GENETICS OF CROSSBREEDING

| Pascal Duenk

Genetics of crossbreeding

Pascal Duenk

Thesis committee

Promotor
Dr M.P.L. Calus
Associate Professor Animal Breeding and Genomics
Wageningen University & Research

Co-promotors
Dr P. Bijma
Assistant Professor Animal Breeding and Genomics
Wageningen University & Research

Dr Y.C.J. Wientjes
Researcher Animal Breeding and Genomics
Wageningen University & Research

Other members
Prof. Dr B. Zwaan, Wageningen University & Research, NL
Prof. Z.G. Vitezica, INRA Toulouse – Laboratoire GenPhySE,
Auzeville-Tolosane, FR
Dr D. Wittenburg, Leibniz-Institut für Nutztierbiologie (FBN), Dummersdorf, DE
Dr N. Duivesteijn, Hendrix Genetics, Boxmeer, NL

This research was conducted under the auspices of the Graduate School Wageningen Institute of Animal Sciences (WIAS).

Genetics of crossbreeding

Pascal Duenk

Thesis

submitted in fulfilment of the requirements for the degree of doctor at Wageningen University

by the authority of the Rector Magnificus

Prof. Dr. A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 13 March 2020

at 4 p.m. in the Aula.

Pascal Duenk
Genetics of crossbreeding
190 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2020)
With references, with summary in English

ISBN: 978-94-6395-254-5
DOI: https://doi.org/10.18174/509783

Abstract

Duenk, P. (2020). Genetics of crossbreeding. PhD thesis, Wageningen University, the Netherlands

In pig and poultry breeding programs, animals from genetically distinct purebred breeding lines are mated to produce crossbred animals, which provide food products to consumers. Although the aim of such breeding programs is to improve the performance of the crossbreds, selection takes place in the purebred lines, and is usually based on purebred performance. This strategy may be suboptimal because the genetic correlation between purebred and crossbred performance (r_{pc}) is usually lower than one. When r_{pc} is lower than one, it may be beneficial to make selection decisions based on information on crossbred performance instead of purebred performance. This is, however, a challenging task, because purebred animals cannot be tested directly for performance at the crossbred level. Now, with the recent developments in genomic prediction, it has become possible to estimate breeding values for crossbred performance of purebred animals. In this thesis, I studied the genetics of crossbreeding, with a focus on genomic prediction for crossbred performance in purebred lines. First, I illustrate how interactions between genes can lead to differences in genetic trait expression between lines, and how such interactions can lead to r_{pc} values that are lower than one. The results show that r_{pc} decreases as the genetic distance between parental lines increases. I derive expressions for r_{pc} based on genetic parameters in the parental lines, which allows breeders to estimate bounds of r_{pc} without having to collect crossbred data. Second, I show that genotype-based models lead to larger estimated r_{pc} with smaller standard errors than pedigree-based models. In contrast to my expectation, considering breed-of-origin of alleles in genotype-based models does not yield different estimates of r_{pc}. Third, I investigate the benefit of training the genomic prediction model with crossbred instead of purebred data. The results show that crossbred data improves the accuracy of breeding values for a trait with an r_{pc} of 0.8, but not for a trait with an r_{pc} of 0.96. Furthermore, taking the breed-of-origin of alleles into account is beneficial for a trait with an r_{pc} of 0.8, but not for a trait with an r_{pc} of 0.96. Finally, I discuss the relationship between r_{pc} and heterosis in the presence of gene interactions, and strategies to estimate breeding values for crossbred performance of purebreds. The results in this thesis improve our understanding of the genetics of crossbreeding, and facilitate the optimization of breeding programs that aim to improve crossbred performance with selection in purebred breeding lines.

Contents

1

General introduction

1.1 Introduction to quantitative genetics and animal breeding

Quantitative genetics studies the genetics of differences between individuals, for traits that are influenced by a large number of genes (i.e. complex traits). An example of such a trait (referred to as phenotype) is egg production in laying hens. Suppose we measure the egg production of all laying hens in the Netherlands, we will observe differences (or, variation) in the number of eggs produced between individual hens. This variation may be caused by differences in external factors such as nutrition, farm management and climate, and by differences in internal factors, such as disease status, energy balance, and genetics.

Animal breeding aims to improve the average performance of a population for certain traits of interest. Such improvements are made by selecting individuals that are genetically superior to become parents of the next generation. This selection process is repeated for many generations, causing an improvement in the average performance of the population in the desired direction. For example, in dairy cattle, selection over a period of 50 years has resulted in an increase of average yearly milk yield of approximately 50% (Hill 2008).

An improvement in average performance of the population is based on the fact that genes are transmitted from parents to offspring. In other words, part of the observed variation in performance of individuals is heritable. The heritable part of an individuals' total genetic merit is called its additive genetic value (A). An improvement in average performance of the population is only guaranteed when the average A of the selected individuals is higher than the average A in the whole population. In animal breeding, A is often called $breeding\ value$ (BV) of an individual, which is half the expected performance of its offspring (Falconer and Mackay 1996). Breeders achieve genetic progress each generation by selecting those animals from the selection candidates that have the highest estimate of their BV.

1.1.1 Estimation of breeding values

Breeding values of individuals cannot be measured or observed directly, but they can be estimated with statistical models that use measurements (i.e. phenotypic records) of the trait of interest. Traditionally, these phenotypes needed to be recorded on the selection candidates themselves or on their close relatives, and these records were connected through a pedigree-based relationship matrix ($\mathbf{A}$) in the statistical model to estimate BV. In the last few decades, the development of DNA marker arrays has made it possible to replace $\mathbf{A}$ with a genomic relationship

matrix ($\mathbf{G}$) that is constructed from the marker genotypes of individuals (VanRaden 2008). The use of marker genotypes has become economically attractive because the costs of genotyping per individual has dropped significantly over the past years (Eggen 2012). The estimation of BV with genomic marker data is known as *genomic prediction* (Meuwissen *et al.* 2001), and is nowadays applied in dairy cattle (Hayes *et al.* 2009b), pig (Knol *et al.* 2016), and poultry breeding (Wolc *et al.* 2016).

With genomic prediction, breeders set up a *reference population* that consists of animals that have both phenotypic records and genotypic records, and use this information to estimate BV of selection candidates that (only) have genotypic records. One of the major benefits of genomic prediction over traditional pedigree-based BV estimation is that young individuals can be selected for breeding without the need for phenotypes recorded on themselves or on their family members, thereby reducing the generation interval. In addition, genomic prediction may improve the accuracy of estimated BV compared to pedigree-based methods (Hayes *et al.* 2009c). Both these advantages may lead to an increased response to selection per year. Genomic prediction is especially beneficial for traits that can only be measured late in life (e.g. fertility, survival, or meat quality related traits), for traits that can only be measured on a single sex (e.g. milk/egg production), or for traits that are expressed only by relatives in a different environment (e.g. crossbreds).

1.2 Crossbreeding

In most pig and poultry breeding programs, animals from different populations (or, purebred lines) are mated to produce crossbred (CB) animals, which are the final production animals. For example, a typical poultry breeding program consists of four purebred (PB) parental lines (A, B, C, and D) that are mated to produce a four-way CB animal (AB)(CD) (Figure 1.1). Selection is only applied in the PB lines, and these lines usually differ in the traits that they are selected for. Dam lines may be selected for female fertility traits, whereas sire lines may be selected for growth or egg production traits, and feed efficiency. Breeders then benefit from breed complementarity in the CB, meaning that the CB animals combine the desirable traits that the PB lines were selected for (e.g. Smith (1964)). Another benefit of crossbreeding is heterosis (or hybrid vigor), which refers to the higher performance of CB individuals compared to the average performance of their PB parental lines (Shull 1952; Morris and Binet 1966; Dickerson 1973).

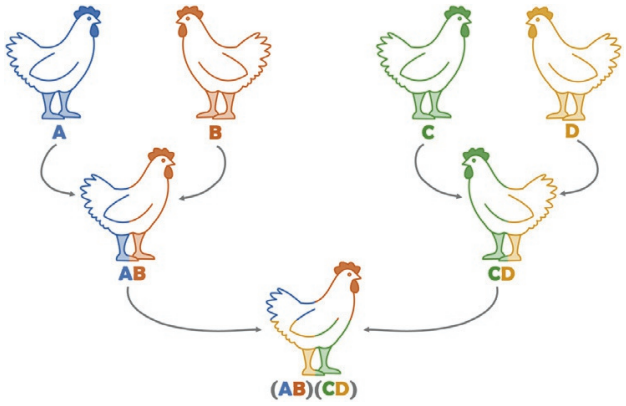

Figure 1.1 Example of a four-way crossbreeding program in poultry, where four parental lines are crossed to produce the final crossbred animal.

1.2.1 Selecting for CB performance

Although the aim of CB breeding programs is to improve the performance of CB animals, selection takes place in the PB lines, because the PB animals will be the parents of the next generation of crossbreds. The breeding values of the PB animals are typically estimated from performance records of PB animals. With such a selection strategy, the response to selection in the CB population is determined by the correlation between BV for PB and CB performance (i.e. the additive genetic correlation between PB and CB performance, r_{pc}). Across species, estimates of r_{pc} are generally lower than one for many trait categories (Wei and van der Werf 1995; Lukaszewicz *et al.* 2015; Wientjes and Calus 2017). When r_{pc} is lower than one, PB selection candidates that have the highest BV for PB performance may not have the highest BV for CB performance. As a result, the response to selection in CB performance may be suboptimal when selection is based on PB performance.

When r_{pc} is lower than one, it may be beneficial to include phenotypic records of CB relatives for the estimation of BV for CB performance (Wei and van der Werf 1994). This approach is known as combined crossbred and purebred selection (CCPS), where selection decisions are based on an index that combines PB and CB performance. CCPS improves the accuracy of EBV for CB performance (Ibanez-Escriche *et al.* 2011), and improves the response to selection in CB performance (Bijma and van Arendonk 1998). However, CCPS also results in more inbreeding (Bijma *et al.* 2001), and requires routine recording of pedigree information in CB animals, which is not always possible. As I discussed in section 1.1.1, the need for pedigree information can be alleviated by the use of genomic marker data. Thus, when the goal is to estimate BV for CB performance, genomic prediction is a promising alternative to pedigree-based prediction.

With genomic prediction for CB performance, the reference population can consist of PB animals, CB animals, or both. As r_{pc} decreases, the advantage of using CB instead of PB animals increases. It is therefore important to find the upper bound of r_{pc} below which there is a benefit to use CB instead of PB information. Simulation studies have shown that CB information can increase the accuracy of estimated breeding values (EBV) for CB performance when r_{pc} is lower than ~0.8 (Dekkers 2007; Esfandyari *et al.* 2015a; Van Grevenhof and Van Der Werf 2015). This result was, however, only found for scenarios where the CB reference population was at least of the same size as the PB reference population, and the PB selection candidates had similar relationships to the PB and CB reference populations. A study with real data from pigs reported an r_{pc} of ~0.9, and lower accuracies of EBV with a CB versus a PB reference population (Hidalgo *et al.* 2016). In this study, the CB reference population was smaller than the PB reference population, and the PB reference population had weaker relationships with the selection candidates than the CB reference population. In contrast to this result, Lopes *et al.* (2017) found higher accuracies of EBV for CB performance with a CB versus a PB reference population for a trait with an r_{pc} of ~0.9. It should be noted, however, that the studies of Hidalgo *et al.* (2016) and Lopes *et al.* (2017) differed in how accuracies were obtained (i.e. how estimated BV for CB performance were validated). An important question that therefore remains is how estimated BV for CB performance should be validated. In addition, there appear to be no studies on the benefit of using CB information for genomic prediction in poultry.

1.3 The genetic correlation between purebred and crossbred performance

One reason that the genetic correlation between purebred and crossbred performance (r_{pc}) can be lower than 1 is the presence of genotype by environment (GxE) interactions (Falconer 1952; Lutaaya *et al.* 2001; Dekkers 2007). Such GxE interactions arise because PB and CB animals are usually kept in different environments: PB animals are typically kept in a nucleus environment under excellent management conditions and high levels of biosecurity (e.g. disease-free), whereas CB animals are kept under field conditions with varying levels of management quality and lower levels of biosecurity. Genes that have no effect or that are harmful for the trait of interest in the nucleus environment may be beneficial in the conventional environment, and vice versa. For example, a PB animal with an above-average growth in the nucleus environment may have a below-average growth in the conventional environment, simply because that animal lacks

genes that are important for disease resistance. Hence, GxE interaction introduces differences in the performance ranking of individuals between environments (Falconer 1952).

1.3.1 Relationship between r_{pc} and non-additive effects

Next to GxE interactions, the r_{pc} can be lower than one due to genotype by genotype (GxG) interactions (i.e. non-additive effects) in combination with differences in allele frequencies between the PB parental lines (McNew and Bell 1971; Wei *et al.* 1991). Non-additive effects can be classified into dominance and epistasis: dominance arises through interactions between alleles at the same locus, whereas epistasis arises through interactions between alleles at different loci (Bateson and Mendel 1909; Falconer and Mackay 1996). Knowledge of the relationship between non-additive effects and r_{pc} may contribute to the understanding of genetic architectures of polygenic traits, and help breeders to predict r_{pc} based on the importance of dominance and epistasis in trait expression, and on the genetic distance between parental lines as expressed in allele frequency differences. However, the impact of non-additive effects and the difference in allele frequencies between PB and CB on r_{pc} is not well understood.

The r_{pc} is expected to decrease with increasing magnitude of non-additive effects, and with increasing differences in allele frequencies between parental lines (Wei *et al.* 1991; Baumung *et al.* 1997). For example, Wei *et al.* (1991) studied the role of dominance in the value of r_{pc}, and found that r_{pc} decreases with increasing magnitude of dominance effects and with increasing difference in allele frequencies between the PB parental lines. In their results, the r_{pc} could become negative when the absolute magnitude of the dominance effect (d) was larger than the absolute magnitude of the additive effect (a) (i.e. overdominance). The role of epistasis in the value of r_{pc} was studied by Baumung *et al.* (1997), who found that epistasis affected the value of r_{pc} much less than dominance did. With pure additive by additive (AxA) epistasis, r_{pc} dropped below 0.6 only when the epistatic effect (ϵ) was larger than the marginal additive effects (a), and the differences in allele frequencies between parental lines was larger than 0.6 for both loci.

The studies of Wei *et al.* (1991) and Baumung *et al.* (1997) considered only two-locus models, where the non-additive effects and difference in allele frequencies between parental lines were both relatively large, and where there was only AxA epistasis. In reality, quantitative traits are influenced by many loci, non-additive effects are usually small (Bennewitz and Meuwissen 2010; Wei *et al.* 2014; Sun and

Mumm 2016), and the differences in allele frequencies between parental lines are a result of both drift and selection. In addition, epistasis interactions may be of a different form than AxA. Hence, there is a need for a more realistic model of line differentiation and non-additive effects to understand their impacts on r_{pc}.

1.3.2 Estimation of r_{pc}

Knowledge of the value of r_{pc} is important because it can help breeders to decide whether data should be collected on CB animals for the estimation of breeding values in PB lines (see subsection 1.2.1). The r_{pc} can be estimated with bivariate models that treat PB and CB performance as two correlated traits (Wei and van der Werf 1995), and the link between PB and CB observations can be established through either a pedigree-based ($\mathbf{A}$), or a genomic relationship matrix ($\mathbf{G}$). Estimating r_{pc} with $\mathbf{A}$ requires a pedigree that can connect the PB and CB animals that have performance records. When PB and CB animals are paternal half-sibs, the accuracy of estimated r_{pc} depends on the number of sires the PB and CB animals have in common, and on the accuracy of estimated breeding values of those sires (Bijma and Bastiaansen 2014). In practice, however, pedigree data is often not recorded in crossbred breeding programs, and the number of sires that have both PB and CB offspring may be limited.

Alternatively, r_{pc} can be estimated with $\mathbf{G}$, thereby alleviating the requirement of strong pedigree relationships between PB and CB animals. In addition, using $\mathbf{G}$ may result in a smaller standard error of the estimate of r_{pc} (Visscher *et al.* 2014; Xiang *et al.* 2016), because the relationships in $\mathbf{G}$ may be a more accurate representation of relationships between individuals at QTL than the relationships in $\mathbf{A}$ (Goddard 2009; Hayes *et al.* 2009c). However, the differences in estimated r_{pc} between models that use $\mathbf{G}$ instead of $\mathbf{A}$ have not been studied before, and the benefit of using $\mathbf{G}$ rather than $\mathbf{A}$ for the estimation of r_{pc} has not been quantified.

1.4 The breed-of-origin of alleles

A diploid individual inherits a single allele of each of its parents. Hence, a two-way CB animal inherits a single allele from each parental line. In the ordinary genomic relationship matrix ($\mathbf{G}$), the relationships between PB selection candidates and CB animals are based on all observed alleles in the CB animals. In other words, the ordinary $\mathbf{G}$ ignores that CB animals only share 1 allele of each locus with the PB line of interest. Recent developments in genotype analyses have made it possible to trace back the breed-of-origin of alleles (BOA) in CB animals, and account for the fact that

CB animals share only 1 allele at each locus with the PB line of interest (Vandenplas *et al.* 2016). Information on the BOA makes it possible to construct a *partial* genomic relationship matrix ($\mathbf{G}_{BOA}$), wherein relationships between PB and CB animals are based on alleles that originated from the same breed (Ibañez-Escriche *et al.* 2009; Christensen *et al.* 2014). Based on theory, it can be shown that considering the BOA is more appropriate than ignoring it, because the relationships in $\mathbf{G}_{BOA}$ should better represent the actual relationships between the PB parental line of interest and the CB animals at QTL than the relationships in $\mathbf{G}$.

1.4.1 Benefits of considering the BOA for genomic prediction

When the BOA is considered, the genomic prediction model allows for marker alleles in the CB animals that originate from different breeds to have different effects on phenotypes. Considering the BOA in genomic prediction for CB performance may therefore be beneficial, because actual marker effects may differ between breeds because of differences in linkage disequilibrium between breeds (de Roos *et al.* 2008; Veroneze *et al.* 2014; Fu *et al.* 2015), or because the actual effects at QTL are different (Fisher 1918; Falconer 1952; Wei *et al.* 1991; Baumung *et al.* 1997). Simulation studies indeed suggested that considering BOA can improve the response to selection, but only for scenarios where the CB reference population was large, the number of markers was small, and the parental lines of the CB animals were distantly related (Ibañez-Escriche *et al.* 2009; Esfandyari *et al.* 2015a). Empirical studies in pigs indicate that considering BOA may yield more accurate EBV than ignoring BOA, but only when the heritability and r_{pc} of the studied trait are low (Lopes *et al.* 2017; Sevillano *et al.* 2017). To my knowledge, there have been no studies on the benefits of considering the BOA for estimation of r_{pc}, or for genomic prediction in poultry.

1.5 Objective of this thesis

The overall objective of this thesis was to study the genetics of crossbreeding in the context of genomic prediction, with a focus on the role of non-additive effects. The thesis specifically addresses the knowledge gaps identified in the above sections 1.2 through 1.4. In chapter 2, the focus is on estimating the average effects of QTL. In this chapter, I investigate the benefits of explicitly modelling the dominance effect for the estimation of average effects at QTL. Chapter 3 presents theoretical work on the genetic correlation between populations (r_g), and investigates how the value of r_g is affected by non-additive effects and differences in allele frequencies between populations. Chapter 4 continues this work by investigating the relationship between r_g of different PB lines, and the genetic correlation between performance in one of

those PB lines and the performance of their CB offspring (r_{pc}). I present simple equations for the value of r_{pc} based on information from the PB parental lines only, for scenarios where there is only dominance or additive by additive epistasis. In chapter 5, I used real data from approximately 4700 PB and 10,500 CB animals to estimate the r_{pc} of two body weight traits in broilers. The aim was to compare estimates of r_{pc} from different methods using either pedigree or genomic data, and where the breed-of-origin of alleles (BOA) in CB animals was either considered or ignored. Lastly, in chapter 6, I use the same data as in chapter 5 to compare accuracies of genomic estimated breeding values for CB performance, using either a PB or CB reference population, and where the BOA in CB animals was either considered or ignored. Furthermore, I compared validation of GEBV for CB performance based on offspring averages with validation based on individual performance records. In the final chapter of this thesis (chapter 7), I discuss two topics. First, I discuss the relationship between non-additive effects and heterosis (i.e. the increased performance of crossbreds compared to the average performance of parental lines), and I show that heterosis is closely related to r_{pc} through the existence of non-additive effects. Second, I discuss strategies to estimate GEBV for CB performance, and highlight strengths and weaknesses of each strategy.

2

Benefits of dominance over additive models for the estimation of average effects in the presence of dominance

Pascal Duenk, Mario P.L. Calus, Yvonne C.J. Wientjes and Piter Bijma

Animal Breeding and Genomics, Wageningen University and Research,
6700 AH Wageningen, the Netherlands

G3: Genes|Genomes|Genetics (2017) 7(10): 3405-3414

Abstract

In quantitative genetics, the average effect at a single locus can be estimated by an additive (A) model, or an additive plus dominance (AD) model. In the presence of dominance, the AD-model is expected to be more accurate, because the A-model falsely assumes that residuals are independent and identically distributed. Our objective was to investigate accuracy of an estimated average effect ($\hat{\alpha}$) in the presence of dominance, using either a single locus A-model or AD-model. Estimation was based on a finite sample from a large population in Hardy-Weinberg equilibrium (HWE), and root mean squared error of $\hat{\alpha}$ was calculated for several broad sense heritabilities, sample sizes, and sizes of the dominance effect. Results show that with the A-model, both sampling deviations of genotype frequencies from HWE frequencies and sampling deviations of allele frequencies contributed to the error. With the AD-model, only sampling deviations of allele frequencies contributed to the error, provided that all three genotype classes were sampled. In the presence of dominance, the root mean squared error of $\hat{\alpha}$ with the AD-model was always smaller than with the A-model, even when the heritability was smaller than one. Remarkably, in the absence of dominance, there was no disadvantage of fitting dominance. In conclusion, the AD-model yields more accurate estimates of average effects from a finite sample, because it is more robust against sampling deviations from HWE frequencies than the A-model. Genetic models that include dominance, therefore, yield higher accuracies of estimated average effects than purely additive models when dominance is present.

2.1 Background

In quantitative genetics, dominance is the phenomenon that the genotypic value of the heterozygote deviates from the mean genotypic value of the two homozygotes (Falconer and Mackay 1996). Dominance has shown to play an important role in production traits of livestock species (Morris and Binet 1966; Sellier 1976; Visscher et al. 2000) and plant crops (Xiao et al. 1995; Stuber 2010; Huang et al. 2016). In livestock genetic improvement, however, research has been focused on the estimation of average effects, because average effects capture all heritable variation (Lynch and Walsh 1998). The average effect of a single gene (α), also known as the allele substitution effect, is defined as the linear regression coefficient of genotypic values on allele counts (Falconer and Mackay 1996). Under Hardy-Weinberg equilibrium, the α at a bi-allelic locus is a function of the additive (a) and dominance (d) part of gene effects, and the population allele frequency p

$$\alpha = a + (1 - 2p)d, \qquad \textbf{2.1}$$

where a is half the difference in genotypic value between both homozygotes, and d is the difference between the genotypic value of the heterozygote and the average genotypic value of both homozygotes. With genomic data, additive (A) models estimate α by linear regression of phenotypes on allele counts (i.e. genotypes). Additive plus dominance (AD) models estimate the additive and dominant gene effects separately, after which $\hat{\alpha}$ can be obtained from Equation 2.1 (Lynch and Walsh 1998). For both models, the part of dominance that is not captured by the average effect is called the dominance deviation (Falconer and Mackay 1996).

When the A-model is used, dominance deviations are not modelled and thus become part of the residual. As a consequence, the residuals are not independent and identically distributed (IID), because dominance deviations are different across genotypes (Ott and Longnecker 2010). The A-model may therefore give inaccurate estimates of α, because it falsely assumes that the residuals are IID. When the AD-model is used, dominance deviations are explicitly modelled, and the residuals will more likely be IID. In the presence of dominance, the AD-model may therefore yield more accurate estimates of α than the A-model. In contrast to the A-model, however, the AD-model requires the estimation of two effects instead of one (for a single locus), which may reduce the accuracy with which these effects are estimated. Additionally, dominance effects are generally smaller and therefore harder to estimate than additive effects (Lynch and Walsh 1998). For these reasons, the AD-model may require more individuals to be sampled for an accurate estimation of α, compared to the A-model. Furthermore, estimating dominance effects when there is very little or no dominance may lead to overfitting (Ott and Longnecker 2010).

Hence, while the AD-model may better fit the data in the presence of dominance, the A-model may be preferred when the sample size is relatively small and dominance is negligible. It is, however, not yet clear how sample size and dominance effect size affect the accuracy of $\hat{\alpha}$ with the A-model versus the AD-model.

The objective of this work, therefore, was to investigate the root mean squared error (RMSE) of the estimated average effect ($\hat{\alpha}$) at a single locus in the presence or absence of dominance, using either an additive (A) model or an additive plus dominance (AD) model. We start with some theory of a single locus model, then derive the expected estimate of α, and calculate RMSE of $\hat{\alpha}$ for several broad sense heritabilities, dominance effects, sample sizes, and allele frequencies. We then calculate the mean RMSE for several degrees of dominance over the distribution of allele frequency, and identify mechanisms that underlie the differences between the A-model and AD-model.

2.2 Theory

Our interest is to estimate the average effect (α) at a single locus in a large population that is in Hardy-Weinberg equilibrium (HWE), from data collected as a finite sample of that population. The average effect will be treated as a fixed effect (as in quantitative genetics), and not a random variable (as in genomic prediction (de los Campos *et al.* 2015)). In quantitative genetics, α at a single locus can be estimated from the sample by linear regression using an additive model (A) or an additive plus dominance (AD) model. The A-model estimates α directly through linear regression of phenotypic values on allele counts,

$$\mathbf{y} = \mathbf{x}\alpha + \mathbf{e}, \qquad \textbf{2.2}$$

where $\mathbf{y}$ is a vector of centered phenotypes, $\mathbf{e}$ is a vector of residuals, and $\mathbf{x}$ is a vector of centered allele counts with $(0 - 2p_s)$ for individuals with 0 copies of the alternative allele, $(1 - 2p_s)$ for individuals with 1 copy, and $(2 - 2p_s)$ for individuals with 2 copies. The term p_s is the allele frequency of the alternative allele, observed in the sample. Throughout this paper, we will use the term genotypes to indicate the three allele count classes, with values of 0, 1, or 2.

With the A-model, the ordinary least squares estimate (LSE) of α is

$$\hat{\alpha}_A = [\mathbf{x}'\mathbf{x}]^{-1}[\mathbf{x}'\mathbf{y}]. \qquad \textbf{2.3}$$

The AD-model estimates the additive (a) and dominant (d) gene effects by multiple linear regression

$$\mathbf{y} = \mathbf{x}a + \mathbf{m}d + \boldsymbol{\varepsilon}, \quad \textbf{2.4}$$

where $\mathbf{m}$ is a dominance indicator vector with $(0 - 2p_s(1 - p_s))$ for homozygous individuals, and $(1 - 2p_s(1 - p_s))$ for heterozygous individuals. Vectors $\mathbf{y}$ and $\mathbf{x}$ are the same as in the A-model, and $\boldsymbol{\varepsilon}$ is a vector of residuals. Note that this is the genotypic parameterization as described by Vitezica *et al.* (2013). With the AD-model, the LSE of a and d are

$$\begin{bmatrix}\hat{a}\\ \hat{d}\end{bmatrix} = \begin{bmatrix}\mathbf{x}'\mathbf{x} & \mathbf{x}'\mathbf{m}\\ \mathbf{m}'\mathbf{x} & \mathbf{m}'\mathbf{m}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{x}'\mathbf{y}\\ \mathbf{x}'\mathbf{m}\end{bmatrix}. \quad \textbf{2.5}$$

The $\hat{\alpha}$ from the AD-model is subsequently calculated as

$$\hat{\alpha}_{AD} = \hat{a} + (1 - 2p_s)\hat{d}. \quad \textbf{2.6}$$

By definition, $\hat{\alpha}$ from both models give an estimate of the average effect in the sample (Falconer and Mackay 1996). Because the size of the sample is finite, genotype and allele frequencies in the sample might deviate from the frequencies in the total population. These deviations might introduce error in the estimation of α. To investigate the effects of finite sample size in the presence of dominance, the estimates from the A-model ($\hat{\alpha}_A$) and the AD-model ($\hat{\alpha}_{AD}$) were compared by computing their RMSE for several scenarios.

2.2.1 Expectation of $\hat{\alpha}$

If we take a random sample of N individuals from a large population in HWE that has allele frequency p, the expectation of $\hat{\alpha}$ can be computed using probabilities and estimates of each possible sample composition. We define c as a set of variables $\{n_0, n_1, n_2\}$ that describe unique sample compositions, where n_0 is the number of individuals with genotype 0, n_1 is the number of individuals with genotype 1, and n_2 is the number of individuals with genotype 2. The probability of sampling c is calculated from the multinomial probability function

$$P(c|N,p) = \frac{N!}{n_0!\,n_1!\,n_2!} g_0^{n_0} g_1^{n_1} g_2^{n_2}. \quad \textbf{2.7}$$

Conditional variables N and p are hereafter omitted to improve readability, so that $P(c|N,p)$ is abbreviated as $P(c)$. The quantities g_0, g_1 and g_2 are the genotype frequencies in the HWE population, and follow from the population allele frequency p ($g_0 = (1-p)^2$; $g_1 = 2p(1-p)$; $g_2 = p^2$).

The expectation of $\hat{\alpha}$ is computed as the sum over all products of probabilities $P(c)$ and corresponding estimates $\hat{\alpha}(c)$,

$$E(\hat{\alpha}) = \sum_{n_0=0}^{N-1} \sum_{n_1=0}^{N-n_0} I(c)P(c)\hat{\alpha}(c), \quad \textbf{2.8}$$

where $\hat{\alpha}(c)$ is the LSE of α given c. The α cannot be estimated when the sample consists of individuals that all have the same genotype, so we use $I(c)$ as an indicator variable to exclude such samples

$$I(c) = \begin{cases} 0, & n_1 = N \\ 0, & n_0 + n_1 = 0 \\ 1, & all\ other. \end{cases} \quad \textbf{2.9}$$

Note that samples including only genotypes 0 are excluded from Equation 2.8, by summing n_0 from 0 to $N-1$, instead of from 0 to N. After excluding samples with $I(c) = 0$, the probabilities $P(c)$ of the remaining samples were rescaled so that they sum to 1.

2.2.2 Root mean squared error

The RMSE is defined as the root of the expected squared difference between the $\hat{\alpha}$ estimated from the sample, and the true value of α

$$RMSE(\hat{\alpha}) = \sqrt{E[(\hat{\alpha} - \alpha)^2]} = \sqrt{\sum_{n_0=0}^{N-1} \sum_{n_1=0}^{N-n_0} I(c)\delta(c)}, \quad \textbf{2.10}$$

where $\delta(c)$ is the contribution of finite sampling deviation to the RMSE

$$\delta(c) = P(c)(\hat{\alpha}(c) - \alpha)^2. \quad \textbf{2.11}$$

We define $\delta(c)$ here because we will later on focus on the contribution of a single finite sample c to the RMSE. The above expressions will be used to investigate the effect of N, H^2, p, and d on the RMSE of $\hat{\alpha}$ with the A-model and the AD-model.

2.3 Methods

We aim to illustrate the effect of sample size (N), broad sense heritability (H^2), allele frequency p, and dominance effect d, on RMSE of estimated average effects ($\hat{\alpha}$). As a base scenario, we chose 1 for both the additive and dominance effect of the gene (e.g. full dominance). The expected value of $\hat{\alpha}$ was calculated for $N \in \{300, 500, 1000\}$, $H^2 \in \{0.01, 0.05, 1\}$, and $p = [0.001 - 0.999]$ (increments of 0.001), with the A-model (Equation 2.3) and AD-model (Equation 2.6). The variation in broad sense heritability was achieved by adding random residuals to the phenotypes ($\mathbf{y}$). In addition, we varied the dominance effect ($d \in \{0, 0.1, 0.2, 0.5\}$) for the scenario where $N = 500$ and $H^2 = 0.05$.

The $\hat{\alpha}(c)$ from the AD-model were computed using the sample allele frequency (p_s) in Equation 2.6 instead of the population allele frequency (p), because the latter is usually unknown. For samples where one of the genotypes was missing, $\hat{\alpha}(c)$ with the AD-model was computed in the same way as with the A-model, because in those cases the vector of genotypes $\mathbf{x}$ was completely confounded with dominance vector $\mathbf{m}$.

Additionally, to quantify the average accuracy of $\hat{\alpha}$, we computed the mean RMSE of $\hat{\alpha}$, assuming a distribution for the allele frequency. For this purpose, we used the RMSE as a function of p and numerically integrated over p using its expected distribution under a drift model,

$$\overline{RMSE} = \int_{\frac{1}{2N_e}}^{1-\frac{1}{2N_e}} f(p)RMSE(p)dp. \qquad \textbf{2.12}$$

Here, N_e is the effective population size, $f(p)$ is the distribution of allele frequencies when mutation is ignored, p ranges from $\frac{1}{2N_e}$ to $1-\frac{1}{2N_e}$ (Wright 1931; Goddard 2009), and

$$f(p) = \frac{k}{2p(1-p)}. \qquad \textbf{2.13}$$

To ensure that $\int f(p)dp = 1$, k was given a value of $\frac{1}{log(2N_e-1)}$. The resulting distribution of allele frequencies is U-shaped, and a low N_e yields a more uniform distribution than a high N_e. We computed the mean RMSE for several N_e (50, 100, and 200), N (200 to 600), and sizes of dominance effect d (0.5, 1 and 1.5). We considered sample sizes up to 600 instead of 1000 to reduce computation time. In these scenarios, both H^2 and the additive gene effect (a) were equal to 1.

The data used can be regenerated exactly following the descriptions of the paper.

2.4 Results

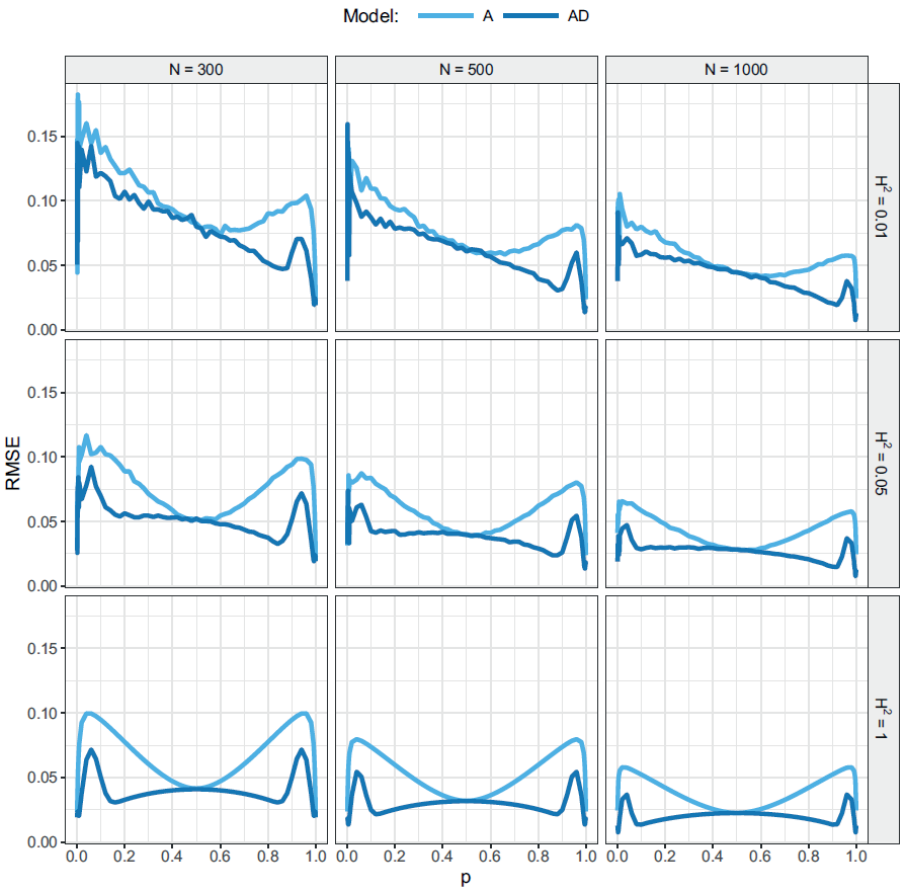

Figure 2.1 Root mean squared error (RMSE) of $\hat{\alpha}$ with the A- and AD-model. Presented as a function of broad sense heritability (H^2), population allele frequency (p) and sample size (N). The additive and dominant effect of the gene were both equal to 1.

2.4.1 Root mean squared error

Figure 2.1 shows the root mean squared error (RMSE) of $\hat{\alpha}$ with the A- or AD-model, for $a = 1$ and $d = 1$. For all scenarios, the RMSE of $\hat{\alpha}$ was smaller with the AD-model than with the A-model.

In scenarios where $H^2 = 1$, RMSE was symmetrical around $p = 0.5$ with both the A- and AD-model. For brevity, we will therefore only describe the pattern for $p < 0.5$. For both models and all N, the RMSE was smallest when p was close to 0, and increased when allele frequency increased. With the A-model, RMSE was largest around $p = 0.04$ and then decreased when p moved towards 0.5. With the AD-

model, RMSE was also largest around $p = 0.04$, then decreased when p moved towards 0.1, after which RMSE slightly increased again until $p = 0.5$.

With $H^2 < 1$, RMSE showed a similar pattern, but was not symmetrical around $p = 0.5$. Compared to $H^2 = 1$, the RMSE was larger for all p, but this contrast decreased when p increased. This asymmetry was a result of fixing H^2 in the simulations, which caused the ratio of the dominance variance and residual variance to increase with p. For all scenarios, RMSE decreased when N increased.

Figure 2.2 shows the RMSE of $\hat{\alpha}$ with the A- or AD-model, for $a = 1$, $N = 500$, $H^2 = 0.05$, and different dominance effects (d). For $d = 0$ and $d = 0.1$, there was almost no difference in RMSE between the A- and AD-model. This indicates that in the absence of dominance, there was no disadvantage of using the AD-model in terms of RMSE. For $d = 0.1$ there was no apparent benefit from using the AD-model. For $d = 0.2$ and $d = 0.5$, however, the AD-model had lower RMSE than the A-model.

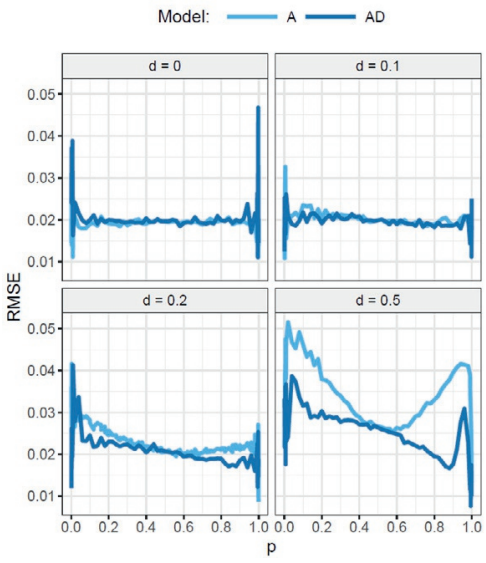

Figure 2.2 Root mean squared error (RMSE) of $\hat{\alpha}$ with the A- and AD-model for several sizes of dominance effect (d). Broad sense heritability is 0.05 and sample size is 500. Presented as a function of population allele frequency (p). Additive effect of the gene was equal to 1.

2.4.2 Contribution of finite sampling deviation to the RMSE

When there is no environmental variance ($H^2 = 1$) and the model is correct, the RMSE of $\hat{\alpha}$ is expected to be zero. The results, however, show that the RMSE is larger than zero with both the A- and AD-model. To gain more insight into the sources of this error, we investigated the contribution of single samples to the RMSE, for one scenario where $H^2 = 1$, $N = 300$, $p = 0.10$, and $a = d = 1$, so that $\alpha = 1.8$. For this purpose, we studied the squared difference between $\hat{\alpha}(c)$ and α (i.e. squared error), as a function of the realized number of individuals with genotype 2

(n_2). The samples have different probabilities of occurring, so that some samples may contribute more to the total RMSE than others. We therefore investigated the contribution of finite sampling deviation to the RMSE ($\delta(c)$), by weighting the squared errors of $\hat{\alpha}$(c) with their probabilities (see Equation 2.11).

2.4.2.1 Additive model

Figure 2.3a shows the squared error as a function of the realized number of individuals with genotype 2, for the A-model. The realized number of individuals with genotype 2 in the sample is expressed as a departure from its expectation (i.e. Δn_2), where the expectation is $E(n_2) = p^2N = 3$. The squared error was smallest when Δn_2 was zero and increased as Δn_2 moved away from zero. The remaining variance in squared error for a given value of Δn_2 (as shown by the boxplots) was due to variation in the difference between p_s and p (i.e. Δp). For example, when $\Delta n_2 = 3$, the allele frequency in the sample can vary, because the number of sampled heterozygotes can vary. This variation in Δp affects $\hat{\alpha}(c)$, except when $\Delta n_2 = -3$. In that case, the number of individuals with genotype 2 was zero (in this example) and $\hat{\alpha}(c)$ was always the slope of a line between two data points.

Figure 2.3b shows the effect of Δn_2 and Δp on $\delta(c)$ for the A-model. The sample where $\Delta n_2 = 0$ and $\Delta p = 0$ did not contribute to the RMSE ($\delta(c) = 0$). Samples where $\Delta n_2 < 0$ had the largest contributions to the RMSE, and samples where $\Delta n_2 > 0$ had somewhat smaller contributions. Figure 2.3b also shows that Δp contributed less to the RMSE than Δn_2, because there were samples where $\Delta p = 0$, but $\delta(c)$ was relatively large.

2.4.2.2 Additive + dominance model

Figure 2.3c shows the squared error as a function of Δn_2, for the AD-model. The squared error was small and about equal for all Δn_2, except for $\Delta n_2 = -3$, where the squared error was largest and exactly the same as with the A-model (see Figure 2.3a), because there were no individuals with genotype 2 in the sample. Similar as with the A-model, the remaining variance (as shown by the boxplots) was due to variation in the difference between p_s and p (Δp).

Figure 2.3d shows the effect of Δn_2 and Δp on $\delta(c)$ for the AD-model. Samples where both $\Delta n_2 \neq -3$ and $\Delta p = 0$, did not contribute to the RMSE ($\delta(c) = 0$). Samples where $\Delta n_2 = -3$ showed the largest contribution, while all other samples showed small $\delta(c)$. Similar to the A-model, Figure 2.3d shows that Δp was not an important source of error.

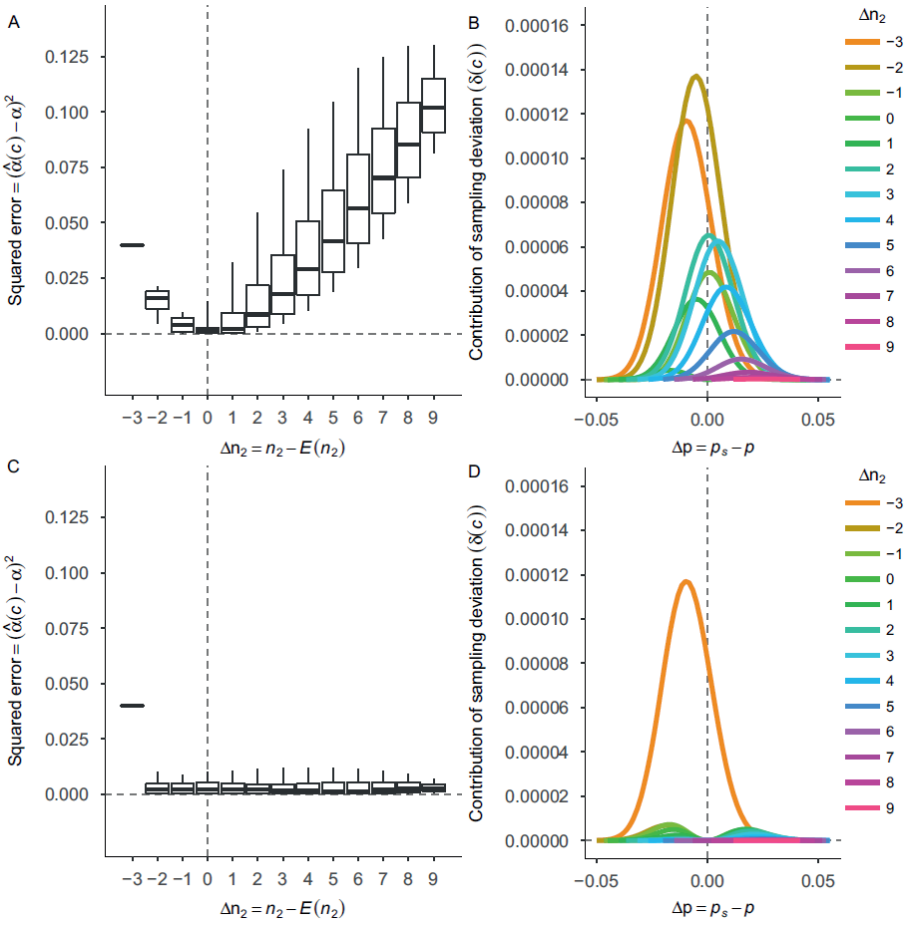

Figure 2.3 Squared errors of $\hat{\alpha}$ and contributions of samples to the RMSE for the A-model (panel A and B) and AD-model (panel B and C), for $N = 300$, $H^2 = 1$, and $p = 0.10$. The additive and dominant effect of the gene were both equal to one. A and C) Squared error of $\hat{\alpha}(c)$ as a function of the departure of n_2 from its expected value under HWE (Δn_2). (B and D) The effect of Δn_2, and deviations of sample allele frequency from population allele frequency, on the contributions of samples to the RMSE of $\hat{\alpha}$.

2.4.2.3 A- versus AD-model

In conclusion, even when the locus explains all variance (i.e., $H^2 = 1$), $\hat{\alpha}$ shows error with both the A- and AD-model when it is based on a finite random sample from a population in HWE and dominance is present. With the A-model, the error originated mainly from sampling deviations of genotype frequencies from expected HWE frequencies (ΔHWE), and to a lesser extent from sampling deviations of allele frequencies (Δp) (Figure 2.3b). With the AD-model, the error originated from Δp only, provided that all three genotype classes were sampled (Figure 2.3d). These

results partly explain the patterns of RMSE in Figure 2.1 (see Appendix A for more detail).

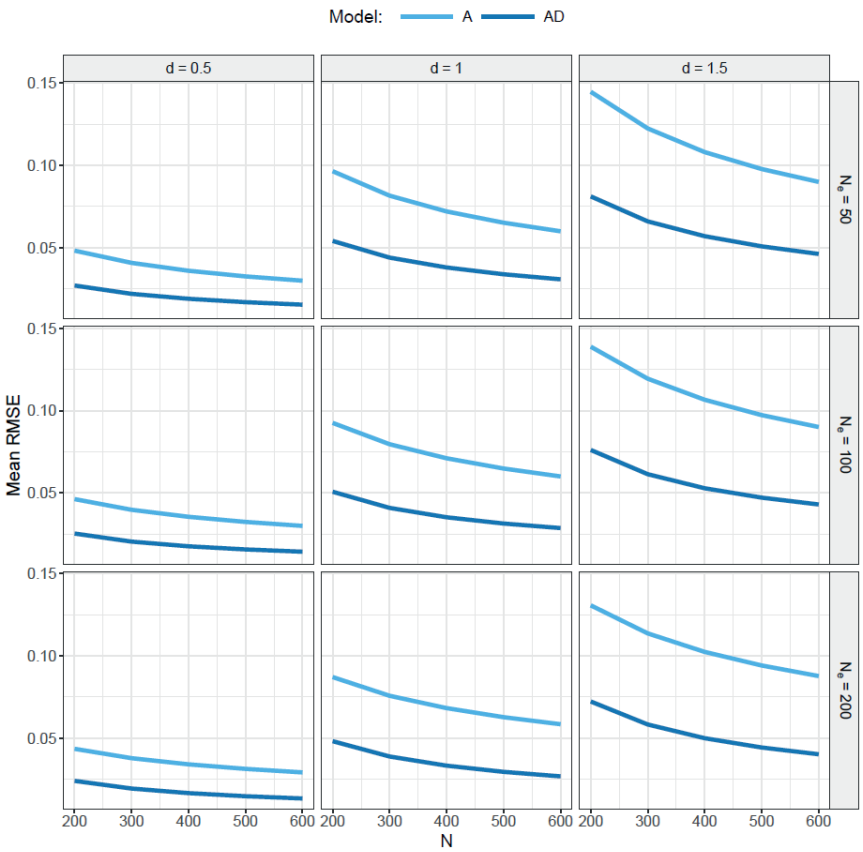

Figure 2.4 Mean RMSE of $\hat{\alpha}$ for the A- and AD-model, averaged over the distribution of $\boldsymbol{p}$. Presented as a function of dominance effect ($\boldsymbol{d}$), sample size ($\boldsymbol{N}$), and effective population size ($\boldsymbol{N_e}$). The distribution of population allele frequencies was assumed to be U-shaped (Equation 2.13). The broad sense heritability and the additive gene effect ($\boldsymbol{a}$) were both equal to one.

