

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Estimation of additive and dominance genetic parameters and
identification of dominance markers for production and fertility
traits in New Zealand dairy cattle.**

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

in
Animal Science

at Massey University, Palmerston North, New Zealand



Dorothy Florence Schmidlin Wilson

2026

Abstract

Traditional genomic evaluation models in dairy cattle only account for additive genetic effects and generally exclude dominance effects. However, dominance effects may contribute to phenotypic variation in economically important traits. The aim of this research was to estimate additive and dominance variance components for milk production traits, milk composition traits, and fertility traits and identify significant dominance marker effects through genome-wide association studies (GWAS). This study used 5,686 lactation records from 1,667 spring-calved cows across two Massey University research herds, and genotype data from 49,171 single nucleotide polymorphisms (SNPs). Genomic restricted maximum likelihood was used to estimate variance components under additive-only and additive + dominance models, with pedigree-based estimates included for comparison. Genome-wide association studies were performed for yields of milk, fat, protein, and lactose, percentages of fat, protein, and lactose, and fertility traits including days from the start of mating to first service (SMFS), pregnancy rate at first service (PRFS), 3-week submission rate (SR21), and 6-week pregnancy rate (PR42) where additive and dominance SNP effects were fitted simultaneously as covariates. Milk yield traits, including milk volume, fat, and protein, had moderate heritability estimates ranging between 0.22–0.28. Milk composition traits, including fat percentage, protein percentage, and lactose percentage, had high heritability estimates ranging between 0.45–0.69, and fertility traits SMFS, PRFS, SR21, and PR42 had low heritability estimates of less than 0.021. Dominance variance contributed between 8–12% of phenotypic variance for milk yield traits, but was negligible for milk composition and fertility traits. These findings suggest that although additive variance accounted for a greater proportion of phenotypic variance, dominance may contribute to variation in economically important yield traits. Genome-wide association analyses identified significant additive associations on chromosome 14 for milk yield traits, and fat and protein percentage, as well as additive association signals on chromosome 5 for milk yield, lactose yield, and fat percentage. Dominance association signals were generally weaker, with a region on chromosome 9 identified for fat yield. Few significant additive and dominance markers were identified for fertility traits. The detection of dominance associations without significant additive associations highlights genomic regions that may not be detected using additive-only models. Although this study was limited by sample size, the results highlight the potential relevance of accounting for dominance effects in future genomic analyses of dairy cattle traits.

Acknowledgements

Firstly, I would like to express my gratitude to my primary supervisor, Professor Nicolas Lopez-Villalobos and my co-supervisor, Dr Swati Jivanji, for their guidance and support throughout this project. Thank you, Nicolas, for sharing your extensive knowledge and enthusiasm for research, which helped make this thesis possible. I am grateful for your continual encouragement throughout my studies and for going above and beyond to support the next stage of my career. Thank you, Swati, for your valuable feedback and discussions which helped shape this thesis.

I would like to acknowledge and express my gratitude for the financial support that I received from the LIC Patrick Shannon Scholarship, C Alma Baker Postgraduate Scholarship, Sir Alan Stewart Postgraduate Scholarship, Catherine Baxter Dairy Scholarship, Helen E Akers Scholarship, Waitawhiti Memorial Scholarship, and Peter Densem Postgraduate Scholarship that helped me to complete this research.

I would also like to thank my fellow Master's students, Jess, Leo, Leith, and Emma, for all the time spent together in and out of the Master's office, the shared discussions about our projects, and making the most of the free food on campus. You made the experience much more enjoyable and I'm grateful I got to share it with you all.

Thank you to Quinn, for your endless patience and support, and for the many hours you spent listening to me talk about my thesis and trying to understand what my research was about. Finally, to my parents, thank you for believing in me and encouraging me to pursue something I am passionate about. I am grateful for your support and everything you have done to help me get here.

Table of contents

Abstract.....	ii
Acknowledgements	v
List of Figures.....	x
List of Tables	xii
List of Abbreviations	xiii
Chapter 1	1
Chapter 2	6
2.1 The New Zealand dairy production system	8
2.2 Breeds and crossbreeding	8
2.3 National genomic evaluation system	10
2.4 Considering dominance genetic effects in the genomic evaluation of dairy cattle	13
2.5 Modelling SNP-level additive and dominance effects in genome-wide association studies	19
2.6 Potential benefits of including dominance in genomic evaluation	21
2.7 Gaps in knowledge.....	23
Chapter 3	26
3.1 Data.....	28
3.2 Phenotypic traits.....	29
3.3. Descriptive statistics and analyses of variance	29
3.4 Pedigree-based variance component estimation	30
3.5 Genotypes and quality control	30
3.6 Genomic variance component estimation	32
3.7 Genome-wide association studies	33
3.8 Estimation of genotype effects for significant SNPs	35
Chapter 4	37
4.1 Descriptive statistics	39
4.2 Effect of fixed effects.....	39
4.3 Variance components and heritability estimates.....	44
4.4 Genome-wide association studies	48
4.5 Genotype effects of top SNPs	57
Chapter 5	61
5.1 Descriptive statistics and fixed effects.....	63
5.2 Genetic variance estimates.....	64

5.3 Genome-wide association studies	67
5.4 Implications.....	71
5.5 Study limitations	72
5.6 Future research.....	73
5.7 Conclusions.....	73
References.....	76

List of Figures

Figure 2.1	Breed composition of the New Zealand dairy cattle population.	10
Figure 4.1	Manhattan plots based on genome-wide association analyses results testing additive and dominance effects for milk yield (MY), fat yield (FY), protein yield (PY) and lactose yield (LY).	49
Figure 4.2	Manhattan plots based on genome-wide association analyses results testing additive and dominance effects markers for fat percentage (FP), protein percentage (PP) and lactose percentage (LP).	52
Figure 4.3	Manhattan plots based on genome-wide association analyses results testing additive and dominance effects markers for the interval between start of mating to first service (SMFS), pregnancy rate to first service in the first 3 weeks of the breeding season (PRFS), 21-day submission rate (SR21), and 42-day in-calf rate (PR42).	54
Figure 4.4	Adjusted mean 305-day milk yield (MY;L), fat yield (FY;kg), protein yield (PY;kg), and lactose yield (LY;kg), with corresponding standard errors, for genotypes (AA, AB, BB) at the most significant dominance-associated SNP for each trait.	57
Figure 4.5	Mean fat, protein, and lactose percentages and corresponding standard errors for genotypes (AA, AB, BB) at the most significant dominance-associated SNP for each trait.	58
Figure 4.6	Mean interval from the start of mating date to the first service (SMFS), pregnancy rate to first service in the first 3 weeks of the breeding season (PRFS), 21-day submission rate (SR21, and 42-day in-calf rate (PR42), and corresponding standard errors for genotypes (AA, AB, BB) at the most significant dominance-associated SNP for each trait.	59

List of Tables

Table 2.1	Estimates of additive and dominance genetic variance using pedigree and genomic information for milk traits in dairy cattle.	16
Table 2.2	Estimates of additive and dominance genetic variance using pedigree and genomic information for fertility traits in dairy cattle.	18
Table 3.1	Number of animals within each herd and breed group.	28
Table 3.2	Number of single nucleotide polymorphisms (SNP) per chromosome after filtering.	31
Table 4.1	Descriptive statistics for milk production, milk composition, and fertility traits of cows from Massey University Dairy 1 and Dairy 4 herds.	39
Table 4.2	Least-squares means and standard errors for milk production and fertility traits by breed within the Massey University Dairy 1 and Dairy 4 herds.	40
Table 4.3	Least-squares means and standard errors for milk yield and composition traits by lactation number within each breed in the Massey University Dairy 1 and Dairy 4 herds.	42
Table 4.4	Least-squares means and standard errors for fertility traits by lactation number in the Massey University Dairy 1 and Dairy 4 herds.	43
Table 4.5	Variance components for milk production and fertility traits estimated with ASReml using pedigree data.	45
Table 4.6	Variance components for milk production and fertility traits estimated using an additive genetic model fitted with GREML in Julia, based on a genomic relationship matrix.	46
Table 4.7	Variance components for milk production and fertility traits estimated using an additive + dominance genetic model fitted with GREML in Julia, based on a genomic relationship matrix.	47
Table 4.8	Most significant single nucleotide polymorphisms for dominance effects identified using a model incorporating additive and dominance effects across milk yield, composition, and fertility traits.	56

List of Abbreviations

BLUP – Best linear unbiased prediction

BW – Breeding worth

DRM – Dominance relationship matrix

BV – Breeding value

FP – Fat percentage

FY – Fat yield

GBLUP – Genomic best linear unbiased prediction

GREML – Genomic restricted maximum likelihood

GRM – Genomic relationship matrix

GWAS – Genome-wide association study

LP – Lactose percentage

LY – Lactose yield

MY – Milk yield

PP – Protein percentage

PR42 – Pregnancy rate at 42 days

PRFS – Pregnancy rate after first service

PY – Protein yield

QTL – Quantitative trait locus

SMFS – Start of mating to first service

SNP – Single nucleotide polymorphism

SR21 – Submission rate at 21 days

Chapter 1

Introduction

The New Zealand dairy industry is a major contributor to the national economy, generating NZ\$27 billion in export revenue in the year to June 2025 (DairyNZ 2025a). The industry operates predominantly under a low-input, pasture-based production system, where calving is aligned with spring pasture growth to maximise feed efficiency. This requires dairy cows to maintain high levels of productivity and reproductive performance under variable environmental conditions. As a result, ongoing genetic improvement is central to maintaining productivity within New Zealand dairy systems.

Genomic selection has contributed substantially to improvement in dairy productivity over recent decades (DairyNZ 2025a). The implementation of the national selection index, Breeding Worth (BW), has allowed for the identification of animals with favourable genetic merit to produce progeny that more efficiently convert feed into profit (DairyNZ 2024a). The development of genomic selection, which predicts genetic merit using genome-wide genotypic information, has further improved accuracy of animal evaluation and can be used to accelerate the rate of genetic gain (Spelman et al. 2010; Schefers and Weigel 2012).

Current dairy cattle evaluations primarily focus on additive genetic effects to predict breeding values (BVs), where alleles contribute independently to the genetic value of an individual. However, non-additive genetic effects, which result from interactions between alleles, may also contribute to phenotypic variation. One important form of non-additive variation is dominance, also referred to as dominance deviation, which occurs when interactions between alleles at the same locus cause the phenotype of a heterozygous genotype to deviate from the average phenotype of the two homozygous genotypes (Falconer 1996). Since dominance effects may contribute to variation in complex traits, incorporating dominance into genomic evaluation models may improve the prediction of total genetic merit of dairy cattle (Aliloo et al. 2016).

Crossbreeding between Holstein-Friesian and Jersey cattle is widely used within the New Zealand dairy industry, with crossbred cows representing 60% of the national herd (DairyNZ and Livestock Improvement Corporation 2024). Crossbreeding utilises breed complementarity and heterosis to enhance animal performance (Sørensen et al. 2008). Heterosis occurs when a crossbred animal outperforms the average performance of its parental breeds and is partly influenced by non-additive genetic effects such as dominance (Sørensen et al. 2008). Therefore, understanding how dominance effects contribute to economically important traits such as milk yield, milk compositions, and fertility-related traits may provide opportunities to improve crossbreeding strategies within New Zealand dairy systems.

Several international studies have reported that dominance variance contributes to phenotypic variation in economically important dairy production traits, although estimates vary across studies (Sun et al. 2014; Aliloo et al. 2016; Schneider et al. 2023). Previous research has also indicated that including dominance effects in genomic evaluation models can improve prediction accuracy, influence the ranking of animals, and support mating allocation strategies to maximise offspring performance (Sun et al. 2013; Ertl et al. 2014; Aliloo et al. 2016). However, research investigating dominance effects in New Zealand populations remains limited.

Given the use of crossbreeding within New Zealand dairy systems, further investigation of dominance effects may improve understanding of the genetic architecture underlying economically important production and fertility traits. Genome-wide association studies incorporating dominance effects may help identify loci contributing to non-additive genetic variation. Therefore, the aims of this study were to compare pedigree-based and genomic estimates of heritability, estimate dominance variance, and identify SNP-level additive and dominance associations through genome-wide association analyses for milk production, milk composition, and fertility traits in New Zealand dairy cattle.

Chapter 2

Literature Review

2.1 The New Zealand dairy production system

The New Zealand dairy industry is a major contributor to the national economy as the country's largest export sector, generating \$27 billion in exports in 2025, and a quarter of total export value (DairyNZ 2025a). New Zealand dairy farming systems are predominantly pasture-based, where seasonal calving is aligned with spring pasture growth to optimise feed efficiency (Verkerk 2003). Compared with total mixed ration systems that are more commonly used internationally, New Zealand's low input model has several advantages including lower enteric methane emissions (O'Neill et al. 2011), reduced production costs (Schären et al. 2016), lower mortality rates (Burow et al. 2011), and improved animal welfare through provision of a "more natural" environment (Verkerk 2003).

However, reliance on pasture for production also presents several challenges. The seasonal feed supply of pasture-based systems requires cows to achieve high reproductive performance. A high proportion of cows must become pregnant early in the mating period in order maintain a compact calving pattern aligned with peak pasture growth (Harris and Montgomerie 2001). Variable conditions such as autumn drought or excessive winter rainfall can reduce pasture growth, thereby limiting feed availability, which can restrict feed intake and lead to lower milk production, loss of body condition score, and fertility issues in cows (Verkerk 2003). Cows fed a pasture-only diet generally have lower dry matter intake than those fed total mixed ration diets, due to physical constraints such as reduced dry matter per bite, lower bite rate, and limited grazing time, which can result in lower energy intake and reduced milk production (Kolver and Muller 1998; O'Neill et al. 2011). In addition to these production constraints, dairy industries globally face increasing pressure to reduce environmental impacts and improve animal welfare under the influence of climate change (Jago et al. 2023). These challenges highlight the need for ongoing genetic improvement in fertility, adaptability, and productivity which are critical for improving economic efficiency across both New Zealand and global dairy industries (Harris 2005).

2.2 Breeds and crossbreeding

New Zealand's national dairy herd consists of approximately 4.7 million cows of which 23.9% are Holstein-Friesian (F), 7.5% are Jersey (J), 60.4% are F×J crossbred, and 8.2% are other breeds (Figure 2.1 (DairyNZ and Livestock Improvement Corporation 2024)). Figure 2.1 illustrates the proportional representation of major dairy breeds, including Holstein-Friesian, Jersey, Holstein-Friesian × Jersey crossbred cows, and other less common breeds. Percentages are based on the most recent national herd recording data (DairyNZ and Livestock

Improvement Corporation 2024). This breed distribution reflects the widespread use of crossbreeding in New Zealand as a strategy to utilise additive and non-additive genetic components of different breeds (Sørensen et al. 2008).

Crossbreeding improves the performance of progeny compared to the average of their purebred parents due to breed complementarity and heterosis (Sørensen et al. 2008; Harris 2022). Heterosis refers to the increased performance of crossbred animals compared with the average of their purebred parental breeds (Sørensen et al. 2008). This improvement results from non-additive genetic effects such as dominance, overdominance, and epistasis which arise from different allele combinations created through the increased heterozygosity between two breeds (Wakchaure et al. 2015). Traits with low heritability, such as reproductive and fitness traits, generally show the greatest response to heterosis, often exhibiting at least 10–15% heterosis (Sørensen et al. 2008). In contrast, production traits have moderately high heritability and exhibit lower heterosis effects of around 3% (Sørensen et al. 2008).

In the New Zealand dairy population, heterosis effects are evident for both production and fertility traits. Industry level comparisons of productivity for crossbred and purebred cows from the 2020 national genetic evaluation showed the F×J cows produced an average of 366 kg milk solids per cow, which is 26.5 kg greater than the mid-parental breed average of 339.5 kg (Harris 2022). Milk solids production per kilogram of liveweight was also higher in F×J cows (0.73) than in F (0.67) or J (0.69) cows (Harris 2022). In a smaller population of New Zealand dairy cattle, Jayawardana et al. (2023b) found that F×J cows achieved up to a 235 kg increase in milk yield, 16 kg increase in fat yield, and 11 kg increase in protein yield, as well as a 5.8% increase in pregnancy rates at 21 and 42 days after the planned start of mating, and 3.7% increase in pregnancy rate after first service compared with the mid-parental breed average. Therefore, crossbreeding is particularly valuable in New Zealand's pasture-based system, where milk production and fertility traits influence dairy farm profitability and productivity. The benefits of crossbreeding demonstrate that non-additive effects are economically relevant, highlighting the need to better understand them in dairy cattle populations.

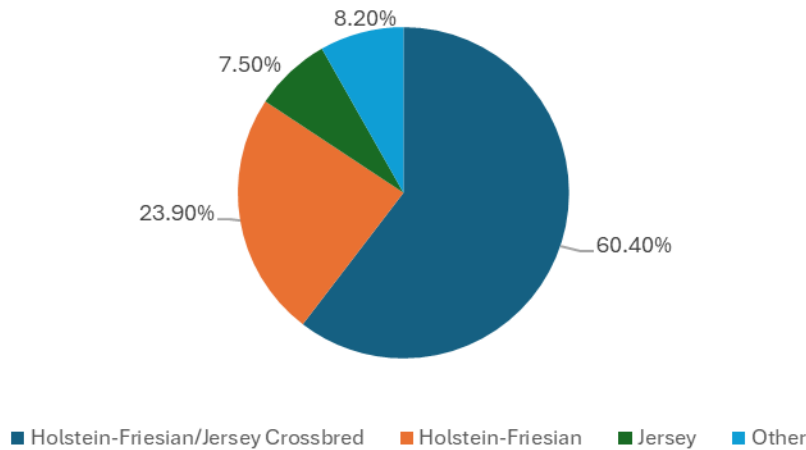


Figure 2.1 Breed composition of the New Zealand dairy cattle population.

2.3 National genomic evaluation system

New Zealand's system for the evaluation of dairy cattle combines milk yield data from herd testing, and conformation and fertility-related trait data to estimate the genetic merit of dairy cattle for genetic selection and breeding decisions. It utilises the national index, Breeding Worth (BW), which ranks dairy cattle based on their ability to meet the national breeding objective of producing cows that efficiently convert feed into profit (DairyNZ 2024a). The BW index incorporates breeding values (BVs) for lactation yields of milk, fat and protein, liveweight, fertility, body condition score, udder conformation, somatic cell score, gestation length, and functional survival, along with the corresponding economic values for each trait (DairyNZ 2024a).

Implementation of the selection index has resulted in the improvement of economically important traits such as feed efficiency and production of milk solids (DairyNZ 2025b). For example, genetic improvement contributes 50% of annual profit and productivity gains, equating to around \$25,000 per farm, and herds in the top 10% for BW achieve approximately \$12,600 more annual profit than median herds due to genetic selection (DairyNZ 2025b). Economic modelling over a 10 year period estimated the present value of these genetic improvements to be approximately \$123,000 per farm after accounting for the time value of money, highlighting the economic value of genetic selection within the New Zealand dairy industry (Lopez-Villalobos et al. 2024). Because BVs are used to construct BW, the effectiveness of the index is dependent on the accuracy of BV estimates. Therefore, improving the accuracy of predictions of animals' genetic merit is essential to the ongoing success of New Zealand's dairy industry.

2.3.1 Pedigree-based evaluations

The BW index is derived from estimated BVs, which predict an animal's genetic merit for different traits and its potential to pass desirable traits onto its offspring (DairyNZ 2024b). Traditionally, these BVs are calculated using pedigree and performance records collected on-farm, such as body weights, herd test results, conformation traits, and calving and mating events (DairyNZ 2014). Best Linear Unbiased Prediction (BLUP) methodology, originally developed by Henderson (1949), uses these records to predict BVs within a mixed model framework that estimates fixed effects and BVs simultaneously (Mrode 2005). The BLUP model provides unbiased linear predictions of an animal's genetic merit by minimising prediction error variance and maximising the correlation between true BVs (the additive genetic value of an animal) and predicted BVs (Mrode 2005). This allows records to be used without being pre-adjusted for environmental effects which are often unknown (Mrode 2005).

The accuracy of BVs is dependent on the amount of information available and therefore will decrease when performance or pedigree data are limited (Van der Werf 2015). Specifically, sires do not express economically important traits such as milk production and therefore lack performance records. To address this limitation, young sires have historically been evaluated through progeny testing, where offspring performance is used as an indirect measure of an animal's genotype, thereby improving the accuracy of BVs (Robertson and Rendel 1950). However, this lengthens the generation interval, which can offset the benefits of increased selection accuracy by slowing the rate of genetic improvement (Robertson and Rendel 1950). In New Zealand, BVs form the basis of the BW index, but the limitations of pedigree-based evaluation have led to the adoption of genomic evaluation to improve the accuracy of dairy cattle selection (Spelman et al. 2010).

