University of Alberta

Patterns of Genomic Variation and Whole Genome Association Studies of Economically Important Traits in Cattle

by

Honghao Li

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Animal Science

Department of Agricultural, Food and Nutritional Science

Fall 2012
Edmonton, Alberta

Abstract

Functionally important genetic variations in cattle have great potential as tools to increase agricultural production once they are identified through association studies of phenotype and genotype. The objectives of this thesis were: 1) to characterize the genome-wide patterns of linkage disequilibrium (LD) and haplotype blocks using the genotyping data obtained from the BovineSNP50 BeadChip (50K) assay; 2) to map functionally important genomic regions and genes via genome-wide association studies for economically important traits in dairy and beef cattle; and 3) to investigate the genetic contribution of epistatic QTL to quantitative traits in cattle. First, genome-wide LD and haplotype block maps were constructed for both dairy and hybrid beef populations, and different genome-wide and regional patterns of haplotype blocks between dairy and beef cattle were compared. Next, whole-genome association studies for several economically important traits were performed. Two approaches, single marker regression and Bayesian regression, were used to detect and fine map QTL for five milk production traits and eight beef carcass traits. Both methods revealed QTL regions and functional candidate genes and their networks in dairy and beef cattle. In addition to the novel QTL regions identified, many of the large effect QTL regions overlap with QTL reported in previous studies, and there were many concordances between the single-marker and Bayesian approaches. Following the one-dimensional genome scan for QTL, genome-wide pair-wise epistatic QTL analyses were carried out for dairy traits by using an empirical Bayes method. We

identified strong additive-by-additive (A × A) epistasis with considerable contribution to the phenotypic variation of analyzed traits. We also observed that epistasis plays different roles in the genetic architectures of different types of traits. The identified A × A epistatic QTL may need to be considered in future breeding programs after further validation studies. Overall, this study will contribute to a better understanding of the genetic basis of phenotypic variation of economically important traits in cattle and identifies markers and genes which may be useful for genetic improvement programs.

Acknowledgments

First of all, I sincerely express my gratitude to my supervisors Drs. Paul Stothard and Zhiquan Wang for their continuous guidance and support throughout my doctoral program. I will never forget that they gave me the opportunity to join the Livestock Gentec research group at the University of Alberta and enter the fields of quantitative genetics, genomics, and bioinformatics. I would also like to acknowledge my supervisory committee members Drs. Stephen Moore and Peng Zhang for their constructive and valuable advice during various stages of the study. Special thanks to Dr. Moore for his continuous encouragement and suggestions not only on my research projects but also on my career development.

I would like to express my sincere thanks to Drs. Rong-Cai Yang and Xin-Sheng Hu for their help in guiding me into the field of population genetics. My grateful acknowledgements to our collaborators, Drs. Flavio Schenkel and Mehdi Sargolzaei from University of Guelph, Dr. Shizhong Xu from University of California, Riverside, and Dr. Dorian Garrick from Iowa State University-their useful suggestions and constructive criticisms made this work a success. Working with them was a great learning experience that I will benefit from throughout my life.

I also want to acknowledge the great research environment and exchange opportunities provided by the Livestock Gentec center. Current and previous members within this center have contributed to my success in this project by providing valuable comments and assistance throughout these years, especially: Dr. Graham Plastow, Dr. Changxi Li, Dr. Guoqing Tang, Dr. Jason Grant, Dr.

Donald Nkrumah, Dr. Jennifer Thomson, Dr. Elisa Marques, Dr. Fidalis Denis Mujibi, Yan Meng and Jackson Mah. I also owe a great deal of appreciation to the Department of Agricultural, Food and Nutrition Science for its facilities and staff members especially Jody Forslund from whom I benefited greatly. I also want to thank all the present and past members in my office at College Plaza for always maintaining a friendly and lively work atmosphere. I gratefully acknowledge the financial support from the University of Alberta.

I would also like to extend my gratitude to all my friends for their continuous encouragement and support throughout these years. Finally, my heartfelt gratitude goes to my parents who have supported me throughout my whole life and to all my family members. Special thanks are due to my dearest husband Wei for his moral support and caring love.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABCG2	ATP-binding cassette, sub-family G, member 2
ADAMTS14	ADAM metallopeptidase with thrombospondin type 1 motif, 14
ADFP	Adipose differentiation-related protein
ANOVA	Analysis of variance
ATP1B2	ATPase, Na+/K+ transporting, beta 2 polypeptide
BLAST	Basic local alignment search tool
BLUP	Best linear unbiased prediction
bp	Base pair(s)
BTA	*Bos taurus* autosome
CABF	Carcass average backfat
CGF	Carcass grade fat
CLCN3	Chloride channel 3
CLMY	Carcass lean meat yield
cM	Centimorgan
CREA	Carcass ribeye area
CRH	Corticotropin releasing hormone
CSN3	Casein kappa
CWT	Carcass weight
CYP11B1	Cytochrome P450, family 11, subfamily B, polypeptide 1
DGAT1	Diacylglycerol O-acyltransferase 1
DNA	Deoxyribonucleic acid