2.4.3 Mean RMSE across allele frequency distribution

We have illustrated the RMSE of $\hat{\alpha}$ as a function of p. Now, we present the mean RMSE averaged over the distribution of p, for $a = 1$ and $H^2 = 1$, assuming a U-shaped distribution of p as a function of N_e, and for different values for N and d (Figure 2.4). For all scenarios, the mean RMSE with the A-model was about twice as large as the mean RMSE with the AD-model.

With both models, the mean RMSE was zero when d was zero (not shown) and increased as d increased. The mean RMSE decreased when N increased. The mean

RMSE decreased a little when N_e increased, which was caused by differences in the U-shaped distribution of allele frequencies. For example, when $N_e = 50$, the percentage of loci with an allele frequency outside the 0.05-0.95 range was 36%, whereas when $N_e = 200$, this percentage was 51%. Loci in this range have a low RMSE (see Figure 2.1 and Appendix A), and therefore a higher N_e results in a lower mean RMSE. The effect of N_e on the mean RMSE decreased as N increased. Results were identical when a was changed, because the mean RMSE scales linearly with the absolute dominance effect d, and not with the dominance coefficient d/a.

2.5 Discussion

We investigated the accuracy (in terms of RMSE) of estimated average effects ($\hat{\alpha}$) in the presence of dominance, using a single locus model including only an additive (A) or an additive plus dominance effect (AD). In the presence of dominance, the A-model falsely assumes that residuals are IID. The AD-model was therefore expected to better fit the data and give more accurate estimates of α, but only when dominance is present and sample size sufficient for dominance effect to be accurately estimated. Our results, however, show that the AD-model was always equally or more accurate than the A-model, even with small sample sizes (i.e. $N = 300$), a heritability lower than one (i.e. $H^2 < 1$), or in the absence of dominance.

With the A-model, both sampling deviations of genotype frequencies from HWE frequencies (ΔHWE) and sampling deviations of allele frequencies (Δp) contributed to the error. With the AD-model, only sampling deviations of allele frequencies contributed to the error, provided that all three genotype classes were sampled. The contribution of Δp to the error was much smaller than the contribution of ΔHWE. The AD-model was therefore more accurate than the A-model. Thus, even when the locus explained all variance (i.e., $H^2 = 1$), the mean RMSE decreased as sample size increased, because with larger sample sizes, deviations from HWE that considerably affect $\hat{\alpha}$ had a lower probability of occurring. Additionally, with larger sample sizes, the chance of missing one of the genotype classes was smaller, which further reduced the RMSE. The (mean) RMSE of $\hat{\alpha}$ was always smaller with the AD-model than with the A-model. The RMSE of $\hat{\alpha}$ scaled linearly with d; if d doubled, the RMSE also doubled. Remarkably, in the absence of dominance, there was no disadvantage of using the AD-model. Hence, the AD-model yielded equally or more accurate estimates of average effects than the A-model for all scenarios considered.

With the A-model, $\hat{\alpha}$ is computed as the linear regression coefficient of genotypic values on allele counts (Fisher 1941), which yields the average effect in

the sample (α_s), rather than the average effect in the whole population (α). Hence, the expectation of $\hat{\alpha}_A$ is equal to

$$E(\hat{\alpha}_A) = \alpha_s = a + (1 - 2p_s)d\left(\frac{1 - F_s}{1 + F_s}\right), \quad \textbf{2.14}$$

where F_s measures the deviation from HWE in the sample (ΔHWE) (Haldane 1954; Falconer 1985). Here, F_s is defined as one minus the ratio between the observed number of heterozygotes and the expected number of heterozygotes based on sample the allele frequency (Haldane 1954; Wright 1969). With the AD-model, $\hat{\alpha}$ is computed from $\hat{a}$ and $\hat{d}$, which are simultaneously estimated from the data. Unlike the A-model (where $E(\hat{\alpha}) = \alpha_s$), the expectation of $\hat{\alpha}_{AD}$ is equal to

$$E(\hat{\alpha}_{AD}) = a + (1 - 2p_s)d, \quad \textbf{2.15}$$

when all three genotype classes are sampled. Comparison of Equations 2.14 and 2.15 shows that the error in $\hat{\alpha}_A$ originates from both ΔHWE and Δp, while the error in $\hat{\alpha}_{AD}$ originates from Δp only, except when one of the genotypes is missing in the sample. When only two genotype classes are sampled, the AD-model reduces to the A-model. With the AD-model, the contrast between the mean genotypic value of the homozygotes and the genotypic value of the heterozygotes (d) does not depend on the number of individuals in these two groups. This is why the AD-model is more robust against deviations from HWE than the A-model.

These results were confirmed by mathematical derivations of the error with the two models (Appendix B). In theory, the error from the A-model can be quantified when p_s, p, F_s and d are known, and from the AD-model when p_s, p, and d are known. In real data, however, p (and also d with the A-model) is not known, and therefore the error cannot be quantified. As a result, the error cannot be removed from either of the two models. In conclusion, the AD-model is preferred for the estimation of average effects when dominance is present, because it yields more accurate estimates than the A-model, particularly when sample sizes are small.

In this study, we used the so-called genotypic parameterization of the AD-model, as opposed to the breeding parameterization (Vitezica *et al.* 2013). The results, however, were identical with the breeding parameterization (results not shown), because the two parameterizations are equivalent.

Additional to the contribution of dominance to additive variance, evidence for the contribution of epistasis is increasing (Mackay 2015; Monnahan and Kelly 2015). Our results show that modelling dominance improves estimated average effects, and it may therefore be tempting to hypothesize that modelling epistasis may also

improve estimates. However, investigating the benefit of modelling epistasis for the accuracy of $\hat{\alpha}$ is not straightforward, because it requires extension to multiple loci.

Taking a finite sample from a large population, which was done in this study, closely resembles a sharp reduction to a small population size, known as a bottleneck. In a small population, genotype frequencies deviate from HWE even under random mating. The expected genotype frequency for heterozygotes is equal to $2p_s(1-p_s)(1-F_s)$ (Haldane 1954). In turn, the expectation of F_s depends on the size of the bottleneck (or sample size, N), and is equal to $-\frac{1}{2N-1}$ with random mating (Kimura and Crow 1963). This indicates that the expected heterozygosity in the sample is larger than the HWE frequency calculated from the sample allele frequency. The effect of ΔHWE on estimated average effects was studied by Wang *et al.* (1998), who focused on the consequences for the additive genetic variance. In agreement with our results, they showed that the average effect was not influenced by ΔHWE when $d = 0$, or when $p \approx 0.5$. Furthermore, the effect of ΔHWE on estimated average effects depended on the size of the bottleneck (or sample size, N) and the size of dominance effect (d) (Wang *et al.* 1998). Because the effects of a bottleneck are very similar to the effects of taking a small sample from a large population, the results of our study also apply to populations in a bottleneck.

We quantified the error in estimates of α that originated from ΔHWE in random finite samples from a population of unrelated individuals. We purposefully used relatively small sample sizes to illustrate the effect. Although sample sizes taken in empirical studies may be larger, effective sample size may be much smaller, because actual populations often have small effective population size (N_e) (Hall 2016). This low N_e is related to the family structure in the population, where many individuals are bred from a limited number of parents, so that $N_e \ll N$. Hence, the effective sample size may be much smaller than N, because the sample will partly consist of related individuals. Because of this relatedness, sampling deviations in allele and genotype frequencies can be larger than expected based on sample size. The sample sizes chosen in this study may therefore be similar to effective sample sizes in empirical studies. As an example, we investigated the standard deviation of F_s across allele frequencies in a dataset of ~3500 pigs (Cleveland *et al.* 2012). The resulting value was comparable to the expected standard deviation of F_s for samples of 500 – 1000 animals (see Appendix C), which supports our expectation that effective number of sampled individuals may be smaller than the actual number of sampled individuals. Furthermore, in many studies that use genotype data, markers are removed if they show a significant deviation from HWE. The significance threshold that is used for HWE filtering, however, is often very liberal (Gondro *et al.*

2013). Consequently, there are still many markers left in the data that deviate from HWE and may give inaccurate estimates of average effects. As a result, we expect that the magnitude of ΔHWE simulated in this study may be similar to ΔHWE in empirical studies.

The estimation of average effects at single loci, as presented in this study, may be relevant for genome-wide association studies (GWAS). In GWAS, a large number of markers spread across the genome are each tested for an association with the observed phenotype (Gondro *et al.* 2013). Most GWAS test these associations by using an additive model which treats the marker genotypes as fixed (Hayes 2013). Only few studies have used the AD-model in GWAS to explicitly estimate a and d (e.g. Lopes *et al.* 2014; Aliloo *et al.* 2015; Huang *et al.* 2015; Bennewitz *et al.* 2017) and, to our knowledge, none have investigated differences in accuracy of estimated average effects between the A-model and AD-model. The effects of sampling genotypes on $\hat{\alpha}$ shown in this study apply to $\hat{\alpha}_m$ in GWAS, because $\hat{\alpha}_m$ are usually estimated by ordinary least squares. Using the AD-model in GWAS will therefore yield more accurate estimates of average effects and explained variance of markers.

The results presented in this study may also be relevant for genomic prediction. In genomic prediction, genomic estimated breeding values (GEBVs) are calculated as the sum of many estimated average effects multiplied by their marker genotypes (Meuwissen *et al.* 2001). Differences in accuracy of GEBVs may therefore be related to differences in accuracy of the estimated average effects. Our results, however, cannot be extrapolated directly to accuracy of GEBVs for several reasons. In this study, we considered a single locus, estimated α as a fixed effect, and assumed known genotypes of the quantitative trait locus (QTL). In contrast, GEBVs are based on many marker loci, for which all α's are estimated simultaneously as random effects (Meuwissen *et al.* 2001). In genomic prediction, the effect of a single QTL is likely to be explained by multiple markers, and errors of individual marker effects may cancel out to some extend when accumulated within individuals to compute their GEBVs. Additionally, random effect models shrink average effects towards zero (Whittaker *et al.* 2000), which may shrink the sampling error as well. In conclusion, to translate our results to accuracy of GEBVs, this research should be extended to the estimation of multiple random effects based on marker genotypes.

Neither GWAS nor genomic prediction are based on the genotypes at QTL directly, but rely on linkage disequilibrium (LD, measured by r) between observed markers and unknown QTL (Lewontin and Kojima 1960). For the additive effect at the QTL, the fraction captured by the marker is proportional to r, whereas for the dominance effect, the fraction captured by the marker is proportional to r^2 (Weir

2008; Zhu *et al.* 2015). The proportion of the signal of the dominance part of α_m that is captured is therefore expected to be smaller than of the additive part, because $r^2 \leq r$. For this reason, a marker should be very close to a QTL to pick up its dominance effect (Wellmann and Bennewitz 2012). As a result, the benefit of dominance models over additive models may be smaller with lower marker densities. We therefore argue that, when dominance is present and markers are able to capture dominance, the dominance model yields more accurate estimates of α than the additive model.

2.6 Conclusions

When a single locus average effect is estimated in a random finite sample from a large population in HWE-equilibrium, both additive (A) and additive plus dominance (AD) models yield error in their estimates, even when the locus explains all variance (i.e., $H^2 = 1$). Estimates from the AD-model, however, are more robust against chance deviations from HWE frequencies than estimates from the A-model. Genetic models that include dominance, therefore, yield higher accuracies of estimated average effects at single loci than purely additive models when dominance is present. In the absence of dominance, there was no penalty for fitting dominance. These results are important for GWAS, and potentially also for genomic prediction.

2.7 Acknowledgements

The authors thank Chris Maliepaard for his suggestions on an earlier draft of the paper. We thank the Netherlands Organisation of Scientific Research (NWO) and the Breed4Food consortium partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin for their financial support.

2.8 Appendices

Appendix A

With the A-model, error originated mainly from sampling deviations of genotype frequencies from expected HWE frequencies (ΔHWE), and to a lesser extent from sampling deviations of allele frequencies (Δp). For all N, the RMSE was smallest at intermediate allele frequencies, and increased when allele frequency moved away from 0.5 (Figure 2.1). This increase was due to the increasing probability of sampling less individuals from the rare genotype class than expected based on HWE. The RMSE decreased again at extreme allele frequencies. This decrease was due to (1) the increasing probability of not sampling the rare homozygous genotype, and (2) the change in α. When one of the homozygous genotypes is not sampled, $\hat{\alpha}$ was always equal to the slope of the regression line between the opposing homozygotes and the heterozygotes ($a + d$ when $n_2 = 0$, or $a - d$ when $n_0 = 0$). As the allele frequency approaches fixation, the probability of missing a homozygous genotype increases, while at the same time, the true α becomes close to the slope of the resulting regression line (see Equation 2.1) ($a + d$ when $p \to 0$, or $a - d$ when $p \to 1$). Hence, the RMSE as a result of missing one genotype class was small at extreme allele frequency and increased as the allele frequency increased. It is good to note that with only one genotype class sampled, the sample was disregarded because $\hat{\alpha}$ could not be estimated.

With the AD-model, error originated from Δp only, provided that all three genotype classes were sampled. For all N, the RMSE with the AD-model was always smaller than with the A-model, and showed a different pattern across allele frequencies (Figure 2.1). The RMSE decreased slightly when allele frequency moved away from 0.5, because the probability of drawing a sample with a very high Δp decreases. RMSE increased again for allele frequencies around 0.1 or 0.9. This increase was due to the increased probability of sampling no individuals from the rare genotype class, in which case the AD-model reduced to the A-model. The RMSE decreased again at even more extreme allele frequencies, for the same reasons as explained for the A-model.

Appendix B

Our aim is to estimate α in the entire population, but with the A-model, the estimate we get from the sample is equal to α_s. So, we can predict what the error will be by deriving the difference between α_s and α.

$$\begin{aligned} error &= E(\hat{\alpha}_A) - \alpha = \alpha_s - \alpha \\ &= \left[a + (1-2p_s)d\left(\frac{1-F_s}{1+F_s}\right)\right] - [a + (1-2p)d] \\ &= (1-2p_s)d\left(\frac{1-F_s}{1+F_s}\right) - (1-2p)d \\ &= \left(\frac{1-F_s}{1+F_s}(1-2p_s) - (1-2p)\right)d \end{aligned}$$

Then, if there is no deviation from HWE ($F = 0$), the error due to deviations of p_s from p is

$$[(1-2p_s) - (1-2p)]d = 2(p-p_s)d.$$

Similarly, if $p_s = p$, the error due to deviations from HWE is equal to

$$\left(\frac{1-F_s}{1+F_s} - 1\right)(1-2p_s)d.$$

When we correct for deviations from HWE, assuming that we know d, we get

$$\begin{aligned} \hat{\alpha}_c = \alpha_s - error &= \left[a + (1-2p_s)d\left(\frac{1-F_s}{1+F_s}\right)\right] - \left[\left(\frac{1-F_s}{1+F_s} - 1\right)(1-2p_s)d\right] \\ &= a + (1-2p_s)d\left(\frac{1-F_s}{1+F_s}\right) - (1-2p_s)d\left(\frac{1-F_s}{1+F_s}\right) + (1-2p_s)d \\ &= a + (1-2p_s)d. \end{aligned}$$

If we subsequently correct for deviations of p, assuming that we know p, we get

$$\begin{aligned} &[a + (1-2p_s)d] - [2(p-p_s)d] \\ &= [a + (1-2p_s)d] - [(1-2p_s) - (1-2p)]d \\ &= a + (1-2p)d, \end{aligned}$$

which is the average effect in the population.

For the AD-model, we can predict the error for samples that have all three genotype classes by

$$\begin{aligned} error &= E(\hat{\alpha}_{AD}) - \alpha = [a + (1-2p_s)d] - [a + (1-2p)d] \\ &= [(1-2p_s) - (1-2p)]d = 2(p-p_s)d, \end{aligned}$$

which is equal to the error due to deviations of p_s from p with the A-model.

Appendix C

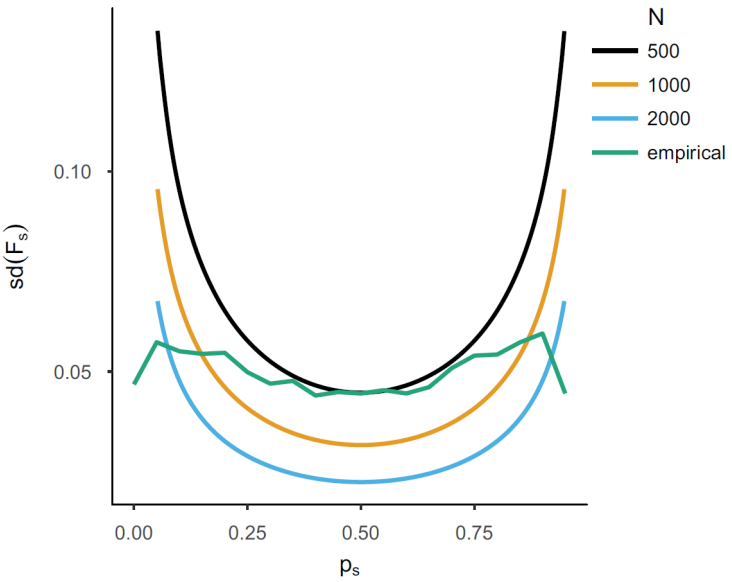

Figure S 2.1 Expected standard deviation of $\boldsymbol{F_s}$ for several sample sizes, compared to the standard deviation of $\boldsymbol{F_s}$ measured in an empirical dataset of pigs

3

The impact of non-additive effects on the genetic correlation between populations

Pascal Duenk[1], Piter Bijma[1], Mario P.L. Calus[1], Yvonne C.J. Wientjes[1]

and Julius H. J. van der Werf[2]

[1] Animal Breeding and Genomics, Wageningen University and Research,

6700 AH Wageningen, the Netherlands

[2] School of Environmental and Rural Science, University of New England,
Armidale 2351 NSW, Australia

Accepted in G3: Genes|Genomes|Genetics

Abstract

Average effects of alleles can show considerable differences between populations. The magnitude of these differences can be measured by the additive genetic correlation between populations (r_g). This r_g can be lower than one due to the presence of non-additive genetic effects together with differences in allele frequencies between populations. However, the relationship between the nature of non-additive effects, differences in allele frequencies, and the value of r_g remains unclear, and was therefore the focus of this study. We simulated genotype data of two populations that have diverged under drift only, or under drift and selection, and we simulated traits where the genetic model and magnitude of non-additive effects were varied. Results showed that larger differences in allele frequencies and larger non-additive effects resulted in lower values of r_g. In addition, we found that with epistasis, r_g decreases with an increase of the number of interactions per locus. For both dominance and epistasis, we found that, when non-additive effects became extremely large, r_g had a lower bound that was determined by the type of inter-allelic interaction, and the difference in allele frequencies between populations. Given that dominance variance is usually small, our results show that it is unlikely that true r_g values lower than 0.80 are due to dominance effects alone. With realistic levels of epistasis, r_g dropped as low as 0.45. These results may contribute to the understanding of differences in genetic expression of complex traits between populations, and may help in explaining the inefficiency of genomic trait prediction across populations.

3.1 Introduction

Populations can differ considerably in the average effects of loci (i.e. α, the difference between average effects of the two alleles, Falconer and Mackay (1996)). For a given genotype (i.e. individual), differences in α between two populations lead to differences in the additive genetic values of that genotype, as expressed in both populations. The magnitude of these differences can be measured by the additive genetic correlation between populations (r_g), defined as the correlation between the additive genetic values of a genotype expressed in population 1 and population 2. In reality, a single genotype cannot belong to two populations at the same time. This means that a trait expressed in two populations can be seen as a pair of traits that cannot be measured on the same individual, analogous to e.g. age at sexual maturity in males and females (Falconer and Mackay 1996). Although no phenotypic correlation exists between such pairs of traits, they can nevertheless be genetically correlated.

The r_g can be lower than one due to genotype by environment interaction (GxE) (Falconer 1952), or due to non-additive genetic effects (GxG-interaction) together with differences in allele frequencies between populations (Fisher 1918). Knowledge of this correlation contributes to the understanding of the genetic architectures of polygenic traits (de Candia *et al.* 2013; Brown *et al.* 2016). Such understanding may lead to improved knowledge of genetics and can facilitate accurate prediction of traits, such as disease risk in humans and yield traits in crops (Forsberg *et al.* 2017). Furthermore, understanding the genetic mechanisms that determine r_g may help in explaining the inefficiency of trait prediction across populations (Wientjes *et al.* 2015).

Following Falconer (1952), we can interpret a metric trait expressed in two populations as two different, genetically correlated traits. The additive genetic value of individual i for the trait expressed in the population that i belongs to (say, population 1) is

$$v_i^{P1} = \mathbf{h}'_{a,i}\boldsymbol{\alpha}^{P1}, \qquad \textbf{3.1}$$

where $\mathbf{h}_{a,i}$ is a column vector of additive genotypes (measured as allele counts, minus the mean allele count in the population) of individual i at quantitative trait loci (QTL), and $\boldsymbol{\alpha}^{P1}$ is a column vector of average effects at those QTL in population 1. The additive genetic value of individual i for another population (say, population 2) is

$$v_i^{P2} = \mathbf{h}'_{a,i}\boldsymbol{\alpha}^{P2}, \qquad \textbf{3.2}$$

where $\boldsymbol{\alpha}^{P2}$ is a column vector of average effects in population 2. Conceptually, this v_i^{P2} can be thought of as the additive genetic value for an individual in population 2 that has the same genotype as individual i. Here we define the additive genetic correlation between population 1 and population 2 (r_g) as the correlation between both additive genetic values for the individuals in population 1,

$$r_g = cor\left(v_i^{P1}, v_i^{P2}\right) = cor\left(\mathbf{h}'_{a,i}\boldsymbol{\alpha}^{P1}, \mathbf{h}'_{a,i}\boldsymbol{\alpha}^{P2}\right). \quad \textbf{3.3}$$

In other words, the r_g is defined for individuals coming from population 1, which may be different from the r_g defined for individuals coming from population 2 (See Discussion).

Equation 3.3 illustrates that the value of r_g depends on the differences in average effects between populations. With non-additive effects, average effects depend on the allele frequencies in the population, and, therefore, larger differences in allele frequencies between populations are expected to result in lower values of r_g.

Note that r_g is the correlation between the additive genetic values, not the genotypic values (i.e. additive plus non-additive genetic values). In the absence of GxE-interaction, the genotypic correlation between both populations is equal to one irrespective of the presence of GxG-interactions, because the genotypic value of a genotype (i.e. individual) is the same in both populations. The additive genetic correlation (r_g) may, however, be smaller than one because the partitioning of genotypic values into additive genetic values, dominance deviations and epistatic deviations depends on the allele frequencies (Fisher 1918; Cockerham 1954; Kempthorne 1954).

A deeper understanding of the relationship between non-additive genetic effects, allele frequencies and r_g may help geneticists to predict the value of r_g based on the importance of dominance and epistasis in the expression of the trait, and the genetic distance between populations. Wei *et al.* (1991) studied the impact of dominance on the additive genetic correlation between a purebred and crossbred population, known as r_{pc}. Using a two-locus model, they showed that r_{pc} indeed depends on both the magnitude of the dominance effect (d), and on the difference in allele frequencies between the populations. We are not aware of any theoretical studies that investigated the relationship between the importance of dominance and r_g between two purebred populations.

With epistasis, r_g is also expected to depend on the magnitude of epistatic effects and on the difference in allele frequencies between populations. Epistasis in

the functional (i.e. biological) sense means that the genotypic values of individuals depend on interactions between alleles or genotypes at different loci (Bateson and Mendel 1909), and there is substantial evidence for the existence of functional epistasis across species (Carlborg *et al.* 2003; Le Rouzic *et al.* 2008; Pettersson *et al.* 2011; Mackay 2015). Epistasis in the statistical sense is measured as the deviation of multi-locus genotypic values from the sum of the marginal effects (i.e. average and dominance effects) of the individual loci (Fisher 1918; Cockerham 1954). Although functional epistatic interactions do not necessarily lead to substantial statistical epistasis (Cheverud and Routman 1995; Hill *et al.* 2008; Maki-Tanila and Hill 2014), epistasis can contribute significantly to the additive genetic variance because average effects of individual loci may capture a substantial part of the functional epistasis (Hill *et al.* 2008; Maki-Tanila and Hill 2014; Monnahan and Kelly 2015). Furthermore, epistatic variance may be 'converted' into additive genetic variance due to genetic drift or due to selection (Cheverud and Routman 1996; Hill 2017). Thus, epistatic interactions modify average effects of individual loci when allele frequencies change, and may therefore play an important role in the value of r_g and its change over time.

In summary, the r_g between populations is affected by non-additive effects in combination with differences in allele frequencies between populations. For populations in the same environment (i.e. in the absence of GxE), r_g is equal to 1 in the absence of non-additive effects or in the absence of allele frequency differences. So far, the relationship between the nature and magnitude of non-additive effects, differences in allele frequencies, and the value of r_g remains unclear. Our objective was therefore to investigate the impact of non-additive effects on r_g for populations that have diverged either under drift only, or under both drift and selection.

3.2 Methods

We aimed to investigate the relationship between non-additive effects and the additive genetic correlation between populations (r_g) with small effective size, as observed in livestock. For this purpose, we simulated genotypes of quantitative trait loci (QTL) for two populations that have diverged for a number of generations under either pure drift, or under drift and selection. The populations were assumed to be kept in the same environment, so there was no GxE. We simulated traits following several scenarios that differed in the type (i.e. genetic model) and the magnitude of non-additive effects (Table 3.1).

Table 3.1 Overview of scenarios with their parameters for distributions of sampled dominance coefficients and epistatic coefficients.

	Small	Intermediate	Large
D	$\mu_\delta = 0.2, \sigma_\delta = 0.30$	$\mu_\delta = 0.2, \sigma_\delta = 0.70$	$\mu_\delta = 0.2, \sigma_\delta = 1.50$
E_{AA}	$\mu_\gamma = 0.0, \sigma_\gamma = 0.16$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.33$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.68$
E_{DD}	$\mu_\gamma = 0.0, \sigma_\gamma = 0.16$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.33$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.68$
E_C	$\mu_\gamma = 0.0, \sigma_\gamma = 0.16$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.33$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.68$
E_M	$\mu_\gamma = 0.0, \sigma_\gamma = 0.16$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.33$	$\mu_\gamma = 0.0, \sigma_\gamma = 0.68$

We considered six genetic models; a basic model with additive effects only (A), which served as a basis for comparison, and five alternative models with non-additive effects: one with only dominance effects (D), and four with only epistatic effects. With epistasis, we simulated interactions between pairs of loci that followed one of the configurations presented in Figure 3.1. We chose these genetic models so that there were scenarios with only dominance variance (D), scenarios with only additive by additive epistatic variance (E_{AA} and E_M), and scenarios with all types of non-additive variance (E_C and E_{DD}). For each genetic model, we considered three magnitudes of non-additive effects, labelled as small, intermediate, and large.

a. Additive x additive (E_{AA})

genotype locus *y* \ genotype locus *l*	YY	Yy	yy
WW	2	1	0
Ww	1	1	1
ww	0	1	2

b. Dominance x dominance (E_{DD})

genotype locus *y* \ genotype locus *l*	YY	Yy	yy
WW	0	1	0
Ww	1	0	1
ww	0	1	0

c. Complementary (E_C)

genotype locus *y* \ genotype locus *l*	YY	Yy	yy
WW	1	1	0
Ww	1	1	0
ww	0	0	0

d. Multiplicative (E_M)

genotype locus *y* \ genotype locus *l*	YY	Yy	yy
WW	4	2	0
Ww	2	1	0
ww	0	0	0

Figure 3.1 Epistatic contrasts for four biological epistatic configurations.

3.2.1 Simulation

We simulated genotypes of two livestock populations (1 and 2) that diverged for 50 generations (Figure 3.2). For divergence, we considered two situations: one where the populations diverged due to drift only, and one where the populations diverged also due to selection in population 1 and drift in population 2.

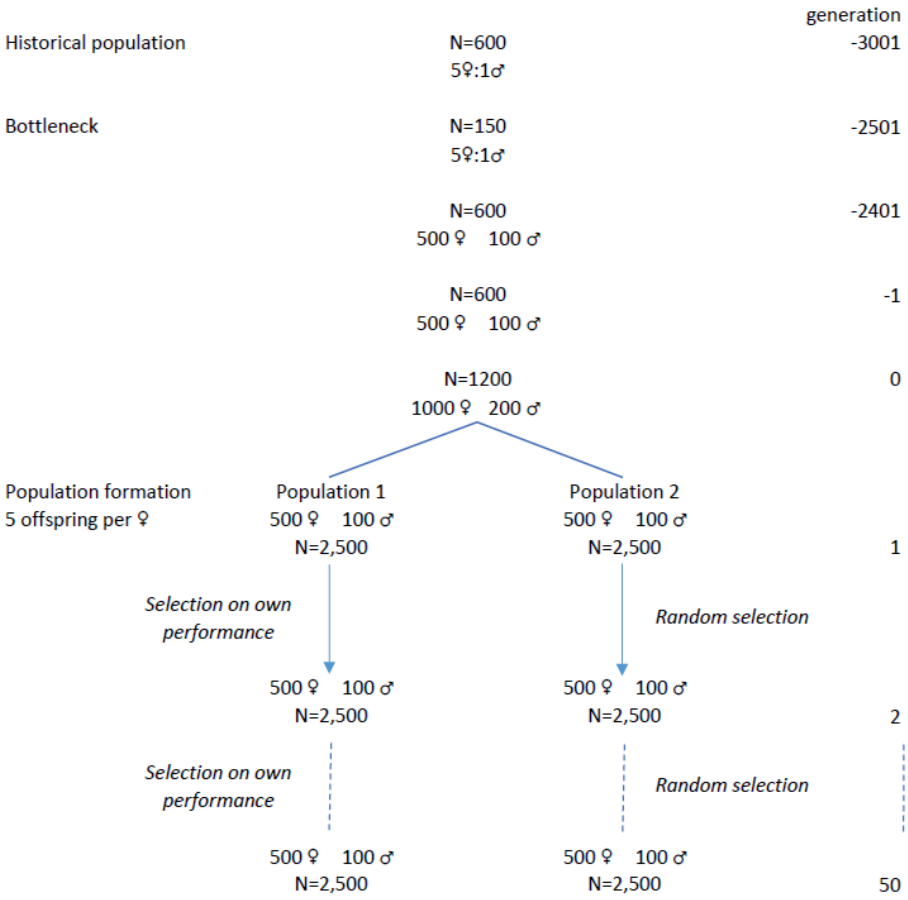

Figure 3.2 Overview of the simulated population structure

3.2.2 Populations

We simulated a historical population with QMSim (Schenkel and Sargolzaei 2009) by randomly mating 100 males and 500 females starting in generation -3001. From generation -3000 to generation -2501, we simulated a bottleneck by gradually decreasing population size to 150 (25 males and 125 females) to create initial linkage disequilibrium (LD), and population size was gradually increased again to 600 during the next 100 generations. The population size remained constant from generation -2400 until -1 to allow for the development of mutation-drift equilibrium. To provide a sufficient number of individuals for the development of populations 1 and 2, we doubled the number of individuals in the last historical generation (generation 0) to 200 males and 1,000 females. This simulation resulted in an average effective population size (N_e) of ~285 generation 0, calculated as the harmonic mean of

$\frac{4N_mN_f}{N_m+N_f}$ in each preceding generation, where N_m is the number of males and N_f is the number of females that become parents in a generation (Falconer and Mackay 1996).

After simulating the historical population, we simulated two current populations (1 and 2). We randomly sampled 100 males and 500 females from the last historical generation to become founders of population 1. The remaining 100 males and 500 females were the founders of population 2. We will refer to the generation of founders as generation 0. Within each population, simulation continued for 50 generations by randomly mating 100 selected males with 500 selected females. Each mating resulted in 5 offspring, resulting in a total of 2,500 offspring (exactly 1,250 males and 1,250 females) in each generation. Generations were non-overlapping, meaning that in each generation, the parents were selected from the previous generation only. In the drift scenario, animals in both populations were randomly selected to become parents of the next generation. Effective population size (N_e) in the drift scenario was ~285 in the two populations. In the selection-drift scenario, animals in population 1 were selected based on their own phenotype (mass selection), while in population 2, selection was random. In this scenario, effective population size (N_e) in population 1 was ~250, which was calculated as $1/(2\Delta F)$, where ΔF is the inbreeding rate estimated from the pedigree (Falconer and Mackay 1996). We simulated selection only in population 1 to reduce computation time.

3.2.3 Genome

The simulated genome consisted of 10 chromosomes of 1 Morgan that each had 200 randomly positioned bi-allelic loci. In the first historical generation (generation -3001), we randomly sampled the allele frequencies of loci from a uniform distribution. Mutation rate was $2.5*10^{-5}$ during the historical generations. In generation 0, the distribution of allele frequencies had evolved to a U-shape, and we randomly selected 500 segregating loci to become QTL, which resulted in low linkage disequilibrium between QTL. There was no mutation from generation 0 to 50, because the QMSim software does not allow for mutation after the last historical generation.

3.2.4 Functional genetic effects

Additive effects (a) of all 500 QTL were sampled from $\sim N(0,1)$. We assumed that the size of the dominance and epistatic effects were proportional to the additive effects of the QTL involved in the interaction (Wellmann and Bennewitz 2011). We

therefore sampled dominance coefficients (δ) for all QTL from $\sim N(\mu_\delta, \sigma_\delta^2)$, from which dominance effects (d) were computed as $\delta|a|$. Similarly, we sampled epistatic coefficients (γ) for all pairwise epistatic interactions from $\sim N(\mu_\gamma, \sigma_\gamma^2)$, from which functional epistatic effects (ϵ) were computed as $\gamma_{kl}\sqrt{|a_k a_l|}$ where k and l denote the QTL involved in the interaction. Each QTL had an epistatic interaction with 5 randomly sampled other QTL, resulting in a total of 1250 pairwise interactions.

For both dominance and epistasis, we considered 3 magnitudes of effects: small, intermediate, and large. For all magnitudes, the mean dominance coefficient (μ_δ) was 0.2, and the mean epistatic coefficient (μ_γ) was 0.0. The magnitude of dominance and epistatic effects were controlled by changing the standard deviation of dominance coefficients (σ_δ) or epistatic coefficients (σ_γ). For dominance, σ_δ was 0.3 with small effects, 0.7 with intermediate effects, and 1.5 with large effects (Table 3.1). The mean and standard deviation of small dominance coefficients were chosen based on empirical results of Bennewitz and Meuwissen (2010) and Sun and Mumm (2016). For epistasis, σ_γ was scaled such that the total functional epistatic variance was comparable to the total functional dominance variance in the scenario with the same magnitude. To this end, σ_γ was computed as $\sqrt{(\mu_\delta^2 + \sigma_\delta^2)/N_\gamma}$, where N_γ is the number of epistatic interactions per QTL, and μ_δ^2 and σ_δ^2 are the squared mean and variance of dominance effects in the scenario with the corresponding magnitude. For example, with small epistatic effects, σ_γ was computed as $\sqrt{(0.2^2 + 0.3^2)/5} \approx 0.16$ (Table 3.1).

3.2.5 From functional dominance and epistatic effects to statistically orthogonal effects

We simulated dominance and epistasis by introducing functional dominance and epistatic effects that are independent of allele and genotype frequencies. Our interest, however, is in statistical average, dominance and epistatic effects of QTL, which do depend on genotype frequencies (Fisher 1918; Cheverud and Routman 1995). We will describe the general procedure to obtain these statistical effects, for a general situation where there can be dominance, epistasis, or both. Note, however, that our scenarios had either dominance or epistasis, but never both. After obtaining statistical effects, we describe how we computed the additive genetic value, genotypic value and phenotype for each individual. Although genotypic values themselves are independent of genotype frequencies, the partitioning of these genotypic values into additive, dominance, and epistatic components does depend on genotype frequencies. Additive genetic values of individuals in population 1 were

needed to compute r_g, and genotypic values and phenotypes were needed because selection in population 1 was based on own performance. In the following, we describe the procedure to obtain the average effects and dominance effects in population 1 ($\boldsymbol{\alpha}^{P1}$). The procedure to obtain these effects in population 2 ($\boldsymbol{\alpha}^{P2}$) follows naturally by replacing the genotype and allele frequencies of population 1 with the frequencies in population 2.

The procedure starts by applying the natural and orthogonal interactions (NOIA) model (Álvarez-Castro and Carlborg 2007) for each epistatic interaction between two QTL. First, functional epistatic values for the 9 possible two-locus genotypes at QTL k and l were collected in a vector $\mathbf{c}_{kl} = \mathbf{t}\epsilon_{kl}$, where ϵ_{kl} is a scalar representing the functional epistatic effect between QTL k and l, and $\mathbf{t}$ is a 9 x 1 vector of epistatic contrasts for the 9 two-locus genotypes, ordered as (WWYY, WwYY, wwYY, WWYy, …, wwyy). The simulated epistatic contrasts in $\mathbf{t}$ followed one of four configurations: additive x additive (E_{AA}), dominance x dominance (E_{DD}), complementary (E_C), or multiplicative (E_M) (Figure 3.1). The contrasts in $\mathbf{t}$ were centred and scaled to a standard deviation of one, so that the contrasts were comparable between configurations. We then used genotype frequencies of QTL k and l to partition the functional epistatic values in $\mathbf{c}_{kl}$ into 9 statistical genetic effects (Álvarez-Castro and Carlborg 2007; Vitezica *et al.* 2017)

$$\mathbf{b}_{kl} = (\mathbf{W}_{kl}'\mathbf{D}_{kl}\mathbf{W}_{kl})^{-\mathbf{1}}\mathbf{W}_{kl}'\mathbf{D}_{kl}\mathbf{c}_{kl}, \qquad \textbf{3.4}$$

where $\mathbf{D}_{kl}$ is a 9x9 diagonal matrix with each of the nine genotype frequencies in the same order as in $\mathbf{t}$. Matrix $\mathbf{W}_{kl} = \mathbf{W}_k \otimes \mathbf{W}_l$, where $\otimes$ denotes the Kronecker product, and $\mathbf{W}_k$ and $\mathbf{W}_l$ are constructed as

$$\mathbf{W}_x = [\mathbf{1} \quad \mathbf{w}_a \quad \mathbf{w}_d] = \begin{bmatrix} 1 & -(-p_{Xx} - 2p_{xx}) & \dfrac{2p_{Xx}p_{xx}}{p_{XX} + p_{xx} - (p_{XX} - p_{xx})^2} \\ 1 & -(1 - p_{Xx} - 2p_{xx}) & \dfrac{4p_{Ww}p_{ww}}{p_{XX} + p_{xx} - (p_{XX} - p_{xx})^2} \\ 1 & -(2 - p_{Xx} - 2p_{xx}) & \dfrac{2p_{Ww}p_{Ww}}{p_{XX} + p_{xx} - (p_{XX} - p_{xx})^2} \end{bmatrix}, \qquad \textbf{3.5}$$

where columns relate to orthogonal contrasts for the mean ($\mathbf{1}$), average effect ($\mathbf{w}_a$), and dominance effect ($\mathbf{w}_d$) of QTL x, and where p_{Xx}, p_{Xx}, and p_{xx} are the genotype frequencies of QTL X. The resulting vector of statistical genetic effects is

$$\mathbf{b}_{kl} = \left[\mu, \alpha_{kl}^k, d_{kl}^k, \alpha_{kl}^l, (\alpha\alpha)_{kl}, (d\alpha)_{kl}, d_{kl}^l, (\alpha d)_{kl}, (dd)_{kl}\right]', \qquad \textbf{3.6}$$

where α_{kl}^k and α_{kl}^l are the terms that contribute to average effects of QTL k and l. The other terms in $\mathbf{b}_{kl}$ contribute to dominance effects (d_{kl}^k, d_{kl}^l) of individual QTL and to epistatic effects of interacting QTL ($(\alpha\alpha)_{kl}, (d\alpha)_{kl}, (\alpha d)_{kl}, (dd)_{kl}$).

We repeated this procedure of partitioning functional epistatic effects into statistical genetic effects for all pairwise interactions between QTL. Each QTL was involved in 5 epistatic interactions and therefore has 5 terms that contribute to its average effect. Following this reasoning, the average effect of QTL k in population 1 with epistasis is

$$\alpha_k^{P1} = a_k + \left(1 - 2p_k^{P1}\right)d_k + \sum_{l \in \mathbb{Z}}^{N_\gamma} \alpha_{kl}^k, \quad \textbf{3.7}$$

where p_k^{P1} is the frequency of the counted allele of QTL k in population 1, $\mathbb{Z}$ is the set of loci that QTL k interacts with, and $N_\gamma = 5$. Note the difference between "additive effect" (a) and "average effect" (α); the additive effect a is half the difference in genotypic value between both opposing homozygotes, whereas the average effect (α) is the (statistical) marginal effect of the QTL. Throughout this manuscript, we will use the term "functional additive effect" to refer to a, and "average effect" (i.e. statistical substitution effect) to refer to α.

In our simulations, we needed to compute phenotypes of selection candidates in each generation, for which we needed the statistical dominance effect (d^*) of each QTL as well. The dominance effect of QTL k in population 1 with epistasis is

$$d_k^{P1^*} = d_k + \sum_{l \in \mathbb{Z}}^{N_\gamma} d_{kl}^k \quad \textbf{3.8}$$

3.2.6 Additive genetic values and phenotypes

We computed additive genetic values ($\mathbf{v}$) of selection candidates in population 1 for the trait expressed in both population 1 and 2. Their genotypic values ($\mathbf{g}$) and phenotypes were only computed for the trait expressed in population 1. The additive genetic value of individual i for the trait expressed in population 1 (2) were computed as $v_i^{P1} = \mathbf{h}'_{a,i}\boldsymbol{\alpha}^{P1}$ ($v_i^{P2} = \mathbf{h}'_{a,i}\boldsymbol{\alpha}^{P2}$), and genotypic values for the trait expressed in population 1 were computed as

$$\begin{aligned} g_i^{P1} = v_i^{P1} + \mathbf{h}'_{d,i}\mathbf{d}^{P1^*} + \mathbf{h}'_{a,i}\otimes\mathbf{h}'_{a,i}(\boldsymbol{\alpha}\boldsymbol{\alpha})^{P1} + \mathbf{h}'_{a,i}\otimes\mathbf{h}'_{d,i}(\boldsymbol{\alpha}\mathbf{d})^{P1} + \mathbf{h}'_{d,i}\otimes\mathbf{h}'_{a,i}(\mathbf{d}\boldsymbol{\alpha})^{P1} \\ + \mathbf{h}'_{d,i}\otimes\mathbf{h}'_{d,i}(\mathbf{d}\mathbf{d})^{P1}, \end{aligned} \quad \textbf{3.9}$$

where $\mathbf{h}_{a,i}$ is a column vector of additive genotype indicators for individual i, and $\mathbf{h}_{d,i}$ is column a vector of dominance genotype indicators for individual i. These indicators were coded following the NOIA parameterization as denoted in the rows of $\mathbf{w}_a$ and $\mathbf{w}_d$ (Equation 3.3) for genotypes XX, Xx, and xx, respectively. Phenotypes with a broad sense heritability of 0.5 were computed as $\mathbf{y}^{P1} = \mathbf{g}^{P1} + \mathbf{e}^{P1}$, where $\mathbf{e}^{P1} \sim N(0, \sigma_e^2)$, and σ_e^2 was equal to the variance of genotypic values (σ_g^2).

3.2.7 Computing parameters of interest

The parameters of interest were (1) the genetic correlation between the trait in population 1 and the trait in population 2 (r_g), and (2) the average absolute difference in allele frequencies between populations ($\overline{\Delta p}$). For each generation, we computed r_g as the Pearson correlation between the additive genetic values of individuals in population 1 for the trait expressed in the two populations (Equation 3.3)[1]. For each generation, we computed $\overline{\Delta p}$ as $\sum(|p_k^A - p_k^B|)\ /500$. We chose this parameter as a measure for population divergence, because we expect that there is a linear relationship between $\overline{\Delta p}$ and r_g. These parameters were computed for generation 1 to 5, and for every 5th generation after generation 5, to limit computation time.

3.2.8 Replicates

We ran the simulation with drift 50 times, resulting in 50 sets of genotypes (i.e. replicates). For each of those replicates, we computed $\overline{\Delta p}$ and r_g for each of the scenarios (i.e. genetic model and magnitude). We ran the simulations with both selection and drift for each scenario separately, because the selection of parents in population 1 depended on the genetic model. To limit computation time, we used 20 replicates for each scenario with selection.

3.2.9 Data availability

The data used in this study can be reproduced with the files and seeds in the following GitHub repository: https://git.wageningenur.nl/duenk002/rg-and-non-additive-effects

3.3 Results

First, for each scenario with selection, we show the change in mean genotypic value ($\bar{g}$) and the change of additive genetic variance (V_A) in population 1 across generations, to illustrate how population 1 evolved over time. Second, we report realized fractions of additive, dominance and epistatic variance in generation 1 and 50. Third, for scenarios with small non-additive effects, we show the effects of the genetic model and of applying selection on the additive genetic correlation (r_g) and

[1] Effectively, this r_g is a weighted correlation between $\boldsymbol{\alpha}^{P1}$ and $\boldsymbol{\alpha}^{P2}$, where the weights depend on the allele frequencies in population 1. Hence, the r_g computed as the correlation of additive genetic values of individuals in population 2 may give different results because the genotypes sampled from population 2 result in different weights than those sampled from population 1 (see Discussion).

the difference in allele frequency ($\overline{\Delta p}$) between populations. Fourth, for each genetic model with selection, we investigate the impact of the magnitude of non-additive effects and the number of generations since divergence. Finally, we investigate the relationship between r_g and $\overline{\Delta p}$ across genetic models and within genetic models. All results presented refer to generation 50 and to scenarios with small non-additive effects, unless otherwise stated.

3.3.1 Mean genotypic value and variance components

With all scenarios, the mean genotypic value expressed in genetic standard deviations ($\bar{g}$) in population 1 increased due to selection (Figure S 3.1). With all genetic models, the increase in $\bar{g}$ was smaller when the magnitude of non-additive effects was larger. This result was expected, because the marginal effects of alleles may change over time in the presence of non-additive effects, reducing the effectiveness of selection. The increase in $\bar{g}$ was largest with model A, and it was smallest with model E_{DD} and large non-additive effects. There were only small differences in $\bar{g}$ between models D, E_{AA}, E_C, and E_M.

The additive genetic variance in population 1 (V_A) decreased due to selection with all scenarios (Table 3.2 and Figure S 3.2). With genetic model A, E_{AA}, E_C and E_M, about 95-98% of V_A was lost after 50 generations of selection, whereas with D and E_{DD}, 88-95% of V_A was lost. A change in magnitude of non-additive effects did not substantially affect the decrease in V_A, except with genetic models D and E_{DD}, where more additive genetic variance was preserved with larger non-additive effects. In the drift scenario, the average loss of V_A was about 7% for all scenarios (results not shown).

In generation 1, scenarios that had only additive genetic (V_A) and epistatic variance (V_I), V_A accounted for the largest, and V_I for the smallest fraction of the total genetic variation (Table 3.2). The largest fraction of V_I was realized with genetic model E_{AA} (max. 0.048), followed by E_{DD} (max. 0.033), E_C (max. 0.024) and E_M (max. 0.017). The largest fraction of dominance variance (V_D) was realized with model E_{DD} (max. 0.364), followed by D (max. 0.298) and E_C (max. 0.105). With genetic models D, E_{DD} and E_C, the fraction V_D increased and V_A decreased across generations, especially with intermediate or large effects (Table 3.2, generation 50). The fraction V_I remained relatively constant across generations with all scenarios.

Table 3.2 Fractions of additive (V_A), dominance (V_D), and epistatic (V_I) variances with respect to the total genetic variance in generation 1 and generation 50 with selection. Reported values are averages of 20 replicates.

		Generation 1			Generation 50		
	Effect size	V_A	V_D	V_I	V_A	V_D	V_I
D	Small	0.961	0.039	0.000	0.924	0.076	0.000
	Intermediate	0.871	0.129	0.000	0.511	0.489	0.000
	Large	0.702	0.298	0.000	0.198	0.802	0.000
E_{AA}	Small	0.992	0.000	0.008	0.997	0.000	0.003
	Intermediate	0.976	0.000	0.024	0.988	0.000	0.012
	Large	0.952	0.000	0.048	0.969	0.000	0.031
E_{DD}	Small	0.910	0.064	0.026	0.703	0.289	0.008
	Intermediate	0.751	0.173	0.076	0.358	0.602	0.040
	Large	0.528	0.333	0.139	0.146	0.752	0.101
E_C	Small	0.985	0.012	0.003	0.947	0.051	0.001
	Intermediate	0.947	0.044	0.009	0.737	0.250	0.013
	Large	0.871	0.105	0.024	0.471	0.511	0.017
E_M	Small	0.998	0.000	0.002	0.999	0.000	0.001
	Intermediate	0.993	0.000	0.007	0.995	0.000	0.005
	Large	0.983	0.000	0.017	0.989	0.000	0.011

3.3.2 Effect of genetic model and of selection on r_g

For all genetic models and small non-additive effects, r_g was lower with selection than with drift only (Figure 3.3). With drift only, r_g was between 0.99 and 1 for all genetic models. After 50 generations of selection, average r_g was lowest with genetic model E_{DD} (0.65), followed by E_{AA} (0.75), D (0.83), E_C (0.83) and finally E_M (0.94). There was a tendency that scenarios with the largest non-additive variance in generation 1 had the smallest r_g in generation 50 (Figure S 3.3). Note that the r_g was always equal to 1 with the additive model (A) (results not shown).