2.3.2 Genomic evaluation

Genomic evaluation uses genome-wide markers to predict the genetic merit of individuals, which is used to select more favourable animals for breeding (Zhang et al. 2011). Advances in molecular genetics, particularly the sequencing of the bovine genome, have led to the development of dense single nucleotide polymorphism (SNP) marker maps and the widespread genotyping of dairy cattle, which have provided the foundation for implementing genomic selection (Hayes et al. 2009). Meuwissen et al. (2001) proposed a model to estimate genomic breeding values (GBVs), in which genome-wide SNP effects are estimated from a reference population containing both genotypic and phenotypic data. This method forms the basis of modern genomic evaluation systems. The use of genome-wide SNP markers ensures that most

quantitative trait loci (QTLs), which are genomic regions that contribute to variation in complex traits, are in linkage disequilibrium with at least one marker (Hayes and Goddard 2010). Linkage disequilibrium refers to the non-random association of alleles at different loci, allowing SNP markers to be used as proxies for causal variants. This enables genomic markers to capture genetic variation contributing to economically important traits.

To incorporate genome-wide SNP information into statistical models, genomic relationships between individuals are estimated using a genomic relationship matrix (GRM). The GRM estimates the relatedness of individuals using genotypic data rather than expected pedigree relationships. Because actual relationships between animals may deviate from expected pedigree relationships due to the effects of Mendelian sampling, GRMs can provide more precise estimates of genetic relationships between animals than pedigree-based relationship matrices (Veerkamp et al. 2011). Therefore, incorporating GRMs into mixed models may improve the accuracy of estimating breeding values and variance components (Veerkamp et al. 2011).

Statistical approaches such as Genomic Best Linear Unbiased Prediction (GBLUP) and Bayesian methods, including BayesA, BayesB, and BayesC, are commonly used to estimate genomic BVs (Wang et al. 2012). The GBLUP approach assumes that all SNPs contribute small and equal effects to a trait, whereas Bayesian methods assume a distribution of marker-specific variances in which some loci have larger effects than others. Genomic relationship matrices can also be applied to estimate variance components and perform genome-wide association studies (GWAS), which aim to identify individual SNPs associated with variation in complex traits and provide insight into their genetic architecture. These approaches enable the integration of genomic information into animal evaluation systems which can improve the accuracy of estimated BVs and support more effective selection decisions in dairy cattle breeding.

2.3.3 Advantages of genomic selection

Genomic selection in dairy cattle breeding provides several advantages over traditional selection based solely on pedigree and phenotypic records (Meuwissen et al. 2001). One major advantage is that the use of genotype information can provide more accurate estimations of BVs, especially for young animals with no performance or progeny records (Scheifers and Weigel 2012). In genomic selection, the genetic merit of young, unproven animals can be predicted based on their genotype. This allows younger, genetically favourable animals to be

selected for breeding programmes. Consequently, the use of genomically selected young animals reduces generation interval and increases selection accuracy, thereby increasing the rate of genetic gain, according to the breeder's equation (Falconer 1996), by an estimated 40–50% (Spelman et al. 2010; Schefers and Weigel 2012).

Furthermore, breeders can use genomic testing, which uses DNA markers across the genome to estimate an animal's genetic merit, to evaluate a much larger group of animals than with traditional progeny testing. This allows more animals to be screened at an earlier age. As a result, selection intensity increases and the costs associated with progeny testing used to evaluate sire genetic merit are reduced, further accelerating genetic improvement within dairy herds (Schaeffer 2006; Matthews et al. 2019).

2.4 Considering dominance genetic effects in the genomic evaluation of dairy cattle

Dairy cattle genomic evaluation models primarily focus on additive genetic effects, where individual alleles contribute independently and linearly to the total genetic value (Rettew et al. 2008). In contrast, non-additive genetic variation is typically excluded from genomic evaluation models as it is not consistently inherited across generations or populations due to changes in the frequency of alleles (Rettew et al. 2008; Varona et al. 2018). As a result, the BW index is based on BVs that reflect only additive variance and does not aim to capture heterosis or dominance variance (Johnson et al. 2018). However, excluding non-additive effects may limit the accuracy of BVs and overlook opportunities to exploit non-additive genetic variation (Varona et al. 2018).

One of the main sources of non-additive genetic variation contributing to heterosis is dominance variance, which arises due to within-locus interactions between alleles, causing the deviation of a heterozygote's genotypic value from the mean value of the two homozygotes (Ertl et al. 2014). Dominance effects may have a large contribution toward the genetic variation of complex traits, and therefore are important to consider in genomic evaluation models (Aliloo et al. 2016).

2.4.1 Pedigree-based estimation of dominance effects

Traditionally, dominance effects have been estimated using pedigree relationship matrices, but results in the literature are variable (Table 2.1). This is because the proportion of dominance variance shared between relatives is often small and confounded by other genetic or environmental effects, making it difficult to obtain accurate and consistent estimates of dominance variances (Fuerst and Sölkner 1994). Most pedigree-based studies have reported

dominance contributions of between 1–5% of phenotypic variance for milk production traits (Miglior et al. 1995; Van Tassell et al. 2000). In contrast, some inflated estimates, such as the 24% reported for fat yield by Tempelman and Burnside (1990), are likely biased upwards due to maternal or environmental effects included in the dam component confounding with dominance variance under the hierarchical sire-dam model.

While these findings highlight the potential role of dominance in the phenotypic variance of key production traits, pedigree-based dominance estimates are rarely included in current dairy cattle evaluations. This is primarily due to the high computational demand required to compute the inverse of dominance relationship matrices in large-scale genetic evaluations (Liu et al. 2022). In addition to this, weak pedigree-based non-additive relationships between individuals make it difficult to separate dominance effects from environmental effects, resulting in low accuracy of dominance estimates (Liu et al. 2022). There are also limited records of full sibling families with strong dominance relationships, which are required to estimate dominance effects from pedigree data (Ertl et al. 2014; Aliloo et al. 2016). Collectively, these constraints reduce the reliability and applicability of pedigree-based approaches for incorporating dominance effects into large-scale genetic evaluations. This highlights the need to use genome-wide genotypic information to more accurately estimate dominance effects.

2.4.2 Genomic estimation of dominance effects

With the increasing availability of genomic information, covariance matrices of genomic dominance relationships among individuals can now be computed using a GBLUP model, providing a more straightforward approach compared to traditional pedigree-based dominance relationship matrices (Ertl et al. 2014; Garcia-Baccino et al. 2020). Using SNP data, heterozygote genotypes can be identified across loci which allows for the separation of additive and dominance effects (Vitezica et al. 2013). This can be used to improve the prediction of breeding values and the genotypic values of future matings (Toro and Varona 2010). Therefore, the availability of genomic information provides new opportunities for incorporating dominance effects into mate allocation strategies and for quantifying their contribution to economically important traits in dairy cattle. As a result, there has been a renewed interest in the prediction of dominance variances in dairy cattle breeding and their potential for integration into genomic evaluation.

Several international studies have quantified the dominance variance estimates for milk production traits using genomic information (Table 2.1), but results vary across populations

and traits (Ertl et al. 2014; Sun et al. 2014; Aliloo et al. 2016; Schneider et al. 2023). When expressed as a proportion of total phenotypic variance, dominance variance accounted for between 3.2–7.1% in Australian Holstein and Jerseys (Aliloo et al. 2016), up to 5% in USA Holsteins and 7% in USA Jerseys (Sun et al. 2014), and between 4.5–7.8% in German Holsteins (Schneider et al. 2023) for milk, fat, and protein yield traits. In New Zealand, Reynolds et al. (2022) reported dominance variance estimates of 4.4–5.9% of phenotypic variance in Holstein-Friesians and 6.4–7.4% in Jerseys, while estimates for milk composition traits were low, ranging from 0–1.5%. These results indicate that while additive effects have the greatest contribution to phenotypic variance, dominance can contribute a measurable proportion of phenotypic variation in some production traits but may vary between populations and studies.

Table 2.1. Estimates of additive and dominance genetic variance using pedigree and genomic information for milk traits in dairy cattle.

Breed and country		Trait	Additive variance as % of phenotypic variance (h ²)	Dominance variance as % of phenotypic variance (d ²)	Reference
Pedigree information					
Holstein	Canada	Milk yield	40.0	6.0	Tempelman and Burnside (1990)
		Fat yield	32.0	24	
	Canada	Milk yield	29.1	3.0	Miglior et al. (1995)
		Fat yield	29.3	2.8	
		Protein yield	24.2	1.1	
		Fat %	55.7	2.1	
		Protein %	68.3	1.1	
	United States	Milk yield	34.2	5.3	Van Tassell et al. (2000)
		Fat yield	33.8	5.0	
Protein yield	31.2	5.2			
Genomic information					
Holstein	United States	Milk yield	27.0	5.1	Sun et al. (2014)
		Fat yield	23.3	5.1	
		Protein yield	20.2	5.3	
		Fat %	46.4	1.1	
		Protein %	48.7	1.3	
	Australia	Milk yield	25.6	3.3	Aliloo et al. (2016)
		Fat yield	18.4	3.8	
		Protein yield	17.4	3.2	
	Germany	Milk yield	34.7	7.8	Schneider et al. (2023)
		Fat yield	27.3	5.3	
		Protein yield	27.6	4.5	
	Holstein- Friesian	New Zealand	Milk yield	29.6	4.4
Fat yield			26.1	5.9	
Protein yield			23.5	5.3	
Fat %			70.0	0.6	
		Protein %	64.2	0	

Table 2.1. Continuation

Jersey	United States	Milk yield	32.2	7.0	Sun et al. (2014)
		Fat yield	19.2	7.2	
		Protein yield	23.0	7.0	
		Fat %	54.0	0.9	
		Protein %	55.6	3.0	
	Australia	Milk yield	30.6	5.5	Aliloo et al. (2016)
		Fat yield	18.6	6.6	
		Protein yield	20.4	7.1	
	New Zealand	Milk yield	31.2	6.4	Reynolds et al. (2022)
		Fat yield	23.2	7.4	
		Protein yield	23.6	7.3	
		Fat %	61.6	1.5	
		Protein %	63.6	0.5	

In contrast to milk production traits, relatively few studies have investigated dominance variance in fertility traits (Table 2.2). Using pedigree information, Hoeschele (1991) estimated both additive and non-additive effects for fertility in American Holstein cattle. Hoeschele (1991) reported that, although additive variance was small, dominance variance was at least as large as the additive component, when using days open (number of days between calving and conception) and service period (number of days between first and last inseminations) as measures of fertility. More recently, Ghiasi et al. (2018) estimated both additive and dominance variance for fertility traits in Iranian Holstein cattle using pedigree-based models, and also reported dominance variance exceeded additive genetic variance for days open, days from calving to first service, and number of services to conception.

Genomic studies have reported lower estimates of dominance variance for fertility traits. Aliloo et al. (2016) found very low dominance variance in calving interval, contributing only 1.2% in Holsteins and close to zero in Jersey cows. Conversely, Sun et al. (2014) found that dominance variance was close to zero for daughter pregnancy rates in USA Holsteins, and 1.2% in USA Jersey cattle. This likely reflects the low heritability of these traits and the large environmental variance, which reduce the accuracy of dominance and additive variance estimates (Aliloo et al. 2016). Overall, these findings indicate that while low, non-additive genetic variance could be relatively important in fertility traits where additive variance is low as they could contribute equally to phenotypic variance (Hoeschele 1991; Druet et al. 2001).

Table 2.2. Estimates of additive and dominance genetic variance using pedigree and genomic information for fertility traits in dairy cattle.

Breed and country		Trait	Additive variance as % of phenotypic variance (h^2)	Dominance variance as % of phenotypic variance (d^2)	Reference
<i>Pedigree information</i>					
Holstein	United States	Days open	2.0	2.3	Hoeschele (1991)
		Service period	0.8	1.4	
	Iran	Days from calving to first service	6.8	26.0	Ghiasi et al. (2018)
		Days open	6.7	19.6	
		Number of services to conception	3.5	23.1	
<i>Genomic information</i>					
Holstein	United States	Daughter pregnancy rate	5.7	0	Sun et al. (2014)
		Calving interval	2.1	1.2	
Jersey	United States	Daughter pregnancy rate	3.0	1.2	Sun et al. (2014)
		Calving interval	2.2	0	
	Australia	Calving interval			Aliloo et al. (2016)

Despite evidence that dominance contributes to phenotypic variation in several economically important traits, there remains a high level of inconsistency in the estimation of dominance variance across studies. This may be because dominance variance is more challenging to estimate than additive variance due to its stronger dependence on linkage disequilibrium (LD) (Wei et al. 2014). Additive variance is proportional to r^2 , where r is the correlation between a marker and the causal variant, whereas dominance variance is proportional to r^4 , or $(r^2)^2$, so detection of dominance effects is dependent on much higher LD than additive effects (Wei et al. 2014). Therefore, higher-density genotypes are required to accurately detect dominance variance compared with additive variance (Wei et al. 2014). Additionally, large datasets are needed to separate dominance from environmental effects to prevent overestimation of dominance variance (Aliloo et al. 2016).

Variation in results between studies may be due to several factors, including the size of datasets used, differences in the definition of traits, levels of inbreeding within breeds or populations, the use of pedigree vs genomic data, and differences in the methods used (Aliloo et al. 2016). Schneider et al. (2023) reported wide highest posterior density (HPD) intervals for estimates of dominance variance, where the HPD interval represents the range containing 95% of the posterior probability distribution for the estimated parameter. For example, the HPD interval for dominance variance as a proportion of phenotypic variance for fat yield ranged from 0.006 to 0.112, compared to 0.226 to 0.326 for additive variance, indicating greater uncertainty in dominance variance estimates (Schneider et al. 2023). This limits their potential use in national genomic evaluation and highlights the need for further analyses to determine the role of dominance effects in improving genomic prediction and mate allocation for dairy cattle production traits (Schneider et al. 2023).

2.5 Modelling SNP-level additive and dominance effects in genome-wide association studies

Genome-wide association studies (GWAS) that integrate dominance effects could provide a more detailed understanding of the genetic architecture influencing economically important traits (Schneider et al. 2023). Although most GWAS focus on identifying additive genetic SNP effects only, GWAS models can be extended to test for dominance deviations at individual loci. This may be useful for investigating the specific loci contributing to non-additive variation in dairy cattle production traits.

Some studies have carried out SNP-level analyses to estimate dominance effects using different approaches. Boysen et al. (2013) carried out a GWAS in German Holstein cattle using genotype probabilities derived from bull genotypes. In their analyses, additive and dominance covariates were fitted as fixed effects in a mixed-model single-marker regression while accounting for relatedness between animals using a pedigree-based additive relationship matrix. This approach identified several genomic regions with significant dominance effects contributing to milk yield and composition traits (Boysen et al. 2013).

A different method was used by Jiang et al. (2019), who applied a large-scale GWAS in USA Holstein cattle using an approximate generalised least squares analysis to test both additive and dominance SNP effects. This approach estimated SNP genotypic values for the three genotype classes and derived additive and dominance contrasts from these estimates for significance testing of SNP effects, using pre-estimated BVs from routine genomic evaluation to correct for

sample stratification. A small number of dominance effects were found for milk, fat, and protein yields, with lower statistical significance than additive effects. However, for fertility traits, dominance effects were found to have levels of significance similar to those observed for additive effects (Jiang et al. 2019).

Sun et al. (2014) investigated additive and dominance SNP effects in USA Holstein and Jersey cattle using SNP-BLUP models that jointly fitted additive and dominance genomic relationship matrices within mixed model equations. Manhattan plots of additive and dominance effects identified several genomic regions with large dominance associations for milk production traits, particularly on chromosome 14, but dominance effects for fertility and health traits were generally small (Sun et al. 2014). However, their study primarily focused on estimating genome-wide genetic effects and evaluating the impact of dominance on genomic prediction accuracy.

Collectively, these studies demonstrate that dominance effects can be incorporated into GWAS using different approaches and may contribute to economically important dairy cattle traits. However, dominance associations are generally weaker, fewer in number, and less consistent across studies than additive effects. Identifying loci where heterozygous individuals outperform homozygotes may improve understanding of the genetic architecture underlying economically important traits and could be useful for breeding strategies, including informing mate allocation decisions designed to exploit specific combining ability (Sun et al. 2014).

Although international studies have investigated the contribution of dominance effects to dairy cattle traits using genomic information, research on SNP-level and genome-wide dominance effects in New Zealand dairy cattle remains limited. Reynolds et al. (2021) carried out a GWAS to measure non-additive effects in the New Zealand dairy cattle population for body weight, stature, and body condition score. They identified eight significant non-additive QTLs for both body weight and stature that were not detected in the additive-only model, indicating that non-additive inheritance can influence important phenotypes in this population (Reynolds et al. 2021). Reynolds et al. (2022) further investigated non-additive genetic effects in a large New Zealand dairy cattle dataset to identify non-additive QTLs affecting lactation traits, including loci with recessive effects associated with potentially deleterious mutations. Their GWAS identified several loci with strong recessive effects on milk yield traits, and some previously reported additive QTLs for milk composition traits that exhibited small dominance effects (Reynolds et al. 2022). Together, these studies demonstrate that non-additive genetic effects

may contribute to traits in New Zealand dairy cattle populations. However, the role of quantitative dominance effects in complex production and fertility traits has not been extensively investigated.

2.6 Potential benefits of including dominance in genomic evaluation

2.6.1 Improve accuracy of genomic prediction models

Results across studies have indicated that if the contribution of dominance variance to the phenotypic variance of a trait is substantial, including dominance effects could improve the predictive ability of genomic evaluation models (Aliloo et al. 2016). Consistent with this, Aliloo et al. (2016) and Sun et al. (2014) reported that estimates of total genetic effects, that include both additive BVs and dominance deviations, generally showed higher correlations with phenotype than predictions based on additive effects alone. Aliloo et al. (2016) observed statistically significant improvements in predictive ability, based on the correlation between predicted genetic values and phenotype, for fat yield ($p < 0.01$) and protein yield ($p < 0.05$) in Holstein cows. In addition, Sun et al. (2014) reported significant improvements in predictive ability ($p < 0.001$) for milk, fat, and protein yields in both Jerseys and Holstein cows, as well as for somatic cell score (SCS) in Holsteins.

Aliloo et al. (2016) also reported smaller mean squared prediction errors, with regression coefficients on estimated genetic values closer to 1 for the additive + dominance models, further indicating better predictive ability than the additive-only model. Maximum likelihood ratio tests also showed that models which included both additive and dominance effects provided a significantly better fit to the observed phenotypic data for milk, fat, and protein yields than additive-only models, as indicated by significantly higher log-likelihood values ($p < 0.01$) (Sun et al. 2014; Aliloo et al. 2016). However, for calving interval and SCS, no significant differences in either goodness of fit or predictive performance were found between the two models (Sun et al. 2014; Aliloo et al. 2016). Conversely, Ertl et al. (2014) found no improvement in predictive ability of the model when including dominance effects for yield traits, but highlighted that this could be attributed to the small proportion of full-siblings and small dominance relationship coefficients between the training and validation datasets. Overall, these findings indicate that including dominance effects has the potential to improve the accuracy of genomic predictions in models for predicting the future performance of animals, but the extent of improvement depends on both the dataset and the trait being selected for.

2.6.2 Crossbreeding

Because dominance is a key contributor to heterosis, including dominance variance in genomic evaluation models is expected to improve selection efficiency over additive-only models when purebred animals are selected for crossbred performance (González-Diéguez et al. 2020). In simulated pig crossbreeding systems, González-Diéguez et al. (2020) reported that mate allocation strategies which included both additive and dominance increased crossbred performance by 0.5–4%. Similarly, Zeng et al. (2013) compared genomic selection models for improving crossbred performance through purebred selection and found that using models which included dominance effects provided superior response to selection than additive-only models. The level of genetic improvement was proportional to the amount of dominance (heterosis) present, but even when dominance effects were minimal, prediction accuracy was similar to that of additive models (Zeng et al. 2013). Therefore, the use of dominance-inclusive models is recommended for crossbreeding programmes even when dominance is not a major source of trait variability, as they offer equal or superior accuracy to traditional additive-only models (Zeng et al. 2013).

