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.

Genetics of flystrike, dagginess and associated traits in New Zealand dual-purpose sheep

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

Doctor of Philosophy

in

Animal Science

at Massey University, Palmerston North, New Zealand

Natalie Kathleen Pickering

2013

"The greatest benefit of dag control or elimination may come from reduced fly strike" V. J. Mackereth (1983), an astute breeder.

Abstract

Pickering, N. K. (2013). Genetics of flystrike, dagginess and associated traits in New Zealand dual-purpose sheep. PhD thesis, Massey University, Palmerston North, New Zealand.

A literature review identified breech bareness, dagginess and fibre traits as potential indirect indicator traits for flystrike. Dagginess (faecal accumulation) had the greatest potential as an indirect indicator, and has been identified as an important trait itself. Therefore flystrike and dagginess were investigated for their associations with fibre and production traits. A genome-wide association study (GWAS) was performed to identify regions under selection and associated with these traits. Finally, a genomic selection (GS) analysis was performed for dagginess and dual-purpose production traits to estimate molecular breeding values (MBVs) and to determine their impact on the New Zealand dual-purpose selection index.

Heritability, genetic and phenotypic parameter estimations were performed on a flystrike case-control dataset collected over 2 years. Flystrike had a heritability of 0.37, and high genetic and phenotypic correlations with dag score and a high genetic correlation with the coefficient of variation of fibre diameter. A similar analysis was performed on an existing New Zealand sheep industry dataset of about 2 million pedigree-recorded animals born between 1990 and 2008. The heritability for dag score at 3 and 8 months (DAG3, DAG8) was 0.34 and 0.31 respectively. There were low or nil genetic and phenotypic correlations of DAG3 and DAG8 with the other standard live weight, fleece weight, reproduction and faecal egg count production traits or breech bareness, fibre and wool traits.

A GWAS performed on an industry dataset of 8,705 genotyped animals, using phenotype information on about 3 million pedigree-recorded animals, identified regions on chromosome 6 and 15 associated with DAG3 and DAG8. The lambs from the flystrike case-control dataset with SNPs imputed from 5K to 50K identified a number of immune, diarrhoea and wool/hair growth genes associated with flystrike, dag score and fibre traits in a GWAS. There were no similarities in the genes identified in the industry or case-control GWAS; however, the SNP on chromosome 15 was re-identified in the GS analysis for DAG8. The GS analysis showed that genomic predictions can be

performed for DAG3 and DAG8 and that using MBVs and modifying generation interval can increase the rate of the genetic gain of the dual-purpose index by 84% per year.

Acknowledgements

At least 4 years ago and maybe back the full 28 years, I have had the guidance of a forward thinking father who did not want to crutch another sheep. Also along the last 6 years I have had the opportunity to work with another forward thinking and supportive, first boss then supervisor, John McEwan. Together I have them to thank for one suggesting the topic and the other in moulding my scientific skills so as to answer the question "Can I breed for no dags and flystrike?"

There are a number of people who I've had the opportunity to work with and learn off during this experience. Firstly my full contingent of supervisors: Hugh Blair, Rebecca Hickson from Massey University and Tricia Johnson and John McEwan from AgResearch, Invermay. Thank you for your contributions to shaping me, my experiments, and my thesis. Also for teaching me 'past and present' tense. I would also like to thank Ken Dodds and Benoit Auvray for their guidance and invaluable help with the statistical component of the project, without them the languages, programs and methods such as R, SAS, ASReml, genomic selection and GWAS would still be foreign outer space 'things'! I've also had the pleasure of sharing my workspace with the Invermay Animal Genomics team, some have come and gone since I first started at AgResearch, but all have contributed to the great atmosphere which has been a pleasure to work in.

