

GENETICS OF SELECTION IN HONEYBEES

Pim Brascamp¹, Aleksandar Uzunov², Piter Bijma¹ and Manuel Du³

¹Animal Breeding and Genomics, Wageningen University & Research, Wageningen, The Netherlands. Evert W. Brascamp <http://orcid.org/0000-0003-2284-0513>, Piter Bijma <http://orcid.org/0000-0002-9005-9131>; ²Ss. Cyril and Methodius University in Skopje, Faculty of Agricultural Sciences and Food, Skopje, Macedonia. <http://orcid.org/0000-0003-1240-868X>; ³Institute for Bee Research Hohen Neuendorf, Hohen Neuendorf, Germany. <http://orcid.org/0000-0003-2264-1610>.



This is non peer reviewed booklet. To cite please use:

Brascamp, E.W., Uzunov, A., Bijma, P. Du, M. (2024) Genetics of selection in honeybees, Wageningen University & Research, The Netherlands. <https://edepot.wur.nl/675067>

Publisher

Copyright © 2024 Wageningen University & Research, The Netherlands

ISBN

This booklet, the texts and figures may be used freely for educational purposes. When you are using parts of this booklet for publication, please read and respect the disclaimer of Wageningen UR (<https://www.wur.nl/en/Disclaimer.htm>).

Edition

First edition October 2024

Preface

About 10 years ago, Aleksandar Uzunov and I began to think about a paper that would introduce the principles of selective breeding in honeybees for a readership of scientists and practical breeders interested in honeybee breeding. This initially resulted in a paper titled “The basic concept of honeybee breeding programs”, published in 2017 in *Bee World*, together with Ralph Büchler. Aleksandar experienced much interest in the paper, and the idea developed to write additional articles discussing the factors addressed in the first paper in greater detail. This resulted in three further papers, also published in *Bee World*, on the initiation and implementation of breeding programs, the relevance of mating control and breeding value estimation. Additional authors were taken on board, such that Manuel Du and Piter Bijma contributed their expertise.

Because of the interest in the papers (for example, 83 citations and over 17,000 reads as reported by ResearchGate by 18 September 2024, for the four papers combined), I suggested to Aleksandar Uzunov that publication of a book on breeding programs in honeybees would be a valuable contribution to the body of knowledge on the subject keeping in mind a readership of non-geneticists interested or involved in honeybee breeding.

The book should contain two approaches to breeding programs. On the one hand, a thorough discussion of implementation aspects and, on the other hand, an introduction to the quantitative genetic basis underlying the design of these programs.

The current chapters focuses on the second aspect and should be considered as work in progress in two ways. Firstly, the intention is to extend the book with aspects of practical implementation, and secondly, we intend to take comments from readers into account to improve the accessibility and usefulness of the text.

Pim Brascamp,
Wageningen, 15 October 2024

Papers in *Bee World*

Uzunov, A., Brascamp, E. W., Büchler, R. (2017). The basic concept of honey bee breeding programs. *Bee World*, 94, 84–87. DOI: 10.1080/0005772X.2017.1345427.

Uzunov, A., Brascamp, E. W., Du, M., Büchler, R. (2022a). Initiation and implementation of honey bee breeding programs. *Bee World*, 99, 50–55. DOI: 10.1080/0005772X.2022.2031545.

Uzunov, A., Brascamp, E. W., Du, M., Büchler, R. (2022b). The relevance of mating control for successful implementation of honey bee breeding programs. *Bee World*, 99(3), 94–98. DOI: 10.1080/0005772X.2022.2088166.

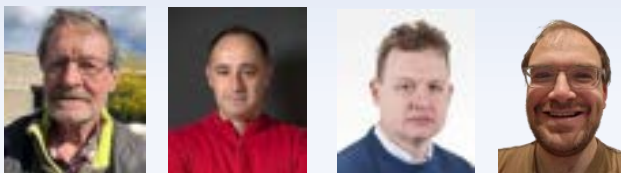
Uzunov, A., Brascamp, E.W., Du, M., Bijma, P., Büchler, R. (2023). Breeding Values in Honey Bees. *Bee World*, 100:1, 9-14, DOI: 10.1080/0005772X.2023.2166737.

Remarks and questions

Please contact pim.brascamp@wur.nl

GENETICS OF SELECTION IN HONEYBEES

Pim Brascamp, Aleksandar Uzunov, Piter Bijma, Manuel Du



Универзитет „Св. Кирил и Методиј“ во Скопје
 Ss. Cyril and Methodius University in Skopje
 Faculty of Agricultural Science and Food

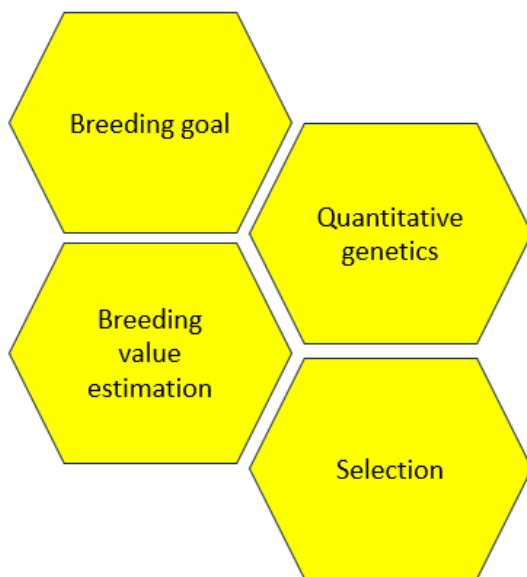
Content

Chapter 1	
Introduction	4
Reference	5
Chapter 2	
Breeding goals and traits to select for	6
Introduction	6
Breeding goal and selection criteria	6
Relevance of traits in the breeding goal	8
Breeding goals with relative weight of traits	9
Theory of breeding goal and selection criterion	12
Breeding goals for different target groups	13
Some final remarks	14
References	15
Chapter 3	
Quantitative genetics	16
DNA-basis of genetics	16
The role of chance	19
Heritability	20
Repeatability	23
Genetic correlation and correlated response to selection	23
Genotype by environment interaction (GxE)	24
Inbreeding	25
Effect of selection: selection program	27
Annual genetic improvement	28
References	29
Chapter 4	
Breeding value estimation	30
Introduction	30
Estimation of breeding values in practice	33
The honeybee pedigree	33
Breeding values for different purposes: to raise virgin queens	
from a colony or to use drones from a colony	34
Practice of breeding value estimation in honeybees	35
References	36
Chapter 5	
Selection	37
Introduction	37
Possibilities to increase the breeding values of selected parents	37
Possibilities to decrease the generation interval	38
Balancing changes in the average breeding value of selected	
parents and the generation interval	40
Balancing annual genetic gain and inbreeding	40
Design of breeding programs in scientific literature	42
Final remarks	43
References	43

Chapter 1

Introduction

This booklet contains four chapters on the theory underlying the design of breeding programs in honeybees. The text intends to be useful for scientists and beekeepers engaged in honeybee breeding without extensive training in genetics. Reference may be made to the textbook "Animal Breeding and Genetics" (Oldenbroek and Calus, 2024) which deals extensively with topics similar to ours in the context of farm animals.



After this introduction, Chapter 2 introduces the *breeding goal*. The breeding goal determines the intended direction of genetic change that is hopefully achieved in the breeding program and, therefore, deserves serious attention. In farm animals, there is a considerable amount of theory and application on the subject, but in honeybees, examples are still limited.

The theory underlying the design of breeding programs - Chapter 3 - is called *quantitative genetics*.

The adjective "quantitative" reflects the intention to quantify gene effects on traits and also the concept that many genes influence most traits

that are subject to selection. Although there are many similarities between the theories of quantitative genetics for farm animals and honeybees, there are several aspects of honeybee biology that request adaptation. The first is that traits generally are not observed on individual worker bees, such as individual dairy cows, but on colonies. Secondly, usually, the workers of a colony descend from a number of different drones and thus split into different patriline. A third aspect is that traits are usually affected not only by the genetic contributions of workers but also by the genetic contributions of their dam, the queen heading a colony. However, this phenomenon is also common in farm animals like swine, where, for example, the weaning weight depends on both the piglet's ability to gain weight and the dam's ability to provide milk.

Recently, a wealth of publications on honeybee-specific quantitative genetics appeared, in particular from the Institute for Bee Research Hohen Neuendorf, Germany and Wageningen University, the Netherlands.

Efficient selection requires a reliable estimate of the value of colonies (and queens) for selection. The most basic estimates are the observed traits on a colony compared with colonies present in the same location. This is the simplest form of *breeding value estimation*, the subject of Chapter 4. More reliable estimates are possible, however, and in farm animals, the introduction of the so-called animal model for breeding value estimation strongly speeded up the response to selection since the

early 1980es. Until recently, the application of the animal model in honeybees has been limited to a few cases, but it deserves serious consideration when running a breeding program.

Chapter 5 deals with *selection*, in particular the design of a breeding program. Annual genetic change, on the one hand, is affected by the degree to which selected parents excel compared to contemporaries and factors affecting this are discussed. A second essential factor is the generation interval, the average age of parents when offspring is born. Possibilities to affect generation intervals are being addressed.

Selection of a limited number of parents inevitably leads to an increase in average inbreeding in the population under selection. There are norms for acceptable levels, and the chapter expounds on them.

Currently, selection programs in farm animals utilize estimates of breeding values, that include the contribution of many thousands of variations in the DNA profile, so-called genomic breeding values. Although there are developments in honeybees in this respect, we do not discuss these in the context of this booklet. For the time being, we do not expect meaningful application in honeybees on a large scale.

Reference

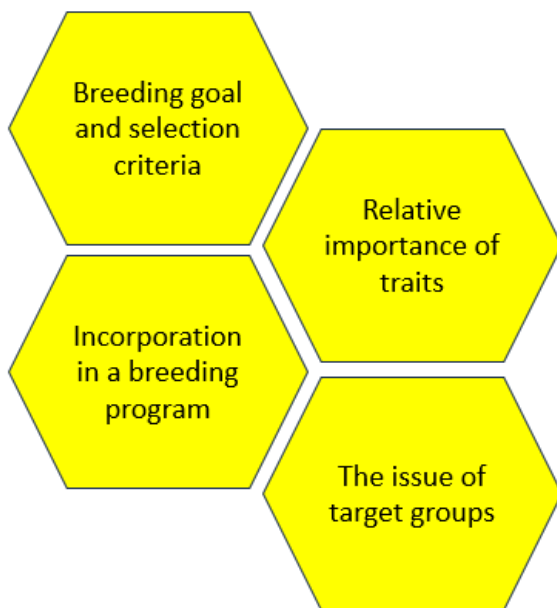
Oldenbroek, K., Calus, M. (2024) Textbook animal breeding and genetics for BSc-students. 2nd Ed. <https://edepot.wur.nl/365432>).

Chapter 2

Breeding goals and traits to select for

Introduction

A basic question in every breeding program is what traits to select for. The answer to this question will be different if an important reason to keep bees is honey yield as compared to the situation where the main reason is the production of royal jelly. Which traits are relevant may also be perceived differently by professional or hobby beekeepers, or if the goal is the conservation of a breed.



In principle, there is a difference between the traits that one likes to improve by selection (the breeding goal) and the characteristics one selects for (the selection criterion). As an example, the breeding goal may include varroa resistance as a trait to be improved by selection, while the selection criterion may comprise characteristics like mite fall, mite infestation, PINtest, suppressed mite reproduction (SMR) and grooming.

The set up of this chapter is as follows.

1. We will start this chapter discussing the difference between breeding goal and selection criterion.
2. Next, we will survey research that aims to set out the relative importance of traits in the breeding goal, by finding out which traits are more important than others.
3. Afterwards, we will discuss how the relative importances of traits can operationally be integrated in a breeding program.
4. The final paragraph of this chapter deals with the issue of target group. The question for what group of beekeepers a selection program is carried out.

Breeding goal and selection criteria

In the introduction, we considered varroa resistance as a breeding-goal trait and characteristics like PINtest as selection criterion. Often, traits in the breeding goal (like gentleness) are characteristics in the selection criterion as well. But it is important to keep in mind that in principle there is a difference. The breeding goal contains traits one likes to change by selection. The selection criterion encompasses the characteristics to be observed such that the observations on colonies can serve to select the better ones. Especially if recording of the trait in the breeding goal is difficult or expensive it will be useful to choose characteristics that are good predictors of the trait but are easily recorded. In his book "Breeding the honeybee" pp 55-65, Brother Adam gives an overview of characteristics in breeding in three categories: primary qualities of performance, secondary qualities and qualities which influence management (Table 1).

Table 1. Traits to select for listed by Brother Adam

THE PRIMARY QUALITIES FOR PERFORMANCE	SECONDARY QUALITIES	QUALITIES WHICH INFLUENCE
<ul style="list-style-type: none"> ▪ Fecundity ▪ Industry or foraging zeal ▪ Resistance to disease ▪ Disinclination to swarm 	<ul style="list-style-type: none"> ▪ Longevity ▪ Wing-power ▪ Keen sense of smell ▪ Instinct to defence ▪ Hardiness and ability to winter ▪ Spring development ▪ Thrift ▪ Instinct for self-provisioning ▪ Arrangement of the honey stores ▪ Wax production and comb building ▪ Gathering of pollen ▪ Tongue-reach 	<ul style="list-style-type: none"> ▪ Good temper ▪ Calm behaviour ▪ Disinclination to propolise ▪ Brace comb ▪ Cleanliness ▪ Honey cappings ▪ Sense of orientation

It is interesting to note that honey yield is not mentioned in Table 1. It is rather a list of characteristics that enable high honey yield per colony linked to a minimum expenditure of time and labour. Interpreted in this way, "honey yield per colony per unit of expenditure for time and labour" may be considered the single trait in the breeding goal and all characteristics in Table 1 jointly are the selection criterion.

One can criticize, however, some of Brother Adam's trait choices. Because in order to be a suitable selection criterion, characteristics need to be as closely related to the breeding goal as possible. For example, the idea behind "tongue-reach" as a selection criterion is the following: Workers with a longer proboscis have an advantage at reaching less accessible nectar sources and can therefore collect greater amounts of honey. But if such a connection really exists is likely dependent on the nectar sources in the environment and not backed by scientific studies. In contrast, it is very easy to directly measure the honey yield of a colony and select for this property. Thereby, one also exercises more direct influence on the actual breeding goal. If the proboscis length is indeed linked to honey yield, a selection for longer tongue-reaches will occur implicitly. So, in short: If your breeding goal is honey yield, you should select for colonies that produce more honey, not for colonies with longer tongues.