This is especially relevant to New Zealand's dairy industry where crossbreeding is commonly used to exploit heterosis within the national herd. In addition, the increasing use of beef-breed sires over dairy-breed cows has become a strategy to increase the value of surplus calves born on dairy farms. Beef sires can be selected for both calving ease and carcass traits to produce crossbred beef × dairy progeny that are suited to the requirements of dairy farmers, beef finishers and processors (Livestock Improvement Corporation 2023). This crossbreeding system therefore provides an opportunity to incorporate dominance variance into genomic evaluation to identify the most suitable beef sires for mating to dairy cows. Given the high heterosis observed in dairy cattle and the widespread use of crossbreeding within the New Zealand herd, accounting for dominance effects may play an important role in accurately predicting the performance of future crossbred individuals (Varona et al. 2022).

2.6.3 Targeted mating

Dominance information can also be used for targeted mate allocation, allowing breeders to selectively pair parent animals that will produce offspring with favourable dominant combinations. With the availability of SNP genotypes, SNP effect estimates can be directly incorporated into mate allocation algorithms to predict the outcomes of planned matings and identify pairings that will generate the highest-performing offspring (Ertl et al. 2014; González-Diéguez et al. 2020). Ertl et al. (2014) calculated the expected genetic values of all possible

matings between 50 bulls and 1,996 Fleckvieh cows and compared the outcomes of selection using additive BVs or total genetic values. Selecting matings based on total genetic value increased the expected offspring genetic value for milk yield from 143.8 kg to 165.2 kg and for protein yield from 3.24 kg to 4.15 kg compared with selection based on additive BVs alone. This corresponds to increases of 14.8% for milk yield and 27.8% for protein yield, demonstrating the potential value of incorporating dominance effects into mate allocation decisions (Ertl et al. 2014).

Similarly, Sun et al. (2013) evaluated genomic mate allocation programmes in USA Holstein and Jersey cattle by comparing mating strategies that incorporated additive effects or both additive and dominance effects when predicting expected progeny value (EPV) for milk yield. They reported that, regardless of sire selection method, mate allocation method, or inbreeding source, including dominance effects increased the EPV relative to additive-only models. For example, under a genomic mate allocation method using top-ranked bulls, inclusion of dominance effects increased EPV from 878 kg to 964 kg in Holsteins and from 680 kg to 732 kg in Jersey cattle (Sun et al. 2013). These findings highlight the potential benefits of incorporating dominance effects into mate allocation to maximise the expected performance of offspring. However, incorporating dominance into national evaluations remains challenging due to the higher computational requirements of estimating dominance effects, the relatively small magnitude of dominance variance for many traits, and the lower stability of dominance effects across generations compared with additive effects.

2.7 Gaps in knowledge

Genome-wide analyses of quantitative dominance effects and SNP-level dominance deviations for economically important traits in New Zealand dairy cattle remain limited. These include milk production and fertility traits that are important for dairy profitability and are relevant to breeding objectives in New Zealand's pasture-based dairy system. The New Zealand dairy system also relies heavily on crossbreeding to exploit heterosis, meaning that non-additive effects such as dominance may play an important role in improving selection accuracy and mating strategies. As the demand for dairy productivity rises and breeding objectives become increasingly complex, exploiting non-additive genetic variation will likely be critical for the future of genetic improvement of dairy cattle (Aliloo et al. 2016). Further research is needed to quantify genome-wide and SNP-level dominance effects for production and fertility traits in New Zealand dairy cattle using genomic data. This will help to determine whether incorporating dominance into genomic evaluation models could be used to improve selection

and mate allocation decisions in New Zealand's dairy system. Therefore, the objectives of this study were to compare pedigree-based and genomic estimates of heritability, estimate dominance variance using a genomic additive + dominance model, and identify SNP-level additive and dominance associations for milk production, composition, and fertility traits in a New Zealand dairy cattle population.

Chapter 3

Methods and Materials

3.1 Data

This study used herd test, reproductive, and pedigree data of spring-calved dairy cows from the Massey University Dairy 1 and Dairy 4 herds, Palmerston North, New Zealand. Mating, calving, pregnancy diagnosis, 305-day lactation yields, and pedigree records were obtained from the Livestock Improvement Corporation (LIC), New Zealand. The dataset contained 6,382 lactation records from 2,103 cows across the 2014–2022 calving seasons. Both herds had at least three herd-test milk records per lactation during the study period. Dairy 1 operated a once-daily milking system with minimal supplementary feeding and a stocking density of 2.1 cows/ha, whereas Dairy 4 used a twice-daily milking system with greater supplementary feeding and a higher stocking density of 2.8 cows/ha (Jayawardana et al. 2023b).

Genotypic data were matched to phenotypic records using cow ID, and only cows with both phenotypic and genotypic records were retained for analysis. The final dataset comprised 1,667 cows with 5,686 lactation records. Parity was categorised into six groups according to lactation number (1, 2, 3, 4, 5, and ≥ 6). Pedigree information, supplied as breed fractions in sixteenths for Holstein-Friesian (F), Jersey (J) and other breeds, was used to determine breed proportions for each cow. Cows with a proportion of $\geq 14/16$ Holstein-Friesian were classified as F, cows with $\geq 14/16$ Jersey were classified as J, cows with $\geq 10/16$ and $\leq 13/16$ Holstein-Friesian were classified as FX, those with $\geq 10/16$ and $\leq 13/16$ Jersey were classified as JX, and those with $< 10/16$ Holstein-Friesian, and $< 10/16$ Jersey were classified as FJ. Because the number of Jersey cows in Dairy 4 was small ($N=25$), these cows were combined with the JX group for analysis to improve model stability. Therefore, no separate J group is shown for Dairy 4 in Table 3.1

Table 3.1. Number of animals within each herd and breed group.

	F	FX	FJ	JX	J	Total
Dairy 1	144	65	116	70	146	541
Dairy 4	478	380	150	118	0	1126
Total	622	445	266	188	146	1667

F = Holstein-Friesian; J = Jersey; FX = Holstein-Friesian cross; FJ = Holstein-Friesian $\times$ Jersey; JX = Jersey cross.

The heterosis coefficient for each cow was calculated as:

$$h^{F \times J} = \alpha_F^S \alpha_J^D + \alpha_J^S \alpha_F^D$$

where $h^{F \times J}$ is the coefficient of expected heterosis between breeds F and J in the progeny, α_F^S and α_J^S are proportions of F and J in the sire, respectively, and α_F^D and α_J^D are proportions of F and J in the dam, respectively (Dickerson 1973).

3.2 Phenotypic traits

Milk production traits analysed were 305-day milk yield (MY; L), fat yield (FY; kg), protein yield (PY; kg), and lactose yield (LY; kg), and fat percentage (FP; %), protein percentage (PP; %), and lactose percentage (LP; %). Fertility traits included interval from start of mating to first service (SMFS; days), proportion of cows that conceived to their first service within 21 days from the start of mating (PRFS; %), proportion of cows first mated within 21 days from the start of mating (SR21; %), and proportion of cows that conceived within the first 42 days from the start of mating (PR42; %).

3.3. Descriptive statistics and analyses of variance

Descriptive statistics for milk production and fertility traits were obtained using the PROC MEANS from Statistical Analysis Software (SAS) version 9.4 (SAS Institute Inc., Cary, NC, USA). Analysis of variance to evaluate the effects of factors on the dependent traits was performed using SAS PROC MIXED with the following mixed linear model:

$$y_{jklm} = \mu + H_k + O_l + L_m + \beta_1 p_{jklm}^F + \beta_2 h_{jklm}^{F \times J} + \beta_3 d_{jklm} + c_j + e_{jklm}$$

where y_{jklm} is the phenotypic record of cow j; μ is the overall mean; H_k is the fixed effect of herd k; O_l is the fixed effect of calving year l; L_m is the fixed effect of lactation number m; p_{jklm}^F is the proportion of Holstein-Friesian genetics in cow j calving in herd k, year l and lactation number m; $h_{jklm}^{F \times J}$ is the coefficient of F×J heterosis for animal j calving in herd k, year l and lactation number m; d_{jklm} is the deviation from median calving date of animal j calving in herd k, year l and lactation number m; β_1, β_2 and β_3 are regression coefficients of the dependent traits on the proportion of Holstein-Friesian, F×J heterosis and deviation from median calving date; c_j is the permanent environmental effect of cow j assumed to follow a normal distribution, with mean zero and variance $\sigma_c^2 \mathbf{I}$, where σ_c^2 is the permanent environmental variance and $\mathbf{I}$ is identity matrix of order equal to the number of cows with records; and e_{jklm} is the residual error associated with observation y_{jklm} assumed to follow a normal distribution with mean zero and variance $\sigma_e^2 \mathbf{I}$, where σ_e^2 is the residual variance and $\mathbf{I}$ is an identity matrix of order equal to the number of total records.

Least squares means and standard errors were obtained and used for multiple mean comparisons using Fisher's least significant difference test.

3.4 Pedigree-based variance component estimation

Variance components and heritabilities for production and fertility traits were estimated using pedigree information in ASReml version 4.2 (Gilmour et al. 2021). The following mixed model was used:

$$y_{jklm} = \mu + H_k + O_l + L_m + \beta_1 p_{jklm}^F + \beta_2 h_{jklm}^{F \times J} + \beta_3 d_{jklm} + a_j + c_j + e_{jklm}$$

Terms in the model were defined in Section 3.3. Random effects were assumed to have the following distributions: a_j is the vector of additive genetic effects of animal j assumed to follow a normal distribution, with mean zero and variance $\sigma_a^2 \mathbf{A}$, where σ_a^2 is the additive variance and $\mathbf{A}$ is the relationship matrix derived from pedigree information; c_j is the permanent environmental effect of animal j assumed to follow a normal distribution, with mean zero and variance $\sigma_c^2 \mathbf{I}$, where σ_c^2 is the permanent environmental variance and $\mathbf{I}$ is an identity matrix of order equal to the number of animals with records; and e_{jklm} is the residual error associated with observation y_{jklm} assumed to follow a normal distribution with mean zero and variance $\sigma_e^2 \mathbf{I}$, where σ_e^2 is the residual variance and $\mathbf{I}$ is an identity matrix of order equal to the number of total records. Heritability was calculated as:

$$h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_c^2 + \sigma_e^2).$$

3.5 Genotypes and quality control

DNA was extracted from ear punch tissue samples and 1,667 cows were genotyped using the Illumina Bovine 50K SNP chip. Genotypes were imputed to the HD platform as previously described by Wang (2020) (N=637,393 SNPs) and recorded in Illumina A/B allele format, where the A allele was the reference allele at any given position, and B was the alternative allele. All genotypes were converted to 0, 1, or 2, according to the number of alternative alleles present at each position. To reduce computational demand, a random subset of approximately 10% of SNPs per autosome was selected before quality control. Sex chromosomes were excluded because standard estimation of genomic relationships for autosomes is not directly applicable to X-chromosome SNPs (Yang et al. 2011). Quality control excluded SNPs with call rate of <0.95, minor allele frequency <0.01 and strong deviation from Hardy-Weinberg equilibrium (chi-squared values >600). In addition, SNPs were required to have genotype class

frequencies ≥ 0.02 to ensure representation of all three genotype classes for analyses of both additive and dominance effects (Lopes et al. 2014). The final genotype dataset comprised 49,171 SNPs (Table 3.2).

Table 3.2. Number of single nucleotide polymorphisms (SNP) per chromosome after filtering.

Chromosome	Number of SNPs after filtering
1	3,012
2	2,587
3	2,405
4	2,377
5	2,309
6	2,390
7	2,152
8	1,843
9	2,088
10	2,061
11	2,238
12	1,724
13	1,379
14	1,387
15	1,718
16	1,604
17	1,550
18	1,400
19	1,352
20	1,496
21	1,457
22	1,316
23	1,069
24	1,301
25	930
26	1,093
27	936
28	952
29	1,045

3.6 Genomic variance component estimation

Variance components for each trait were estimated using genomic restricted maximum likelihood (GREML) by fitting two linear mixed models: an additive model (including only additive genetic effects) and an additive + dominance model (including both additive and dominance genetic effects). Fixed effects included in the model were herd, calving year, and lactation number, with covariates for proportion of Holstein-Friesian, heterosis, and deviation from median calving date. Random effects included additive genetic effects, dominance genetic effects, permanent environmental effects of cows, and residual effects. Model equations used for variance calculations are as follows.

For the additive model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Zc} + \mathbf{e}$$

For the additive + dominance model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Zd} + \mathbf{Zc} + \mathbf{e}$$

where $\mathbf{y}$ is the vector of phenotypic records in each trait; $\mathbf{b}$ is the vector of fixed effects (herd, calving year, lactation number as class effects and proportion of Holstein-Friesian genes, coefficient of F×J heterosis, and deviation from median calving date as covariates); $\mathbf{a}$ is the vector of additive genetic effects, assumed to be distributed as $\mathbf{a} \sim N(\mathbf{0}, \mathbf{G}\sigma_a^2)$, where σ_a^2 is the additive variance and $\mathbf{G}$ is the genomic relationship matrix (GRM); $\mathbf{d}$ is the vector of dominance genetic effects of animals, assumed to be distributed as $\mathbf{d} \sim N(\mathbf{0}, \mathbf{D}\sigma_d^2)$, where σ_d^2 is the dominance variance and $\mathbf{D}$ is the dominance relationship matrix (DRM); $\mathbf{c}$ is the vector of cow permanent environmental effects, with $\mathbf{c} \sim N(\mathbf{0}, \mathbf{I}\sigma_c^2)$ where σ_c^2 is the permanent environmental variance; $\mathbf{e}$ is the vector of random residual effects of the trait, with $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$ where σ_e^2 is the residual variance; $\mathbf{X}$ is the design matrix for fixed effects and $\mathbf{Z}$ is the incidence matrix relating records to cows. The variance components σ_a^2 , σ_d^2 , σ_c^2 , and σ_e^2 were estimated in the Julia programming language (Bezanson et al. 2017) using the GREMLModels.jl package (Eilertsen 2021).

The GRM was constructed using genome-wide SNP genotypes as described by Van Raden (2008). SNP genotypes were coded as $M_{ji} \in \{0,1,2\}$ for genotypes AA, AB and BB, respectively, carried by animal j at SNP i . The GRM was calculated as:

$$\mathbf{G} = \frac{\mathbf{ZZ}'}{\sum 2p_i q_i}$$

where $\mathbf{Z}$ is the centred genotype matrix with elements $-2p_i$, $1-2p_i$, and $2-2p_i$ for genotypes AA, AB and BB, respectively; p_i is the frequency of the B allele at SNP i , and $q_i=1-p_i$.

The DRM was calculated according to Vitezica et al. (2013) as:

$$\mathbf{D} = \frac{\mathbf{WW}'}{\sum (2p_i q_i)^2}$$

where the dominance matrix $\mathbf{W}$ was coded as $-2p_i^2$, $2p_i q_i$, and $-2q_i^2$ for genotypes AA, AB and BB, respectively. This parameterisation ensures that, under Hardy–Weinberg equilibrium in a non-inbred population, the covariance between additive and dominance genetic effects is zero, and prevents confounding between additive and dominance variance components (Vitezica et al. 2013).

For the additive model, heritability was calculated as:

$$h^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_c^2 + \sigma_e^2).$$

For the additive + dominance model, additive heritability was calculated as:

$$h_a^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_d^2 + \sigma_c^2 + \sigma_e^2).$$

The proportion of dominance variance was calculated as:

$$h_d^2 = \sigma_d^2 / (\sigma_a^2 + \sigma_d^2 + \sigma_c^2 + \sigma_e^2).$$

3.7 Genome-wide association studies

Genome-wide association studies (GWAS) were carried out for all traits to estimate additive and dominance effects for individual SNPs using a mixed model fitted to include all records. To reduce the effects of proximal contamination between the tested SNP and the genomic relationship matrix, a leave-one-chromosome-out (LOCO) approach was implemented (Widmayer et al. 2022). For each chromosome k , an additive GRM was constructed using all quality-controlled SNPs except those located on chromosome k .

Variance components were re-estimated separately for each chromosome using a GREML model including additive genetic, permanent environmental, and residual effects, as described

in Section 3.6. These variance component estimates were then held constant for SNPs tested within that chromosome.

For each SNP i , the following model was fitted:

$$y_{jklm} = \mu + H_k + O_l + L_m + \beta_1 p_{jklm}^F + \beta_2 h_{jklm}^{F \times J} + \beta_3 d_{jklm} + \beta_4 s_i^a + \beta_5 s_i^d + a_j + c_j + e_{jklm}$$

where s_i^a is the additive genotype covariate for SNP i ; s_i^d is the dominance genotype covariate for SNP i ; β_4 is the additive SNP effect; β_5 is the dominance SNP effect; a_j is the vector of additive genetic effects of animal j assumed to follow a normal distribution, with mean zero and variance $\sigma_a^2 \mathbf{G}_{-k}$ where $\mathbf{G}_{-k}$ is the genomic relationship matrix constructed excluding SNPs located on chromosome k ; c_j is the permanent environmental effect of animal j assumed to follow a normal distribution, with mean zero and variance $\sigma_c^2 \mathbf{I}$, where $\mathbf{I}$ is an identity matrix of order equal to the number of animals with records, and e_{jklm} is the residual error associated with observation y_{jklm} assumed to follow a normal distribution, with mean zero and variance $\sigma_e^2 \mathbf{I}$, where $\mathbf{I}$ is an identity matrix of order equal to the number of total records.

For each SNP with genotype $g \in \{0,1,2\}$ and allele frequency p_i , additive and dominance covariates were constructed using the same parameterisation applied in the construction of the GRM and DRM in Section 3.6 (Vitezica et al. 2013). A dominance relationship matrix was not included in the GWAS model because of the additional computational demand. Dominance effects were evaluated as fixed SNP-specific dominance covariates.

For each SNP, the additive and dominance covariates were fitted jointly with the fixed effects matrix $\mathbf{X}$, and mixed model equations were constructed using the estimated variance components. Solutions for β_4 and β_5 were obtained by solving the mixed model equations using variance components fixed at their previously estimated values.

The additive and dominance SNP estimates from the GWAS were represented in Manhattan plots where $-\log_{10}(\text{p-values})$ for each SNP were plotted against genomic position for each autosomal marker. Manhattan plots were generated using the qqman package (Turner 2014) in R software version 4.4.3 (R Core Team 2025). The significance threshold values were based on the Bonferroni significance threshold, calculated as $0.05/m$, where m is the total number of SNPs tested ($0.05/49,171 = 1.02 \times 10^{-6}$) and the suggestive significance threshold, calculated as $1/m$ ($1/49,171 = 2.03 \times 10^{-5}$). On the $-\log_{10}(\text{p-value})$ scale, the Bonferroni threshold was

plotted at 5.99, and the suggestive threshold at 4.69, and these horizontal lines were added to the Manhattan plots to indicate genome-wide and suggestive significance levels.

3.8 Estimation of genotype effects for significant SNPs

The SNP with the strongest dominance association signal for each trait was selected, and genotype means were calculated for the three genotype classes. Each candidate SNP was excluded from the SNP dataset to construct the GRM, and the SNP genotype was fitted as a class variable with three levels (AA, AB, and BB) rather than as additive and dominance covariates used in the previous GWAS model:

$$y_{jklm} = \mu + H_k + O_l + L_m + \beta_1 p_{jklm}^F + \beta_2 h_{jklm}^{F \times J} + \beta_3 d_{jklm} + S_{it} + a_j + c_j + e_{jklm}$$

where S_{it} is the effect of genotype t (AA, AB, and BB) at SNP i fitted as a fixed class effect. All other model terms are as defined in the previous single-SNP association model in Section 3.7.

Chapter 4

Results

4.1 Descriptive statistics

Descriptive statistics for the milk production, composition, and fertility traits of cows from Massey University Dairy 1 and Dairy 4 herds are presented in Table 4.1. Milk production traits were expressed as 305-day lactation yields. The average 305-day milk yield (MY) was 5,020 L, and average fat yield (FY), protein yield (PY), and lactose yield (LY) were 234 kg, 190 kg, and 253 kg, respectively. For milk composition traits, cows produced an average of 4.75% fat, 3.83% protein, and 5.05% lactose per L of total milk. The average start of mating to first service period (SMFS) was 9.2 days, and submission rate at 21 days (SR21) was 93.2%. Additionally, the average pregnancy rate to first service (PRFS) was 53.7% and the average 42-day pregnancy rate (PR42) was 78.9%.