Secondly, the financial support of Ovita and AgResearch for funding the research and for the provision of my doctoral stipend. Also the IVABS Travel Award, Early Career Scientist Travel Bursary, and the Charles Elgar Scholarship awards which have allowed me to attend and/or present my work at 9th World Congress on Genetics Applied to Livestock Production (2010), The Applied Genomics for Sustainable Livestock Breeding Conference (2011), New Zealand Society of Animal Production (2012), and the 32nd and 33rd International Society of Animal Genetics (2010, 2012) conferences.

Thirdly, this project would not be possible without the support of the farmers that have allowed me access to their flocks via SIL and have allowed me to look at ~ 30,000 dirty and clean lamb bottoms. Also the efforts of those that collected samples and measurements of flystruck lambs for me, a not very pleasant job, were very much

appreciated. A big thank you for those that provide me a bed and great home cooking during my road trips around the country collecting measurements; these provided some of my most memorable moments during my project.

A big thank you to my family; Mum, Dad and brothers for their amazing support throughout my life, and helping me through the PhD by being a willing participant in the case-control study. Finally to Jason, who has been a rock for me for the past 3 years, putting up with my stressed out moments and encouraging me to try new things like scuba diving.

I thank you all.

Table of contents

Abstract iii

Acknowledgements v

Table of contents vii

List of tables viii

List of figures xiii

List of abbreviations xix

Introduction xxi

Chapter 1: Review of literature 25

Chapter 2: Case-control experiment: estimation of genetic parameters and a summary of two flystrike seasons 71

Chapter 3: Genetic parameters for production traits in New Zealand dual-purpose sheep, with an emphasis on dagginess 94

Chapter 4: Genetic relationships between dagginess, breech bareness and wool traits in New Zealand dual-purpose sheep 113

Chapter 5: Evaluation of sampling, paternity parentage and imputation 133

Chapter 6: Genome-wide association study: flystrike case – controls 160

Chapter 7: Genome-wide association study: SIL industry data including bare breech and fibre traits 187

Chapter 8: The impact of genomic selection on dual-purpose selection index including dagginess 228

Chapter 9: Concluding discussion 271

References 285

Appendices 318

List of tables

Table 1.1: Distribution of strikes, in 3 separate studies from the United Kingdom (UK), New Zealand (NZ), and Australia (AUS) 28

Table 1.2: Heritability (h^2) of potential indirect selection traits, and their genetic, phenotypic and between flock correlations with flystrike in Australian Merino 57

Table 2.1: Number of flystrike cases (n), incidence rate (I, %), proportion of flystrike located on the breech (B, %) and preventative insecticide use for lambs born 2009 and 2010 per flock 83

Table 2.2: Summary of strikes by area for each flystrike season 84

Table 2.3: Least square means of each trait for cases and controls 84

Table 2.4: Final mixed models and fixed effects used for individual trait analysis 85

Table 2.5: ASReml estimates of heritabilities (diagonal), phenotypic (above diagonal), genotypic (below diagonal) correlations, and phenotypic standard deviation (σ_p) ± s.e 87

Table 2.6: Insecticide use by farmers (A to K) for the prevention and treatment of flystrike. IGR: insect growth regulator, OP: organophosphates, SP: synthetic pyrethoids 88

Table 3.1: Final mixed models and fixed effects used for individual trait analysis 99

Table 3.2: Least square means and ANOVA summary for traits 100

Table 3.3: Estimates of heritabilities (diagonal), phenotypic (above diagonal), genotypic (below diagonal) correlations, and phenotypic standard deviations (σp) and repeatability estimates (last row) ± s.e 102

Table 3.4: Estimates of permanent environmental correlations (above diagonal) and genetic plus permanent environmental correlations (below diagonal) ± s.e 103

Table 3.5: Source estimates of current SIL parameters 106

Table 3.6: Summary of estimated breeding values (EBVs) results for 3 flocks run with the current Sheep Improvement Limited (SIL) and new dag models 109

Table 4.1: Summary of 2009/2010 progeny dataset before data cleaning, transformation and addition of dam information. 118