Interestingly, the ability to deal with the varroa mite is not among the traits in Table 1, caused by the background of the book. It describes the experiences and insights gained from the start of Brother Adam at Buckfast Abbey in the thirties until early eighties in the past century, when varroa was on the edge of being the most important threat of the honeybee. It is an excellent illustration of the fact that breeding goals have no eternal value. New traits may come up, but also some traits may lose importance. For example, a hot topic of Brother Adam's times was the "Isle of Wight disease" (hence the inclusion of "resistance to disease" among the traits of importance), which does not play a big role in honey bee breeding anymore. Likewise, if, for example, gentleness (or as Brother Adam called it: "good temper") is perceived to have reached a sufficiently high level, the trait may get less weight in selection.

Relevance of traits in the breeding goal

Most research about breeding goals in honeybees investigates the relevance of traits in the breeding goal but does not lead to relative weights given to these traits. To illustrate the difference, we look at an example: Assume that a survey among a group of breeders revealed that most beekeepers judge that a good PINtest result is more important for them than a high honey yield. Then this by itself can, of course, already be valuable information. We call such information "qualitative information". Even more informative, however, is "quantitative information", which allows to put results in perspective with each other. The pure qualitative information that the PINtest is more important than honey yield, does not tell us whether a colony with a PINtest result of 80% and 30 kg honey yield is to be preferred over a colony with 70% PINtest and 40 kg honey. Or over one with 70% PINtest and 50 kg honey? We will start in this section with practical information of the qualitative kind and turn towards quantitative information in the following section.

We start by Figure 1, which is the result of an online survey on expectations and satisfaction about marketed queens by 396 beekeepers from different European countries. The survey was part of the EurBest project.



Figure 1. Importance of and satisfaction with four traits from a survey among 396 beekeepers from different European countries in the context of EurBest (Büchler *et al.*, 2022).

The bars for the four traits represented in the left half of the figure can be interpreted as the importance that is attributed to each trait in the breeding goal. The figure shows that there is concurrence among the breeders but no strong agreement. As an example, about 65% of the respondents considered resistance (primarily to varroa) the most important trait, but a bit more than 5% of them considered it least important. Although the figure provides relevant information for defining the breeding goal, it is difficult – or even impossible – to derive relative weights for the four traits.

Comparing importance and satisfaction, it appears that both show agreement. Respondents are least satisfied with traits that are considered most important. Nevertheless, the comparison illustrates that when surveying to provide information about the breeding goal it is very important what questions are being asked and how.

Another example is a study on the breeding goal for a small population of Swiss dark honeybees with beekeepers that prefer locally selected bees with broad genetic diversity (Guichard *et al.*, 2019). Among other things the beekeepers were asked to score the relevance of eight traits (Table 2) on a scale from 1 to 5. The traits include the breed's conformity and genetic diversity, which usually are not considered as traits in the breeding goal. The table shows the results for 99 beekeepers. For all traits there were beekeepers that scored 5, but the lowest score given varied. Brood health clearly was considered the most important trait, rated with a 5 by all respondents. For the other traits the scores varied between respondents, but across respondents, honey yield was least important as the median score was 3.

Table 2. Importance of eight traits, each scored on relevance on a scale from 1 (not important) to 5 (very important) for 99 respondents to a survey. The median score is the score below and above which the number of respondents are in balance. The highest (lowest) score is the highest (lowest) one as given by at least one respondent.

TRAIT	MEDIAN	HIGHEST	LOWEST
Honey yield	3	5	2
Defensive behaviour	4	5	1
Swarming drive	4	5	2
Brood health	5	5	5
Resistance to varroa	4	5	3
Calmness during inspection	4	5	2
Racial conformity	4	5	3
Genetic diversity	4	5	3

Beekeepers were also asked their opinions about the relevance of varroa resistance compared to other traits. They were invited to imagine that resistant bees were successfully bred in Switzerland (no treatment needed) and asked at what expense that might go in terms of less results for other traits. In that way, they evaluated the trade-off between selection for varroa resistance and other traits. The beekeepers in the survey preferred resistant bees even if these would produce less honey, swarm more often or would be less gentle. But as in the previous example, the information remained mostly qualitative, because it is not clear to which extent they would sacrifice these traits in favour of disease resistance.

Breeding goals with relative weight of traits

We pick up once more our example from the previous section where beekeepers rated PINtest results to be more important than honey yield. From this purely qualitative information, it remains unclear, how much more honey yield would be needed in order to compensate for a 10% loss in PINtest. Operationally, for a breeding goal it is important to find out where such a break-even point is. Or in other words, in the example above, how much weight to give to 1% improved PINtest and how much weight to 1 kg improved honey yield. Table 3 illustrates the situation where breeders give a weight of 3 per 1% PINtest result and a weight of 1 per kg honey. In effect, a colony with a PINtest

result of 80% and 30 kg honey is judged as equally valuable as a colony with only 70% PINtest but 60 kg honey yield.

Table 3. Overall value of a colony when the relative weights for 1% PINtest and 1 kg honey are 3 and 1, respectively.

	WEIGHT PER UNIT	COLONY 1	COLONY 2	COLONY 3	COLONY 4	COLONY 5
PINtest	3	80	70	70	70	70
Honey yield	1	30	40	50	60	61
Overall value		270	250	260	270	271

Beebreed (www.beebreed.eu) is a service that estimates breeding values for colonies that are tested by breeders mainly in Germany and Austria. In Chapter 4, we discuss backgrounds of breeding values, but here we discuss their relative weights in the breeding goal. The traits in Table 4 are honey yield, defensive behaviour (gentleness), calmness during inspection, swarming drive and the varroa-index. The varroa-index is a combination of PINtest and mite infestation and it is taken to be the trait in the breeding goal that represents varroa resistance. In Beebreed all breeding values are standardized in two ways. The average of all breeding values in the database of the last five years is made to equal 100 and their standard deviation is made to equal 10. In the Beebreed system there are default weights for the five traits, leading to a “Total breeding value” (TBV). Table 4 illustrates this. The weights were not derived in a formal way – as in the example that will be discussed later – but expresses that varroa resistance is considered far more important than each of the other four traits. The Total Breeding Value is computed by multiplication of 108 for honey yield with 0.15, 120 of gentleness with 0.15, etcetera and summation of the 5 results. If you check this, you will find 113.15 for the Total Breeding Value. Doing this sum for all colonies in the database the resulting standard deviation would be less than 10. But also for the Total Breeding Value the standard deviation is standardized to 10, leading to spreading out of the results.

Table 4. An example of a colony with breeding values for honey yield, gentleness, calmness, swarming behaviour and varroa-index and their weights (in %), and the resulting Total Breeding Value.

BREEDING VALUES (AVERAGE OF LAST FIVE YEARS EQUALS 100)						
	HONEY YIELD	GENTLENESS	CALMNESS	SWARMING BEHAVIOUR	VARROA- INDEX	TOTAL BREEDING VALUE
Weight in %	15	15	15	15	40	
Breeding value	108	120	117	108	113	118

When selecting colonies (and at the same time future queens raised from those colonies) for Total Breeding Value individual bee breeders can change the weights according to their insights and needs. In practice, however, most breeders use the standard weights and for the breeding program as a whole it is desirable that the weights given by different breeders do not contradict. On top of these five traits, breeders can enter data on other traits, also leading to breeding value estimates. Of course, a breeder can use those when selecting colonies, even though these traits are not included in

the Total Breeding Value. These additional traits are robustness in winter, development in spring, colony strength, chalkbrood, chronic bee paralysis (CPV), nosemosis, SMR, recapping of inspected cells and recapping of infested cells. It illustrates that often selection is based upon formal breeding-goal traits as well as additional observations or considerations.

The trait weights in BeeBreed were given based on broad guesses or at best some back-of-the-envelope calculations. In theory, it is, however, possible to derive optimal trait weights on a more formal basis, as extensively discussed by Nielsen, Byrne and Amer (2014).

In farm animals, the weights for traits in the breeding goal are often referred to as 'economic weight'. This is because farm animals usually are kept for economic profit. Initially these weights were mostly derived trait-by-trait. For each trait the question was asked: "what are the costs to increase the level of the trait with one unit and what are the benefits". As an example, the question may be "what are the costs to increase a dairy cow's milk yield per lactation by one kg, and what are the associated costs?". In this approach one must be careful not to count a trait more than once. Part of costs associated with increased milk yield, for example, are feeding costs. But if feed intake is included as a trait in the breeding goal as well, this causes problems. The solution for this was that a farm was taken as the entity to look at and a model was built to describe the profit of such a farm by input-output equations, usually a set of equations describing the processes taking place at the farm. Together, such a set of equations should describe all inputs and outputs and associated costs and returns for the farm, resulting in the overall profit for the farm. One then can study the effect on profit of the change by one unit of each of the traits in the breeding goal. In the course of time, the way to look at farming became more comprehensive, including animal welfare, societal acceptance and sustainability, affecting breeding goals. Such considerations in the breeding goal cannot be approached by input-output equations as associated traits usually are subject to opinions that will differ between interest groups such as producers, consumers, chain stores and the civil society. A way to arrive at the weights for different traits in such a situation may be the use of questionnaires and surveys, to get insight in the relative importance of different traits for the interest groups. Such approaches, using questionnaires, probably are more relevant for breeding goals in honeybees than equations of inputs and outputs.

Theory of breeding goal and selection criterion.

In common theory the breeding goal, usually called aggregate genotype, is a list of traits (breeding values) to be improved by selection, with their weights. The selection criterion (usually called selection index) is a list of characteristics (phenotypes) to be observed with their weighing factors. These weighing factors can be derived by selection-index theory (Hazel, 1943). Here we discuss a few points illustrating the theory. Suppose that the breeding goal has two traits with equal weights and that these traits also are the characteristics in the selection criterion. If the first is highly heritable (the observed phenotypic variation is mainly caused by genetics) and the second has a moderate heritability, then in the index the first will get a higher weighing factor than the second. To put this to the extreme: suppose that one trait is very important but not heritable at all, then it will receive zero weight in the selection index. Apart from heritabilities, also the correlations between traits will affect the conversion from weights in the breeding goal to weighing factors in the selection index.

The selection index may not only contain observations on the colony for which the value for the breeding goal is to be computed, but also observations on relatives. Then in addition also the genetic relationships among the colony and its relatives play a role.

When breeding values are being estimated as discussed in Chapter 4, jointly estimated for all traits, then the weights can be directly applied to the estimated breeding values because heritabilities, correlations, and genetic relationships are accounted for.

A promising method using questionnaires was applied for example for slaughter lambs by Byrne *et al.* (2012). The basis of this approach is that stakeholders are asked to express their preferences for a set of pairwise comparisons. We illustrate this method with an example for honeybees, where 21 members of the Working Group Beebreed Netherlands answered 46 questions on the relevance of different traits. Here we discuss the method and provide some results. A paper (Brascamp, Bijma and Van der Lans, 2025) is in preparation. Figure 2 gives an example of one of these questions.

honey % as it is	no preference	honey % 20% more
PINtest 20% higher		PINtest as it is
Put a cross in one of the green boxes		

Figure 2. One of 46 questions on the relevance of different traits

In this question a choice between two cases is presented. The first (left) is the case that as a result of selection in the future honey yield at the apiary still is as it is now, while the PINtest in average is 20% higher. The second (right) is the situation that in the future honey yield is 20% higher than it presently is, but PINtest remains unaltered. Each of the 21 members answered this question alongside 45 other questions, each time balancing the relevance of two traits in a quantitative manner. As an example, suppose that one of the participants puts a cross in the left green box then the answer to another question involving a PINtest that is 10% higher instead of 20% higher still may result in a cross in the left green box; however, it may be that for this participant this is just a tipping point which makes him put a cross in the middle green box or the right one. Answers to the 46 questions given by all participants provide information about the quantitative weights to be given to each of the traits. The traits included were honey yield, gentleness and calmness, swarming, and PINtest.

Table 5 gives results for honey yield, gentleness and PINtest, as these three traits turned out to be most relevant. Probably due to the small number of 21 respondents, for calmness and swarming no credible weights could be estimated together with the other three traits.

Table 5. Weights for honey yield, gentleness and PINtest

TRAIT	UNIT IN THE QUESTIONNAIRE	WEIGHT PER UNIT	UNIT IN PRACTICE	WEIGHT PER UNIT
Honey yield	% relative to mean	0.30	kg	0.25
Gentleness	% highest score	3.00	points 1-4	2.00
PINtest	% cleaned cells	0.10	% cleaned cells	0.10

The unit used in the questionnaire for honey yield was percentage relative to the apiary mean and not kg. Average honey yields differ greatly between apiaries such that an additional kg for one beekeeper would mean far less than for another. Percentage was assumed to give a better picture of relevance. For gentleness the % highest score (=4) was used and not the actual points 1-4, considering that the former is easier to interpret.

In Table 5, the resulting weights were converted in weights for the commonly used scores kg, points and %. As an example, if honey yield would be 5 kg above the mean, gentleness 0.1 point below the mean and PINtest 5% above the mean, the total value would be $0.25 \times 5 - 2.00 \times 0.1 + 0.10 \times 5 = 1.55$. Obviously, the weights derived in this way reflect the present aggregated preferences of the 21 members of the working group, and not necessarily those of this group in the future, or of other groups. Also, the preferences of individual group members may deviate.

Breeding goals for different target groups

The realisation that the weights in Table 5 reflect the aggregated preferences of this group of 21 beekeepers emphasises the importance of a proper definition of the target group for which the weights, and therefore the breeding goal, holds. The question to be answered is: "which is the target group the selection program is working for". Breeding goals obviously will be different when the target group is interested in conservation of a breed, or in high honey yield, or in calm colonies in an urban environment or in royal jelly production, but also less extreme variation will exist between groups of beekeepers.

