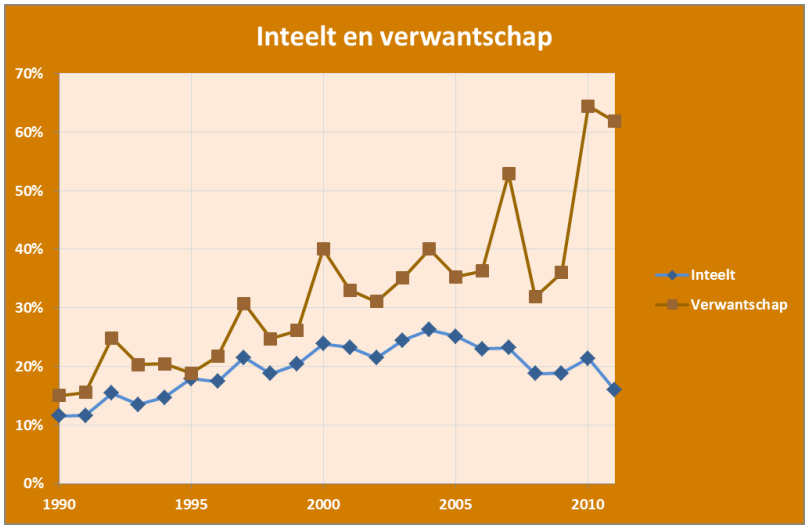
For a reliable calculation of the inbreeding rate it is important that pedigrees are complete over 5 generations of ancestors. Incompleteness leads to an underestimation of inbreeding coefficients and inbreeding rates.

Chapter 14.6.5: Relationship between the additive genetic relationship and inbreeding

In the same way as for the inbreeding rate, the rate in additive genetic relationship can be calculated. On the population level the inbreeding coefficients are also equal to half the additive relationships of the parents. When in a population random mating takes place the average additive relationship is twice the average inbreeding coefficient. But when voluntary inbreeding is practised, the average inbreeding coefficient is higher than half the average additive genetic relationship.

When inbreeding is avoided in the population, the average inbreeding coefficient is lower than the average additive genetic relationship. The latter is often seen when inbreeding is marked as an issue in a population. The logical reaction of breeders is to mate less related parents. Inbreeding coefficients decrease, but in the population the average additive relationship does not change and might even increase. And after a few generations the possibility to mate less related parents does no longer exists. The best mating strategy will be outlined later in this chapter. This phenomenon is illustrated below in a dog breed:

Figure 1: inbreeding and relationship in a dog breed, where from 2005 onwards the average inbreeding coefficient decreases by mating less related individuals but where the average relationship still increases due to the popularity of a few sires**.**



Chapter 14.7: Prevention of inbreeding

In the long term the average additive relationship in the population determines the constrained inbreeding. Therefore prevention of (forced) inbreeding highly depends of the methods to manage relationships among animals in the population. In commercial breeding populations a lot of effort is applied to manage this, although random effects (e.g. a selected animal produces no offspring) might disturb the intentions of the mating program. In less controlled breeding programs it is very difficult to realize management programs to control relatedness. Then it is wise to give breeders inside in the implications of mating relatives and to stimulate individual breeders to apply matings that all together have a favourable effect on the average additive relationship. A mating advice (which sires can be mated to this individual dam to minimise the increase in relationship in the population) might be of great help.

What measures will help to cease down in a population the increase of the average additive relationship and have therefore a favourable effect on the rate of inbreeding? Three measures might be effective:

1. Expansion of the size of the effective population
2. Restrictions in the number of offspring per parent
3. Mating schemes to control and manage relationships

Chapter 14.7.1: Expansion of the size of the population

The first expansion measure is to decrease the selection intensity in sires and dams: more sires and dams are selected as parents for the next generation. It facilitates the inclusion of sires and dams from the full population guaranteeing the presence of the total variation in pedigrees. A high selection intensity works in the opposite direction: it may easily lead to a limited number of selected parents not representing the total variation in pedigrees. Especially

a limited number of selected of sires, as is often the case in less controlled breeding programs (horses and dogs), increases the future relatedness in the population and leads to future (constrained) matings between offspring of such sires resulting in inbreeding. Thus the use of more sires and dams has a favourable effect of the average additive genetic relationship, but it should be realized that it gives less genetic improvement.

The second expansion measure is to import in a native population animals from the same breed living in other countries. In foreign populations of the same breed you may find animals with a pedigree containing ancestors not or less present in the pedigrees of the native animals. Thanks to the development of reproduction techniques it is possible to import semen or embryos of foreign animals and to produce offspring in the native population. Thanks to these foreign pedigrees the average additive genetic relationship in the native population decreases. This method can sometimes be used in dogs and horses, but when deep pedigrees are available, often the same early ancestors pop up. The breed is composed out of a limited number of ancestors and their offspring is spread over countries.