As expected, $\overline{\Delta p}$ was larger with selection than with drift, and was the same across all genetic models with drift (0.05; Figure 3.4). With selection, $\overline{\Delta p}$ with non-additive models was very similar (around 0.20) to the value with an additive model.

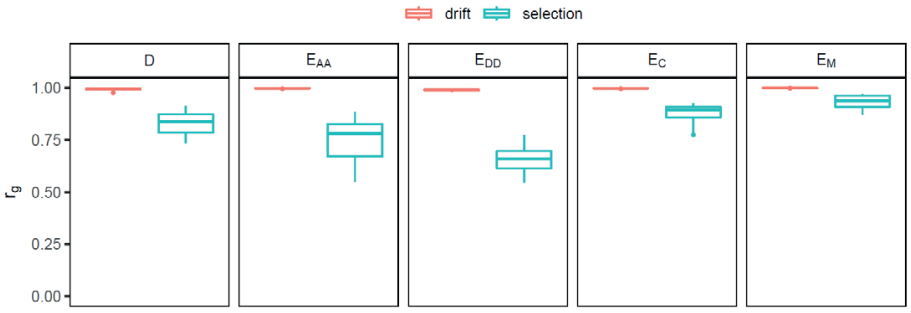

Figure 3.3 Effect of genetic model on r_g with small non-additive effects, under drift only, or under drift and selection.

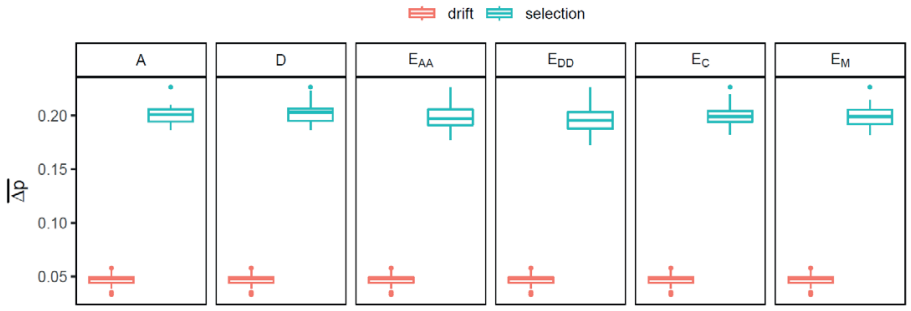

Figure 3.4 Effect of genetic model on the difference in allele frequencies between populations, under drift only, or under drift and selection.

3.3.3 Effect of the magnitude of non-additive effects

For all genetic models and with selection, r_g decreased with increasing magnitude of non-additive effects (Figure 3.5). With genetic model D, r_g dropped about 31% from small to intermediate, and about 27% from intermediate to large dominance effects. With all epistatic models, the drop in r_g with increasing magnitude was smaller (16-23%) than with D.

For all genetic models with selection, the average absolute difference in allele frequency between lines ($\overline{\Delta p}$) decreased with increasing magnitude of non-additive effects, especially with D and E_{DD} (Figure 3.6). With model D, $\overline{\Delta p}$ was 0.18 with intermediate dominance effects, and 0.141 with large effects. With E_{DD}, $\overline{\Delta p}$ was 0.162 with intermediate epistatic effects, and 0.130 with large effects. With the other epistatic models (E_{AA}, E_C and E_M), the effect of an increase in magnitude on $\overline{\Delta p}$ was much smaller (~0.19 with intermediate and ~0.18 with large effects).

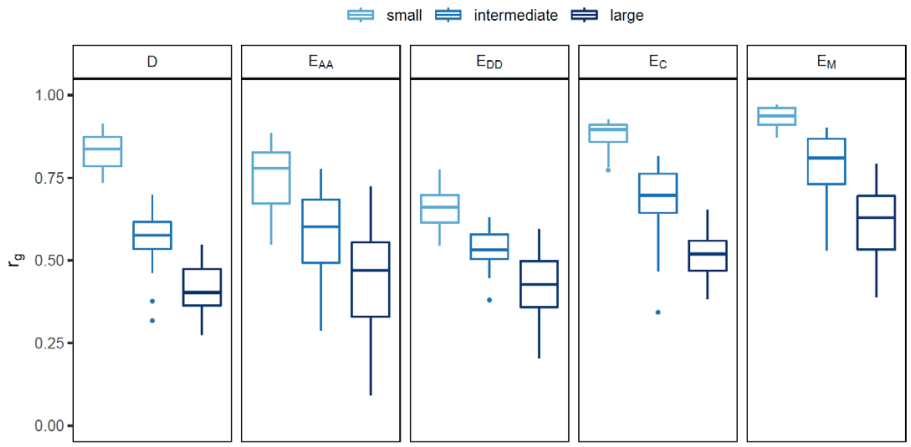

Figure 3.5 Effect of magnitude of non-additive effects on r_g, where population 1 was selected and population 2 was not selected.

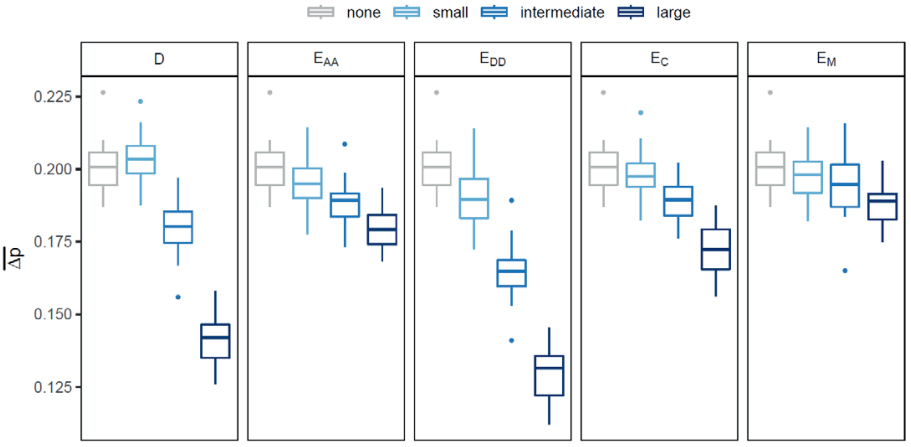

Figure 3.6 Effect of magnitude of non-additive effects on the difference in allele frequencies between populations under selection. In each subplot, the additive scenario (A) was included for reference (i.e. magnitude "none").

3.3.4 Effect of number of generations since divergence

With all scenarios, r_g decreased with the number of generations since divergence, and the rate of decrease was relatively small during the first five generations ($r_g > 0.94$), especially when the non-additive effects were small ($r_g > 0.98$) (Figure 3.7). After the first five generations, the rate of decrease in r_g differed across genetic models. There was a considerable difference between genetic models, the E_M model showed the smallest decline of r_g over time, and the E_{DD} model showed the largest decline. With large non-additive effects, models E_M and E_{AA}

tended to show an accelerated decrease in r_g across generations, whereas models D, E_C and E_{DD} tended to show a decelerated decrease in r_g (Figure 3.7).

With all scenarios, the average absolute difference in allele frequency between lines ($\overline{\Delta p}$) increased with the number of generations since divergence (Figure 3.8). In contrast to the result of the genetic correlation with small non-additive effects (Figure 3.7), $\overline{\Delta p}$ was remarkably similar between the genetic models (Figure 3.8). With large effects, models D and E_{DD} showed a smaller $\overline{\Delta p}$ than models E_M, E_{AA}, and E_C.

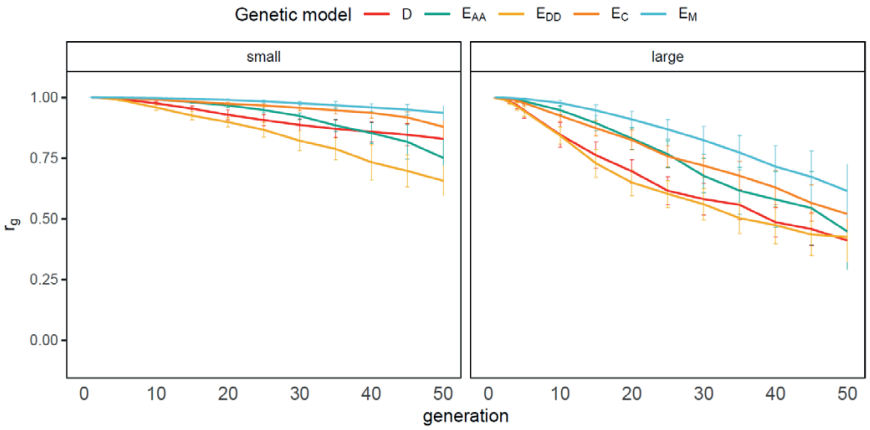

Figure 3.7 Effect of number of generations since divergence on r_g for all genetic models with small (left) or large (right) non-additive effects.

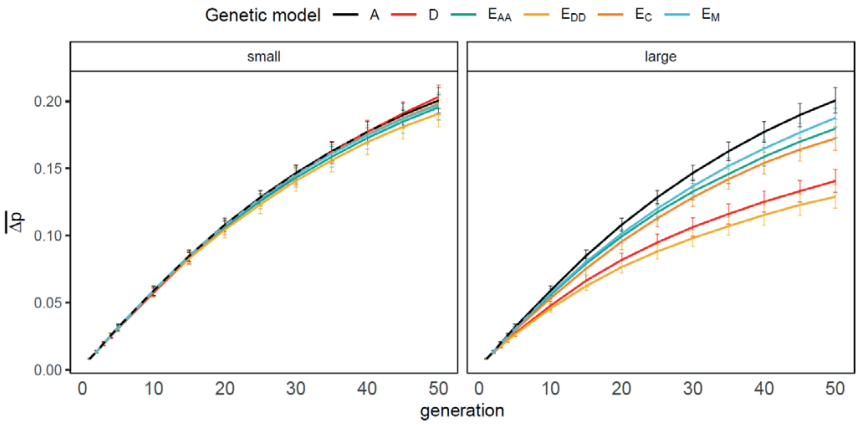

Figure 3.8 Effect of number of generations since divergence on the difference in allele frequencies between populations, for all genetic models and small (left) or large (right) non-additive effects.

In summary, for each genetic model, r_g was smallest with selection, large non-additive effects and many generations since divergence. Overall, the smallest realized value of r_g after 50 generations of divergence was achieved with genetic model D or E_{DD} ($r_g \approx 0.41$ for both).

3.3.5 Relationship between r_g and $\overline{\Delta p}$

For all genetic models, there was a clear negative relationship between $\overline{\Delta p}$ and r_g (Figure 3.9), and the relationship was strongest for genetic models showing the strongest decline of r_g with time (Figure 3.7). This result suggests that differences between genetic models in the decline of r_g over time originate from different impacts of $\overline{\Delta p}$ on r_g, and not from differences in $\overline{\Delta p}$ per se. For example, with small non-additive effects and after 50 generations of divergence, the value of r_g was different between genetic models, whereas the realized $\overline{\Delta p}$ was very similar (Figure 3.9). In other words, r_g is a function of $\overline{\Delta p}$ and of genetic architecture.

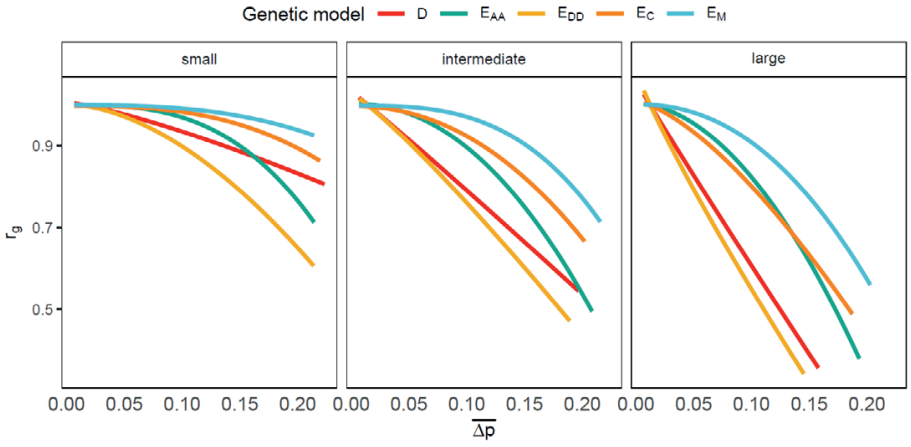

Figure 3.9 Relationship between r_g and the difference in allele frequencies between populations for all scenarios.

3.4 Discussion

Our objective was to investigate the relationship between non-additive effects, differences in allele frequencies between populations ($\overline{\Delta p}$), and the genetic correlation between populations (r_g). We simulated genotype data of two populations that have diverged for a number of generations under drift only, or drift and selection, and we simulated traits where the genetic model and magnitude of non-additive effects were varied.

We computed r_g as the correlation between additive genetic values of individuals in population 1, for the trait expressed in population 1 and 2. Effectively, this r_g is a weighted correlation between average effects in population 1 ($\boldsymbol{\alpha}^1$) and 2 ($\boldsymbol{\alpha}^2$), where the weights depend on the sample of genotypes that were used to compute the additive genetic values. This suggests that different values of r_g could have been obtained when using the additive genetic values of individuals in population 2, because of differences in genotype frequencies between populations. We chose, however, to focus on population 1 because we were also interested in the change of allele frequencies over time due to selection. This approach leads to values of r_g that indicate whether information from an unselected population (population 2) can be used to predict additive genetic values in a selected population (population 1).

3.4.1 Realized variance components

Because little is known about the quantity and magnitude of dominance and epistatic effects in reality, we considered a range of functional non-additive effect sizes and epistatic configurations. Realized proportions in our simulations (Table 3.2) did not always match with those observed in real data. For example, with large dominance effects, the fraction of dominance variance was 30%, which is uncommon in real data (Ertl *et al.* 2014; Lopes *et al.* 2016; Moghaddar and van der Werf 2017; Joshi *et al.* 2018). Similarly, scenario E_{DD} also resulted in more dominance variance than expected in real populations, especially with large epistatic effects (33%). Empirical studies on livestock (Bennewitz and Meuwissen 2010) and crops (Sun and Mumm 2016) found that approximately 0.3% of loci show overdominance, which is comparable to our scenario with small dominance effects (0.5% overdominance). Furthermore, the scenario with small dominance effects resulted in a small proportion of dominance variance, and might therefore be most realistic for actual populations.

In contrast to our realized proportions of dominance variance, proportions of epistatic variance were lower (max. 5%) than estimates from an empirical study on litter size in pigs (about 26%) (Vitezica *et al.* 2018), though the standard error of that estimate was large (about 22%). Further evidence of statistical epistatic effects is scarce, probably because methods used for the detection of statistical epistasis are frequently underpowered (Wei *et al.* 2014). Furthermore, it has been suggested that incomplete LD between genomic markers and QTL may create the illusion of epistasis, making inference about the importance of epistasis from genome-wide regression studies difficult (Wei *et al.* 2014; Zan *et al.* 2018; de los Campos *et al.* 2019). In contrast to the lack of evidence of statistical epistasis, there is substantial

evidence that physiological epistasis is abundant in several classes of organisms (Carlborg *et al.* 2003; Le Rouzic *et al.* 2008; Pettersson *et al.* 2011; Mackay 2015). Nevertheless, large epistatic effects between pairs of loci are believed to be unlikely (Wei *et al.* 2014), and the contribution of epistatic variance to the total genetic variance is expected to be small (Hill *et al.* 2008).

In summary, among the scenario's we studied here, scenarios D and E_{DD} with small effects, and scenarios E_{AA}, E_C and E_M are probably most realistic, because these scenarios always resulted in little dominance (max. 7%) and epistatic (max. 5%) variance.

3.4.2 Effect of genetic model on r_g

For the dominance model (D), we observed that r_g decreased with increasing size of dominance effects and with increasing difference of allele frequencies between populations. In some cases, the r_g can be negative due to dominance alone, as shown for a two-locus model (Wei *et al.* 1991). Such low values of r_g were, however, only obtained with scenarios where both loci showed substantial overdominance, and where the difference in allele frequencies between the two populations was at least 0.3 for one of the loci. In our study, we considered many loci and the distributions of dominance effects was based on empirical results (Bennewitz and Meuwissen 2010; Sun and Mumm 2016). These distributions resulted in only a fraction of loci showing overdominance (i.e. 0.5% for small effects, 16% for intermediate effects, and 51% for large effects). Furthermore, our simulations resulted in U-shaped distributions of allele frequencies in the last generation of the historical population, which agrees with expectations based on neutral theory (Kimura and Crow 1964; Goddard 2001). After the two populations separated, allele frequency differences between populations were a result of drift and/or selection. We therefore believe that our simulations represent a more realistic model of quantitative traits and population divergence than those in Wei *et al.* (1991). In conclusion, given that dominance variance is usually small and overdominance does not occur frequently, our results show that it is unlikely that true r_g values lower than 0.80 are due to dominance effects alone.

In another simulation study, where the fraction of loci showing overdominance was 12%, realized r_g was 0.78 (Esfandyari *et al.* 2015a). Although the fraction of loci showing overdominance in that study was comparable to our scenario with intermediate dominance effects, our realized r_g in that scenario was much lower (0.57). This difference is likely due to the smaller number of generations that populations diverged in the study of Esfandyari *et al.* (2015a).

With epistasis, r_g decreased with increasing size of epistatic effects and with increasing difference of allele frequencies between populations, and the value of r_g depended on the nature of the epistatic interaction (i.e. configuration). In addition, there was a tendency for configurations that resulted in large initial non-additive variance to result in smaller values of r_g (Figure S 3.3). Even though large epistatic effects are unlikely and epistatic variance is expected to be small, r_g could be as low as 0.45 for supposedly realistic epistatic scenarios.

To our knowledge, the relationship between the nature of epistasis and r_g has not been studied before. The mechanism behind differences in r_g between epistatic models can be illustrated with an example of two interacting loci. Suppose that both loci have an additive effect (a) of 1, an epistatic coefficient (γ) of 0.5, and the allele frequency at locus 1 (p_1) is the same in both populations (here we use 0.10). Then, we study the effect of allele frequency difference between populations at locus 2 (Δp_2) on the difference in average effects between populations ($\Delta\alpha$) for locus 1 and 2. Results show that E_{AA} and E_M interactions only affect the α of the locus with fixed p (locus 1), whereas E_{DD} and E_C interactions affect the α at both loci (Figure S 3.4). Note that this result was the same with different values for a, γ, or p_1. This shows that, in general, E_{AA} and E_M interactions create a dependency of α at a locus on the allele frequency of all loci it interacts with, whereas E_{DD} and E_C interactions also create a dependency of α on the allele frequency of the locus itself. These mechanisms may contribute to the differences in r_g between genetic models, because the interplay between differences in allele frequencies and r_g depends on the genetic model.

3.4.3 Effect of magnitude of non-additive effects on r_g

As expected, an increase in magnitude of dominance effects resulted in a lower r_g, which is in line with results from Wei *et al.* (1991). Similarly, an increase in magnitude of epistatic effects also resulted in a lower r_g. An important question is whether this decrease of r_g due to an increase in magnitude continues until the theoretical limit of $r_g = -1$ is reached. Additional analyses revealed that r_g appears to asymptote with increasing magnitude of non-additive effects. In these analyses, we repeated our original simulations of genetic models D and E_{AA}, using non-additive effects that were multiplied by 100 for all magnitudes. Results from those simulations showed that the difference in r_g between "small", "intermediate", or "large" effects had indeed disappeared (Figure S 3.5), and that the lower bound of realized values for r_g was ~0.25 with scenario D and ~0.36 with scenario E_{AA}.

To show the mechanism behind this result, we again consider a two-locus model where, like before, both loci have an additive effect (a) of 1, the allele frequency of locus 1 (p_1) is 0.10 in both populations and $\Delta p_2 = 0.20$. We studied the effect of the magnitude of the epistatic effect (γ) on the absolute difference in average effects between populations, relative to the absolute value of α in population 1 ($\Delta\alpha/\alpha_A$). We observed that for all epistatic models, especially for larger values of γ, both $\Delta\alpha$ and α_A increase roughly linearly with γ, and that therefore $\Delta\alpha/\alpha_A$ stops increasing with large values of γ (Figure S 3.6). Note that the same mechanism was observed with dominance when p_2 was the same in both populations and $\Delta p_1 = 0.20$. Hence, a change in magnitude equally affects the variance of α's in the two populations, and the covariance between them. As a result, r_g is unaffected by a change in size of non-additive effects when non-additive effects are already large. In conclusion, when non-additive effects are very large, r_g no longer depends on the magnitude of non-additive effects relative to the magnitude of functional additive effects. At that point, there is a lower bound of r_g that is determined by the nature of the non-additive effects (i.e. type of inter-allelic interaction) and by the difference in allele frequencies between populations.

3.4.4 Number of epistatic interactions

In the epistatic scenarios, we assumed that each locus interacted with 5 other loci. Because little is known about the number of interactions per locus (N_γ) in reality, we tested whether our results were sensitive to a change in N_γ. For that purpose, we repeated all simulations of epistatic scenarios with $N_\gamma = 100$. Note that the total functional epistatic variance with $N_\gamma = 100$ was the same as with $N_\gamma = 5$, because the epistatic coefficients were scaled with N_γ, so that the product $N_\gamma\sigma_\gamma^2$ is constant. This analyses resulted in values of r_g that were very similar to those of our original simulations (results not shown), suggesting that, in our simulations, the value of r_g depends on the level of total functional epistatic variance, which scales similarly with N_γ or σ_γ^2.

3.4.5 Effect of selection on r_g

Non-additive effects and selection create a complex interplay between average effects, the difference in allele frequencies between populations ($\overline{\Delta p}$) over time, and their effects on r_g. For a trait with small dominance effects under selection, we observed that $\overline{\Delta p}$ was almost the same as for an additive trait (Figure 3.8). We expected, however, that directional dominance would reduce $\overline{\Delta p}$, because the average effect at a locus can become smaller or even switch sign when the frequency

of the favourable dominant allele increases (Falconer and Mackay 1996). This change in average effects would affect the change in allele frequencies over time due to selection in population 1, because the selection pressure at loci may change. A reduction in $\overline{\Delta p}$ with small dominance effects was not observed, probably because only a small fraction of loci showed full- or over-dominance. Indeed, with large dominance coefficients (so that the fraction of loci showing over-dominance was much larger compared to with small dominance coefficients) $\overline{\Delta p}$ was smaller (Figure 3.6). In real data, however, we do not expect a large fraction of loci that show full- or over-dominance (Wellmann and Bennewitz 2011). It is therefore unlikely that dominance significantly affects the change in allele frequencies over time due to selection, compared to a purely additive trait.

For a trait with epistatic effects under selection, we observed that $\overline{\Delta p}$ was a bit smaller than that for a trait with only additive effects (Figure 3.8). Similar to the models with only dominance effects, this reduction in $\overline{\Delta p}$ was expected because the average effect at a locus can become smaller or switch sign over time in the presence of epistasis. How epistasis affects the change in allele frequencies due to selection depends on the directionality of the epistatic interaction effect. Theory suggests that, compared to pure additivity, positive interactions (i.e. in the same direction as the additive effects) will promote the selection of favourable alleles, whereas negative interactions (i.e. in the opposite direction from the additive effects) will suppress the selection of favourable alleles (Carter *et al.* 2005; Hansen 2013; Paixão and Barton 2016). We chose to simulate both positive and negative interactions with equal probabilities, because empirical studies suggest that epistatic interactions are not biased in being either positive or negative (i.e. they are non-directional) (Mackay 2014). Our results showed that, for a trait with intermediate epistatic effects, the net effect of having both positive and negative interactions was a decrease in fixation rate of favourable alleles (i.e. with a positive a), and an increase in fixation rate of unfavourable alleles (i.e. with a negative a) compared to an additive trait (Figure S 3.7). Similar results were found by Esfandyari *et al.* (2017). In conclusion, epistatic effects may affect r_g through two related mechanisms. First, with an epistatic model and when selection takes place in one of the populations, the difference in allele frequencies between populations may be smaller compared to an additive model. This reduction occurs because negative interactions decrease the fixation rates of favourable alleles, and increase those of unfavourable alleles. Second, for given allele frequency differences, the value of r_g depends on the nature of the epistatic interaction.

3.4.6 Loss of additive genetic variance

Selection experiments in Drosophila, maize, and *Escherichia coli* have shown that additive genetic variation (V_A) can be maintained for at least 100 generations (Hill 2016). Some researchers suggested that this preservation of V_A may be due to the conversion of non-additive genetic variance to additive genetic variance (Cheverud and Routman 1996; Hallander and Waldmann 2007; Hill 2017). Simulation studies, however, have failed to show a preservation of V_A due to this conversion (Carter *et al.* 2005; Esfandyari *et al.* 2017). Similarly, our simulations showed little conversion of non-additive genetic variance to V_A with genetic models E_{AA} and E_M, and no conversion with other genetic models (Table 3.2). As a result, almost all additive genetic variance was lost after 50 generations (Figure S 3.2).

The large loss of additive genetic variance in our simulations may be due to two reasons. First, there was little epistatic variance in generation one that could be 'converted' to V_A in subsequent generations (Hill *et al.* 2008; Maki-Tanila and Hill 2014). This was largely because the allele frequency distribution was strongly U-shaped in generation one. Second, mutational variance was zero because there were no mutations simulated after the historical generation. Even though these mechanisms may explain some of the loss of V_A in our simulations, the issue still remains that, to date, simulations have failed to convincingly reproduce the conservation of V_A observed in reality (Johnson and Barton 2005; Walsh and Lynch 2018).

3.4.7 Practical relevance

In our simulations, there was selection in only one of the populations, while the other population was unselected. In reality, populations may have been divergently selected (e.g. Friesian Holstein vs Angus cattle), resulting in larger differences in allele frequencies than simulated here. Hence, r_g between divergently selected populations may be smaller than observed in our simulations.

In this study, we assumed that there were no genotype x environment interactions (GxE), so that r_g values smaller than one were only due to non-additive effects. In reality, both non-additive effects and GxE may contribute to r_g values being smaller than one. The relative importance of non-additive effects and GxE can be inferred from the difference between estimated r_g, from a design where the populations were tested in different environments, and from a design where one of the populations was tested in the environment of the other population. This approach is similar to what was proposed by Wientjes and Calus (2017) to dissect the components of the genetic correlation between purebred and crossbred

performance. However, to our knowledge, there are no studies that have used this approach to disentangle the effects of non-additive effects and GxE on r_g. This study shows that, even without GxE, the r_g can be substantially smaller than one, and sometimes even close to zero.

Estimated genetic correlations between two populations ($\hat{r}_g$) may differ across traits (e.g. Lund *et al.* 2011; Karoui *et al.* 2012; Porto-Neto *et al.* 2015). For example, in dairy cattle, $\hat{r}_g$ of fertility traits tended to be lower than those of fat yield and milk production (Karoui *et al.* 2012). The results from the present study suggest that such differences in $\hat{r}_g$ may indicate differences in the underlying genetic model between traits (i.e., in the importance of non-additive effects). Although this may be the case, differences in $\hat{r}_g$ between traits can arise through other mechanisms as well. First, $\hat{r}_g$ often include a component due to GxE interactions. Such GxE interactions may be more important for some traits than for others, resulting in differences in $\hat{r}_g$ between traits. Second, different traits are influenced by (at least partly) different QTL, and some traits may have been under stronger selection than others. As a result, the differences in allele frequencies at QTL between populations may vary across traits. These mechanisms may result in differences in $\hat{r}_g$ between traits, even when the underlying genetic models of those traits are similar. It is therefore questionable whether inferences can be made about differences in genetic model among traits, based on differences in $\hat{r}_g$.

The results in this study may be relevant for the prediction of additive genetic values across populations using genomic information. In this strategy, termed across-population genomic prediction, average effects at markers are estimated in one population, and used to compute additive genetic values in another population (de Roos *et al.* 2009; Hayes *et al.* 2009a). It has been suggested that the inefficiency of across-population genomic prediction is partly due to differences in linkage disequilibrium between markers and QTL. This insight has inspired the use of whole-genome sequence (WGS) data, because in WGS data, genotypes of the QTL themselves are included (Iheshiulor *et al.* 2016; Raymond *et al.* 2018a; Raymond *et al.* 2018b). The results of the current study suggest, however, that even when QTL genotypes are known and their average effects are accurately estimated in one population, across-population genomic prediction may be inefficient, because r_g can differ considerably from one, even when genetic variance is mostly additive. This view is supported by the results of Raymond *et al.* (2018b), who reported that although the r_g estimated from putative QTL was higher than the estimate from regular marker data, it was still lower than one.

Similar to across populations, genomic prediction from current to future generations may be inefficient because of changes in allele frequencies, and the subsequent changes in average effects at QTL. In other words, two different generations can be considered as two populations that have a genetic correlation between them that may be lower than unity. The results of this study may therefore partly explain the need for frequent retraining of genomic prediction models to achieve constant accuracy across generations (Sonesson and Meuwissen 2009; Wolc *et al.* 2011). We expect, however, that the change in allele frequency at a single QTL is relatively small across a few (4-5) generations, especially for traits that are highly polygenic. As a result, r_g may be relatively high across a few generations. Nevertheless, the relative contribution of non-additive effects to the decline of genomic prediction accuracy across generations is currently unknown, and would be an interesting topic for future research.

3.5 Conclusion

Our findings show that the genetic correlation between populations (r_g) is partly determined by the difference in allele frequencies between populations and the magnitude of non-additive effects. Large differences in allele frequencies and large non-additive effects resulted in low values of r_g. For both dominance and epistasis, when non-additive effects become extremely large, r_g has a lower bound that is determined by the nature of non-additive effects, and the difference in allele frequencies between populations. In addition, we found that with epistasis, r_g depends on the level of total functional epistatic variance, which is a function of epistatic effect size and the number of interactions per locus. Given that dominance variance is usually small and there is not much overdominance, we expect that it is unlikely that values of r_g below 0.8 are due to dominance alone. With supposedly realistic epistasis, r_g could be as low as 0.45. These results may contribute to the understanding of differences in genetic expression of complex traits between populations, and may help in explaining the inefficiency of genomic prediction across populations.

3.6 Acknowledgements

The authors thank the Netherlands Organisation of Scientific Research (NWO) and the Breed4Food consortium partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin for their financial support. The use of the HPC cluster has been made possible by CAT-AgroFood (Shared Research Facilities Wageningen UR). The authors also thank Mehdi Sargolzaei for providing an improved version of QMSim.

3.7 Appendix

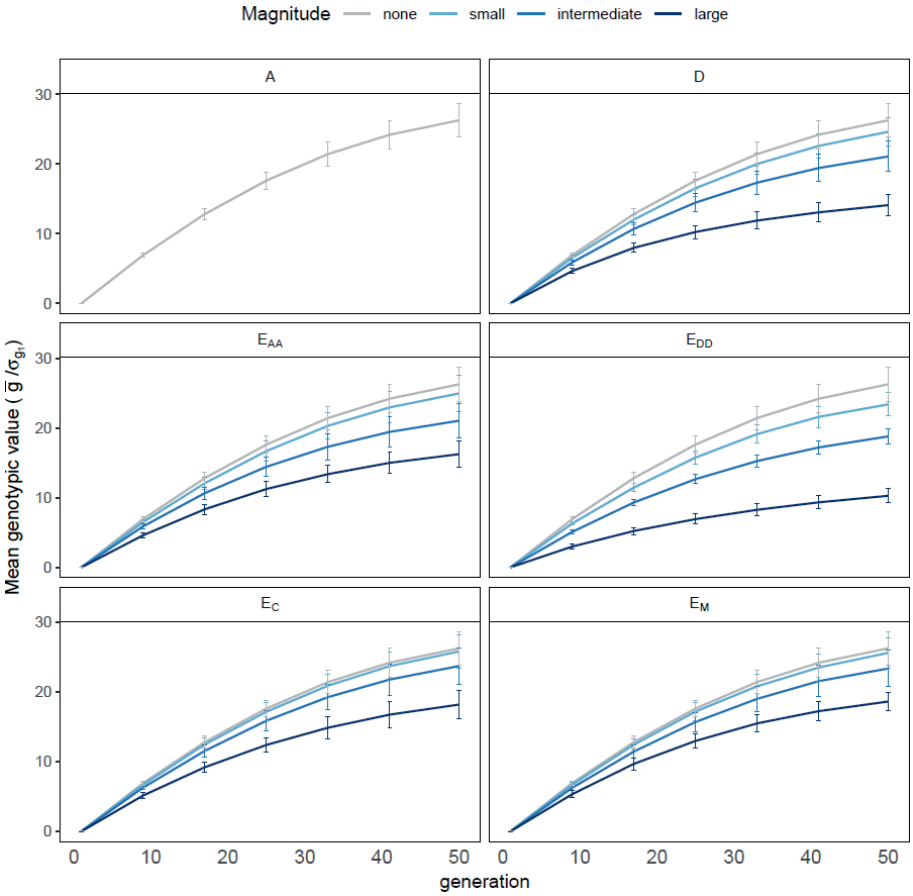

Figure S 3.1 Mean genotypic value across generations for different genetic models and magnitudes. In each subplot, the additive scenario (A) was included for reference (i.e. magnitude "none").

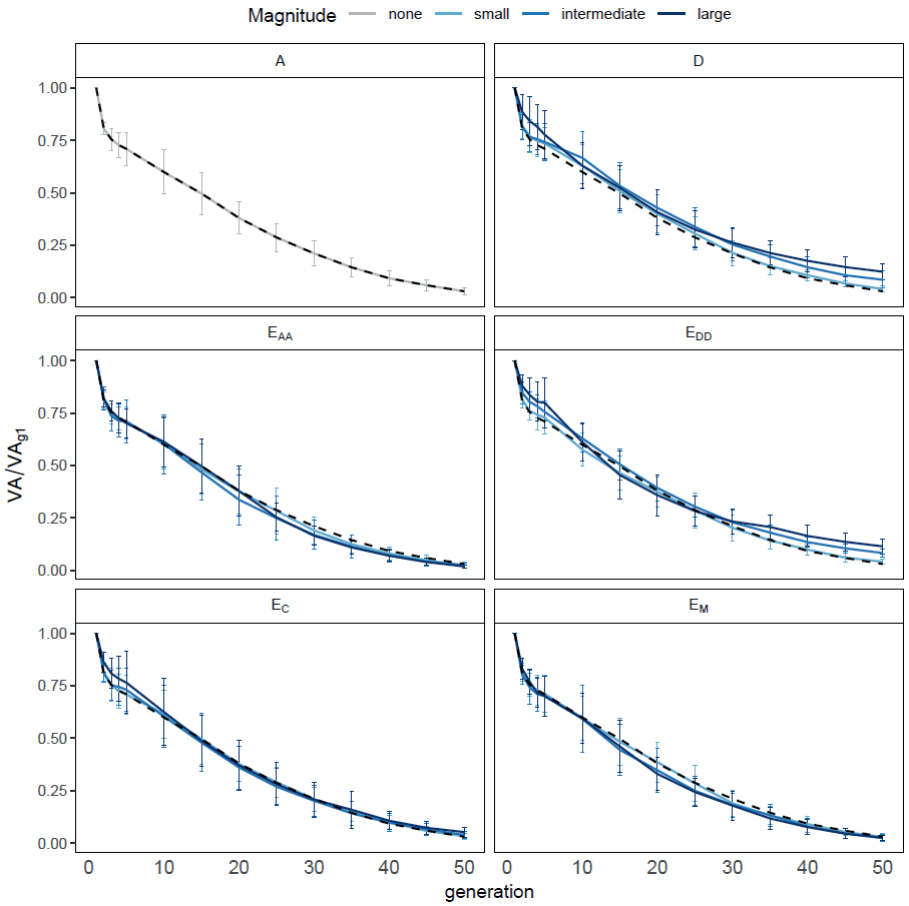

Figure S 3.2 Loss of additive genetic variance over time under different genetic models and magnitudes. The dotted line represents the additive genetic variance under an additive model.

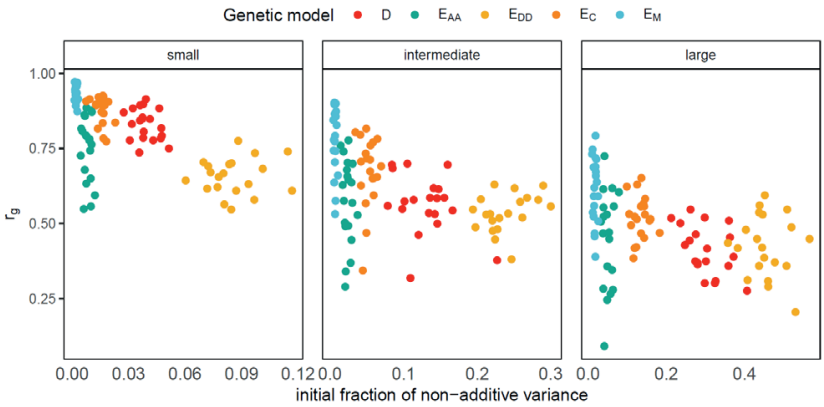

Figure S 3.3 Relationship between the fraction of inital non-additive variance (i.e. sum of dominance and epistatic variance divided by the total genetic variance in generation 1) and r_g in generation 50, for all scenarios.

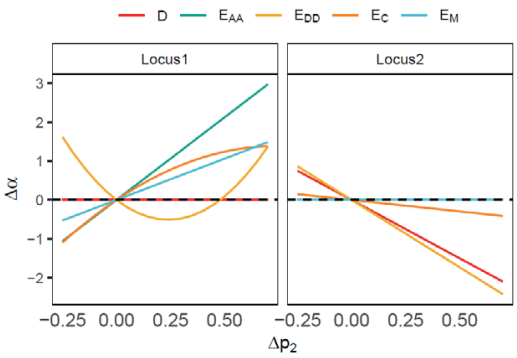

Figure S 3.4 Effect of difference in allele frequency between two populations at locus 2 (x-axis) on the difference in average effects at locus 1 (left column) and at locus 2 (right column). Colours of lines indicate the epistatic interaction model.

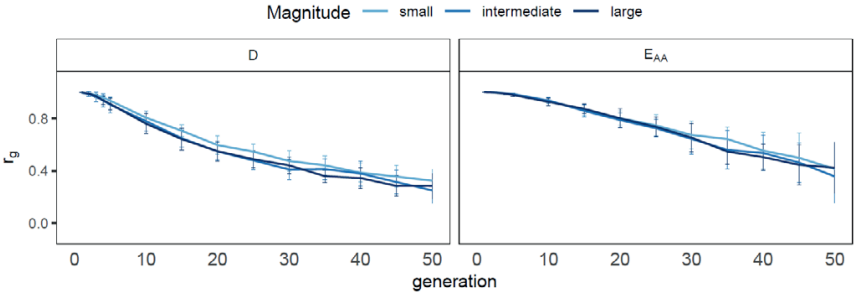

Figure S 3.5 Effect of magnitude of non-additive effects on r_g across generations, using non-additive effects that were 100 times larger than in the original simulations. The left plot is for genetic model D and the right plot for genetic model E_{AA}. Colours of lines indicate the magnitude of non-additive effects.

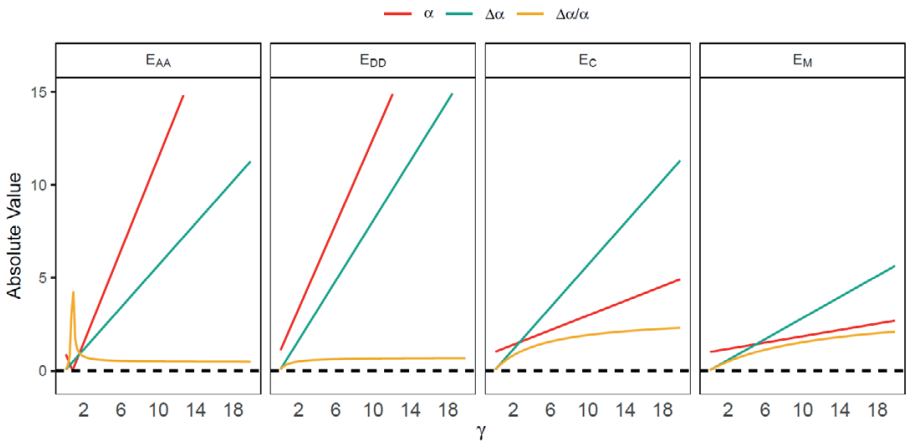

Figure S 3.6 Effect of magnitude of epistatic effects (γ) on the average effect of locus 2 (red), the difference in average effects at locus 2 between populations (green), and the ratio between them (yellow).

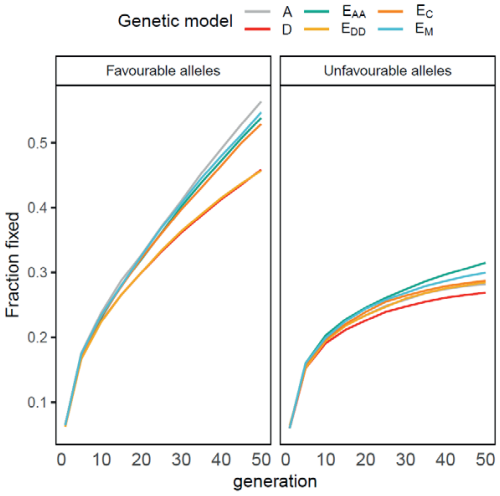

Figure S 3.7 Effect of genetic model (indicated by coloured lines) on the rate of fixation of favourable (left plot) and unfavourable (right plot) alleles across generations (x-axis).

4

Predicting the purebred-crossbred genetic correlation from genetic variances within, and covariance between parental lines

Pascal Duenk, Piter Bijma, Yvonne C.J. Wientjes and Mario P.L. Calus

Animal Breeding and Genomics, Wageningen University and Research,

6700 AH Wageningen, the Netherlands

In preparation

Abstract

Pig and poultry breeders benefit from heterosis and breed complementarity by mating individuals from genetically distinct purebred lines to produce crosses. While the aim of such breeding programs is to improve crossbred performance, selection usually takes place in the purebred lines based on purebred performance. The response to selection in crossbred performance therefore depends on the genetic correlation between purebred and crossbred performance (r_{pc}). The r_{pc} can be lower than one due to non-additive effects in combination with differences in allele frequencies between parental lines. This suggests that r_{pc} can be expressed as a function of parameters in the parental lines. In this study, we derive expressions for r_{pc} based on genetic variances within, and genetic covariance between parental lines. These expressions were derived for a genetic model with additive and dominance effects (D), and additive and epistatic additive by additive effects (E_{AA}). We validated our expressions using simulations of purebred parental lines and their crosses, where the parental lines were either selected or not. Finally, using these simulations, we investigated the value of r_{pc} for other genetic models, for which expressions could not be derived. Results show that our expressions provide exact estimates of r_{pc} for models D and E_{AA}, and accurate upper and lower bounds of r_{pc} for other genetic models, using information of parental lines only. In conclusion, our work demonstrates the impact of non-additive effects on r_{pc}, and aids in determining bounds of r_{pc}.

4.1 Introduction

Pig and poultry breeders benefit from heterosis and breed complementarity by mating animals from genetically distinct purebred parental lines to produce crossbred animals (Smith 1964; Dickerson 1973). The aim of such breeding programs is usually to maximize crossbred (CB) performance, while selection decisions are made in the parental lines, usually based on measurements of purebred (PB) performance. As a result, the response to selection in CB performance depends partly on the genetic correlation between PB and CB performance (r_{pc}), which is generally lower than one across species and trait categories (Wei and van der Werf 1995; Besbes and Gibson 1999; Newman *et al.* 2002; Lukaszewicz *et al.* 2015; Mulder *et al.* 2016; Wientjes and Calus 2017; Duenk *et al.* 2019b). Hence, r_{pc} is an important parameter in breeding programs of pig and poultry.

The r_{pc} can be estimated with models that use phenotypic information on both PB and CB performance. Such models require either pedigree information that links the purebreds to crossbreds (Wei and van der Werf 1995; Lutaaya *et al.* 2001), or genotype information on both PB and CB animals (Wientjes *et al.* 2017). Tracking the pedigree in a crossbreeding system is often impractical, and collecting phenotypes and genotypes on crossbreds may be difficult and costly. Another issue in the optimization of a CB breeding program involves avoiding crosses between parental lines that result in low r_{pc}, because r_{pc} partly determines the response in CB performance with PB selection. However, breeding companies usually maintain many different breeding lines, and mating all possible combinations of lines and estimating the resulting r_{pc}'s is costly and time-consuming. To overcome these issues, it would be beneficial to breeders if the value of r_{pc} could be predicted based on data from their parental PB lines, instead of estimating r_{pc} from CB data.

The r_{pc} can be lower than one due to differences in trait definition between PB and CB performance (Lo *et al.* 1997; Zumbach *et al.* 2007), due to genotype by environment interactions (GxE) (Falconer 1952; Lutaaya *et al.* 2001), and due to genotype by genotype interactions (GxG) in combination with allele frequency differences between parental lines (Wei *et al.* 1991; Baumung *et al.* 1997; Duenk *et al.* 2019b). GxG interactions are also known as non-additive genetic effects (i.e. dominance and epistasis). Here we consider only the impact of GxG interactions on r_{pc}, assuming absence of GxE interactions. The impact of non-additive effects on r_{pc} has been studied by Wei *et al.* (1991) and Baumung *et al.* (1997), who derived expressions for r_{pc} in terms of known additive, dominance, and epistatic genetic effects of loci, and as a function of the differences in allele frequencies between parental lines. These expressions were, however, derived for one- and two-locus

models, and can therefore not be used to predict r_{pc} for traits that are highly polygenic. Furthermore, genetic effects and allele frequencies at QTL are usually unknown. Hence, for prediction of r_{pc}, there is a need for expressions of r_{pc} for polygenic traits, based on parameters in parental lines.

Previously, we investigated the impact of non-additive effects on the additive genetic correlation between lines (r_g) (Duenk *et al.* 2019a). Results showed that r_g decreases with increasing size of non-additive effects, and with increasing differences in allele frequencies between lines. In the current study, we will investigate the impact of non-additive effects on the relationship between r_g and r_{pc}. We show that r_{pc} is closely related to the r_g between parental lines, and that r_{pc} can be expressed in terms of genetic variances in and covariance between parental lines.

While the estimation of the genetic correlation between populations is relatively straightforward with a genomic relationship matrix (Wientjes *et al.* 2018), the interpretation of this correlation requires careful consideration. Following Duenk *et al.* (2019a), we define the r_g for line 1, because our interest is in the r_{pc} for line 1. Hence, we defined r_g between line 1 and line 2 as the correlation between additive genetic values of the individuals in line 1, for the trait expressed in line 1 and 2. In other words, suppose we know average effects of all QTL in lines 1 and 2. Then we can calculate two additive genetic values for the individuals in line 1; one based on the average effects in line 1, and one based on the average effects in line 2. The r_g for line 1 is the correlation between these two additive genetic values,

$$r_g = \frac{\sigma_{1,1(2)}}{\sqrt{\sigma_1^2}\sqrt{\sigma_{1(2)}^2}}. \qquad \textbf{4.1}$$

The σ_1^2 in the denominator is the ordinary additive genetic variance for purebred performance in line 1. The $\sigma_{1(2)}^2$ is the additive genetic variance *in line 1*, for the trait *expressed in line 2*; thus $\sigma_{1(2)}^2$ depends on the allele frequencies in line 1 and the average effect for performance in line 2. Similarly, the numerator, $\sigma_{1,1(2)}$, is the additive genetic covariance *in line 1*, between the trait expressed *in line 1*, and expressed *in line 2*; thus $\sigma_{1,1(2)}$ depends on the allele frequencies in line 1, and the average effects for performance in line 1 and 2.

We focus here on the r_g for line 1. One can also define the r_g for line 2, but this is a different parameter because it depends on the allele frequencies in line 2. In other words, $\sigma_{1(2)}^2$ differs from the ordinary additive genetic variance for purebred performance in line 2 (denoted as σ_2^2). Similarly, the covariance also differs between

lines 1 and 2, $\sigma_{1,1(2)} \neq \sigma_{2,2(1)}$; while both covariances depend on the average effect in both lines, the $\sigma_{1,1(2)}$ depends on the allele frequencies in line 1, while $\sigma_{2,2(1)}$ depends on the allele frequencies in line 2.