A special issue emerges from the fact that the performance of colonies for most traits is influenced by the genetics of the queen as well as by the genetics of the workers. As an example, for honey yield the so-called queen effect may relate to egg laying capacity and pheromonal influence while the worker effect may relate to traits like foraging zeal or wing-power. Figure 3 illustrates the relative weight in the breeding goal of queen effect and worker effect for two situations. The first is that the target group concerns beekeepers that use controlled mating, and the other that the target group concerns beekeepers that use open mating. Both target groups work with queens that derive from the breeding program for which we define the breeding goal.

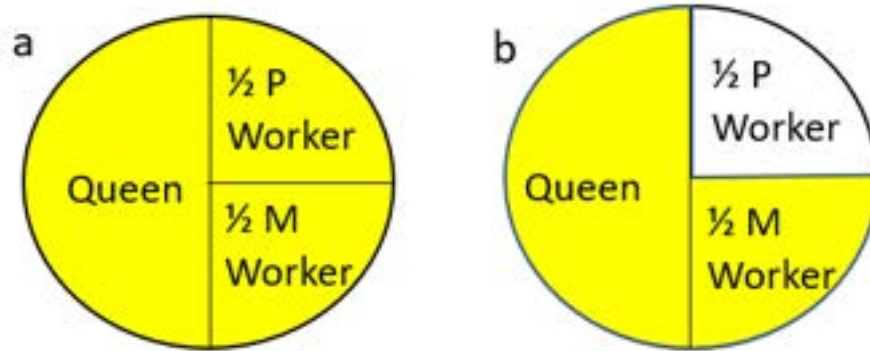


Figure 3. Contribution of a queen and of workers with their parents to the performance of a colony for controlled mating (a) and open mating (b). Yellow relates to genetics contributed by a selection program, white to genetics contributed by chance (induced by open mating).

Figure 5a illustrates a closed population of honeybees from which the target group utilises controlled mated queens. Looking at the performance of a colony, the left part of the circle illustrates that the queen effect contributes half of the total to the breeding value for any trait. The worker effect contributes the other half, be it that half of that originates from the mate of the queen in the colony (paternal contribution, P) and half from the queen herself (maternal contribution, M). Both the paternal as the maternal contribution to the worker effect originate from queens, producing either sperm cells through their drones or egg cells. In principle each queen may contribute from the maternal side to workers and from the paternal side. Therefore, the worker effect gets the same weight in the breeding goal as the queen effect. Now, consider a situation where beekeepers continuously buy virgin queens from queen breeders and that these queens are open mated. Then one may argue that these beekeepers with their open mated queens are the target group relevant for the selection program. Figure 5b represents the situation. The queen breeders do not influence the quality of the paternal part to the worker effect in the colonies and consequently the breeding values for queen effect should receive a weight that is twice that of the worker effect. It is likely that the situation in Figures 5a and 5b jointly exist in practice. One then might weigh worker effect and queen effect according to the relative numbers of colonies of both situations.

In practice it is impossible to distinguish between worker effect and queen effect on performance traits unless statistical techniques are used to estimate breeding values as discussed in Chapter 4. When breeding values are estimated in this way, the weighing of queen and worker effect deserves consideration.

Some final remarks

A breeding goal describes the traits to be changed by selection. For an operational breeding goal not only the traits to be changed should be decided but also their relative weights. In that way every colony gets an overall value and colonies can be ranked for this overall value. It should be noted that in principle these weights apply for breeding values of the traits in the breeding goal and not for the values of characteristics observed on colonies. The reason is that traits differ in degree of heritability and their mutual relationships should be considered as well. For practical purposes, however, these

weights may be applied for the values as observed as usually the information needed to convert weights of breeding values into weights for observed values is not available.

References

- Brascamp, E.W., Bijma, P., Van der Lans, I. (2025) Derivation of weights for traits in a breeding goal for honeybees. In preparation.
- Brother Adam, 1987. Breeding the honey bee. Northern Bee Book, Scout Bottom Farm, Mytholmroyd, Hebden Bridge, West Yorkshire. This book is a translation of Züchtung der Honigbiene, 1982.
- Büchler, R., Uzunov, A., Costa, C., Meixner, M., Le Conte, Y., Mondet, F., Kovacic, M., Andonov, S., Carreck, N.L., Dimitrov, L., Basso, B., Bienkowska, M., Dall'Olio, R., Hatjina, F., Wirtz, U. (2022) EurBeST – A pilot study testing varroa-resistant bees under commercial beekeeping conditions. *American Bee Journal* 162(2), 213-215. DOI 10.3896/IBRA.1.52.1.07
- Byrne, T.J., Amer, P.R., Fennessy, P.F., Hansen, P., Wickham, B.W. (2011) A preference-based approach to deriving breeding objectives: applied to sheep breeding. *Animal* 6:5, 778–788. DOI:10.1017/S1751731111002060
- Guichard, M., Neuditschko, M., Fried, P., Soland, G., Dainat, B. (2019) A future resistance breeding strategy against Varroa destructor in a small population of the dark honey bee, *Journal of Apicultural Research*, 58:5, 814-823, DOI: 10.1080/00218839.2019.1654966
- Hazel, L.N. (1943) The genetic basis for constructing selection indexes. *Genetics* 28, 476-490.
- Nielsen, H.M., Amer, P.R., Byrne, T.J. (2014) Approaches to formulating practical breeding objectives for animal production systems. *Acta Agriculturae Scand Section A*, 2014 Vol. 64, No. 1, 2-12. DOI:10.1080/09064702.2013.82723

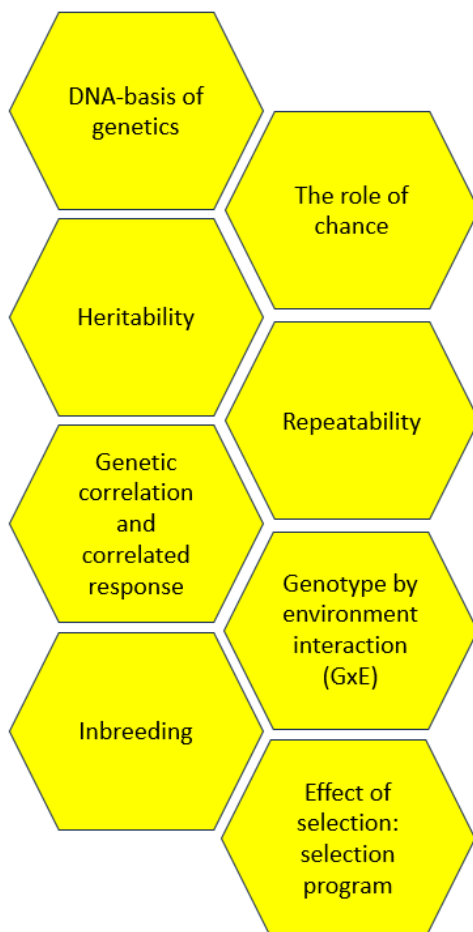
Chapter 3

Quantitative genetics

Introduction

The intention of selection is to achieve genetic improvement. Selection involves the choice of better colonies to rear a new generation of queens to be fertilized by superior drones as well. The intention is that this new generation of queens should head colonies that outperform those of the previous generation. However, it is often not easy to observe if the new generation actually outperforms the older one, and it may even be that genetically, the next generation is not better than the previous one at all. Why is that so? Why is it difficult to detect improvements between generations, and how can a generation that is reared from superior parents be genetically not better than the previous one?

The purpose of this chapter is to get a hold on these issues. To understand the challenge to distinguish genetics and environment when judging colonies' performances. And to gain insight into the role of chance in the inheritance of genetic merit. Our experience shows that breeders usually overrate the role of genetics when comparing the differences in the performances of colonies. Also, breeders tend to emphasize when progeny resembles parents and ignore when this is not the case.



We will first briefly explain the DNA-basis of genetics and then move to various aspects of quantitative genetics. What does the word quantitative refer to? In our view, it relates to the fact that we are concerned with complex traits. That is traits that build on many physiological processes, each likely affected by various genes. Consequently, complex traits are affected by many genes, and not by just one or two.

The DNA-basis of genetics and how DNA inherits from parents to offspring will clarify that chance plays a significant role. Later, we will see that offspring on average, resembles parents, but that variation between offspring is the rule. Going into various aspects of quantitative genetics allows quantification of the degree of inheritance and how it enables us to predict the degree to which the offspring of selected parents on average outperforms the generation of parents.

DNA-basis of genetics

Terms like "genetics" or "DNA", may give rise to associations with many other expressions like "genes", "alleles", "double helix", "chromosomes" and so on. With all these technical terms, it is easy to get confused. In this

section, we will explain what all these expressions mean, how they work together and what their connection to honeybee breeding is.

It turns out that the genetics of drones is simpler than that of queens and workers, which is why we start our explanation by this caste. The DNA of a drone can be envisaged as a very long string of information. Actually, this "string" is organized in the form of the famous double-helix, but for our purposes, this geometrical aspect does not play a role, and we may think of the string as flat. By "very long", we mean that the DNA of a drone consists of approximately 250 million base units. Except for small disturbances, this length of the DNA is the same for all honeybee drones. Collections of neighboring base units (typically several thousand) form functional units, the so-called genes. Genes can code for different proteins and thereby influence biochemical processes. The vast majority of genes will carry the same information for all honeybee drones. This is because they define, for example, that the drone is a honeybee as opposed to, say, a banana. However, the small fraction of genes, but still about 10,000, for which different drones can actually differ, gives rise to great differences between individuals simply because the total number of base units and genes is so high.

So, instead of seeing the DNA of a drone as a string of 250 million base units, it is often more practical to imagine it as a string of about 10,000 genes. Because a gene consists of many base units, it can take a large number of different states; the majority of genes can thus very well differ in detail for different drones. We call the different states in which a specific gene can appear its *alleles*. An even coarser look at the DNA leads to the notion of chromosomes. In fact, the drone's DNA turns out not to be a single connected string of information but to be separated into 16 disjoint segments: the chromosomes. During cell division, these 16 chromosomes coil up to an X-shaped geometry that one might have in mind when thinking of chromosomes.

The difference between female bees (*i.e.* queens and workers) and male bees (*i.e.* drones) is that in female bees, everything is doubled. Instead of one string, one should imagine the DNA as two strings. Instead of base units, one has base pairs; instead of 16 chromosomes, one has 16 pairs of chromosomes, and each gene carries two alleles. If those two alleles are identical, we say that the queen or worker is homozygous for that gene; otherwise, we call her heterozygous. In general, organisms with two DNA strains are called diploid, while organisms with only one strain, like drones, are called haploid. As a whole, honeybees are called haplodiploid.

The reason for the fundamental role of genetics in honeybee breeding is that DNA is inherited. When a queen produces an unfertilized egg that will turn into a drone, for each gene the drone will inherit one of the two alleles of the queen. So, if a queen has 10,000 genes with two alleles each, in theory there are 2^{10000} possible genetic set-ups that a drone can receive from one queen. This is an unimaginable large number! It has to be slightly reduced because the queen will be homozygous at some loci so that it does not make a difference which of the two alleles she will pass on. Furthermore, there is a tendency for neighboring genes to be linked, so they are likely to be transmitted from the same chromosome copy. Nevertheless, the possibilities of Mendelian sampling (*i.e.* passing on diverse genetics) are huge and drones that come from the same queen can be genetically very dissimilar.

When an egg is fertilized to develop into a female bee, one of the two alleles at each locus is inherited from the dam (= mother) queen just as in drones. Which of the two alleles is inherited is a random process, called Mendelian sampling, after Gregor Mendel who studied the inheritance of traits mainly in peas (but also in honeybees). The other allele is received from the father drone. Because the father drone is haploid, there is no room for Mendelian sampling on this path. Sampling takes place when the drone is formed. Two worker bees from the same colony that have the same drone father will have genetic differences only from the sampling effects of maternal inheritance. But of course, a further driver of genetic variance within a colony is that different workers may have different drone fathers.

As we have explained above, genes are functional units within the DNA, and their biochemical properties can affect several traits of the bee. It may happen that a particular gene fully determines a specific characteristic. If for example a diploid larva is homozygous at the sex-determining gene (*csd*), it starts to develop into a diploid drone rather than a worker or a queen. Under natural conditions, such diploid drones are quickly killed by the nurse bees. However, most interesting traits in honeybees, like foraging behavior or gentleness, are determined by many genes. The general idea is that each allele of a gene may have a small positive or negative influence on the trait and that the genetic quality of a bee (its so-called true breeding value) is the sum of all individual allele effects. In fact, this general idea is a simplification of reality because allele effects do not always behave additively. Instead, there may be interactions between different genes. Interaction effects between the two alleles of a gene are called dominance effects, while interactions between different genes are named epistasis effects.

Now, the general idea of breeding is that by the selection of superior queens and drones for reproduction, more and more favorable alleles will accumulate in the population. In general, this is indeed to be expected. Note, however, that this idea relies fundamentally on the concept that most allele effects behave additively. If a queen's phenotypic superiority comes from the interplay of different genes (i.e. dominance or epistasis effects), there is no guarantee that an offspring who only inherits half of its dam's genes, will also benefit from such effects. This is part of the reason why superior parents may have inferior offspring.

The role of chance

In this paragraph, we want to emphasize the role of chance. Looking at an individual queen, she receives a random half of alleles of her dam that arrived in the dam's egg. The dam's egg fuses with a sperm from a drone that contains a random half of alleles of the queen that produced the drone. The figure illustrates the vast role of chance. The flattest curve (blue) represents the distribution of all queens' genetic values (true breeding values) in a generation. By definition, the queens have an average true breeding value of 0 kg of honey. Some 32% of the queens have true breeding values above 5 kg of honey. Now consider a virgin queen that is a daughter of a dam with a true breeding value of 0 kg mated to a single drone with a zero true breeding value as well. Of course, we expect that the true breeding value of that virgin queen is 0 kg as well. There is chance variation around that zero, however. This queen developed from an egg with a random set of alleles from the dam (fused with a fixed sperm cell). The grey curve represents the likelihood of true breeding values for the virgin queen. This can also be interpreted as the distribution of true breeding values of many virgin queens, all from the same dam and drone. Unsurprisingly, the distribution is sharper, but not as sharp as you, the reader, might expect – or hope. In between is the orange curve. It represents the case where a queen descends from a drone that is a random sample from all drones produced by 10 full-sib drone-producing queens.

Looking at these curves, you may lose all confidence in inheritance and selection. You should not, but please be aware of the necessity to work with larger numbers. Otherwise, as a breeder you will often be disappointed by your lack of sustainable results.