Table 4.2: Final mixed models and fixed effects including contemporary groups used for individual trait analysis 122

Table 4.3: Means and ANOVA summary for traits after cleaning, transformation and addition of dam information for WWT and LWAU 123

Table 4.4: Estimates of heritabilities (diagonal), phenotypic (above diagonal), genotypic (below diagonal) correlations, and phenotypic standard deviations (σ_p) ± s.e 126-127

Table 5.1: Sample types collected as part of the case-control experiment 136

Table 5.2: Average, standard deviation (SD), minimum (Min) and maximum (Max) amount (µg) of DNA from tissue samples successfully extracted 142

Table 5.3: Sample type and number of animals that were genotyped by the Sequenom® assay 143

Table 5.4: Genotyping success rate of the flystrike case-controls run on the 5K SNP Chip and the sires run on the 50K SNP Chip 146

Table 5.5: Average proportion of errors for each imputed SNP and imputed progeny for each program option. 148

Table 5.6: Comparison of paternity testing programs. Number of paternity assigned progeny per program on the diagonal (percentage). Below the diagonal actual values of paternity tested progeny with the same result in 2 tests (percentage), and number of total assigned progeny per SNP platform, number of unassigned progeny per test (percentage) and number of progeny that passed quality control on the 2 platforms; Sequenom (Cervus and Partial Pedigree) and 5K SNP Chip 149

Table 5.7: Computational time taken to impute each chromosome (Chr) in days hours (hrs) and minutes (mins) 151

Table 6.1: Power of analysis for given sample size (n), and parameters for a dominant model 165

Table 6.2: Power of analysis for given sample size (n), and parameters for a recessive model 166

Table 6.3: Estimate of lambda (slope), and their standard error (s.e.) of the linear regression of the observed -$\log_{10}(P)$ on the expected -$\log_{10}(P)$ (QQplot) for each trait 168

Table 6.4: The trait, number of animals (n), marker name, chromosome (Chr) and chromosome position (base pairs, bp) of the best 5 SNPs for the traits flystrike (as $\log_{10}$(scaled flystrike +1)), dag score (DAG), breech bareness (BBREECH), wool bulk (BULK) and wool length (LENGTH). The *P* value and -$\log_{10}(P)$ are shown for each marker 175

Table 6.5: The trait, number of animals (n), marker name, chromosome (Chr) and chromosome position (base pairs, bp) of the best 5 SNPs for the fibre traits; mean fibre diameter (MFD), standard deviation of MFD (FDSD), coefficient of variation of MFD (FDCV), curvature (CURV) and proportion medullation (MED%). The *P* value and -$\log_{10}$(*P*) are shown for each marker 176

Table 6.6: List of SNPs identified with immune, inflammatory response and diarrhoea roles (greyed) for the traits flystrike and dag score, which reached nominal significant threshold *P*<0.001 178

Table 6.7: List of SNPs identified with wool/hair roles (greyed) for breech bareness (BBREECH), mean fibre diameter standard deviation (FDSD) and proportion medullation (MED%) for chromosome 3, 7 and 8 which reached the nominal significant threshold *P*<0.001 179

Table 6.8: List of SNPs identified with wool/hair roles (greyed) for mean fibre diameter standard deviation (FDSD) and proportion medullation (MED%) on chromosome 13 and 23 which reached the nominal significant threshold *P*<0.001 180

Table 7.1: Summary table of the number of animals (n), heritability (h^2), reliability cut off ($0.8h^2$), average and maximum reliability (Av Rel, Max Rel) of breeding values used in the analysis 196

Table 7.2: The breed genotypic variance explained by the first 6 principle components (PC) for each trait 197

Table 7.3: Estimate of lambda (slope), and their standard error (s.e.) of the linear regression of the observed -$\log_{10}$(*P*) on the expected -$\log_{10}$(*P*) (QQplot) for each trait 199