DYD	Daughter yield deviations
EBVs	Estimated breeding values
EHH	Extended haplotype homozygosity
EM	Expectation-maximization
EXT1	Exostosin 1
FABP4	Fatty acid binding protein 4, adipocyte
FAM13A1	Family with sequence similarity 13, member A1
FASN	Fatty acid synthase
FDR	False discovery rate
FP	Fat percentage
FY	Fat yield
GALNT3	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3
GH1	Growth hormone 1
GHR	Growth hormone receptor
GOLT1B	Golgi transport 1B
GPT	Glutamic-pyruvate transaminase (alanine aminotransferase)
GS	Genomic selection
GWAS	Genome-wide association studies
HERC1	HECT and RLD domain containing E3 ubiquitin protein ligase family member 1
HERC6	HECT and RLD domain containing E3 ubiquitin protein ligase family member 6

HWE	Hardy-Weinberg equilibrium
IAPP	Islet amyloid polypeptide
IBD	Identity by descent
IGF1	Insulin like growth factor-1
IL8	Interleukin 8
kb	Kilo base pairs
K-S	Kolmogorov-Smirnov
LAP3	Leucine aminopeptidase 3
LCORL	Ligand dependent nuclear receptor corepressor-like
LD	Linkage disequilibrium
LDLA	Linkage disequilibrium and linkage analysis
LEP	Leptin
MACE	Multiple across-country evaluation
MAF	Minor allele frequency
MAPK15	Mitogen-activated protein kinase 15
MAS	Marker assisted selection
Mb	Mega base pairs
MCMC	Markov chain Monte Carlo
MED28	Mediator complex subunit 28
ML	Maximum likelihood
MSTN	Myostatin
MY	Milk yield
NCAPG	Non-SMC condensin I complex, subunit G

NCBI	National Center for Biotechnology Information
OLR1	Oxidized low density lipoprotein receptor 1
OPN	Osteopontin
PCBD1	Pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha
PIC	Polymorphic information content
PKD2	Polycystic kidney disease 2
PP	Protein percentage
PPARGC1A	Peroxisome proliferator-activated receptor gamma, coactivator 1 alpha
PPIB	Peptidylprolyl isomerase B (cyclophilin B)
PRL	Prolactin
PY	Protein yield
QTL	Quantitative trait loci
REML	Restricted maximum likelihood
RPL8	Ribosomal protein L8
SCD	Stearoyl-CoA desaturase
SGPL1	Sphingosine-1-phosphate lyase 1
SLC29A3	Solute carrier family 29 (nucleoside transporters), member 3
SNP	Single nucleotide polymorphism
SNX1	Sorting nexin 1
SPP1	Secreted phosphoprotein 1
SSE	Residual sums of squares or sum of squared errors of prediction

TG	Thyroglobulin
TGS1	Trimethylguanosine synthase 1
UBF	Ultrasound backfat
UMAR	Ultrasound marbling
UREA	Ultrasound ribeye area

Chapter 1. General Introduction

1.1. Introduction

Animal agriculture is an important source of food for humans-animals provide one-sixth of the food energy and more than one-third of the protein consumed globally (BRADFORD 1999). Cattle production has been an essential part of agriculture for more than 300 years. Modern cattle breeds were created through thousands of years of selective breeding, some of which are specialized for milk production and others for beef production. These breeds can serve as unique resources for understanding the genetic basis of phenotypic variation of economically important traits (ANDERSSON and GEORGES 2004).

Breeding programs aimed at improving milk and beef quantity as well as quality have been in place for many years using phenotypic measurements from pedigreed herds, and have resulted in significant genetic gains (DEKKERS and HOSPITAL 2002). In dairy cattle, selection of dairy bulls for their superior milking ability has been achieved due to the development of three technologies: the systematic milk recording system in 1990, the artificial insemination and progeny test in 1950, and the adaptation of an advanced statistical genetic evaluation technique known as the best linear unbiased prediction (BLUP) to estimate the genetic merit of individual animals in 1970 (BROTHERSTONE and GODDARD 2005). These together resulted in a nearly doubled milk production per lactation of Holstein cows during the past 40 years (DEKKERS and HOSPITAL 2002). However, up until recently selection has been based on phenotype only, which

hinders the rate of genetic improvement for traits that are difficult and expensive to measure, traits measured in one sex, traits that are measured after animals are slaughtered, and traits that have a low heritability. In addition, phenotype based selection has been done without knowledge of the genetic architecture of the selected traits. Most of the economically important traits targeted in animal breeding programs are complex traits with their phenotypic variation determined by a large number of genes together with the environmental effects. This complexity will make it very difficult to fully understand the genetic mechanisms underlying the processes of milk and beef production.

Knowledge of the variation in the bovine genome sequence is key to understanding the genetic basis of phenotypic variation. Genetic studies on the association of genotype and phenotype have the potential to increase agricultural production (DEKKERS and HOSPITAL 2002; GODDARD and HAYES 2009). With the completion of the bovine genome sequence assembly and the availability of high density single nucleotide polymorphisms (SNPs) (ELSIK *et al.* 2009; MATUKUMALLI *et al.* 2009), quantitative trait loci (QTL) mapping studies and genome-wide association studies (GWAS) make the identification of the mutations that underlie the economically important traits a realistic and important endeavor. Knowledge of QTL, genetic markers, genes, and biological pathways could be incorporated into breeding programs to accelerate genetic improvement of dairy and beef herds through marker assisted selection (MAS) (DEKKERS 2004; DEKKERS and HOSPITAL 2002). The development of DNA chips that can

simultaneously genotype tens of thousands of SNPs has opened up the era of genomic selection (GS) (HAYES *et al.* 2009a; MEUWISSEN *et al.* 2001).

