

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.

Genome-wide association study for stature 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

Mun Ee, Tan

2013

Abstract

The objective of this thesis was to perform a genome-wide association study (GWAS) to identify single nucleotide polymorphism (SNP) associated with stature in New Zealand dairy cattle. The phenotype data set used for this study contained the animal key, sire code of the bull, birth date, breed code, proportion of Holstein-Friesian genes, proportion of Jersey genes, percentage of North American Holstein genes, estimated breeding values (EBV) for live weight and stature and their reliabilities of 3140 bulls. The genotype data set contained the genotype of 692,598 SNPs for every bull and another file contained the name and position of the SNPs.

The GWAS was performed on Holstein-Friesian, Jersey and Holstein-Friesian \times Jersey crossbred bulls using PLINK software version 1.07. Stature EBV was used as the phenotype. The phenotypes were adjusted for percentage of Holstein-Friesian, Jersey, North American Holstein genes and year of birth using multiple regression. Manhattan plots and multi Manhattan plots of P -values adjusted to genomic control against the chromosomes were plotted to identify top SNPs with the most significant P -values above the significant threshold line.

Based on the top 50 SNPs according to the P -value, this study identified nine chromosomes or BTA in the HF population with SNPs significantly associated with stature, BTA2, 3, 4, 5, 6, 11, 12, 14 and 24. SNPs with significant effect on stature were detected in six chromosomes, BTA9, 10, 12, 18, 19 and 25 in the JE population while the SNPs determined to be significantly associated with stature were located on eleven chromosomes, BTA1, 3, 4, 5, 7, 9, 10, 14, 18, 22 and 24 in the XB population. Several SNPs located above the suggestive threshold in the Manhattan plots were also inspected and kept in view for future studies.

The results from this study suggest that the highlighted SNPs with significant associations to stature can serve as candidate SNPs for further investigation to determine the regions of QTLs and ultimately the exact genes that affect stature with other correlated traits in dairy cattle.

Acknowledgements

This section is to acknowledge and express my gratitude for the help, assistance and encouragements given by many people who made this completion of thesis possible. First of all, I thank God for this given opportunity to study this program in Massey University, and for providing everything, from tangible support to kind people I met here.

Secondly, a special thanks to my supervisor Associate Professor Dr. Nicolas Lopez-Villalobos. Offering me the opportunity to handle this study, giving excellent advices with statistical analysis and spending his time to adjusting the data models were just few endeavours he had done for me throughout this year. His invaluable help is much appreciated. To my other supervisor Dr. Richard Spelman from Livestock Improvement Co., his comments and reviews on the results further enhanced this thesis. His help in providing the data set for this study is also acknowledged.

Not to forget Richard Sherlock, Frances Creagh and Kathryn Sanders from Livestock Improvement Corporation either. Their advices and support with the PLINK program and Manhattan plots package in R have helped me very much in this thesis.

My work partner, Sibyl Chan is an irreplaceable friend especially these two years of study. Her help has also made this completion of thesis possible. Her humour and jokes made this period of time easier to endure too. Acknowledging also my fellow post-graduates and friends at Massey, I cannot make it through this time without them. Especially Judy Lau, a post-graduate fellow and an awesome friend.

To my parents and siblings, thank you for believing in me and support me here in New Zealand to further my studies. Thank you for providing the financial, physical, emotional and spiritual support. I cannot possibly have performed this stint without you all.

Last but not least, I want to give thanks to Zachary, my boyfriend for his constant love, encouragements and prayers especially during the final phase of my thesis.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iii
List of tables and figures	vi
List of abbreviations	viii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	3
2.1 Introduction	3
2.2 Linear score of stature	4
2.3 Economic importance of stature	5
2.4 Heritability of stature	7
2.5 Correlations between stature and other traits of economic importance	9
2.6 Finding quantitative trait loci for stature	15
2.6.1 Definition of quantitative trait loci	15
2.6.2 Genetic markers used to detect quantitative trait loci	16
2.6.3 Linkage disequilibrium	17
2.7 Genome-wide association studies	18
2.7.1 Genetic markers associated with stature	19
2.7.2 Genetic markers associated with live weight	21
2.8 Genomic selection	22
2.9 Genome imputation	24
2.10 Impacts of genomic selection	26
CHAPTER 3: MATERIALS AND METHODS	30
3.1 Data	30
3.1.1 Phenotypes	30