Table 4.1. Descriptive statistics for milk production, milk composition, and fertility traits of cows from Massey University Dairy 1 and Dairy 4 herds.

Trait ¹	No. of records	Mean	Standard Deviation	Coefficient of Variation	Minimum	Maximum
MY	5,686	5,020	1,240	24.7	387	8,861
FY	5,686	234	50	21.4	22	398
PY	5,686	190	42	21.9	19	311
LY	5,686	253	63	24.9	11	449
FP	5,686	4.75	0.69	14.6	2.92	7.38
PP	5,686	3.83	0.32	8.37	3.00	5.14
LP	5,686	5.05	0.17	3.48	2.84	5.48
SMFS	5,514	9.2	7.5	81.5	0	66
PRFS	5,530	53.7	49.9	92.8	0	100
SR21	5,530	93.2	25.3	27.1	0	100
PR42	5,530	78.9	40.8	51.8	0	100

¹ MY = 305-day milk yield (L); FY = 305-day fat yield (kg); PY = 305-day protein yield (kg); LY = 305-day lactose yield (kg); FP = fat percentage; PP = protein percentage; LP = lactose percentage; SMFS = start of mating to first service (d); PRFS = cow conceived to first service in the first 3 weeks of the breeding season (%); SR21 = cow inseminated in the first 3 weeks from the start of mating (%); PR42 = in-calf rate at 42 days from the start of mating (%).

4.2 Effect of fixed effects

4.2.1 Breed

Least-squares means ($\pm$ standard error) for milk production, composition, and fertility traits by breed within farm are presented in Table 4.2. Breed differences were observed for several traits

within farms ($p < 0.05$). In both Dairy 1 and Dairy 4, Holstein-Friesian (F) cows had the highest 305-day MY, whereas Jersey (J) cows in Dairy 1 and Jersey cross (JX) cows in Dairy 4 had the lowest MY. In Dairy 1, crossbred cows produced the highest FY, while no significant differences were observed between breeds for FY in Dairy 4.

In Dairy 1, J cows had the highest FP and PP, whereas F cows had the lowest. LP was also higher in J and crossbreds in Dairy 1 compared with F. Dairy 4 had similar trends for milk composition traits, where JX cows had the highest FP and PP, and F had the lowest FP and PP.

For fertility traits, there were no significant differences between breeds for SMFS in either farm. In Dairy 1, there were also no significant breed differences for PRFS or SR21. However, in Dairy 4, JX cows had higher PRFS and SR21 than F, FX, and FJ cows. For PR42, F, FX, and FJ cows had higher in-calf rates than JX and J cows, but there were no significant differences between breeds in Dairy 4.

Table 4.2. Least-squares means and standard errors for milk production and fertility traits by breed within the Massey University Dairy 1 and Dairy 4 herds.

Trait ¹	Dairy 1					Dairy 4			
	F	FX	FJ	JX	J	F	FX	FJ	JX
MY	4700±52 ^A	4623±76 ^A	4332±67 ^B	4329±76 ^B	3701±43 ^C	5749±29 ^A	5531±32 ^B	5205±48 ^C	4739±57 ^D
FY	219±2.4 ^{BC}	224±3.5 ^{AB}	223±2.6 ^{AB}	228±3.5 ^A	213±2.4 ^C	245±1.3 ^A	248±1.5 ^A	249±2.2 ^A	243±2.6 ^A
PY	180±1.8 ^A	182±2.6 ^A	175±1.9 ^B	179±2.6 ^{AB}	157±1.8 ^C	208±1.0 ^A	206±1.1 ^A	197±1.7 ^B	186±2.0 ^C
LY	231±2.7 ^A	229±3.9 ^A	215±2.9 ^B	216±3.9 ^B	184±2.7 ^C	290±1.5 ^A	280±1.7 ^B	265±2.5 ^C	241±3.0 ^D
FP	4.7±0.04 ^E	4.9±0.06 ^D	5.2±0.04 ^C	5.5±0.05 ^B	5.8±0.04 ^A	4.3±0.02 ^D	4.5±0.02 ^C	4.8±0.04 ^B	5.2±0.04 ^A
PP	3.9±0.02 ^E	3.9±0.03 ^D	4.0±0.02 ^C	4.2±0.03 ^B	4.3±0.02 ^A	3.6±0.01 ^D	3.7±0.01 ^C	3.8±0.02 ^B	3.9±0.02 ^A
LP	4.9±0.01 ^B	4.9±0.02 ^{AB}	5.0±0.01 ^A	5.0±0.02 ^A	5.0±0.01 ^A	5.1±0.01 ^B	5.1±0.01 ^{AB}	5.1±0.01 ^A	5.1±0.01 ^{AB}
SMFS	11.5±0.4	12.1±0.5	11.5±0.4	10.9±0.5	11.1±0.4	9.3±0.2	9.3±0.2	9.3±0.3	8.5±0.4
PRFS	58.9±2.5	55.0±3.6	51.1±2.7	55.2±3.8	54.9±2.6	52.8±1.4 ^B	52.5±1.5 ^B	51.9±2.2 ^B	60.0±2.6 ^A
SR21	90.2±1.6	89.1±1.8	91.9±1.3	92.7±1.9	90.9±1.3	91.7±0.7 ^B	91.0±0.7 ^B	91.1±1.1 ^B	94.6±1.3 ^A
PR42	78.5±2.0 ^A	77.4±2.8 ^A	82.9±2.1 ^A	74.9±3.0 ^B	75.9±2.0 ^B	79.5±1.1	79.6±1.2	78.4±1.8	81.2±2.0

¹MY = 305-day milk yield (L); FY = 305-day fat yield (kg); PY = 305-day protein yield (kg); LY = 305-day lactose yield (kg); FP = fat percentage; PP = protein percentage; LP = lactose percentage; SMFS = interval from start of mating to first service (d); PRFS = pregnancy rate to first service in the first 3 weeks of the breeding season; SR21 = 21-day submission rate; PR42 = 42-day in-calf rate. F = Holstein-Friesian; J = Jersey; FX = Holstein-Friesian cross; FJ = Holstein-Friesian × Jersey; JX = Jersey cross. Least-squares means were obtained from the mixed model and are presented as mean ± standard error (SE).

A, B, C, D, E Means with different superscripts within farm are significantly different ($p < 0.05$).

4.2.2 Parity

Least-squares means ($\pm$ SE) for milk production, composition, and fertility traits by lactation number within breed are presented in Tables 4.3 and 4.4. Lactation number significantly affected most production traits across all breeds ($p < 0.05$). Across breeds, MY, FY, PY, and LY were generally greater in later lactations. Cows in their first lactation generally produced

the lowest yields across all breeds. In contrast, first-lactation cows generally had the highest FP, PP, and LP, which generally decreased as lactation number increased across most breeds, except for Jerseys who showed no significant differences between parities for FP and PP.

For SMFS and SR21, there were no significant differences between parities across all breeds. Significant differences for PRFS between parity were only observed in Holstein-Friesian, with the highest in cows in their first, third, and fourth lactations, and PR42 tended to decline after the fifth lactation across most breeds.

Table 4.3. Least-squares means and standard errors for milk yield and composition traits by lactation number within each breed in the Massey University Dairy 1 and Dairy 4 herds.

		MY (L)	FY (kg)	PY (kg)	LY (kg)	FP (%)	PP (%)	LP (%)
F	1	3,928±36 ^D	179±1.6 ^D	149±1.3 ^D	199±1.9 ^D	4.60±0.02 ^A	4.69±0.02 ^A	5.07±0.007 ^A
	2	4,989±38 ^C	221±1.7 ^C	187±1.4 ^C	250±2 ^C	4.48±0.02 ^B	4.50±0.02 ^B	5.00±0.008 ^B
	3	5,515±42 ^B	241±1.9 ^B	206±1.5 ^B	275±2.2 ^B	4.44±0.02 ^C	4.44±0.02 ^C	4.98±0.008 ^C
	4	5,690±47 ^A	249±2.2 ^A	210±1.7 ^{AB}	282±2.5 ^A	4.45±0.02 ^C	4.45±0.03 ^C	4.96±0.009 ^{CD}
	5	5,774±54 ^A	253±2.5 ^A	211±2 ^A	287±2.8 ^A	4.44±0.03 ^C	4.44±0.03 ^C	4.96±0.01 ^{CD}
	≥6	5,780±54 ^A	250±2.5 ^A	207±1.9 ^B	286±2.8 ^A	4.38±0.03 ^D	4.38±0.03 ^D	4.95±0.01 ^D
FX	1	3,741±44 ^D	179±2 ^D	146±1.6 ^D	191±2.2 ^D	4.81±0.03 ^A	4.80±0.03 ^A	5.09±0.009 ^A
	2	4,821±45 ^C	225±2 ^C	185±1.6 ^C	242±2.3 ^C	4.74±0.03 ^B	4.74±0.03 ^B	5.01±0.009 ^B
	3	5,349±48 ^B	248±2.2 ^B	206±1.7 ^B	270±2.5 ^B	4.67±0.03 ^C	4.68±0.03 ^C	5.00±0.009 ^{BC}
	4	5,469±52 ^A	251±2.4 ^B	208±1.9 ^{AB}	273±2.7 ^{AB}	4.64±0.03 ^C	4.64±0.03 ^C	4.98±0.01 ^{CD}
	5	5,576±60 ^A	257±2.7 ^A	211±2.2 ^A	278±3.1 ^A	4.67±0.03 ^C	4.67±0.03 ^C	4.97±0.01 ^D
	≥6	5,592±58 ^A	253±2.7 ^A	208±2.1 ^{AB}	278±3 ^A	4.54±0.03 ^D	4.55±0.03 ^D	4.97±0.01 ^D
FJ	1	3,424±57 ^D	174±2.6 ^D	134±2 ^E	174±3 ^D	5.14±0.03 ^A	5.15±0.03 ^A	5.08±0.01 ^A
	2	4,585±61 ^C	228±2.8 ^C	191±2.2 ^D	232±3.2 ^C	5.01±0.03 ^B	5.01±0.03 ^B	5.05±0.01 ^B
	3	4,840±61 ^B	240±2.8 ^B	190±2.3 ^C	244±3.2 ^B	4.96±0.03 ^{BC}	4.96±0.03 ^{BC}	5.03±0.01 ^{BC}
	4	5,279±64 ^A	261±3 ^A	208±2.4 ^A	265±3.3 ^A	4.94±0.04 ^C	4.95±0.03 ^C	5.00±0.01 ^C
	5	5,309±66 ^A	259±3 ^A	208±2.4 ^A	266±3.4 ^A	4.89±0.04 ^D	4.89±0.03 ^D	5.00±0.01 ^C
	≥6	5,223±59 ^A	256±2.7 ^A	199±2.1 ^B	261±3 ^A	4.91±0.04 ^C	4.91±0.03 ^D	4.99±0.01 ^C
JX	1	3,239±68 ^E	174±3.1 ^D	131±2.4 ^D	164±3.5 ^E	5.44±0.04 ^A	5.44±0.04 ^A	5.06±0.01 ^A
	2	4,129±72 ^D	218±3.3 ^C	170±2.6 ^C	208±3.7 ^D	5.38±0.04 ^B	5.38±0.04 ^B	5.03±0.01 ^B
	3	4,621±78 ^C	241±3.6 ^B	188±2.9 ^B	233±4.1 ^C	5.34±0.04 ^B	5.34±0.04 ^{BC}	5.04±0.02 ^{AB}
	4	4,802±81 ^{BC}	254±3.7 ^A	196±2.9 ^A	242±4.2 ^{BC}	5.33±0.04 ^B	5.33±0.04 ^{BC}	5.03±0.02 ^{AB}
	5	5,029±86 ^A	262±4 ^A	202±3.1 ^A	252±4.4 ^A	5.29±0.04 ^B	5.29±0.04 ^C	5.01±0.01 ^B
	≥6	4,955±80 ^{AB}	260±3.7 ^A	194±2.9 ^A	250±4.2 ^{AB}	5.34±0.05 ^B	5.34±0.05 ^{BC}	5.04±0.01 ^{AB}
J	1	3,064±77 ^C	167±3.5 ^C	124±2.8 ^C	157±4 ^D	5.70±0.05	5.70±0.05	5.07±0.01 ^A
	2	3,924±79 ^B	217±3.6 ^B	162±2.9 ^B	199±4.1 ^C	5.69±0.05	5.69±0.04	5.06±0.02 ^{AB}
	3	4,340±87 ^A	240±4.0 ^A	179±3.2 ^A	220±4.5 ^{AB}	5.70±0.05	5.70±0.05	5.05±0.02 ^{ABC}
	4	4,440±101 ^A	245±4.7 ^A	183±3.7 ^A	224±5.2 ^A	5.67±0.05	5.67±0.05	5.03±0.02 ^{BC}
	5	4,185±115 ^A	235±5.3 ^A	172±4.2 ^A	210±6 ^B	5.69±0.06	5.69±0.06	5.01±0.02 ^C
	≥6	4,314±114 ^A	238±5.2 ^A	177±4.1 ^A	217±5.9 ^{AB}	5.62±0.06	5.62±0.06	5.01±0.02 ^C

MY = milk yield; FY = fat yield; PY = protein yield; LY = lactose yield; FP = fat percentage; PP = protein percentage; LP = lactose percentage. F = Holstein-Friesian; FX = Holstein-Friesian cross; FJ = Holstein-Friesian × Jersey; JX = Jersey cross; J = Jersey. Least-squares means are presented as mean ± standard error (SE) from the mixed model.

^{A, B, C, D, E} Means with different superscripts among lactation numbers are significantly different ($p < 0.05$).

Lactation ≥6 includes lactations six and above.

Table 4.4. Least-squares means and standard errors for fertility traits by lactation number within each breed in the Massey University Dairy 1 and Dairy 4 herds.

		SMFS (days)	PRFS (%)	SR21 (%)	PR42 (%)
F	1	10.2±0.35	61.5±2.4 ^A	92.4±1.2	85.3±1.9 ^A
	2	9.8±0.37	53.7±2.5 ^B	92.5±1.3	80.5±2.0 ^{AB}
	3	10.5±0.42	57.9±2.8 ^{AB}	89.6±1.4	78.1±2.3 ^{BC}
	4	10.7±0.47	54.1±3.2 ^A	87.9±1.6	81.4±2.6 ^{AB}
	5	10.3±0.55	49.7±3.7 ^B	92.7±1.9	76.0±3.1 ^{BC}
	≥6	11.2±0.45	53.5±3.1 ^B	91.1±1.6	71.9±2.5 ^C
FX	1	9.9±0.42	55.8±2.8	92.8±1.5	82.4±2.3 ^A
	2	11.0±0.44	56.7±2.9	89.6±1.5	82.8±2.4 ^A
	3	10.1±0.47	52.6±3.2	90.8±1.6	81.2±2.6 ^A
	4	10.7±0.50	57.7±3.4	90.4±1.7	80.4±2.8 ^A
	5	10.8±0.61	53.3±4.1	88.9±2.1	75.9±3.3 ^{AB}
	≥6	10.9±0.50	52.0±3.4	88.7±1.7	71.0±2.7 ^B
FJ	1	10.1±0.56	53.3±3.8	90.7±1.9	75.3±3.1 ^{BC}
	2	10.4±0.60	61.2±4.1	90.9±2.1	84.2±3.3 ^{AB}
	3	10.2±0.61	62.7±4.1	93.5±2.1	82.4±3.4 ^{AB}
	4	11.5±0.66	64.7±4.5	90.5±2.3	88.0±3.7 ^A
	5	10.2±0.66	55.0±4.5	93.5±2.3	81.5±3.7 ^{ABC}
	≥6	10.4±0.49	45.6±3.4	89.8±1.7	72.6±2.7 ^C
JX	1	9.1±0.66	63.9±4.4	94.7±2.3	82.0±3.6
	2	9.6±0.71	63.8±4.7	95.8±2.4	83.5±3.9
	3	9.4±0.78	59.5±5.3	94.9±2.7	82.1±4.3
	4	9.9±0.87	58.6±5.8	93.0±2.9	72.9±4.8
	5	9.7±0.87	53.6±5.8	95.6±2.9	76.8±4.8
	≥6	10.5±0.69	53.1±4.7	89.7±2.4	74.2±3.8
J	1	9.3±0.73	59.0±4.9	92.3±2.5	79.9±4.0 ^A
	2	8.8±0.76	58.8±5.1	94.6±2.6	84.0±4.2 ^A
	3	10.9±0.86	49.0±5.8	87.4±2.9	79.1±4.7 ^A
	4	12.1±1.01	42.3±6.8	87.7±3.5	68.9±5.6 ^{AB}
	5	8.1±1.17	49.6±7.9	94.1±4.0	74.5±6.5 ^{AB}
	≥6	10.2±1.05	52.1±7.1	91.8±3.6	64.2±5.8 ^B

SMFS = start of mating to first service (d); PRFS = pregnancy rate to first service in the first 3 weeks of the breeding season (%); SR21 = 21-day submission rate (%); PR42 = 42-day in-calf rate (%). F = Holstein-Friesian; FX = Holstein-Friesian cross; FJ = Holstein-Friesian × Jersey; JX = Jersey cross; J = Jersey. Least-squares means are presented as mean ± standard error (SE) from the mixed model.

^{A, B, C} Means with different superscripts within a breed and trait among lactation numbers are significantly different ($p < 0.05$).

Lactation ≥6 includes lactations six and above.

4.3 Variance components and heritability estimates

Estimates of variance components and heritability from the pedigree-based additive model, genomic-based additive-only model and the additive + dominance model are summarised in Tables 4.5–4.7.

Under the additive pedigree-based model, heritability estimates for yield traits were moderate, ranging from 0.182 (FY) to 0.215 (MY), and milk composition traits showed high heritability, ranging from 0.649 (FP) to 0.687 (PP). Heritability of LP was lower at 0.269. For fertility traits, the additive pedigree model showed low heritabilities of 0.00, 0.054, 0.017, and 0.063 for SMFS, PRFS, SR21, and PR42 respectively.

In comparison, heritability estimates for milk traits were higher in the additive genomic model ranging from 0.222 (PY) to 0.284 (LY) for yield traits, and 0.454 (LP) to 0.693 (FP) for milk composition traits. However, fertility traits exhibited lower heritability, with 0.021 for PRFS and close to zero for SMFS, SR21, and PR42. Permanent environmental variances were generally higher in the pedigree model compared to the genomic models. However, total phenotypic variance and residual variance were similar between pedigree and genomic models for all traits.

Inclusion of dominance variance had minimal effect on the additive and residual variance estimates for all traits. Additive heritability estimates were similar to those obtained under the additive genomic model, ranging from 0.216 to 0.282 for yield traits, 0.453 to 0.693 for composition traits, and remained low for fertility. Dominance variance contributed to yield traits, accounting for between 8.4% (MY) and 12.1% (FY) of total phenotypic variance. For these traits, the permanent environmental variance decreased when dominance variance was accounted for, but total phenotypic variance was similar between additive and additive + dominance models.

For composition traits, dominance was close to zero, with only LP showing a small amount of dominance variance contributing to phenotypic variance (4.4%). Permanent environmental variance for LP decreased from 0.002 to 0.0006 when dominance was included. For fertility traits, dominance variance was negligible, and estimates were associated with relatively large standard errors (see Table 4.7).

Table 4.5. Variance components for milk production and fertility traits estimated with ASReml using pedigree data.

Trait ¹	σ^2_a	σ^2_c	σ^2_e	σ^2_p	h^2
MY	130100 (24640)	118946 (22962)	355103 (7867)	604149 (14249)	0.215 (0.039)
FY	231 (49.2)	251 (46.8)	786 (17.4)	1267 (29.1)	0.182 (0.038)
PY	143 (28.3)	119 (26.6)	509 (11.3)	772 (17.2)	0.186 (0.035)
LY	345 (66.5)	323 (62.0)	952 (21.1)	1620 (38.2)	0.213 (0.039)
FP	0.168 (0.017)	0.045 (0.013)	0.046 (0.001)	0.260 (0.009)	0.649 (0.055)
PP	0.040 (0.004)	0.008 (0.003)	0.011 (0.000)	0.059 (0.002)	0.687 (0.053)
LP	0.007 (0.001)	0.006 (0.001)	0.013 (0.001)	0.025 (0.001)	0.269 (0.050)
SMFS	0	1.166 (0.598)	51.11 (1.121)	52.27 (0.998)	0.000 (0.000)
PRFS	0.013 (0.003)	0	0.226 (0.005)	0.239 (0.005)	0.054 (0.012)
SR21	0.001 (0.001)	0.001 (0.001)	0.060 (0.001)	0.062 (0.001)	0.017 (0.017)
PR42	0.010 (0.002)	0	0.150 (0.003)	0.160 (0.003)	0.063 (0.012)

¹MY = 305-day milk yield (L); FY = 305-day fat yield (kg); PY = 305-day protein yield (kg); LY = 305-day lactose yield (kg); FP = fat percentage; PP = protein percentage; LP = lactose percentage; SMFS = start of mating to first service (d); PRFS = cow conceived to first service in the first 3 weeks of the breeding season (%); SR21 = cow inseminated in the first 3 weeks from the start of mating (%); PR42 = in-calf rate at 42 days from the start of mating (%).