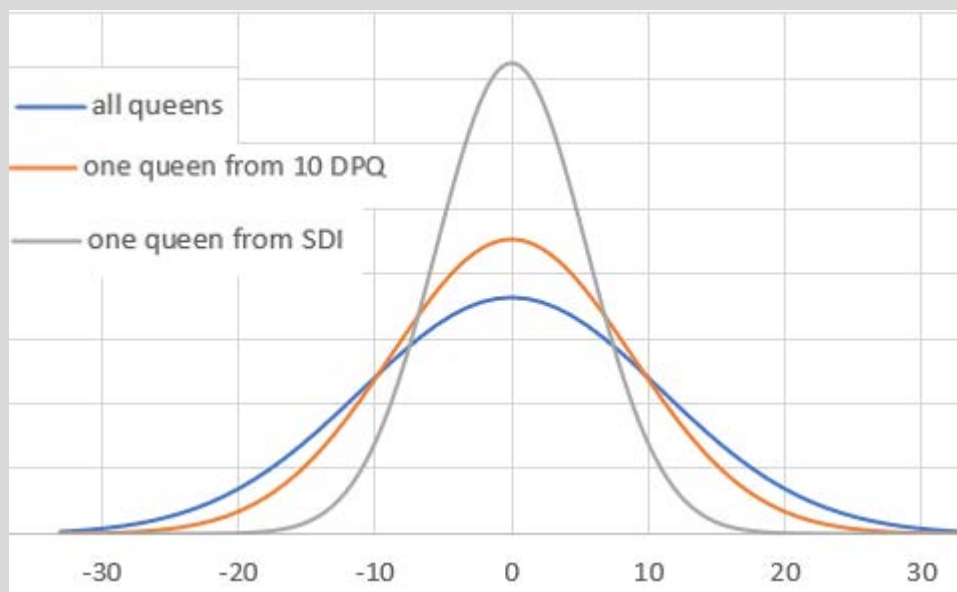


Figure 1. Distribution of true breeding values (kg honey yield) of all queens in a generation (blue line), of all possible true breeding values of a single queen raised from a colony where the dam was mated to 10 drone-producing queens (orange) and where the dam was inseminated with sperm of one drone (grey).

Heritability

Many traits in honeybees, like honey yield or mite infestation, are primarily shaped by the environment. Evidently, if a colony is placed in an area with an abundance of melliferous plants, it will generate a higher honey yield than when it is located in a barren landscape. The first step when assessing a trait is therefore always to filter out environmental effects. Even in simple selection strategies that rely on phenotypical selection, the most attractive colonies are not necessarily those that show the best phenotype but those that outperform the other colonies *of the same environment* by the greatest margin.

When colonies are placed at the same apiary and thus are confronted with the same environment, different genetics can be a main source of observed differences between colonies. But there are, of course, also other influences. For example, aggressive behaviour during inspection of a colony may be triggered by a random action of the beekeeper or the random presence of predators. Together, all these non-genetic random effects are called the residual effects of a trait. So, in total, the performance of a colony is determined by the environment, its genetics, and the residual effects. The coefficient of heritability for a trait like gentleness or mite infestation measures for colonies in the same environment, to which extent their phenotype is determined by genetics. The term "heritability" is chosen, because this coefficient measures how well a trait can be inherited from generation to generation. If a colony performs better than another colony in the same environment at a trait with high heritability, this means that the outperformance is likely due to superior genes. And as genes are inherited to the offspring, it is likely that daughters from the first colony will also perform better than daughters of the second colony. In contrast, if the heritability is low, then the better performance of the first colony was likely caused by random residual effects and there is no guarantee that the offspring will also perform well.

Typically, the heritability is a number between 0 and 1, where 0 indicates that the trait is not at all influenced by the genes, whereas a heritability of 1 means that there are no residual effects and all differences between colonies in the same environment have purely genetic causes. In reality, heritabilities of 0 and 1 barely occur, and the values for different traits fall somewhere in between. For example, Hoppe *et al.* (2020) estimated the heritability for gentleness at 0.28 and the heritability for mite infestation development at 0.05. This means gentle behaviour can be transmitted easier than mite infestation development across generations. In practice, several competing notions of "heritabilities" exist in honey bees (see text box), and some of them even allow for values greater than one.

We recapitulate: If the heritability is higher, one expects offspring more to resemble parents. Figure 2 shows the relationship between the honey yields of colonies of queens and honey yields of colonies of the queens' dams in a dataset of Dutch bee breeders. Each dot is a combination of the yield of the queen's colony and her dam's colony. In the figure, a line is drawn that best fits the linear relation between both, the so-called regression line. On average, each kilogram increase in honey yield of dam's colonies results in 0.17 kg increase in daughter's colonies' honey yield. We should emphasize the word average. As can be seen from the vertical alignments of dots in the figure, most dams have many daughters and the variation in honey yield between daughters of the same dam is considerable. Nevertheless, in average the honey yield of daughter colonies increases when the honey yield of dam's colonies increases.

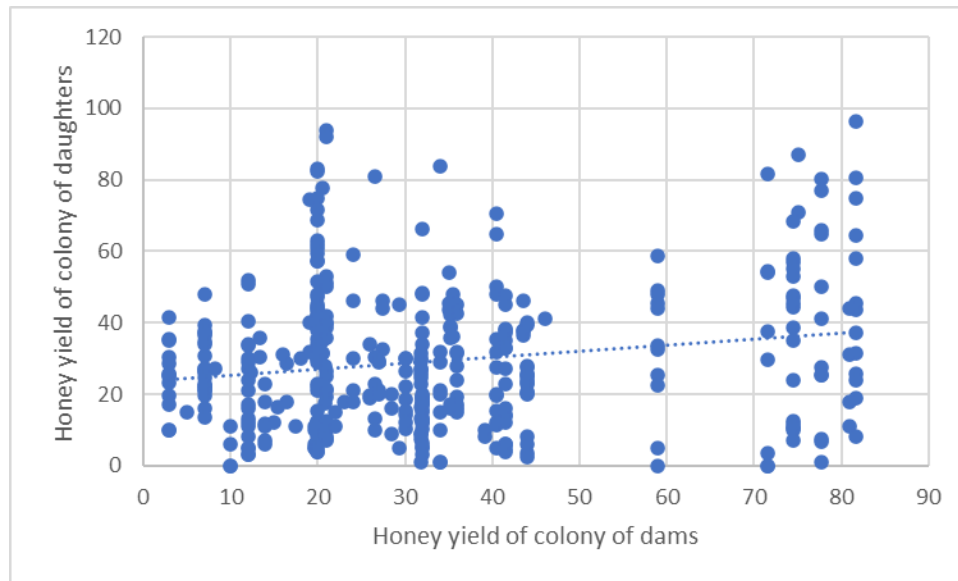


Figure 2. Honey yield of colonies of daughters (vertical axis) plotted against honey yields of dams (horizontal axis), both in kg.

An average increase of 0.17 kg honey of daughter's colony with an increase of 1 kg honey of dam's colony may seem small but consider the following: Firstly, honey yield is not only affected by genetics but also by environmental and residual effects. Secondly, the genetic merit of a daughter is affected not only by her mother but also by the drones the mother is mated to. Consequently, the theoretically expected maximum of the slope of the regression line is 0.50, in case the trait is fully heritable and all colonies are in the same environment. Assuming that all colonies faced the same environmental conditions, we can estimate the heritability from Figure 2 as twice the slope of the line, $2 * 0.17 = 0.34$. The factor 2 comes from the fact that the dam contributes half to the genetics of workers in the colony, the other half originating from the mating drones. In practice, the heritability is estimated with more advanced statistical methods that take not only the comparison dam-daughter into account but also all possible comparisons that follow from the pedigree. Current standard is the so-called animal model methodology, used for estimating breeding values, as discussed in Chapter 4.

An example of estimates of heritability in a large dataset is in Table 1.

Table 1. Heritabilities for honey yield, gentleness, calmness, swarming, PINtest and mite-population-growth (MPG) estimated by Hoppe *et al.* (2020) in the dataset of Beebreed (www.beebreed.eu).

HONEY YIELD	GENTLENESS	CALMNESS	SWARMING	PIN TEST	MPG
0.14	0.28	0.28	0.14	0.21	0.05

The estimate of the heritability for honey yield in Table 1 differs from that from Figure 2. This is not a contradiction. The Dutch dataset is a minor part of the larger dataset mainly from Germany and Austria. The estimate from Figure 2 is based upon the regression of daughter's colony performance on dam's colony performance, while the estimate in Table 1 results from the animal model analysis, considering all relationships. The table shows that the heritability estimates are low (for mite-

population growth) to moderate (honey yield). Estimates are not very high, implying that judging a colony's genetic merit is not easy and reliable selection response requires large populations.

Heritability

Heritability is a measure for the degree to which differences between individuals are heritable. If the heritability for example equals 0.34, 34% of the observed differences between individuals in the same environment are caused by genetic differences, and 66% by random other differences. In Figure 2, the slope of the regression line is 0.17. The interpretation of this can be that the heritability equals $2 \times 0.17 = 0.34$, because only half of the differences between daughters is caused by mothers, and the other half by the drones that mated to the mothers.

The degree to which differences between individuals are heritable is called the heritability coefficient, or briefly, the heritability. To define the heritability in statistical terms we need the term variance, which is a statistical representation of the common notion variation. And in equation:

$$\text{heritability} = (\text{genetic variance}) / (\text{genetic variance} + \text{residual variance})$$

The denominator, the sum of the genetic variance and the residual variance, is usually called the phenotypic variance.

Above, we took as an estimate of the heritability twice the slope of the regression line. For honeybees this interpretation is not fully correct. There are several reasons for this.

The first reason is not honeybee specific. But as we explained in the beginning of this paragraph, a first step always has to be to filter out influences of the environment. So, when we estimate the heritability directly from the regression, we implicitly assume that all colonies were confronted with the same or at least very similar environments. From the data set as it is presented here, it is not clear, to which extent this is the case, but it is to be expected that colonies are kept at various test locations with differing environments

Another reason is that traits like honey yield are not only affected by the genetics of workers but also by the genetics of the queen in the colony. For example, workers' ability to find nectar sources plays a role and the way the queen steers the colony behaviour through her pheromones as well. Genetically distinct traits, that can be attributed to workers and queens, respectively. Analyses allow to separately estimate the genetic queen effect on traits as well as the worker effect. One usually finds that both effects are negatively related. The simple regression analysis cannot untangle these effects and will thus be biased.

Furthermore, traits are measured on colonies and not on individual bees. Usually, the factor 2 used to estimate the heritability from the slope is not far off, but because of both phenomena, not perfectly correct.

To derive the heritability from the slope is only one way to estimate it. Current standard is the so-called animal model methodology, in which both the genetic effects of queens and workers are estimated. However, successful use of this methodology requires large datasets with many generations of data. Hoppe *et al.* (2020) is a good example of this. The value of the heritability depends on the method used to estimate the parameter and it is therefore important to know how reported heritabilities are estimated what exactly is meant by the term.

Repeatability

The coefficient of repeatability is a measure for the degree to which repeated observations resemble each other. If one for example counts mite infestation repeatedly, a high repeatability means that colonies that initially have a high mite infestation also have that later. When there are two observations per colony, repeatability can simply be computed as the correlation coefficient between the first and the second (the degree to which the dots are close to the regression line when plotting the second on the first measurement). If one assumes that repeated measurements represent the same trait, then the repeatability will be larger than the heritability, with the degree to which residual differences between colonies are permanent. But if all residual differences between colonies are non-permanent and differ all the time, then the expected estimate of the repeatability equals the heritability. Or in other words: The estimate of the repeatability is the upper limit of the heritability. If there are not two measurements for each colony but more, there are techniques to compute a repeatability that represents an average correlation between each pair of measurements.

Genetic correlation and correlated response to selection

The correlation coefficient measures the degree to which two characteristics are linked and thus the strength of the tendency that they will show simultaneous behaviour. For example, if one scores 100 colonies for gentleness and calmness, colonies with higher scores for gentleness generally also have a higher score for calmness. It's an example of a strong tendency. However, if one measures mite infestation of 100 colonies in summer and scores colony strength the following spring, a fairly weak tendency may be observed, perhaps a negative one. High mite infestation often coincides with low winter survival.

In general, the correlation coefficient can vary from minus 1 to plus 1. For example, a value of 0.8 between gentleness and calmness and -0.3 between mite infestation and spring colony strength may appear plausible. The correlations between gentleness and calmness and between mite infestation and spring colony strength are between observations on colonies. These partly will have genetic causes and partly environmental or residual causes. The genetic correlation is a measure for the degree of relatedness between the genetic merits for two traits. For selection, the genetic correlation between traits is much more important than the general, phenotypic, correlation. The genetic correlation between gentleness and calmness is high. Consequently, if one selects for gentleness and ignores calmness, one can expect that not only gentleness changes by selection but calmness as well: the so-called correlated response. In a selection program, selection is always for a limited number of traits and one should be aware of possible undesirable correlated responses. For example, a correlated response of selection for honey yield may be lower resistance to diseases. This is a very difficult area, however, because genetic correlations are not easy to detect and may differ for different populations. To know if honey yield and disease resistance are related, one needs to estimate the genetic correlation between these traits and therefore, observations are required. But if these observations are available, one could also directly select for resistance to diseases. It is kind of a vicious circle. Caution is justified for undesirable correlated responses, but it is very difficult to quantify these, and if one can quantify them then the problem ceases to exist as one can include the traits concerned in the selection process.

There are statistical techniques to estimate the size of genetic correlations and the separation of the observed correlation in the genetic, environment and residual parts finds its basis in the relation

between observed correlations between relatives. A warning is justified. A large dataset is required to estimate genetic correlation with acceptable precision, and this is rarely the case for honeybees. Published genetic correlations in the literature generally are with low precision or, in statistical language, with high standard errors.

Table 2 lists genetic correlations between the traits in Table 1, from the same analysis.

Table 2. Genetic correlations between honey yield, gentleness, calmness, swarming, PINtest and mite-population growth (MPG) estimated by Hoppe *et al.* (2020), in the dataset of Beebreed (www.beebreed.eu).