Our aim is to derive expressions for the prediction of r_{pc} in a two-way crossbred breeding program, based on genetic variances within parental lines, and genetic covariance between parental lines (i.e the terms in Equation 4.1). We assumed absence of GxE interaction between PB and CB performance, i.e. that purebreds and crossbreds were kept in the same environment. The expressions were derived for two genetic models; a genetic model with additive and dominance effects (D), and a genetic model with additive, and additive by additive (AxA) epistatic effects (E_{AA}). We validated our expressions using simulations of PB parental lines and their crosses, where the parental lines were either selected or not. Finally, using simulations, we also investigated the value of r_{pc} for two genetic models for which expressions could not be derived; a model with both dominance and AxA epistatic effects (D+E_{AA}), and a model with complementary epistatic effects (E_C). We compared results from simulation of these models with our predictions of r_{pc} under models D and E_{AA}.

4.2 Theory

We consider two PB parental lines (1 and 2) that are mated to produce CB individuals, and we assume that the PB and CB individuals are housed in the same environment, so that there is no genotype by environment interaction between PB and CB performance. The additive genetic correlation between PB and CB performance (r_{pc}) for line 1 is defined as the correlation between additive genetic values for PB and CB performance of members of line 1 (Wei *et al.* 1991; Falconer and Mackay 1996). For PB performance, the additive genetic value of a member of line 1 (individual i) is

$$v_{i,1} = \mathbf{h}_i'\boldsymbol{\alpha}_1, \qquad \textbf{4.2}$$

where $\mathbf{h}_i$ is a column vector of genotypes of individual i at QTL (measured as allele counts minus the average allele count in population 1), and $\boldsymbol{\alpha}_1$ is a column vector of average effects for PB performance at QTL in line 1. Similarly, the additive genetic value of individual i for CB performance is

$$v_{i,C} = \mathbf{h}_i'\boldsymbol{\alpha}_{1(C)}, \qquad \textbf{4.3}$$

where $\boldsymbol{\alpha}_{1(C)}$ is a column vector of average effects for CB performance at QTL in line 1. The r_{pc} for line 1 is the correlation between additive genetic values in Equations 4.2 and 4.3.

$$r_{pc} = cor(v_{i,1}, v_{i,C}) = \frac{cov(v_{i,1}, v_{i,C})}{\sqrt{var(v_{i,1})}\sqrt{var(v_{i,C})}} = \frac{\sigma_{1,1(C)}}{\sigma_1 \sigma_{1(C)}}, \quad \textbf{4.4}$$

where σ_1 is the additive genetic standard deviation for PB performance in line 1, $\sigma_{1(C)}$ is the additive genetic standard deviation for CB performance in line 1, and $\sigma_{1,1(C)}$ is the additive genetic covariance between PB and CB performance in line 1.

Our aim is to express r_{pc} in terms of genetic parameters in the parental lines. First, we derive expressions for α_1 and $\alpha_{1(C)}$, for a genetic model with additive and dominance effects (D), and for a model with additive, and additive by additive epistatic effects (E_{AA}). Second, we express the $\alpha_{1(C)}$ in terms of average effects for PB performance in the parental lines (α_1 and α_2). Third, we express $\sigma_{1,1(C)}$ and $\sigma_{1(C)}$, and finally the r_{pc}, in terms of genetic parameters in parental lines.

4.2.1 Derivation of average effects for PB and CB performance

There are several ways in which average effects can be defined. Here, we use the definition of average value of transmitted alleles to the offspring (Falconer 1985). Following this definition, the average effect of a locus is equal to the difference between the average effects of its alleles, where the average effect of an allele is the mean genotypic value of offspring produced by transmitting that allele. This definition is convenient here because we are interested in the effect of alleles in line 1 when mated to line 2, i.e., the value that is transmitted to a CB offspring. Strictly speaking, this is the definition of average excess, but it is equivalent to the average effect under random mating (Falconer 1985). Hence, if individuals of line 1 are mated at random to individuals of line 2, then average effect and average excess are identical, even though the resulting CB population is not in Hardy-Weinberg equilibrium. Throughout this study, we assume that functional additive, dominance, and epistatic effects of alleles are the same for PB and CB performance, and that for CB performance, functional effects of alleles are independent of line origin. Statistical additive, dominance and epistatic effects, however, are line dependent due to differences in allele frequency. In other words, there is GxG-interaction.

4.2.1.1 Dominance model (D)

Consider a locus that has an additive effect (a), a dominance effect (d), and no epistatic interactions. The average effect of this locus for PB performance in line 1 under genetic model D is equal to

$$\alpha_1^D = a + (1 - 2p_1)d, \quad \textbf{4.5}$$

where p_1 is the frequency of the focal allele in line 1 (Fisher 1941). The derivation of α_1^D using the definition of transmitted alleles can be found in Falconer and Mackay (1996).

The average effect of this locus for CB performance in line 1 when mated to line 2 can be derived in a similar way. In contrast to alleles transmitted to PB animals, line 1 alleles transmitted to crossbreds will always pair with an allele from line 2. The average effect for CB performance in line 1 under genetic model D therefore depends on the allele frequency in line 2 (Pirchner and Mergl 1977; Dekkers 1999),

$$\alpha_{1(C)}^D = a + (1 - 2p_2)d, \qquad \textbf{4.6}$$

where p_2 is the frequency of the focal allele in line 2. Hence, under genetic model D, the average effect for CB performance in line 1 when mated to line 2 is equal to the average effect for PB performance in line 2 (α_2^D) (see also Zeng *et al.* (2013)).

4.2.1.2 Additive by additive epistasis model (E_{AA})

With additive by additive (AxA) epistasis (i.e. genetic model E_{AA}), the average effect at a locus depends not on the allele frequency at the locus itself, but also on the allele frequencies at the locus it interacts with. Consider a locus F with alleles F and f, that has an additive effect (a), and an AxA epistatic interaction with locus G with alleles G and g, and the epistatic effect between F and G is denoted as ϵ. For simplicity of presentation, we assume in the following derivation that locus G has no additive effect[2]. Table 4.1 shows the genotypic values of the two-locus genotypes (e.g.,Wade 2002).

Table 4.1 Genotypic values of two locus genotypes with additive by additive (AxA) epistasis (model E_{AA}).

		P^F	H^F	Q^F
		FF	Ff	ff
P^G	GG	$a + \epsilon$	0	$-a - \epsilon$
H^G	Gg	a	0	$-a$
Q^G	gg	$a - \epsilon$	0	$-a + \epsilon$

[2] The result does not depend on the additive effect at locus G, and in the simulations used to validate the theoretical predictions, we simulated additive effects at all loci.

The average effect of locus F for PB performance in line 1 can be derived by computing the difference between the average effects of alleles F and f. The average effect of an allele is the mean genotypic value of offspring produced by transmitting that allele. For allele F, the average effect in line 1 is

$$\begin{aligned}\alpha^F &= p_1^F\left(P_1^G(a+\epsilon) + H_1^G(a) + Q_1^G(a-\epsilon)\right) + p_1^f\left(P_1^G(0) + H_1^G(0) + Q_1^G(0)\right)\\ &= P_1^G p_1^F \epsilon - Q_1^G p_1^F \epsilon + p_1^F a\\ &= p_1^F\left((P_1^G - Q_1^G)\epsilon + a\right),\end{aligned}$$

where P_1^G, H_1^G, and Q_1^G are genotype frequencies of locus G in line 1 (Table 4.1), p_1^F is the frequency of allele F in line 1, and $p_1^f = 1 - p_1^F$. Similarly, for allele f, the average effect is

$$\begin{aligned}\alpha^f &= p_1^F\left(P_1^G(0) + H_1^G(0) + Q_1^G(0)\right) + p_1^f\left(P_1^G(-a-\epsilon) + H_1^G(-a) + Q_1^G(a+\epsilon)\right)\\ &= -P_1^G p_1^f \epsilon + Q_1^G p_1^f \epsilon - p_1^f a\\ &= -p_1^f\left((P_1^G - Q_1^G)\epsilon + a\right)\end{aligned}$$

The average effect of locus F for PB performance in line 1 under genetic model E_{AA} is

$$\begin{aligned}\alpha_1^{AA} &= \alpha^F - \alpha^f\\ &= (p_1^F + p_1^f)\left((P_1^G - Q_1^G)\epsilon + a\right)\\ &= a + (P_1^G - Q_1^G)\epsilon\\ \alpha_1^{AA} &= a - (1 - 2p_1^G)\epsilon,\end{aligned} \quad \textbf{4.7}$$

where p_1^G is the frequency of allele G in line 1. The average effect for PB performance in line 2 can be obtained by using the line 2 allele frequency in Equation 4.7.

The average effect for CB performance of line 1 can be derived similarly using the expected genotype frequencies in the crossbreds at locus G. This results in

$$\alpha_{1(C)}^{AA} = a - \left(1 - 2p_C^G\right)\epsilon, \quad \textbf{4.8}$$

where p_C^G is the expected frequency of allele G in the crossbreds. Given the expressions for α_1^{AA} and α_2^{AA} (Equations 4.7 and 4.8), and using $p_C^G = 0.5(p_1^G + p_2^G)$, the average effect of CB performance in line 1 under genetic model E_{AA} can be written as the mean of average effects for PB performance in line 1 and line 2

$$\alpha_{1(C)}^{AA} = 0.5(\alpha_1^{AA} + \alpha_2^{AA}). \quad \textbf{4.9}$$

4.2.2 Derivation of r_{pc}

In the following, we use our derivations of α_1 and $\alpha_{1(C)}$ for genetic models D and E_{AA} to derive the additive genetic variance for CB performance in line 1 ($\sigma^2_{1(C)}$), and the additive genetic covariance between PB and CB performance in line 1 ($\sigma_{1,1(C)}$). Regardless of the genetic model, we will define the additive genetic variance for PB performance in line 1 as

$$\sigma_1^2 = var(\mathbf{h}_i'\boldsymbol{\alpha}_1) = \mathbf{h}_i'\mathbf{h}_i\sigma^2_{\alpha_1}. \quad \textbf{4.10}$$

4.2.2.1 Dominance model (D)

With dominance (model D), average effects for CB performance in line 1 are equal to average effects for PB performance in line 2 (Equation 4.6). Hence, with the dominance model, the additive genetic variance for CB performance in line 1 can be written as

$$\sigma^2_{1(C)} = var(\mathbf{h}_i'\boldsymbol{\alpha}_2) = \mathbf{h}_i'\mathbf{h}_i\sigma^2_{\alpha_2} = \sigma^2_{1(2)}, \quad \textbf{4.11}$$

where $\sigma^2_{\alpha_2}$ is the variance of average effects for PB performance in line 2. As mentioned in the introduction, it is important to note that $\sigma^2_{1(2)}$ is the additive genetic variance *in line 1*, for the trait *expressed in line 2*. This can be seen from the notation $\mathbf{h}_i'\mathbf{h}_i\sigma^2_{\alpha_2}$, where $\mathbf{h}$ includes genotypes of line 1 and $\sigma^2_{\alpha_2}$ refers to average effects expressed in line 2 (see also Introduction and Discussion). The additive genetic covariance between PB and CB performance in line 1 can be written as

$$\sigma_{1,1(C)} = cov(\mathbf{h}_i'\boldsymbol{\alpha}_1, \mathbf{h}_i'\boldsymbol{\alpha}_2) = \mathbf{h}_i'\mathbf{h}_i\sigma_{\alpha_{1,2}} = \sigma_{1,1(2)}, \quad \textbf{4.12}$$

where $\sigma_{\alpha_{1,2}}$ is the covariance between average effects for PB performance in line 1 and 2, and $\sigma_{1,1(2)}$ is the additive genetic covariance for individuals in line 1, between the trait expressed in line 1 and 2. As a result, with model D, the r_{pc} can be written as

$$r^D_{pc} = cor(\mathbf{h}_i'\boldsymbol{\alpha}_1, \mathbf{h}_i'\boldsymbol{\alpha}_2) = \frac{\sigma_{1,1(2)}}{\sigma_1\sigma_{1(2)}}. \quad \textbf{4.13}$$

Hence, with model D, the r_{pc} in line 1 is equal to the genetic correlation between additive genetic values of individuals in line 1, for the trait expressed in parental lines 1 and 2.

4.2.2.2 Additive by additive epistasis (model E_{AA})

With AxA epistatic interactions (model E_{AA}), the average effect for CB performance in line 1 is equal to the mean of the average effects in the PB lines (Equation 4.9). The additive genetic variance for CB performance in line 1 can therefore be written as

$$\begin{aligned}\sigma^2_{1(C)} &= var\left(\mathbf{h}_i'\boldsymbol{\alpha}_{1(C)}\right) = \mathbf{h}_i'\mathbf{h}_i\ var\left(0.5\ (\alpha_1^{AA} + \alpha_2^{AA})\right)\\ &= 0.25\mathbf{h}_i\mathbf{h}_i'(\sigma^2_{\alpha_1} + \sigma^2_{\alpha_2} + 2\sigma_{\alpha_{1,2}})\\ &= 0.25\left(\mathbf{h}_i\mathbf{h}_i'\sigma^2_{\alpha_1} + \mathbf{h}_i\mathbf{h}_i'\sigma^2_{\alpha_2} + 2\mathbf{h}_i\mathbf{h}_i'\sigma_{\alpha_{1,2}}\right)\\ &= 0.25(\sigma_1^2 + \sigma^2_{1(2)} + 2\sigma_{1,1(2)}) \end{aligned} \quad \textbf{4.14}$$

The additive genetic covariance between PB performance and CB performance in line 1 is

$$\begin{aligned}\sigma_{1-1(C)} &= cov\left(\mathbf{h}_i'\boldsymbol{\alpha}_1, \mathbf{h}_i'\boldsymbol{\alpha}_{1(C)}\right) = \mathbf{h}_i'\mathbf{h}_i\ cov(\alpha_1, 0.5\alpha_1 + 0.5\alpha_2)\\ &= \mathbf{h}_i'\mathbf{h}_i\left(0.5\sigma^2_{\alpha_1} + 0.5\sigma_{\alpha_{1,2}}\right)\\ &= 0.5\ \mathbf{h}_i'\mathbf{h}_i\sigma^2_{\alpha_1} + 0.5\ \mathbf{h}_i'\mathbf{h}_i\sigma_{\alpha_{1,2}}\\ &= 0.5\sigma_1^2 + 0.5\sigma_{1,1(2)} \end{aligned} \quad \textbf{4.15}$$

Hence, the r_{pc} with genetic model E_{AA} is equal to

$$r_{pc}^{AA} = \frac{0.5\sigma_1^2 + 0.5\sigma_{1,1(2)}}{\sigma_1\sqrt{0.25\left(\sigma_1^2 + \sigma^2_{1(2)} + 2\sigma_{1,1(2)}\right)}} = \frac{\sigma_1^2 + \sigma_{1,1(2)}}{\sigma_1\sqrt{\left(\sigma_1^2 + \sigma^2_{1(2)} + 2\sigma_{1,1(2)}\right)}}, \quad \textbf{4.16}$$

showing that r_{pc} can be expressed in terms of additive genetic covariance in line 1, between the trait expressed in line 1 and line 2 ($\sigma_{1,1(2)}$), and the additive genetic variances in line 1, for the trait expressed in line 1 (σ_1^2) and line 2 ($\sigma^2_{1(2)}$).

4.2.3 Expressions as bounds of r_{pc}

It can be seen from the foregoing that the r_{pc} depends on the difference between average effects for PB and CB performance (Δ_α). With model D, Δ_α increases by $2(p_1 - p_2)$ per unit increase in the magnitude of non-additive effects, i.e. this is the difference between Equations 4.5 and 4.6. This is because with model D, the $\alpha_{1(C)}$ at a locus depends on the allele frequency in the mated line, whereas α_1 depends on the allele frequency in line 1. With model E_{AA}, in contrast, Δ_α increases by $2(p_1 - p_C) = (p_1 - p_2)$ per unit increase in non-additive effects, i.e. this is the difference between Equations 4.7 and 4.8. This is because with model E_{AA}, the $\alpha_{1(C)}$ of a locus depends on the allele frequency of the interacting locus in the cross, rather than the mated line. Since $p_C = 0.5(p_1 + p_2)$, the allele frequency difference between line 1 and the cross is only half that between line 1 and 2.

Because any non-additive interaction involves either dominance, epistasis, or both, the models D and E_{AA} may be seen as extremes, where r_{pc} either depends on (1) the difference in allele frequency between the parental lines (model D), or on the difference in allele frequency between the parental line and the cross (model E_{AA}). With other genetic models, r_{pc} may depend on (1), (2) or both. It seems, however, unlikely that other genetic models will lead to lower r_{pc} than with model D, because the maximum Δ_α is bounded by the difference in allele frequencies between parental lines. In addition, it seems unlikely that other genetic models will lead to higher r_{pc} than predicted with model E_{AA}, because the minimum Δ_α is bounded by the difference in allele frequencies between line 1 and the cross. Hence, we can expect that, with other genetic models, r_{pc} lies somewhere between r_{pc}^{D} and r_{pc}^{AA}. We therefore tested whether r_{pc}^{D} represents a lower bound for r_{pc}, and r_{pc}^{AA} represents an upper bound of r_{pc} for different genetic models.

a. Additive x additive (E_{AA})

		genotype locus F		
		FF	Ff	ff
genotype locus G	GG	+ε	0	-ε
	Gg	0	0	0
	gg	-ε	0	+ε

b. Complementary (E_C)

		genotype locus F		
		FF	Ff	ff
genotype locus G	GG	+ε	+ε	0
	Gg	+ε	+ε	0
	gg	0	0	0

Figure 4.1 Epistatic contrasts for two functional epistatic configurations. The epistatic effect between loci F and G is denoted with ϵ.

4.3 Methods

We validated the expressions for r_{pc} (i.e. Equations 4.13 and 4.16) with simulations. For that purpose, we simulated 7 purebred lines that were either positively (P), negatively (N), or randomly selected (R). Both positive and negative selection was considered, so that there were combinations of lines selected in the same direction (convergent) and in opposite direction (divergent), resulting in pairs of lines with small and large differences in allele frequencies. We considered four scenarios that differed in the type of non-additive effects present; one scenario with only dominance effects (D), one with only additive by additive (AxA) epistatic effects (E_{AA}), one with both dominance and AxA epistatic effects (D+E_{AA}), and one with complementary epistatic effects (E_C, Figure 4.1). We considered model E_C because it is expected to result in substantial non-additive variance of all types (i.e. dominance, additive by additive, dominance by additive, additive by dominance, and dominance by dominance) (Hill *et al.* 2008). For each scenario and each pairwise cross between

parental lines, we computed the realized (i.e. true) r_{pc} and compared it with predicted r_{pc} from Equations 4.13 and 4.16.

4.4 Simulation

4.4.1 Population

We simulated genotypes of animals from 7 breeding lines that originated from a common historical population using QMSim (Schenkel and Sargolzaei 2009), such that the number of generations of separation between pairs of lines varied from 20 to 100 (Figure 4.2). First, a historical population was simulated by randomly mating 600 females with 100 males for 200 generations. During the following 200 generations, population size was gradually decreased to 300 males and 50 females, to generate LD. Then, mating continued with constant population size for another 200 generations. In the last historical generation (generation 0), the population size was increased to 1500 males and 1500 females by creating litters of 10 offspring per mating. The effective population size (N_e) between generations -600 and 0 was ~234, calculated as the harmonic mean of $\frac{4N_mN_f}{N_m+N_f}$ in each historical generation, where N_m is the number of males and N_f is the number of females that become parents in a generation (Falconer and Mackay 1996).

From the last historical generation, 3 breeding lines (P50, R and N50) were created by sampling 300 females and 50 males for each of the lines, without replacement. Within each line, mating continued for 50 generations, selecting 50 males and 300 females in each generation based on own performance records with a broad sense heritability of 0.3. In line P50, selection was for high performance (positive selection, P), in line N50 for low performance (negative selection, N), and in line R, selection was random. Similar to lines P50 and N50, two additional lines (P25 and N25) were created by randomly sampling and mating 50 males and 300 females from line R in generation 25, again without replacement. Within each of these lines, mating continued for 25 generations with positive (P25) or negative (N25) selection. Finally, another two lines (P10 and N10) were created by randomly sampling and mating 50 males and 300 females from line R in generation 40. Within these lines, mating continued for 10 generations with positive (P10) or negative (N10) selection. Litter size was kept constant at 10 offspring in each of the breeding lines, and mating was always random. The average effective population size within the breeding lines was around ~115, which was calculated as $1/(2\Delta F)$, where ΔF is the inbreeding rate estimated from the pedigree (Falconer and Mackay 1996).

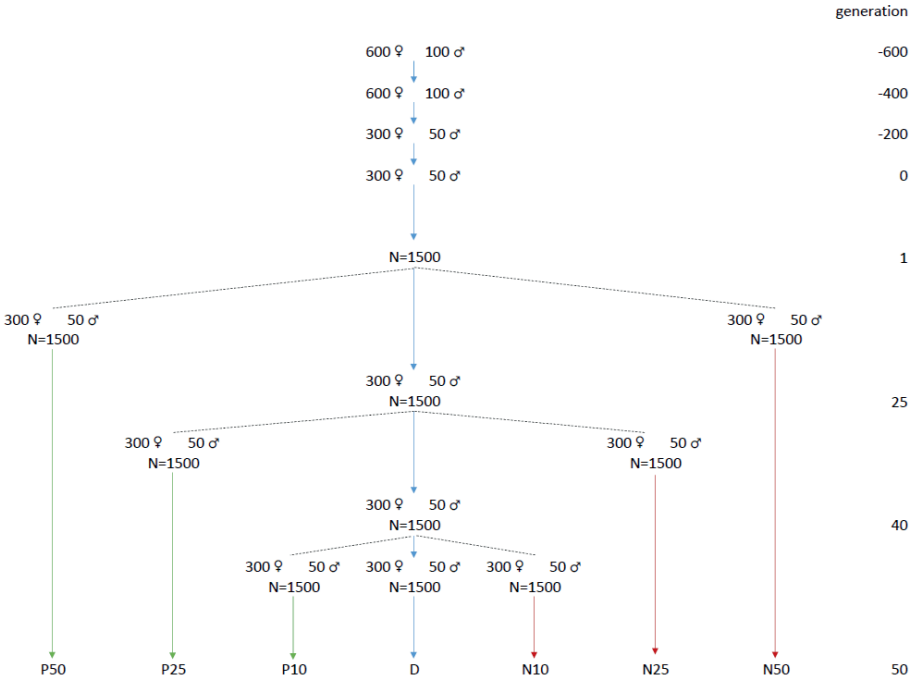

Figure 4.2 Overview of the simulation of 7 breeding lines. Green lines indicate positive selection based on own performance records, and red lines indicate negative selection based on own performance. Blue lines indicate random selection.

4.4.2 Genome

The genome consisted of 10 chromosomes of 1 Morgan each, and each chromosome had 5000 randomly positioned bi-allelic loci. In the first historical generation, allele frequencies of these loci were sampled from a uniform distribution. During the historical generations, mutation rate was $5.0*10^{-5}$, whereas after the historical generations, there were no mutations. In generation 0, the distribution of allele frequencies was U-shaped, and we randomly selected 1000 loci from the ones that segregated to become quantitative trait loci (QTL). We did not simulate markers, because our interest was in the true value of r_{pc}, not in its estimation.

4.4.3 Functional genetic effects

Additive effects ($\mathbf{a}$) of all 1000 QTL were sampled from a normal distribution with mean 0 and variance 1. The size of non-additive effects was assumed to depend on the size of additive effects of the loci involved. We first sampled dominance and epistatic coefficients. Dominance coefficients (δ) were sampled from a normal

distribution with mean of 0.2 and a standard deviation of 0.3, following empirical observations in Bennewitz and Meuwissen (2010) and Sun and Mumm (2016). Dominance effects (d) were subsequently computed as $\delta|a|$. Epistatic interactions were always between two loci, but each QTL had an epistatic interaction with 5 other, randomly sampled QTL. The interactions between pairs of loci followed either the additive by additive (E_{AA}) configuration, or the complementary (E_C) configuration (Figure 4.1), depending on the scenario. Epistatic coefficients ($\boldsymbol{\gamma}$) were sampled from a normal distribution with a mean of 0 and a standard deviation of 0.16. We chose this standard deviation so that the total functional epistatic variance per locus was comparable to the total functional dominance variance per locus with scenario D (i.e. $\sqrt{(0.2^2 + 0.3^2)/5} \approx 0.16$). Epistatic effects ($\epsilon$) were computed as $\gamma_{kl}\sqrt{|a_k a_l|}$ for all pairwise interactions between QTL k and l.

4.4.4 Average effects and additive genetic values

For a single locus, the average effect for PB performance in line 1 (α_1) can be computed from the functional genetic effects (a, d, and ϵ), and genotype frequencies in that line, as described in Duenk *et al.* (2019a), using the natural and orthogonal interactions (NOIA) model (Álvarez-Castro and Carlborg 2007; Vitezica *et al.* 2017). Similarly, the average effect for CB performance in line 1 when mated with line 2 ($\alpha_{1(C)}$) can be computed by the same procedure, with a small adjustment, as explained in the Appendix.

Additive genetic values for PB performance in line 1 were computed as

$$\mathbf{v}_1 = \mathbf{H}_1\boldsymbol{\alpha}_1, \qquad \textbf{4.17}$$

where $\mathbf{H}_1$ is a $(n\ x\ m)$ QTL genotype matrix of animals in line 1, and $\boldsymbol{\alpha}_1$ is a $(m\ x\ 1)$ column vector of average effects for PB performance in line 1, where n is the number of animals and m is the number of QTL. Genotypes in $\mathbf{H}_1$ were coded as in $\mathbf{h}_i$ (Equation 4.2) with elements

$$h_{i,j} = \begin{cases} 2 - 2p_j \\ 1 - 2p_j \\ 0 - 2p_j \end{cases} \quad \text{for genotypes} \quad \begin{cases} GG \\ Gg \\ gg \end{cases} \qquad \textbf{4.18}$$

for individual i at QTL j, where p_j is the frequency of allele G. Additive genetic values for crossbred performance of animals in line 1 (when mated to line 2) ($\mathbf{v}_C$) were computed by replacing $\boldsymbol{\alpha}'_1$ with $\boldsymbol{\alpha}'_{1(C)}$ in Equation 4.17.

4.4.5 Parameters of interest

The r_{pc} in line 1 when mated to line 2 was computed as the correlation between additive genetic values for PB ($\mathbf{v}_1$) and CB performance ($\mathbf{v}_C$) of animals in line 1. Hence, we calculated the true value of r_{pc} from the simulated true effects; we did not estimate r_{pc} from the simulated data. Note that this is not the same as the r_{pc} in line 2 when mated to line 1, because differences in allele frequencies between the lines lead to differences in contributions of QTL to the (co)variance of additive genetic values. In addition, average effects for CB performance in line 1 may be different from those in line 2 (e.g. with genetic model D, Equation 4.5 and 4.6). We therefore computed r_{pc} for all $7 * (7 - 1) = 42$ combinations of breeding lines. All simulations were replicated 20 times, resulting in $42 * 20 = 840$ realized r_{pc} values for each scenario.

We compared each of the realized r_{pc} values with predicted r_{pc} under genetic models D (r_{pc}^{D}, Equation 4.13) and E_{AA} (r_{pc}^{AA}, Equation 4.16). We expected that r_{pc}^{D} would exactly predict r_{pc} in scenario D, and that r_{pc}^{AA} would exactly predict r_{pc} in scenario E_{AA}. For other scenarios (D+E_{AA} and E_C), r_{pc} could not be expressed in terms of genetic parameters in the parental lines. However, we expected that r_{pc}^{D} represents a lower bound, and r_{pc}^{AA} represents an upper bound of realized r_{pc}. In addition, we tried to predict realized r_{pc} with a multiple linear regression model that uses both r_{pc}^{D} and r_{pc}^{AA} as covariates, for each scenario separately.

4.5 Results

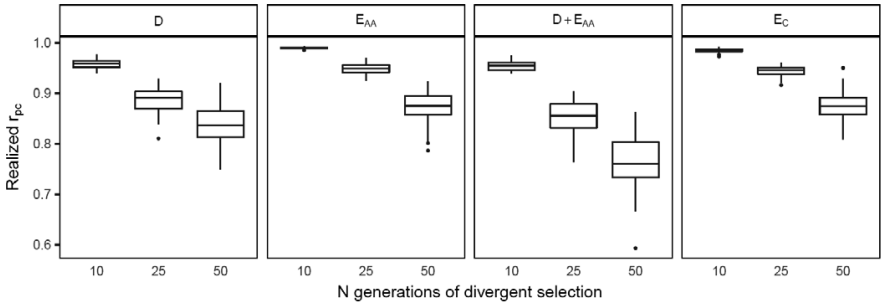

Figure 4.3 Realized r_{pc} (y-axis) for crosses between lines that were divergently selected for 10 (P10-N10), 25 (P25-N25), or 50 generations of selection (P50-N50) (x-axis). Panels refer to the genetic model simulated.

Figure 4.3 shows the realized r_{pc} for all crosses between lines that were divergently selected for 10, 25, or 50 generations. Results were very similar for positive and negatively selected lines, so we merged those lines in Figure 4.3. For all

scenarios (i.e. simulated genetic models), realized r_{pc} decreased with increasing generations of divergent selection, as expected. The lowest realized r_{pc} was observed when both dominance and epistasis were simulated (scenario D+E_{AA}), and the highest realized r_{pc} was observed when only epistasis was simulated (scenarios E_{AA} and E_C).

Figure 4.4 shows the predicted r_{pc} from our expressions, plotted against the realized r_{pc} from our simulations, for all replicates and for all combinations of parental lines within replicate. With simulated scenarios D and E_{AA}, our expressions for r_{pc} based on parameters in the purebred parental lines provided exact predictions of realized r_{pc} (left two panels in Figure 4.4). With scenario D+E_{AA}, our expressions for r_{pc} provide an upper bound (r_{pc}^{AA}) and lower bound (r_{pc}^{D}) for realized r_{pc}. With simulated scenario E_C, our expressions for r_{pc} provided an upper bound (r_{pc}^{AA}) and an approximate lower bound (r_{pc}^{D}) for realized r_{pc}. Here, realized r_{pc} was lower than the lower bound (r_{pc}^{D}) in 12% of the cases, in which realized r_{pc} was about ~0.01 lower than r_{pc}^{D}. For both scenarios D+E_{AA} and E_C, the gap between the predicted lower and upper bound increased with decreasing realized r_{pc}.

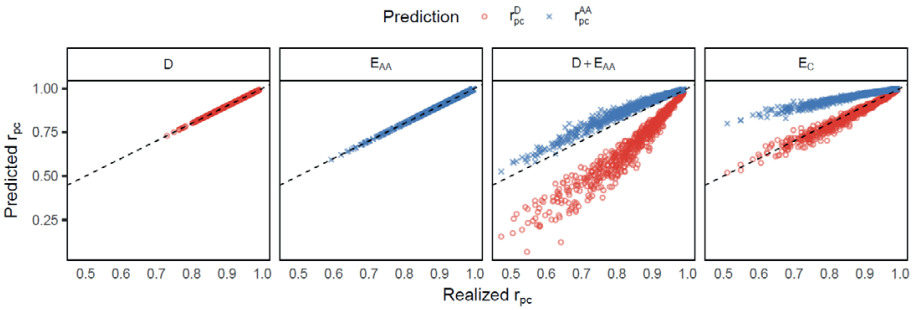

Figure 4.4 Predicted r_{pc} (y-axis) for a genetic model with only dominance (r_{pc}^{D}, red circles) and with only additive by additive epistasis (r_{pc}^{AA}, blue crosses), plotted against the realized r_{pc} from simulated scenarios, indicated by panels. The dotted line indicates y=x, so that that red circles below, and blue circles above this line indicate proper bounds.

With scenarios D+E_{AA} and E_C, we could not express realized r_{pc} in terms of genetic parameters in the purebred parental lines. We therefore fitted realized r_{pc} using a multiple linear regression model that uses both r_{pc}^{D} and r_{pc}^{AA} as covariates, for each scenario separately. Table 4.2 shows the estimated regression coefficients and coefficients of determination (R^2). Realized r_{pc} with scenarios D+E_{AA} and E_C could be accurately predicted by the regression model (Figure 4.5), because the R^2 was >0.99. With scenario D+E_{AA}, prediction of r_{pc} was based on both r_{pc}^{D} and r_{pc}^{AA} (with

regression coefficients of 0.34 and 0.65), whereas with scenario E_C, prediction of r_{pc} was mostly based on r_{pc}^{D} (with regression coefficients of 0.93 and 0.07).

Table 4.2 Estimated regression coefficients from multiple linear regression of realized r_{pc} on r_{pc}^{D} and r_{pc}^{AA}, for each scenario. Standard errors of estimated regression coefficients were all smaller than 0.01. The last column indicates the variance explained by the regression model (R^2).

scenario	r_{pc}^{D}	r_{pc}^{AA}	R^2
D	1.0	0.0	1.0
E_{AA}	0.0	1.0	1.0
D+E_{AA}	0.34	0.65	0.99
E_C	0.93	0.07	0.99

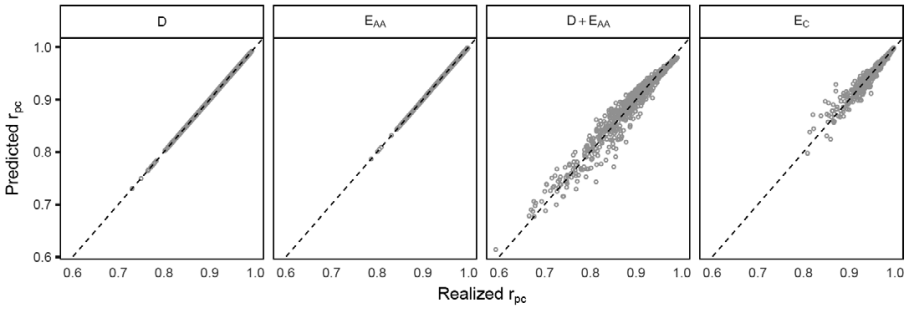

Figure 4.5 Predicted r_{pc} (y-axis) from a multiple linear regression model that fits both r_{pc}^{D} and r_{pc}^{AA}, plotted against the realized r_{pc} from simulated scenarios, indicated by panels. The dotted line indicates y=x.

4.6 Discussion

The aim of this study was to derive expressions of r_{pc} in a purebred line when mated to another purebred line, based on genetic variances within, and genetic covariance between parental lines. These expressions were derived for a genetic model with additive and dominance effects (D), and a genetic model with additive, and additive by additive (AxA) epistatic effects (E_{AA}). The results showed that our expressions provide exact predictions of r_{pc} for simulated scenarios D and E_{AA}. For scenarios with both dominance and AxA epistasis (D+E_{AA}), and models with complementary epistasis (E_C), our expressions provide upper and lower bounds for r_{pc}. With D+E_{AA}, realized r_{pc} always fell between these bounds, whereas with E_C, realized r_{pc} was lower than the lower bound in 12% of the cases.

The results of our simulations showed that realized r_{pc} decreased when the number of generations of divergent selection in parental lines increased, for all scenarios. This was in line with our expectations, because with divergent selection, the difference in allele frequencies between parental lines increases, causing an increase in the differences between average effects for PB and CB performance (Wei *et al.* 1991; Baumung *et al.* 1997). Across scenarios, the realized r_{pc} varied from 0.6 to 0.99, covering a large range of what is observed empirically (Wei and van der Werf 1995; Besbes and Gibson 1999; Newman *et al.* 2002; Lukaszewicz *et al.* 2015; Mulder *et al.* 2016; Wientjes and Calus 2017; Duenk *et al.* 2019b).

We expected that the predicted r_{pc} under the dominance model (r_{pc}^{D}) would yield an absolute lower bound of realized r_{pc}, because the difference between average effects for PB and CB performance at a single locus is maximized under model D. Our simulations showed that this lower bound was indeed correct in most cases, apart from a few replicates with complementary epistasis (model E_C). This is probably because, with model E_C, the expression for the average effect involves an interaction term between the allele frequency of the same locus in the mated line, and the allele frequencies of the interacting loci in the cross. In contrast, with model D, the expression for the average effect only involves the allele frequency of the same locus in the mated line. As a consequence, a difference in allele frequency between parental lines at a single locus can result in larger differences in average effects between PB and CB performance with model E_C, than with model D.

In reality, r_{pc} may be lower than the predicted lower bound given by our expression, because the predicted bounds do not account for any genotype by environment (GxE) interactions that may contribute to $r_{pc} < 1$ in reality. Despite this issue, the results of this study may be used to evaluate the contribution of GxE to the value of r_{pc}, by comparing estimates of r_{pc} from purebred and CB data with the predicted lower bound. For example, when the estimated r_{pc} is much lower than the predicted lower bound of r_{pc}, it is likely that the contribution of GxE is large compared to the contribution of non-additive effects and differences in allele frequencies.

The predictions of r_{pc} presented here are valid for prediction of r_{pc} for a purebred parental line that produces a two-way crossbred (i.e. resulting from mating with another purebred parental line). In practice, however, commercial animals are usually three- or four-way crossbreds. It may be possible to derive expressions of r_{pc} for three- and four-way CB performance, using a similar approach as in this study. The mathematics of such derivations may, however, become quite tedious, and we therefore chose to focus on two-way crossbreds only.

The results of this study can be used to predict the impact of non-additive effects on the r_{pc}, based on genetic variances within and covariance between parental lines. Recent developments in genome-wide marker panels have made it feasible to accurately estimate the variances within and covariances between distantly related lines (Karoui *et al.* 2012; Carillier *et al.* 2014; Wientjes *et al.* 2018). To our knowledge, there is only one study that presents both estimated r_{pc} from PB and CB data, and genetic variances within and covariances between parental lines (Xiang *et al.* 2017). In that study, estimated r_{pc} in the Yorkshire breed when mated to the Landrace breed was 0.67. Based on the parameters presented in that paper, the predicted lower bound of r_{pc} from our expressions was 0.30 and the predicted upper bound was 0.84. Hence, for that study, our expressions provide correct lower and upper bounds for r_{pc} based on parameters in the parental lines.

It should be noted that estimated genetic variances within and covariances between parental lines from empirical data are usually different from the ones used in our expressions (Equation 4.13 and 4.16). For example, when a bivariate model is used to estimate genetic parameters within and between two parental lines (say line 1 and 2), the resulting estimate for variance in line 2 refers to the variance in line 2, for the trait expressed in line 2 (i.e. σ_2). However, when the aim is to predict the bounds of r_{pc} in line 1, we need the variance in line 1, for the trait expressed in line 2 (i.e. $\sigma_{1(2)}$, Equation 4.13 and 4.16). In other words, we need to combine the allele frequencies in line 1 with the average effects in line 2. Similarly, the covariance between line 1 and 2 estimated from data ($\sigma_{1,2}$), is not the same as the covariance used in our expressions ($\sigma_{1,1(2)}$). Hence, before estimates of variances and covariances are used in the expressions for r_{pc}, they need to be transformed to a proper scale.

4.7 Acknowledgements

The authors thank the Netherlands Organisation of Scientific Research (NWO) and the Breed4Food consortium partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin for their financial support. The use of the HPC cluster has been made possible by CAT-AgroFood (Shared Research Facilities Wageningen UR). The authors also thank Mehdi Sargolzaei for providing an improved version of QMSim.

4.8 Appendix

With non-additive effects (i.e. dominance and epistasis), average effects are a function of genotype frequencies and functional genetic effects. The procedure to obtain average effects for purebred performance from known genotype frequencies and functional additive, dominance, and epistatic effects, was described in (Duenk *et al.* 2019a). In short, the procedure involves applying the natural and orthogonal interactions (NOIA) model (Álvarez-Castro and Carlborg 2007) for each epistatic interaction between two loci in a population with given allele frequencies, resulting in statistically orthogonal terms that contribute to the average effects of the two loci. Consider for example two loci, k and l, that have an epistatic interaction between them. The functional epistatic values for each possible two-locus genotype can be partitioned into 9 statistical genetic effects, using Equation 2 in Duenk *et al.* (2019a). For average effects for purebred performance, the frequencies used in $\mathbf{D}_{kl}$, $\mathbf{W}_k$ and $\mathbf{W}_l$ are those in the purebred line. The procedure immediately leads to two terms (α^k_{kl} and α^l_{kl}) that contribute to the average effects of loci k and l. For example, in a two-locus model where locus k only has an interaction with locus l, the average effect for purebred performance of locus k in line 1 is

$$\alpha^k_1 = a^k + (1 - 2p^k_1)d^k + \alpha^k_{kl},$$

where a^k is the functional additive effect of locus k, p^k_1 is the allele frequency of locus k in line 1, and d^k is the functional dominance effect of locus k. For purebred performance, the average effect of locus k depends on genotype frequencies of k in line 1, because alleles of locus k transmitted to purebred animals always pair with an allele from the same line origin. Furthermore, the average effect of k depends on genotype frequencies of locus l in line 1, because alleles of locus k transmitted to purebred animals will be expressed in the genetic background of line 1.

The same procedure can be used to obtain average effects for crossbred performance in line 1, by making a small alteration. The alleles of locus k transmitted to crossbreds always pair with an allele from line 2, and they will be expressed in the genetic background of crossbreds. Hence, the average effect for crossbred performance of locus k depends on genotype frequencies of k in line 2, and on genotype frequencies of l in the crossbreds. Thus, to obtain the average effect of locus k for crossbred performance, $\mathbf{W}_k$ should be constructed using the genotype frequencies of k in line 2, $\mathbf{W}_l$ should be constructed using the genotype frequencies of l in crossbreds, and $\mathbf{D}_{kl}$ should be a (9x9) diagonal matrix of two-locus genotype probabilities, constructed using genotype frequencies of k in line 2, and frequencies of l in crossbreds. Then, the average effect of locus k for crossbred performance in line 1 is

$$\alpha_{1(C)}^{k} = a^{k} + (1 - 2p_{2}^{k})d^{k} + \alpha'^{k}_{kl},$$

where p_2^k is the allele frequency of locus k in line 2. For the same epistatic interaction, the procedure should be repeated for locus l, because the average effect of locus l for crossbred performance depends on genotype frequencies of l in line 2, and on genotype frequencies of k in the crossbreds.

5

Estimating the purebred-crossbred genetic correlation of body weight in broiler chickens with pedigree or genomic relationships

Pascal Duenk[1], Mario P.L. Calus[1], Yvonne C.J. Wientjes[1], Vivian P. Breen[2], John M. Henshall[2], Rachel Hawken[2] and Piter Bijma[1]

[1]Animal Breeding and Genomics, Wageningen University and Research, 6700 AH Wageningen, the Netherlands

[2]Cobb-Vantress Inc., Siloam Springs, AR 72761-1030, USA

Genetics Selection Evolution (2019) 51:6

Abstract

Background

In pig and poultry breeding programs, the breeding goal is to improve crossbred (CB) performance, whereas selection in the purebred (PB) lines is often based on PB performance. Thus, response to selection may be suboptimal, because the genetic correlation between PB and CB performance (r_{pc}) is generally lower than 1. Accurate estimates of the r_{pc} are needed, so that breeders can decide if they should collect data from CB animals. r_{pc} can be estimated either from pedigree or genomic relationships, which may produce different results. With genomic relationships, the r_{pc} estimate could be improved when relationships between purebred and crossbred animals are based only on the alleles that originate from the PB line of interest. This work presents the first comparison of estimated r_{pc} and variance components of body weight in broilers, using pedigree-based or genotype-based models, where the breed-of-origin of alleles was either ignored or considered. We used genotypes and body weight measurements of PB and CB animals that have a common sire line.

Results

Our results showed that the r_{pc} estimates depended on the relationship matrix used. Estimates were 5 to 25% larger with genotype-based models than with pedigree-based models. Moreover, r_{pc} estimates were similar (max. 7% difference) regardless of whether the model considered breed-of-origin of alleles or not. Standard errors of r_{pc} estimates were smaller with genotype-based than with pedigree-based methods, and smaller with models that ignored breed-of-origin than with models that considered breed-of-origin.

Conclusions

We conclude that genotype-based models can be useful for estimating r_{pc}, even when the PB and CB animals that have phenotypes are closely related. Considering breed-of-origin of alleles did not yield different estimates of r_{pc}, probably because the parental breeds of the CB animals were distantly related.

5.1 Background

In pig and poultry breeding programs, the breeding goal is to improve crossbred (CB) performance, whereas selection in the purebred (PB) lines is often based on PB performance. Thus, response to selection in CB performance may be suboptimal, because the genetic correlation between PB and CB performance (r_{pc}) is generally lower than 1 (Wei and van der Werf 1995; Lukaszewicz *et al.* 2015; Wientjes and Calus 2017). An r_{pc} lower than 1 can be caused by genotype-by-environment interactions (Lutaaya *et al.* 2001; Dekkers 2007), by genotype-by-genotype interactions in combination with allele frequency differences between the two parental breeds (Wei *et al.* 1991), and by differences in trait definitions between PB and CB performance (Lo *et al.* 1997; Zumbach *et al.* 2007). With a low r_{pc}, the use of CB instead of PB data may improve response to selection for CB performance (Wei and van der Werf 1994; Bijma and van Arendonk 1998; Dekkers 2007; Van Grevenhof and Van Der Werf 2015). Thus, accurate estimates of the r_{pc} are needed, so that breeders can decide if they should collect data from CB animals.

r_{pc} is the additive genetic correlation between breeding values for PB and CB performance, and is defined as

$$r_{pc} = \frac{\sigma_{A_{PB,CB}}}{\sigma_{A_{PB}}\sigma_{A_{CB}}}, \qquad \textbf{5.1}$$

where $\sigma_{A_{PB,CB}}$ is the additive genetic covariance between breeding values for PB and CB performance and $\sigma_{A_{PB}}$ ($\sigma_{A_{CB}}$) is the additive genetic standard deviation in purebreds (crossbreds) (Wei *et al.* 1991; Falconer and Mackay 1996). To estimate r_{pc}, phenotypic data from both PB and CB animals are needed. When these data are available, r_{pc} can be estimated with a pedigree-based animal or sire model (Mrode 2005). Such models treat PB and CB performance as correlated traits and use a pedigree-based relationship matrix ($\mathbf{A}$) to link PB and CB observations (Wei and van der Werf 1995). To estimate r_{pc} with $\mathbf{A}$, pedigree data should be available for both PB and CB individuals, and provide a link between PB and CB individuals. When the CB individuals are paternal half-sibs of the PB individuals, the accuracy of r_{pc} estimated with $\mathbf{A}$ depends on the number of common sires between the PB and CB animals, and the accuracy of the estimated breeding values of the sires (Bijma and Bastiaansen 2014). However, in practice, pedigree information is often not recorded in CB populations and the number of sires that have both PB and CB offspring with phenotypes may be limited.

These requirements for estimating r_{pc} with pedigree information can be alleviated by replacing $\mathbf{A}$ with a multi-breed genomic relationship matrix ($\mathbf{G}$) (VanRaden 2008; Wientjes *et al.* 2017). An advantage of this approach is that the r_{pc} can then also be estimated when the PB and CB animals are more distantly related, or when pedigree information is not recorded. In addition, genomic relationships may be more accurate than pedigree relationships (Goddard 2009; Hayes *et al.* 2009c), which results in a smaller standard error of the estimate of r_{pc} (Visscher *et al.* 2014; Xiang *et al.* 2016).

Usually, the r_{pc} between the CB and one of the PB parental lines is estimated. As such, genomic relationships between PB and CB animals should ideally be based on alleles that originate from that PB parental line only. However, the ordinary $\mathbf{G}$ is based on both alleles of an individual, which in the case of CB individuals, also include those originating from the other PB line. For example, when r_{pc} is estimated between CB and its PB sire line, the ordinary $\mathbf{G}$ matrix is also based on alleles that originated from the dam line. An alternative for $\mathbf{G}$ is a genomic partial relationship matrix ($\mathbf{G}_{BOA}$) that is based on the breed-of-origin of the alleles in the CB animals (Ibañez-Escriche *et al.* 2009; Christensen *et al.* 2014). Recently, a method to determine the breed-of-origin of alleles (BOA) based on phased genotypes was developed, allowing $\mathbf{G}_{BOA}$ to be constructed (Vandenplas *et al.* 2016). In $\mathbf{G}_{BOA}$, relationships between PB and CB animals are expected to be more accurate than in $\mathbf{G}$, because relationships in $\mathbf{G}_{BOA}$ are based on marker alleles that originated from the same breed. This approach was successfully applied to estimate variance components from data of three-way crossbred pigs, where 93% of the alleles of the crossbreds could be assigned a breed-of-origin (Sevillano *et al.* 2016; Sevillano *et al.* 2017). However, empirical studies in other species are lacking and, to date, no studies have compared r_{pc} estimates and their standard errors from pedigree-based models to those from genotype-based models. In addition, it is not yet clear how r_{pc} estimates and their standard errors are affected by the model used. Thus, our objective was to compare estimates of r_{pc} and variance components obtained from pedigree-based and genotype-based models. In addition, we compared models that either consider or ignore the breed-of-origin of alleles. We analysed body weight in broilers, using genotypes and measurements of PB and CB animals that have a common sire line.

5.2 Methods

5.2.1 Data

Data were collected on male and female broilers from a PB sire line (A) and on a three-way cross between this sire line and crossbred dams (BC), where lines B and C are dam lines. The dam lines were selected on egg production and the sire line on male fertility, along with standard traits, i.e., growth, yield, and feed efficiency. The three parental lines (A, B, and C) were genetically distant, as shown by the principal component analysis plot (Figure 5.1).