Table 7.4: The trait, number of animals (n), marker name, chromosome (Chr) and chromosome position (base pairs, bp) of the best 3 SNPs for traits with SIL calculated BVs. The *P* value and -$\log_{10}$(*P*) are shown for each marker 200

Table 7.5: The trait, number of animals (n), marker name, chromosome (Chr) and chromosome position (base pairs, bp) of the top 3 SNPs for the wool and fibre traits. The *P* value and -$\log_{10}$(*P*) are shown for each marker 204

Table 7.6: Candidate genes within 100kbp of the best SNPs for the SIL traits, by chromosome (Chr) 205

Table 7.7: Correlation (r) of -$\log_{10}$(*P*) values between case-control GWAS (C6, Chapter 6) and industry GWAS (C7, this Chapter) 207

Table 7.8: List of significant peaks and the candidate genes within the regions. The greyed genes are discussed in more detail 210-211

Table 8.1: The year of birth of the first animals placed in the validation set and number (n) of animals in training and validation sets for each breed 234

Table 8.2: Phenotypic standard deviation (σ_p), heritability (h^2), repeatability (Rep), economic weights (EW), and number of records used for each trait in the dual-purpose selection index for selecting a 2 year old ram 239

Table 8.3: The breed genotypic variance explained by the first 6 principal components (PC) for each trait 241

Table 8.4: Variance components; additive (σ_u^2) and corrected residual (σ_e^2) variance, and the heritabilities used; (h^2) from Sheep Improvement Limited (SIL) and the effective h^2 for each trait 243

Table 8.5: The weighted correlations from G1 and G6z analysis, between mBVs and dependent variables in the 7 validation breeds 244

Table 8.6: Accuracies of the weighted correlation between MBV and dependent variable (r_A), the weighted average individual accuracy (r_I) and the combined accuracy (r_C) calculated for the 16 traits in the 7 validation breeds 245

Table 8.7: The lower limit of the 90% confidence interval for the weighted correlation between MBV and dependent variable (r_A), for all traits and all 7 validation breeds 246

Table 8.8: Proportion of variance explained by MBVs in 7 validation breeds 247

Table 8.9: The k values for traits in the 7 validation breeds 248

Table 8.10: The best 5 SNPs for the traits: weaning weight (WWT), fleece weight at 12 months (FW12), dag score at 3 and 8 months (DAG3, DAG8), with chromosome (Chr), Chr position (base pairs, bp), and the *P* value and $-\log_{10}(P)$ value 249

Table 8.11: The correlation (r), number of SNPs in top 20 for both analyses (Top20) and summary of the $-\log_{10}(P)$ values from the 2 analyses (GWAS and GS) 250

Table 8.12: Estimate of lambda (slope), and their standard error (s.e.) of the linear regression of the observed $-\log_{10}(P)$ on the expected $-\log_{10}(P)$ (QQ plot) for each trait 254

Table 8.13: The heritability (h^2), repeatability (rep) and genomic selection accuracy (GS acc), effective number of progeny (n prog) and their single trait selection accuracy (Seln acc) estimated from the breed combined-accuracy (r_C) and the individual accuracy (r_I) 256

Table 8.14: Response per selection round per trait (change (Δ) in units and dollars ($)) and accuracies, and overall index (standard deviation (SD) index in dollars) and accuracy for a dual-purpose Romney index 258

Table 8.15: Response per selection round (SD index) and accuracy (acc), conversion factor (i/L, selection intensity / generation interval), and rate of genetic gain (ΔG), for each scenario 259

Table 8.16: Number of animals to genotyped to reach accuracy of ~0.70 if have own measurement or 126 progeny measurements 262

List of figures

Figure 1.1: Images of the 4 main fly species with larvae inset (not to scale): *L. cuprina* (A), *L. sericata* (B), *C. stygia* (C) and *C. rufifacies* (D). The two *Lucilia* species are separated by the green colouration of the fore femora of *L. cuprina* and the increased number of hairs of the humeral calli and notopleuron area of *L. sericata* (Lang et al., 2001) 28