σ^2_a = additive variance; σ^2_c = permanent environmental variance; σ^2_e = residual variance; σ^2_p = phenotypic variance; h^2 = heritability. Standard errors (SE) in parentheses.

Table 4.6. Variance components for milk production and fertility traits estimated using an additive genetic model fitted with GREML in Julia, based on a genomic relationship matrix.

Trait ¹	σ^2_a	σ^2_c	σ^2_e	σ^2_p	h^2
MY	163102 (19542)	73309 (13222)	354173 (7816)	590585 (15374)	0.276 (0.029)
FY	297 (40.7)	212 (30.3)	782 (17.3)	1290 (32.6)	0.230 (0.028)
PY	173 (23.3)	98 (17.1)	506 (11.2)	777 (18.9)	0.222 (0.027)
LY	455 (53.7)	198 (35.9)	949 (20.9)	1601 (42.1)	0.284 (0.029)
FP	0.174 (0.013)	0.031 (0.006)	0.046 (0.001)	0.252 (0.010)	0.693 (0.029)
PP	0.039 (0.003)	0.007 (0.001)	0.010 (0.000)	0.057 (0.002)	0.691 (0.029)
LP	0.013 (0.001)	0.002 (0.000)	0.013 (0.000)	0.027 (0.001)	0.454 (0.029)
SMFS	0.366 (0.447)	0.772 (0.702)	50.98 (1.114)	52.12 (0.996)	0.007 (0.008)
PRFS	0.005 (0.003)	0.006 (0.003)	0.227 (0.005)	0.238 (0.005)	0.021 (0.011)
SR21	0	0.002 (0.001)	0.060 (0.001)	0.062 (0.001)	0
PR42	0	0.005 (0.002)	0.155 (0.003)	0.160 (0.003)	0

¹MY = 305-day milk yield (L); FY = 305-day fat yield (kg); PY = 305-day protein yield (kg); LY = 305-day lactose yield (kg); FP = fat percentage; PP = protein percentage; LP = lactose percentage; SMFS = start of mating to first service (d); PRFS = cow conceived to first service in the first 3 weeks of the breeding season (%); SR21 = cow inseminated in the first 3 weeks from the start of mating (%); PR42 = in-calf rate at 42 days from the start of mating (%).

σ^2_a = additive variance; σ^2_c = permanent environmental variance; σ^2_e = residual variance; σ^2_p = phenotypic variance; h^2 = heritability. Standard errors (SE) in parentheses.

Table 4.7. Variance components for milk production and fertility traits estimated using an additive + dominance genetic model fitted with GREML in Julia, based on a genomic relationship matrix.

Trait ¹	σ^2_a	σ^2_d	σ^2_c	σ^2_e	σ^2_p	h^2_a	h^2_d
MY	160877 (19394)	49707 (30522)	24241 (31754)	354256 (7817)	589081 (15303)	0.273 (0.028)	0.084 (0.052)
FY	294 (40.3)	155 (70.3)	53 (72.1)	782 (17.3)	1285 (32.4)	0.229 (0.028)	0.121 (0.054)
PY	167 (22.9)	82 (40.0)	18 (41.0)	506 (11.1)	773 (18.8)	0.216 (0.027)	0.106 (0.051)
LY	451 (53.4)	141 (83.1)	58 (86.2)	949 (21.0)	1598 (41.9)	0.282 (0.029)	0.088 (0.052)
FP	0.174 (0.013)	0	0.031 (0.014)	0.046 (0.001)	0.252 (0.010)	0.693 (0.029)	0
PP	0.039 (0.003)	0	0.007 (0.003)	0.010 (0.000)	0.057 (0.002)	0.691 (0.029)	0
LP	0.012 (0.001)	0.001 (0.001)	0.0006 (0.001)	0.013 (0.000)	0.027 (0.000)	0.453 (0.029)	0.044 (0.047)
SMFS	0.366 (0.447)	0	0.772 (1.731)	51 (1.114)	52 (0.996)	0.007 (0.008)	0
PRFS	0.005 (0.003)	0	0.006 (0.008)	0.227 (0.005)	0.238 (0.005)	0.021 (0.011)	0
SR21	0	0	0.001 (0.002)	0.060 (0.001)	0.061 (0.001)	0	0
PR42	0	0.0015 (0.005)	0.003 (0.005)	0.155 (0.003)	0.160 (0.003)	0	0.009 (0.032)

¹ MY = 305-day milk yield (L); FY = 305-day fat yield (kg); PY = 305-day protein yield (kg); LY = 305-day lactose yield (kg); FP = fat percentage; PP = protein percentage; LP = lactose percentage; SMFS = start of mating to first service (d); PRFS = cow conceived to first service in the first 3 weeks of the breeding season (%); SR21 = cow inseminated in the first 3 weeks from the start of mating (%); PR42 = in-calf rate at 42 days from the start of mating (%).

σ^2_a = additive variance; σ^2_d = dominance variance; σ^2_c = permanent environmental variance; σ^2_e = residual variance; σ^2_p = phenotypic variance; h^2_a = additive heritability; h^2_d = proportion of phenotypic variance explained by dominance variance. Standard errors (SE) in parentheses.

4.4 Genome-wide association studies

Figure 4.1 presents Manhattan plots for genome-wide association analyses investigating additive and dominance effects on MY, FY, PY, and LY. All Manhattan plots for additive effects showed a significant peak on chromosome 14 for milk yield traits. Additionally, a peak on chromosome 5 was observed for MY and LY. For dominance effects, most SNP markers were distributed below the genome-wide significance threshold; however, a small number of SNPs exceeded this threshold across the yield traits, with a clear peak on chromosome 9 for FY, and a cluster of markers above the significance threshold on chromosome 15 for PY.

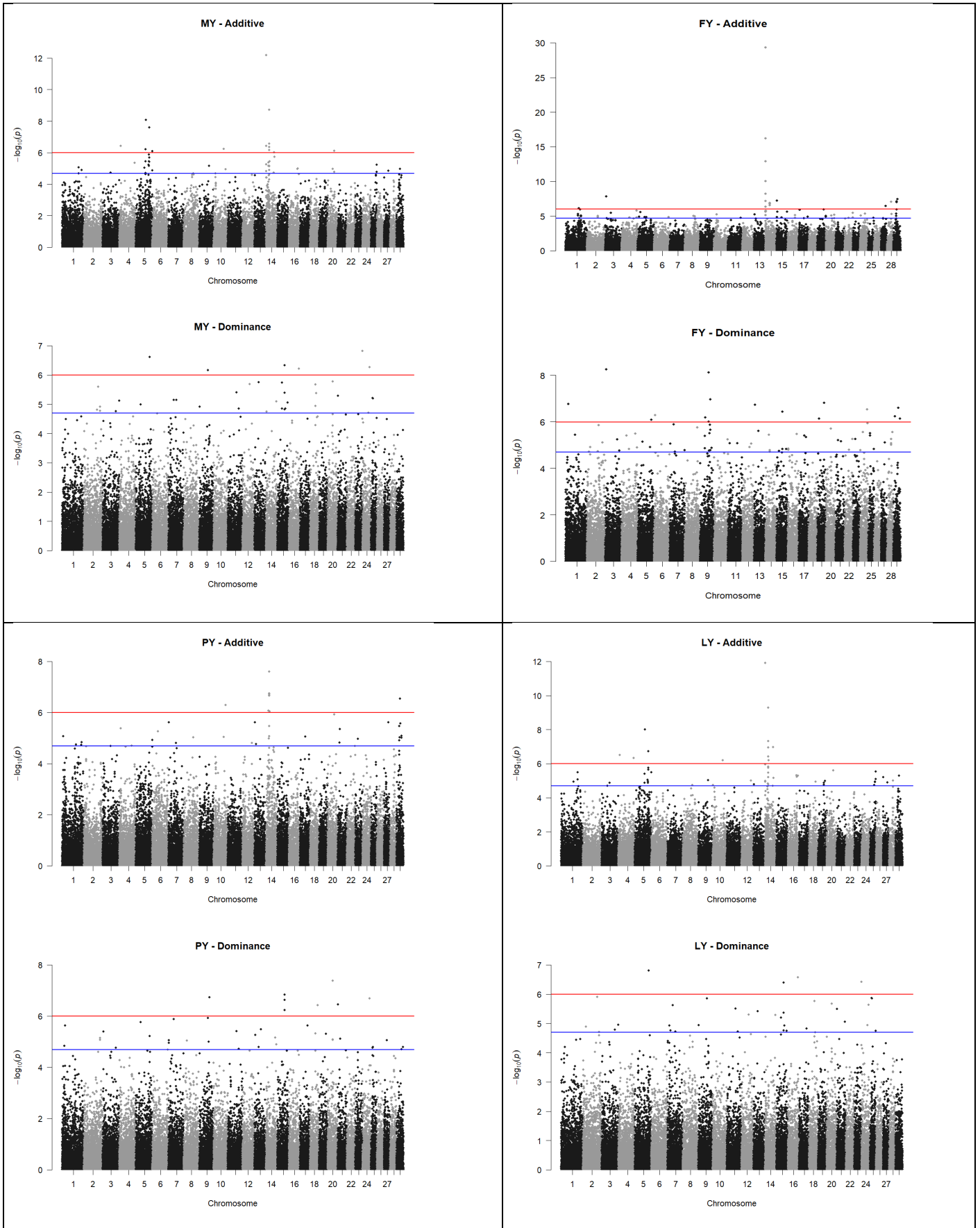


Figure 4.1. Manhattan plots based on genome-wide association analyses results testing additive and dominance effects for milk yield (MY), fat yield (FY), protein yield (PY), and lactose yield (LY). The

genome-wide significance threshold is represented by the horizontal red line at $-\log_{10}(p) = 5.99$ ($p = 1.02 \times 10^{-6}$), and the suggestive significance threshold is represented by the horizontal blue line at $-\log_{10}(p) = 4.69$ ($p = 2.04 \times 10^{-5}$).

Figure 4.2 presents Manhattan plots for the genome-wide association analyses investigating additive and dominance effects on FP, PP, and LP. For additive effects, FP and PP had significant peaks on chromosomes 5, 14, and 19. In contrast, most SNP markers for LP sat below the genome-wide significance threshold ($p = 1.02 \times 10^{-6}$), but small peaks were found on chromosome 19 and chromosome 6 that exceeded the suggestive significance threshold ($p = 2.04 \times 10^{-5}$). For the association analyses testing for dominance effects, the majority of SNPs were not found to have significant associations with the traits investigated ($p > 2.04 \times 10^{-5}$). However, a small number of markers slightly exceeded the threshold for FP, PP, and LP.

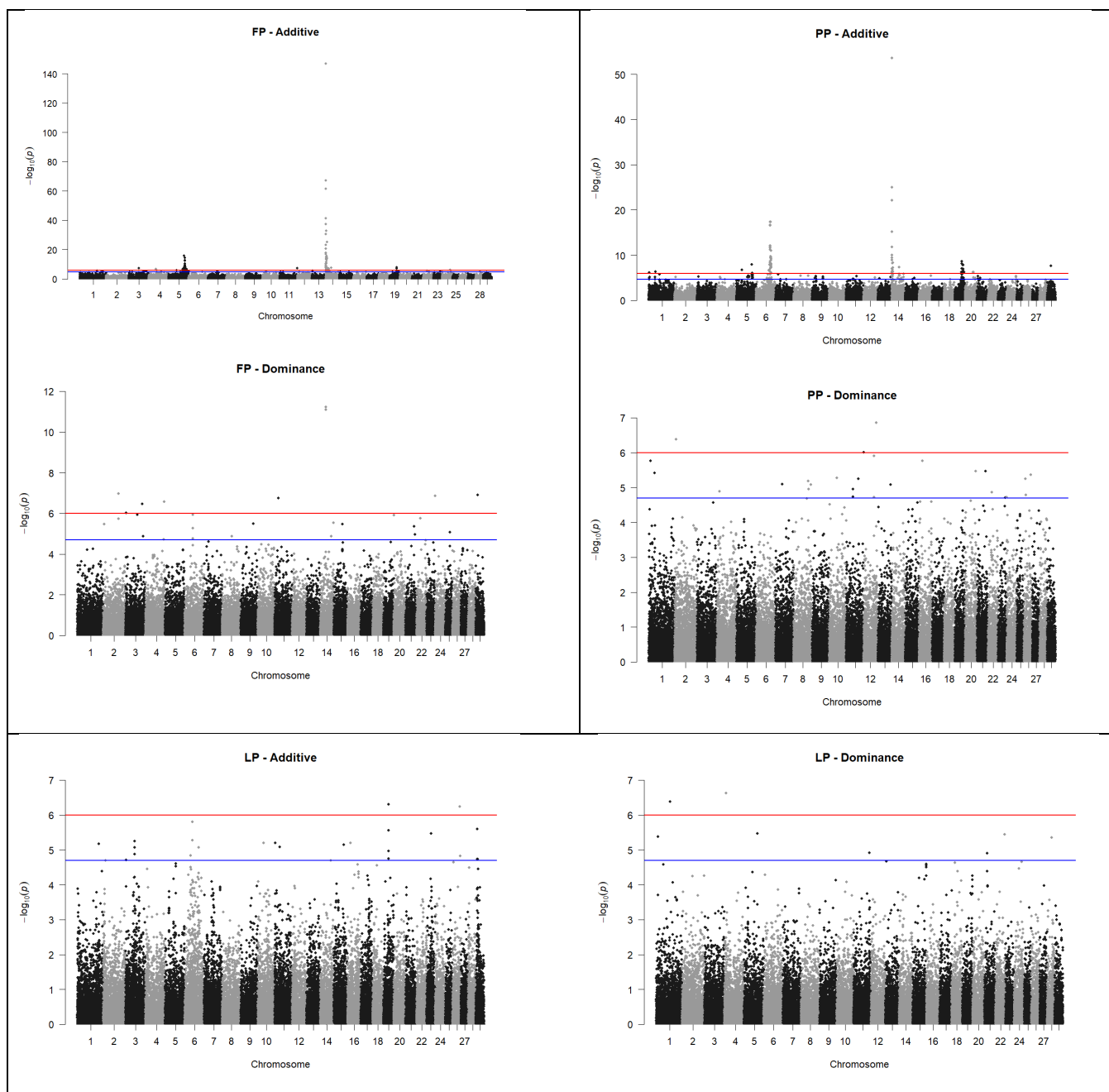


Figure 4.2. Manhattan plots based on genome-wide association analyses results testing additive and dominance effects markers for fat percentage (FP), protein percentage (PP), and lactose percentage (LP). The genome-wide significance threshold is represented by the horizontal red line at $-\log_{10}(p) = 5.99$ ($p = 1.02 \times 10^{-6}$), and the suggestive significance threshold is represented by the horizontal blue line at $-\log_{10}(p) = 4.69$ ($p = 2.04 \times 10^{-5}$).

Figure 4.3 presents Manhattan plots for genome-wide association analyses investigating additive and dominance effects on fertility traits. For both additive and dominance effects, most SNPs were not significantly associated with the fertility traits investigated ($p > 2.04 \times 10^{-5}$). However, for additive effects, a small number of SNPs were located close to, or slightly above, the genome-wide significance threshold, and some association peaks were visible above the suggestive significance threshold. This can be seen for SR21, where an association signal is visible on chromosome 7, and for PR42, where a peak is visible on chromosome 5. For dominance effects, a few SNPs approached the suggestive threshold across the fertility traits, however only one SNP exceeded the genome-wide significance level across all fertility traits, which was observed for SR21.

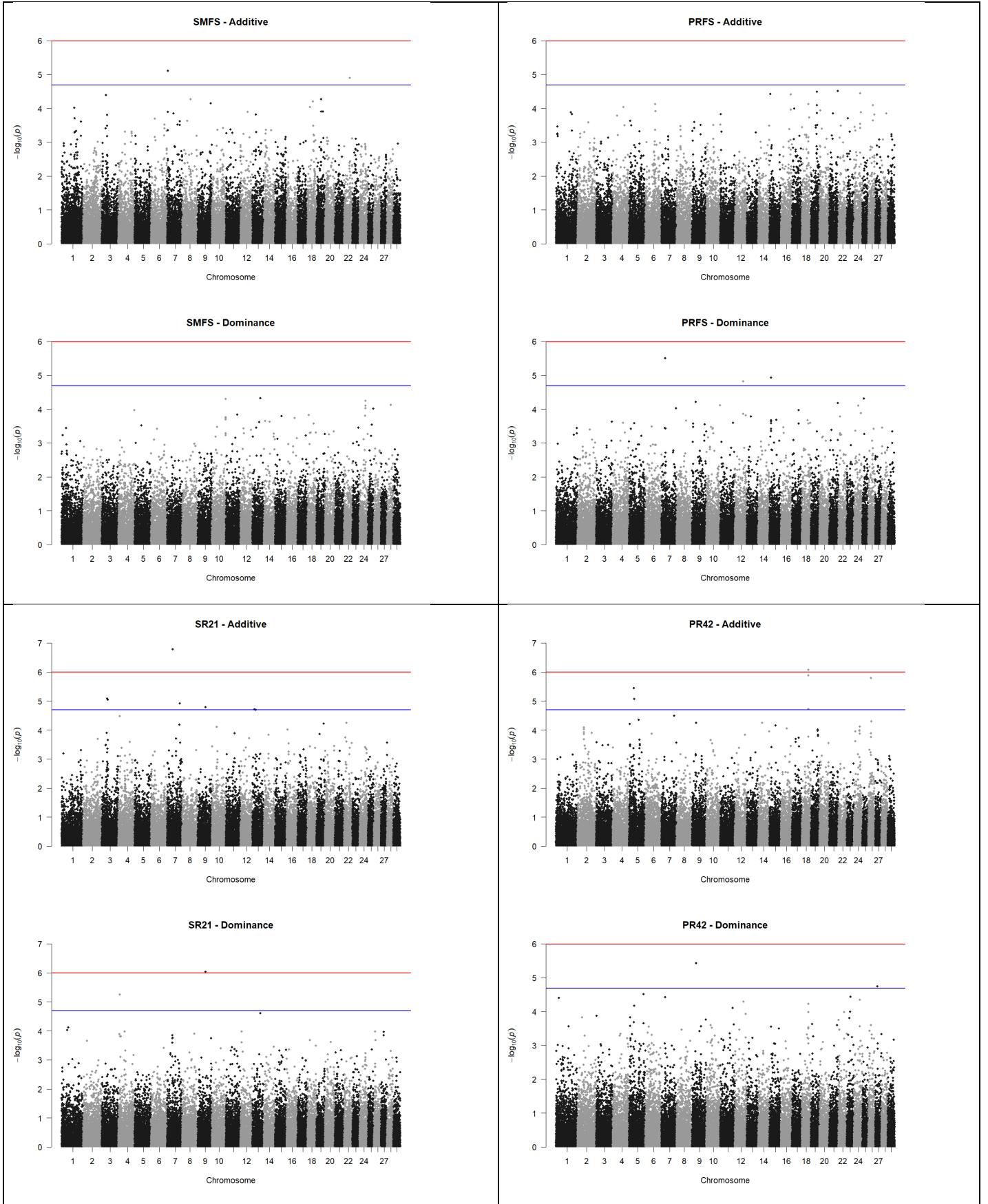


Figure 4.3. Manhattan plots based on genome-wide association analyses results testing additive and dominance effects markers for the interval between start of mating to first service (SMFS), pregnancy

rate to first service in the first 3 weeks of the breeding season (PRFS), 21-day submission rate (SR21), and 42-day in-calf rate (PR42). The genome-wide significance threshold is represented by the horizontal red line at $-\log_{10}(p) = 5.99$ ($p = 1.02 \times 10^{-6}$), and the suggestive significance threshold is represented by the horizontal blue line at $-\log_{10}(p) = 4.69$ ($p = 2.04 \times 10^{-5}$).