	GENTLENESS	CALMNESS	SWARMING	MPG
Honey yield	0.32	0.40	0.15	
Gentleness		0.89	0.31	
Calmness			0.31	
PINtest				0.48

The table contains an estimate of the genetic correlation of PINtest with mite-population growth (MPG) but not with the other traits. These were not estimated. The genetic correlation between PINtest and MPG is favourable because low MPG is scored as positive. The genetic correlation between gentleness and calmness is very high. This implies, as stipulated earlier, that selection for gentleness without paying attention to calmness would lead to a considerable correlated response to selection. It is also worth noting that all genetic correlations are favourable. For these traits there seem to be no antagonisms, as for gentleness, calmness and swarming positive scores are favourable. To get an impression of antagonism between honey yield, gentleness, calmness and swarming on the other hand and PINtest and MPG on the other hand also correlations were estimated between the phenotypic observations, combinations of genetic and environmental effects.

Genotype by environment interaction (GxE)

The interaction between genotype and environment is a well-known phenomenon when discussing local adaptation of subspecies or populations of honeybees. Subspecies of bees tend to perform better in their native environments and relatively poorly anywhere else. To be more precise, take two subspecies, say *Apis mellifera ligustica* (the Italian honeybee) and *Apis mellifera mellifera* (the dark honeybee from north of the Alps). A large European experiment (Büchler *et al.* 2014) has shown a clear tendency that in Italy the colonies of *A. m. ligustica* will outperform the dark honeybees in terms of resilience and winter survival rates, whereas in Denmark, it is the other way round. In this sense, it is not possible to generally say that one of the two subspecies is superior to the other but rather that both are ideally adapted to their respective environments. However, it can also occur that one breed is superior to another in respect to a trait in all environments. For example, Iberian honeybees (*A. m. iberiensis*) will likely exhibit greater aggressive behaviour than the Carniolian bee (*A. m. carnica*) anywhere in the world. But this does not necessarily mean that no genotype by environment interaction exists. Still, it may be that the differences in aggressiveness are enhanced in some places and attenuated in others. Also this is then a form of GxE.

Just as one can look at GxE between breeds and environments, one can also look at GxE for the genetic merit of individual colonies and environments. To quantify this is clearly more difficult than

with breeds as a colony can only perform in one environment. This problem can be solved, for example, by studying full-sib colonies in different environments. A statistical technique is to consider a trait measured in two environments as two different traits such that one trait only is measured in the first environment and the second only in the other. Then, the genetic correlation between these two traits is estimated. The general idea is that if this genetic correlation exceeds 0.8, there is no need to bother about GxE in the selection program. But if it is lower, and certainly if considerably lower, two selection programs should be considered, a separate one for each of the environments. In the scientific literature there is only one example of estimates of genetic correlations like this (Brascamp et al, 2022). The reason likely is that very large datasets are needed to get reliable estimates. In Austrian data (the breeding program administered by Biene Österreich) various genetic correlations were estimated, for example, between a region with very high honey yields in the northeast of Austria and other regions. For honey yield the genetic correlation was 0.55 and for calmness 0.64. The estimate for gentleness was very unreliable, and it was not estimable for swarming. This supports that estimation requires large datasets but nevertheless suggests that separate selection for the one region and the others might be considered. The analysis, however, also showed considerable GxE on a far smaller scale, and the authors concluded that selection for general adaptability was a better option than for local adaptation.

Inbreeding

A general notion is that inbreeding is the result the mating of close relatives. This focus on close relatives probably is because in that case the effect of inbreeding – inbreeding depression – can often easily be observed. The degree of inbreeding can be quantified by the coefficient of inbreeding and in principle this can vary between zero (no inbreeding at all) and unity (complete inbreeding, the individual is homozygous for all genes). In honeybees, diploid drones are a consequence of inbreeding. Diploid drones are workers that are homozygous for the sex gene, and they are removed from the colony early in life. There is a limited number of sex alleles, so in many colonies there will be a small fraction of diploid drones. With inbreeding, this fraction increases because it is more likely that the queen in the colony and the drones she mates with, carry the same sex allele.

There are two ways to look at inbreeding.

The first is the coefficient of inbreeding of an **individual** queen or average worker in the colony. The size of this coefficient can be *calculated* when the pedigree of both parents is known. If there are ancestors common to both parents, the coefficient of inbreeding of the progeny will be larger than zero. Let us look again at who the parents of a queen are. These are the mother (= dam) of the queen and the mother's mate, the mate being the queen that produced the drone that mated to the mother. This is because the mother produces the egg and the mate the sperm cell that, after uniting, lead to the progeny queen. The word "calculated" is important here because clearly, the calculated coefficient of inbreeding potentially becomes larger if more generations of the pedigree happen to be known. With a limited number of generations included in the calculations, the inbreeding coefficient gives the degree of inbreeding built up in these generations.

The second is the **change in average** coefficient of inbreeding in a population from one year to the next, the so-called rate of inbreeding. This rate can be calculated from the average coefficients of

inbreeding of all breeding queens born in different years. Figure 3 gives average inbreeding coefficients for workers for the years of birth of the queens in the colonies. There are three lines in the figure, all representing the same honeybee population but with different pedigree depths. The line starting in 2002 represents the situation where the eldest queens in the pedigree were raised in 2002. Similarly, for 2007 and 2012 the eldest queens were raised in 2007 and 2012, respectively. For all graphs, the trend is similar, an increase of about 0.18% per year, but if a longer pedigree is included, the average inbreeding coefficient is higher.

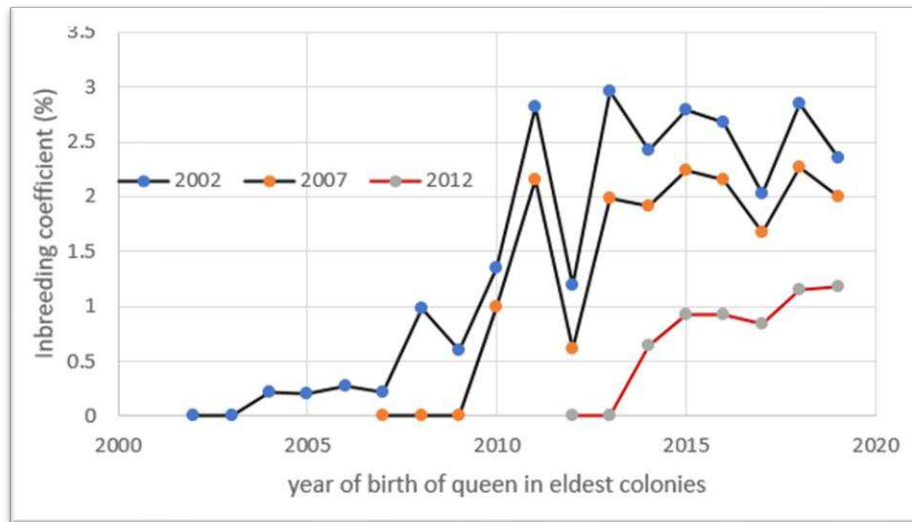


Figure 3. Inbreeding coefficients of workers computed in data from the breeding program administered by Biene Österreich (Brascamp, 2024, unpublished).

The curve for 2002 represents a pedigree that starts with queens raised in 2002 with their dam and mate. Similarly for 2007 and 2012.

For a population the *realized* rate of inbreeding thus can be calculated in retrospect, but also the *expected* rate of inbreeding can be *predicted* as a function of the number of breeding queens and the number of drone lines. The expression ‘drone line’ is a bit vague, but it intends to capture different situations. For example, the case where breeding queens are being inseminated with sperm from single queens (then that number of queens is relevant), or when queens are mating with drones from full-sib drone producing queens (then the number of dams that produce those groups of drone-producing queens is most relevant). In selection programs the number of breeding queens usually is a lot larger than the number of drone lines. Then the predicted rate of inbreeding depends mostly on the number of drone lines, which should be about 40 to keep the rate of inbreeding below the consensus 0.5 to 1% per generation.

In the bee community it is often thought that the number of patriline in a colony bears a relationship with inbreeding. That is a misconception which easiest can be seen by the realization that the inbreeding coefficient of workers depends upon the relatedness of the queen in the colony and the drones she mated with and does not depend on the number of drones she mated with or the inter-relatedness of the drones.

A consequence of inbreeding is inbreeding depression, usually expressed as the percentage with which a trait deteriorates when the coefficient of inbreeding is 1% higher. A study by Bienefeld *et al.* (1989) indicated that inbreeding leads to gentler and calmer colonies with a lower tendency to swarm but also with reduced honey yield.

Inbreeding depression often is explained by the increased proportion of homozygous genes, particularly genes that are deleterious when homozygous. One would think that this explanation is not applicable for honeybees because drones are haploid, such that the presence of one deleterious allele already is deleterious which suppresses the frequency of the allele. Instead, for a trait like colony strength, inbreeding depression may be due to many diploid drones.

Effect of selection: selection program

First, let us look at a colony with a queen that is reared from a very good colony and mated with very good drones. One justifiably can expect a very good colony with that queen. As pointed out earlier in this chapter, however, this queen is the result of the fusion of one single egg and one single sperm cell, the quality of both being subject to a lot of variation in terms of the proportion of favourable alleles they contain. The average quality of the multitude of workers then will be positively or negatively affected by the quality of the queen and the variation among the limited number of drones the queen mated with. Consequently, the expectation of a very good colony is justified but at the same time in each specific case colony performance can be disappointing or surprisingly good. Note that this statement refers to genetic variation alone. The actual performance of the colony in addition is subject to environmental and residual effects that can affect colony performance positively but negatively as well. The expected contribution of residual effects is just average, neither positive nor negative compared to the mean. Figure 2 can be used to illustrate this, as most sets of vertical dots represent the honey yield of colonies that are headed by queens reared from one dam. Although variation is the most important feature of the performance of an individual progeny of selected parents, the effect of this diminishes when we talk about a selection program, where many top colonies are selected to produce the next generation of queens.

For a selection program counts the average performance of the colonies in a particular year compared to the previous year, usually expressed as the annual genetic improvement. Though this estimate is still subject to chance, it will be more precise if the number of colonies per generation or year increases. In the remainder of this paragraph, we will consider the effect of selection in a selection program and the effect of selection for larger groups of selected parents or, in statistical terms, the expectation of the selection effect.

Annual genetic improvement

The annual genetic improvement in a population is often denoted by ΔG (the Greek letter Δ (Delta) for "difference" and G for "genetics"). In general, there are two main influences on ΔG : the genetic superiority of selected parents and the generation interval.

As we have explained earlier, a colony's performance is influenced by its environment, genotype and residual factors. So, if a colony produced 40 kg of honey, it may be that 30 kg is due to the environment, 5 kg due to favourable genes and another 5 kg due to positive residual effects. But genetic and residual effects can also take on negative values, so if another colony in the same environment only produces 25 kg of honey, the environment would, of course, still allow for 30 kg, but maybe bad genes caused the colony to produce 4 kg less, and bad luck (i.e. negative residual effects) led to another kg of honey lacking. The genetic influence on the trait is what is called the breeding value. At least at the beginning of a selection program, the average breeding value among all colonies will be 0. But the goal of selection is to choose genetically superior individuals for reproduction, so the average breeding values of selected parents are usually positive.

Now, who are those "parents". The question of who the mothers (or "dams") are is easily answered. They are simply those queens from which the next generation of queens is reared. But what about the fathers (or "sires"). Spontaneously, one might think of the drones that mate with the mothers. But this view would come with a host of practical difficulties. The main reason is that the breeder is usually not able to make any reasonable selection directly for favourable drones. During a mating flight, it is impossible to ensure that a particular drone participates in the mating process. On the other hand, it is possible to influence which *colonies* the mating drones come from. And such drone producing colonies can also be meaningfully selected on the basis of colony performances.

Therefore, it is usually the drone producing queens rather than the drones themselves that are seen as the sires in a honeybee population. In the case of instrumental insemination, the sire may be a single drone-producing queen. In the case of mating stations, the sire is a group of drone-producing queens, often full sisters. Now, let us assume that the genetic superiorities (breeding values) of selected dams and sires are quantified as \bar{A}_{dam} and \bar{A}_{sire} , respectively. Then, because each offspring inherits half of its genes from its dam and half from its sire, the average breeding value in the next generation will be $\frac{\bar{A}_{dam} + \bar{A}_{sire}}{2}$.

This value marks the genetic improvement per generation. However, we defined ΔG as the *annual* genetic improvement. So, the other defining value is the generation interval, meaning the time that passes between two generations. The generation interval corresponds to the average age of parents at the time their offspring is born. Evidently, a shorter generation interval leads to faster genetic progress and, thus a larger value for ΔG . For dams, the typical cycle in a breeding program is that they undergo a performance test in the year after they hatched. Based on their performances, they are then selected for reproduction when they are two years old. Some strategies exist to shorten the maternal generation interval to one year. On the paternal side, the generation interval depends on how mating is organized – on isolated mating stations or via instrumental insemination. Typically, it is between two and three years. Overall, the total generation interval is the average of maternal and paternal generation intervals: $\frac{\bar{L}_{dam} + \bar{L}_{sire}}{2}$.

Now, ΔG can be calculated as the quotient between genetic improvement per generation and the average generation interval:

$$\Delta G = \frac{(\bar{A}_{dam} + \bar{A}_{sire})/2}{(\bar{L}_{dam} + L_{sire})/2}$$

Of course, this fraction can be simplified to $\Delta G = \frac{\bar{A}_{dam} + \bar{A}_{sire}}{\bar{L}_{dam} + \bar{L}_{sire}}$.

As a breeding program strives to obtain large values for ΔG , one aims for large genetic superiorities of selected parents and short generation intervals. However, these are often conflicting goals. If for example the performance tests are simplified and shortened, it is possible to reduce the generation intervals. On the other hand, the selection based on such incomplete performance tests is less accurate and therefore, the genetic superiority of parents will be limited. Optimizing ΔG is consequently a highly non-trivial task and in Chapter 5 we will discuss this in further detail.

We started this chapter with the role of chance when looking at the performance of a colony. Chance plays an important role in inheritance: on average, the true breeding value of an individual equals the average of its parents. And that is what we use when we estimate breeding values. However, the true breeding value will deviate upward or downward, just by chance, whether the egg and sperm produced by its parents contain more favourable or unfavourable alleles. Furthermore, the environment in which the colony is kept influences performance. In estimating breeding values, we work with the expectation. The most likely outcome is that the offspring receives an average of the favourable alleles harboured by dam and sire, and that the environment is average. When breeding programs are small, the chance is playing a big role. When the size of breeding programs in terms of parents increases, realized results will come close to expectations.