The last two decades have seen a dramatic increase in the number of livestock QTL mapping studies. However, most of the mapped QTL were located within large confidence intervals with tens of centimorgans that may contain thousands of genes (GRAPES *et al.* 2004), and only a few genes with conclusive effects have been identified (CASAS *et al.* 1998; COHEN-ZINDER *et al.* 2005; GRISART *et al.* 2002; SETOGUCHI *et al.* 2009). The major genes that regulate milk production in dairy cattle and meat production and quality traits in beef cattle are still largely unknown, as are the gene networks and pathways involved.

In addition to the undiscovered major genes, neglected QTL interactions (epistasis) might be another reason for the current small proportion of explained genetic variance from reported marginal (additive and dominance) effects of QTL for most of the complex traits (HAYES and GODDARD 2010). Epistasis has been demonstrated to play a prominent role in the genetic architecture of complex traits (CARLBORG and HALEY 2004; SHAO *et al.* 2008) and is being characterized by the recently developed methods of epistatic QTL mapping in different livestock species (ANKRA-BADU *et al.* 2010; CARLBORG *et al.* 2003; GROSSE-BRINKHAUS *et al.* 2010; UEMOTO *et al.* 2009; XU and JIA 2007).

QTL mapping studies rely on the extent of linkage disequilibrium (LD) between the QTL and the markers, thus quantifying the extent of LD becomes the first important step in determining the number of markers and samples required

for an association study. The extent and patterns of LD are not only important for the design and application of association studies in specific cattle populations, but can also provide information on the evolutionary and selection history of a population and on evolutionary forces shaping certain chromosome regions (ARDLIE *et al.* 2002; GOLDSTEIN and WEALE 2001; SLATKIN 2008). Archaeological and genetic data suggest that the domestication and artificial selection of *Bos taurus* occurred approximately 8,000 to 10,000 years ago in the Near East (BRADLEY *et al.* 1998; BRUFORD *et al.* 2003; DIAMOND 2002). No or little differentiation between different cattle breeds occurred until recent specialization for milk or beef production using strong artificial selection (HAYES *et al.* 2009b). In cattle, the patterns of LD have been used to detect historical bottleneck signatures and to infer effective population sizes during breed domestication and formation (DE ROOS *et al.* 2008; GIBBS *et al.* 2009; HAYES *et al.* 2003). In addition, the characterization of block-like patterns of LD, called haplotype blocks, may facilitate the design of association studies and the identification of genetic variants underlying complex traits (GABRIEL *et al.* 2002; KHATKAR *et al.* 2007; VILLA-ANGULO *et al.* 2009; ZHANG *et al.* 2002; ZHAO *et al.* 2003).

1.2. Research Hypothesis and Objectives

This thesis describes the use of the BovineSNP50 BeadChip (50K) as the DNA marker panel to characterize the structure of the bovine genome in terms of LD and haplotype block patterns, and to detect novel QTL regions associated with

milk production and meat quality traits in dairy and beef cattle. This research also aims to identify epistatic QTL contributing genetic variance to economically important traits of dairy cattle. We hypothesize that: (1) different LD and haplotype block structures exist between modern dairy and beef cattle; (2) considerable novel genetic variants exist for milk production traits and meat quantity and quality traits in dairy and beef cattle, respectively, which can be identified by genome-wide association studies with high-density genome-wide SNP assays and advanced statistical methods; and (3) epistasis is an important genetic component of phenotypic variation in quantitative traits in dairy cattle.

One Canadian Holstein population and one commercial hybrid beef population are used in this study. The Holstein population has experienced strong directional selection for milk production and conformation with pure-breeding. The hybrid beef cattle population was derived by mixing three synthetics developed at the Kinsella beef cattle research station, University of Alberta (BERG *et al.* 1990). The major objective of this breeding population has been crossbreeding and moderate selection for performance and productivity under commercial management conditions similar to typical beef operations in Alberta and elsewhere in Canada. The main components of this thesis are:

1. Comparative assessment of LD and haplotype block maps between beef and dairy cattle. The results are presented in Chapter 3 of this thesis.

2. Detection of QTL and candidate genes for milk production traits in Canadian Holstein cattle by GWAS using single marker mixed model regression and Bayesian regression. Chapter 4 presents the findings of this study.

3. Detection of QTL and candidate genes for carcass traits in beef cattle using GWAS approaches similar to those used for the milk production traits. The results are described in Chapter 5.

4. Genome-wide analysis of epistatic QTL for quantitative traits in dairy cattle. Chapter 6 describes this work.

1.3. Literature Cited

ANDERSSON, L., and M. GEORGES, 2004 Domestic-animal genomics: deciphering the genetics of complex traits. Nature Reviews Genetics **5:** 202-212.

ANKRA-BADU, G. A., D. SHRINER, E. LE BIHAN-DUVAL, S. MIGNON-GRASTEAU, F. PITEL *et al.*, 2010 Mapping main, epistatic and sex-specific QTL for body composition in a chicken population divergently selected for low or high growth rate. BMC Genomics **11:** 107.