Figure 1.2: Classification of sheep body regions in strike identification. 1: Breech, 2: Body, 3: Shoulders, 4: Head, 5: Belly (and Pizzle in males), 6: Foot. Modified from French et al.(1995) 29

Figure 1.3: Lifecycle of the fly 30

Figure 1.4: Dag scoring system: 0 no dags to 5 most daggy. As proposed by Sheep Improvement Limited (www.sil.co.nz/Files/Tech-Notes/DagScoreV2.aspx) 46

Figure 1.5: Breech bareness scoring system: 1: fleece cover extends to margins of anus, to 5: an extensive bare area either side of the anus. As proposed by Sheep Improvement Limited (www.sil.co.nz/getdoc/a9bb121b-2016-4b34-a163-399ae4b28471/Doc-ID-000010-GW-Bare-Points-sheep.aspx) 52

Figure 2.1: The 6-month flystrike season overlapped by the timing of weaning and live weight at 6 months (LW6) recording, when the majority of flystruck lambs were observed and measured by the farmer 74

Figure 2.2: Map of experimental sites. Properties involved in case-control test; commercial farmers (●) and Breeders (◆) 75

Figure 2.3: Classification of sheep body regions in strike identification. 1: Breech, 2: Body, 3: Shoulders, 4: Head, 5: Belly (and Pizzle in males), 6: Foot. Modified from French et al. (1995) 76

Figure 2.4: Total rainfall (mm) from July 2009 to 30th June 2010 (top left), and over the flystrike season, October 2009 to 31st March 2010 (top right), July 2010 to 30th June 2011 (bottom left), and over the flystrike season, October 2010 to 31st March 2011 (bottom right) 81

Figure 2.5: Average maximum (red line) and minimum (blue line) temperature (°C) for the 2009/10 season (left) and 2010/11 season (right) across all farms 82

Figure 3.1: Plot of standard deviation versus mean for contemporary groups (flock-year-sex) for new dag score at 3 months model (A) and current Sheep Improvement Limited dagginess model (B), after scaling for all 3 flocks tested 110

Figure 3.2: Plot of current Sheep Improvement Limited dagginess EBV value (x axis) versus new dag score at 3 months (DAG3; A), and at 8 months (DAG8; B) EBV values (y axis) for flock A 111

Figure 4.1: Map of experimental sites; properties involved in the progeny test (▲) 116

Figure 5.1: Frequency distribution of DNA concentrations (µg) from tissue extractions. Arrows indicate 1 spleen sample at 322 µg and 389 µg. 142

Figure 5.2: Gel picture of DNA from TypiFix™ ear tissue extractions. Samples were run in duplicate, on a 1.5% agarose gel with 1kb plus ladder 143

Figure 5.3: Frequency distribution of call rate (%) of all samples genotyped through the Sequenom® assay derived from the 115 SNPs genotyped 144

Figure 5.4: Frequency distribution of call rate (%) of all 115 Sequenom® SNPs across all animals 144

Figure 5.5: Minor allele frequency (MAF) of SNPs genotyped over the progeny on the 5K SNP Chip. The average MAF noted by arrow 146

Figure 5.6: Minor allele frequency (MAF) of SNPs genotyped over the sires on the 50K SNP Chip. The average MAF noted by arrow 147

Figure 5.7: Density graph of error rates for Option1: animals (A), SNPs (B), and Option 5: animals (C), SNPs (D). Option 1: LinkPHASE, DAGPHASE and BEAGLEv2.1.3: 287 progeny and 129 sires including pedigree information. Option 5: BEAGLEv3.0.4: 227 progeny and 129 sires ('pairs' file) plus the 60 unrelated progeny and 1130 unrelated animals (unrelated file) 148