References

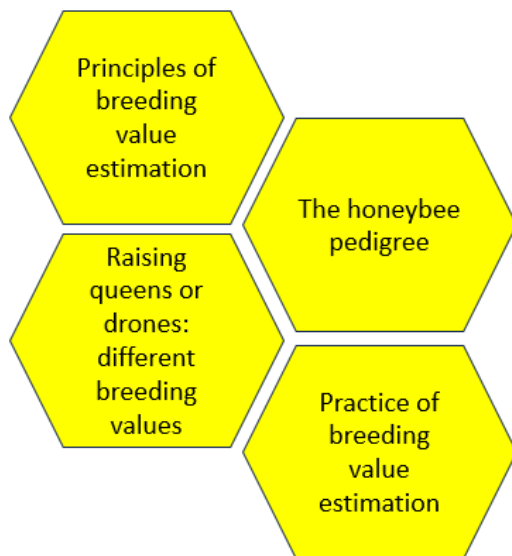
- Büchler, R., Costa, C., Hatjina, F., Andonov, S., Meixner, M.D., Le Conte, Y., Uzunov, A., Berg, S., Bienkowska, M., Bouga, M., Drazic, M., Dyrba, W., Kryger, P., Panasiuk, B., Pechhacker, H., Petrov, P., Kezić, N., Korpela, S., Wilde, J. (2014). The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe. *Journal of Apicultural Research* 53: 205-214 (2014). DOI: 10.3896/IBRA.1.53.2.03.
- Bienefeld, K., Reinhardt, F., Pirchner, F. (1989). Inbreeding effects of queen and workers on colony traits in the honey bee. *Apidologie*, 20, 439-450. DOI: 10.1051/apido:19890509.
- Brascamp, E.W., Rubinigg, M., Veerkamp, R.F., Bijma, P. (2022). Very local genotype by environment interaction in Austrian honey bees. *Proc. World Congress on Genetics Applied to Livestock Production*. Rotterdam 3-8 July 2022, 2588-2591. DOI: 10.3920/978-90-8686-940-4_626.
- Hoppe, A., Du, M., Bernstein, R., Tiesler, F.-K., Kärcher, M., Bienefeld, K. (2020) Substantial genetic progress in the international *Apis mellifera carnica* population since the implementation of genetic evaluation. *Insects* 11, 768. DOI:10.3390/insects11110768.

Chapter 4

Breeding value estimation

Introduction

When a bee breeder likes to select the best colony of the apiary, there are two challenges. One challenge is to decide about the relevance of the different traits the breeder likes to pay attention to. This issue was discussed in Chapter 2 on the breeding goal. The second challenge is to separate genetics from the residual effects. As elaborated in Chapter 3 on quantitative genetics, the observation on a colony (the phenotype) is affected by the environment, its genotype, and random residual factors. Within an apiary one can try to eliminate environmental differences between colonies as far as possible, but in practice that does not work out very well. Though serious attempts



are being made, differences between colonies caused by slight environmental disturbances remain large. Because the underlying environmental influences are often not detectable (otherwise one could eliminate them), these effects are counted as residual effects. This is for example illustrated by the heritabilities estimated for different traits as shown in Table 1 in Chapter 3. These values represent the fraction of phenotypic differences between colonies caused by genetics estimated *within* apiaries. That is, accounting for environmental differences *between* apiaries as well as possible. When a breeder likes to compare performances of colonies not only *within* a single apiary, but also compare them with performances of

colonies at other apiaries, an additional difficulty arises, in that the environments between the apiaries will differ, but also that the average genetic level may differ between apiaries. The way to deal with these difficulties is to estimate breeding values for colonies. These breeding values aim to be the best estimate of the genetic merit of the colony and the best prediction of the genetic quality of the queens raised from the colony.

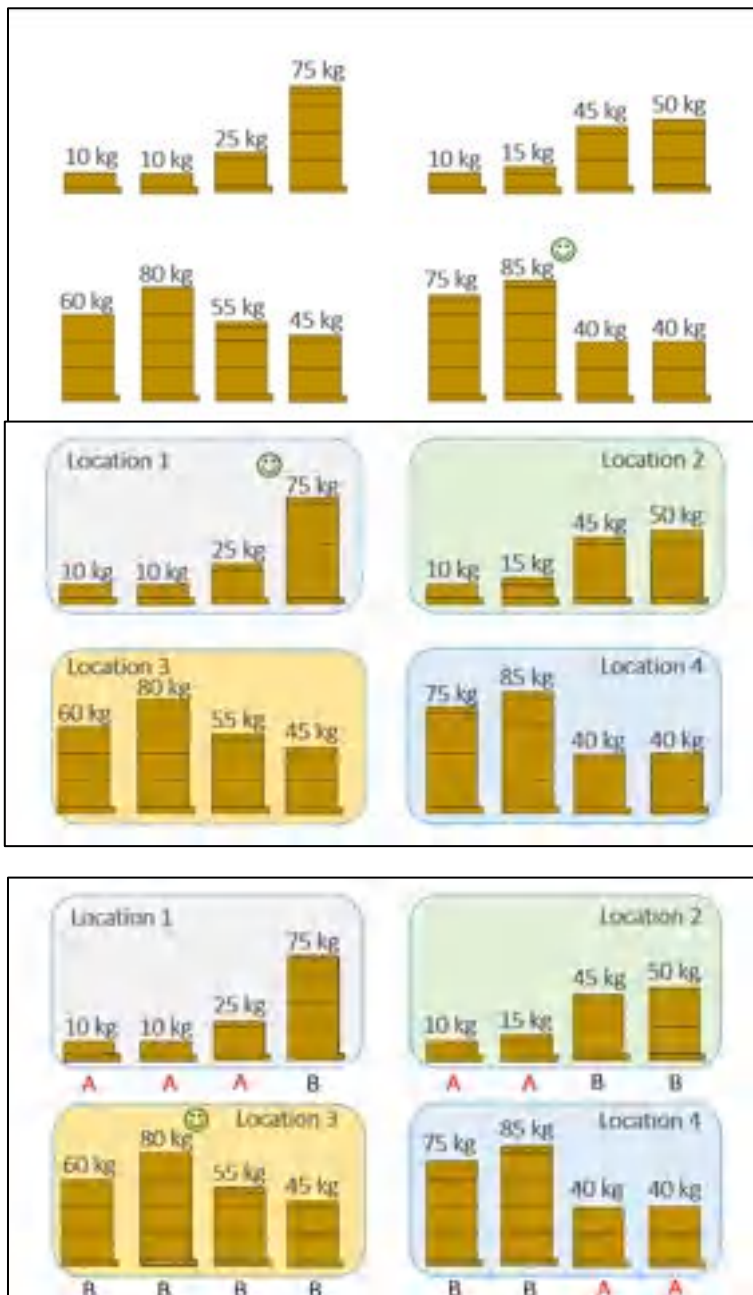
Now, if we compare queens on one apiary, we may expect that they all face the same environmental effects. So, our best guess for the non-environmental influences on a trait for a colony is the difference between the colony's performance P and the average performance \overline{Pt} of colonies in the apiary. But remember that this difference is only partly caused by genetic effects and the rest is due to residual effects. The fraction that genetics is responsible for is exactly h^2 , so that a reasonable estimate for the genetic value (breeding value) of a colony would be

$$\hat{A} = h^2(P - \overline{Pt}).$$

As a matter of fact, in many applications this is what is done for honey yield, by expressing it as a percentage of the average honey yield of the apiary. In that case a colony with 25 kg honey in an apiary averaging 20 kg of honey would have the same adjusted value as a colony with 50 kg honey in a 40-kg apiary.

An objection against this equation is that apiaries perhaps, or likely, do not only differ by environmental reasons, but also because the genetic level can differ between apiaries. This may particularly be the case if different lines of honeybees are placed on different apiaries.

We illustrate the increase in complexity to estimate breeding values as discussed above in Figure 1.



Case 1 illustrates 16 colonies with their annual honey yields ranging from 10 kg to 85 kg. The average honey yield equals 45 kg. The colony with the smiley has the highest yield. Nothing else taken into account, this colony is the preferred colony.

Now consider case 2 where the same 16 colonies are located on four locations. Locations 1 and 2 average 30 kg, locations 3 and 4 average 60 kg. When the differences in averages are taken to be caused by differing environments, the colony with the smiley is best: 45 kg above the location average.

Case 3 illustrates the situation where the colonies happen to belong to two different sister groups, A and B. Clearly sister group B outperforms A genetically, as it has higher honey yields in each of the four locations. The best colony now likely could be the one which deviates most from the family-B average, within location, indicated with a smiley.

In case 4, apart from honey yields, also scores for gentleness

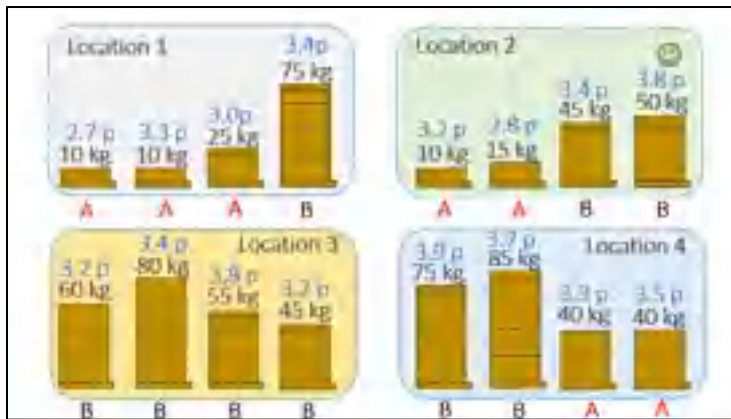


Figure 1. Four cases to illustrate increasing complexity to decide which colony is best genetically (modified after Uzunov et al., 2023).

(scored from 1 to 4 points) are known. Taking these scores at face value, the first colony at location 4 would be considered best for gentleness. Taking into account the test-location averages this is no longer the case; then the colony with the smiley is the winner, with 0.5 points above the location average. Compared with the sister-group average within location, still another colony

would be chosen: the third in location 3.

In the first case, we take it that all 16 colonies are at the same test location. If this is the only information, the colony that is expected genetically best is the colony with highest honey yield. Case 2 illustrates the situation that the colonies actually are at four different test locations. When there is no information about genetic differences between these test locations, the test-location averages best can be considered to be caused by environmental differences. Then the deviation of colony's honey yield from the test-location average provides the best impression of its genetic value. In practice, when breeding values are not being estimated as discussed below, this is the way to go. Case 3 introduces a next step of complexity where the 16 colonies actually belong to two sister groups, A and B. At each location sister group B has the higher honey yields which justifies the conclusion that genetically group B is the better sister group. To pick the best colony it looks attractive to take the colony which deviates most from the group average, within location. Here this is the 80-kg colony at location 3. Now location 3 seems to combine the highest genetic value with the best environment, but perhaps that is not the case. And certainly, if the colony with the largest deviation from the group average was in location 2, one might hesitate to take that one. To best disentangle the contribution of genetics and environment to the location averages, heritability should be taken into account, and also genetic relationships between parents and offspring, cousins, etcetera.

Case 4 concerns the usual situation that selection is not for one single trait. Now apart from honey yields also scores for gentleness are given. To choose the best colony for gentleness again is to look at the deviations of location averages, but as in case of honey yield, these may differ genetically for gentleness as well, which again changes the picture. In Chapter 2 about the breeding goal, we discussed how to combine information on honey yield and gentleness such that an overall ranking of colonies arises. In reality, there are more than two traits of interest and not only two sister groups as in our example, and genetic relationships may vary from negligible to strong. When breeding values are estimated professionally, as in the Beebreed program (www.beebreed.eu) and in farm animals, heritabilities and genetic relationships are taken into account to get the most reliable estimates of

breeding values and the environmental contributions of test locations (Bienefeld *et al.*, 2007; Brascamp and Bijma, 2014). To calculate the family relationships, pedigrees are needed.

Estimation of breeding values in practice

In practice, breeding value estimation in farm animals is carried out using specialized computer programs. The same holds for breeding value estimation in honeybees via Beebreed. For research purposes, but also for selection programs that do not use such specialized computer programs, general computer programs are available that are designed to estimate so-called variance components. Those components can be used to estimate heritabilities, and in a way estimated breeding programs are by-products. These programs cannot be used directly for the estimation of breeding values in honeybees, however. This is due to peculiarities of the honeybee pedigree. Observations are not done on individuals but on colonies that can be interpreted as groups of workers. Furthermore, the sire usually is not a single queen but a group. There is a computer program freely available (Brascamp and Bijma, 2019) to construct the proper pedigree from data that contain the identification of the queen in a colony, of her dam and of her mate and to produce the proper input for a general-purpose program. This program can handle different modes of mating control and different numbers of drones to mate with a queen. The ideal model to estimate breeding values contains both the contribution of workers' breeding value to colony performance and queen's contribution. The combined breeding value then is the sum of both. Small datasets with few generations of data usually do not allow to estimate both breeding values separately because the pedigrees of both workers and queen overlap to a large extent. If the full model does not lead to results, a reduced model can be applied, with worker effect or queen effect only. Research of Du *et al.* (2022) suggests that then the queen model is often the best choice.

The honeybee pedigree

In farm animals, pedigree entries usually fall into one of two categories: females and males. In honeybees, an additional category is needed when observations are not on queens but on worker groups. Consequently, the honeybee pedigree contains three types of entries: queens, their worker groups, and sires. To build a pedigree, for each of the three entries the dam and sire should be known.

The input to build the pedigree usually is a queen certificate. Such a certificate contains the identification of the queen, that of her dam and information about the mating. The mating can concern drones from a single drone-producing queen (DPQ) –in case of instrumental insemination – or drones of a group of full-sib DPQ. How entries in the pedigree look like is illustrated in Figure 2 for two cases. The case with single DPQ and the case with a group of DPQ. In practice both will occur in the same pedigree.