ARDLIE, K. G., L. KRUGLYAK and M. SEIELSTAD, 2002 Patterns of linkage disequilibrium in the human genome. Nature Reviews Genetics **3:** 299-309.

BERG, R. T., M. MAKARECHIAN and P. F. ARTHUR, 1990 The University of Alberta beef breeding project after 30 years - A review. Annual Feeders' Day Report **69:** 65-69.

BRADFORD, G. E., 1999 Contributions of animal agriculture to meeting global human food demand. Livestock Production Science **59:** 95-112.

BRADLEY, D. G., R. T. LOFTUS, P. CUNNINGHAM and D. E. MACHUGH, 1998 Genetics and domestic cattle origins. Evolutionary Anthropology **6:** 79-86.

BROTHERSTONE, S., and M. GODDARD, 2005 Artificial selection and maintenance of genetic variance in the global dairy cow population. Philosophical Transactions of the Royal Society B-Biological Sciences **360:** 1479-1488.

BRUFORD, M. W., D. G. BRADLEY and G. LUIKART, 2003 DNA markers reveal the complexity of livestock domestication. Nature Reviews Genetics **4:** 900-910.

CARLBORG, O., and C. S. HALEY, 2004 Epistasis: too often neglected in complex trait studies? Nature Reviews Genetics **5:** 618-U614.

CARLBORG, O., S. KERJE, K. SCHUTZ, L. JACOBSSON, P. JENSEN *et al.*, 2003 A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome Research **13:** 413-421.

CASAS, E., J. W. KEELE, S. D. SHACKELFORD, M. KOOHMARAIE, T. S. SONSTEGARD *et al.*, 1998 Association of the muscle hypertrophy locus with carcass traits in beef cattle. Journal of Animal Science **76:** 468-473.

COHEN-ZINDER, M., E. SEROUSSI, D. M. LARKIN, J. J. LOOR, A. EVERTS-VAN DER WIND *et al.*, 2005 Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Research **15:** 936-944.

DE ROOS, A. P., B. J. HAYES, R. J. SPELMAN and M. E. GODDARD, 2008 Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics **179:** 1503-1512.

DEKKERS, J. C., 2004 Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. Journal of Animal Science **82 E-Suppl:** E313-328.

DEKKERS, J. C. M., and F. HOSPITAL, 2002 The use of molecular genetics in the improvement of agricultural populations. Nature Reviews Genetics **3:** 22-32.

DIAMOND, J., 2002 Evolution, consequences and future of plant and animal domestication. Nature **418:** 700-707.

ELSIK, C. G., R. L. TELLAM, K. C. WORLEY, R. A. GIBBS, A. R. R. ABATEPAULO *et al.*, 2009 The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution. Science **324:** 522-528.

GABRIEL, S. B., S. F. SCHAFFNER, H. NGUYEN, J. M. MOORE, J. ROY *et al.*, 2002 The structure of haplotype blocks in the human genome. Science **296:** 2225-2229.

GIBBS, R. A., J. F. TAYLOR, C. P. VAN TASSELL, W. BARENDSE, K. A. EVERSOIE *et al.*, 2009 Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. Science **324:** 528-532.

GODDARD, M. E., and B. J. HAYES, 2009 Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics **10:** 381-391.

GOLDSTEIN, D. B., and M. E. WEALE, 2001 Population genomics: Linkage disequilibrium holds the key. Current Biology **11:** R576-R579.

GRAPES, L., J. C. M. DEKKERS, M. F. ROTHSCHILD and R. L. FERNANDO, 2004 Comparing linkage disequilibrium-based methods for fine mapping quantitative trait loci. Genetics **166:** 1561-1570.

GRISART, B., W. COPPIETERS, F. FARNIR, L. KARIM, C. FORD *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research **12:** 222-231.

GROSSE-BRINKHAUS, C., E. JONAS, H. BUSCHBELL, C. PHATSARA, D. TESFAYE *et al.*, 2010 Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc x Pietrain population. Genetics Selection Evolution **42**.

HAYES, B., and M. GODDARD, 2010 Genome-wide association and genomic selection in animal breeding. Genome **53:** 876-883.

HAYES, B. J., P. J. BOWMAN, A. J. CHAMBERLAIN and M. E. GODDARD, 2009a Invited review: Genomic selection in dairy cattle: progress and challenges (vol 92, pg 433, 2009). Journal of Dairy Science **92:** 1313-1313.

HAYES, B. J., A. J. CHAMBERLAIN, S. MACEACHERN, K. SAVIN, H. MCPARTLAN *et al.*, 2009b A genome map of divergent artificial selection between Bos taurus dairy cattle and Bos taurus beef cattle. Animal Genetics **40:** 176-184.

HAYES, B. J., P. M. VISSCHER, H. C. MCPARTLAN and M. E. GODDARD, 2003 Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Research **13:** 635-643.

KHATKAR, M. S., K. R. ZENGER, M. HOBBS, R. J. HAWKEN, J. A. L. CAVANAGH *et al.*, 2007 A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in Holstein-Friesian cattle. Genetics **176:** 763-772.