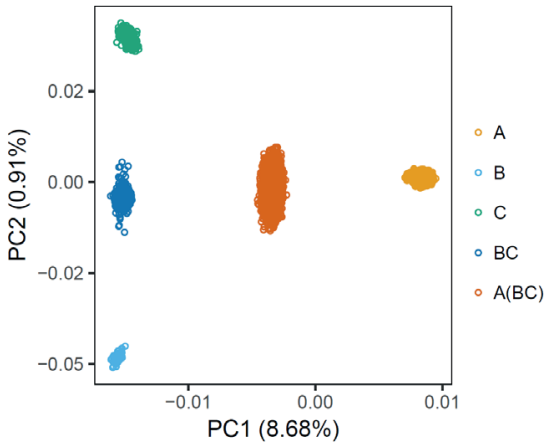

Figure 5.1
Principal component plot of the genotype data for the parental lines and the crossbreds. Values for principal component 1 (x-axis) are plotted against values for principal component 2 (y-axis). Between brackets is the variance explained by each principal component. Colours indicate genetic group.

PB and CB animals were weighed between 6 and 8 days of age (BW7) and between 33 and 36 days of age (BW35). We chose these phenotypes because they are easy to measure proxies for growth, which is an important trait for breeding companies (Cobb; 2018 personal communication). Phenotype recording was done in five consecutive trials of similar size, which each included both PB and CB animals. All animals were housed in the same environment, in a barn located in Herveld, The Netherlands. The distribution of animals across trials and pens is in Table 5.1. Each pen had an approximately equal number of males and females. Offspring of a given sire were housed mostly in the same pen but each pen had offspring of multiple sires. Pens mostly had either PB or CB animals. An outlier analysis was done separately for PB and CB animals and separately for each day of measurement. Observations with standard deviations more than 3.5 away from the mean were considered as outliers and removed, which resulted in 4687 PB and 10,585 CB records on BW7 and 4471 PB and 10,272 CB records on BW35 (Table 5.2). The number of animals with observations (N_{PB} for PB and N_{CB} for CB animals) was smaller for BW35 than for BW7 because some animals did not survive until 35 days.

Table 5.1 Distribution of animals across trials and pens for body weight measured around 7 days (BW7)

Trial	Pen 1	Pen 2	Pen 3	Pen 4	NA*	Total
1	654	235	404	627	0	1920
2	821	0	829	860	0	2510
3	1281	1117	1122	1225	55	4800
4	1275	662	514	895	0	3346
5	1187	204	213	1092	0	2696
Total	5218	2218	3082	4699	55	15272

The distribution of animals across trials and pens for body weight measured around 35 days (BW35) was very similar. *Number of animals with unknown pen

Table 5.2 Summary statistics for body weight measured around 7 (BW7) and 35 days (BW35)

		Number (N)	N sires	N dams	Mean	SD
BW7 (kg)	Purebreds	4687	142	628	176	25
	Crossbreds	10,585	156	1028	179	23
	Total	15,272	161*	1656		
BW35 (kg)	Purebreds	4471	140	623	2066	303
	Crossbreds	10272	156	1027	2090	302
	Total	14743	161*	1650		

Statistics are presented for PB and CB data, separately. SD: standard deviations
*Total number of sires for all purebred and crossbred animals.

Markers with more than 1% inconsistent genotypes between derived parent-offspring pairs were removed and any remaining inconsistencies were set to missing. All missing genotypes of PB and CB animals were imputed simultaneously with FImpute (Sargolzaei *et al.* 2014). Genotypes of the parents and grandparents were used to assign the breed-of-origin of alleles in the CB animals but were not included in the trait analyses. After assigning breed-of-origin, we removed markers if they had a minor allele frequency lower than 0.005 in either the genotype file or in the breed-of-origin file. These edits resulted in 50,960 markers that were used in the trait analyses.

5.2.2 Assigning breed-of-origin of alleles

The breed-of-origin of alleles in the A(BC) crossbreds was derived with the BOA approach (Sevillano *et al.* 2016; Vandenplas *et al.* 2016). In short, the BOA approach consists of (1) simultaneously phasing genotypes of PB and CB animals with

AlphaPhase 1.1 by using pedigree information (Hickey *et al.* 2011), (2) collecting a library of haplotypes for each line using phased haplotypes of the PB lines, and (3) assigning the breed-of-origin of alleles in the CB animals. With this approach, 49.5% of the alleles were assigned to sire line A, which is close to the expected 50%. The full procedure and results are described in Calus *et al.* (2018).

5.2.3 Statistical model

The BW7 and BW35 phenotypes were analysed separately with a bivariate model that treats PB and CB performance as separate but correlated traits. We compared four models that differed in the relationship matrix used. The general bivariate model can be written as (Wei and van der Werf 1995; Karoui *et al.* 2012):

$$\begin{bmatrix}\mathbf{y}_{\mathrm{PB}}\\ \mathbf{y}_{\mathrm{CB}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}_{\mathrm{PB}} & 0\\ 0 & \mathbf{X}_{\mathrm{CB}}\end{bmatrix}\begin{bmatrix}\mathbf{b}_{\mathrm{PB}}\\ \mathbf{b}_{\mathrm{CB}}\end{bmatrix} + \begin{bmatrix}\mathbf{L_{PB}} & 0\\ 0 & \mathbf{L_{CB}}\end{bmatrix}\begin{bmatrix}\mathbf{m_{PB}}\\ \mathbf{m_{CB}}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{\mathrm{PB}} & 0\\ 0 & \mathbf{Z}_{\mathrm{CB}}\end{bmatrix}\begin{bmatrix}\mathbf{u}_{\mathrm{PB}}\\ \mathbf{u}_{\mathrm{CB}}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_{\mathrm{PB}}\\ \mathbf{e}_{\mathrm{CB}}\end{bmatrix}, \quad \textbf{5.2}$$

where $\mathbf{y}$ is a vector of phenotypes, $\mathbf{b}$ is a vector of fixed effects (breed × trial × pen × sex × age at measurement), with 85 (BW7) and 103 (BW35) levels, $\mathbf{X}$ is the design matrix of fixed effects, $\mathbf{m}$ is a vector of length equal to the total number of BC dams that contains (non-genetic) maternal effects with incidence matrix $\mathbf{L}$, $\mathbf{u}$ is a vector of length $(N_{PB} + N_{CB})$ that contains additive genetic effects with incidence matrix $\mathbf{Z}$, and $\mathbf{e}$ is a vector of random residuals. Subscripts denote whether the terms relate to PB or CB performance. The distribution of maternal effects was $\begin{bmatrix}\mathbf{m}_{\mathrm{PB}}\\ \mathbf{m}_{\mathrm{CB}}\end{bmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathbf{I}\sigma^2_{m,PB} & \mathbf{0}\\ 0 & \mathbf{I}\sigma^2_{m,CB}\end{bmatrix}\right)$, where $\sigma^2_{m,PB}$ ($\sigma^2_{m,CB}$) is the maternal variance in the PB (CB) animals, and $\mathbf{I}$ is an identity matrix. Note that these maternal effects are not genetic effects, but permanent environmental effects. The distribution of additive genetic effects for PB ($\mathbf{u}_{\mathrm{PB}}$) and CB performance ($\mathbf{u}_{\mathrm{PB}}$) was:

$$\begin{bmatrix}\mathbf{u}_{\mathrm{PB}}\\ \mathbf{u}_{\mathrm{CB}}\end{bmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\sigma^2_{a,PB} & \sigma_{PB,CB}\\ \sigma_{PB,CB} & \sigma^2_{a,CB}\end{bmatrix} \otimes \mathbf{K}\right), \quad \textbf{5.3}$$

where $\sigma^2_{a,PB}$ ($\sigma^2_{a,CB}$) is the additive genetic variance in the PB (CB) animals, $\sigma_{PB,CB}$ is the genetic covariance between PB and CB performance, and $\mathbf{K}$ is the relationship matrix between all animals, which differed between models. This parameterization yields additive genetic effects for both PB and CB performance of all animals. The distribution of residuals was $\begin{bmatrix}\mathbf{e}_{\mathrm{PB}}\\ \mathbf{e}_{\mathrm{CB}}\end{bmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathbf{I}\sigma^2_{e,PB} & \mathbf{0}\\ 0 & \mathbf{I}\sigma^2_{e,CB}\end{bmatrix}\right)$, where $\sigma^2_{e,PB}$ ($\sigma^2_{e,CB}$) is the residual variance in the PB (CB) animals. Concerning the fixed effects, we used the full interaction between effects (breed × trial × pen × sex × age at

measurement), because males and females (in PB and CB animals) may have different growth rates (breed × sex × age at measurement), pens may have housed different groups of animals across trials (trial × pen), and the number of degrees of freedom (maximum 103) needed was acceptable for the size of this dataset.

Variance components were estimated by restricted maximum likelihood (REML) using the MTG2 software (Lee and van der Werf 2016). From the estimated variance components (indicated by ^), the estimate of r_{pc} was computed as:

$$\hat{r}_{pc} = \frac{\hat{\sigma}_{PB,CB}}{\hat{\sigma}_{a,PB}\hat{\sigma}_{a,CB}} \quad \textbf{5.4}$$

We compared estimates obtained from four models that use different relationship matrices, and we assessed model performance by comparing the standard errors and likelihoods of these models.

5.2.4 Relationship matrices

We compared four models that use different relationship matrices (i.e., that replace $\mathbf{K}$ in Equation 5.3: (1) based on pedigree ($\mathbf{A}$; PED), (2) based on pedigree ignoring dams of CB animals ($\mathbf{A}_{BOA}$; PED_BOA), (3) based on marker genotypes ($\mathbf{G}$; GEN), and (4) based on marker alleles with sire origin ($\mathbf{G}_{BOA}$; GEN_BOA). We included PED_BOA because it only fits the additive genetic effects for CB performance that are contributed by the sire line.

The $\mathbf{A}$ and $\mathbf{A}_{BOA}$ matrices were constructed from pedigree information, which was available for all animals with phenotypes, up to the generation of their grandparents. A single base population was assumed for all PB lines (i.e. no genetic groups were included). With $\mathbf{A}$, the full pedigree was used, whereas with $\mathbf{A}_{BOA}$, the dams of CB animals were set to missing. In addition, we set all the self-relationships of CB animals in $\mathbf{A}_{BOA}$ equal to 0.5 (Garcia-Cortes and Toro 2006). As such, PED_BOA is the pedigree equivalent of GEN_BOA. The $\mathbf{G}$ matrix was constructed following the multi-breed genomic relationship matrix of Wientjes *et al.* (2017)

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{\mathrm{PB}} & \mathbf{G}_{\mathrm{PB-CB}} \\ \mathbf{G}_{\mathrm{PB-CB}} & \mathbf{G}_{\mathrm{CB}} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\mathbf{M}_{\mathrm{PB}}\mathbf{M}'_{\mathrm{PB}}}{\sum 2p_j^{PB}(1-p_j^{PB})} & \dfrac{\mathbf{M}_{\mathrm{PB}}\mathbf{M}'_{\mathrm{CB}}}{\sqrt{\sum 2p_j^{PB}(1-p_j^{PB})}\sqrt{\sum 2p_j^{CB}(1-p_j^{CB})}} \\ \dfrac{\mathbf{M}_{\mathrm{CB}}\mathbf{M}'_{\mathrm{PB}}}{\sqrt{\sum 2p_j^{PB}(1-p_j^{PB})}\sqrt{\sum 2p_j^{CB}(1-p_j^{CB})}} & \dfrac{\mathbf{M}_{\mathrm{CB}}\mathbf{M}'_{\mathrm{CB}}}{\sum 2p_j^{CB}(1-p_j^{CB})} \end{bmatrix}, \quad \textbf{5.5}$$

where $\mathbf{M}_{\mathrm{PB}}$ ($\mathbf{M}_{\mathrm{CB}}$) is a centred marker genotype matrix of PB (CB) animals, and p_j^{PB} (p_j^{CB}) is the allele frequency of marker j in PB (CB) animals. We used the line-specific allele frequencies to separately centre the genotype matrices $\mathbf{M}_{\mathrm{PB}}$ and $\mathbf{M}_{\mathrm{CB}}$. The $\mathbf{G}_{BOA}$-matrix was constructed following Sevillano *et al.* (2017) as

$$\mathbf{G}_{BOA} = \begin{bmatrix} \mathbf{G}_{BOA,\mathrm{PB}} & \mathbf{G}_{BOA,\mathrm{PB-CB}} \\ \mathbf{G}_{BOA,\mathrm{PB-CB}} & \mathbf{G}_{BOA,\mathrm{CB}} \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{M}_{\mathrm{PB}}\mathbf{M}'_{\mathrm{PB}}}{\sum 2p_j(1-p_j)} & \frac{\mathbf{M}_{\mathrm{PB}}\mathbf{T}'_{\mathrm{CB}}}{\sum 2p_j(1-p_j)} \\ \frac{\mathbf{T}_{\mathrm{CB}}\mathbf{M}'_{\mathrm{PB}}}{\sum 2p_j(1-p_j)} & \frac{\mathbf{T}_{\mathrm{CB}}\mathbf{T}'_{\mathrm{CB}}}{\sum 2p_j(1-p_j)} \end{bmatrix}, \quad \textbf{5.6}$$

where $\mathbf{T}_{\mathrm{CB}}$ is a centred marker allele matrix of CB animals, with a value of $(0 - p_j)$ if the reference allele was inherited from the PB line, and a value of $(1 - p_j)$ if the alternative allele was inherited, where p_j is the frequency of the alternative allele at marker j, which was calculated as the total number of alternative alleles in the PB and CB animals that were inherited from the PB line, divided by the total number of PB alleles in these animals. Note that the resulting $\mathbf{G}_{BOA}$ matrix is similar to the marker-based partial relationship matrix of Christensen *et al.* (2014), with a scaling factor of $\sum 2p_j(1-p_j)$.

The expected value of diagonal elements for CB animals in $\mathbf{G}_{BOA}$ and $\mathbf{A}_{BOA}$ is 0.5. The phenotypic variance of CB performance with PED_BOA and GEN_BOA was therefore computed as $0.5\sigma^2_{a,CB} + \sigma^2_{m,CB} + \sigma^2_{e,CB}$.

5.2.5 Scaling of relationship matrices

With pedigree-based models, the population to which the variance components refer is the population of the founders of the pedigree. However, with genotype-based models, the reference population is, in most cases, the group of genotyped individuals, because $\mathbf{G}$ and $\mathbf{G}_{BOA}$ were constructed using the allele frequencies in the genotyped group. Thus, estimated variance components from pedigree- and genotype-based models are not directly comparable, because they refer to a different population (Legarra 2016). To let the variance components from different models refer to the same (arbitrary) population, all relationship matrices were adjusted as

$$\mathbf{K}' = \begin{bmatrix} \frac{\mathbf{K}_{11}}{D_{k_1}} & \frac{\mathbf{K}_{12}}{\sqrt{D_{k_1}}\sqrt{D_{k_2}}} \\ \frac{\mathbf{K}_{21}}{\sqrt{D_{k_1}}\sqrt{D_{k_2}}} & \frac{\mathbf{K}_{22}}{D_{k_2}} \end{bmatrix},$$

where $\mathbf{K}_{11}$ denotes relationships among the PB animals, $\mathbf{K}_{22}$ denotes relationships among the CB animals, and $\mathbf{K}_{12}$ and $\mathbf{K}_{21}$ denote the relationships between PB and

CB animals, as defined in Equations (5) and (6). Scalar D_{k_1} (D_{k_2}) is the scaling factor of PB (CB) animals, which was defined as

$$D_{k_x} = \overline{Diag(\mathbf{K}_x)} - \overline{\mathbf{K}}_x,$$

where $\overline{Diag(\mathbf{K}_x)}$ is the mean of off-diagonals in $\mathbf{K}_x$ and $\overline{\mathbf{K}}_x$ is the mean of all elements in $\mathbf{K}_x$. This scaling procedure is equivalent to multiplying estimated variance components from models with unscaled relationship matrices by the appropriate scaling factors, as proposed by Legarra (2016). For models that considered the breed-of-origin of alleles, the expected value of D_{k_2} was close to 0.5, so we used $2D_{k_2}$ instead of D_{k_2} as a scaling factor in these models.

5.3 Results

Detailed information and estimates from all models are in Table A 5.1 (Appendix). The phenotypic variance for BW7 was around 363 g^2 for PB performance and 291 g^2 for CB performance, whereas for BW35, it was around 37,048 g^2 for PB performance and 33,455 g^2 for CB performance. The estimated phenotypic variance was similar across models, thus we present variances instead of their ratio to phenotypic variance. Estimates of r_{pc} and of the additive genetic covariance were larger for BW35 than for BW7 (Figure 5.2). For BW7, the estimate of the additive genetic variance was smaller for PB performance than for CB performance, except with PED_BOA (Figure 5.3). For BW35, estimates of the additive genetic variance and heritabilities were consistently larger for PB performance. Differences in estimates between models, were roughly similar for BW7 and BW35. For the sake of brevity, in the following, the description of results applies to both traits, unless stated otherwise. We will refer to PED and PED_BOA as pedigree-based models, and to GEN and GEN_BOA as genotype-based models. In addition, we will refer to PED and GEN as models that ignore breed-of-origin, and to PED_BOA and GEN_BOA as models that consider breed-of-origin.

5.3.1 Pedigree versus genomic relationship information

Estimates of r_{pc} were larger with genotype-based models than with pedigree-based models, particularly for BW7 (Figure 5.2, Table 5.3). However, estimates of the additive genetic covariance were smaller with genotype-based models than with pedigree-based models, except for BW7 and GEN_BOA versus PED_BOA. For PB performance, estimates of the additive genetic variance were smaller with genotype-based models than with pedigree-based models. For CB performance,

estimates of the additive genetic variance were similar with genotype-based models and pedigree-based models, except for BW7 and GEN_BOA versus PED_BOA, for which the additive genetic variance was larger with GEN_BOA (Figure 5.3). Because of these differences in estimates of additive genetic variance, the product of estimates of additive genetic standard deviations in the denominator of r_{pc} was smaller with genotype-based models than with pedigree-based models. Estimates of the maternal variance of PB performance were larger with genotype-based models than with pedigree-based models, while for CB performance, estimates of maternal variance were similar for both types of models (Figure 5.3).

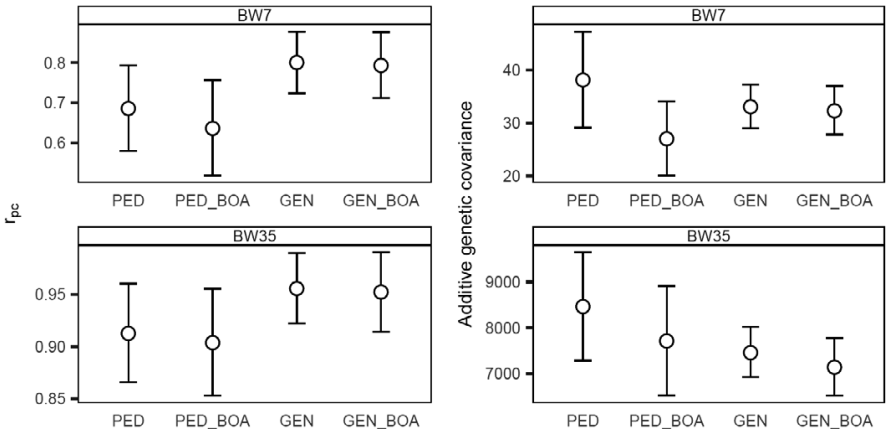

Figure 5.2 Estimates of r_{pc} (left) and additive genetic covariance (right) from four models. Traits are body weight measured around 7 (BW7) and 35 days (BW35). Error bars represent standard errors reported by the MTG2 software.

Table 5.3 Estimates of the purebred-crossbred genetic correlation (r_{pc}) and their standard errors for body weight measured around 7 (BW7) and 35 days (BW35), from four models.

	BW7		BW35	
Model	Estimate	Standard error	Estimate	Standard error
PED	0.69	0.11	0.91	0.05
PED_BOA	0.64	0.12	0.90	0.05
GEN	0.80	**0.08**	0.96	**0.03**
GEN_BOA	0.79	**0.08**	0.95	0.04

The standard errors are those reported by MTG2. The smallest standard errors per trait are in bold. Models used a relationship matrix based on pedigree (PED), based on pedigree ignoring dams of CB animals (PED_BOA), based on marker genotypes (GEN), or based on marker alleles with sire origin (GEN_BOA).

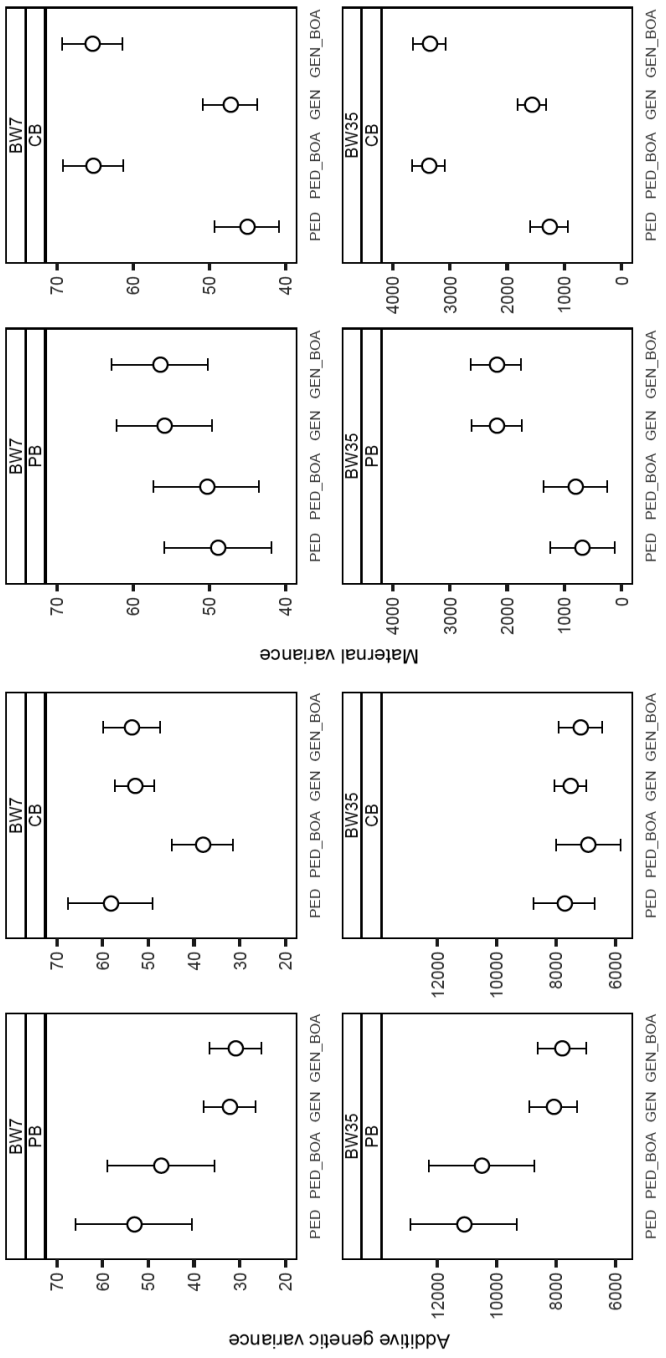

Figure 5.3 Estimates of additive genetic variance (left) and maternal variance (right) for purebred and crossbred body weight. Traits are body weight measured around 7 (BW7) and 35 days (BW35). Estimates are shown from four models. Error bars represent standard errors reported by the MTG2 software.

5.3.2 Ignoring versus considering breed-of-origin

Estimates of r_{pc} from models that ignored or considered breed-of-origin were similar (Figure 5.2, Table 5.3). With pedigree-based models, estimates of the additive genetic covariance were smaller when breed-of-origin was considered, whereas with genotype-based models, estimates were similar. For PB performance, there were almost no differences in estimates of additive genetic variance and maternal variance (Figure 3.3) between models that considered or ignored breed-of-origin. However, for CB performance, models that considered breed-of-origin had a larger estimate of maternal variance than models that did not. With pedigree-based models, the estimate of additive genetic variance of CB performance was smaller when breed-of-origin was considered than when it was not. However, with genotype-based models the estimate of additive genetic variance of CB performance was similar between models that considered or ignored breed-of-origin.

Table 5.4 Likelihoods from four models for body weight measured around 7 (BW7) and 35 days (BW35)

	BW7		BW35	
Model	Likelihood	relative to PED	Likelihood	relative to PED
PED	-50303.6064		-83374.8368	
PED_BOA	-50324.9275	-21.321	-83397.1397	-22.303
GEN	**-50143.4092**	160.197	**-83116.5065**	258.330
GEN_BOA	-50254.9253	48.681	-83275.2876	99.549

The likelihoods are those reported by MTG2. The largest likelihoods per trait are in bold. Models used a relationship matrix based on pedigree (PED), based on pedigree ignoring dams of CB animals (PED_BOA), based on marker genotypes (GEN), or based on marker alleles with sire origin (GEN_BOA).

5.3.3 Likelihoods and standard errors

For both traits, model GEN had the largest likelihood, followed by GEN_BOA, PED, and PED_BOA (Table 5.4). Likelihoods were larger for genotype-based methods than for pedigree-based methods, while considering breed-of-origin unexpectedly reduced likelihoods compared to ignoring breed-of-origin. In addition to the best fit, model GEN also gave the smallest standard error of estimates of r_{pc}, followed by GEN_BOA, PED, and PED_BOA (Figure 5.2; Table 5.3). In general, the standard errors of estimates or r_{pc} were smaller with genotype-based methods than with pedigree-based methods. The standard errors of estimates of variance components of CB performance were slightly larger with models that considered breed-of-origin compared to models that did not, while there were no differences in standard errors

for estimates of variance components of PB performance (Figure 5.3 and see Table A 5.1, Appendix).

5.4 Discussion

This study aimed at comparing models that estimate PB and CB genetic parameters of body weight in broiler chicken. We were particularly interested in the estimation of r_{pc}, because the value of r_{pc} allows breeders to determine whether the use of CB information in the breeding program will increase genetic gain of CB performance, compared to a situation where only PB information is used. Our results showed that, for our population, r_{pc} estimates were 5 to 25% larger with genotype-based models than with pedigree-based models. Moreover, r_{pc} estimates were similar (max. 7% difference) with models that consider breed-of-origin and for models that ignore breed-of-origin. Genotype-based models had larger likelihoods and estimates with smaller standard errors than pedigree-based models, which was in line with expectations. This suggests that, although our results are not conclusive, r_{pc} was underestimated with pedigree-based models in this study.

Estimates of r_{pc} were between 0.64 and 0.80 for BW7 and between 0.90 and 0.96 for BW35. To our knowledge, this is the first time that r_{pc} are estimated for body weight in broilers. It should be noted that, in this study, PB and CB animals were housed in the same environment. As such, our estimates provide an upper bound for values of r_{pc} in situations where PB animals are housed in a breeding nucleus environment and CB animals in a commercial herd environment. Nevertheless, our estimates are similar to estimates from the literature on egg production traits in laying hens, for which estimates of r_{pc} ranged from 0.62 to 0.83 (Wei and van der Werf 1995). In pigs, the average estimate of r_{pc} for growth-related traits was lower (~0.6) (Wientjes and Calus 2017).

With an r_{pc} larger than ~0.7, the accuracy of predicting breeding values for CB performance is not expected to substantially improve when CB data instead of PB data is used (Van Grevenhof and Van Der Werf 2015). An empirical study on pigs also showed that, with an r_{pc} of about 0.90, replacing PB data with CB data did not improve prediction accuracy (Hidalgo *et al.* 2016). However, these results cannot be extrapolated directly to the current study, because differences in accuracy also depend on the number of phenotypic records available from the PB and CB populations and on the strength of relationships between the reference population and selection candidates (Van Grevenhof and Van Der Werf 2015; Hidalgo *et al.* 2016; Xiang *et al.* 2016). In addition, information on PB performance may be more

valuable than information on CB information, because the former may have been measured on the selection candidates themselves, whereas the latter can only be measured on relatives. Nevertheless, we expect that the use of CB instead of PB data will not substantially increase the accuracy of predicted breeding values for CB body weight in the current dataset, due to the high r_{pc}. A detailed investigation of the benefit of using CB instead of PB data for the accuracy of predicted breeding values will be investigated in a follow-up study.

Heritability estimates ranged from 0.09 to 0.20 for BW7 and from 0.21 to 0.30 for BW35. To our knowledge, heritability estimates for body weight at seven days of age (BW7) have not been reported before. Our heritability estimates for BW35 were similar to those reported by Koerhuis and Thompson (1997), Mulder *et al.* (2009) and Maniatis *et al.* (2013). In contrast, our heritability estimates for BW35 were lower than those reported by Kapell *et al.* (2012) and Rekaya *et al.* (2013). Estimates of the ratio of maternal to phenotypic variance (m^2) ranged from 0.13 to 0.22 for BW7 and from 0.02 to 0.10 for BW35. These results match with the general belief that maternal effects decrease with age. Estimates of m^2 for BW35 from the literature ranged from 0.02 to 0.05 (Koerhuis and Thompson 1997; Kapell *et al.* 2012; Maniatis *et al.* 2013; Rekaya *et al.* 2013) and were somewhat smaller than our estimates, which may be due to the use of models that consider breed-of-origin in our study, where part of the genetic variance that is not captured moves to the non-genetic maternal variance.

5.4.1 Pedigree versus genomic relationship information

Estimates of r_{pc} were larger with genotype-based than with pedigree-based models, but the estimate of additive genetic covariance was often smaller with genotype-based models than with pedigree-based models, so the difference in r_{pc} estimates was the result differences in both additive genetic variances and covariance. The estimate of the additive genetic variance of PB performance was slightly larger with pedigree-based than with genotype-based models, while the estimate of the maternal variance of PB performance was smaller with pedigree-based models. First, the difference in variance estimates for PB performance may be due in part to bias in the genomic relationships that are estimated with markers (Yang *et al.* 2010). To account for sources of bias when $\mathbf{G}$ is used, Goddard *et al*. (Goddard *et al.* 2011) proposed to regress $\mathbf{G}$ towards $\mathbf{A}$. However, for our data, this procedure neither changed the relationships in $\mathbf{G}$ substantially, nor changed the estimates of variance components (results not shown). Furthermore, in contrast to PB performance, the additive genetic variance of CB performance was similar with

pedigree-based and genotype-based models. Thus, we chose not to regress $\mathbf{G}$ towards $\mathbf{A}$. Second, the estimate of maternal variance may be more accurate with genotype-based than with pedigree-based models because genotype-based models may be more efficient at disentangling non-genetic maternal effects from the maternal component of an individual's additive genetic effect (Lee *et al.* 2010; Berenos *et al.* 2014). However, in contrast to PB performance, estimates of the additive genetic and maternal variances for CB performance were similar with pedigree-based and genotype-based models. Thus, it remains unclear why the differences in estimates of variances for PB performance between genotype-based and pedigree-based models were not observed for CB performance.

5.4.2 The effect of considering breed-of-origin of alleles

For PB performance, estimates of variance components from models that ignored or considered breed-of-origin of alleles were similar, which is not surprising, because relationships between PB animals are the same regardless of whether breed-of-origin is considered or not. However, for CB performance, the estimate of the maternal variance was much larger with models that considered breed-of-origin. In these models, only alleles inherited from the sires were used to describe the variation in relationships between CB offspring. Thus, the genetic part of these models only captured the additive genetic variance of CB performance that is caused by the PB sire line. As a result, the non-genetic maternal effect absorbed most of the genetic variance caused by the BC dams. In contrast, with models that ignore breed-of-origin, alleles inherited from dams describe additional genetic covariation between CB offspring. Thus, the genetic components of these models also capture some of the additive genetic variance caused by the BC dams and, as a result, the variance explained by the maternal effect was smaller with models that ignored breed-of-origin than with models that considered breed-of-origin.

In spite of the differences in estimates of maternal variance between GEN and GEN_BOA, the estimate of additive genetic variance in CB performance ($\sigma^2_{a,CB}$) was similar between GEN and GEN_BOA. Thus, we hypothesized that either (1) the contribution of alleles that originated from the sire line to $\sigma^2_{a,CB}$ is equal to the contribution of alleles that originated from the BC dams, or (2) the relationships between the sires and BC dams contributed little to the estimate of $\sigma^2_{a,CB}$ (because the sires and BC dams were distantly related) and the paternal relationships dominated the estimate of $\sigma^2_{a,CB}$.

To test the first hypothesis, we analysed CB performance with a univariate model that fitted random sire and random dam effects separately, each with their

own BOA matrix. This model yielded two estimates of $\sigma^2_{a,CB}$, one for the sire line and one for the BC dams, which showed that the contribution of the BC dams to the estimate of $\sigma^2_{a,CB}$ was larger than the contribution of the sire line (see Table A 5.2, Appendix). The first hypothesis was therefore rejected. To test the second hypothesis, we compared estimates of $\sigma^2_{a,CB}$ from the aforementioned univariate BOA approach with estimates from a univariate GEN approach using only CB performance (GEN_CB). With GEN_CB, the estimate of $\sigma^2_{a,CB}$ also depends on genetic covariances between sires and dams because GEN_CB merges alleles from both lines into a single $\mathbf{G}$ matrix. However, we observed that the average of the $\sigma^2_{a,CB}$ estimates of the sire line and BC dams from the BOA approach was close to the estimate of $\sigma^2_{a,CB}$ with GEN_CB (8389 vs 8410; [see Table A 5.2, Appendix]) , which suggests that relationships between sires and dams contributed little to the likelihood or to the estimate of $\sigma^2_{a,CB}$ with GEN_CB. Indeed, there was almost no variance in genomic relationships between sires and dams and, as a result, relationships between sires and between dams dominated the estimate of $\sigma^2_{a,CB}$ with GEN_CB. Similar to $\sigma^2_{a,CB}$, the estimate of the additive genetic covariance between PB and CB performance was the same with GEN and GEN_BOA. Hence, the estimate of additive genetic covariance with GEN is probably dominated by variation in relationships between sires and between dams. Of these, we believe that paternal relationships dominated the estimate of $\sigma^2_{a,CB}$ because the model included a non-genetic maternal effect, which is strongly confounded with the maternal part of the genetic covariance between full sibs. Hence, covariances in the BC dams that are informative for $\sigma^2_{a,CB}$ originated mainly from more distant relationships, which have a smaller impact on the likelihood than, e.g., paternal half-sib relationships. In addition, the standard error of the estimate of $\sigma^2_{a,CB}$ was larger when using dam alleles than when using sire alleles (see Table A 5.2, Appendix), which suggests that paternal relationships dominated the estimate of $\sigma^2_{a,CB}$.

5.4.3 Model usefulness

This study focused on the estimation of variance components and r_{pc} using different models based on estimated standard errors and model fit. However, it should be noted that the model with the best fit does not necessarily yield the most accurate predicted breeding values (Erbe *et al.* 2012), which shall be investigated in a follow-up study. Nevertheless, results showed that genotype-based models had a better model-fit and smaller estimated standard errors than pedigree-based models. Thus, genotype-based models may be preferred over pedigree-based models to estimate r_{pc}, even when the PB and CB animals are closely related. The benefit of

genotype-based models may be slightly larger when the PB and CB animals are less related or when pedigree information is difficult to obtain. However, reported standard errors of estimates should be used with care. For example, the assumption in model GEN that all alleles in the CB animals originate from the same line is incorrect, which can lead to unreliable estimates of standard errors.

Models that consider breed-of-origin of alleles had smaller likelihoods than models that ignore it, which is somewhat unexpected for GEN versus GEN_BOA. With GEN, relationships between PB and CB animals are based on alleles from both the sires and dams and alleles of the dams in the PB and CB animals are assumed to have the same origin. Thus, the PB-CB relationships in $\mathbf{G}$ may be less accurate than the PB-CB relationships in $\mathbf{G}_{BOA}$, which may decrease the likelihood of GEN. Nevertheless, estimates of r_{pc} with GEN and GEN_BOA were similar, which suggests that violation of model assumptions with GEN had only minor effects on the estimate of r_{pc}. In addition, GEN may have an advantage over GEN_BOA, because the assignment of the BOA is probably not without error, which may affect estimates of variance components.

The GEN_BOA model that we used in this study does not explicitly fit a genetic component for the maternal alleles in the CB animals and, hence, does not allow for a covariance of allele effects from the dams with those from the sire line. In addition, we did not use phenotypes from the BC dams, such that we were not able to estimate r_{pc} between the A(BC) crossbreds and the BC dams. A more complete model would use phenotypes and phased genotypes from the A(BC) crossbreds and its three parental lines (A, B and C), model these phenotypes as four separate traits, and allow covariances between these traits (Christensen *et al.* 2015). Although such a model is more sophisticated and complete, we do not expect that it would result in different estimates of r_{pc} between the CB and its sire line, because the parental lines were genetically distant.

In spite of differences in standard errors and likelihoods between models, we were not able to establish which estimates were closest to the true values (i.e., the genetic correlation at the causal loci) because this value is unknown. Gianola *et al.* (2015) showed that estimates of genetic correlations using marker information may not necessarily reflect the true genetic correlation at causal loci because of imperfect linkage disequilibrium between markers and QTL. However, simulation studies have suggested that genotype-based models result in unbiased estimates of genetic correlations when relationships at causal loci are accurately predicted by the markers (Wientjes *et al.* 2018). Further research is needed to establish whether

these results also apply to estimation of r_{pc} and which of the models presented in this study yields the most accurate estimate of r_{pc}.

5.5 Conclusions

This work presents the first comparison of estimated r_{pc} and variance components of body weight in broilers, using pedigree-based and genotype-based models, where the breed-of-origin of alleles was either ignored or considered. Estimates of r_{pc} ranged from 0.64 to 0.80 for BW7 and from 0.90 to 0.96 for BW35. Genotype-based models resulted in larger estimates of r_{pc} than pedigree-based models and are preferred for estimating r_{pc} because they resulted in smaller standard errors of estimates and had better model fit than pedigree-based models. Considering breed-of-origin of alleles did not affect estimates of r_{pc}, probably because the parental breeds of the CB animals were distantly related but could result in different estimates of r_{pc} when the parental breeds are more closely related, or when the amount of data is limited.

5.6 Acknowledgements

The authors thank the Netherlands Organisation of Scientific Research (NWO) and the Breed4Food consortium partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin for their financial support. Special thanks go to Cobb Europe for providing the data. The use of the HPC cluster was made possible by CAT-AgroFood (Shared Research Facilities Wageningen UR).

5.7 Appendix

Table A 5.1 Estimates (est) of variance components and purebred-crossbred genetic correlations of body weight around 7 (BW7) and 35 days (BW35) for four models.

		BW7				BW35			
	model	PB		CB		PB		CB	
		est	se	est	se	est	se	est	se
σ^2_m	PED	49	7	45	4	690	563	1270	326
	PED_BOA	50	7	65	4	808	562	3371	287
	GEN	56	6	47	4	2188	436	1573	248
	GEN_BOA	57	6	65	4	2193	439	3358	283
σ^2_a	PED	53	13	58	9	11115	1799	7737	1030
	PED_BOA	47	12	38	7	10517	1767	6924	1082
	GEN	32	6	53	4	8104	811	7529	538
	GEN_BOA	31	6	54	6	7815	818	7197	734
σ^2_e	PED	262	9	186	6	25122	1162	24438	671
	PED_BOA	265	9	206	4	25423	1150	26577	505
	GEN	275	7	189	3	27064	701	24367	415
	GEN_BOA	276	7	202	3	27152	706	26539	423
σ^2_p	PED	364	NA	290	NA	36928	NA	33445	NA
	PED_BOA	362	NA	291	NA	36748	NA	33410	NA
	GEN	364	NA	289	NA	37356	NA	33469	NA
	GEN_BOA	363	NA	294	NA	37160	NA	33495	NA
m^2	PED	0.13	0.018	0.16	0.014	0.02	0.015	0.04	0.010
	PED_BOA	0.14	0.018	0.22	0.012	0.02	0.015	0.10	0.008
	GEN	0.15	0.016	0.16	0.011	0.06	0.011	0.05	0.007
	GEN_BOA	0.16	0.016	0.22	0.012	0.06	0.012	0.10	0.008
h^2	PED	0.15	0.034	0.20	0.030	0.30	0.045	0.23	0.029
	PED_BOA	0.13	0.032	0.13	0.021	0.29	0.045	0.21	0.028
	GEN	0.09	0.015	0.18	0.014	0.22	0.019	0.23	0.014
	GEN_BOA	0.09	0.015	0.18	0.018	0.21	0.020	0.21	0.018

		BW7		BW35	
		est	se	est	se
$\sigma_{PB,CB}$	PED	38	9	8467	1183
	PED_BOA	27	7	7714	1195
	GEN	33	4	7466	548
	GEN_BOA	32	5	7143	630
r_{pc}	PED	0.69	0.107	0.91	0.047
	PED_BOA	0.64	0.119	0.90	0.051
	GEN	0.80	0.077	0.96	0.034
	GEN_BOA	0.79	0.082	0.95	0.038

Table A 5.2 Estimates of variance components for CB performance of BW35 from models that either fit a single G matrix (GEN_CB), or that separately fit a genetic sire component and a genetic dam component with two BOA matrices.

		2 BOA		GEN_CB	
		estimate	se	estimate	se
σ_a^2	sire	6931	787	***8410***	635
	dam	9847	974		
σ_m^2		1443	257	1596	252
σ_e^2		23770	438	24414	422
Likelihood		-57922.9749		-57925.4123	
average σ_a^2 sire and dam:		***8389***			

6

Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles

Pascal Duenk[1], Mario P.L. Calus[1], Yvonne C.J. Wientjes[1], Vivian P. Breen[2], John M. Henshall[2], Rachel Hawken[2] and Piter Bijma[1]

[1]Animal Breeding and Genomics, Wageningen University and Research, 6700 AH Wageningen, the Netherlands

[2]Cobb-Vantress Inc., Siloam Springs, AR 72761-1030, USA

Genetics Selection Evolution (2019) 51:38

Abstract

Background

Pig and poultry breeding programs aim at improving crossbred (CB) performance. Selection response may be suboptimal if only purebred (PB) performance is used to compute genomic estimated breeding values (GEBV) because the genetic correlation between PB and CB performance (r_{pc}) is often lower than 1. Thus, it may be beneficial to use information on both PB and CB performance. In addition, the accuracy of GEBV of PB animals for CB performance may improve when the breed-of-origin of alleles (BOA) is considered in the genomic relationship matrix (GRM). Thus, our aim was to compare scenarios where GEBV are computed and validated by using (i) either CB offspring averages or individual CB records for validation, (ii) either a PB or CB reference population, and (iii) a GRM that either accounts for or ignores BOA in the CB individuals. For this purpose, we used data on body weight measured at around 7 (BW7) or 35 (BW35) days in PB and CB broiler chickens and evaluated the accuracy of GEBV based on the correlation GEBV with phenotypes in the validation population (validation correlation).

Results

With validation on CB offspring averages, the validation correlation of GEBV of PB animals for CB performance was lower with a CB reference population than with a PB reference population for BW35 (r_{pc} = 0.96), and about equal for BW7 (r_{pc} = 0.80) when BOA was ignored. However, with validation on individual CB records, the validation correlation was higher with a CB reference population for both traits. The use of a GRM that took BOA into account increased the validation correlation for BW7 but reduced it for BW35.

Conclusions

We argue that the benefit of using a CB reference population for genomic prediction of PB animals for CB performance should be assessed either by validation on CB offspring averages, or by validation on individual CB records while using a GRM that accounts for BOA in the CB individuals. With this recommendation in mind, our results show that the accuracy of GEBV of PB animals for CB performance was equal to or higher with a CB reference population than with a PB reference population for a trait with an r_{pc} of 0.8, but lower for a trait with an r_{pc} of 0.96. In addition, taking BOA into account was beneficial for a trait with an r_{pc} of 0.8 but not for a trait with an r_{pc} of 0.96.

6.1 Background

In pig and poultry breeding programs, purebred (PB) animals from different lines or breeds are mated to produce crossbred (CB) production animals. Although the aim of such breeding programs is to improve CB performance, typically, breeding values of PB selection candidates are estimated using only information on PB performance. As a result, response to selection in CB performance may be suboptimal because the genetic correlation between PB and CB performance (r_{pc}) is often lower than 1 (Wei and van der Werf 1995; Lukaszewicz *et al.* 2015; Wientjes and Calus 2017). A low r_{pc} may be due to genotype-by-environment interactions (Lutaaya *et al.* 2001; Dekkers 2007), genotype-by-genotype interactions (i.e., dominance and epistasis) in combination with differences in allele frequencies between the purebred parental lines (Wei *et al.* 1991), and/or differences in the definition of PB and CB performance traits (Lo *et al.* 1997; Zumbach *et al.* 2007).

When the r_{pc} is lower than 1, it may be beneficial to use information on both PB and CB performance to estimate breeding values of PB selection candidates. For this strategy, breeders need to be able to connect observations on CB performance to the PB selection candidates. This connection can be established with a pedigree-based relationship matrix. However, in a CB breeding scheme, breeders do not always routinely record pedigree information. In such cases, the pedigree-based relationship matrix can be replaced by a genomic relationship matrix (GRM) that is based on observed marker genotypes (VanRaden 2008). This GRM enables breeders to use a reference population that consists of animals with phenotypes and genotypes to estimate genomic estimated breeding values (GEBV) of selection candidates that only have records on genotypes (Meuwissen *et al.* 2001). When pedigree information is available, replacing the pedigree-based relationship matrix by a GRM may increase the accuracy of estimated breeding values (Hayes *et al.* 2009c). As such, this method, called genomic prediction, allows breeders to use a CB reference population to compute GEBV for CB performance of PB selection candidates (Dekkers 2007).

Simulation studies have suggested that a CB reference population may yield more accurate GEBV for CB performance than a PB reference population when the r_{pc} is lower than 0.8 (Dekkers 2007; Esfandyari *et al.* 2015a; Van Grevenhof and Van Der Werf 2015). This result was shown for situations for which the CB reference population had at least the same size as the alternative PB reference population and the selection candidates had similar relationships to the CB and the PB reference populations. In agreement with these simulation studies, Hidalgo *et al.* (2016), using

real data in pigs, found that for a trait with a high r_{pc} (~0.90), the accuracy of GEBV of PB animals for CB performance was lower with a reference population of CB compared to PB pigs. These results were not only due to a high r_{pc}, but also to the smaller number of CB pigs compared to PB pigs in the reference population, and weaker relationships of the PB selection candidates with the CB reference population than with the PB reference population (Hidalgo *et al.* 2016). In summary, the expected benefit of using a CB reference population instead of a PB reference population increases with (1) lower r_{pc}, (2) stronger relationships of the CB reference population with PB selection candidates, and (3) larger sizes of the CB reference population.

When a CB reference population is used to estimate GEBV of PB selection candidates, relationships in the GRM (i.e. $\mathbf{G}$) are often constructed while ignoring the breed-of-origin of alleles (BOA) of the CB animals. Thus, one assumes that the apparent effects of markers are the same for alleles that originate from the sire breed and the dam breed. Thus, apparent effects of markers are assumed to be equal across breeds, which may not be valid because of differences in linkage disequilibrium (LD), and/or in allele frequencies between the parental breeds (de Roos *et al.* 2008; Veroneze *et al.* 2014; Fu *et al.* 2015; Wientjes *et al.* 2015; Pengelly *et al.* 2016). In addition, actual effects at causal loci may differ between breeds due to genotype-by-environment interactions (Lutaaya *et al.* 2001; Dekkers 2007) and/or the presence of non-additive effects in combination with differences in allele frequencies (Fisher 1918; Falconer 1952). Thus, considering BOA when constructing the GRM may lead to more accurate GEBV.

Recently, a method has been developed that allows the BOA in CB animals to be determined based on phased genotypes, while taking advantage of the known crossbreeding structure (Vandenplas *et al.* 2016). This allows the construction of a partial genomic relationship matrix ($\mathbf{G}_{BOA}$) (Ibañez-Escriche *et al.* 2009; Christensen *et al.* 2014), in which relationships that involve CB animals are based only on alleles that originate from the line of selection candidates for which GEBV are estimated. Simulation studies suggested that genomic prediction models that take BOA into account may outperform models that ignore it (Ibañez-Escriche *et al.* 2009; Esfandyari *et al.* 2015a). However, this benefit of considering BOA was only observed when the CB reference population was large (4000), the number of markers was small (500), and the parental lines of CB animals were distantly related. Moreover, empirical studies on pigs suggested that taking BOA into account may increase the accuracy of GEBV only when r_{pc} and heritability are low (Lopes *et al.* 2017; Sevillano *et al.* 2017).

In summary, the use of CB information instead of PB information and taking BOA into account may be beneficial for genomic evaluation of PB animals for CB performance. Such benefits are expected when r_{pc} is low but, to date, this hypothesis has not been tested in broiler breeding programs. Furthermore, it is not yet clear how such benefits should be evaluated, i.e. how GEBV from such models should be validated. Thus, the aim of our study was to compare scenarios in which GEBV of PB animals for CB performance are computed and validated by using (i) either CB offspring averages or individual CB records for validation, (ii) either a PB or CB reference population, and (iii) a GRM that either accounts for or ignores BOA in the CB individuals. Scenarios were compared based on the correlation of GEBV with validation records (hereafter called the validation correlation) and based on the regression coefficient of validation records on GEBV (i.e. bias). For this purpose, we used data on body weight measured at around 7 (BW7) or 35 days (BW35) of age in PB and CB broilers.