Queen certificate		
Queen	Dam	Mating
A	B	single DPQ: C
B	D	single DPQ: E

Queen certificate		
Queen(s)	Dam	Mating
A	B	group of DPQ: gC of C
B	D	group of DPQ: gE of E
C	F	group of DPQ: gG of G

Entries to build the pedigree		
Entry	Dam	Sire
queen A	B	E
worker group of A	A	C
sire A	B	E

Entries to build the pedigree		
Entry	Dam	Sire
queen A	B	gE
worker group of A	A	gC
sire gC	C	gG

Figure 2. Queen certificates for the case of one DPQ and a group of DPQ and the entries to build the pedigree derived from those. Capital letters refer to the identification of queens, and capital letters preceded by a g refer to the identification of a group of DPQ. For example, gC refers to a group of DPQ that is raised from the colony of queen C.

To derive the entries to build a pedigree, in the case of single DPQ only two queen certificates are needed. The one of a queen and the one of her dam. For the entry of queen A, the dam can be found on the queen certificate of queen A, but to know A's sire, the queen certificate of A's dam is needed. For the entry of the worker group of queen A only the queen certificate of queen A is needed, as the workers descend from the queen in the colony and her mate. As an example of a sire, we take queen A. In practice there will be only one entry for queen A, because she can serve as dam and as sire as well.

In the case of groups of DPQ we need three queen certificates because now the sire needs an additional entry. Again, for the entry of the queen we need two queen certificates (of the queen and her dam), and for the worker group only one. For the entry of the sire, we also need the birth certificate of the dam of the group of DPQ, in the table of queen C.

Breeding values for different purposes: to raise virgin queens from a colony or to use drones from a colony

Usually, breeding values are published along with the identification of a queen. This commonly is the queen in the colony. But is it really the estimated breeding value of that queen that we are interested in? When we select a queen, we do so because we hope for her queen daughters to be genetically superior. But the daughters' genetics is not only influenced by the genetics of the queen (i.e. her own breeding value) but also that of the drone mates. In fact, on average, a daughter queen is genetically not different from a random worker and only differs phenotypically because as a larva she is fed with royal jelly. This means that the breeding value published along with a queen in fact should not be her own breeding value, but that of her worker group as that is the best estimate of the breeding value of a virgin queen raised from that colony. To be explicit: this breeding value published with the identification of colony's queen is the estimated breeding value for her virgin daughters.

Now consider the case that drones are taken from the colony, instead of raising virgin queens. In that case the estimated breeding value of the worker group is irrelevant as workers contain genetics of the queen in the colony and genetics of her mate, while drones only contain genetics of the queen. In that case therefore, the relevant estimated breeding value in fact is that of the queen herself.

Estimated breeding values can be used to predict the breeding value of a planned mating as it simply is the mean of the breeding value of the dam and the mate. The result then is the breeding value of the group of workers of the planned colony, at the same time the breeding value of virgin queens one might raise from that planned colony. We already addressed the estimated breeding value of the queen in the planned mating. That is the breeding value published along with her dam relevant for her virgin daughters. Now what about the estimated breeding value of the sire? If the sire is a single queen, we need the breeding value of the queen herself, relevant for the drones she produces. If the sire, however, is a group of DPQ, the breeding value of the group equals that of the group of workers in the colony they are raised from, or, in other words, the breeding value published of the dam of the DPQs relevant for her virgin queens. Why her virgin queens? Because the mating of the DPQ is irrelevant for the genetics of the drones they produce.

Practice of breeding value estimation in honeybees

There are only a few breeding programs where estimation of breeding values is used. Usually, selection is based upon colony performance, perhaps with a slanted eye on the performance of full-sister colonies. Sometimes also, the genetic value of a queen is judged based on the performance of colonies of her daughters. We are only aware of three practical examples dealing properly with the peculiarities of honeybees. These are breeding value estimation carried in the framework of Beebreed (www.beebreed.eu), in the breeding program administered by Biene Österreich, and in the breeding program of the French Royal Jelly Association (Basso *et al.*, 2024). This is in contrast to breeding programs in farm animals where selection based on estimated breeding values is the rule. Because successful estimation of breeding values requires a substantial dataset with several generations, in practice its application may not be possible. But on the other hand, for many breeding programs breeding value estimation has a lot to offer. Not only because performances of relatives are included in each estimated breeding value, but also – perhaps *even more* so – because it allows for a proper separation of genetic and environmental differences between apiaries.

Estimated breeding values as we discussed them here are simply the part of the phenotype that is attributed to the genetics. Consequently, they have the same unit as the selection trait. This means that a breeding value for honey may be +6 kg or –10 kg, a breeding value for calmness (judged from 1 to 4) might be 0.12 scoring points, etc. It is easy to detect whether such a breeding value is positive or negative but it is difficult to decide if a positive value, like for example +4 kg honey yield is a high or a low value. To facilitate such comparisons and provide a better overview, in practice breeding values are standardized so that they become numbers with average 100 and a standard deviation of 10. In that way, also breeding values for several different traits can be compared and combined as has already previously be explained in Chapter 2 on the breeding goal. Thus, after the entire procedure of breeding value estimation, one ends up with a ranked list of queens like in Table 1, from which one can then select the best for reproduction.

Table 1. Each queen obtains two standardized estimated breeding values (EBV), one for herself and one for her workers. Because queen and workers are related, high breeding values for queens and worker groups often come together, but the detailed ranking may differ. When selecting queens for the production of new queens, one should rely on the worker breeding values, when selecting queens for drone production, the queen breeding value is relevant.

	Queen 1	Queen 2	Queen 3	Queen 4	Queen 5	...	Queen n
Queen EBV	121	120	112	111	111	...	79
Worker EBV	112	120	113	105	106	...	83

References

- Basso, B., Kistler, T., Phocas, F. (2024). Genetic parameters, trends, and inbreeding in a honeybee breeding program for royal jelly production and behavioral traits. *Apidologie* 55, 11. DOI: 10.1007/s13592-023-01055-3.
- Bienefeld, K., Erhardt, K., Reinhardt, F. (2007). Genetic evaluation in the honey bee considering queen and worker effects – A BLUP-Animal Model approach. *Apidologie* 38, 77-85. DOI: 10.1051/apido:2006050
- Brascamp, E.W., Bijma, P. (2014). Methods to estimate breeding values in honey bees. *Genetics Selection Evolution* 46, 53. DOI: 10.1186/s12711-014-0053-9.
- Brascamp, E.W., Bijma, P. (2019). Software to facilitate estimation of genetic parameters and breeding values for honey bees. *Proceedings of the 46th Apimondia Canada: Montréal*. Software and manual at <https://zenodo.org/records/7951334>.
- Du, M., Bernstein, R., Hoppe, A., Bienefeld, K. (2022). Influence of model selection and data structure on the estimation of genetic parameters in honeybee populations. *G3 Genes/Genomes/Genetics*, 12, jkab450. DOI: 10.1093/g3journal/jkab450.
- Uzunov, A., Brascamp, E.W., Du, M., Bijma, P., Büchler, R. (2023). Breeding Values in Honey Bees. *Bee World*, 100, 9-14, DOI: 10.1080/0005772X.2023.2166737.

Chapter 5

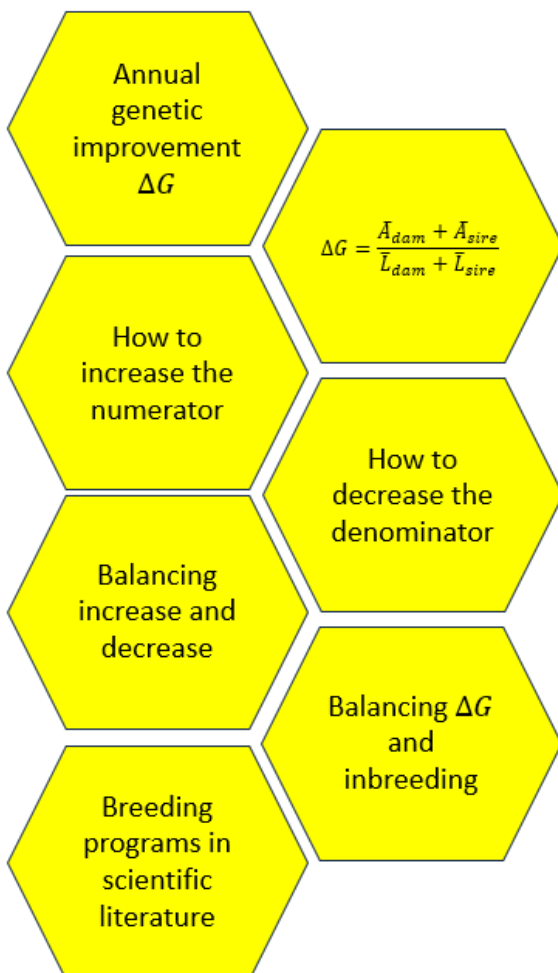
Selection

Introduction

This chapter discusses the design of breeding programs, or to be more precise: the effect of the design on annual genetic improvement. Here, annual genetic improvement stands for the difference in average breeding values of colonies between subsequent years of birth. In Chapter 3, an equation was introduced that predicts the annual genetic gain

$$\Delta G = \frac{\bar{A}_{dam} + \bar{A}_{sire}}{\bar{L}_{dam} + \bar{L}_{sire}}$$

Each year, dams and sires are selected to produce the next generation. The numerator of the equation is the sum of the average breeding value of selected dams and sires, and the denominator the average generation interval of dams and sires.



The purpose of this chapter is to explain the meaning of the numerator and the denominator, but especially the possibilities to increase ΔG by increasing the numerator and by decreasing the denominator. We first discuss these two possibilities separately. Often, however, a decrease in the denominator leads to a decrease of the numerator as well and in those cases a proper balance should be found. Attention will also be paid to the increase in inbreeding which inevitably results from selection. That is, the average inbreeding coefficient of queens and workers increases from one year to another due to the selection of a limited number of parents. This increase is displayed as ΔF , so the design of the breeding program should balance ΔG and ΔF . Finally, some examples will be given of comparisons of designs of breeding programs in the scientific literature.

Possibilities to increase the breeding values of selected parents (numerator)

The average breeding value of selected parents depends on two factors. The first is the accuracy of the estimated breeding values. The second is the selection intensity. The average breeding value of selected parents is proportional to the product of the accuracy and the selection intensity.

Let us look at both accuracy and selection intensity in more detail!

Estimated breeding values are always estimates. If the job is done properly, on average the estimated breeding values equal the true ones, but for individual colonies or queens the estimate may deviate from the true value. These deviations vary between colonies and between queens, but the less variation the larger the accuracy.

The simplest estimated breeding value for a characteristic is its phenotype, usually as a deviation from the apiary average. In that case the accuracy equals the square root of the heritability, approximately. The accuracy can be increased including also phenotypes of relatives, closer relatives in particular. As an example, the more full-sister colonies are tested the higher the accuracy of the estimated breeding value of each of the full sisters. In general, the accuracy of estimated breeding values will be increased by meticulous record-keeping, consequent mating control, and the distribution of sister groups on several testing apiaries. The intensity of selection is a function of the selected proportion. If, for example the upper 50% of colonies are selected from which to raise the next generation, the selection intensity equals 0.8, while if the top 5% are selected, it equals 2.1. Figure 1 provides the relationship between selected proportion and selection intensity.

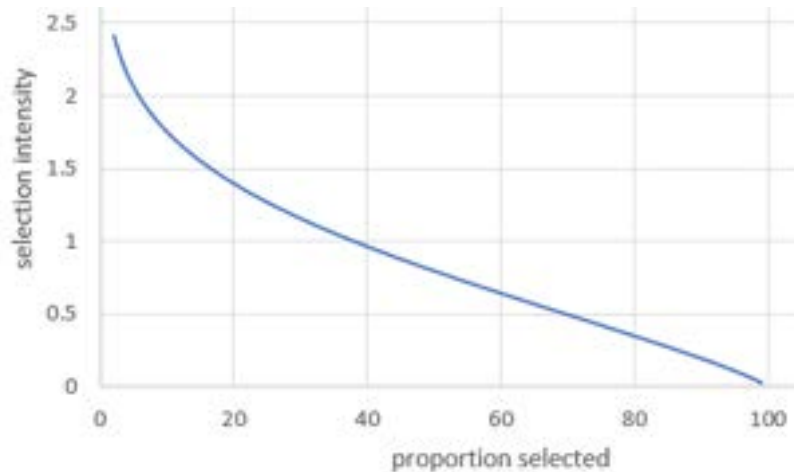


Figure 1. The relationship between selected proportion and selection intensity.

Now consider the design of a breeding program with a given number of dams and a given number of mating stations, each of these occupied with a given number of drone-producing colonies. The simplest way to increase the selection intensity then is to increase the number of tested colonies. In doing so, ΔG will increase while ΔF will roughly remain unaltered. The disadvantage of such an increase lies in the costs of test facilities. Because of these costs it will be tempting to leave the volume of testing unaltered but to decrease the number of dams and mating stations. In that case ΔG is expected to increase as before, but now ΔF will increase as well, and because of this, on the longer run decreasing the number of parents is not attractive.

Possibilities to decrease the generation interval (the denominator)

These possibilities will be discussed starting with a design of a breeding program which is implemented frequently when mating stations are used for controlled mating. Figure 2 illustrates this.

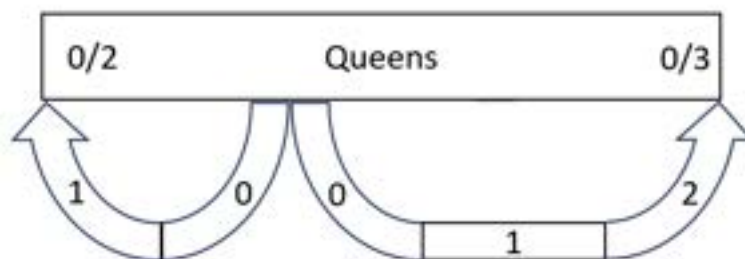


Figure 2. Representation of a breeding program where queens are selected in year 0 and tested in year 1 (maternal half circle to the left). Through the maternal path, in year 2 new queens are reared. Among the tested queens a smaller number is selected to produce drone-producing queens (DPQ) in year 2 (the paternal path to the right) and, because the DPQ are then used in the following year, through the paternal path in year 3 queens are reared.