The most significant SNPs identified from genome-wide association analyses investigating dominance effects, and their corresponding additive effects, across milk production, composition, and fertility traits, are presented in Table 4.8. Across all traits, SNPs identified as significant for dominance effects generally showed stronger dominance association signals than their corresponding additive association signals. Only seven of these SNPs had additive association signals that exceeded the suggestive significance threshold, and these additive signals were weaker than their corresponding dominance association signals. Several dominance association signals were seen across multiple traits, including MY, PY, and LY. The most significant dominance association was observed for FP on chromosome 14.

Table 4.8. Most significant single nucleotide polymorphisms for dominance effects identified using a model incorporating additive and dominance effects across milk yield, composition, and fertility traits.

Trait	SNP Location		Additive SNP effect		Dominance SNP effect	
	Chromosome	Position (base pairs)	a	-Log ₁₀ (P-value)	d	-Log ₁₀ (P-value)
Yield traits						
Milk yield	24	163270	-90.21	2.26	183.68	6.81
	5	100079902	42.56	1.45	119.22	6.60
	15	53343785	-50.18	1.04	161.04	6.33
	24	51733234	5.38	0.09	122.29	6.26
	16	71619975	-142.30	4.65	183.22	6.22
Fat yield	3	4954740	-7.34	7.83**	8.51	8.24
	3	4959543	-7.34	7.83**	8.51	8.24
	3	4967420	-7.34	7.83**	8.51	8.24
	9	63181088	-3.37	3.49	6.28	8.12
	9	74993010	-0.19	0.09	5.46	6.96
Protein yield	20	35672395	-2.67	2.51	5.28	7.38
	15	53762431	-3.54	1.94	7.89	6.83
	9	74993010	-0.48	0.32	4.25	6.73
	24	51733234	0.01	0.01	4.54	6.69
	15	53343785	-1.83	1.09	5.91	6.63
Lactose yield	5	100079902	2.15	1.38	6.32	6.80
	16	71619975	-7.98	5.29*	9.84	6.57
	24	163270	-4.28	1.93	9.28	6.42
	15	53343785	-2.63	1.04	8.45	6.39
	2	103555336	1.56	0.90	-5.74	5.91
Composition traits						
Fat percentage	14	33142286	-0.10	7.66**	0.13	11.22
	14	33150554	-0.11	7.77**	0.13	11.10
	2	90977880	-0.03	1.72	0.08	6.98
	29	11834909	-0.06	2.97	0.10	6.91
	24	163270	0.07	2.83	-0.12	6.85
Protein percentage	12	77193270	-0.02	1.95	0.04	6.86
	2	7604327	0.01	1.41	-0.04	6.38
	11	105577071	-0.01	0.41	0.04	6.00
	12	62886090	-0.03	3.21	0.04	5.90
	16	21099960	-0.01	0.72	-0.03	5.77
Lactose percentage	4	9278932	-0.01	0.03	0.03	6.63
	1	85102888	0.01	-0.03	-0.03	6.38
	5	82257181	0.00	0.02	0.02	5.47
	22	55038557	0.01	-0.03	-0.03	5.45
	1	10551823	0.01	-0.03	-0.03	5.37
Fertility traits ¹						
PRFS (%)	7	27306372	-1.11	0.48	6.77	5.50
	15	11245881	0.35	0.13	-6.01	4.93
	12	59356240	4.95	2.17	-9.30	4.82
SR21 (%)	9	58119693	-3.89	4.78*	5.22	6.04
	4	9198537	-4.07	4.47	5.27	5.24
PR42 (%)	9	28985157	-7.70	4.24	10.16	5.42
	9	28991358	-7.70	4.24	10.16	5.42
	27	14522048	0.29	0.14	4.63	4.74

d = dominance effect; a = additive effect; ** indicates additive effects exceeding genome-wide significance threshold $-\log_{10}(p) = 5.99$, * indicates additive effects exceeding suggestive-significance threshold $-\log_{10}(p) = 4.69$.

¹ PRFS = pregnancy rate to first service in the first 3 weeks of the breeding season; SR21 = 21-day submission rate; PR42 = 42-day in-calf rate.

4.5 Genotype effects of top SNPs

The SNPs with the strongest dominance association signal for each trait are presented in Figures 4.5-4.7. Genotypes were fitted as a 3-level class effect, and mean phenotypes were adjusted for fixed effects in the model described in section 3.8.

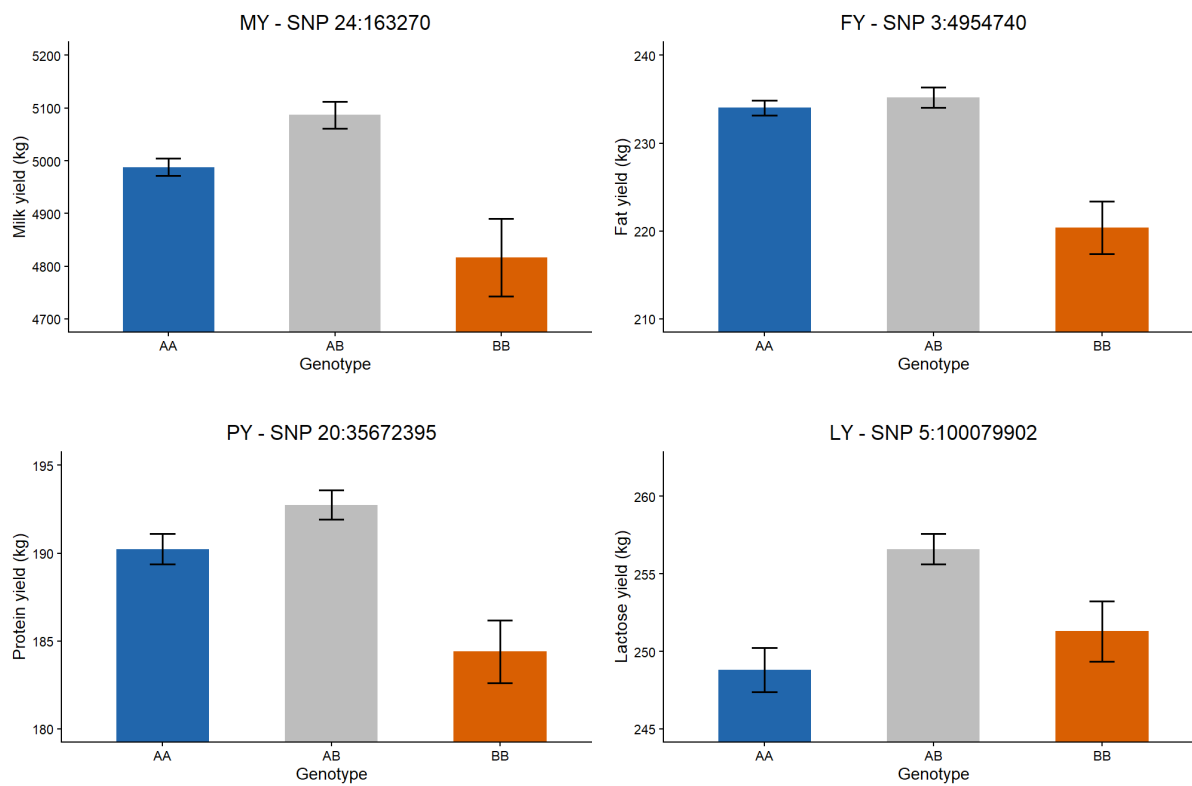


Figure 4.4. Adjusted mean 305-day milk yield (MY;L), fat yield (FY;kg), protein yield (PY;kg), and lactose yield (LY;kg), with corresponding standard errors, for genotypes (AA, AB, BB) at the most significant dominance-associated SNP for each trait. Values are corrected for fixed effects included in the statistical model. Sample sizes for each genotype group were: Milk yield (AA = 1,097, AB = 509, BB = 61); fat yield (AA = 1,032, AB = 545, BB = 90); protein yield (AA = 808, AB = 626, BB = 233); and lactose yield (AA = 567, AB = 796, BB = 304).

The adjusted phenotypic means for the genotype classes at the SNPs with the strongest dominance association signal for each yield trait are presented in Figure 4.4. For MY,

animals with the AB genotype showed the highest adjusted mean MY ($5,087 \pm 25$ L), followed by AA ($4,988 \pm 17$ L), and BB ($4,816 \pm 73$ L). Similarly, for FY, AB animals had the highest mean (235 ± 1.15 kg), followed by AA (234 ± 0.84 kg), and BB (220 ± 3.0 kg). For PY, AB animals again showed the highest mean (193 ± 0.83 kg), followed by AA (190 ± 0.83 kg), and BB (184 ± 1.79 kg). For LY, AB animals had the highest mean (257 ± 0.98 kg), followed by BB (251 ± 1.95 kg), while AA animals had the lowest mean LY (249 ± 1.43 kg).

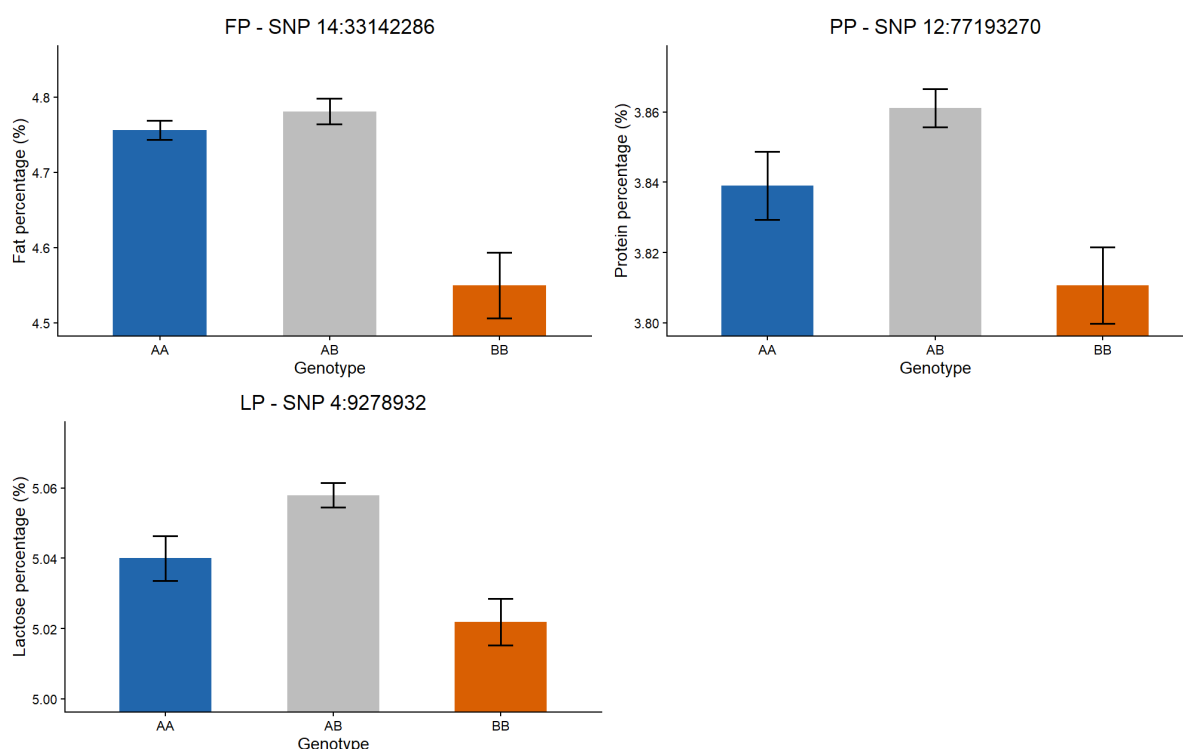


Figure 4.5. Adjusted mean fat (FP), protein (PP), and lactose (LP) percentages and corresponding standard errors for genotypes (AA, AB, BB) at the most significant dominance-associated SNP for each trait. Values are corrected for fixed effects included in the statistical model. Sample sizes for each genotype group were: Fat percentage (AA = 992, AB = 570, BB = 105); protein percentage (AA = 484, AB = 802, BB = 381); and lactose percentage (AA = 397, AB = 879, BB = 394).

The adjusted phenotypic means for the genotype classes at the SNPs with the strongest dominance association signal milk for composition traits are presented in Figure 4.5. For FP, animals with the AB genotype showed the highest mean ($4.78 \pm 0.17\%$), followed by AA ($4.76 \pm 0.13\%$), and BB ($4.55 \pm 0.44\%$). A similar pattern was observed for PP, where AB animals had the highest mean ($3.86 \pm 0.005\%$), followed by AA ($3.85 \pm 0.009\%$), and BB ($3.81 \pm 0.011\%$). For LP, AB animals again showed the highest mean ($5.06 \pm 0.003\%$), followed by AA ($5.04 \pm 0.006\%$), and BB ($5.02 \pm 0.007\%$).

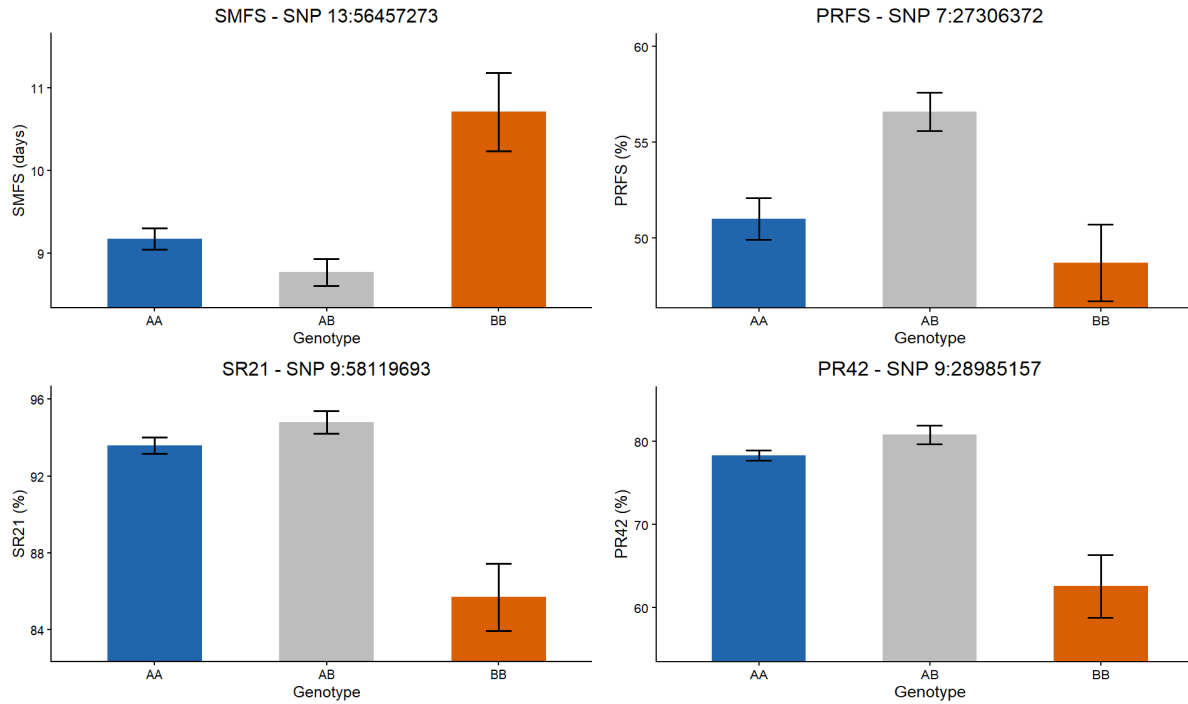


Figure 4.6. Mean interval from the start of mating date to the first service (SMFS), pregnancy rate to first service in the first 3 weeks of the breeding season (PRFS), 21-day submission rate (SR21), and 42-day in-calf rate (PR42), and corresponding standard errors for genotypes (AA, AB, BB) at the most significant dominance-associated SNP for each trait. Values are corrected for fixed effects included in the statistical model. Sample sizes for each genotype group were: SMFS (AA = 1,023, AB = 572, BB = 64); PRFS (AA = 705, AB = 729, BB = 226); SR21 (AA = 1,047, AB = 544, BB = 69); and PR42 (AA = 1,234, AB = 384, BB = 42).

The adjusted phenotypic means for the genotype classes at the SNPs with the strongest dominance association signal for fertility traits are presented in Figure 4.6. For SMFS, animals with the AB genotype had the shortest mean interval (8.8 ± 0.16 days), followed by AA (9.2 ± 0.13 days), and BB (10.7 ± 10.7 days). For PRFS, AB animals showed the highest mean ($56.6 \pm 1\%$), followed by AA ($51.0 \pm 1\%$), and BB ($48.7 \pm 2\%$). A similar pattern was observed for SR21, where AB animals had the highest mean ($94.3 \pm 0.6\%$), followed by AA ($93.0 \pm 0.4\%$), and BB ($85.1 \pm 1.8\%$). For PR42, AB animals again showed the highest mean ($80.8 \pm 1.1\%$), followed by AA ($78.3 \pm 0.6\%$), and BB ($62.6 \pm 4.0\%$).

Chapter 5

Discussion

5.1 Descriptive statistics and fixed effects

This study investigated phenotypic variation, genetic parameters, and genome-wide associations for milk production and fertility traits in 1,667 dairy cows from Massey University Dairy 1 and Dairy 4 herds.

Overall, cows in this population showed higher averages for milk production and composition traits, as well as higher 42-day in-calf (PR42) and 21-day submission rates (SR21), compared with national averages reported for New Zealand dairy herds in 2014–2022 (DairyNZ and Livestock Improvement Corporation 2022). These results are consistent with those reported in previous analyses of the same population of Massey University Dairy 1 and Dairy 4 cows (Jayawardana et al. 2023b).

Breed had a significant effect on milk production traits, with Holstein-Friesian (F) cows producing the highest milk yield (MY), protein yield (PY), and lactose yield (LY), while Holstein-Friesian $\times$ Jersey cross (FJ) cows produced the highest fat yield (FY). In contrast, Jersey (J) and Jersey cross (JX) cows had the lowest yields. For milk composition, J cows had the highest fat percentage (FP), protein percentage (PP), and lactose percentage (LP) compared with other breeds. These findings are consistent with established breed trends in New Zealand dairy cattle (DairyNZ and Livestock Improvement Corporation 2022), where Holstein-Friesians have been selected to have higher milk yield while Jersey cattle generally produce milk with higher milk solids concentrations (Lembeye et al. 2016). In contrast, fertility traits showed limited and inconsistent differences between breeds in this study. Breed did not have a significant effect on several traits and there was variation between herds, which suggests that fertility is not strongly influenced by breed within this population.

Within breed, parity had a significant effect on milk production traits, with yields generally increasing with lactation number and reaching their peak after the fourth lactation. This is consistent with previous studies showing that mean yield increases with parity up to parity 5 (Berry and Downing 2025). This increase in yield may be due to physiological maturity, increased body size, and development of the mammary gland in older cows (Ridhowi et al. 2025). Cows in their first lactation had the highest FP, PP, and LP, which declined with increasing parity across most breeds. This may be because first-lactation cows produce lower MY than multiparous cows, resulting in a higher concentration of milk solids (Ridhowi et al. 2025).

Within most breeds, fertility traits showed little variation across parities, indicating that lactation number had limited influence on reproduction in this population. However, PR42 was significantly influenced by parity across most breeds, with fertility generally decreasing with increasing parity. This is consistent with previous studies, where lower pregnancy rates have been found in cows with higher parity compared to lower-parity cows and may be caused by changes in follicular development, hormone concentrations, and uterine environment in multiparous cows (Balendran et al. 2008).

5.2 Genetic variance estimates

5.2.1 Additive genetic variance

5.2.1.1 Milk production and composition traits

In this population, estimates of genetic parameters obtained using the pedigree-based model showed moderate heritability (0.18–0.22) in yield traits, high heritability (0.65–0.69) in FP and PP, moderate heritability (0.27) in LP, and low heritability (0.00–0.06) in fertility traits. This is consistent with the findings of Jayawardana et al. (2023b) and Lembeye et al. (2016) who reported similar heritabilities for these traits for twice-daily-milked dairy cattle in New Zealand.