6.2 Methods

Previously, in Duenk *et al.* (2019b), we estimated genetic parameters for BW7 and BW35 with data from PB and CB animals that were housed in the same environment and that originated from a common group of sires. The estimated heritability of BW7 was 0.09 for PB performance and 0.18 for CB performance, and that of BW35 was 0.22 for PB performance and 0.23 for CB performance. The estimates of r_{pc} for BW7 and BW35 were 0.80 and 0.96, respectively. Furthermore, for the CB animals in this dataset, BOA were derived by Calus *et al.* (2018), which allowed us to consider BOA for genomic prediction. In the current study, we will use the same data as from the previous study to estimate GEBV of PB animals for CB performance, with a reference population of either PB or CB animals, and validate those GEBV with either CB offspring averages or individual CB records.

We used phenotype data on body weight from male and female offspring from a PB sire line (A), and from a three-way crossbred population (A(BC)). The three-way crossbred offspring resulted from mating sires of line A with F1 dams that were a cross between dam lines B and C (BC). All PB and CB offspring came from the same generation and were generated using the same PB line A sires in order to create sufficient links between the PB and CB offspring to enable accurate estimation of r_{pc}. The dam lines used (B and C) have been selected on egg production traits, whereas the sire line A has been selected on male fertility traits, along with growth, yield, and feed efficiency. The three parent lines (A, B, C) were genetically distant, as shown by the principal component analysis in Duenk *et al.* (2019b).

Our aim was to investigate the effect of the validation records used (CB offspring averages or CB individual records) on the validation correlation and bias based on linear regression of validation records on GEBV. Our first strategy was to validate PB sire GEBV for CB performance with CB offspring averages (scenarios –A, Table 6.1). However, because the number of sires was small (161), we expected a relatively large standard error of the resulting validation correlation. Thus, our second strategy was to validate GEBV for CB performance with individual CB records (scenarios –I, Table 6.1), following Xiang *et al.* (2016). For both these validation methods, we compared the validation correlation and bias for GEBV obtained using either a PB reference population (scenarios PB-A and PB-I, Table 6.1) or a CB reference population (scenarios CB-A and CB-I, Table 6.1). With a CB reference population, we also investigated the benefit of considering the BOA (CB-A-BOA and CB-I-BOA, Table 6.1). Note that, in this study, we did not use own performance records of the purebred selection candidates, because we wanted to compare the predictive value of a CB reference population with that of PB reference population, both consisting of animals that are not closely related to the selection candidates.

Table 6.1 Overview of scenarios with information on the types of reference population, validation records, and genomic relationship matrix (GRM) that were used

Scenario[a]	Reference population	Prediction	Validation	GRM
PB-A	PB offspring	sire GEBV	Offspring averages	$\mathbf{G}$
CB-A	CB offspring	sire GEBV	Offspring averages	$\mathbf{G}$
CB-A-BOA	CB offspring	sire GEBV	Offspring averages	$\mathbf{G}_{BOA}$
PB-I	PB offspring	CB GEBV	Individual records	$\mathbf{G}$
CB-I	CB offspring	CB GEBV	Individual records	$\mathbf{G}$
CB-I-BOA	CB offspring	CB GEBV	Individual records	$\mathbf{G}_{BOA}$

[a]In the abbreviation of the scenarios, the first element indicates the reference population (PB or CB), the second element the validation record (CB offspring averages indicated by A or individual offspring records indicated by I), and a third element "BOA" is added for scenarios that consider BOA.

6.2.1 Phenotype data

For recording phenotype data, a single generation of offspring were weighed at around 7 (BW7) and 35 (BW35) days of age in five consecutive batches of similar size, with both PB and CB offspring in every batch. The five batches followed each other directly, and together spanned less than five months. Birds from the first batch hatched in June 2014, and those from the last batch hatched in November 2014. Animals that belonged to the offspring generation in one of the batches were not parents of birds in any of the other batches. Within each batch, the PB and CB

offspring were housed in three to five pens. For 16 out of 20 pen-batch combinations, at least 90% of the animals in the pen were from the same genetic group (i.e. PB or CB animals), while for the remaining pens, between 53 and 77% of the animals in the pen were from the same genetic group. Each pen had a near equal number of males and females. Each sire had most of its offspring housed in the same pen, and each pen had offspring of multiple sires. Outlier analysis was done separately per day of recording, and separately for PB and CB animals. Observations that deviated more than 3.5 standard deviations from the mean were removed. After outliers were removed, 4687 PB and 10,585 CB records remained for BW7, and 4471 PB and 10,272 CB records remained for BW35 (Table 6.2). The number of animals with observations was smaller for BW35 than for BW7, because some animals did not survive until 35 days.

Table 6.2 Summary statistics for body weight measured around 7 (BW7) and around 35 days of age (BW35)

		Number	Number of sires	Number of dams	Mean (g)	sd (g)
BW7	Purebreds	4687	142	628	176	25
	Crossbreds	10,585	156	1028	179	23
	Total	15,272	161[a]	1656		
BW35	Purebreds	4471	140	623	2066	303
	Crossbreds	10,272	156	1027	2090	302
	Total	14,743	161[a]	1650		

[a]Total number of sires for all purebred and crossbred animals.

6.2.2 Genotype data

Genotypes were collected from all PB and CB offspring with phenotypes, as well as from their potential parents and from most of their potential grandparents. Marker positions were determined based on the *Gallus gallus* 4.0 (galGal4) reference assembly. Genotype markers were removed if they were located on sex chromosomes or on the mitochondrial genome, had unknown locations, or a call rate lower than 90%. Animals were removed from the genotype data if they had a call rate lower than 90%. The remaining genotypes were used to reconstruct the pedigree, so that pedigree information was available up to the generation of the grandparents. Genotypes of the grandparents were only used to assign BOA for the animals with phenotypes. In total, there were 161 unique PB sires from line A, of which 135 sires had both PB and CB offspring, five sires had only PB offspring, and

21 sires had only CB offspring (Table 6.2). The PB offspring had 628 unique dams, whereas the CB offspring had 1028 unique dams.

We used the reconstructed pedigree to check the genotypes of each marker for Mendelian inheritance inconsistencies between all parent-offspring pairs. Markers with more than 1% inconsistent genotypes between parent-offspring pairs were removed, and for the remaining identified inconsistencies, the genotypes of parent and offspring were set to missing. No animal had more than 1% of inconsistencies across markers. All missing genotypes were imputed with FImpute (Sargolzaei *et al.* 2014). After assigning BOA, we removed markers if they had a minor allele frequency lower than 0.005 in either the genotype file or the BOA file. After these edits, 50,960 markers remained for analysis.

6.2.3 Assigning breed-of-origin of alleles

For all markers, the BOA in the CB offspring were derived with the BOA approach (Sevillano *et al.* 2016; Vandenplas *et al.* 2016). In short, the BOA approach consists of (1) simultaneously phasing genotypes of PB and CB animals with AlphaPhase 1.1 using pedigree information (Hickey *et al.* 2011), (2) collecting a library of haplotypes for each line using these phased haplotypes, and (3) assigning the BOA in the CB animals. Steps 2 and 3 were performed using in-house software. This approach resulted in 49.5% of the alleles being assigned to sire line A, which is close to the expected 50%. The full procedure and results of assigning BOA in these data are described in Calus *et al.* (2018).

6.2.4 Data selection

The available number of CB animals with phenotypes and genotypes was more than twice as large as the number of PB animals (Table 6.2). However, our aim was to compare the use of a PB reference population to that of a CB reference population of similar size. Thus, we randomly selected a set of ~4500 CB animals to be used in the analyses, while aiming for a comparable family structure in the PB data and the selected set of CB animals. To this end, we counted the number of PB full-sib families of size s (ranging from 1 to 11) and we randomly selected the same number of CB full-sib families of size s. If the available number of CB families of size s was smaller than the number of PB families of size s, all CB families of this size were selected (Table 6.3). As a result, the number of CB offspring in the selected set was 4655 for BW7 and 4445 for BW35. These numbers were only slightly smaller than the corresponding numbers of PB offspring (4687 for BW7 and 4471 for BW35).

An initial analysis revealed that the validation correlation from using a CB reference population differed substantially between randomly selected sets of CB animals. To reduce the impact of this variability on the outcome of the study, we independently sampled 100 different sets of CB animals using the procedure described above. The average fraction of CB animals that two sets had in common for each family size is in Table 6.3; the overall average fraction was 0.47.

Table 6.3 Number of full-sib families in the PB and CB offspring by family size

Family size	Number of PB families	Number of CB families		Average fraction overlap[a]
		Total	Selected	
1	1699	4406	1699	0.39
2	653	1610	653	0.40
3	276	607	276	0.46
4	117	177	117	0.66
5	46	60	46	0.77
6	13	14	13	0.93
7	6	3	3	1.00
8	2	2	2	1.00
9	1	1	1	1.00
10	0	0	0	-
11	1	0	0	-

[a]The average fraction of CB animals that two randomly selected sets of CB animals (replicates) had in common, computed per family size.

6.2.5 Genomic prediction and cross-validation populations

We ran all scenarios for each of the 100 sets of ~4500 CB animals separately, resulting in 100 replicates for each scenario. For every replicate, we used the selected set of CB animals or all PB animals with phenotypes to create the reference and validation population following the cross-validation strategy explained in the next paragraph. For scenarios denoted by CB-A and CB-A-BOA, the selected CB set was used to create reference populations, and CB offspring averages of sires were used for validation; for scenarios CB-I and CB-I-BOA, the selected CB set was used to create both the reference and validation populations; for scenario PB-I, the selected CB set was used to create the validation populations, and all PB offspring were used to create the reference populations; for scenario PB-A, all PB offspring were used to create the reference populations and CB offspring averages of sires were used for validation.

For each replicate, our aim was to minimise relationships between animals in the reference and animals in the validation population by creating five cross-validation (CV) groups. The CV groups were created so that animals in the validation population did not have offspring or paternal-half sibs in the reference population. Thus, we randomly assigned the 156 PB sires that had CB offspring to these CV groups, such that four groups had 32 sires and one group had 33 sires. All offspring were then assigned to the same CV group as their sire. For each CV group, either the sires (for validation on CB offspring averages) or the CB animals (for validation on individual CB records) in this group were used as the validation population, while either the PB or CB offspring in the remaining CV groups were used as the reference population (Table 6.1). The PB offspring of sires without CB offspring were always included in the PB reference population.

GEBV were predicted separately for BW7 and BW35 with the following univariate model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Lm} + \mathbf{Za} + \mathbf{e}, \qquad \textbf{6.1}$$

where $\mathbf{y}$ is a vector of phenotypes, $\mathbf{b}$ is a vector of fixed effects (batch × pen × sex × age at measurement) with design matrix $\mathbf{X}$, $\mathbf{m}$ is a vector of permanent environmental (maternal) effects with incidence matrix $\mathbf{L}$, $\mathbf{a}$ is a vector of additive genetic effects with incidence matrix $\mathbf{Z}$, and $\mathbf{e}$ is a vector of random residuals. The distribution of permanent environmental (maternal) effects was assumed $\mathbf{m} \sim N(0, \mathbf{I}_m\sigma_m^2)$, where σ_m^2 is the permanent environmental variance and $\mathbf{I}_m$ is an identity matrix. The distribution of additive genetic effects was assumed $\mathbf{a} \sim N(0, \mathbf{G}\sigma_a^2)$, where σ_a^2 is th0e additive genetic variance and $\mathbf{G}$ is a multi-breed genomic relationship matrix that either ignores or considers BOA ($\mathbf{G}_{BOA}$). The distribution of residuals was assumed $\mathbf{e} \sim N(0, \mathbf{I}_r\sigma_e^2)$, where σ_e^2 is the residual variance and $\mathbf{I}_r$ is an identity matrix.

For scenarios CB-A and PB-I, matrix $\mathbf{G}$ was constructed following Wientjes *et al.* (2017):

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{\mathrm{PB}} & \mathbf{G}_{\mathrm{PB-CB}} \\ \mathbf{G}_{\mathrm{CB-PB}} & \mathbf{G}_{\mathrm{CB}} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\mathbf{M}_{\mathrm{PB}}\mathbf{M}'_{\mathrm{PB}}}{\sum 2p_j^{PB}(1-p_j^{PB})} & \dfrac{\mathbf{M}_{\mathrm{PB}}\mathbf{M}'_{\mathrm{CB}}}{\sqrt{\sum 2p_j^{PB}(1-p_j^{PB})}\sqrt{\sum 2p_j^{CB}(1-p_j^{CB})}} \\ \dfrac{\mathbf{M}_{\mathrm{CB}}\mathbf{M}'_{\mathrm{PB}}}{\sqrt{\sum 2p_j^{PB}(1-p_j^{PB})}\sqrt{\sum 2p_j^{CB}(1-p_j^{CB})}} & \dfrac{\mathbf{M}_{\mathrm{CB}}\mathbf{M}'_{\mathrm{CB}}}{\sum 2p_j^{CB}(1-p_j^{CB})} \end{bmatrix}, \qquad \textbf{6.2}$$

where $\mathbf{M_{CB}}$ and $\mathbf{M_{PB}}$ are a centred marker genotype matrix of CB animals and PB animals, respectively, by subtracting $\mathbf{2p_j^{CB}}$ (for $\mathbf{M_{CB}}$) or $\mathbf{2p_j^{PB}}$ (for $\mathbf{M_{PB}}$) from all genotypes of marker $\boldsymbol{j}$, where $\boldsymbol{p_j^{CB}}$ and $\boldsymbol{p_j^{PB}}$ are the allele frequency of marker $\boldsymbol{j}$ in the CB and PB animals, respectively. For scenarios PB-A and CB-I, either PB or CB animals were involved, so the $\mathbf{G}$ matrix in Equation 6.2 reduced to the genomic relationship matrix for a single breed: $\mathbf{G} = \frac{\mathbf{MM'}}{\sum 2p_j(1-p_j)}$ (VanRaden 2008).

When BOA was considered and validation was based on offspring averages (CB-A-BOA), the genomic relationships in $\mathbf{G}_{BOA}$ were constructed by using only the alleles that came from sire line A as:

$$\mathbf{G}_{BOA} = \begin{bmatrix} \mathbf{G}_{BOA,\mathrm{PB}} & \mathbf{G}_{BOA,\mathrm{PB-CB}} \\ \mathbf{G}_{BOA,\mathrm{PB-CB}} & \mathbf{G}_{BOA,\mathrm{CB}} \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{M}_{\mathrm{PB}}\mathbf{M}'_{\mathrm{PB}}}{\sum 2p_j(1-p_j)} & \frac{\mathbf{M}_{\mathrm{PB}}\mathbf{T}'_{\mathrm{CB}}}{\sum 2p_j(1-p_j)} \\ \frac{\mathbf{T}_{\mathrm{CB}}\mathbf{M}'_{\mathrm{PB}}}{\sum 2p_j(1-p_j)} & \frac{\mathbf{T}_{\mathrm{CB}}\mathbf{T}'_{\mathrm{CB}}}{\sum 2p_j(1-p_j)} \end{bmatrix}, \quad \textbf{6.3}$$

where $\mathbf{T}_{\mathrm{CB}}$ is the centred marker allele matrix for CB animals, with a value of $(1-p_j)$ if the counted allele was inherited from the PB sire line, and a value of $(0-p_j)$ if the other allele was inherited (Sevillano *et al.* 2017), where p_j denotes the frequency of the counted allele at marker j. The latter was calculated as the total number of counted alleles in the PB sires and in the CB offspring that were inherited from these sires, divided by the total number of PB alleles in these animals. Note that the $\mathbf{G}_{BOA}$ is similar to the marker-based partial relationship matrix from Christensen *et al.* (2014), with a scaling factor of $\sum 2p_j(1-p_j)$. As a result, the expected value of the diagonal elements for CB animals in $\mathbf{G}_{BOA}$ is 0.5. For scenario CB-I-BOA, only CB animals were involved, so $\mathbf{G}_{BOA}$ from Equation 6.3 reduced to a genomic relationship matrix between CB animals where only alleles from sire line A were considered, i.e. $\mathbf{G}_{BOA} = \frac{\mathbf{T_{CB}T'_{CB}}}{\sum 2p_j(1-p_j)}$.

6.2.6 Validation records, validation correlation and bias

Phenotypic records corrected for systematic environmental effects were used for validation and were obtained from the following model, separately for BW7 and BW35:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Lm} + \mathbf{Ts} + \mathbf{e} \quad \textbf{6.4}$$

where $\mathbf{y}$ is a vector of all available CB phenotypes, $\mathbf{s}$ is a vector of random sire effects with incidence matrix $\mathbf{T}$, and all other terms are the same as in Equation 6.1. The distribution of sire effects was assumed $\mathbf{s} \sim N(0, \mathbf{I}_s\sigma_s^2)$, where σ_s^2 is the sire

variance and $\mathbf{I}_s$ is an identity matrix. From the solutions of this model, corrected phenotypes were computed as $\mathbf{y}_c = \mathbf{T}\hat{\mathbf{s}} + \hat{\mathbf{e}}$.

The validation correlation and bias were evaluated for each replicate separately, using the GEBV and validation records of validation animals from all CV groups. For validation on individual CB records, the validation correlation was calculated as the correlation between GEBV and corrected individual CB records ($\mathbf{y}_c$) and the bias was calculated by regressing $\mathbf{y}_c$ on GEBV. For validation on CB offspring averages, the validation correlation was calculated as the weighted correlation between the sire GEBV and the average of corrected phenotypes of their CB offspring ($\bar{\mathbf{y}}_c$) and bias was calculated by weighted regression of $\bar{\mathbf{y}}_c$ on sire GEBV, with the weighted regression coefficient multiplied by two because the offspring average represents half the breeding value of the sire. The weights used in these analyses were the reliabilities of $\bar{\mathbf{y}}_c$, which were computed as $\frac{\frac{1}{4}nh^2_{CB}}{1+\frac{1}{4}(n-1)h^2_{CB}}$ (Cameron 1997), where h^2_{CB} is the estimated heritability of CB performance and n is the number of CB offspring. Note that the resulting validation correlations are not equal to but are proportional to the accuracies of the GEBV for a given validation population, defined as the correlation between GEBV and true breeding values in validation. The validation correlations therefore allow for a comparison between scenarios, which was the aim of our study.

6.3 Results

6.3.1 PB versus CB reference population

For BW7 and with validation on offspring averages, the PB and CB reference populations yielded a similar mean validation correlation (both equal to 0.16; Table 6.4). With validation on individual CB records, however, the CB reference population yielded a higher mean validation correlation than the PB reference population (0.13 vs. 0.05; Table 6.4). For BW35 and with validation on CB offspring averages, the CB reference population yielded a lower mean validation correlation than the PB reference population (0.26 vs. 0.36; Table 6.4). With validation on individual CB records, the CB reference population yielded a higher mean validation correlation than the PB reference population (0.16 vs. 0.13; Table 6.4).

Table 6.4 Mean validation correlations for BW7 and BW35

Scenario	Reference	Validation	BW7		BW35	
			Mean[a]	sd[b]	Mean[a]	sd[b]
PB-A	PB	Offspring averages	0.16	0.032	*0.36*	0.032
CB-A	CB	Offspring averages	0.16	0.058	0.26	0.060
CB-A-BOA	CB	Offspring averages	*0.20*	0.058	0.22	0.059
PB-I	PB	Individual records	0.05	0.014	0.13	0.014
CB-I	CB	Individual records	*0.13*	0.020	*0.16*	0.020
CB-I-BOA	CB	Individual records	0.08	0.025	0.09	0.025

[a]Reported values are means of 100 replicates. Highest mean validation correlations per validation record and per trait are in italics.
[b]Reported values are standard deviations of validation correlations of 100 replicates.

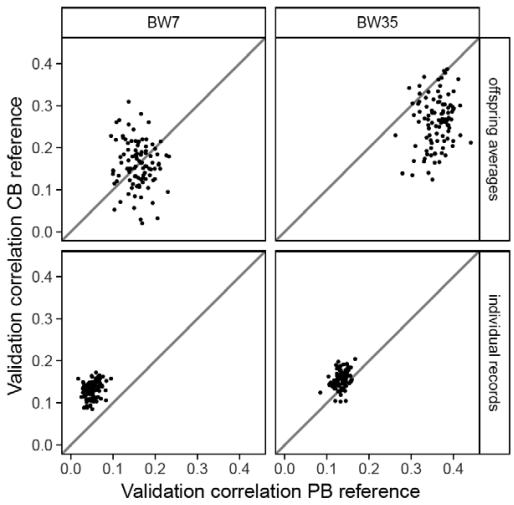

Figure 6.1 Validation correlations when validation was on CB offspring averages or individual CB records, using a PB or a CB reference population. The x-axis represents the validation correlation using a PB reference population and the y-axis represents the validation correlation using a CB reference population. Panels refer to validation on CB offspring averages or individual CB records across rows, and to body weight measured at around 7 (BW7) or 35 (BW35) days across columns. Dots represent individual validation correlations of 100 replicates and straight lines indicate x = y.

The differences between mean validation correlations were not always larger than their standard errors and, thus, we examined if these observed differences were consistent for individual validation correlations of replicates. For BW7 and with validation on CB offspring averages, there was no clear difference between a PB and a CB reference population (Figure 6.1, top-left); in 51% of the replicates, the validation correlation was higher for the CB reference population. However, with validation on individual CB records, the validation correlation was higher with a CB reference population for all replicates (Figure 6.1, bottom-left). For BW35 and with validation on CB offspring averages, the PB reference population yielded a higher validation correlation than a CB reference population for 93% of the replicates

(Figure 6.1, top-right). However, with validation on individual CB records the CB reference population mostly yielded a higher validation correlation (86% of the replicates; Figure 6.1, bottom-right).

Table 6.5 Mean regression coefficients of GEBV on validation records for BW7 and BW35

Scenario	Reference	Validation	BW7		BW35	
			Mean[a]	sd[b]	Mean[a]	sd[b]
PB-A	PB	Offspring averages[c]	0.51	0.105	*0.73*	0.069
CB-A	CB	Offspring averages[c]	0.36	0.133	0.64	0.158
CB-A-BOA	CB	Offspring averages[c]	*0.55*	0.171	0.59	0.167
PB-I	PB	Individual records	0.51	0.147	*0.77*	0.080
CB-I	CB	Individual records	0.55	0.070	0.67	0.073
CB-I-BOA	CB	Individual records	*0.67*	0.202	0.64	0.169

[a]Reported values are means of 100 replicates. Mean regression coefficients that are closest to 1 per validation record and per trait are in italics.
[b]Reported values are standard deviations of regression coefficients of 100 replicates.
[c]Reported regression coefficients were multiplied by 2 because offspring averages represent half the breeding value of sires.

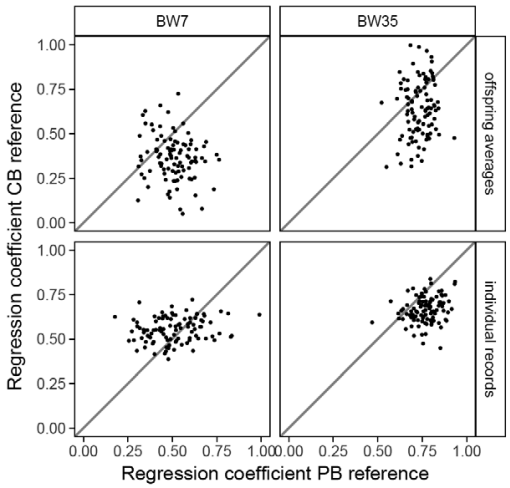

Figure 6.2 Regression coefficients of validation records on GEBV when validation was on CB offspring averages or individual CB records, using a PB or a CB reference population. The x-axis represents the regression coefficient using a PB reference population and the y-axis represents the regression coefficient using a CB reference population. Panels refer to validation on CB offspring averages or individual CB records across rows, and to body weight measured at around 7 (BW7) or 35 (BW35) days across columns. Dots represent individual regression coefficient of 100 replicates, and straight lines indicate x = y.

There were no clear differences in bias of GEBV between using a PB or CB reference population, except for BW35 and with validation on individual offspring records. For that scenario, GEBV from the PB reference population were less biased in 87% of the replicates (Figure 6.2, bottom-right), with a mean regression coefficient

of 0.77 for the PB reference population and 0.67 for the CB reference population (Table 6.5).

6.3.2 Ignoring versus considering BOA

With validation on offspring averages, considering BOA increased the mean validation correlation for BW7 (0.20 vs. 0.16; Table 6.4), but decreased the mean validation correlation for BW35 (0.22 vs. 0.26; Table 6.4). With validation on individual CB records, considering BOA decreased the mean validation correlation for both BW7 (0.08 vs. 0.13; Table 6.4) and BW35 (0.09 vs. 0.16; Table 6.4). Again, we examined whether the observed differences in mean validation correlations were consistent for individual replicates. With validation on CB offspring averages, taking BOA into account almost always increased the validation correlation for BW7 (93% of the replicates; Figure 6.3, top-left), whereas for BW35, it almost never increased it (3% of the replicates; Figure 6.3, top-right). With validation on individual CB records, taking BOA into account never increased the validation correlation for either BW7 or BW35 (Figure 6.3, bottom).

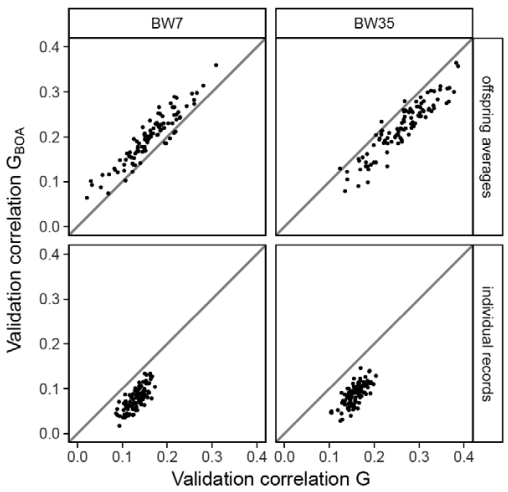

Figure 6.3 Validation correlations when validation was on CB offspring averages or individual CB records, the reference population consisted of CB animals, and BOA was ignored or considered. The x-axis represents the validation correlation when ignoring BOA and the y-axis represents the validation correlation when considering BOA. Panels refer to validation on CB offspring averages or individual CB records across rows, and to body weight measured at around 7 (BW7) or 35 (BW35) days across columns. Dots represent individual validation correlations of 100 replicates, and straight lines indicate x = y.

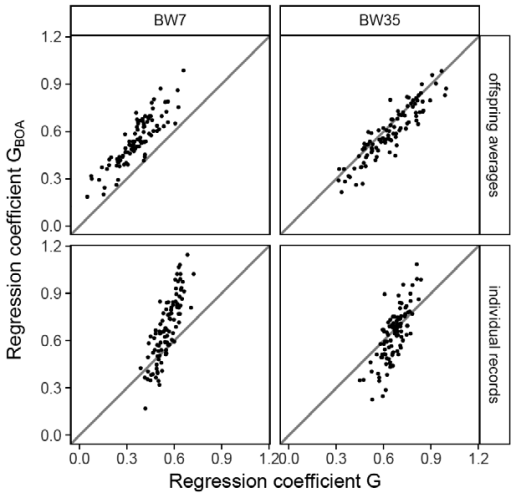

Figure 6.4 Regression coefficients of validation records on GEBV when validation was on CB offspring averages or individual CB records, the reference population consisted of CB animals, and BOA was ignored or considered. The x-axis represents the regression coefficient when ignoring BOA and the y-axis represents the regression coefficient when considering BOA. Panels refer to validation on CB offspring averages or individual CB records across rows, and to body weight measured at around 7 (BW7) or 35 (BW35) days across columns. Dots represent individual validation correlations of 100 replicates, and straight lines indicate x = y.

For BW35, there were no clear differences in bias of GEBV between models that considered or ignored BOA. For BW7 and with validation on offspring averages, GEBV from models that considered BOA were less biased (0.55; Table 6.5) than those from models that ignored BOA (0.36; Table 6.5) in 99% of the replicates (Figure 6.4, top-left). For BW7 and with validation on individual CB records, GEBV from models that considered BOA were less biased (0.67; Table 6.5) than those from models that ignored BOA (0.55; Table 6.5) in 77% of the replicates (Figure 6.4, bottom-left).

6.4 Discussion

We compared the validation correlation and bias of GEBV of PB animals for CB performance using either CB offspring averages or individual CB records as validation records. Our aim was to investigate the effect of using either a PB or CB reference population, and the effect of either ignoring or considering BOA.

It should be noted that the PB and CB animals in this study were housed in the same environment, whereas in practice, PB animals are housed in a nucleus facility and CB animals are housed in a commercial environment. As such, the estimates of r_{pc} obtained here provide an upper bound for r_{pc} in practical situations, where genotype-by-environment interactions may be present (Duenk *et al.* 2019b). Consequently, the benefit of using CB information may be larger in practical situations than found here. Thus, our results on differences in validation correlations

between scenarios should not be associated with the body weight traits per se, but with the value of the r_{pc}.

We investigated bias of GEBV by computing weighted regression coefficients of validation records on GEBV. The average coefficients across replicates were substantially lower than 1 for all scenarios, which indicates a strong bias (over-dispersion of GEBV). This bias may be due to family structure in the data and imprecision of GEBV, which may lead to a theoretical expectation of the true regression coefficient being smaller than 1 (Legarra and Reverter 2018). Regardless, our results show that for BW7, taking BOA into account reduced bias in almost all the replicates. For all other comparisons, differences in regression coefficients were not statistically significant because of large standard deviations of estimates across replicates To date, no other studies have evaluated the impact of considering BOA on the bias of GEBV, and therefore, it remains unclear whether considering BOA generally reduces bias or not.

6.4.1 Purebred versus crossbred reference populations

As expected, our results suggest that with validation on CB offspring averages, the difference in validation correlation between using a PB and a CB reference population partly depends on the r_{pc}. With an r_{pc} of 0.96 (BW35), the validation correlation was lower with a CB reference population than with a PB reference population, while validation correlations were similar for the CB and PB reference populations with an r_{pc} of 0.80 (BW7). These results are in line with studies based on simulated (Esfandyari *et al.* 2015a; Van Grevenhof and Van Der Werf 2015) and real data (Hidalgo *et al.* 2016), thus confirming that the benefit of a CB reference population is larger for smaller values of r_{pc}. However, with validation on individual CB records, the validation correlation was higher with a CB reference population, regardless of the r_{pc} (i.e., for both traits), which agrees with Lopes *et al.* (2017), who analysed traits with an r_{pc} of about 0.9 and also validated on CB offspring records. In addition, two other studies have shown that genotyping CB animals improves the accuracy of CB offspring GEBV using single-step genomic best linear unbiased prediction (GBLUP) (Lourenco *et al.* 2016; Xiang *et al.* 2016). In the following sections, we will discuss the two validation strategies and give reasons that explain why they can result in different conclusions about the benefit of using CB information for genomic prediction.

6.4.2 Validation on offspring averages

With validation on CB offspring averages, differences in validation correlations between using a PB (PB-A) vs. a CB (CB-A) reference population can result from two mechanisms: (i) with an r_{pc} less than 1, a CB reference population has an advantage over a PB reference population; (ii) in the CB reference population, only half of the alleles originate from the sire line (Moghaddar *et al.* 2014; Van Grevenhof and Van Der Werf 2015), whereas all alleles originate from the sire line in the PB reference population. When the sire and dam lines are unrelated, the maternal alleles in the CB reference population introduce noise in estimation of the sire-line genetic component because the sire-line alleles in the CB reference population explain only half of the genetic variance, whereas sire-line alleles in the PB reference population explain the full genetic variance. This results in a disadvantage for a CB reference population compared to a PB reference population. However, when the sire and dam lines are somewhat related, the dam-line allelic effects in the CB reference population may have some predictive value for the sire-line allelic effects. This would increase the accuracy of sire-line GEBV, and thus reduce the disadvantage for the CB reference population compared to the PB reference population when using a model that ignores BOA.

Observed differences in validation correlations between PB-A and CB-A depend on the balance between the aforementioned two mechanisms. To quantify the predictive value of dam-line allelic effects for sire offspring averages, we estimated sire GEBV by using only the alleles in the CB reference population that originated from the dam line. For BW7, the mean validation correlation from this model was equal to 0.03, with a standard deviation of 0.07 across replicates, whereas for BW35, the mean validation correlation was equal to 0.14 with a standard deviation of 0.07. These results indicate that the dam alleles in the CB animals may have some predictive value for sire offspring averages, which is supported by the observation that considering BOA (i.e. removing the dam alleles) decreased the validation correlation for BW35 (as discussed in later sections). For BW7, the effects of the two mechanisms resulted in similar validation correlations for PB-A and CB-A. For BW35, for which r_{pc} was closer to 1, the effects of the two mechanisms resulted in a lower validation correlation with CB-A than with PB-A.

6.4.3 Validation on individual offspring records

With validation on individual offspring records, differences in validation correlations between a PB reference population (PB-I) and a CB reference population (CB-I) observed in this study may be due to the same two mechanisms described

above. However, the predictive value of the dam alleles is higher with validation on individual crossbred records than with validation on crossbred offspring averages of sires, for two reasons: (i) the prediction of individual CB records is partly (i.e., half) based on the dam-line alleles of those CB individuals, and (ii) an individual record may have a residual genetic dam component. Thus, the CB-I validation correlations are prone to overestimate GEBV accuracies due to the contribution of dam alleles to the prediction of individual records, which contain a residual genetic dam component. For both traits (BW7 and BW35), the effects of these two mechanisms resulted in higher validation correlations with CB-I than with PB-I, but this difference was smaller for BW35 than for BW7, which was probably due to the higher r_{pc} of BW35.

6.4.4 Choice of validation records

As discussed in the previous sections, the difference in validation correlations of genomic predictions between using a CB and a PB reference population depend not only on the value of r_{pc} but also on the choice of validation records (CB offspring averages or individual CB records). We even observed that the ranking of validation correlations with a PB versus a CB reference population changed when a different validation record was used, which raises the question which validation record is most relevant. In practice, breeders usually aim to identify PB selection candidates that, on average, produce the best CB offspring. Thus, the relevant validation correlation is the correlation of the GEBV of sires and their CB offspring averages. Validation on offspring averages may not be possible when the number of genotyped PB sires with phenotyped CB offspring is too small. In those cases, validation of GEBV from CB animals on their individual records may provide an alternative. However, with validation on individual records, the apparent superiority of a CB over a PB reference population will likely be inflated because, as discussed above, validation correlations from models that use a CB reference population and ignore BOA are contaminated with the predictive value of dam alleles for the residual genetic dam component in the validation records. Indeed, this inflation was reflected in a higher validation correlation with validation on individual records (0.29 for BW7 and 0.33 for BW35, Table A 6.1) instead of on offspring averages (0.18 for BW7 and 0.30 for BW35, Table A 6.1), when the validation correlations were compared on the same scale (i.e. scaled by the square root of the heritability and of the mean reliability, respectively). This mechanism may explain why, for traits with similar r_{pc}, Lopes et al. (2017) found that the validation correlation was higher with a CB reference population (with validation on individual CB records), but Hidalgo et al. (2016) found that the validation correlation was higher with a PB reference population (with validation on CB

offspring averages). Thus, when genomic predictions using a PB versus a CB reference population are compared, validation of sire GEBV on CB offspring averages are preferred.

In the previous paragraph, we argued that, with validation on individual offspring records and when BOA is ignored, validation correlations may be inflated due to the predictive value of dam alleles. However, when BOA is considered, the dam alleles of CB animals are removed from the explanatory variables of the model and the validation correlation is not expected to be inflated. So, when validating on individual records, the benefit of using a CB reference population is better evaluated by comparing a model that uses PB information with a model that uses CB information while considering BOA. This comparison for our data showed that that the CB reference population yielded a higher validation correlation than the PB reference population for BW7 (0.08 vs. 0.05) but not for BW35 (0.09 vs. 0.13). Furthermore, for this comparison, GEBV were less biased with a CB reference population than with a PB reference population for BW7 but not for BW35, although differences in regression coefficients were not statistically significant.

6.4.5 Considering versus ignoring BOA

We compared the validation correlation of models that ignored (CB-A and CB-I) or considered BOA (CB-A-BOA and CB-I-BOA). With validation on offspring averages, the difference in validation correlations between considering and ignoring BOA depended on the predictive value of dam alleles in the CB animals for sire offspring averages. As shown before, this predictive value was close to zero for BW7 but larger than zero for BW35. In other words, the dam alleles introduced noise in the estimation of the genetic sire component for BW7 but this noise was less for BW35, resulting in a higher validation correlation when BOA was considered for BW7 but lower for BW35. These results suggest that taking BOA into account was beneficial for a trait with an r_{pc} of 0.8 but not for a trait with an r_{pc} of 0.96, which agrees with results of Sevillano et al. (2017) and Lopes et al. (2017), who also found that the benefit of considering BOA decreased with increasing r_{pc} and heritability. It has been argued that considering BOA may improve the validation correlation when the estimated r_{pc} from a model that takes BOA into account is different from a model that ignores it (Sevillano 2018). Our study neither confirmed nor contradicted this hypothesis because, although we observed a benefit of considering BOA for BW7, the estimate r_{pc} from models that ignored or considered BOA were the same in this dataset (Duenk *et al.* 2019b).

6.4.6 Implementation of BOA in practice

To our knowledge, information on BOA is currently not used in commercial crossbred evaluations. One reason may be that the algorithm to derive BOA is computationally demanding for large datasets. However, phasing algorithms are continuously being improved in terms of computational requirements (Loh *et al.* 2016) and computation power keeps increasing (Denning and Lewis 2016). In the long term, we expect that implementation of BOA-models will depend mainly on their benefit for genomic prediction, because computing costs will be relatively small compared to other costs of a breeding program. The results of this study and those of others (Lopes *et al.* 2017; Sevillano *et al.* 2017) suggest that considering BOA can improve the accuracy of genomic predictions for traits with a low r_{pc} and low heritability. Furthermore, as discussed above, the value of CB information for genomic prediction accuracy may be over-predicted when validation is on individual offspring records and BOA is ignored.

6.4.7 Practical relevance

In this study, we investigated whether GEBV of PB animals for CB performance should be computed based on PB or CB performance measured on animals that have comparable relationships with the selection candidates. Thus, own performance records of selection candidates were ignored. In practice, however, selection candidates may have an own performance record for PB performance. For those cases, it may be more useful to compare scenarios that use only PB records with those that combine PB and CB records in a single reference population. However, some traits cannot be measured on selection candidates (e.g. carcass traits) and, as a result, GEBV can only be computed based on information from relatives. For those cases, our results provide valuable insight into the benefit of CB over PB information.

6.5 Conclusions

Our findings show that the difference in validation correlations between using a PB or CB reference population not only depends on the r_{pc} of the trait evaluated but also on the choice of the validation record. With a CB reference population, the validation correlation from validation on individual CB records can be inflated because CB offspring records contain a substantial residual genetic dam component that can be predicted by the dam alleles of CB animals. Thus, we argue that, whenever possible, validation correlations for GEBV of PB animals for CB performance should be obtained from validation on CB offspring averages, because the interest usually lies in the identification of PB animals that, on average, produce

the best CB offspring. When validation on offspring averages is not possible and validation is on individual CB records, the actual benefit of using a CB reference population should be assessed by comparing the use of a PB reference population with the use of a CB reference population with BOA considered. For this comparison, our results show that the validation correlation with a CB reference population was equal to or higher than with a PB reference population for a trait with an r_{pc} of 0.8 but lower for a trait with an r_{pc} of 0.96. In addition, in our population, taking BOA into account was beneficial for a trait with an r_{pc} of 0.8 but not for a trait with an r_{pc} of 0.96.

6.6 Acknowledgements

The authors thank the Netherlands Organisation of Scientific Research (NWO) and the Breed4Food consortium partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin for their financial support. Special thanks go to Cobb Europe for providing the data. The use of the HPC cluster was made possible by CAT-AgroFood (Shared Research Facilities Wageningen UR).

6.8 Appendix

Table A 6.1 Scaled[a] mean validation correlations for BW7 and BW35

Scenario	Reference	Validation	BW7		BW35	
			Mean[b]	sd[c]	Mean[b]	sd[c]
PB-A	PB	Offspring averages	0.18	0.037	*0.41*	0.036
CB-A	CB	Offspring averages	0.18	0.067	0.30	0.069
CB-A-BOA	CB	Offspring averages	*0.23*	0.066	0.25	0.067
PB-I	PB	Individual records	0.11	0.032	0.28	0.028
CB-I	CB	Individual records	*0.29*	0.044	*0.33*	0.042
CB-I-BOA	CB	Individual records	0.17	0.056	0.18	0.051

[a]Scaled correlations were computed as the unscaled validation correlation divided by the square-root of the heritability (for validation on individual records), or divided by the square-root of the weighted mean reliability (for validation on offspring averages). Across scenarios, the average weighted mean reliability was 0.77 for both BW7 and BW35.
[b]Reported values are means of 100 replicates. Highest mean validation correlations per validation record and per trait are in italics.
[c]Reported values are standard deviations of validation correlations of 100 replicates.

7

General discussion

7.1 Introduction

Pig and poultry breeding programs usually mate different purebred (PB) lines to produce crossbred (CB) animals, so that farmers benefit from breed complementarity and heterosis. The ultimate aim of such programs is to improve the performance of CB animals, while selection usually takes place in the PB lines based on PB performance. Selection on PB performance enables the improvement of CB performance through a correlated response. This strategy may, however, be suboptimal because the performance of CB animals can be genetically different from the performance of PB animals. This difference is quantified by the genetic correlation between PB and CB performance (r_{pc}), which can be lower than one due to genotype by environment interactions (GxE), and due to non-additive genetic effects in combination with differences in allele frequency between the parental lines (GxG). The r_{pc} is an important parameter in crossbred breeding programs, because the value of r_{pc} partly determines whether it is useful to collect data on CB animals instead of PB animals for the estimation of breeding values. The genetic mechanisms that cause r_{pc} to be lower than one are, however, not fully understood. Moreover, it is not yet clear whether estimates and standard errors of r_{pc} differ between estimation models.

To improve the response to selection in CB performance, breeders may choose to select PB animals based on breeding values for CB performance instead of PB performance. With genomic prediction, it is possible to obtain genomic estimated breeding values (GEBV) for CB performance of PB selection candidates. This approach, however, requires the collection of information (phenotypes and genotypes) on CB animals, which may be costly and difficult in practice. In addition, the benefit of using CB over PB information depends on r_{pc}, which may differ across species, breeds, and traits. Finally, it is not yet clear how breeding values for CB performance should be validated properly. These and other questions were addressed in this thesis.

The overall objective of this thesis was to study the genetics of crossbreeding with a focus on the role of non-additive effects, and on genomic prediction for crossbred performance. Chapter 2 focused on the estimation of average effects at QTL, and I showed that a model that explicitly models dominance yielded more accurate estimates of average effects than an additive model. The dominance model had a benefit over the additive model, because the dominance model was more robust against sampling deviations from Hardy-Weinberg equilibrium (HWE). In chapter 3, I showed that the genetic correlation between populations (r_g) decreases

with an increasing size of non-additive effects (i.e. both dominance and epistasis), and with increasing differences in allele frequencies between populations. In chapter 4, I investigated whether the true value of r_{pc} can be predicted based only on information from the two parental lines without using information on crossbreds. I showed that with only dominance, r_{pc} in line 1 is equal to the correlation between additive genetic values of individuals in line 1, for the trait expressed in parental lines 1 and 2, which is similar to the genetic correlation between parental lines (r_g). With only additive by additive (AxA) epistasis, r_{pc} can be predicted exactly from the genetic covariance between, and genetic variance within parental lines. In chapter 5, I showed that estimates of r_{pc} from genotype-based models were somewhat larger and had smaller standard errors than those from pedigree-based models. Estimates of r_{pc} were similar between models that either ignored or considered the breed-of-origin of alleles (BOA). Finally, in chapter 6, I showed that GEBV for CB performance should be validated with CB offspring averages rather than individual CB records, because validation on individual CB records led to inflated accuracies. Furthermore, the accuracy of GEBV for CB performance was equal to or higher with a CB reference population than with a PB reference population for a trait with an r_{pc} of 0.8, but lower for a trait with an r_{pc} of 0.96. In addition, taking BOA into account resulted in higher accuracy for a trait with an r_{pc} of 0.8 but not for a trait with an r_{pc} of 0.96.

In this final chapter, I will discuss two topics. First, I will show that r_{pc} and heterosis are closely related through the existence of non-additive genetic effects. I show that heterosis can occur because the response in CB performance can be larger than the response in PB performance, even when selection is on PB performance and r_{pc} is lower than one. Second, I will discuss different strategies to estimate breeding values for CB performance. For each strategy, I will discuss their strengths and weaknesses, and I hypothesize that using a CB reference population and considering the BOA will optimize the accuracy of genomic prediction models.

7.2 Heterosis

One of the benefits of crossbreeding is heterosis. Heterosis is the phenomenon where the average performance of a cross exceeds the average performance of its PB parental lines (Shull 1952; Dickerson 1973). For a two-way CB population (i.e. an F1 cross), heterosis can be expressed as $H = CP - BA$, where CP is the average performance of the CB, and BA is the average performance of the two parental lines (i.e. breed average) (Falconer and Mackay 1996). Breeders may benefit from mating lines that result in large values of H. At the same time, breeders may benefit from

avoiding crosses between lines that result in a low r_{pc}, because the value of data on PB performance for the response to selection in CB performance is higher when r_{pc} is high. Together, strong heterosis and high r_{pc} will result in improved CB performance when selection is on PB performance. From this point of view, it would be interesting to determine whether traits exist that show both strong heterosis and high r_{pc}.

In this section, I investigate the relationship between r_{pc} and heterosis, and discuss the relevance of this relationship for crossbred animal breeding. First, I will review the genetic basis of heterosis in the context of simple single- or two-locus models. Second, I will investigate heterosis under different genetic models (i.e. different types of non-additive effects being present), using the simulated data from chapter 4. Then, I will use these simulations to investigate the relationship between r_{pc} and heterosis. Finally, I will discuss how selection on PB performance affects r_{pc}, the response in CB performance, and heterosis.

7.2.1 Genetic basis of heterosis

There are three hypotheses of how heterosis occurs. Two of these hypotheses are based on the existence of dominance effects and the increased frequency of heterozygotes in CB compared to the parental lines. First, the *overdominance hypothesis* states that CB performance exceeds that of the breed average because heterozygote genotypes at QTL outperform both of the two homozygous genotypes (East 1908; Shull 1908). Second, the *dominance hypothesis* states that the effects of deleterious alleles in CB that are inherited from one of the parents, are masked by (partially) dominant alleles inherited from the other parent (Davenport 1908). From empirical and theoretical studies on heterosis, it seems likely that heterosis is mainly due to the masking of deleterious alleles, because there is only little evidence for many overdominant genes, and the existence of overdominance is not required to explain continued heterosis observed in empirical studies (Crow 1999; Charlesworth and Willis 2009).

The third hypothesis for why heterosis occurs relies on the existence of positive and negative epistatic effects (as explained in the next paragraph). While the importance of dominance for heterosis is widely recognized, the role of epistasis is often overlooked, probably because its effect on heterosis is expected to be smaller than that of dominance (Schnell and Cockerham 1992; Goodnight 1999). Empirical research has suggested, however, that epistasis may play a significant role for heterosis in *Arabidopsis* (Melchinger *et al.* 2007a), rice (Shen *et al.* 2014), and pigs (Bidanel 1993). Of course, the three hypotheses of heterosis are not mutually

exclusive, and I expect that in reality, dominant, overdominant, and epistatic effects contribute to heterosis.

For a single locus model with directional dominance, heterosis is proportional to the size of the dominance effect, and to the squared difference in allele frequency between parental lines (Falconer and Mackay 1996). As a result, with dominance, an increase in allele frequency differences between parental lines always leads to increased (and positive) heterosis in their cross. For a two-locus model with only an AxA epistatic interaction, heterosis is proportional to the size of the epistatic effect, and to the *product of absolute differences in allele frequency* between lines at the interacting loci (Hill 1982; Willham and Pollak 1985; Melchinger *et al.* 2007b). The direction of heterosis (i.e. positive or negative) due to AxA epistasis depends on the sign of differences in allele frequencies between parental lines at interacting loci: epistasis can lead to positive heterosis only when the differences in allele frequency between lines for the two interacting loci are of opposite sign, and the epistatic effect is positive (Minvielle 1987; Schnell and Cockerham 1992). The idea is that favourable allele combinations across loci are not present in either of the PB lines, and that these favourable combinations increase in frequency in the cross between those lines. Such a situation may occur with composite traits that result from taking the product of two component traits (resulting in multiplicative epistasis in the component trait), and the parental lines are selected for different components (Box 1). In all other cases, epistasis leads to negative heterosis, because favourable allele combinations in either of the PB lines are lost with crossbreeding (Kinghorn 1982; Minvielle 1987; Melchinger *et al.* 2007b). An example of negative heterosis due to epistasis is when the trait of interest is under direct positive selection in at least one of the parental lines, as we will see in the following.