Figure 5.8: Comparison of error rate for the imputed progeny per SNP with the minor allele frequency in the imputed data set for each SNP (0-1 line, red). Option 1: LinkPHASE, DAGPHASE and BEAGLEv2.1.3: 287 progeny and 129 sires including pedigree information (left). Option 5: BEAGLEv3.0.4: 227 progeny and 129 sires ('pairs' file) plus the 60 unrelated progeny and 1130 unrelated animals (unrelated file) (right) 149

Figure 5.9: Minor allele frequencies (MAF) for unrelated best guess (A), unrelated genotype probabilities (B), pairs (C), and the correlation between best guess and genotype probabilities MAF 152

Figure 5.10: Diagonal values of the G matrix using the genotype probabilities for the 253 unrelated progeny (left of red line), and the pairs phased genotypes for the paternity assigned progeny (right of red line) 153

Figure 5.11: Diagonal values of the G matrix using the best guess genotypes for the 253 unrelated progeny (left of red line), and the pairs phased genotypes for the paternity assigned progeny (right of red line) 154

Figure 5.12: The comparison of the diagonal values of the G matrix from the genotype probabilities and the best guess genotypes. Including the paternity pairs phased genotypes 154

Figure 6.1: Plot of the power for a dominant (blue) and recessive (red) model against the heritability (h^2) of the QTL 166

Figure 6.2: Plot of the first 4 principal components (PC) for flystrike case and control derived from genotype information. Coloured by farmer, see Chapter 2. Farmer A (orange), B (yellow), C (Dark blue), D (green), E (blue), F (pink), G (red), H (light grey), I (grey) and K (brown) 167

Figure 6.3: QQ plots for flystrike (left) and dag score (right) $-\log_{10}(P)$ values. The 0-1 line is in black and the linear regression in red 169

Figure 6.4: Graph of $-\log_{10}(P)$ values of SNPs for flystrike, ordered on ovine genome v1 map, P<0.001 (red dash line) 170

Figure 6.5: Graph of $-\log_{10}(P)$ values of SNPs for dag score, ordered on ovine genome v1 map, P<0.0001 (solid red line), P<0.001 (red dash line) 170

Figure 6.6: Graph of $-\log_{10}(P)$ values of SNPs for breech bareness score, ordered on Ovine genome v1 map, P<0.0001 (solid red line), P<0.001 (red dash line) 171

Figure 6.7: Graph of $-\log_{10}(P)$ values of SNPs for wool bulk, ordered on ovine genome v1 map, P<0.001 (red dash line) 171

Figure 6.8: Graph of $-\log_{10}(P)$ values of SNPs for wool length, ordered on ovine genome v1 map, P<0.0001 (solid red line), P<0.001 (red dash line) 172

Figure 6.9: Graph of $-\log_{10}(P)$ values of SNPs for mean fibre deviation, ordered on ovine genome v1 map, P<0.001 (red dash line) 172

Figure 6.10: Graph of $-\log_{10}(P)$ values of SNPs for standard deviation of mean fibre deviation, ordered on ovine genome v1 map, P<0.001 (red dash line) 173

Figure 6.11: Graph of $-\log_{10}(P)$ values of SNPs for coefficient of variation of mean fibre diameter score, ordered on ovine genome v1 map, P<0.0001 (solid red line), P<0.001 (red dash line) 173

Figure 6.12: Graph of -$\log_{10}(P)$ values of SNPs for curvature, ordered on ovine genome v1 map, P<0.0001 (solid red line), P<0.001 (red dash line) 174

Figure 6.13: Graph of -$\log_{10}(P)$ values of SNPs for proportion medullation score, ordered on ovine genome v1 map, P<0.00001 (black dash line), P<0.0001 (solid red line), P<0.001 (red dash line) 174

Figure 6.14: Plot of the power for a dominant (blue/green) and recessive (red/purple) model when 1,000 (diamond) or 667 (triangle) cases are collected against the heritability (h^2) of the QTL 182