MATUKUMALLI, L. K., C. T. LAWLEY, R. D. SCHNABEL, J. F. TAYLOR, M. F. ALLAN *et al.*, 2009 Development and Characterization of a High Density SNP Genotyping Assay for Cattle. PLoS One **4**.

MEUWISSEN, T. H. E., B. J. HAYES and M. E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157:** 1819-1829.

SETOGUCHI, K., M. FURUTA, T. HIRANO, T. NAGAO, T. WATANABE *et al.*, 2009 Cross-breed comparisons identified a critical 591-kb region for bovine carcass weight QTL (CW-2) on chromosome 6 and the Ile-442-Met substitution in NCAPG as a positional candidate. BMC Genetics **10**.

SHAO, H. F., L. C. BURRAGE, D. S. SINASAC, A. E. HILL, S. R. ERNEST *et al.*, 2008 Genetic architecture of complex traits: Large phenotypic effects and pervasive epistasis. Proceedings of the National Academy of Sciences of the United States of America **105:** 19910-19914.

SLATKIN, M., 2008 Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics **9:** 477-485.

UEMOTO, Y., S. SATO, C. OHNISHI, S. TERAI, A. KOMATSUDA *et al.*, 2009 The effects of single and epistatic quantitative trait loci for fatty acid composition in a Meishan x Duroc crossbred population. Journal of Animal Science **87:** 3470-3476.

VILLA-ANGULO, R., L. K. MATUKUMALLI, C. A. GILL, J. CHOI, C. P. VAN TASSELL *et al.*, 2009 High-resolution haplotype block structure in the cattle genome. BMC Genetics **10**.

XU, S., and Z. JIA, 2007 Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics **175:** 1955-1963.

ZHANG, K., P. CALABRESE, M. NORDBORG and F. SUN, 2002 Haplotype block structure and its applications to association studies: power and study designs. American Journal of Human Genetics **71:** 1386-1394.

ZHAO, H., R. PFEIFFER and M. H. GAIL, 2003 Haplotype analysis in population genetics and association studies. Pharmacogenomics **4:** 171-178.

Chapter 2. Literature Review

2.1. Linkage Disequilibrium

Linkage disequilibrium (LD) or more appropriately gametic phase disequilibrium refers to the nonrandom association of alleles at different loci within gametes (LEWONTIN and KOJIMA 1960). LD is a sensitive indicator of the evolutionary and selection history of a population, and is important in mapping QTL regions or genes that are associated with quantitative traits or diseases (GOLDSTEIN and WEALE 2001; SLATKIN 2008).

2.1.1. Measures of LD

Different measurements have been defined for characterizing LD, including *D*, *D* ´and r^2. These measures of LD depend on the quantity *D*, or disequilibrium coefficient, which quantifies disequilibrium as the difference between the observed frequency of a two-locus haplotype and the expected frequency if the alleles are segregating at random (HILL 1981). For a pair of markers *A* and *B* with two alleles at each locus (*A*, *a*) and (*B*, *b*), *D* is calculated as:

$$D_{AB} = p_{AB} - p_A p_B,$$

in which p_{AB} is the observed frequency of gamete *AB*, p_A is the frequency of allele *A*, and p_B is the frequency of allele *B*. If the two loci *A* and *B* are assorted independently, then the expected frequency of haplotype $p_{AB} = p_A p_B$. If

$p_{AB} \neq p_A p_B$, then it indicates that these two loci are non-randomly associated and tend to be segregating jointly. The extent of LD in a population decreases at a rate that depends on the time (t) and recombination fraction (r) between the two loci according to the following formula:

$$D_t = (1-r)^t D_0,$$

where D_0 is the extent of disequilibrium at some starting point and D_t is the extent of disequilibrium t generations later. The D statistic is very dependent on the frequencies of the individual alleles. It is not the best statistic to use when comparing the extent of LD among multiple pairs of loci because the range of possible values of D for each pair is constrained by the allele frequencies (ARDLIE *et al.* 2002).

D', also called Lewontin's D ´, is the absolute value of D standardized by its maximum value, given the allele frequencies at the two loci (LEWONTIN 1964),

$$D'_{AB} = |D| / D_{\max} = \begin{cases} \dfrac{|D_{AB}|}{\min(p_A p_B, p_a p_b)} & D_{AB} < 0 \\ \dfrac{|D_{AB}|}{\min(p_A p_b, p_a p_B)} & D_{AB} > 0 \end{cases}.$$

When $D' = 1$, it indicates that the two markers are in complete LD. When $D' < 1$, it means that the ancestral LD between two loci has been disrupted by historical recombination. However, the relative magnitude of values of $D' < 1$ has no clear interpretation and D' tends to be strongly inflated with small sample sizes or low

allele frequencies (ARDLIE *et al.* 2002; MCRAE *et al.* 2002). It is also difficult to compare the strength of LD between studies using different samples because of the dependence of D' on sample size.