The third expansion is to cross a limited number of selected parents with selected parents from another breed. In most situations it is a handsome to buy semen from a limited number of selected sires of another breed and to inseminate a selected number of dams of your breed with this semen. The choice of the “foreign” breed is crucial: when the difference in conformation and size, in adaptive traits and in breeding goal traits is large, it will take many generations to obtain acceptable offspring and it will be difficult to get support from individual breeders for crossbreeding. In a lot of species and breeds “breed purity” is a real issue to account for and the breeding standard should not be threatened.

The second and third method of expansion maybe hampered by large genetic differences between the populations in the breeding goal traits (level and combination of traits).

The appropriate crossbreeding scheme to be applied has the structure of the introgression scheme (see chapter on crossbreeding):

A * B

F_1 (AB) * A

F_2 * A etc

Where breed A is the original pure breed and B is the selected “foreign” breed. Animals from breed B are only used to produce the F_1. In the F_1 and F_2 the animals are as much as possible selected for the breeding goal traits at stake in the original pure breed A. It might be wise that the breeding organization keeps full control over the use of the F_1 and F_2 animals. When the traits of these animals appear to be far outside the breeding goal of breed A, than it should always be possible to terminate the introgression of their genes in breed A.

This crossbreeding method is only recommended when the rate of inbreeding in the population is that high that genetic defects are a real threat for a “melt down” of the population. In a few dog breeds the introgression is considered and in the past in the Dutch Gelderlander horse breed introgression has taken place a few times.

Why are all these three measures increasing the effective population size? All three result in more variation in ancestors in the pedigree of the selected animals and therefore to less constrained inbreeding in their progeny. Introgression of animals from another breed is very effective in this respect. Then, sires and dams are mated with no common ancestors in their pedigrees resulting in a sharp decrease in the average additive relationship and in an inbreeding coefficient of zero in the F_1 animals.

Chapter 14.7.2: Restrictions in the use of parents

In well set up and controlled breeding schemes it is aimed to use the selected sires and dams with the same intensity. In the next generation they get an equal number of offspring. In that way the genetic variation in the population is maintained. All ancestors in the pedigrees of the parents return all in the pedigrees of the animals in the next generation. Such breeding schemes are optimal and sustainable: in future generations all opportunities for selection are still present. In less controlled populations we have a lot of examples (in breeds of dairy cattle, horses and dogs) of sires that were heavily used in the past. Such excessive use of a few breeding animals has a marked increasing effect of the average additive relationship in the population causing inbreeding problems in future generations. In populations they create genetic bottlenecks in populations.

The excessive use of popular sires often lead to a limited use of other sires or even neglecting sires that were selected for use. That contributes to the effects of genetic bottlenecks.

The first reaction to avoid the excessive use of a few selected animals is to restrict their use: an example is the maximum number of matings that may be performed by a sire. In less controlled breeding programs this often leads to obsession of individual breeders and owners of sires. As a rule of thumb a sire should not produce more than 5 % of the total number of offspring in the next generation.

A positive reaction to the excessive use is to make and propagate a plan in which all selected sires get an equal number of matings. They approach the optimal and sustainable schemes of well controlled breeding programs.

Chapter 14.7.3: Mating schemes to control and manage relationship

In controlled breeding programs mating programs are in place that will be used over many generations. They are characterized by two principles: 1) each sire and each dam will deliver offspring from which at least one individual (male or female) will be selected as parent in the next generation and 2) circular mating will be applied which implies that, when 25 sires are

used, it will take 25 generations before his offspring will be mated mutually and inbreeding on this sire starts. Such systems are applied in commercial pig and poultry breeding to maintain and to develop the purebred selection lines.

In less controlled breeding programs, e.g. in heath sheep, *circular mating* is applied among the participating flocks: ram circles. In these flocks of heath sheep a large number of females are kept together with a number of rams. This implies that from individual heath sheep the sire is unknown. Below you find the illustration of the circle:

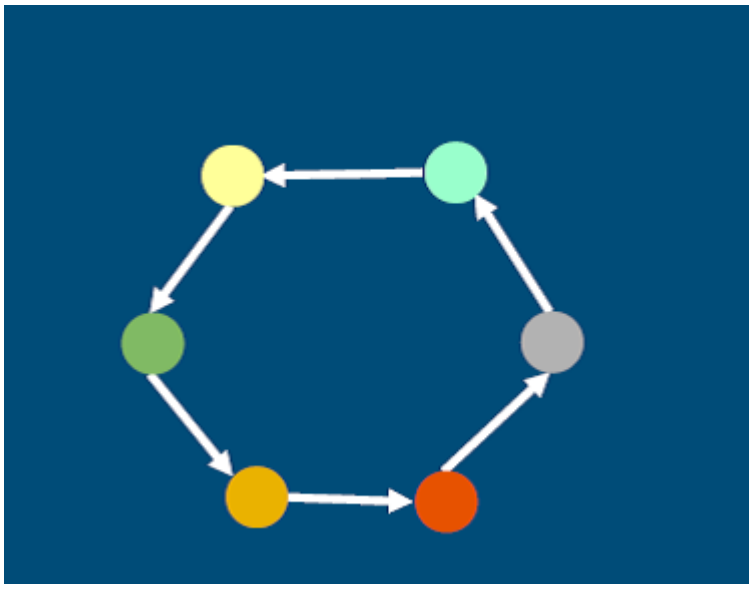


Figure 2: Example of a ram circle. Each coloured spot is a flock. The red flock always (annually) gets young rams from the dark yellow flock and delivers annually young rams to the grey flock, etc.