The figure above represents selection through the maternal pathway (the left half circle) and through the paternal pathway (the path to the right). The numbers in the figure represent years. In year 0 queens are reared from selected colonies. The queens form colonies of their own and are tested in year 2 for characteristics like honey yield and PINtest. In year 2, a part of the tested colonies is selected and in the same year queens are reared from these selected colonies. The generation interval for the maternal path therefore equals 2 years. For the paternal pathway, a smaller part of the colonies is selected from which drone-producing queens are reared for mating stations. These queens are reared in year 2 and produce drones in year 3. Offspring queens therefore are reared in year 3 as well, and the generation interval for the paternal path equals three years. It might be noted that usually the drone-producing queens are full sisters and an unselected sample of the queens reared for this purpose, as they are not tested for characteristics like honey yield.

Are there ways to limit the paternal generation interval to two years? Schematically this is represented by Figure 3.

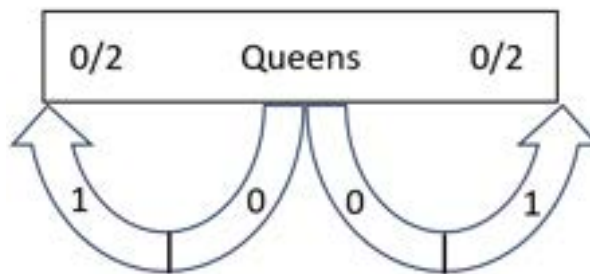


Figure 3. A breeding program where both the maternal and the paternal path take two years.

In the breeding program depicted by Figure 3, queens in a selected set of colonies that are tested in year 1 are used as drone-producing queens. The difference with Figure 2 is that colonies are not selected to rear drone-producing queens, but the queens in the colonies themselves are used as drone-producing queens. This reduces the paternal generation interval to two years, and the average generation interval from $(2+3)/2=2.5$ to $(2+2)/2=2$ years, leading to a reduction of the average generation interval by 20%. In the following paragraph we will discuss that this reduction also leads to a reduction of the numerator, but if this reduction is less than 20%, the breeding program in Figure 3 leads to higher ΔG than the one in Figure 2.

In essence, there are two ways to utilize the colonies tested in year 1 for the paternal path. The first is to utilize several full sisters and perhaps several selected full sisters. These may be placed on a mating station but used for instrumental insemination as well. The second way is to use several queens for single-queen insemination.

In principle it is also possible to reduce the maternal generation interval as depicted by Figure 4.

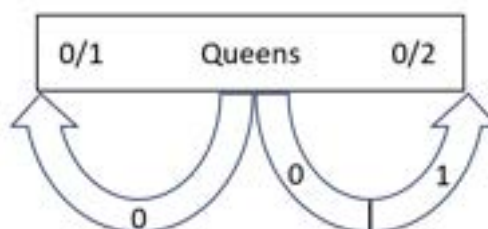


Figure 4. A breeding program where the maternal generation interval is reduced to one year.

Reduction of the maternal generation interval to 1 year implies that only characteristics can be selected for that do not need a full season to be recorded. It is likely that measuring a characteristic like SMR fruitfully can be carried out but not a characteristic like honey yield. The average generation interval in Figure 1 is $(1 + 2)/2 = 1.5$ years, a reduction by 40%, so clearly considerable reduction in the numerator is acceptable to still achieve higher ΔG in the program depicted by Figure 4 as compared to Figures 2 or 3.

Balancing changes in the average breeding value of selected parents and the generation interval

We have already mentioned that reducing the average generation interval by half a year in Figure 3 or to by 1 year in Figure 4, both compared to Figure 2, goes at the expense of the average breeding value of selected parents.

First, let us consider the case of Figure 3. In that case several full sisters are taken as drone-producing queens instead of selecting their mother (as in Figure 2) to produce drone-producing queens. As an example, suppose that in year 1 there are tested 400 colonies, consisting of 40 full-sister groups. Further suppose that there are 20 mating stations each occupied by a set of 10 full sisters. Then in the case of Figure 2, 20 out of 40 full-sister groups are selected as sets of drone-producing queens while in Figure 1, 20 out of 400 colonies are selected, from which drone-producing queens are raised. Obviously 20 out of 40 leads a lower average breeding value of selected parents than selection 20 out of 400 and a rough calculation learns that in this example the setup in Figure 2 leads to higher ΔG . But in other examples the setup of Figure 3 will outperform Figure 2.

In this example we assumed that the accuracy of estimated breeding values was similar in Figure 2 and Figure 3. We only considered a difference in selected proportion and resulting selection intensity.

In the comparison of the setup of Figure 4 with that in Figure 3, differences in proportion selected may play a minor role as the only difference is that between year 0 and year 1 there is some loss of colonies such that in Figure 3 there are less colonies to select from. It may well be that the lower accuracy of estimated breeding values in Figure 4 as compared to Figure 3 is the more important difference, as some characteristics will not be observed at all.

Balancing ΔG and ΔF

As pointed out in Chapter 3, the rate of inbreeding, ΔF , can be predicted based on the number of dams and the number of sires. The lower the number of dams and the number of sires, the higher ΔF .

An equation that predicts ΔF per generation equals

$$\Delta F = \frac{1}{4N_{dams}} + \frac{1}{4N_{sires}}$$

with N_{dams} and N_{sires} the numbers of dams and sires selected each year, respectively. This formula was developed for mammals and due to the biological peculiarities of the honeybee it does not fit perfectly for this species. Nevertheless, the formula can serve well to get a general idea about dynamics of inbreeding development and its influences, also for honeybees. Note that ΔF is the increment of average inbreeding per generation, from one generation to the other. On an annual basis, ΔF per year equals ΔF per generation divided by the average generation interval.

The prediction of ΔF can be used when planning a breeding program. An important message from the equation is that the level ΔF is affected most by the parent with the smallest number. If, for example, N_{sires} equals 10, ΔF equals approximately 2.5% irrespective whether N_{dams} equals 100 or 1000.

As discussed in Chapter 3, for a running selection program, ΔF can be computed from the pedigree and it is useful to check the level of ΔF as soon as the program is actually running. The Food and Agricultural Organisation of the United Nations (FAO) advises for a sustainable selection program to keep ΔF below 1% per generation. A discussion on this can be found in Bijma (2000).

For the equation, a closed population is assumed. This means that all dams and all sires are descendants from parents belonging to the population concerned. If, for example, it is common practice to introduce genetic material through open mating, ΔF will be very low, even if all dams are from the population. Assuming of course, that the drones contributing to the open mating are produced by a large pool of fairly unrelated queens, which is a realistic scenario under common beekeeping conditions today.

The equation also assumes that in the population selection takes place. Without selection, and random mating of dams and sires, ΔF would be roughly halved. Actually, it is the other way around. The equation for ΔF was derived for random mating without selection, and with selection ΔF is about twice as large. Another assumption is that there is no extreme variation in the numbers of progeny per dam and also not in the variation of progeny per sire. If, for example, a small number of sires contributes strongly to the numbers of offspring and a large number contributes very little, the small number is a far better approximation of N_{sires} than the total number.

In honeybees, N_{dams} equals the number of selected queens that contribute to the next generation. For single-queen insemination, also N_{sires} equals the number of selected queens that produce the drones used for insemination. The situation is less simple in case of mating stations where the queens heading the drone-producing colonies are full-sisters. The total number of drone-producing queens then strongly overestimates N_{sires} because of the relatedness of the queens on each mating station. Du *et al.* (2023) derived approximations and concluded that in that case N_{sires} equals a bit more than twice the number of dams of drone-producing queens. For 10 drone-producing queens per mating station N_{sires} was 2.2 times the number of dams of drone-producing queens and for 30 drone-producing queens it was 2.4 times.

Table 1. Predicted rate of inbreeding per generation (ΔF) for breeding programs using single-queen insemination and for breeding programs with mating stations with drone-producing colonies headed by full-sister queens.

Single-queen insemination									
Number of dams	20	20	40	40	40	60	60	60	60
Number of single queens ¹⁾	10	20	10	20	40	10	20	40	60
ΔF (%)	3.8	2.5	3.1	1.9	1.3	2.9	1.7	1.0	0.8
Mating stations									
Number of dams	20	20	40	40	40	60	60	60	60
Number of dams of drone-producing queens ²⁾	10	20	10	20	40	10	20	40	60
ΔF (%)	2.5	1.9	1.9	1.3	0.9	1.7	1.0	0.7	0.6

¹⁾ This is the number of queens per year that produce semen for single-queen insemination

²⁾ ΔF was computed taking N_{sires} as twice the number of dams of drone-producing queens

For single-queen insemination in principle the number of sires can be large, even equal to the number of dams. Table 1 suggests that some 40-60 sires per year are needed to keep ΔF below 1% per generation when the number of dams exceeds 60.

In a breeding program with mating stations, the number of dams of drone-producing queens equals the number of mating stations, assuming that every year the drone-producing queens descend from a fresh set of selected dams. Table 1 suggests that some 20 mating stations are needed to keep ΔF below 1%. This is a

challenge for small breeding programs such that for small breeding programs single-queen insemination looks promising.

It should be stressed that these are approximations. Often, in practice, numbers of offspring per selected parent vary greatly. Also other factors, like for example the genetic correlation between direct and maternal effect, play a role. Furthermore, selection based on breeding values including family information increases ΔF as compared to selection based on phenotypes; the reason is that in the former case it is more likely that close relatives are selected. When designing a breeding program, it is therefore wise to choose the safe side and start with larger numbers than suggested in Table 1.

Design of breeding programs in scientific literature

The equation for ΔG allows to approximate the value of annual genetic improvement for various designs of breeding programs. In previous paragraphs we gave some examples.

In many instances, the differences between alternative breeding programs deserve more careful analysis than quick back-of-the-envelope calculations. This usually is done by computer simulation. The basis of one approach is that for colonies phenotypes are generated taking the sum of half the breeding value of the dam, half the breeding value of the sire, a random factor which accounts for the effect of Mendelian sampling (see Chapter 3), and a random addend that accounts for residual effects. Out of the generated phenotypes then the best are selected to produce the next generation. Another approach simulates one level deeper, on the level of the effect of individual genes. The first approach in fact assumes a very large number of genes affecting the traits (infinite in principle) and the second approach allows smaller numbers. Both approaches do not differ very much in terms of conclusions when the time horizon of the simulation is 20 years or so, but on the longer run the first approach overestimates reality – if the second approach is the more realistic one (Plate *et al.* 2019a).

In the following, we present four examples of such simulations.

Du *et al.* (2023) studied the effect of the use of single queen insemination as discussed above (Figure 3). Often the limitation of the number of mating stations hampers the sustainability of breeding programs as inbreeding accumulates too fast. More sires for single queen insemination avoids this.

Plate *et al.* (2019b) studied the importance of controlled mating. As opposed to the simple approach in Chapter 3 they did not assume that with open mating the genetic level of the drone population doesn't change. They account for various alternatives where the passive population (colonies of beekeepers outside the breeding program) benefits of improvement in the selection program by simple mixing of drones in open air or because beekeepers in the passive population purchase improved stock from the breeding program. They showed that genetic improvement requires controlled mating. They also showed that the passive population benefits from the activities in the breeding program to such an extent that on the long run ΔG in the passive population equals that in the breeding program, be it with a lag in time. This lag in time can be many years if the transmission of genes from the breeding population to the passive population is limited.

Du *et al.* (2024) compared the use of pooled semen insemination with single queen insemination, either selecting on phenotypes or on breeding values. When selection is on phenotypes, increase in inbreeding turns out to be very similar. Also ΔG is similar for both alternatives. On the contrary, selecting for estimated breeding values shows a large advantage for single queen insemination in terms of ΔG , but also ΔF will be larger. The authors point out that the latter can be avoided by proper design of the breeding program. They also point out that pooled semen insemination can enhance the vitality of colonies because of a far larger diversity of

patrilines. Further, they suggest that a split of the bee population between a breeding program and a passive population that leans on the results of the former, would benefit from single queen insemination in the breeding program and pooled semen insemination in the passive population.

Kistler *et al.* (2024) compared a program according to Figure 4 with a program as in Figure 3 by simulation. Figure 3 represented a basic program where in the maternal path colonies were phenotyped for honey yield and hygienic behavior, while in the alternative program according to Figure 3 honey yield was not phenotyped. They concluded that in nearly all programs studied the alternative program outperformed the basic program, yielding up to 45% greater genetic gain. However, this genetic gain was accompanied by a relative increased mean inbreeding of about 20% to 35%.

Final remarks

This chapter provided possibilities to increase ΔG either by affecting the numerator and denominator of its prediction equation. The accuracy of estimated breeding values can be increased applying breeding value estimation as opposed to selection based on phenotypes. The selection intensity can safely be increased by enlarging the number of tested colonies as the decrease of the selected parents leads to increased ΔF . The chapter also made clear that careful consideration of the design of the breeding program in terms of generation intervals is commendable as the effect of their decrease usually are considerable.

References

- Bijma, P. (2000). Long-term Genetic Contributions Prediction of Rates of Inbreeding and Genetic Gain in Selected Populations. PhD-thesis Wageningen University. <https://edepot.wur.nl/198962>, pages 205-206
- Du, M., Bernstein, R., Hoppe, A. (2023) The Potential of instrumental Insemination for sustainable honeybee breeding. *Genes* 14, 1799-. DOI: 10.3390/genes14091799.
- Du, M., Bernstein, R., Hoppe, A. (2024). Comparison of pooled semen insemination and single colony insemination as sustainable honeybee breeding strategies. *R. Soc. Open Sci.* 11: 231556. DOI: 10.1098/rsos.231556.
- Kistler, T., Brascamp, E.W., Basso, B., Phocas, F., Bijma, P. (2024) How partial phenotyping to reduce generation intervals can help to increase annual genetic gain in selected honeybee populations. Non-refereed preprint. Submitted. DOI: 10.1101/2024.10.30.621079.
- Plate, M., Bernstein, R., Hoppe, A., Bienefeld, K. (2019a). Comparison of infinitesimal and finite locus models for long-term breeding simulations with direct and maternal effects at the example of honeybees. *PLoS ONE* 14,, e0213270. DOI: 10.1371/journal.pone.0213270
- Plate, M., Bernstein, R., Hoppe, A., Bienefeld, K. (2019b). The importance of controlled mating in honeybee breeding. *Genet Sel Evol* 51:74. DOI: 10.1186/s12711-019-0518-y.