The most frequently used measure of disequilibrium is the square of the correlation coefficient (r^2) between markers (HILL and ROBERTSON 1968),

$$r^2 = \frac{D_{AB}^2}{p_A(1-p_A)p_B(1-p_B)}.$$

r^2 is the correlation of alleles at the two sites, and is formed by dividing D^2 by the product of the four allele frequencies at the two loci. It removes the arbitrary sign of D and is less dependent on the allele frequencies. In addition, r^2 can provide information on the required marker density and sample size for a genome scan of QTL. The r^2 value between a genotyped marker and an unobserved QTL is the proportion of phenotypic variation caused by the alleles at the QTL that can be observed using the marker. Thus r^2 is a key parameter determining the power of LD mapping to detect a QTL. The sample size should be increased by a factor of $1/r^2$ to detect the ungenotyped QTL with the same power as an experiment observing the QTL directly (CARLSON *et al.* 2003; PRITCHARD and PRZEWORSKI 2001).

2.1.2. Mechanisms that Generate and Erode LD

The extent and distribution of LD in a population are affected by many demographic factors including mutation, genetic drift, migration, selection, small

finite population size, and recombination (ARDLIE *et al.* 2002; LANDER and SCHORK 1994). Genetic drift refers to changes in gene and haplotype frequency in a population due to the random sampling of gametes during the production of a finite number of offspring (ARDLIE *et al.* 2002). Genetic drift will often result in the loss of some haplotypes in small populations and alone can create LD (TERWILLIGER *et al.* 1998). Similar increases in LD can be caused by inbreeding, which increases haplotype sharing.

LD can be created through population admixture and migration (gene flow), where the mixed populations have haplotypes that occur in different frequencies. The extent of LD created depends heavily on the time since admixture occurred and the allele frequency differences in the parental populations (GODDARD 1991; GREENWOOD *et al.* 2004).

Selection is an important force to create LD through the genetic hitchhiking effect, by which the frequency of nucleotides in neighboring DNA of a positively selected gene will be rapidly increased. In addition to the selection on favorable variants, background selection, caused by occasional purging of non-deleterious alleles due to spatial proximity to deleterious variants, can also inflate LD. Besides the LD created by selection of physically linked loci, epistatic selection for combinations of alleles at two or more loci far from each other on the same chromosome can result in LD among the syntenic markers at different loci (ARDLIE *et al.* 2002; FARNIR *et al.* 2000). The amount of LD created by selection depends on both the selection intensity and the generation interval in the species,

and is not averaged throughout the genome since the selection is localized around specific genes.

Finite population size is indicated as the most important cause of genome-wide LD in livestock populations since the effective population size for most livestock species is relatively small (ANDERSSON and GEORGES 2004; GODDARD 1991; HAYES *et al.* 2003). In cattle, the effective population size was over 50,000 at 10,000 generations ago and was reduced to a few thousand around 1,000 generations ago (MACEACHERN *et al.* 2009). Recently the effective population size has been decreased to ~100 over the last 50 generations by breed formation and artificial breeding techniques (DE ROOS *et al.* 2008; MACEACHERN *et al.* 2009).

LD is continuously eroded primarily by recombination. The extent of LD is expected to vary in negative relation to the local recombination rate (GREENWOOD *et al.* 2004). Stronger LD is observed across non-recombining regions and weaker LD is observed at localized recombination hot spots (JEFFREYS *et al.* 2001).

Mutation is another factor affecting LD (SUNYAEV *et al.* 2003). No or little LD was observed between SNPs located in the CpG islands with a high mutation rate and markers in close proximity (ARDLIE *et al.* 2002). On the other hand, selection on favorable or deleterious mutations could create LD between the mutation and the neighboring loci. For quantitative traits affected by a large

number of alleles, the amount of LD created by such functional mutations is likely to be small since the individual effect of the mutation is generally small.

Overall, LD can be affected by recurrent factors including gene drift (inbreeding), recurrent migration, selection, and recombination, while it is sporadically affected by punctual factors including mutation, one-time migration or admixture, and population bottleneck or founder effects.

2.1.3. LD Studies in Cattle

LD has been extensively examined in a variety of cattle populations. The first genome-wide LD study in Dutch black-and-white dairy cattle was generated by Farnir *et al.* (2000) using a few hundred microsatellite markers. In this previous study, the syntenic marker pairs disclosed high levels of LD (measured using Lewontin's normalized D') that extended over several tens of centimorgans. Similar results on extensive LD were also observed in several subsequent studies, for example, in US Holstein cattle (VALLEJO *et al.* 2003), United Kingdom Holstein cattle (TENESA *et al.* 2003), Japanese Black and Japanese Brown beef cattle (ODANI *et al.* 2006), and Australian Holstein-Friesian cattle (KHATKAR *et al.* 2006b).

However, studies using different measures of LD (D' or r^2), types of genetic markers (microsatellite markers or SNPs) and marker densities yielded quite different conclusions in terms of the strength of LD in cattle (BOHMANOVA *et al.* 2010; DE ROOS *et al.* 2008; KHATKAR *et al.* 2008; MARQUES *et al.* 2008;