Compared with the pedigree-based model, the additive genomic model produced higher heritability estimates and lower standard errors for milk yield and composition traits. The lower standard errors associated with the genomic estimates suggest that using the genomic relationship matrix (GRM) provided more precise estimates. This is consistent with other studies that found GRM-based heritability estimates were more accurate than those based on pedigree relationship matrices (Lee et al. 2010; Veerkamp et al. 2011; Haile-Mariam et al. 2013; Manzanilla-Pech et al. 2023). However, the difference between genomic and pedigree-based estimates has not been consistent across studies. For example, Veerkamp et al. (2011) reported lower genetic variance estimates for milk yield, body weight, and dry matter intake based on genomic relationships, and Haile-Mariam et al. (2013) obtained similar findings for milk production, fertility, survival, temperament, and stature traits, when compared to pedigree-based estimates. In contrast, Manzanilla-Pech et al. (2023) reported slightly higher estimates when using a GRM to estimate heritability for body weight. These differences may be due to different SNP density, base populations, and reference population sizes, which can affect the amount of genetic variance accounted for by the GRM (Veerkamp et al. 2011; Haile-Mariam et al. 2013). However, the lower standard errors indicated that genomic-based

estimates were more precise than the corresponding pedigree-based estimates across these studies. Lee et al. (2010) showed that using a GRM can reduce confounding of environmental effects with pedigree relationships and therefore can produce more accurate estimates of genetic variance components. This is likely because GRMs use genotypic information to account for genetic relationships more precisely than expected pedigree relationships (Veerkamp et al. 2011).

Heritability estimates from the genomic models in the present study were broadly consistent with previous genomic studies in dairy cattle (Table 2.1). For example, estimates for milk traits were consistent with those of Reynolds et al. (2022) who reported heritability estimates using a GRM-based model of 0.23–0.31 for yield traits, and 0.62–0.70 for FP and PP in New Zealand Holstein-Friesian and Jersey cows. Estimates for MY, FY, and PY were similar to those reported by Sun et al. (2014) in USA Holstein and Jersey, but were higher than their estimates for FP and PP.

Differences in heritability estimates between populations may be due to variation in environments, selection, and inbreeding levels (Visscher et al. 2008). New Zealand dairy cattle tend to have lower inbreeding levels than many international countries, with Porteous (2022) reporting an inbreeding coefficient of 1.9% for cows from Massey University Dairy 1 and Dairy 4. In comparison, Sun et al. (2014) reported average inbreeding levels of 5.66% and 6.63% for USA Holstein and Jersey cattle, respectively. Because an increase in inbreeding can reduce genetic variance in a population (Visscher et al. 2008), these differences may contribute to the higher heritability estimates for milk composition traits compared to those in international studies. Furthermore, environmental and management differences between New Zealand and the USA may also influence heritability estimates through differences in environmental variance (Visscher et al. 2008).

5.2.1.2 Fertility traits

Heritability estimates from the pedigree-based model for SR21 and start of mating to first service (SMFS) were consistent with those reported by Jayawardana et al. (2023b) but were higher for pregnancy rate to first service (PRFS) and PR42. This difference may be due to the larger dataset used by Jayawardana et al. (2023b) which included approximately 490,000 first-lactation cows, and may have allowed more stable population-level estimates for traits with low heritability.

Heritability estimates from the genomic model were lower for fertility traits than for the pedigree-based model, with SR21 and PR42 estimates being close to zero. For both pedigree and genomic models, standard errors were high relative to heritability, indicating low precision. This reflects the difficulty of estimating genetic variance for fertility traits and may be due to the binary nature of pregnancy-related traits, which can make it difficult to estimate underlying additive and non-additive genetic variance (Cammack et al. 2009). It is also more difficult to partition variance components due to the low heritability and high residual variance of these traits (Martinez-Castillero et al. 2020). There are limited genomic studies estimating heritability for these fertility traits, making direct comparisons difficult. However, the low estimates observed in the present study are consistent with previous findings for fertility traits in New Zealand dairy cows based on pedigree models (Jayawardana et al. 2023b).

5.2.2 Dominance variance

The contribution of dominance variance to phenotypic variance was smaller than that of additive genetic variance, but non-negligible dominance effects were observed for yield traits. However, these had relatively large standard errors (0.051–0.054), indicating low precision. In contrast, the proportion of dominance variance to phenotypic variance was negligible for milk composition traits, with estimates of zero for FP and PP, and a very small estimate (0.044) for LP that was associated with a large standard error (0.047). This indicates that dominance effects for these traits are minimal or could not be reliably detected in the present study.

Estimates for the proportion of dominance variance to phenotypic variance for yield traits were higher than those reported by Reynolds et al. (2022) in New Zealand dairy cattle (Table 2.1). Dominance variance for FP and PP was zero, which was consistent with results from Reynolds et al. (2022), which reported low dominance variance of 0–0.015 for New Zealand Holstein-Friesian and Jersey cattle, but were lower than estimates of 0.009–0.03 reported by Sun et al. (2014) in USA Holstein and Jersey cattle.

The higher dominance variance observed for yield traits in the present study may be due to differences in population structure compared with those examined by Reynolds et al. (2022), Schneider et al. (2023), Sun et al. (2014), and Aliloo et al. (2016). The population in this study included Holstein-Friesian, Jersey, and crossbred cattle, whereas the previous studies analysed purebred Holstein or Holstein-Friesian and purebred Jersey populations separately. This may affect the dominance variance estimates, as crossbred populations typically exhibit higher levels of heterozygosity than purebred populations (Iversen et al. 2019) which can increase the

expression of dominance deviations. This is consistent with the study by Tusell et al. (2019) which reported higher dominance variance in crossbred pigs compared with purebred populations across multiple traits, and a greater number of traits showing significant dominance effects in crossbreds.

Differences among studies may also be due to variation in dataset size and modelling approaches. Previous studies used much larger genotyped populations than in the present study, ranging from 13,000 to 39,000 cows (Sun et al. 2014; Aliloo et al. 2016; Reynolds et al. 2022), which may have allowed for more stable dominance variance estimates. In contrast, the smaller dataset in the present study may have limited the ability to accurately separate dominance variance from permanent environmental effects. When sample size is limited, these components may be confounded and dominance variance can be over-estimated if permanent environmental variance is absorbed (Aliloo et al. 2016). Additionally, the present study fitted fixed effects directly within the mixed model, whereas Reynolds et al. (2022), Aliloo et al. (2016), and Sun et al. (2014) used pre-corrected phenotypes. This might influence variance component estimates because fixed effects are estimated separately from the model rather than within a single framework. While the inclusion of breed proportion and heterosis as fixed effects allows the separation of breed differences from dominance effects, dominance variance estimates may still be partially confounded with other sources of variation due to the relatively small dataset.

Dominance variance was estimated as zero for SMFS, PRFS, and SR21, and 0.009 for PR42 with a very high relative standard error of 0.032. While these traits have not been examined in other studies calculating dominance variance, these results are consistent with studies that have used measures of fertility such as calving interval which showed very low dominance variance (0–0.012) (Druet et al. 2001; Aliloo et al. 2016). This is likely due to the low heritability of reproductive traits, making it difficult to partition genetic variance, including dominance, from environmental and residual variance components, especially in a smaller dataset like in this study.

5.3 Genome-wide association studies

5.3.1 Yield traits

The genome-wide association analysis identified significant additive associations across the yield traits. Association signals were observed on chromosomes 14 and 5 for both MY and LY. For FY and PY, a signal was also observed on chromosome 14. These results are consistent

with previous studies that have identified a major quantitative trait locus (QTL) for milk production traits on chromosome 14, primarily due to the *diacylglycerol O-acyltransferase 1* (*DGATI*) gene, which has been widely reported to have significant effects on milk, fat, protein, and lactose production in dairy cattle (Grisart et al. 2002; Spelman et al. 2002; Bovenhuis et al. 2015; Li et al. 2021).

A QTL for LY has been reported on chromosome 5, potentially due to the presence of the *lactalbumin alpha* gene, which encodes α -lactalbumin, the regulatory subunit of the lactose synthase complex involved in lactose synthesis (Lopdell et al. 2017). Significant associations with MY on chromosome 5 have also been previously identified (Kemper et al. 2014; Iso-Touru et al. 2016; Maddahi et al. 2025). This may be due to the role of lactose in regulating milk volume, as increased lactose synthesis is expected to increase MY through osmotic effects (Kemper et al. 2014; Gargiulo et al. 2025). Overall, the additive association signals identified for yield traits are consistent with previously reported regions with significant additive SNP associations. This suggests that the additive component of the model captured important genetic signals influencing these traits.

In contrast to additive effects, fewer and weaker dominance associations were detected for yield traits across the genome. For MY, six SNPs exceeded the genome-wide significance threshold, but no distinct association peaks were observed. The top five most significant SNPs associated with dominance effects are presented in Table 4.8, and were located on chromosomes 24, 5, 15, and 16. For FY, several SNPs exceeded the genome-wide threshold, and as seen in Figure 4.1, a distinct peak was observed on chromosome 9, with other significant SNPs more dispersed across the genome. Boysen et al. (2013) used a genome-wide association approach using genotype probabilities based on German Holstein bull genotypes to investigate dominance effects for milk production traits. They reported significant dominance associations on chromosome 9 for MY, FY, and PY, including several SNPs with high degrees of dominance where additive effects were not significant (Boysen et al. 2013). This is consistent with the results in the present study for FY, where the association signal observed on chromosome 9 for FY did not correspond to significant additive associations, suggesting that this region may contain loci with non-additive contributions to genetic variation. The three most highly significant SNPs showing dominance association signals for FY were located on chromosome 3, which also showed significant additive associations, indicating that both additive and dominance effects may contribute to this signal.

For PY, eight SNPs exceeded the genome-wide significance threshold, with a small peak also observed on chromosome 9, further supporting the findings of Boysen et al. (2013). For LY, only four SNPs exceeded the genome-wide significance threshold, with no clear association peaks observed. These SNPs were located on chromosomes 5, 16, 24, and 14 and were identical to those identified for MY, indicating a shared genetic influence on these traits. This is expected, as lactose regulates milk volume through osmotic mechanisms and is closely linked to overall milk production and component synthesis (Gargiulo et al. 2025). Across all yield traits, the corresponding additive associations at most loci were not significant, which suggests that the detected signals may reflect non-additive genetic variation.

Examination of genotype means for the most significant dominance SNP for 305-day MY, FY, PY, and LY, adjusted for fixed effects (Figure 4.4), showed patterns consistent with overdominance for these SNP markers across these traits. For example, for MY, heterozygote individuals deviated from the midpoint of the homozygotes (4,902 L) by an average of 185 L and exceeded the higher-producing homozygote by 99 L. This overdominance was observed for other yield traits, where heterozygotes produced approximately 1 kg, 2 kg, and 6 kg more than the higher-producing homozygote for FY, PY, and LY, respectively. Overall, these results indicate that yield traits are predominantly influenced by additive genetic effects, with dominance contributing to a smaller extent. However, regions showing dominance association signals in the absence of corresponding additive associations highlight loci that would not be detected using additive-only models, indicating that accounting for dominance effects for yield traits in New Zealand crossbred populations may improve the capture of total genetic variation.

5.3.2 Milk composition traits

For FP, a strong additive association signal was observed on chromosome 14 in the present study, with a secondary signal on chromosome 5. For PP, significant additive associations were also identified on chromosome 14, with an additional signal on chromosome 6. LP had a small peak on chromosome 6, and a signal was observed on chromosome 19 for FP, PP, and LP. These findings are consistent with those reported by Jayawardana et al. (2023a), who identified significant marker effects for milk composition traits in cows from the same Massey University Dairy 1 and Dairy 4 herds.

The signal on chromosome 14 for both FP and PP reflects the widely established role of this region on milk composition, including the effects from genes such as *DGAT1*. Similarly, the detection of signals on chromosome 5 for FP and chromosome 6 for PP is consistent with

previous GWAS, where candidate genes such as *MGST1* and *CSN3* have been identified for milk fat and protein synthesis, respectively (Jayawardana et al. 2023a). Previous studies have reported QTLs on chromosome 19 associated with milk composition traits, including lactose concentration, with candidate genes such as *STAT5A* and *STAT5B* involved in milk protein gene expression (Bouwman et al. 2011; Lopdell et al. 2017). Additionally, a major QTL on chromosome 6 has been reported for lactose concentration, driven by variation in the *ABCG2* gene, which encodes a membrane transporter involved in osmotic regulation (Lopdell et al. 2017). This supports the signal observed on chromosome 6 in the present study for LP.

The results for dominance components showed few significant associations for milk composition traits, with limited signals exceeding the genome-wide significance threshold. PP and LP each had two SNPs that surpassed this threshold but did not form clear association peaks so may represent isolated signals. In contrast, FP showed stronger evidence of dominance associations, with ten SNPs above the genome-wide significance threshold, including two SNPs on chromosome 14 that were well above it. This is consistent with previous studies, such as Sun et al. (2014) and Boysen et al. (2013), which reported significant dominance associations for FY on chromosome 14 in the *DGATI* region. However, in the present study, these signals did not form clear association peaks, and no strong dominance signals were detected for yield traits on chromosome 14, as reported in earlier studies. This may reflect differences in sample size, statistical power, or breed and population structure which can influence the detection of non-additive genetic effects (Aliloo et al. 2016). Overall, these findings indicate that milk composition traits are predominantly influenced by additive genetic effects, while dominance effects are minimal. This is consistent with the variance component estimates (Table 4.7), which showed that dominance explained a negligible proportion of the phenotypic variance for composition traits.

5.3.3 Fertility traits

Relatively few strong associations were identified for fertility traits in the present study. This is consistent with findings by Jayawardana et al. (2023a) on the same fertility traits, where only limited genome-wide significant associations were detected. This likely reflects the low heritability of fertility traits and the small contribution of individual QTLs to the phenotypic variance, which reduces the power to detect associations (Jayawardana et al. 2023a). However, some smaller peaks exceeded the suggestive significance threshold, with a peak around chromosome 7 for SR21, which was not detected by Jayawardana et al. (2023a), but is consistent with previously reported regions linked to reproductive traits such as early

pregnancy (Irano et al. 2016). Additionally, an additive association on chromosome 5 for PR42 was observed in the present study, although this region has not been consistently reported in previous studies as an important region for fertility traits.

Few significant dominance associations were detected for fertility traits, which is consistent with low heritability and difficulty of partitioning non-additive effects. No significant dominance SNPs were identified for SMFS, while only a small number of SNPs exceeded the suggestive significance threshold for PRFS, SR21, and PR42. Despite the limited number of associations, the presence of dominance signals in the absence of corresponding additive signals may be important, particularly for PRFS where no significant additive associations were identified. This suggests some loci associated with fertility traits may show dominance association signals that are not captured by additive-only models. This is further supported by genotype means at specific loci (Figure 4.6), where heterozygotes had higher phenotypic values than both homozygotes, consistent with overdominance. Although these findings indicate that dominance association signals for fertility traits were generally weak, some loci showed dominance associations where additive associations were minimal or absent.

5.4 Implications

The findings of this study contribute to a clearer understanding of the role of dominance variance underlying economically important traits in New Zealand dairy cattle. The genetic parameter estimates demonstrate that dominance effects are non-negligible for milk yield traits, with higher estimates than those reported in previous studies. The crossbred structure of the population may have influenced the dominance variance estimates through differences in heterozygosity, allele frequencies, and the expression of dominance deviations. These findings have practical implications for dairy cattle breeding in New Zealand, as current genomic evaluation systems focus on additive effects and therefore may underestimate total genetic value for traits influenced by dominance variance.

Inclusion of dominance effects may affect the ranking of animals when genomic evaluations are based on total genetic values rather than only on additive breeding values. Aliloo et al. (2016) compared the rankings of animals based on their additive breeding values with rankings based on total genetic value and reported a slight re-ranking of animals, especially among the top-ranking individuals. Although their study did not directly evaluate the performance of re-ranked animals, the improved predictive ability observed for some traits indicates that incorporating dominance deviations into genomic evaluations may provide a more complete

prediction of total individual performance. Therefore, incorporating dominance effects into genomic evaluation may improve identification and selection of animals with high total genetic merit and contribute to improved overall herd performance (Aliloo et al. 2016).

Additionally, the GWAS identified loci exhibiting dominance association signals in the absence of significant additive associations, suggesting that these loci may not be detected in an additive-only model. This highlights the potential value of including dominance information in genomic selection models. Identified dominance markers could be used in mate allocation strategies by optimising specific combining ability and allowing breeders to select mating pairs that maximise favourable allele combinations, reduce the expression of inbreeding depression, and utilise potential overdominance at a SNP level (Wellmann and Bennewitz 2012).

In fertility traits, although additive and dominance variance as a proportion of phenotypic variance were low, the presence of some dominance associations suggests that non-additive genetic variation may play a role where additive genetic variation does not. In these traits, where selection response is limited, incorporating dominance information may provide an alternative pathway to improve reproductive performance. This could be relevant in designing mating strategies that enhance heterozygote advantage and improve outcomes such as pregnancy rates, as seen for PRFS (Figure 4.6).

Overall, these findings indicate that dominance effects may be relevant for some traits in New Zealand dairy cattle. Therefore, incorporating genomic information that accounts for dominance effects into genomic evaluation models could provide a more complete estimate of total genetic value, which may be useful for improving future selection and mating decisions.

5.5 Study limitations

A primary limitation of this study is the relatively small population size used compared with previous studies. Estimation of dominance variance requires large datasets to accurately partition variance components, and smaller datasets may result in confounding of these components, leading to reduced precision and potentially inflated or underestimated dominance variance estimates (Aliloo et al. 2016). This is reflected in the high standard errors observed in the present study for dominance variance estimates. The relatively small sample size in the GWAS also reduces statistical power to detect significant associations, which may explain why some dominance-associated regions detected in other studies were not detected in the present study. Another limitation of this study was that the genotypes were reduced to 10% of the original SNP dataset to reduce the computational demands of the analyses. This may

have reduced the power to detect dominance effects in this study as fewer SNPs decrease the likelihood that markers are in linkage disequilibrium with causal variants (Zhu et al. 2015).

The structure of the population analysed may also present a limitation, as it consisted of animals from two research herds that may not fully represent the genetic diversity present in the New Zealand dairy population. In addition to this, variance components were estimated using a combined population of purebred and crossbred animals. Because dominance effects depend on allele frequencies and heterozygosity, which differ between breeds and crossbreds (Vitezica et al. 2016), the estimated dominance variance reflects the genetic structure of the Dairy 1 and Dairy 4 herds. Therefore, the findings from this study may not be representative of other commercial dairy populations in New Zealand and should be validated in larger populations.

5.6 Future research

Future research should expand on the current findings by applying GWAS to larger and more representative New Zealand dairy cattle populations. Incorporating larger datasets, such as from the New Zealand national herd, would allow for greater statistical power and validate the findings in this study in other commercial herds. Additionally, future studies should compare conventional additive-only models with models that incorporate both additive and dominance effects to estimate total genetic values in New Zealand populations. As demonstrated in international studies, including dominance effects may provide a more complete representation of the genetic merit of the animals (Sun et al. 2014; Aliloo et al. 2016). This could help determine the value of incorporating non-additive genetic effects in genomic selection models for improving the selection of animals and mating decisions where dominance effects can be exploited.

5.7 Conclusions

This research aimed to estimate additive and dominance variance components for milk production and fertility traits, and to identify genomic regions associated with these effects using genome-wide association analysis. Additive genetic variance contributed to a greater proportion of phenotypic variance than dominance variance for milk yield and composition traits, indicating that additive effects were the primary driver of phenotypic variation. However, dominance contributed to phenotypic variance for milk, fat, protein, and lactose yields but was close to zero for milk composition traits. Fertility traits showed low heritability and negligible dominance variance.

Genome-wide association analyses fitting both additive and dominance SNP effects simultaneously identified significant additive associations including regions on chromosomes 14 and 5 for yield and composition traits. In comparison, fewer and weaker dominance associations were detected across the genome for all traits, including a region on chromosome 9 for fat yield, where several markers showed dominance associations in the absence of significant additive association signals. The limited number of significant markers detected for fertility traits was consistent with the low heritability and polygenic nature of these traits.

Overall, these findings indicate that dominance effects may contribute to variation in some economically important milk production traits, but additive effects remain the primary source of genetic variation. Incorporating both additive and non-additive genetic effects into future genomic studies may improve understanding of trait architecture in New Zealand dairy cattle.