7.2.2 Heterosis under different genetic models

The manifestation and amount of heterosis depends on the divergence between parental lines in terms of allele frequency differences, and on the type of non-additive effects that affect the trait (i.e. the genetic model). To study the effect of line divergence and genetic model on heterosis, I used the simulated data from chapter 4, consisting of six PB parental lines that had been separated from the randomly selected line (R) for 10, 25, or 50 generations under either positive (P) or negative selection (N) on PB performance. Here, I only consider crosses between line R and one of the six positively or negatively selected lines. First, I computed the expected genotype frequencies in each cross, based on the allele frequencies in their parental lines. Then, I used these genotype frequencies and the simulated genetic

effects for each genetic model to compute the expected average performance of each cross. Finally, I computed expected heterosis as the difference between the expected average performance of the cross, and the average performance of its parental lines, for each cross and each genetic model.

Box 1 Positive heterosis due to epistasis in composite traits

Positive heterosis due to epistasis can arise when genes act in a multiplicative manner (i.e. there is multiplicative epistasis), for example with composite traits that result from taking the product of two component traits (Schnell and Cockerham 1992; Charlesworth and Willis 2009). An example of such a trait is total seed number (NS) in plants, which is the product of number of flowers (NF) and number of seeds per flower (S/F). Suppose NF is controlled additively by locus A (with a positive effect of allele A_1), whereas S/F is controlled additively by locus B (with a positive effect of allele B_1). As a result, the trait NS results from an multiplicative epistatic interaction between locus A and B, in the following manner (blue cells indicate marginal additive effects of loci A and B).

			Locus *NF*		
			A_1A_1	A_1A_2	A_2A_2
			2	*1*	*0*
Locus *S/F*	B_1B_1	*2*	4	2	0
	B_1B_2	*1*	2	1	0
	B_2B_2	*0*	0	0	0

Now assume that line 1 has been selected for NF, and line 2 has been selected for S/F. Looking at the composite trait NS, the average performance of a cross between line 1 and 2 will outperform the average performance of the two parental lines, because the cross will have higher frequencies of favorable genotype combinations (e.g. $A_1A_1B_1B_1$) than the parental lines on average. For example, two fully inbred lines with genotypes $A_1A_1B_2B_2$ and $A_2A_2B_1B_1$ both have a genotypic value of 0, whereas the cross between these lines ($A_1A_2B_1B_2$) has a genotypic value of 1, resulting in heterosis. Note that this specific example can be considered a case of breed complementarity that leads to positive heterosis. In general, for any type of epistasis that introduces additive by additive interaction, there will be positive heterosis when the first parental line has a higher frequency of the positive allele at locus A (allele A_1) than the second line, and the second line has a higher frequency of the positive allele at locus B (allele B_1) than the first.

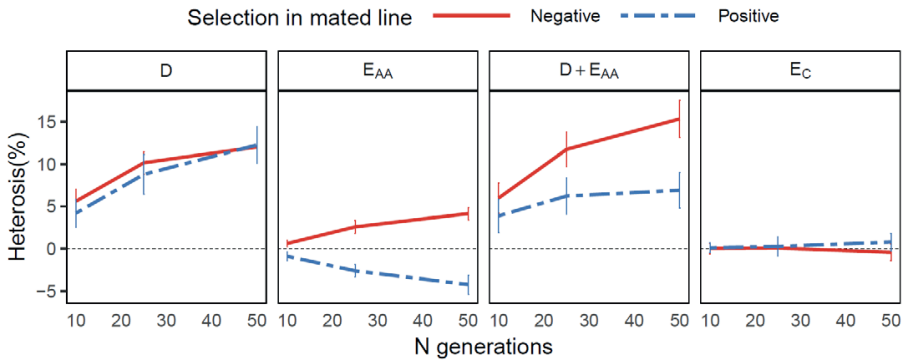

Figure 7.1 Expected heterosis (expressed in percentage points relative to parent average, y-axis) for crosses between a randomly selected line (R) and one of six mated lines. The six mated lines were either positively (blue lines) or negatively selected (red lines) for N generations, where N was equal to 10, 25 or 50 (x-axis). Column facets indicate the genetic model, where D refers to the presence of dominance effects, E_{AA} to additive by additive epistatic effects, and E_C refers to complementary epistatic effects. Lines represent means across 20 replicates, and error bars represent standard errors of those means.

The results showed that, with only (directional) dominance (D), heterosis increases with the number of generations of divergence, regardless of the direction of selection[3] (Figure 7.1). This result was expected, because with only dominance, heterosis is proportional to the square of allele frequency differences between parental lines (Falconer and Mackay 1996). In contrast, with only AxA epistasis (E_{AA}), heterosis increases with negative selection and decreases with positive selection in the mated line[3]. This result was also expected, because favourable (unfavourable) allele combinations in a positively (negatively) selected PB line are lost in the CB genotypes. With both dominance and AxA epistasis ($D + E_{AA}$), the expected heterosis is the sum of the heterosis observed with D and E_{AA}. This resulted in increasing heterosis across generations with negative selection, and stable heterosis after 25 generations with positive selection. Heterosis stabilized with positive selection in this scenario because the increase of heterosis due to dominance was counteracted by the decrease in heterosis due to AxA epistasis. Finally, with complementary epistasis (E_C), heterosis was very small. These results are in line with results of Melchinger *et al.* (2007b), who showed that heterosis increases with increasing line divergence and directional dominance, and that heterosis can either increase or decrease with AxA interactions, through the breakdown of unfavourable or favourable allele combinations in the crossbreds. In addition, these results agree with results from

[3] Note that, with a trait that is negatively selected, positive heterosis is unfavourable when the aim is to decrease the trait for CB performance.

other simulation studies that considered dominance effects, where heterosis was still observed after 5 to 40 generations of selection on PB performance (Esfandyari *et al.* 2015b; Esfandyari *et al.* 2018).

7.2.3 The relationship between heterosis and r_{pc}

Heterosis and r_{pc} both depend on the size of non-additive effects and on the difference in allele frequencies between parental lines (chapter 4; Falconer and Mackay 1996), and therefore heterosis and r_{pc} are expected to be closely related. To my knowledge, the relationship between heterosis and r_{pc} has never been described, and is therefore the focus of this section.

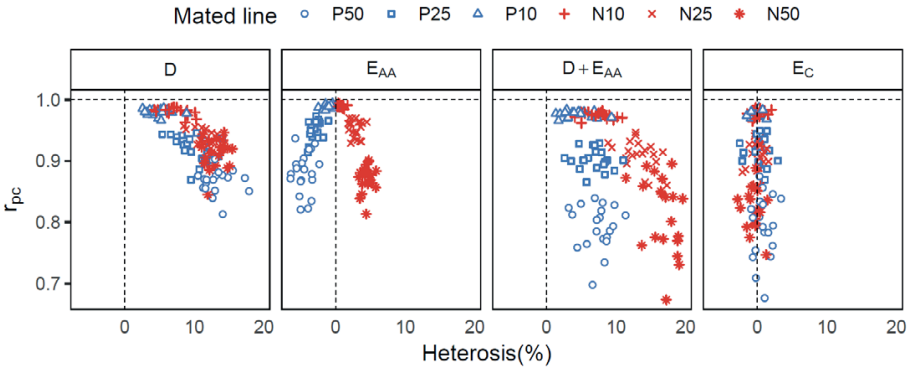

Figure 7.2 The purebred-crossbred genetic correlation (r_{pc}, y-axis) plotted against expected heterosis (x-axis), for crosses between a randomly selected line (R) and one of six mated lines. The six mated lines were either positively (P) or negatively (N) selected for 10, 25, or 50 generations (indicated with shapes and colours). Column facets indicate the genetic model, where D refers to dominance effects, E_{AA} to additive by additive epistatic effects, and E_C refers to complementary epistatic effects.

I studied the relationship between heterosis and r_{pc} under different genetic models, using the same simulations as before (section 7.2.2). Again, I considered only crosses between the randomly selected line (R) and one of the six positively (P) or negatively (N) selected lines. The realized r_{pc} values were computed as explained in chapter 4, using additive genetic values of individuals in the mated (i.e. selected) line[4]. Results showed that with only dominance effects (D), heterosis and r_{pc} were negatively correlated, so that greater heterosis goes together with lower r_{pc} (Figure 7.2). With only AxA effects (E_{AA}), heterosis and r_{pc} were also negatively correlated

[4] Note that for this discussion, I simulated stronger epistatic effects than in chapter 4, so that the relationship between heterosis and r_{pc} with epistasis in Figure 7.2 is clearly visible. This led to smaller values of r_{pc} with epistasis than in chapter 4.

when heterosis was positive (i.e. when the mated line was negatively selected), and they were positively correlated when heterosis was negative (i.e. when the mated line was positively selected). With both dominance and AxA epistatic interactions (D + E_{AA}), heterosis and r_{pc} were always negatively correlated, regardless of the direction of selection in the mated line. The slope of the regression line between heterosis and r_{pc} was, however, steeper with positive selection than with negative selection. This difference in slope is caused by the AxA epistatic interactions that are added to the dominance effects, which reduce heterosis with positive selection, and increase heterosis with negative selection. With only complementary epistasis (E_C), there was no clear relationship between r_{pc} and heterosis, because complementary interactions contribute only little to heterosis.

From the above it can be seen that, for the genetic models considered, there exists a positive correlation between heterosis and r_{pc} only when heterosis itself is negative. Negative heterosis in F1 crosses between lines of the same species is, however, rarely observed and therefore unlikely (Minvielle 1987; Lynch 1991). I therefore expect that, in practice, heterosis and r_{pc} are usually negatively correlated. This result makes sense, because when the genetic distance between parental lines increases, heterosis is expected to increase, while r_{pc} is expected to decrease. In conclusion, breeders are probably unable to benefit from both strong positive heterosis and high r_{pc}, because an increase in allele frequency differences between parental lines has opposite effects on r_{pc} and heterosis.

7.2.4 Selection, heterosis and r_{pc}

When selection in the parental lines is on PB performance, the improvement in CB performance occurs through a correlated response. When r_{pc} is lower than one, it is expected that this response in CB performance is lower than the response in PB performance. Hence, as selection proceeds for multiple generations, the cumulative response in CB performance is expected to be smaller than in PB performance. It can therefore be expected that heterosis reduces from one generation to the next, which may eventually even lead to no or negative heterosis (Brian Kinghorn, pers. comm.). In contrast to these expectations, selection experiments have shown that positive heterosis can persist for at least 5 to 40 generations with selection on PB performance (Sheridan and Randall 1977; Ayyagari *et al.* 1982; Minvielle *et al.* 1999; Yang *et al.* 1999). Furthermore, my results in the previous section also showed that after 50 generations of PB selection, positive heterosis was still observed (Figure 7.1), and that heterosis in fact increases (instead of decreases) when r_{pc} decreases (Figure 7.2). This observation implies that there is

a discrepancy between how we expect that heterosis changes with selection on PB performance, and what is observed in real data and simulations. In this subsection, I will discuss how heterosis may arise and persist across generations with selection for PB performance.

With selection on PB performance, the difference in response to selection between PB and CB performance not only depends on r_{pc}, but also on the ratio between additive genetic standard deviations for CB and PB performance,

$$R_{CB} = r_{pc}\frac{\sigma_{CB}}{\sigma_{PB}}R_{PB}, \qquad \textbf{7.1}$$

where R_{CB} (R_{PB}) is the response in CB (PB) performance, and parameters σ_{CB} and σ_{PB} are the additive genetic standard deviations for PB and CB performance *in the PB line under selection* (Wientjes and Calus 2017). It is important to note that the differences between σ_{CB} and σ_{PB} are only due to differences in the standard deviations of average effects for PB ($\sigma_{\alpha,PB}$) and those for CB performance ($\sigma_{\alpha,CB}$), because the allele frequencies used in the expressions for σ_{PB} and σ_{CB} are the same. In the following paragraph, I will show that the ratio σ_{CB}/σ_{PB} can increase due to selection when there is directional dominance, which can lead to heterosis.

For illustration purposes, consider a cross between a randomly selected dam line (R), and a sire line that has been positively selected for a number of generations (P). Selection in line P causes the frequency of favourable alleles in line P (p) to increase. With directional dominance (i.e. the dominance effects d are predominantly positive), an increase in p beyond 0.5 causes average effects for PB performance (α_{PB}) to become smaller, because the dominance term ($(1-2p)d$) decreases when p increases. This change in α_{PB} across generations leads to a reduction in $\sigma_{\alpha,PB}$. In contrast, $\sigma_{\alpha,CB}$ remains largely unchanged across generations, because with dominance, the average effects for CB performance (α_{CB}) depend heavily on the allele frequencies in the unselected line (R) (e.g. chapter 4). As a result, the ratio σ_{CB}/σ_{PB} becomes increasingly larger than 1 with each generation of selection. At the same time, r_{pc} decreases with each generation of selection, because the differences in allele frequencies between lines P and R increase. When σ_{CB}/σ_{PB} is equal to $1/r_{pc}$, the response to selection in CB performance is equal to the response in PB performance (Equation 7.1). In those cases, there is no expected change in absolute heterosis from the current generation to the next. However, heterosis is expected to increase when σ_{CB}/σ_{PB} is larger than $1/r_{pc}$, and heterosis is expected to decrease when σ_{CB}/σ_{PB} is smaller than $1/r_{pc}$. These expectations were

supported by my simulations, because the observed change in heterosis in Figure 7.1 could be fairly well predicted by Equation 7.1 (results not shown).

A change in heterosis was also observed with additive by additive epistatic effects (model $E_{AA,}$ Figure 7.1). Epistatic effects were assumed to have no direction, so that there were as many interactions with a positive effect ($\epsilon > 0$) as with a negative effect ($\epsilon < 0$). Individual α_{PB} at loci can either become smaller or larger with any change in allele frequencies at the interacting locus (p_i), depending on the sign of the epistatic effect. For example, with positive selection and $p_i > 0.5$, the epistatic term in α_{PB} ($-(1-2p_i)\epsilon$) increases when $\epsilon > 0$, and decreases when $\epsilon < 0$. With negative selection and $p_i < 0.5$, the epistatic term in α_{PB} decreases when $\epsilon > 0$ and increases when $\epsilon < 0$. As a result, both selection strategies result in an increase in $\sigma_{\alpha,PB}$, because half of the α_{PB} are expected to increase, and the other half of the α_{CB} are expected to decrease. At the same time, $\sigma_{\alpha,CB}$ also increases, but at a slower rate. Hence, the ratio σ_{CB}/σ_{PB} becomes smaller than one with both positive and negative selection. When the direction of selection is positive, this reduced response leads to negative heterosis, whereas when the direction of selection is negative, it leads to positive heterosis (Figure 7.1).

In my simulations, the increase of σ_{CB}/σ_{PB} was caused by selection on PB performance in one of the lines. In reality, such an increase may also arise through other mechanisms. For example, genes for disease resistance may be important for CB performance, but not for PB performance, because the PB lines are kept in pathogen-free environments. As a result, the genes that are important for PB performance ('PB genes') are different from those important for CB performance ('CB genes'). If selection is on PB performance, there is little selection pressure on CB genes, so that the allele frequencies of those genes remain relatively unchanged. However, the allele frequencies of PB genes change due to selection, resulting in a decrease in σ_{PB} (with directional dominance and positive selection, see above). Hence, this mechanism also results in an increase of σ_{CB}/σ_{PB} with selection on PB performance, which may lead to heterosis.

In summary, with selection on PB performance, the reduction of response in CB performance due to $r_{pc} < 1$ may be overcome by the larger additive genetic standard deviation for CB performance compared to PB performance, leading to stable or increased heterosis over time. For genetic models that included both dominance and AxA epistatic effects (D+E_{AA}), this mechanism resulted in positive heterosis, and heterosis was observed for at least 50 generations, regardless of the direction of selection on PB performance. These findings are in line with, and may at

least partly explain, the observed persistence of heterosis in a number of empirical (Sheridan and Randall 1977; Ayyagari *et al.* 1982; Minvielle *et al.* 1999; Yang *et al.* 1999) and simulation studies (Esfandyari *et al.* 2015b; Esfandyari *et al.* 2018).

7.2.5 Concluding remarks

The results presented in this section suggest that selection on PB performance can result in positive heterosis in the long-term. In other words, the response to selection can be larger for the correlated trait (CB performance) than for the trait under selection (PB performance). Note, however, that this (possibly counterintuitive) result holds only for the specific scenarios simulated here. I simulated selection on a single trait in only one of the parental lines, and assumed a certain genetic model (i.e. type and size of non-additive effects). There are a number of assumptions underlying the results that deserve to be mentioned.

First, the genetic distance between lines in my simulations may be different from genetic distances between PB lines in reality. Genetic distance increases when lines are divergently selected, resulting in larger differences in allele frequencies between lines. In my simulations, I considered scenarios where there was selection only in one of the parental lines, while the other line was unselected. In reality, each parental line would be selected for multiple traits that are combined in a selection index. Complicating things even further, the traits in such indices may differ between lines. Hence, if both lines in my simulations would have been selected in the same direction, allele frequency differences would be smaller, resulting in smaller observed heterosis. My results showed, however, that the combined effect of dominance and epistasis resulted in positive heterosis for any distance between PB lines. I therefore expect that, for traits that are under dominant and AxA epistatic gene action, heterosis will be present in various crosses between differentially selected PB lines.

Second, the actual genetic model of traits is largely unknown. For example, little is known about the importance of epistasis in the expression of complex traits. In my simulations, I included models with dominance, additive by additive epistasis, or both, because it has been shown that these types of non-additive effects contribute to heterosis in F1 crosses (Hill 1982; Willham and Pollak 1985; Melchinger *et al.* 2007b). Furthermore, parental lines were closely to moderately related. However, as parental lines become more distantly related, the (negative) contribution of AxA epistasis to heterosis becomes larger, relative to the (positive) contribution of dominance, which can result in reduced heterosis or even outbreeding depression (Lynch 1991). Hence, for lines that are more distantly

related, heterosis may be smaller than observed in my simulations. For closely or moderately related lines, the results from my simulations may provide valuable insight into the genetic basis of heterosis and r_{pc}.

Finally, in reality, the CB animals are housed in a different environment than the PB lines, introducing two issues. First, differences in environment between PB and CB animals may lead to genotype by environment interaction (GxE). The presence of GxE results in lower values of r_{pc} and therefore a smaller response to selection in CB performance when selection is on PB performance, reducing realized heterosis. Second, when measured in real data, heterosis may appear lower than the actual heterosis arising from non-additive effects, because the CB animals are usually kept under more challenging conditions than the PB animals. It may therefore be difficult to quantify heterosis in real data when PB and CB animals are housed in different environments.

7.3 Genomic prediction for crossbred performance

The ultimate aim of a crossbred breeding program is to improve the performance of commercial CB animals. Hence, selection decisions in the PB lines should ideally be based on breeding values for CB performance ($\mathbf{a}_{CB}$). In this section, I will investigate and discuss strategies for estimating $\mathbf{a}_{CB}$.

Throughout this section, I will consider a crossbreeding program where animals from line A are mated to animals from line B to produce two-way CB animals (AB). For selection candidates in line A, breeding values for CB performance can be denoted as

$$\mathbf{a}_{CB} = \mathbf{Z}_{PB}\boldsymbol{\alpha}_{CB}, \qquad \textbf{7.2}$$

where $\mathbf{Z}_{PB}$ is a QTL genotype matrix with allele counts (i.e. genotypes) of selection candidates in line A, and $\boldsymbol{\alpha}_{CB}$ is a vector of average effects for CB performance at those QTL in line A. For a single locus, α_{CB} is the average effect of a locus in line A, given that line A is mated to line B to produce crossbred animals (AB) in the next generation. There are many ways of defining average effects. For CB performance, it is convenient to define α_{CB} in terms of transmitted value of alleles, because interest is in the value of alleles in the next generation of crossbreds. This definition states that the average effect *of an allele* is the average genotypic value of CB offspring produced by transmitting that allele (Falconer 1985). For a locus with two alleles, α_{CB} is equal to the average effect of the alternative allele minus the average effect of the reference allele at that locus, where the decision of which allele is considered alternative is arbitrary. This procedure is equivalent to performing a linear regression

of genotypic values of CB animals on the number of alternative alleles these CB animals inherited from line A (coded as {0, 1}). The resulting regression coefficient is equal to α_{CB}.

In practice, the genotypic values of CB animals cannot be directly observed, and the genotypes at quantitative trait loci (QTL) are usually unknown. It is therefore impossible to get true values of α_{CB} and to compute $\mathbf{a}_{CB}$. However, $\mathbf{a}_{CB}$ can be estimated from data with an approach called *genomic prediction.* Genomic prediction uses a so-called reference population that consists of animals that have both phenotype and marker genotype data. The phenotypes act as a proxy for genotypic values, and the marker genotypes act as a proxy for QTL genotypes. In general, the first step is to estimate average effects of all markers ($\hat{\boldsymbol{\alpha}}$) simultaneously, using multiple linear regression of phenotypes on the number of alternative alleles at each marker (i.e. allele counts). The $\hat{\boldsymbol{\alpha}}$ at markers capture the average effects at QTL, because the markers are believed to be in linkage disequilibrium (LD) with the QTL. These $\hat{\boldsymbol{\alpha}}$ are used to compute genomic estimated breeding values (GEBV) of selection candidates that have marker genotypes available (Meuwissen *et al.* 2001). For selection candidates in line A, GEBV for CB performance can be denoted as

$$\mathbf{GEBV}_{CB} = \mathbf{M}_{PB}\hat{\boldsymbol{\alpha}}_{CB}, \qquad \textbf{7.3}$$

where $\mathbf{M}_{PB}$ is a marker genotype matrix with allele counts, and $\hat{\boldsymbol{\alpha}}_{CB}$ is a vector of estimated average effects for CB performance at those markers. As I discussed before, α_{CB} at QTL are equal to the regression coefficient of CB genotypic values on the number of alternative QTL alleles they inherited from line A. This suggests that $\hat{\alpha}_{CB}$ at markers should be estimated by linear regression of CB phenotypes on the number of alternative marker alleles those crossbreds inherited from line A. This approach requires that the alleles in the CB animals can be traced back to their line of origin (i.e. the breed-of-origin of alleles (BOA)). Hereafter, I will call this genomic prediction model *the BOA model* (abbreviated as MA-BOA).

When selection decisions are based on $\mathbf{GEBV}_{CB}$, the response to selection partly depends on the accuracy of $\mathbf{GEBV}_{CB}$ (ρ_{CB}), which is the correlation between $\mathbf{GEBV}_{CB}$ and $\mathbf{a}_{CB}$. The ρ_{CB} depends on how accurate $\hat{\boldsymbol{\alpha}}_{CB}$ are, which at least partly depends on the genomic prediction strategy that is used. The foregoing suggests that, in theory, strategy MA-BOA maximizes ρ_{CB}, because that strategy is parallel to the definition of average effects for CB performance. Alternative strategies may, however, result in similar or slightly lower ρ_{CB}, while those strategies may be more easily implemented in practice and less costly. For example, the collection of genotype and phenotype data on CB animals may not be optimal in terms of using

available resources, because CB animals themselves are not part of the breeding program. Furthermore, the MA-BOA approach requires that the BOA is determined, which can be computationally challenging and may not be fully accurate (Vandenplas *et al.* 2016). Hence, alternative strategies may be preferred over MA-BOA when the expected benefit of MA-BOA does not outweigh the extra resources needed.

In the following subsections, I will assume that the aim is to estimate $\mathbf{GEBV}_{CB}$ for selection candidates in only one of the parental lines of a two-way cross (as opposed to estimating $\mathbf{GEBV}_{CB}$ for both parental lines simultaneously). I will list alternative genomic prediction strategies for estimating $\mathbf{GEBV}_{CB}$, and discuss their strengths and weaknesses. These strategies differ in the type of animals in the reference population (i.e. PB or CB animals), and in the way the data are modelled (Table 7.1). I will start with the most basic strategy that is easy to implement in terms of data collection and modelling, and continue with strategies that require the collection of new data or more complex modelling.

Table 7.1 Strategies for estimating GEBV for CB performance. The first column indicates the type of animals in the reference population, the second column indicates the model used. The four right-most columns indicate whether the strategies (in theory) account for dominance, epistasis, GxE, and differences in LD phase between parental lines.

		Accounts for[1]			
Reference population	Model[2]	dominance	epistasis	GxE	LD
PB	MA-PB				✓
	MAD-PB	✓			✓
CB	MA-CB	~	~	✓	
	MA-BOA	✓	✓	✓	✓

[1] ✓ = correctly accounts for, ~ = partially accounts for. [2]MA = additive model, MAD = additive plus dominance model, -PB = using a PB reference population, -CB = using a CB reference population, -BOA = considering the breed-origin of alleles in CB.

7.3.1 Using PB data

The first strategy for estimating $\mathbf{GEBV}_{CB}$ is to use a reference population that consists of PB animals that are from the same line as the selection candidates, and to apply a standard within-line additive genomic prediction model (MA-PB). Using MA-PB effectively results in GEBV for PB performance ($\mathbf{GEBV}_{PB}$). Hence, MA-PB assumes that the average effects for CB performance (α_{CB}) are the same as those for PB performance (α_{PB}). The accuracy of $\mathbf{GEBV}_{CB}$ (ρ_{CB}) with MA therefore depends on the correlation between these average effects, which is measured by

the genetic correlation between PB and CB performance (r_{pc})[5]. The r_{pc} can be lower than one due to the following reasons:

a. Genotype by genotype interaction (GxG, i.e. non-additive genetic effects): In the presence of non-additive effects (i.e. dominance and epistasis), the effects of alleles depend on the genetic background they are expressed in. Effectively, this results in a dependency of the average effects of loci on the allele frequencies in the population. In a two-way CB breeding program, the two purebred parental lines may have different allele frequencies, and in such case the allele frequencies in the parental lines differ from those in the CB as well. This mechanism may result in differences in average effects for PB and CB performance in the parental line. The relationship between non-additive effects, differences in allele frequencies between parental lines, and r_{pc} was studied in chapter 4.

b. Genotype by environment interaction (GxE): In addition to dependency on genetic background, the effects of alleles may depend on the environment they are expressed in. Usually, PB animals are kept in a nucleus environment under excellent management conditions and high levels of biosecurity (e.g. disease-free), whereas CB animals are kept under field conditions with varying levels of management quality and lower levels of biosecurity. Alleles that have no effect or that are harmful for the trait of interest in the nucleus environment may be beneficial in the conventional environment, and vice versa.

In practice, both GxG and GxE interactions lead to values of r_{pc} lower than one. In a review article on estimates of r_{pc} in pigs, the average r_{pc} was ~0.63, and the majority of estimates was above 0.5 (Wientjes and Calus 2017). Similar estimates of r_{pc} were found in beef cattle, where r_{pc} for body weight ranged from 0.64 to 0.84 (Lukaszewicz *et al.* 2015), and for growth and carcass traits from 0.48 to 1.00 (Newman *et al.* 2002). In laying hens, r_{pc} estimates were in a similar range, with estimates from 0.56 to 0.99 for egg production traits (Wei and van der Werf 1995; Besbes and Gibson 1999), and around 0.70 for eggshell color (Mulder *et al.* 2016).

In chapter 5, I estimated r_{pc} for body weight at 7 (BW7) and 35 days (BW35) in broilers, using models based on either pedigree or genomic information. For BW7,

[5] Strictly speaking, r_{pc} is the correlation between breeding values for PB and CB performance, but this correlation is equivalent to the correlation between average effects under some assumptions (e.g. Wientjes *et al.* (2017))

r_{pc} estimates ranged from 0.64 to 0.80, whereas for BW35, estimates ranged from 0.90 to 0.96. It should be noted, however, that the PB and CB animals in this study were housed in the same environment, so r_{pc} was lower than one due to GxG interactions only. For PB animals placed in a PB and CB environment, estimated genetic correlations for body weight between environments were around 0.5, suggesting a strong effect of GxE interaction (Chu *et al.* 2019). In summary, it seems that most estimates of r_{pc} across species and traits are lower than one, and usually larger than 0.5. When r_{pc} is substantially lower than one, strategy MA-PB may lead to suboptimal ρ_{CB}, and reduce the response to selection in CB performance. In such cases, it may be beneficial to refine the genomic model to (partially) account for $r_{pc} < 1$, as explained in the following, or to collect data on CB animals.

7.3.1.1 Dominance model

One of the reasons for $r_{pc} < 1$ is the existence of dominance effects in combination with differences in allele frequency between parental lines (Wei *et al.* 1991; chapter 4). Strategy MA-PB can therefore be improved by accounting for dominance in the model. With only dominance, the average effect of a locus depends on the allele frequency of that locus in the mated individuals (Falconer and Mackay 1996). This introduces differences between average effects for PB (α_{PB}) and CB performance (α_{CB}), because α_{PB} depends on the allele frequency in the line of selection candidates (line A), whereas α_{CB} depends on allele frequency in the mated line (line B) (chapter 4, Pirchner and Mergl (1977), Dekkers (1999)). Using PB data, ρ_{CB} may therefore be improved by estimating the additive (a) and dominance (d) effects separately, and then use the allele frequency in line B to compute $\hat{\alpha}_{CB}$. This approach is hereafter called the dominance model (abbreviated as MAD-PB). It has been shown with simulations that MAD-PB results in a higher cumulative response to selection than MA-PB when the amount of dominance resembles what is observed in real data (Kinghorn 2010; Zeng *et al.* 2013; Esfandyari *et al.* 2015b). Furthermore, in a study on real data of litter size in pigs, MAD-PB resulted in higher ρ_{CB} than MA-PB (Esfandyari *et al.* 2016).

7.3.2 Using CB data

Another reason for $r_{pc} < 1$ are genotype by environment interactions (GxE). When estimation of $\mathbf{GEBV}_{CB}$ is based on PB data that is collected in a nucleus environment, it is not possible to account for GxE. A solution may be to test some PB animals in a commercial crossbred environment to account for GxE (e.g. Chu *et al.* 2018). However, PB animals tested in a commercial environment cannot be used as

selection candidates anymore, because they cannot be returned to the nucleus environment for bio-security reasons. Furthermore, when a group of PB animals is tested in a commercial environment, the selection intensity should be reduced, resulting in a smaller response to selection (Chu *et al.* 2018). Alternatively, the ρ_{CB} may be improved by using a CB instead of a PB reference population. With a CB reference population and an additive genomic prediction model (MA-CB), α_{CB} can be estimated in the environment the transmitted alleles are expressed in. This approach therefore accounts for differences between α_{PB} and α_{CB} due to GxE interactions (Table 7.1). In addition, the MA-CB model accounts for differences between α_{PB} and α_{CB} due to epistatic interactions that create a dependency between average effects and the allele frequencies of other loci, such as additive by additive and multiplicative epistasis (chapter 4). MA-CB does not, however, correctly account for dominance, because with dominance, α_{CB} depends on the allele frequency in the mated line.

In simulations, a CB reference population yielded higher ρ_{CB} than a PB reference population when r_{pc} was lower than ~0.8 (Dekkers 2007; Esfandyari *et al.* 2015a; Van Grevenhof and Van Der Werf 2015). In agreement with this result, a CB reference population yielded lower ρ_{CB} than a PB reference population for a trait with an r_{pc} of 0.90 in pigs (Hidalgo *et al.* 2016). Note, however, that in this latter study, the PB reference population was larger than the CB reference population. A fair comparison between MA-PB and MA-CB requires that the number of PB and CB animals used for the reference population are similar, and that the relationship between selection candidates and the two reference populations are comparable. These two requirements were met in my study on broiler chicken (chapter 6), where a CB reference population yielded higher ρ_{CB} than a PB reference population for a trait with an r_{pc} of 0.8, but not for a trait with an r_{pc} of 0.96. In summary, the results from these studies suggest that MA-CB can improve ρ_{CB} compared to MA-PB for traits with an r_{pc} lower than ~0.8 (as already suggested by Robertson (1959)).

With MA-CB, $\hat{\boldsymbol{\alpha}}_{CB}$ are estimated by linear regression of CB phenotypes on the number of alternative alleles those CB animals have. MA-CB therefore results in estimated average effects for CB performance in the *crossbreds*, whereas we need the average effects for CB performance in the *purebreds* (see Equation 7.2). In addition, MA-CB does not consider the origin of alleles, and therefore assumes that the effects of alleles in CB animals are independent of line origin. This assumption may not hold because (1) with dominance or some forms of epistasis, effects of alleles for CB performance within a parental line depend on the allele frequencies in the mated line instead of those in the CB, and (2) parental lines may differ in linkage

disequilibrium (LD) between markers and QTL, introducing differences in the apparant effects of marker alleles from different origins[6]. Hence, when dominance is the main cause of $r_{pc} < 1$, MAD-PB may yield higher ρ_{CB} than MA-CB. Despite this shortcoming, MA-CB may still be preferred over MA-PB when GxE and epistasis are important, and when LD between lines is consistent.

7.3.2.1 Considering the breed-origin of alleles (BOA)

In the previous section, I discussed that MA-CB does not allow for breed specific effects of marker alleles, while in reality, the effects of marker alleles in crossbreds may depend on their origin. The model may therefore be improved by considering the breed-origin of alleles (BOA), leading to the BOA model (MA-BOA). In theory, MA-BOA accounts for all factors that result in an r_{pc} lower than one, and assumes that LD in parental lines are uncorrelated (i.e. the correlation of LD phase between parental lines is zero) (Table 7.1). Theoretically, this strategy may therefore yield a higher ρ_{CB} than other strategies that use either a PB or CB reference population, because MA-BOA agrees best with genetic reality. Compared to MA-CB, MA-BOA is expected to improve ρ_{CB} when (1) there is substantial dominance and epistasis in combination with large differences in allele frequency between lines, or (2) there are substantial differences in LD between parental lines (i.e. low correlation of LD phase) (Table 7.1). In other words, MA-BOA is expected to improve ρ_{CB} when the parental lines are distantly related.

Ibañez-Escriche *et al.* (2009) illustrated the benefits of considering the BOA with simulations where GxE and GxG interactions were absent ($r_{pc} = 1$), and breeds only differed in allele frequencies and in LD between markers and QTL. This study showed that ρ_{CB} was higher with MA-BOA than with MA-CB, for scenarios where marker density was low (500 markers per Morgan), the number of CB animals in the reference population was relatively large (4,000), and when breeds were distantly related. In other simulation studies where dominance effects were simulated, ρ_{CB} was higher and the response in CB performance was larger with MA-BOA compared to MA-CB (Kinghorn 2010; Esfandyari *et al.* 2015a). In the study of Esfandyari *et al.* (2015a), a benefit of MA-BOA over MA-CB was only observed when the parental lines were distantly related and the CB reference population was small. In a study on three-way CB pigs, a benefit of MA-BOA over MA-CB was observed for a trait with an r_{pc} of about 0.30, but not for traits with an r_{pc} between 0.55 and 0.73 (Sevillano *et*

[6] Note that I assume that functional additive, dominance, and epistatic genetic effects (i.e. a, d, and ϵ) are independent of line origin, and that CB performance is affected by the same QTL in both parental lines.

al. 2017). These results agree with my results from chapter 6, where a benefit of MA-BOA was observed for a trait with an r_{pc} of 0.8, but not for a trait with an r_{pc} of 0.96. In summary, these results suggest that considering the BOA may increase ρ_{CB} when parental lines are distantly related, and when traits have a relatively low r_{pc}.

The results above suggest that a low r_{pc} may indicate a possible benefit of MA-BOA over MA-CB. However, the benefit of MA-BOA depends on the reason for low r_{pc}. For example, when r_{pc} is lower than one due to GxE interactions alone, MA-BOA is not expected to yield higher ρ_{CB} than MA-CB, because GxE interactions do not necessarily lead to breed-specific effects of alleles in the crossbreds (Ibañez-Escriche *et al.* 2009). Instead, Sevillano (2018) suggested that a benefit of MA-BOA over MA-CB can be expected when there is a difference in estimated r_{pc} between models that either ignore or consider the BOA. In theory, there is indeed a connection between differences in estimated r_{pc} and the benefit of MA-BOA, because such a difference indicates that parental lines are distantly related, and that therefore a benefit of considering the BOA may be expected. However, the results in this thesis neither supported nor contradicted this theory: in chapter 5, there was no difference in estimated r_{pc} between MA-BOA and MA-CB, while MA-BOA did result in higher ρ_{CB} than MA-CB (chapter 6). A possible explanation for this result is that the estimation of covariance between and variances of PB and CB performance with MA-CB is already dominated by the alleles originating from the line of interest, especially when data size is sufficient and parental lines are distantly related. This may result in similar estimates of r_{pc} with MA-CB and MA-BOA. However, the CB alleles that originate from the mated line may still introduce noise in the estimation of $\hat{\boldsymbol{\alpha}}_{CB}$ with MA-CB, leading to a higher ρ_{CB} with MA-BOA than with MA-CB.

In some studies that investigated the benefit of considering the BOA, average effects for CB performance were estimated for all parental lines simultaneously (Ibañez-Escriche *et al.* 2009; Lopes *et al.* 2017; Sevillano *et al.* 2017). In those cases, it has been argued that ignoring the BOA may yield the same or higher ρ_{CB} than considering the BOA, because considering the BOA requires the estimation of twice as many parameters (in a two-way crossbred setting) (Ibañez-Escriche *et al.* 2009). In other words, the estimation of effects when the BOA is considered is based on half the information, compared to when the BOA is ignored (Sevillano 2018). Although these arguments may be true, they do not necessarily lead to a benefit of ignoring the BOA over considering it, because a model that considers the BOA is expected to better fit the true genetic model than a model that ignores the BOA. Furthermore, in chapter 6, I used a model that considered the alleles from only one of the parental lines (i.e. MA-BOA), and ignored the alleles from the second parental line. Still, I

observed that MA-CB resulted in higher ρ_{CB} than MA-BOA, while the number of parameters estimated was the same in both models. I therefore expect that observed benefits of MA-CB over MA-BOA are not due to increased model complexity of MA-BOA compared to MA-CB. In the following paragraph, I will give a possible explanation for why the performance of MA-CB is very competitive, despite the expected superiority of MA-BOA based on theoretical arguments.

The difference in ρ_{CB} between MA-BOA and MA-CB may be determined by two mechanisms. On the one hand, MA-BOA may have a benefit over MA-CB, because MA-CB assumes that allele effects are independent of line origin. As discussed before, this assumption introduces noise in the estimation of $\widehat{\boldsymbol{\alpha}}_{CB}$ with MA-CB due to differences in LD and allele frequencies between parental lines. On the other hand, with MA-CB, $\widehat{\boldsymbol{\alpha}}_{CB}$ results from regression on genotypes (i.e. 0, 1, 2), whereas with MA-BOA, $\widehat{\boldsymbol{\alpha}}_{CB}$ results from regression on haplotypes (i.e. 0, 1). As a result, MA-CB may yield more accurate $\widehat{\boldsymbol{\alpha}}_{CB}$, because the variance explained by the regression is proportional to $2pq$, whereas with MA-BOA, the variance explained is proportional to pq. In general, it holds that the more variance is explained with regression, the greater the accuracy of the estimated regression coefficient (i.e. $\widehat{\boldsymbol{\alpha}}_{CB}$). The benefit of MA-BOA over MA-CB depends on whether the amount of noise that is introduced in $\widehat{\boldsymbol{\alpha}}_{CB}$ with MA-CB is large enough to overrule the advantage of MA-CB, due to the larger proportion of variance explained. As discussed before, the noise introduced with MA-CB depends on the correlation of LD phase and the differences in allele frequencies between parental lines, which depend on the genetic distance between those lines. In conclusion, MA-BOA may yield higher ρ_{CB} than MA-CB when lines are distantly related, whereas MA-BOA may yield lower ρ_{CB} than MA-CB when lines are more closely related.

The explanation for the competitive performance of MA-CB compared to MA-BOA in the previous paragraph is in agreement with the results in this thesis. In chapter 6, MA-CB led to higher ρ_{CB} than MA-BOA for a trait with high r_{pc} (0.96) and where parental lines were believed to be distantly related. In that study, I showed that alleles in the crossbreds that originated from the dam line had predictive value for $\mathbf{GEBV}_{CB}$ in the sire line. This result suggests that although the parental lines were separated for many generations, marker allele effects for CB performance in the two parental lines are correlated, and that therefore LD phase between marker and QTL are similar as well. To further test the similarity of LD phase, I correlated $\mathbf{GEBV}_{CB}$ of sires from a model that considered only dam alleles in a CB reference population, with GEBV of sires from a model that used a PB reference population ($\mathbf{GEBV}_{PB}$). I expected that this correlation would be zero when LD phase between parental lines

was uncorrelated. The results showed, however, that the average correlation across replicates was ~0.15, suggesting that the LD patterns are somewhat similar between parental lines. As a result, the noise introduced with MA-CB was not large enough to overrule the benefit of MA-CB over MA-BOA due to the larger proportion of variance explained, resulting in competitive performance of MA-CB.

7.3.3 Practical relevance

From the results of this thesis and previous studies, it seems that the accuracy of $\mathbf{GEBV}_{CB}$ can be improved by using CB data instead of PB data, for traits with an r_{pc} lower than ~0.8. In practice, r_{pc} may be lower than 0.8 for many species and traits, especially when GxE interactions are involved. I therefore think that crossbred breeding programs will benefit from the collection of phenotype and genotype data of CB animals, especially considering the fact that information on a single CB animal can contribute to estimation of $\mathbf{GEBV}_{CB}$ in all its parental lines. In addition, the collection of data on CB animals is necessary for traits that cannot be measured on PB selection candidates (e.g. disease traits). When collection of CB data is too costly or challenging, $\mathbf{GEBV}_{CB}$ should be estimated with models that use a PB reference population and account for dominance (MAD-PB). The benefit of MAD-PB over MA-PB has, however, not been studied extensively in real data, and should be investigated further.

At the beginning of this discussion, I hypothesized that a model that considers the BOA would result in higher ρ_{CB} than alternative strategies, because such a model agrees best with genetic reality. In practice, however, ignoring the BOA may sometimes yield a higher ρ_{CB} than considering it. I expect that there is a benefit of considering the BOA only when parental lines are distantly related (i.e. low correlation of LD phase), or when non-additive effects (especially dominance) are important components of a low r_{pc}. When lines are somewhat related, and non-additive effects are relatively unimportant, I expect no benefit of considering the BOA. In conclusion, despite the fact that considering the BOA improves the model's agreement with genetic reality, it does not necessarily lead to improved predictions.

In the previous sections, I have focused on strategies for estimating $\mathbf{GEBV}_{CB}$ with either a PB or CB reference population. In most situations, however, genotypes and phenotypes of PB animals are already available when a breeder decides to collect data on CB animals. In such cases, it may be beneficial to use both PB and CB information to estimate $\mathbf{GEBV}_{CB}$, instead of using only PB or CB information. With both PB and CB information, a bivariate model can be used that treats PB and CB performance as two separate, but correlated traits (Christensen *et al.* 2014). Such a

model leads to additive genetic values of PB selection candidates for both PB and CB performance. In general, it is expected that the benefit of combining PB and CB information instead of using only PB information decreases with increasing r_{pc} (Van Grevenhof and Van Der Werf 2015). In addition, considering the BOA is expected to be beneficial when parental lines are distantly related, possibly leading to differences between effects of alleles coming from different line origins (Sevillano *et al.* 2017). Hence, the benefits of *adding* CB information to PB information are probably similar to those of using CB *instead* of PB information, because of the same reasons described in the previous subsections. Similarly, the benefits of considering the BOA when PB and CB information is combined are expected to be similar to the benefits of MA-BOA over MA-CB. When CB information is available, I recommend to both use PB and CB information, and use a bivariate model that treats PB and CB performance as two separate, but correlated traits. In addition, I recommend to consider the BOA when dominance effects are expected to be large, and when parental lines show low correlation of LD phase and large differences in allele frequencies.

In this discussion, I considered a situation where two purebred parental lines were mated to produce a CB (i.e. a two-way crossbred). In reality, however, the final CB animals are usually a cross between three or four parental lines (i.e. a three- or four-way CB). For example, in pig breeding, breeding programs typically consist of a sire line (A) and two dam lines (B and C), and the commercial CB product is a result from mating line A with crossbreds BC. In general, the results and conclusions from this section apply to situations where $\mathbf{GEBV}_{CB}$ are computed in parental lines that are responsible for 50% of the genes in the commercial CB animals (e.g. in the sire line (A) of a three-way crossbred, A(BC)). With a four-way CB, however, parental lines are responsible for only 25% of the genes in the commercial CB animals. Hence, for estimation of $\mathbf{GEBV}_{CB}$ in one parental line, the predictive value of a two-way CB is larger than the predictive value of a four-way CB, because the proportion of genes shared with the parental line is twice as large for a two-way CB as for a four-way CB. However, for the breeding program as a whole, information on a four-way CB may be more valuable than information on a two-way CB, because the four-way CB can improve the estimation of $\mathbf{GEBV}_{CB}$ in four parental lines, whereas the two-way CB can improve $\mathbf{GEBV}_{CB}$ in only two parental lines. Despite these differences, it holds that the benefit of a four-way CB reference population over a PB reference population increases with decreasing r_{pc}, and the benefit of considering the BOA over ignoring it increases with increasing genetic distance between parental lines.

References

References

Aliloo H., J.E. Pryce, O. González-Recio, B.G. Cocks, and B.J. Hayes, 2015 Validation of markers with non-additive effects on milk yield and fertility in Holstein and Jersey cows. BMC Genet 16:89.

Álvarez-Castro J.M., and Ö. Carlborg, 2007 A Unified Model for Functional and Statistical Epistasis and Its Application in Quantitative Trait Loci Analysis. Genetics 176:1151-1167.

Ayyagari V., S.C. Mohapatra, P. Renganathan, D.C. Johri, T.S. Thiagasundaram, and B.K. Panda, 1982 Cross performance as influenced by selection in pure lines. British Poultry Science 23:469-479.

Bateson W., and G. Mendel, 1909 *Mendel's principles of heredity*. Cambridge [Eng.]: University Press.

Baumung R., J. Solkner, and A. Essl, 1997 Correlation between purebred and crossbred performance under a two-locus model with additive by additive interaction. J Anim Breed Genet 114:89-98.

Bennewitz J., C. Edel, R. Fries, T.H.E. Meuwissen, and R. Wellmann, 2017 Application of a Bayesian dominance model improves power in quantitative trait genome-wide association analysis. Genet Sel Evol 49:7.

Bennewitz J., and T.H.E. Meuwissen, 2010 The distribution of QTL additive and dominance effects in porcine F2 crosses. J Anim Breed Genet 127:171-179.

Berenos C., P.A. Ellis, J.G. Pilkington, and J.M. Pemberton, 2014 Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. Mol Ecol 23:3434-3451.

Besbes B., and J.P. Gibson, 1999 Genetic variation of egg production traits in purebred and crossbred laying hens. Animal Science 68:433-439.

Bidanel J.P., 1993 Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. III. Dominance and epistatic components of heterosis on reproductive traits. Genet Sel Evol 25:263.

Bijma P., and J.W. Bastiaansen, 2014 Standard error of the genetic correlation: how much data do we need to estimate a purebred-crossbred genetic correlation? Genet Sel Evol 46:79.

Bijma P., and J.A.M. van Arendonk, 1998 Maximizing genetic gain for the sire line of a crossbreeding scheme utilizing both purebred and crossbred information. Anim Sci 66:529-542.

Bijma P., J.A. Woolliams, and J.A.M. Van Arendonk, 2001 Genetic gain of pure line selection and combined crossbred purebred selection with constrained inbreeding. Animal Science 72:225-232.

Brown B.C., C.J. Ye, A.L. Price, and N. Zaitlen, 2016 Transethnic Genetic-Correlation Estimates from Summary Statistics. Am J Hum Genet 99:76-88.

Calus M.P.L., J. Vandenplas, B. Hulsegge, R. Borg, J.M. Henshall, and R. Hawken, 2018 Derivation of parentage and breed-origin of alleles in a crossbred broiler dataset in *Proceedings of the 11th World Congress on Genetics Applied to Livestock Production*, Auckland, New Zealand.

Cameron N.D., 1997 *Selection indices and prediction of genetic merit in animal breeding*. Wallingford: CAB INTERNATIONAL.

Carillier C., H. Larroque, and C. Robert-Granié, 2014 Comparison of joint versus purebred genomic evaluation in the French multi-breed dairy goat population. Genet Sel Evol 46:67.

Carlborg O., S. Kerje, K. Schutz, L. Jacobsson, P. Jensen, and L. Andersson, 2003 A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome Res 13:413-421.

Carter A.J., J. Hermisson, and T.F. Hansen, 2005 The role of epistatic gene interactions in the response to selection and the evolution of evolvability. Theor Popul Biol 68:179-196.

Charlesworth D., and J.H. Willis, 2009 The genetics of inbreeding depression. Nat Rev Genet 10:783.

Cheverud J.M., and E.J. Routman, 1995 Epistasis and its contribution to genetic variance components. Genetics 139:1455-1461.

Cheverud J.M., and E.J. Routman, 1996 Epistasis as a source of increased additive genetic variance at population bottlenecks. Evolution 50:1042-1051.