In this example in figure 2 6 different flocks participate in the circle. That implies that it takes 6 generations before a ram with 1/6 of the genes from the red flock is used again via rams born in the dark yellow flock in the Red flock and inbreeding will take place for the first time. When more flocks participate it takes more generations before inbreeding starts and the level of inbreeding decreases. It is a very effective scheme to keep inbreeding at a low level and to create a low inbreeding rate. When considering implementing a scheme genetic differences between the flocks and the fixed order of exchanging rams should be thoroughly studied and discussed. A breeder has to accept that he always gets rams from the same flock.

For selection and mating in the nucleus of dairy cattle breeding program the method of *optimal contributions* is developed. It outweighs the breeding value of the sires and dams with the relatedness of the sires and dams with the average of the nucleus. For each individual the programs gives the number of matings that should be performed. The mating partners are paired according to their mutual relationship aiming at a low relationship between the animals of each pair. Outside the nucleus sires and dams are combined in a sire advice program. This program aims at *compensatory matings*: weak traits in the breeding value of the dam are compensated by strong traits in the breeding value.

For individual matings a practical directive is not to mate sires and dams that share common ancestors in three generations. This means that the additive relationship between sires and dams is always lower than 12,5 % and the offspring will have an inbreeding coefficient lower than 6.25 %.

Note: the value of a gene bank in mating programs

For small populations gene banks offer the opportunity to use sires again when the initial use did not (accidently) result in offspring that can be used for further breeding. E.g. in Sweden since the introduction of artificial insemination an amount of semen from each bull entering AI is stored in the gene bank. There, AI-studs have the opportunity to fall back on a sire when it becomes relevant.

Chapter 14.8: Key issues on genetic diversity

1. Genetic diversity is the set of differences between species, breeds within species, and individuals within breeds expressed as a consequence of differences in their DNA.
2. In farm animals, the variation between breeds in performance is important in the start of a breeding program or farm activities. What will be the best breed given the production circumstances and which breed fits the best to our defined breeding goal? The concentration worldwide on a limited number of breeds leads to an increasing number of breeds being considered unprofitable and consequently leads to risk of extinction, reducing the variation between breeds.
3. The variation between breeds is the result of random drift, migration, selection and mutation. Standardized breeds were created out of Landrace breeds and selection lines out of standardized breeds, based on crossbreeding and subsequent selection.
4. The *in vivo* conservation of a breed requires a well-designed breeding program that accounts for the small number of animals and that is strictly followed by the breeders and evaluated frequently. The main objective for *in vivo* conservation is to facilitate the use of the breeds in the rural area for 1) nature management, 2) production of regional products with a high additive value and 3) the maintenance of cultural historic activities.
5. *In vivo* and *in vitro* conservation is complementary: when you do both, you guarantee flexibility of breeding programs by safeguarding rare breeds and have the possibility to use them for present activities. Gene banks can have an important role in the support of small rare breeds and can help to make these populations viable again.

6. A major source of the reliable information on the genetic variation in a breed is the pedigree, i.e. a record of sire and dam for each individual, accumulated over generations. There is a need to identify these relationships through recording the pedigrees of animals, at least in sufficient depth to identify sires and dams that are responsible for these relationships among individual animals. The deeper the pedigree, the better the true relatedness can be calculated.
7. Informative DNA markers can therefore help the measurement of genetic diversity in two ways. The first way is to overcome the problem that in some species it may be impossible or very costly to observe pedigree directly. By genotyping a small number of polymorph markers on all offspring and all possible parents, it is possible to identify the sires and dams of almost all the offspring. The second involves the extensive genotyping across all chromosomes of the genome in order to estimate the actual proportion of DNA shared by sibs or other relatives more precisely than simply using the expectation for sharing DNA between relatives that is provided by the pedigree.
8. Breeding populations should be monitored for population size, the rate of inbreeding and the generation interval
9. The rate of inbreeding can be decreased by expansion of the size of the effective population, restrictions in the number of offspring per parent and mating schemes to control and manage relationships. An outcross is very effective in this respect.
10. For small populations gene banks offer the opportunity to use sires again when the initial use did not (accidently) result in offspring that can be used for further breeding.