Figure 7.1: Minor allele frequency (MAF) for the Ovine 50K SNP chip, following removal, for quality control checks of the 8579 animals retained for GWAS. Arrow indicates mean MAF 196

Figure 7.2: The first 4 principal components (PC) calculated from the G matrix for weaning weight. Romney (blue), Coopworth (green), Texel (yellow), Perendale (purple) and others (grey) 198

Figure 7.3: QQ plot for weaning weight (A), fleece weight at 12 months (B), dag score at 3 months (C) and 8 months (D) -$\log_{10}(P)$ values. The 0-1 line is in black and the linear regression in red 199

Figure 7.4: Manhattan plot of -$\log_{10}(P)$ values of SNPs for weaning weight. Ordered on the ovine genome v2 map, P<0.001 (red dash line) 201

Figure 7.5: Manhattan plot of -$\log_{10}(P)$ values of SNPs for fleece weight at 12 months. Ordered on the ovine genome v2 map 201

Figure 7.6: Manhattan plot of -$\log_{10}(P)$ values of SNPs for dag score at 3 months. Ordered on the ovine genome v2 map, P<0.0001 (solid red line), P<0.001 (red dash line) 202

Figure 7.7: Manhattan plot of -$\log_{10}(P)$ values of SNPs for dag score at 8 months. Ordered on the ovine genome v2.0 map, P<0.0001 (solid red line), P<0.001 (red dash line) 202

Figure 7.8: Plot of -$\log_{10}(P)$ values for individual SNPs from the case-control GWAS (y axis) for dag score versus the industry GWAS (x axis) for dag score at 3 months (DAG3, left) and at 8 months (DAG8, right). The 0-1 line is plotted in red 207

Figure 7.9: Animals clustered on the basis of principal components 1 and 2 (PC 1, PC 2). Romney and Arapawa breeds shown in blue and green respectively 208

Figure 7.10: A Manhattan plot of the moving window of 5 (WIN5) F_{ST} values between the Arapawa and high dag score breeding value Romneys. Ordered on the ovine genome v2 map, WIN5 F_{ST} = 0.3 (red dash line) 209

Figure 7.11: Schematic of significant region on chromosome 2, SNPs (diamonds), scaffold (green lines), and bovine ref sequences (black line) 211

Figure 7.12: Schematic of significant region on chromosome 10, SNPs (diamonds), scaffold (green lines) and bovine ref sequences (black line) 211

Figure 7.13: Schematic of significant region 1 on chromosome 13, SNPs (diamonds), scaffold (green lines) and bovine ref sequences (black line) 212

Figure 7.14: Schematic of significant region 2 on chromosome 13, SNPs (diamonds), scaffold (green lines) and bovine ref sequences (black line) 212

Figure 7.15: Schematic of significant region 3 on chromosome 13, SNPs (diamonds), scaffold (green lines) and bovine ref sequences (black line) 213

Figure 7.16: Schematic of significant region 4 on chromosome 13, SNPs (diamonds), scaffold (green lines) and bovine ref sequences (black line) 213

Figure 7.17: Schematic of significant region (45.32 to 46.86 Mbp) on chromosome X, SNPs (diamonds), scaffolds (green lines) and bovine ref sequences (black line) 214

Figure 7.18: Comparison of the F_{ST} values from the Romney versus Arapawa and the average F_{ST} value from the ovine HapMap study (Kijas et al., 2012a). The 0-1 line is plotted in red. 215

Figure 7.19: Plot of the QQ slope against the selection intensity for each trait; weaning weight direct and maternal (♦, ■), live weight at 8 months (▲), fleece weight at 12 months (+), dag at 3 and 8 months and breech bareness (●), wool bulk and length (▬) and fibre traits (♦) 217

Figure 8.1: The first 4 principal components (PC) calculated from all animals for weaning weight. Romney (blue), Coopworth (green), Texel (yellow), Perendale (purple) and others/composites (grey) 242