References

- Aliloo H, Pryce JE, González-Recio O, Cocks BG, Hayes BJ. 2016. Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. *Genetics Selection Evolution*. 48:8.
- Balendran A, Gordon M, Pretheeban T, Singh R, Perera R, Rajamahedran R. 2008. Decreased fertility with increasing parity in lactating dairy cows. *Canadian Journal of Animal Science*. 88:425-428.
- Berry DP, Downing K. 2025. Relationship between cow parity and maternal parity on dairy cow lactation performance. *JDS Communications*. 6:544-547.
- Bezanson J, Edelman A, Karpinski S, Shah VB. 2017. Julia: A fresh approach to numerical computing. *SIAM Review*. 59:65-98.
- Bouwman AC, Bovenhuis H, Visker MH, Van Arendonk JA. 2011. Genome-wide association of milk fatty acids in Dutch dairy cattle. *BMC Genetics*. 12:43.
- Bovenhuis H, Visker MHPW, Van Valenberg HJF, Buitenhuis AJ, Van Arendonk JAM. 2015. Effects of the DGAT1 polymorphism on test-day milk production traits throughout lactation. *Journal of Dairy Science*. 98:6572-6582.
- Boysen T-J, Heuer C, Tetens J, Reinhardt F, Thaller G. 2013. Novel use of derived genotype probabilities to discover significant dominance effects for milk production traits in dairy cattle. *Genetics*. 193:431-442.
- Burow E, Thomsen PT, Sørensen JT, Rousing T. 2011. The effect of grazing on cow mortality in Danish dairy herds. *Preventive Veterinary Medicine*. 100:237-241.
- Cammack KM, Thomas MG, Enns RM. 2009. Reproductive traits and their heritabilities in beef cattle. *The Professional Animal Scientist*. 25:517-528.
- DairyNZ. 2014. Technical Series July 2014 - Genetic Improvement https://www.dairynz.co.nz/media/gashsy2k/technical_series_july_2014.pdf [8 June 2025]
- DairyNZ. 2024a. Breeding Worth Retrieved 8 June 2026 <https://www.dairynz.co.nz/animal/breeding-decisions/breeding-worth/> [8 June 2026]
- DairyNZ. 2024b. Breeding Values. Retrieved 11 June 2026 11 June 2026]
- DairyNZ. 2025a. Annual Report 2024-2025. <https://www.dairynz.co.nz/media/ha4fg0cg/annual-report-2024-2025-v2.pdf> [8 June 2026]
- DairyNZ. 2025b. Genetic Trends in the National Herd. Retrieved 8 June 2026 <https://www.dairynz.co.nz/animal/breeding-decisions/genetic-trends-in-the-national-herd/> [8 June 2026]
- DairyNZ, Livestock Improvement Corporation. 2022. New Zealand Dairy Statistics 2021-2022. <https://www.dairynz.co.nz/media/uzeekwgr/nz-dairy-statistics-2021-22-web.pdf> [8 June 2026]
- DairyNZ, Livestock Improvement Corporation. 2024. New Zealand Dairy Statistics 2023-24. <https://www.dairynz.co.nz/media/bywm13d4/dairy-statistics-2023-24.pdf> [8 June 2026]
- Dickerson GE. 1973. Inbreeding and heterosis in animals. *Journal of Animal Science*. 1973:54-77.
- Druet T, Sölkner J, Groen AF, Gengler N. 2001. Additive and dominance genetic variance of fertility by method $\mathfrak{R}$ and preconditioned conjugate gradient. *Journal of Dairy Science*. 84:987. e1-987. e16.
- Eilertsen EM 2021. GREMLModels.jl – A Julia package for fitting variance component models with relationship matrices. Version 0.1.0.
- Ertl J, Legarra A, Vitezica ZG, Varona L, Edel C, Emmerling R, Götz K-U. 2014. Genomic analysis of dominance effects on milk production and conformation traits in Fleckvieh cattle. *Genetics Selection Evolution*. 46:40.

- Falconer DS. 1996. Introduction to quantitative genetics, Pearson Education India.
- Fuerst C, Sölkner J. 1994. Additive and nonadditive genetic variances for milk yield, fertility, and lifetime performance traits of dairy cattle. *Journal of Dairy Science*. 77:1114-1125.
- Garcia-Baccino CA, Lourenco DAL, Miller S, Cantet RJC, Vitezica ZG. 2020. Estimating dominance genetic variances for growth traits in American Angus males using genomic models. *Journal of Animal Science*. 98.
- Gargiulo JI, Garcia SC, Hovey RC. 2025. Sources of variation underlying the production of lactose by dairy cows. *Journal of Dairy Science*. 108:4403-4421.
- Ghiasi H, Abdollahi-Arpanahi R, Razmkabir M, Khaldari M, Taherkhani R. 2018. Estimation of additive and dominance genetic variance components for female fertility traits in Iranian Holstein cows. *The Journal of Agricultural Science*. 156:565-569.
- Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R. 2021. ASReml User Guide Release 4.2 Functional Specification. www.vsni.co.uk
- González-Diéguez D, Tusell L, Bouquet A, Legarra A, Vitezica ZG. 2020. Purebred and crossbred genomic evaluation and mate allocation strategies to exploit dominance in pig crossbreeding schemes. *G3 Genes|Genomes|Genetics*. 10:2829-2841.
- Grisart B, Coppieters W, Farnir F, Karim L, Ford C, Berzi P, Cambisano N, Mni M, Reid S, Simon P, Spelman R, Georges M, Snell R. 2002. Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine *DGATI* Gene with Major Effect on Milk Yield and Composition. *Genome Research*. 12:222-231.
- Haile-Mariam M, Nieuwhof GJ, Beard KT, Konstatinov KV, Hayes BJ. 2013. Comparison of heritabilities of dairy traits in Australian Holstein-Friesian cattle from genomic and pedigree data and implications for genomic evaluations. *Journal of Animal Breeding and Genetics*. 130:20-31.
- Harris BL. 2005. Breeding dairy cows for the future in New Zealand. *New Zealand Veterinary Journal*. 53:384-389.
- Harris BL. 2022. Genomic evaluations for crossbred dairy cattle. *JDS Communications*. 3:152-155.
- Harris BL, Montgomerie WA. 2001. Fertility breeding values for seasonal dairying. *Interbull Bulletin*:139-139.
- Hayes B, Goddard M. 2010. Genome-wide association and genomic selection in animal breeding. *Genome*. 53:876-883.
- Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*. 92:433-443.
- Henderson CR. 1949. Estimation of changes in herd environment. *J. Dairy Sci*. 32:706-706.
- Hoeschele I. 1991. Additive and nonadditive genetic variance in female fertility of Holsteins. *Journal of Dairy Science*. 74:1743-1752.
- Irano N, De Camargo GMF, Costa RB, Terakado APN, Magalhães AFB, Silva RMDO, Dias MM, Bignardi AB, Baldi F, Carvalheiro R, De Oliveira HN, De Albuquerque LG. 2016. Genome-wide association study for indicator traits of sexual precocity in nellore cattle. *PLOS ONE*. 11:e0159502.
- Iso-Touru T, Sahana G, Guldbbrandtsen B, Lund MS, Vilkki J. 2016. Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants. *BMC Genetics*. 17: 55.
- Iversen MW, Nordbø Ø, Gjerlaug-Enger E, Grindflek E, Lopes MS, Meuwissen T. 2019. Effects of heterozygosity on performance of purebred and crossbred pigs. *Genetics Selection Evolution*. 51: 8.
- Jago J, Beukes P, Cuttance E, Dalley D, Edwards JP, Griffiths W, Saunders K, Shackleton L, Schütz K. 2023. Strategies to minimise the impact of climate change and weather

- variability on the welfare of dairy cattle in New Zealand and Australia. *Animal Production Science*. 64.
- Jayawardana JMDR, Lopez-Villalobos N, McNaughton LR, Hickson RE. 2023a. Genomic regions associated with milk composition and fertility traits in spring-calved dairy cows in New Zealand. *Genes*. 14:860.
- Jayawardana JMDR, Lopez-Villalobos N, Hickson RE, McNaughton LR. 2023b. Estimation of genetic parameters and individual and maternal breed, heterosis, and recombination loss effects for production and fertility traits of spring-calved cows milked once daily or twice daily in New Zealand. *Journal of Dairy Science*. 106:364-380.
- Jiang J, Ma L, Prakapenka D, Van Raden PM, Cole JB, Da Y. 2019. A large-scale genome-wide association study in U.S. Holstein cattle. *Frontiers in Genetics*. 10:412.
- Johnson T, Eketone K, McNaughton L, Tiplady K, Voogt J, Sherlock R, Anderson G, Keehan M, Davis SR, Spelman RJ, Chin D, Couldrey C. 2018. Mating strategies to maximize genetic merit in dairy cattle herds. *Journal of Dairy Science*. 101:4650-4659.
- Kemper K, Vander Jagt C, Bowman P, Reich C, Mason B, Hayes B, Goddard M. 2014. Mapping QTL in Australian dairy cattle using Genomic Selection Methodologies. *Proceedings of the 10th World Congress on Genetics Applied to Livestock Production*, Vancouver, BC, Canada. 17-22.
- Kolver ES, Muller LD. 1998. Performance and nutrient intake of high producing Holstein cows consuming pasture or a total mixed ration. *Journal of Dairy Science*. 81:1403-1411.
- Lee SH, Goddard ME, Visscher PM, Van Der Werf JH. 2010. Using the realized relationship matrix to disentangle confounding factors for the estimation of genetic variance components of complex traits. *Genetics Selection Evolution*. 42:22.
- Lembeye F, Lopez-Villalobos N, Burke JL, Davis SR. 2016. Estimation of genetic parameters for milk traits in cows milked once- or twice-daily in New Zealand. *Livestock Science*. 185:142-147.
- Li Y, Zhou H, Cheng L, Edwards GR, Hickford JGH. 2021. Effect of DGAT1 variant (K232A) on milk traits and milk fat composition in outdoor pasture-grazed dairy cattle. *New Zealand Journal of Agricultural Research*. 64:101-113.
- Liu T, Luo C, Ma J, Wang Y, Shu D, Qu H, Su G. 2022. Including dominance effects in the prediction model through locus-specific weights on heterozygous genotypes can greatly improve genomic predictive abilities. *Heredity*. 128:154-158.
- Livestock Improvement Corporation. 2023. Selecting super beef sires for dairy. Retrieved 8 June 2026 <https://www.lic.co.nz/news/selecting-super-beef-sires-for-dairy/> [8 June 2026]
- Lopdell TJ, Tiplady K, Struchalin M, Johnson TJJ, Keehan M, Sherlock R, Couldrey C, Davis SR, Snell RG, Spelman RJ, Littlejohn MD. 2017. DNA and RNA-sequence based GWAS highlights membrane-transport genes as key modulators of milk lactose content. *BMC Genomics*. 18: 968.
- Lopes MS, Bastiaansen JW, Harlizius B, Knol EF, Bovenhuis H. 2014. A genome-wide association study reveals dominance effects on number of teats in pigs. *PloS one*. 9:e105867.
- Lopez-Villalobos N, Wiles P, Udy G. 2024. The Value of Genetic Improvement Evaluated Using a Whole of Enterprise Market Model. *Dairy*. 5:372-383.
- Maddahi N, Sadeghi M, Sarghale AJ, Saatchi M, Davar Siar MK, Kholghi M. 2025. Identification of candidate gene networks affecting the number of somatic cells count and milk production in Iranian Holstein cows using Genome-wide association study. *Scientific Reports*. 15.

- Manzanilla-Pech CIV, Stephansen RB, Lassen J. 2023. Genetic parameters for feed intake and body weight in dairy cattle using high-throughput 3-dimensional cameras in Danish commercial farms. *Journal of Dairy Science*. 106:9006-9015.
- Martinez-Castillero M, Toledo-Alvarado H, Pegolo S, Vazquez AI, De Los Campos G, Varona L, Finocchiario R, Bittante G, Cecchinato A. 2020. Genetic parameters for fertility traits assessed in herds divergent in milk energy output in Holstein-Friesian, Brown Swiss, and Simmental cattle. *Journal of Dairy Science*. 103:11545-11558.
- Matthews D, Kearney JF, Cromie AR, Hely FS, Amer PR. 2019. Genetic benefits of genomic selection breeding programmes considering foreign sire contributions. *Genetics Selection Evolution*. 51:40.
- Meuwissen TE, Hayes BJ, Goddard ME. 2001. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics*. 157:1819-1829.
- Miglior F, Burnside EB, Kennedy BW. 1995. Production traits of Holstein cattle: estimation of nonadditive genetic variance components and inbreeding depression. *Journal of Dairy Science*. 78:1174-1180.
- Mrode RA. 2005. Linear models for the prediction of animal breeding values, CABI.
- O'Neill BF, Deighton MH, O'Loughlin BM, Mulligan FJ, Boland TM, O'Donovan M, Lewis E. 2011. Effects of a perennial ryegrass diet or total mixed ration diet offered to spring-calving Holstein-Friesian dairy cows on methane emissions, dry matter intake, and milk production. *Journal of Dairy Science*. 94:1941-1951.
- Porteous T. 2022. Effect of genomic inbreeding on lactation yields of milk, fat and protein and fertility of grazing dairy cows in two New Zealand farms. Master's thesis, Massey University, Palmerston North, New Zealand.
- R Core Team 2025. R: A Language and Environment for Statistical Computing. Version 4.4.3. Vienna, Austria, R Foundation for Statistical Computing.
- Rettew DC, Rebollo-Mesa I, Hudziak JJ, Willemsen G, Boomsma DI. 2008. Non-additive and additive genetic effects on extraversion in 3314 Dutch adolescent twins and their parents. *Behavior Genetics*. 38:223-233.
- Reynolds EGM, Lopdell T, Wang Y, Tiplady KM, Harland CS, Johnson TJJ, Neeley C, Carnie K, Sherlock RG, Couldrey C. 2022. Non-additive QTL mapping of lactation traits in 124,000 cattle reveals novel recessive loci. *Genetics Selection Evolution*. 54:5.
- Reynolds EGM, Neeley C, Lopdell TJ, Keehan M, Dittmer K, Harland CS, Couldrey C, Johnson TJJ, Tiplady K, Worth G, Walker M, Davis SR, Sherlock RG, Carnie K, Harris BL, Charlier C, Georges M, Spelman RJ, Garrick DJ, Littlejohn MD. 2021. Non-additive association analysis using proxy phenotypes identifies novel cattle syndromes. *Nature Genetics*. 53:949-954.
- Ridhowi A, Chang H-L, Liang C-L, Wu M-C, Thiasari N, Hsia L-C, Lee J-W. 2025. Effects of Season and Parity on Milk Production and Milk Composition of Holstein Cows. *BIO Web of Conferences*. 00009.
- Robertson A, Rendel JM. 1950. The use of progeny testing with artificial insemination in dairy cattle. *Journal of Genetics*. 50:21-31.
- Schaeffer LR. 2006. Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics*. 123:218-223.
- Schären M, Jostmeier S, Ruesink S, Hütther L, Frahm J, Bulang M, Meyer U, Rehage J, Isselstein J, Breves G, Dänicke S. 2016. The effects of a ration change from a total mixed ration to pasture on health and production of dairy cows. *Journal of Dairy Science*. 99:1183-1200.
- Scheffers JM, Weigel KA. 2012. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. *Animal Frontiers*. 2:4-9.

- Schneider H, Heise J, Tetens J, Thaller G, Wellmann R, Bennewitz J. 2023. Genomic dominance variance analysis of health and milk production traits in German Holstein cattle. *Journal of Animal Breeding and Genetics*. 140:390-399.
- Sørensen M, Norberg E, Pedersen J, Christensen L. 2008. Invited review: Crossbreeding in dairy cattle: A Danish perspective. *Journal of Dairy Science*. 91:4116-4128.
- Spelman RJ, Ford CA, McElhinney P, Gregory GC, Snell RG. 2002. Characterization of the DGAT1 Gene in the New Zealand Dairy Population. *Journal of Dairy Science*. 85:3514-3517.
- Spelman RJ, Arias J, Keehan MD, Obolonkin V, Winkelman AM, Johnson DL, Harris BL. 2010. Application of genomic selection in the New Zealand dairy cattle industry. *Proceedings of the 9th world congress on genetics applied to livestock production*. 1-6.
- Sun C, VanRaden PM, Cole JB, O'Connell JR. 2014. Improvement of prediction ability for genomic selection of dairy cattle by including dominance effects. *PloS one*. 9:e103934.
- Sun C, Vanraden PM, O'Connell JR, Weigel KA, Gianola D. 2013. Mating programs including genomic relationships and dominance effects. *Journal of Dairy Science*. 96:8014-8023.
- Tempelman RJ, Burnside EB. 1990. Additive and nonadditive genetic variation for production traits in Canadian Holsteins. *Journal of Dairy Science*. 73:2206-2213.
- Toro MA, Varona L. 2010. A note on mate allocation for dominance handling in genomic selection. *Genetics Selection Evolution*. 42:33.
- Turner SD. 2014. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *Biorxiv*:005165.
- Tusell L, Gilbert H, Vitezica ZG, Mercat MJ, Legarra A, Larzul C. 2019. Dissecting total genetic variance into additive and dominance components of purebred and crossbred pig traits. *Animal*. 13:2429-2439.
- Van der Werf J. 2015. Principles of estimation of breeding values. *Genetic evaluation and breeding program design*. Armidale, Australia: University of New England:1-17.
- Van Raden PM. 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science*. 91:4414-4423.
- Van Tassell CP, Misztal I, Varona L. 2000. Method R estimates of additive genetic, dominance genetic, and permanent environmental fraction of variance for yield and health traits of Holsteins. *Journal of Dairy Science*. 83:1873-1877.
- Varona L, Legarra A, Toro MA, Vitezica ZG. 2018. Non-additive effects in genomic selection. *Frontiers in Genetics*. 9:78.
- Varona L, Legarra A, Toro MA, Vitezica ZG. 2022. *Genomic Prediction Methods Accounting for Nonadditive Genetic Effects*. Springer US. Pp. 219-243.
- Veerkamp RF, Mulder HA, Thompson R, Calus MPL. 2011. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. *Journal of Dairy Science*. 94:4189-4197.
- Verkerk G. 2003. Pasture-based dairying: challenges and rewards for New Zealand producers. *Theriogenology*. 59:553-561.
- Visser PM, Hill WG, Wray NR. 2008. Heritability in the genomics era — concepts and misconceptions. *Nature Reviews Genetics*. 9:255-266.
- Vitezica ZG, Varona L, Legarra A. 2013. On the Additive and Dominant Variance and Covariance of Individuals Within the Genomic Selection Scope. *Genetics*. 195:1223-1230.
- Vitezica ZG, Varona L, Elsen J-M, Misztal I, Herring W, Legarra A. 2016. Genomic BLUP including additive and dominant variation in purebreds and F1 crossbreds, with an application in pigs. *Genetics Selection Evolution*. 48: 6.

- Wakchaure R, Ganguly S, Praveen PK, Sharma S, Kumar A, Mahajan T, Qadri K. 2015. Importance of heterosis in animals: a review. *International Journal of Advanced Engineering Technology and Innovative Science*. 1:1-5.
- Wang C-L, Ma P-P, Zhang Z, Ding X-D, Liu J-F, Fu W-X, Weng Z-Q, Zhang Q. 2012. Comparison of five methods for genomic breeding value estimation for the common dataset of the 15th QTL-MAS Workshop. *BMC Proceedings*. 6:S13.
- Wang Y. 2020. Evaluating the accuracy of imputed whole-genome sequence data in admixed dairy cattle. *Proceedings of the 6th International Conference of Quantitative Genetics*. 427
- Wei W-H, Hemani G, Haley CS. 2014. Detecting epistasis in human complex traits. *Nature Reviews Genetics*. 15:722-733.
- Wellmann R, Bennewitz J. 2012. Bayesian models with dominance effects for genomic evaluation of quantitative traits. *Genetics Research*. 94:21-37.
- Widmayer SJ, Evans KS, Zdravljec S, Andersen EC. 2022. Evaluating the power and limitations of genome-wide association studies in *Caenorhabditis elegans*. *G3 Genes|Genomes|Genetics*. 12:jkac114.
- Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics*. 88:76-82.
- Zeng J, Toosi A, Fernando RL, Dekkers JC, Garrick DJ. 2013. Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. *Genetics Selection Evolution*. 45:11.
- Zhang Z, Zhang Q, Ding X. 2011. Advances in genomic selection in domestic animals. *Chinese Science Bulletin*. 56:2655-2663.
- Zhu Z, Bakshi A, Anna, Hemani G, Sang, Ilja, Jana, Snieder H, Esko T, Milani L, Mägi R, Metspalu A, William, Bruce, Michael, Peter, Yang J. 2015. Dominance Genetic Variation Contributes Little to the Missing Heritability for Human Complex Traits. *The American Journal of Human Genetics*. 96:377-385.