Christensen O., A. Legarra, M. Lund, and G. Su, 2015 Genetic evaluation for three-way crossbreeding. Genet Sel Evol 47:98.

Christensen O.F., P. Madsen, B. Nielsen, and G. Su, 2014 Genomic evaluation of both purebred and crossbred performances. Genet Sel Evol 46:23.

Chu T.T., S.W. Alemu, E. Norberg, A.C. Sørensen, J. Henshall, R. Hawken, and J. Jensen, 2018 Benefits of testing in both bio-secure and production environments in genomic selection breeding programs for commercial broiler chicken. Genet Sel Evol 50:52.

Chu T.T., J.W.M. Bastiaansen, P. Berg, H. Romé, D. Marois, J. Henshall, and J. Jensen, 2019 Use of genomic information to exploit genotype-by-environment interactions for body weight of broiler chicken in bio-secure and production environments. Genet Sel Evol 51:50.

Cleveland M.A., J.M. Hickey, and S. Forni, 2012 A Common Dataset for Genomic Analysis of Livestock Populations. G3: Genes|Genomes|Genetics 2:429-435.

Cockerham C.C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. Genetics 39:859.

Crow J.F., 1999 Dominance and Overdominance, pp. 49-58 in *The genetics and exploitation of heterosis in crops*, edited by J.G. Coors and S. Pandey. American Society of Agronomy, Crop Science Society of America, Madison, WI.

Davenport C.B., 1908 Degeneration, albinism and inbreeding. Science 28:454-455.

de Candia T.R., S.H. Lee, J. Yang, B.L. Browning, P.V. Gejman, D.F. Levinson, B.J. Mowry, J.K. Hewitt, M.E. Goddard, M.C. O'Donovan, S.M. Purcell, D. Posthuma, P.M. Visscher, N.R. Wray, and M.C. Keller, 2013 Additive genetic variation in schizophrenia risk is shared by populations of African and European descent. Am J Hum Genet 93:463-470.

de los Campos G., D. Sorensen, and D. Gianola, 2015 Genomic Heritability: What Is It? PLoS Genet 11:e1005048.

de los Campos G., D.A. Sorensen, and M.A. Toro, 2019 Imperfect Linkage Disequilibrium Generates Phantom Epistasis (&Perils of Big Data). G3: Genes|Genomes|Genetics:g3.400101.402019.

de Roos A.P., B.J. Hayes, and M.E. Goddard, 2009 Reliability of genomic predictions across multiple populations. Genetics 183:1545-1553.

de Roos A.P., B.J. Hayes, R.J. Spelman, and M.E. Goddard, 2008 Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics 179:1503-1512.

Dekkers J.C., 2007 Marker-assisted selection for commercial crossbred performance. J Anim Sci 85:2104-2114.

Dekkers J.C.M., 1999 Breeding values for identified quantitative trait loci under selection. Genet Sel Evol 31:421.

Denning P.J., and T.G. Lewis, 2016 Exponential laws of computing growth. Commun. ACM 60:54-65.

Dickerson G.E., 1973 Inbreeding and heterosis in animals. J Anim Sci 1973:54-77.

Duenk P., P. Bijma, M.P.L. Calus, Y.C.J. Wientjes, and J.H.J. van der Werf, 2019a The impact of non-additive effects on the genetic correlation between populations.

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, V.P. Breen, J.M. Henshall, R. Hawken, and P. Bijma, 2019b Estimating the purebred-crossbred genetic correlation of body weight in broiler chickens with pedigree or genomic relationships. Genet Sel Evol 51:6.

East E., 1908 Inbreeding in corn, pp. pp. 419–428 in *Reports of the Connecticut Agricultural Experiment Station for Years 1907–1908*, Connecticut Agricultural Experiment Station, New Haven, CT.

Eggen A., 2012 The development and application of genomic selection as a new breeding paradigm. Animal Frontiers 2:10-15.

Erbe M., B.J. Hayes, L.K. Matukumalli, S. Goswami, P.J. Bowman, C.M. Reich, B.A. Mason, and M.E. Goddard, 2012 Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J Dairy Sci 95:4114-4129.

Ertl J., A. Legarra, Z.G. Vitezica, L. Varona, C. Edel, R. Emmerling, and K.U. Gotz, 2014 Genomic analysis of dominance effects on milk production and conformation traits in Fleckvieh cattle. Genet Sel Evol 46:40.

Esfandyari H., P. Berg, and A.C. Sørensen, 2018 Balanced selection on purebred and crossbred performance increases gain in crossbreds. Genet Sel Evol 50:8.

Esfandyari H., P. Bijma, M. Henryon, O.F. Christensen, and A.C. Sorensen, 2016 Genomic prediction of crossbred performance based on purebred Landrace and Yorkshire data using a dominance model. Genet Sel Evol 48:40.

Esfandyari H., M. Henryon, P. Berg, J.R. Thomasen, P. Bijma, and A.C. Sorensen, 2017 Response to Selection in Finite Locus Models with Nonadditive Effects. J Hered 108:318-327.

Esfandyari H., A.C. Sorensen, and P. Bijma, 2015a A crossbred reference population can improve the response to genomic selection for crossbred performance. Genet Sel Evol 47:76.

Esfandyari H., A.C. Sorensen, and P. Bijma, 2015b Maximizing crossbred performance through purebred genomic selection. Genet Sel Evol 47:16.

Falconer D., and T. Mackay, 1996 *Introduction to quantitative genetics*. Harlow: Pearson Education Limited.

Falconer D.S., 1952 The Problem of Environment and Selection. The American Naturalist 86:293-298.

Falconer D.S., 1985 A note on Fisher's 'average effect' and 'average excess'. Genet Res (Camb) 46:337-347.

Fisher R.A., 1918 The Correlation between Relatives on the Supposition of Mendelian Inheritance. Trans. R. Soc. Edinburgh 52:399-433.

Fisher R.A., 1941 Average excess and average effect of a gene substitution. A Eug 11:53-63.

Forsberg S.K.G., J.S. Bloom, M.J. Sadhu, L. Kruglyak, and Ö. Carlborg, 2017 Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. Nat Genet 49:497-503.

Fu W., J.C. Dekkers, W.R. Lee, and B. Abasht, 2015 Linkage disequilibrium in crossbred and pure line chickens. Genet Sel Evol 47:11.

Garcia-Cortes L.A., and M.A. Toro, 2006 Multibreed analysis by splitting the breeding values. Genet Sel Evol 38:601-615.

Gianola D., G. de los Campos, M.A. Toro, H. Naya, C.C. Schon, and D. Sorensen, 2015 Do Molecular Markers Inform About Pleiotropy? Genetics 201:23-29.

Goddard M., 2009 Genomic selection: prediction of accuracy and maximisation of long term response. Genetica 136:245-257.

Goddard M.E., 2001 The validity of genetic models underlying quantitative traits. Livest Prod Sci 72:117-127.

Goddard M.E., B.J. Hayes, and T.H.E. Meuwissen, 2011 Using the genomic relationship matrix to predict the accuracy of genomic selection. J Anim Breed Genet 128:409-421.

Gondro C., J. van der Werf, and B. Hayes, 2013 *Genome-Wide Association Studies and Genomic Prediction*: Humana Press.

Goodnight C.J., 1999 Epistasis and Heterosis, pp. 59-68 in *The genetics and exploitation of heterosis in crops*, edited by J.G. Coors and S. Pandey. American Society of Agronomy, Crop Science Society of America, Madison, WI.

Haldane J.B.S., 1954 An exact test for randomness of mating. J Genet 52:631-635.

Hall S.J.G., 2016 Effective population sizes in cattle, sheep, horses, pigs and goats estimated from census and herdbook data. Animal 10:1778-1785.

Hallander J., and P. Waldmann, 2007 The effect of non-additive genetic interactions on selection in multi-locus genetic models. Heredity 98:349.

Hansen T.F., 2013 Why epistasis is important for selection and adaptation. Evolution 67:3501-3511.

Hayes B., 2013 Overview of Statistical Methods for Genome-Wide Association Studies (GWAS) in *Genome-Wide Association Studies and Genomic Prediction*. Humana Press.

Hayes B.J., P.J. Bowman, A.C. Chamberlain, K. Verbyla, and M.E. Goddard, 2009a Accuracy of genomic breeding values in multi-breed dairy cattle populations. Genet Sel Evol 41:51.

Hayes B.J., P.J. Bowman, A.J. Chamberlain, and M.E. Goddard, 2009b Invited review: Genomic selection in dairy cattle: progress and challenges. J Dairy Sci 92:433-443.

Hayes B.J., P.M. Visscher, and M.E. Goddard, 2009c Increased accuracy of artificial selection by using the realized relationship matrix. Genet Res (Camb) 91:47-60.

Hickey J.M., B.P. Kinghorn, B. Tier, J.F. Wilson, N. Dunstan, and J.H.J. van der Werf, 2011 A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. Genet Sel Evol 43:12-12.

Hidalgo A.M., J.W. Bastiaansen, M.S. Lopes, M.P. Calus, and D.J. de Koning, 2016 Accuracy of genomic prediction of purebreds for cross bred performance in pigs. J Anim Breed Genet 133:443-451.

Hill W.G., 1982 Dominance and epistasis as components of heterosis. Zeitschrift für Tierzüchtung und Züchtungsbiologie 99:161-168.

Hill W.G., 2008 Estimation, effectiveness and opportunities of long term genetic improvement in animals and maize. Lohmann information 43:3-20.

Hill W.G., 2016 Is Continued Genetic Improvement of Livestock Sustainable? Genetics 202:877-881.

Hill W.G., 2017 “Conversion” of epistatic into additive genetic variance in finite populations and possible impact on long-term selection response. J Anim Breed Genet 134:196-201.

Hill W.G., M.E. Goddard, and P.M. Visscher, 2008 Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet 4:e1000008.

Huang X., S. Yang, J. Gong, Q. Zhao, Q. Feng, Q. Zhan, Y. Zhao, W. Li, B. Cheng, J. Xia, N. Chen, T. Huang, L. Zhang, D. Fan, J. Chen, C. Zhou, Y. Lu, Q. Weng, and B. Han, 2016 Genomic architecture of heterosis for yield traits in rice. Nature 537:629-633.

Huang X., S. Yang, J. Gong, Y. Zhao, Q. Feng, H. Gong, W. Li, Q. Zhan, B. Cheng, J. Xia, N. Chen, Z. Hao, K. Liu, C. Zhu, T. Huang, Q. Zhao, L. Zhang, D. Fan, C. Zhou, Y. Lu, Q. Weng, Z.X. Wang, J. Li, and B. Han, 2015 Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. Nat Commun 6:6258.

Ibañez-Escriche N., R.L. Fernando, A. Toosi, and J.C. Dekkers, 2009 Genomic selection of purebreds for crossbred performance. Genet Sel Evol 41:12.

Ibanez-Escriche N., J. Reixach, N. Lleonart, and J.L. Noguera, 2011 Genetic evaluation combining purebred and crossbred data in a pig breeding scheme. J Anim Sci 89:3881-3889.

Iheshiulor O.O.M., J.A. Woolliams, X. Yu, R. Wellmann, and T.H.E. Meuwissen, 2016 Within- and across-breed genomic prediction using whole-genome sequence and single nucleotide polymorphism panels. Genet Sel Evol 48:1-15.

Johnson T., and N. Barton, 2005 Theoretical models of selection and mutation on quantitative traits. Philos Trans R Soc Lond B Biol Sci 360:1411-1425.

Joshi R., J.A. Woolliams, T. Meuwissen, and H.M. Gjoen, 2018 Maternal, dominance and additive genetic effects in Nile tilapia; influence on growth, fillet yield and body size traits. Heredity 120:452-462.

Kapell D.N., W.G. Hill, A.M. Neeteson, J. McAdam, A.N. Koerhuis, and S. Avendano, 2012 Genetic parameters of foot-pad dermatitis and body weight in purebred broiler lines in 2 contrasting environments. Poult Sci 91:565-574.

Karoui S., M.J. Carabaño, C. Díaz, and A. Legarra, 2012 Joint genomic evaluation of French dairy cattle breeds using multiple-trait models. Genet Sel Evol 44:39.

Kempthorne O., 1954 The Correlation between Relatives in a Random Mating Population. Proc R Soc Lond Ser B Biol sci 143:103-113.

Kimura M., and J.F. Crow, 1963 The Measurement of Effective Population Number. Evolution 17:279-288.

Kimura M., and J.F. Crow, 1964 The number of alleles that can be maintained in a finite population. Genetics 49:725-738.

Kinghorn B., 1982 Genetic effects in crossbreeding. Zeitschrift für Tierzüchtung und Züchtungsbiologie 99:59-68.

Kinghorn B.P.H., J.M.; van der Werf, J.H.J., 2010 Reciprocal recurrent genomic selection (RRGS) for total genetic merit in crossbred individuals in *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Leipzig.

Knol E.F., B. Nielsen, and P.W. Knap, 2016 Genomic selection in commercial pig breeding. Animal Frontiers 6:15-22.

Koerhuis A., and R. Thompson, 1997 Models to estimate maternal effects for juvenile body weight in broiler chickens. Genet Sel Evol 29:225.

Le Rouzic A., J.M. Alvarez-Castro, and O. Carlborg, 2008 Dissection of the genetic architecture of body weight in chicken reveals the impact of epistasis on domestication traits. Genetics 179:1591-1599.

Lee S.H., M.E. Goddard, P.M. Visscher, and J.H. van der Werf, 2010 Using the realized relationship matrix to disentangle confounding factors for the estimation of genetic variance components of complex traits. Genet Sel Evol 42:22.

Lee S.H., and J.H.J. van der Werf, 2016 MTG2: an efficient algorithm for multivariate linear mixed model analysis based on genomic information. Bioinformatics 32:1420-1422.

Legarra A., 2016 Comparing estimates of genetic variance across different relationship models. Theor Popul Biol 107:26-30.

Legarra A., and A. Reverter, 2018 Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. Genet Sel Evol 50:53.

Lewontin R.C., and K. Kojima, 1960 The evolutionary dynamics of complex polymorphisms. Evolution 14:458-472.

Lo L.L., R.L. Fernando, and M. Grossman, 1997 Genetic evaluation by BLUP in two-breed terminal crossbreeding systems under dominance. J Anim Sci 75:2877-2884.

Loh P.-R., P. Danecek, P.F. Palamara, C. Fuchsberger, Y. A Reshef, H. K Finucane, S. Schoenherr, L. Forer, S. McCarthy, G.R. Abecasis, R. Durbin, and A. L Price, 2016 Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet 48:1443-1448.

Lopes M.S., J.W.M. Bastiaansen, B. Harlizius, E.F. Knol, and H. Bovenhuis, 2014 A Genome-Wide Association Study Reveals Dominance Effects on Number of Teats in Pigs. PLoS One 9:e105867.

Lopes M.S., J.W.M. Bastiaansen, L. Janss, E.F. Knol, and H. Bovenhuis, 2016 Genomic prediction of growth in pigs based on a model including additive and dominance effects. J Anim Breed Genet 133:180-186.

Lopes M.S., H. Bovenhuis, A.M. Hidalgo, J.A.M. van Arendonk, E.F. Knol, and J.W.M. Bastiaansen, 2017 Genomic selection for crossbred performance accounting for breed-specific effects. Genet Sel Evol 49:51.

Lourenco D.A.L., S. Tsuruta, B.O. Fragomeni, C.Y. Chen, W.O. Herring, and I. Misztal, 2016 Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices. J Anim Sci 94:909-919.

Lukaszewicz M., R. Davis, J.K. Bertrand, I. Misztal, and S. Tsuruta, 2015 Correlations between purebred and crossbred body weight traits in Limousin and Limousin-Angus populations. J Anim Sci 93:1490-1493.

Lund M.S., A.P.W. de Roos, A.G. de Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbrandtsen, Z. Liu, R. Reents, C. Schrooten, F. Seefried, and G. Su, 2011 A common reference population from four European Holstein populations increases reliability of genomic predictions. Genet Sel Evol 43:43-43.

Lutaaya E., I. Misztal, J.W. Mabry, T. Short, H.H. Timm, and R. Holzbauer, 2001 Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model. J Anim Sci 79:3002-3007.

Lynch M., 1991 The genetic interpretation of inbreeding depression and outbreeding depression. Evolution 45:622-629.

Lynch M., and B. Walsh, 1998 *Genetics and analysis of quantitative traits*: Sinauer Sunderland, MA.

Mackay T.F.C., 2014 Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat Rev Genet 15:22.

Mackay T.F.C., 2015 Epistasis for Quantitative Traits in Drosophila, pp. 47-70 in *Epistasis: Methods and Protocols*, edited by J.H. Moore and S.M. Williams. Springer New York, New York, NY.

Maki-Tanila A., and W.G. Hill, 2014 Influence of gene interaction on complex trait variation with multilocus models. Genetics 198:355-367.

Maniatis G., N. Demiris, A. Kranis, G. Banos, and A. Kominakis, 2013 Genetic analysis of sexual dimorphism of body weight in broilers. J Appl Genet 54:61-70.

McNew R.W., and A.E. Bell, 1971 The nature of the purebred-crossbred genetic covariance. Genet Res 18:1-7.

Melchinger A.E., H.-P. Piepho, H.F. Utz, J. Muminovic, T. Wegenast, O. Törjék, T. Altmann, and B. Kusterer, 2007a Genetic basis of heterosis for growth-related traits in Arabidopsis investigated by testcross progenies of near-isogenic lines reveals a significant role of epistasis. Genetics 177:1827-1837.

Melchinger A.E., H.F. Utz, H.-P. Piepho, Z.-B. Zeng, and C.C. Schön, 2007b The Role of Epistasis in the Manifestation of Heterosis: A Systems-Oriented Approach. Genetics 177:1815-1825.

Meuwissen T.H.E., B.J. Hayes, and M.E. Goddard, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Minvielle B.F., J.-L. Monvoisin, J. Costa, A. Frenot, and Y. Maeda, 1999 Changes in heterosis under within-line selection or reciprocal recurrent selection: an experiment on early egg production in Japanese quail. J Anim Breed Genet 116:363-377.

Minvielle F., 1987 Dominance is not necessary for heterosis: a two-locus model. Genet Res 49:245-247.

Moghaddar N., A.A. Swan, and J.H.J. van der Werf, 2014 Comparing genomic prediction accuracy from purebred, crossbred and combined purebred and crossbred reference populations in sheep. Genet Sel Evol 46:58.

Moghaddar N., and J.H.J. van der Werf, 2017 Genomic estimation of additive and dominance effects and impact of accounting for dominance on accuracy of genomic evaluation in sheep populations. J Anim Breed Genet 00:1-10.

Monnahan P.J., and J.K. Kelly, 2015 Epistasis Is a Major Determinant of the Additive Genetic Variance in Mimulus guttatus. PLoS Genet 11:e1005201.

Morris J.A., and F.E. Binet, 1966 Heterosis and selection in the domestic fowl. Aust J Biol Sci 19:575-585.

Mrode R.A., 2005 *Linear models for the prediction of animal breeding values*. Wallingford: CABI Publishing.

Mulder H.A., W.G. Hill, A. Vereijken, and R.F. Veerkamp, 2009 Estimation of genetic variation in residual variance in female and male broiler chickens. Animal 3:1673-1680.

Mulder H.A., J. Visscher, and J. Fablet, 2016 Estimating the purebred–crossbred genetic correlation for uniformity of eggshell color in laying hens. Genet Sel Evol 48:39.

Newman S., A. Reverter, and D.J. Johnston, 2002 Purebred-crossbred performance and genetic evaluation of postweaning growth and carcass traits in Bos indicus x Bos taurus crosses in Australia. J Anim Sci 80:1801-1808.

Ott L., and M. Longnecker, 2010 *An introduction to statistical methods and data analysis*: Brooks/Cole, Cengage Learning.

Paixão T., and N.H. Barton, 2016 The effect of gene interactions on the long-term response to selection. PNAS 113:4422-4427.

Pengelly R.J., A.A. Gheyas, R. Kuo, E. Mossotto, E.G. Seaby, D.W. Burt, S. Ennis, and A. Collins, 2016 Commercial chicken breeds exhibit highly divergent patterns of linkage disequilibrium. Heredity (Edinb) 117:375-382.

Pettersson M., F. Besnier, P.B. Siegel, and O. Carlborg, 2011 Replication and explorations of high-order epistasis using a large advanced intercross line pedigree. PLoS Genet 7:e1002180.

Pirchner F., and R. Mergl, 1977 Overdominance as cause for heterosis in poultry. Zeitschrift für Tierzüchtung und Züchtungsbiologie 94:151-158.

Porto-Neto L., W. Barendse, J. Henshall, S. McWilliam, S. Lehnert, and A. Reverter, 2015 Genomic correlation: harnessing the benefit of combining two unrelated populations for genomic selection. Genet Sel Evol 47:84.

Raymond B., A.C. Bouwman, C. Schrooten, J. Houwing-Duistermaat, and R.F. Veerkamp, 2018a Utility of whole-genome sequence data for across-breed genomic prediction. Genet Sel Evol 50:27.

Raymond B., A.C. Bouwman, Y.C.J. Wientjes, C. Schrooten, J. Houwing-Duistermaat, and R.F. Veerkamp, 2018b Genomic prediction for numerically small breeds, using models with pre-selected and differentially weighted markers. Genet Sel Evol 50:49.

Rekaya R., R.L. Sapp, T. Wing, and S.E. Aggrey, 2013 Genetic evaluation for growth, body composition, feed efficiency, and leg soundness. Poult Sci 92:923-929.

Robertson A., 1959 The Sampling Variance of the Genetic Correlation Coefficient. Biometrics 15:469-485.

Sargolzaei M., J.P. Chesnais, and F.S. Schenkel, 2014 A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478.

Schenkel F.S., and M. Sargolzaei, 2009 QMSim: a large-scale genome simulator for livestock. Bioinformatics 25:680-681.

Schnell F.W., and C.C. Cockerham, 1992 Multiplicative vs. arbitrary gene action in heterosis. Genetics 131:461-469.

Sellier P., 1976 The basis of crossbreeding in pigs; A review. Livest Prod Sci 3:203-226.

Sevillano C.A., 2018 Genomic evaluation considering the mosaic genome of the crossbred pig. Wageningen University, Wageningen.

Sevillano C.A., J. Vandenplas, J.W. Bastiaansen, and M.P. Calus, 2016 Empirical determination of breed-of-origin of alleles in three-breed cross pigs. Genet Sel Evol 48:55.

Sevillano C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus, 2017 Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet Sel Evol 49:75.

Shen G., W. Zhan, H. Chen, and Y. Xing, 2014 Dominance and epistasis are the main contributors to heterosis for plant height in rice. Plant Science 215-216:11-18.

Sheridan A.K., and M.C. Randall, 1977 Heterosis for egg production in white leghorn Australorp crosses. British Poultry Science 18:69-77.

Shull G.H., 1908 The composition of a field of maize. J Hered:296-301.

Shull G.H., 1952 Beginnings of the heterosis concept, pp. 14-48 pp. in *Heterosis*, edited by J.W. Gowen. Iowa State College Press, Ames.

Smith C., 1964 The use of specialised sire and dam lines in selection for meat production. Animal Science 6:337-344.

Sonesson A.K., and T.H. Meuwissen, 2009 Testing strategies for genomic selection in aquaculture breeding programs. Genet Sel Evol 41:37.

Stuber C.W., 2010 Heterosis in Plant Breeding, pp. 227-251 in *Plant breeding reviews*. John Wiley & Sons, Inc.

Sun X., and R.H. Mumm, 2016 Method to represent the distribution of QTL additive and dominance effects associated with quantitative traits in computer simulation. BMC Bioinformatics 17:73-73.

Van Grevenhof I.E., and J.H. Van Der Werf, 2015 Design of reference populations for genomic selection in crossbreeding programs. Genet Sel Evol 47:14.

Vandenplas J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen, 2016 Assigning breed origin to alleles in crossbred animals. Genet Sel Evol 48:61.

VanRaden P.M., 2008 Efficient methods to compute genomic predictions. J Dairy Sci 91:4414-4423.

Veroneze R., J.W. Bastiaansen, E.F. Knol, S.E. Guimaraes, F.F. Silva, B. Harlizius, M.S. Lopes, and P.S. Lopes, 2014 Linkage disequilibrium patterns and persistence of phase in purebred and crossbred pig (Sus scrofa) populations. BMC Genet 15:126.

Visscher P., R. Pong-Wong, C. Whittemore, and C. Haley, 2000 Impact of biotechnology on (cross)breeding programmes in pigs. Livest Prod Sci 65:57-70.

Visscher P.M., G. Hemani, A.A.E. Vinkhuyzen, G.-B. Chen, S.H. Lee, N.R. Wray, M.E. Goddard, and J. Yang, 2014 Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. PLoS Genet 10:e1004269.

Vitezica Z.G., A. Legarra, M.A. Toro, and L. Varona, 2017 Orthogonal Estimates of Variances for Additive, Dominance and Epistatic Effects in Populations. Genetics.

Vitezica Z.G., A. Reverter, W. Herring, and A. Legarra, 2018 Dominance and epistatic genetic variances for litter size in pigs using genomic models. Genet Sel Evol 50:71.

Vitezica Z.G., L. Varona, and A. Legarra, 2013 On the additive and dominant variance and covariance of individuals within the genomic selection scope. Genetics 195:1223-1230.

Wade M.J., 2002 A gene's eye view of epistasis, selection and speciation. J Evol Biol 15:337-346.

Walsh B., and M. Lynch, 2018 Maintenance of Quantitative Genetic Variation in *Evolution and Selection of Quantitative Traits*. Oxford University Press, Oxford.

Wang J., A. Caballero, and W.G. Hill, 1998 The effect of linkage disequilibrium and deviation from Hardy-Weinberg proportions on the changes in genetic variance with bottlenecking. Heredity 81:174-186.

Wei M., H.A.M. van der Steen, J.H.J. van der Werf, and E.W. Brascamp, 1991 Relationship between purebred and crossbred parameters. J Anim Breed Genet 108:253-261.

Wei M., and J.H. van der Werf, 1995 Genetic correlation and heritabilities for purebred and crossbred performance in poultry egg production traits. J Anim Sci 73:2220-2226.

Wei M., and J.H.J. van der Werf, 1994 Maximizing genetic response in crossbreds using both purebred and crossbred information. Anim Sci 59:401-413.

Wei W.-H., G. Hemani, and C.S. Haley, 2014 Detecting epistasis in human complex traits. Nat Rev Genet 15:722-733.

Weir B.S., 2008 Linkage disequilibrium and association mapping. Annu Rev Genomics Hum Genet 9:129-142.

Wellmann R., and J. Bennewitz, 2011 The contribution of dominance to the understanding of quantitative genetic variation. Genet Res (Camb) 93:139-154.

Wellmann R., and J. Bennewitz, 2012 Bayesian models with dominance effects for genomic evaluation of quantitative traits. Genet Res (Camb) 94:21-37.

Whittaker J.C., R. Thompson, and M.C. Denham, 2000 Marker-assisted selection using ridge regression. Genet Res (Camb) 75:249-252.

Wientjes Y.C.J., P. Bijma, J. Vandenplas, and M.P.L. Calus, 2017 Multi-population Genomic Relationships for Estimating Current Genetic Variances Within and Genetic Correlations Between Populations. Genetics 207:503-515.

Wientjes Y.C.J., and M.P.L. Calus, 2017 BOARD INVITED REVIEW: The purebred-crossbred correlation in pigs: A review of theory, estimates, and implications. J Anim Sci 95:3467-3478.

Wientjes Y.C.J., M.P.L. Calus, P. Duenk, and P. Bijma, 2018 Required properties for markers used to calculate unbiased estimates of the genetic correlation between populations. Genet Sel Evol 50:65.

Wientjes Y.C.J., R.F. Veerkamp, P. Bijma, H. Bovenhuis, C. Schrooten, and M.P.L. Calus, 2015 Empirical and deterministic accuracies of across-population genomic prediction. Genet Sel Evol 47:5.

Willham R.L., and E. Pollak, 1985 Theory of Heterosis. J Dairy Sci 68:2411-2417.

Wolc A., J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, R. Preisinger, D. Habier, R. Fernando, D.J. Garrick, and J.C. Dekkers, 2011 Persistence of accuracy of genomic estimated breeding values over generations in layer chickens. Genet Sel Evol 43:23.

Wolc A., A. Kranis, J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, A. Avendano, K.A. Watson, J.M. Hickey, G. de los Campos, R.L. Fernando, D.J. Garrick, and J.C.M. Dekkers, 2016 Implementation of genomic selection in the poultry industry. Animal Frontiers 6:23-31.

Wright S., 1931 Evolution in Mendelian Populations. Genetics 16:97-159.

Wright S., 1969 *The Theory of Gene Frequencies*. Chicago: The University of Chicago Press.

Xiang T., O.F. Christensen, and A. Legarra, 2017 Technical note: Genomic evaluation for crossbred performance in a single-step approach with metafounders. J Anim Sci 95:1472-1480.

Xiang T., B. Nielsen, G. Su, A. Legarra, and O.F. Christensen, 2016 Application of single-step genomic evaluation for crossbred performance in pig. J Anim Sci 94:936-948.

Xiao J., J. Li, L. Yuan, and S.D. Tanksley, 1995 Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. Genetics 140:745-754.

Yang J., B. Benyamin, B.P. McEvoy, S. Gordon, A.K. Henders, D.R. Nyholt, P.A. Madden, A.C. Heath, N.G. Martin, G.W. Montgomery, M.E. Goddard, and P.M. Visscher, 2010 Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42:565-569.

Yang N., E. Dunnington, and P. Siegel, 1999 Heterosis following long-term bidirectional selection for mating frequency in male Japanese quail. Poult Sci 78:1252-1256.

Zan Y., S.K.G. Forsberg, and Ö. Carlborg, 2018 On the Relationship Between High-Order Linkage Disequilibrium and Epistasis. G3: Genes|Genomes|Genetics.

Zeng J., A. Toosi, R.L. Fernando, J.C. Dekkers, and D.J. Garrick, 2013 Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. Genet Sel Evol 45:11.

Zhu Z., A. Bakshi, A.E. Vinkhuyzen, G. Hemani, S.S.H. Lee, I.M. Nolte, J.V. van Vliet-Ostaptchouk, H. Snieder, T. Esko, L. Milani, R. Mägi, A. Metspalu, W.G. Hill, Bruce S. Weir, M.E. Goddard, P.M. Visscher, and J. Yang, 2015 Dominance Genetic Variation Contributes Little to the Missing Heritability for Human Complex Traits. Am J Hum Genet 96:377-385.

Zumbach B., I. Misztal, S. Tsuruta, J. Holl, W. Herring, and T. Long, 2007 Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. J Anim Sci 85:901-908.

Summary

In pig and poultry breeding programs, animals from genetically distinct purebred lines are mated to produce crossbred animals, which are the final production animals. The aim of such breeding programs is to improve the performance of crossbred animals every generation, by selecting genetically superior individuals in the purebred lines to become parents of the next generation. The purebred lines usually differ in the traits that they are selected for. For instance, dam lines may be selected for female fertility and health traits, whereas sire lines may be selected for growth, egg production traits, and feed efficiency. The breeders then benefit from breed complementarity in the crossbreds, meaning that the crossbred animals combine the desirable traits that the purebred lines were selected for. Another benefit of crossbreeding is heterosis, which refers to the better average performance of crossbreds compared to the average performance of their parental lines.

Although the aim of crossbred breeding programs is to improve the performance of crossbreds, selection takes place in the purebred lines. When these selection decisions are based on performance of the purebreds, improvement in crossbred performance follows from a correlated response. This strategy may be suboptimal because the performance of crossbred animals can be genetically different from the performance of purebred animals. This difference is quantified by the genetic correlation between purebred and crossbred performance (r_{pc}), which is usually lower than one. The r_{pc} is an important parameter for crossbreeding programs, because it partly determines the response in crossbred performance when selection is based on purebred performance.

One of the reasons for r_{pc} values lower than one are genotype by genotype (GxG) interactions. Such interactions (also known as non-additive effects) can occur between genes at the same location of the genome (termed dominance), or between genes at different locations (termed epistasis). Although dominance and epistatic effects can be strong and abundant, most of these effects translate into the 'average effects' of genes. As a result of non-additive effects, these gene effects can differ between two purebred populations, or between performance in purebreds and performance in crossbreds. Taken together, the differences in gene effects between populations can be summarized into a single number, known as the genetic correlation. In chapter 3, we investigated how the genetic correlation between two populations (r_g) is affected by non-additive effects and differences in allele frequencies between populations. We simulated two populations that had diverged under drift only, or under drift and selection, and we simulated traits where the

genetic model and magnitude of non-additive effects varied. Results showed that larger differences in allele frequencies and larger non-additive effects resulted in lower values of r_g between populations. In addition, we found that, when the non-additive effects became extremely large, r_g had a lower limit that was determined by the type of interactions between alleles, and the difference in allele frequencies between populations. With realistic dominance, r_g values did not drop below 0.8, whereas with realistic epistasis, r_g values dropped to as low as 0.45. These insights contribute to the understanding of differences in genetic expression of complex traits between populations.

In chapter 4, we focused on the relationship between non-additive effects and the genetic correlation between purebred and crossbred performance (r_{pc}). We hypothesized that r_{pc} could be predicted when we have information on r_g in the parental lines. The aim of this chapter was therefore to derive expressions for r_{pc} in a two-way crossbred breeding program, based on genetic parameters in parental lines. We derived these expressions for a genetic model with additive and dominance effects (D), and a model with additive, and additive by additive epistatic effects (E_{AA}). We validated our expressions with simulations, and showed that our expressions provide exact predictions of r_{pc} for models D and E_{AA}, and that they provide upper and lower bounds for r_{pc} in scenarios with other types of non-additive effects. Breeders may be able to use the expressions derived in this chapter to predict the impact of non-additive effects on r_{pc} for specific crosses, without having to collect data on crossbreds.

When r_{pc} is lower than one, it may be beneficial to make selection decisions based on information on crossbred performance instead of purebred performance. This is, however, a challenging task, because purebred animals cannot be tested directly for performance at the crossbred level. Now, with the recent developments in genomic prediction, it has become possible to estimate breeding values for crossbred performance for purebred animals. Genomic prediction makes use of a reference population that consists of individuals that have both phenotypic records and genotypic marker records. This information is used to estimate effects for all genotype markers simultaneously. These estimated marker effects can then be used to compute genomic estimated breeding values (GEBV) for crossbred performance of purebred animals that have only genotypic marker records available.

In chapter 2, we investigated the accuracy of estimated gene effects, when data is collected on a finite sample from a larger population, and when dominance is present. We compared a model that explicitly models dominance with a model

that does not. Our results showed that the dominance model estimates gene effects more accurately. We showed that errors in the estimation of gene effects arose because of differences between the sample and the whole population in terms of genetic composition. A model that includes dominance was more robust against these deviations. Furthermore, in the absence of dominance, there was no penalty of modelling dominance. These results may suggest that modelling dominance is beneficial for estimating gene effects across the whole genome, and possibly for genomic prediction.

The objective of chapter 5 was to compare estimates of r_{pc} obtained from pedigree-based and genotype-based models. In addition, we compared estimates from genotype-based models that either consider or ignore the breed-of-origin of alleles (BOA). For this purpose, we analysed body weight in broiler chicken, using genomic and phenotypic data collected from purebred and crossbred animals that had a common sire line. Our results showed that estimated r_{pc} was 5 to 25% larger with genotype-based models than with pedigree-based models, and that standard errors of r_{pc} were smaller with genotype-based models than with pedigree-based models. Considering the BOA did not result in different r_{pc} estimates compared to ignoring the BOA, probably because the parental lines of the crossbred animals were distantly related. We concluded that genotype-based models can be useful for estimating r_{pc}, even when the purebred and crossbred animals that have phenotypes are closely related.

In chapter 6, we used the same data as in chapter 5 to investigate the benefit of using a crossbred instead of purebred reference population for the estimation of GEBV for crossbred performance. In addition, when using a crossbred reference population, we investigated the benefit of considering the BOA. These benefits were determined with two alternative validation strategies: one where validation was based on crossbred offspring averages, and one where validation was based on individual crossbred records. Our results showed that there were large differences between validation strategies in the observed benefits of crossbred data and of considering the BOA. We argued that when GEBV are validated with individual crossbred records, the accuracy of models that use a crossbred reference population and ignore the BOA are inflated. Thus, we recommended that, whenever possible, GEBV for crossbred performance should be validated with crossbred offspring averages. With this validation strategy, our results showed that the accuracy of GEBV was higher with a crossbred reference population than with a purebred reference population for a trait with an r_{pc} of 0.8, but lower for a trait with an r_{pc} of 0.96.

Similarly, taking the BOA into account was beneficial for a trait with an r_{pc} of 0.8, but not for a trait with an r_{pc} of 0.96.

In chapter 7, I first discussed the relationship between non-additive effects and heterosis (i.e. the increased performance of crossbreds compared to the average performance of parental lines). I showed that heterosis is closely related to r_{pc} through the existence of non-additive effects. I show that the response to purebred selection in crossbred performance can be larger than the response in purebred performance, even when r_{pc} is lower than one. My results may partly explain how heterosis can persists across many generations of selection on purebred performance. Second, I discussed strategies to estimate GEBV for crossbred performance. I hypothesized that using a crossbred reference population and considering the BOA would optimize the accuracy of genomic prediction models. I concluded, however, that even though this model agrees best with genetic reality, it does not necessarily lead to the highest prediction accuracy.

Curriculum Vitae

About the author

Publications

Contributions to conferences

Training and eduation

About the author

Pascal Duenk was born on the 7th of November 1989 in Aalten, the Netherlands. From a young age, he was fascinated by animals, and he had a special interest for birds of prey. After studying Biology at Utrecht University for a year, Pascal enrolled in the Animal Science bachelor program at Wageningen University in 2009. In the year that he obtained his BSc degree (2013), he started his master at the same university. He majored in Animal Behaviour, and wrote an MSc thesis on the relationship between personality traits and trainability in young assistance dogs. His minor thesis was at the Animal Breeding and Genomics group, where he worked on estimating the heritability of resilience in dairy cattle using item response theory. This thesis kickstarted his interest in statistics and computer programming, and it was evaluated very positively. Pascal graduated in 2014, and decided to pursue a career in quantitative genetics. His minor thesis supervisor offered him a temporary position at the research company Genetwister in Wageningen, where he further developed his skills and knowledge on genetics and genetic analyses. He stayed at Genetwister for almost a year, after which he obtained a PhD position at Wageningen University to work on genomic prediction for crossbred performance in 2015. The results of that project are presented in this thesis. As part of his PhD, Pascal visited the University of New England in Armidale (Australia) for 6 months to work on the impact of non-additive effects on the genetic correlation between populations. Since November 2019, Pascal works as a postdoctoral researcher at Wageningen University, and focuses on the optimization of crossbred breeding programs.

Publications

Duenk P., P. Bijma, Y.C.J. Wientjes, and M.P.L. Calus, *Predicting the purebred-crossbred genetic correlation from genetic variances within, and covariance between parental lines*. In preparation.

van Bergen G.H.H., **P. Duenk**, C.A. Albers, P. Bijma., M.P.L. Calus, Y.C.J. Wientjes, and H.J. Kappen, 2020 *Bayesian neural networks with variable selection for prediction of genotypic values*. In review at Genet Sel Evol.

Kamphuis C., **P. Duenk**, B. Visser, G. Singh, A. Nigsch, R.M. de Mol, M. L. W. J. Broekhuijse, 2020 *Machine learning to further improve prediction of boar semen fertility*. Accepted in Theriogenology.

Duenk P., P. Bijma, M.P.L. Calus, Y.C.J. Wientjes, and J.H.J. van der Werf, 2019 *The Impact of Non-additive Effects On the Genetic Correlation Between Populations*. Accepted in G3: Genes|Genomes|Genetics

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, V.P. Breen, J.M. Henshall, R. Hawken, and P. Bijma, 2019 *Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles*. Genet Sel Evol 51:38.

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, V.P. Breen, J.M. Henshall, R. Hawken, and P. Bijma, 2019 *Estimating the purebred-crossbred genetic correlation of body weight in broiler chickens with pedigree or genomic relationships*. Genet Sel Evol 51:6.

Wientjes Y.C.J., M.P.L. Calus, **P. Duenk**, and P. Bijma, 2018 *Required properties for markers used to calculate unbiased estimates of the genetic correlation between populations*. Genet Sel Evol 50:65.

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, and P. Bijma, 2017 *Benefits of Dominance over Additive Models for the Estimation of Average Effects in the Presence of Dominance*. G3: Genes|Genomes|Genetics 7:3405-3414.

Contributions to conferences

Duenk P., P. Bijma, M.P.L. Calus, Y.C.J. Wientjes, and J.H.J. van der Werf, 2019 *The impact of non-additive effects on the genetic correlation between populations*. European Association of Animal Production (EAAP). Ghent, Belgium.

Calus M. P. L., C.A. Sevillano, Y.C.J. Wientjes, **P. Duenk**, L.M.G. Verschuren, D. Schokker, P. Bijma, J. Ten Napel, R.F. Veerkamp, and J. Vandenplas, 2019 *Genomic prediction in animal breeding – focus on crossbred performance.* Phenotypic Production Workshop. University of Florida, Gainesville (FL), United States.

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, V.P. Breen, J.M. Henshall, R. Hawken, and P. Bijma, 2018 *Accuracy of genomic estimated breeding values for crossbred performance in broilers using a purebred or crossbred reference population.* World Congress on Genetics Applied to Livestock Production (WCGALP). Auckland, New Zealand.

Wientjes Y.C.J., M.P.L. Calus, J. Vandenplas, **P. Duenk**, and P. Bijma, 2018 *A multi-population genomic relationship matrix to estimate the genetic correlation between populations.* World Congress on Genetics Applied to Livestock Production (WCGALP). Auckland, New Zealand.

Kamphuis C., B. Visser, **P. Duenk**, G. Singh, A. Nigsch, R. de Mol, R.F. Veerkamp, M.L.W.J. Broekhuijse 2018 *Hacking CASA to predict boar semen fertility.* European Association of Animal Production (EAAP). Dubrovnik, Croatia.

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, and P. Bijma, 2017 *Bias and accuracy of estimated average effects due to dominance, for two statistical models*. WIAS Science Day. Wageningen University, Wageningen, the Netherlands.

Wientjes Y.C.J., P. Bijma, **P. Duenk**, and M.P.L. Calus, 2017 *Estimating the genetic correlation between populations.* Gordon Research Conference. Galveston (Texas), United States.

Duenk P., M.P.L. Calus, Y.C.J. Wientjes, and P. Bijma, 2017 *Benefits of Dominance over Additive Models for the Estimation of Average Effects in the Presence of Dominance*. European Association of Animal Production (EAAP). Tallinn, Estonia.

Duenk P. M.P.L. Calus, Y.C.J. Wientjes, and P. Bijma, 2016 *Bias and MSE of estimated allele substitution effects due to dominance, for two statistical models.* International Conference of Quantitative Genetics (ICQG). Madison (Wisconsin), United States.

Wientjes Y.C.J., P. Bijma, **P. Duenk**, and M.P.L. Calus, 2016 *Factors affecting the estimation of the genetic correlation between populations*. International Conference of Quantitative Genetics (ICQG). Madison (Wisconsin), United States.

Training and education

The Basic Package (1.8 credits)	
WIAS Introduction Day	2016
Course on philosophy of science and/or ethics	2017

Disciplinary Competences (18.5 credits)	
Statistical models for genomic prediction in animals and plants (Arhus, Denmark)	2015
Quantitative Genetics Discussion Group (Wageningen)	2015-2019
Fortran Discussion Group (Wageningen)	2016-2017
Writing of a literature review (Wageningen)	2016
ASReml course (Wageningen)	2016
Genomic prediction in the era of genome sequencing (Madison (Wisconsin), USA)	2016
Short course on evolutionary quantitative genetics (Edinburgh, Scotland)	2016
Design of breeding programs with Genomic Selection (Wageningen)	2017
Introduction to Graphical Models with applications to Quantitative Genetics (Armidale, Australia)	2019

Professional Competences (5.5 credits)	
Techniques for Writing and Presenting a Scientific Paper	2016
Brain Training	2016
WGS PhD Workshop Carousel	2016
Teaching and supervising Thesis students	2017
Scientific Writing	2018
Workshop creative research by Bas Haring	2018
Presenting with Impact 4	2019

Presentation Skills (6 credits)	
ICQG5 (Poster)	2016
WIAS Science day (Oral)	2017
EAAP 2017 (Oral)	2017
Fokkerij & Genetica Dagen 2017 (Oral)	2017
11th WCGALP (Oral)	2018
EAAP 2019 (Oral)	2019

Teaching competences (8.6 credits)	
Supervision of MSc thesis students (2 times)	2015-2016
Assisting Animal Breeding and Genetics BSc course	2016
Assisting Genetics course at Veterinary School Utrecht	2017
Mini Lecture (45 min) High school Laren	2017
Mini Lecture (45 min) Open day High school students	2017
Assisting Animal Breeding and Genetics BSc course	2017
Review 2 Proposals for Research Master Cluster	2019

Total credits	40.4

Acknowledgements

Throughout the four years that it took to complete this thesis, there have been many people that helped me along the way, in one way or another. You are with too many to name, so I want to thank every single one of you for your contribution, no matter how small. I would like to thank a couple of individuals in particular that have played an important part in my development as a PhD student. I will start with those that have been closely involved in my daily work, and in finalizing this thesis.

Piter, my daily supervisor, thank you for all our insightful discussions and inspiring meetings. I especially admire your writing skills: it's incredible how you can improve a text with a few (seemingly) simple changes. You were a great mentor and teacher, and I hope that after this project ends, we can continue working together from time to time. Mario, my supervisor turned promotor, thank you for everything you have taught me. I am amazed by your work efficiency and presentation skills, and I admire the way you supervise your students. I am very proud to be the first PhD student with you as promotor. Thanks for all the good conversations and story-sharing while having a few beers, hopefully there are many more to come! Yvonne, in these four years you not only became my (official) supervisor, but also a very good friend. Thank you for everything you have taught me, and for all the fun times outside the office. You have a very good eye for details, which has helped me to think critically about my work. I am proud to have you there with me on stage during my defence.

The cover of this thesis was designed by my sister. Mariska, bedankt dat je de moeite hebt genomen om voor mij dit design te maken. Je hebt er enorm veel werk in gestoken, en ik ben super trots op jou en op het resultaat!

The greatest thing about being a PhD student is that you get to share so many great experiences with your fellow students and colleagues. There are many people that I have crossed paths with during these four years. Thanks to everyone with whom I have shared a coffee, a discussion, a paper, an office, a foosball game, a conference trip, a holiday, a dinner, a dance, a beer (or two), a drinking game, and so much more! You have made my time as a PhD student so much fun, and I am forever grateful to have had such an amazing group of people around me.

There are two people that I would like to thank for fulfilling their roles as paranymphs. I am proud to have them at my side during my defence, and there is no doubt in my mind that these guys will make the party unforgettable. Renzo, in a short time you have become a very close friend. I admire your friendly and open personality, your honesty, and your passion for Italian cuisine. Thank you for all the fun we had together, and I am sure there is much more to come! María, I am super

proud to have you as my friend and paranymph. I have lost count of the number of dinners, parties, shots, and dances we have shared. You're one of the most caring people I know, and I admire your energy and enthusiasm. Thanks so much for your support throughout all these years.

During my PhD, I had the privilege to work abroad for six months in Armidale, Australia. I would like to thank all the people that have made that experience unforgettable, and I want to mention a few in particular. Julius, thank you for inviting me and for being a great host. During my stay, you and Sam (thanks!) have helped me (re)discover the fun in doing science. I especially enjoyed our discussions during long walk/coffee breaks. Jim and Jackie, thanks for letting us stay in your 'villa', you made us feel at home from day one!

Besides thanking my colleagues, I would like to thank some people who have not been involved with my work directly, but they have been important for getting where I am now. Pa en ma, bedankt voor jullie steun in de keuzes die ik maakte voor mijn studie en loopbaan. Jullie hebben me altijd gezegd dat ik moet doen wat ik leuk vind, de rest komt dan vanzelf. Bedankt dat de deur altijd openstond als we een logeeradres nodig hadden in Aalten. Tot slot, allerliefste Anne, ik ben je ontzettend dankbaar dat je voor mij naar Wageningen (en Bennekom) wilde verhuizen. Bij jou voel ik me thuis, en kan ik de werkdag even van me af zetten. Het is heerlijk om 's avonds naast je op de bank te ploffen en mijn gedachten met je te delen. Dankjewel voor het aanhoren van al mijn zorgen, problemen, en gedachtekronkels, zeker tijdens de laatste fase van dit project. Ik had het niet kunnen doen zonder jou, en ik hoop nog veel van dit soort mijlpalen met je te delen!

Colophon

This research was financially supported by the Netherlands Organization for Scientific Research (NWO) and the Breed4Food consortium partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin, through the GenoMiX project. The research visit of Pascal Duenk to the University of New England, Armidale (Australia) was funded by the Wageningen Institute for Animal Sciences (WIAS).

The cover of this thesis was designed by Mariska Duenk.

Printed by: DigiForce | Proefschriftmaken.nl