Figure 8.2: Manhattan plot of $-\log_{10}(P)$ values of SNPs for weaning weight. Ordered on the ovine genome v2 map, P<0.0001 (solid red line), P<0.001 (red dash line) 251

Figure 8.3: Manhattan plot of $-\log_{10}(P)$ values of SNPs for fleece weight at 12 months. Ordered on the ovine genome v2 map, P<0.0001 (solid red line), P<0.001 (red dash line) 251

Figure 8.4: Manhattan plot of -$\log_{10}(P)$ values of SNPs for dag score at 3 months. Ordered on the ovine genome v2 map, P<0.0001 (solid red line), P<0.001 (red dash line) 252

Figure 8.5: Manhattan plot of -$\log_{10}(P)$ values of SNPs for dag score at 8 months. Ordered on the ovine genome v2 map, P<0.0001 (solid red line), P<0.001 (red dash line) 252

Figure 8.6: Comparison of GWAS and GS -$\log_{10}(P)$ values for weaning weight (A), weaning weight maternal (B), live weight at 8 months (C), fleece weight at 12 months (D), dag score at 3 (E) and 8 (F) months. The 0-1 line also plotted (red line) 253

Figure 8.7: Comparison of dag score at 3 and 8 months (DAG3, DAG8) -log10(P) values from GS analysis. The 0-1 line also plotted (red line) 254

Figure 8.8: QQ plot for weaning weight (A), fleece weight at 12 months (B), dag score at 3 months (C) and 8 months (D) –$\log_{10}(P)$ values. The 0-1 line is in black and the slope in red 255

Figure 8.9: Comparison of the observed breed combined-accuracies (r_C) versus the theoretical accuracy using Goddard's equation for each trait, and for the breeds Romney (green), Coopworth (blue) and Perendale (purple). The line of equality is in grey 261

List of abbreviations

AFEC	adult faecal egg count
AFW	adult fleece weight
BBREECH	breech bareness
BULK	wool bulk
BV	breeding value
chr	chromosome
CURV	curvature
CW	carcass weight
DAG3	dag score at 3 months of age
DAG8	dag score at 8 months of age
dNTP	deoxynucleoside triphosphate
EWT	adult ewe live weight
FDCV	coefficient of variation of mean fibre diameter
FDSD	standard deviation of mean fibre diameter
FE	facial eczema
FEC1	faecal egg count in summer
FEC2	faecal egg count in autumn
FW12	fleece weight at 12 months
GBLUP	genomic best linear unbiased prediction
GBV	genomic breeding value
GC	genotype call score
GWAS	genome-wide association study
GS	genomic selection
HapMap	haplotype map
ISGC	International Sheep Genomic Consortium
IWTO	International Wool Textile Organisation
LENGTH	wool length

LFW	lamb fleece weight
LW6	live weight at 6 months
LW8	live weight at 8 months
MAF	minor allele frequency
MBV	molecular breeding value
MED%	proportion of medullated fibres
MFD	mean fibre diameter
MT-EBV	multi trait-estimated breeding value
NEM1	nematodirus egg counts in summer
NEM2	nematodirus egg counts in autumn
NLB	number of lambs born
NZWTA	New Zealand Wool Testing Authority
OFDA100	Optical Fibre Diameter Analyser 100
OMIM	Online Mendelian Inheritance in Man
PC	principal components
PETA	People for the Ethical Treatment of Animals
PEV	prediction error variance
QQ	quantile-quantile
RRS	reduced representational sequencing
SAP	shrimp alkaline phosphate
SIL	Sheep Improvement Limited
ST-EBV	single trait-estimated breeding value
SURV	lamb survival
SURVm	survival maternal
TBE	Tris base boric acid and ethylenediaminetetra-acetic acid buffer
TBV	true breeding value
WWT	weaning weight at 3 months
WWTm	maternal weaning weight