MCKAY *et al.* 2007; SARGOLZAEI *et al.* 2008). McKay *et al.* (2007) constructed whole genome LD maps for eight cattle breeds from the *Bos taurus* and *Bos indicus* subspecies using 2,670 SNPs. The results showed that the extent of LD (measured using r^2) was no more than 0.5 Mb in all eight cattle breeds. In comparison with previous studies in cattle showing that LD extended several tens of centimorgans, this study indicated that LD persisted over much more limited distances and that the extent of LD available for association analyses did not significantly exceed 500 kb. This study suggested that 50,000 SNPs is the minimum requirement for whole genome association studies in cattle based on the extent of LD. Later on, similar results on the extent of genome-wide LD using high density SNPs were obtained in Australian Holstein-Friesian cattle (KHATKAR *et al.* 2008) and North American Holstein cattle (BOHMANOVA *et al.* 2010; KIM and KIRKPATRICK 2009; SARGOLZAEI *et al.* 2008). In addition, De Roos *et al.* (2008) characterized the persistence of LD phase across multiple cattle populations using genome-wide SNPs genotyped in Dutch and Australian Holstein-Friesian bulls, Australian Angus cattle, and New Zealand Friesian and Jersey cows. The correlation of LD between populations for the same marker pairs decreased with increasing marker distance and the extent of divergence between populations. This study on the average r^2 suggested ~50,000 SNPs for genomic selection within breeds and ~300,000 markers for genomic selection across divergent cattle breeds.

2.1.4. Haplotype Blocks

Generally, the term "haplotype block" refers to sizeable genomic regions that show little evidence for historical recombination and low haplotype diversity (GABRIEL *et al.* 2002a). Such blocks have been described in humans (DALY *et al.* 2001; GABRIEL *et al.* 2002a; GABRIEL *et al.* 2002b; PATIL *et al.* 2001; REICH *et al.* 2001) and many other species including cattle (KHATKAR *et al.* 2007), dog (LINDBLAD-TOH *et al.* 2005), pig (AMARAL *et al.* 2008) and rat (GURYEV *et al.* 2006).

2.1.5. Mechanisms for Generating or Maintaining Haplotype Blocks

Previous studies found that haplotype blocks are mainly shaped by recombination, mutation, selection, and population demographic history (GABRIEL *et al.* 2002a; GREENAWALT *et al.* 2006; GURYEV *et al.* 2006; JEFFREYS *et al.* 2001; PHILLIPS *et al.* 2003; WANG *et al.* 2002). These demographic factors are similar to those affecting LD. Recombination rate can be a strong contributor to haplotype block structure based on the evidence of a striking negative correlation between block length and recombination rate (GREENWOOD *et al.* 2004). However, recombination was not considered as an important factor for maintaining the haplotype block structures since recombination evolves rapidly and will not leave its full imprint on haplotype diversity (GURYEV *et al.* 2006; JEFFREYS *et al.* 2005; PHILLIPS *et al.* 2003).

Population history is another factor affecting haplotype block size. Shared genealogical history was found to be the primary determinant of haplotype block patterns in the human genome (REICH *et al.* 2002). In addition, human populations with a relatively heterogeneous founder population had small blocks with a median size of 45 kb, while inbred populations of laboratory mice, which experienced a recent genetic bottleneck during domestication, had large blocks spanning hundreds of kilobases (GURYEV *et al.* 2006).

Natural selection, in the form of selective sweeps or background selection, can also create long-range LD (PHILLIPS *et al.* 2003). The specific haplotypes under selection within a population is mainly driven by the beneficial or deleterious effect of an individual polymorphism in the haplotype block. Selection was found to govern the conservative haplotype blocks by comparing the orthologous rat, mouse, and human haplotype structure of a 5 Mb region from rat chromosome 1, where haplotype block structure was found conserved across mammals, most prominently in genic regions. The results suggested the existence of an evolutionary selection process that drives the conservation of long-range allele combinations (GURYEV *et al.* 2006).

Genetic drift may be an additional mechanism to create haplotype blocks. A previous simulation study found that block patterns were observed in models where recombination crossovers are randomly or uniformly distributed, and those blocks were demonstrated to be generated by genetic drift (ZHANG *et al.* 2003).

Furthermore, mutation was found to jointly dictate haplotype block characteristics with population demographic history and recombination (WANG *et al.* 2002).

2.1.6. Haplotype Block Studies in Cattle

There is increasing interest in understanding patterns of haplotype blocks in the bovine genome. Khatkar *et al.* (2006) identified 40 blocks on chromosome 6 in 433 Australian dairy bulls using 220 SNPs. These blocks accounted 41% of the chromosome 6 and was estimated based on the definition that a block is a set of loci with zero LD units increase over the span of the loci.

Khatkar *et al.* (2007) presented a haplotype block map in 1,000 Holstein-Friesian bulls using 15,036 SNPs with an intermarker spacing of 251.8 kb. Based on the definition of a block as a region in which 95% of the combinations of markers within the region were in high LD (GABRIEL *et al.* 2002a), 727 haplotype blocks covering 2.18% of the autosomal bovine genome were identified. The average block size identified from this study was 69.7 kb and was ~ 5 – 10 times larger than the blocks in human. This study provided very limited block coverage of the bovine genome and suggested ~250,000 SNPs, or ~75,000 – 100,000 tag SNPs for tracking all important haplotype blocks in the bovine genome.

Kim and Kirkpatrick (2009) analyzed the haplotype block structure for 200 North American Holstein cattle using 7,119 SNPs. The results showed that the maximum haplotype block size was over 1 Mb with the mean block size of 26

– 113 kb by various definitions. The results also confirmed that the haplotype block size in Holstein cattle was larger than that observed in human (~10 kb).