3.1.2	Genotypes	30
3.2	Statistical analysis	31
3.3	Genome-wide association studies	32
3.3.1	Adjustment of phenotype	32
3.3.2	Estimation of single nucleotide polymorphism effects	32
3.3.3	Manhattan plots	34
CHAPTER 4: RESULTS		35
4.1	Descriptive statistics	35
4.2	Genome-wide association analysis	42
4.2.1	Manhattan plots	42
4.3	Top 50 single nucleotide polymorphisms	51
CHAPTER 5: DISCUSSION		55
5.1	Correlation between stature and live weight	55
5.1.1	Effects of percentage of North American Holstein genes on stature and live weight estimated breeding value	56
5.2	Effects on feed conversion efficiency when selecting for live weight	57
5.3	Identification of significant single nucleotide polymorphisms affecting stature	58
5.3.1	Significant single nucleotide polymorphisms	59
5.3.2	Reported significant single nucleotide polymorphisms from literature	63
5.4	Differences in genome-wide association studies	66
5.5	Threshold probability values	67
5.6	Genomic selection	68
CHAPTER 6: CONCLUSIONS		70
REFERENCES		71

LIST OF TABLES AND FIGURES

CHAPTER 2:

Table 2.1	Estimates of heritability and standard errors for stature in beef and dairy cattle.....	8
Table 2.2	Estimates of genetic parameters for stature and traits of economic importance in dairy cattle.....	10
Table 2.3	Genetic markers linked with stature in beef and dairy cattle.....	19
Table 2.4	Genetic markers linked with live weight in beef and dairy cattle.....	21

CHAPTER 4:

Table 4.1	Descriptive statistics for EBV for live weight and stature of HF, JE, XB and OT bulls.....	35
Table 4.2	Estimates of regression coefficients of EBV for live weight on HF_{pct} , JE_{pct} and NAH_{pct} and year of birth.....	40
Table 4.3	Estimates of regression coefficients of EBV for stature on HF_{pct} , JE_{pct} and NAH_{pct} and year of birth.....	40
Table 4.4	SNPs located above the genome-wide significant and suggestive threshold for HF population.....	50
Table 4.5	SNPs located above the genome-wide significant and suggestive threshold for JE population.....	50
Table 4.6	SNPs located above the genome-wide significant and suggestive threshold for XB population.....	51
Table 4.7	Top 50 SNPs according to the <i>P</i> -value for HF population.....	52
Table 4.8	Top 50 SNPs according to the <i>P</i> -value for JE population.....	53
Table 4.9	Top 50 SNPs according to the <i>P</i> -value for XB population.....	54
Figure 4.1	Average live weight EBV per year of birth for HF, JE and XB population.....	37
Figure 4.2	Average stature EBV per year of birth for HF, JE and XB population.....	38
Figure 4.3	Correlation between stature and live weight EBV of bulls from all breeds.....	39

Figure 4.4	Scatter plot of stature EBV against NAH_{pct}	41
Figure 4.5	Genome-wide Manhattan plot of all associated SNP with adjusted stature EBV for HF.....	44
Figure 4.6	Genome-wide Manhattan plot of all associated SNP with adjusted stature EBV for JE.....	45
Figure 4.7	Genome-wide Manhattan plot of all associated SNP with adjusted stature EBV for XB.....	46
Figure 4.8	Multi Manhattan plots for BTA2, 3, 4, 5, 6, 7, 8, 11, 12, 14 and 24 in HF population.....	47
Figure 4.9	Multi Manhattan plots for BTA5, 6, 9, 10, 12, 18, 19, 23 and 25 in JE population.....	48
Figure 4.10	Multi Manhattan plots for BTA1, 3, 5, 7, 9, 10, 14, 22 and 24 in XB population.....	49

LIST OF ABBREVIATIONS

- ATP : Adenosine triphosphate
- BLUP : Best linear unbiased prediction
- BTA : *Bos Taurus* autosome
- BV : Breeding value
- BW : Breeding worth
- cm : Centimetres
- DNA : Deoxyribonucleic acid
- EBV : Estimated breeding value
- FCE : Feed conversion efficiency
- GEBV : Genomic estimated breeding value
- GWAS: Genome-wide association study
- HF : Holstein-Friesian
- HF_{pct} : Holstein-Friesian genes percentage
- ID : Identification
- IGF1 : Insulin-like growth factor 1
- JE : Jersey
- JE_{pct} : Jersey genes percentage
- kg : Kilograms
- LD : Linkage disequilibrium
- Mbp : Mega base pairs
- NAH : North American Holstein genes
- NAH_{pct}: North American Holstein genes percentage
- Others : OT
- PLAG1*: Pleiomorphic adenoma gene 1
- SNP : Single nucleotide polymorphism
- TOP : Traits other than production

QTL : Quantitative trait loci

QTN : Quantitative trait nucleotide

XB : Holstein-Friesian and Jersey crossbred