Villa-Angulo *et al.* (2009) presented the first high-resolution analysis of haplotype block structure in 501 animals sampled from 19 worldwide taurine and indicine breeds, plus two outgroup species. This study focused on 101 targeted genomic regions on chromosomes 6, 14, and 25, spanning ~7.6 Mb, with an average intermarker spacing of 4 kb. The results indicated an average block size of 10.3 kb (~30 bp – 75 kb) and high similarities in LD and haplotype block structure between cattle and humans on the scale of 1 – 100 kb. This study also observed similar haplotype block structure between beef and dairy breeds, and suggested to use ~30,000 uniformly distributed SNPs for the complete LD map and ~580,000 SNPs for the complete haplotype block structure across the cattle genome.

2.2. QTL Mapping

Most of the economically important traits in livestock are quantitative in nature with their genetic variation determined by a large number of genes plus environmental factors and their interactions. Identification of the genetic variants that underlie these traits is an important and challenging goal in animal genetics studies. QTL refer to chromosomal regions likely to contain genes affecting the genetic variation of quantitative traits (TANKSLEY 1993). Without prior knowledge of the actual genes in the genome, QTL mapping studies can be carried out, which involve identifying associations between specific regions of the

genome and phenotypic traits using molecular markers as anchors. The genome locations and estimated effects of the QTL, particularly those QTL with moderate to large effects, can be used to increase selection accuracy when choosing genetically superior animals (DEKKERS and HOSPITAL 2002).

2.2.1. Statistical Methods of QTL Detection

To date, several statistical methods have been developed for QTL mapping studies. These methods can be generally classified into four categories: (1) regression methods; (2) maximum likelihood methods; (3) mixed model methods based on variance components analysis; and (4) Bayesian methods via Markov chain Monte Carlo (MCMC).

2.2.1.1. Regression Methods

Analysis of variance (ANOVA) using single marker genotypes (a) or multiple marker genotypes (b) can be carried out to detect genetic markers linked to a QTL using the models,

$$y = \mu + MG + e \qquad \text{(a)},$$

$$y = \mu + MG_1 + MG_2 + ... + MG_n + e \qquad \text{(b)},$$

where *MG* represents the marker genotype and is fitted as a fixed effect (SOLLER *et al.* 1976). The multiple marker method does not take into account of the recombination rates between markers, or between QTL and markers.

The probabilistic distribution of a QTL genotype can be inferred from the genotypes of the flanking markers. Lander and Botstein (1989) proposed the first interval mapping method of using two markers only each time to infer the genotype of an internal putative QTL locus. In this method, every possible position within the interval is evaluated. For a given marker haplotype, the probability of inheriting the *Q* or *q* allele for a QTL from the sire can be calculated. The phenotype can then be regressed on the QTL probability conditional on the marker haplotype using the model,

$$y = \mu + \alpha \cdot x + e .$$

Where *y* is the observed phenotype, *x* is the probability of having inherited a paternal *Q* or *q* given the observed marker haplotypes and the marker/QTL positions. The residual sum of squares, also called sum of squared errors of prediction (SSE), is determined for each pseudo QTL position and the true QTL location is where the SSE is minimum.

Haley and Knott (1992) presented a more general model where QTL genotypes are dependent on marker genotypes rather than dealing with the marker haplotypes using model,

$$y = \mu + \alpha \cdot x_1 + \beta \cdot x_2 + e .$$

Where $x_1 = P(QQ \mid M_i) - P(qq \mid M_i)$ and $x_2 = P(Qq \mid M_i)$ are probabilities of QTL genotypes conditional on the flanking marker genotypes, the regression coefficient α represents the difference between additive effects of homozygote

QTL genotypes *QQ* and *qq*, the coefficient β represents the QTL dominance effect. An approximate likelihood ratio (LR) test is proposed for significance testing as,

$$LR = n\ln(\frac{SSE_{reduced}}{SSE_{full}}).$$

The test statistic is the ratio of the residual SSE in a model with the QTL (full model) to the SSE in a model without the QTL (reduced model). This method is also called Haley-Knott regression.

Later, the composite interval mapping (ZENG 1994) or multiple-QTL mapping (JANSEN 1993; JANSEN and STAM 1994) approaches were developed to improve the estimates of a QTL's location and effect. A number of markers outside the tested interval on the same chromosome or on other chromosomes are fitted to the model as cofactors to absorb background noise caused by other QTL or polygenic variation. The improvement in mapping resolution was dramatic by using the composite mapping method in contrast to the standard interval mapping (KUITTINEN *et al.* 1997). In addition, a multipoint implemented interval mapping method was developed in which all markers can be used simultaneously to infer the genotype of any putative QTL locus (JIANG and ZENG 1997). Furthermore, interval mapping was extended to consider multiple QTL by using the multiple interval mapping (MIM) approach, where multiple marker intervals are simultaneously used to fit multiple putative QTL in the model, and the precision and power of QTL mapping was improved (KAO *et al.* 1999).

2.2.1.2. Maximum Likelihood Methods

For a normally distributed phenotypic trait *y* with mean μ and standard deviation σ, the probability density function (PDF) is,

$$f(y_i \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\frac{1}{2}(y-\mu)^2}{\sigma^2}}.$$

The likelihood function describes the probability of parameters given the observed data,

$$L(\mu, \sigma \mid y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\frac{1}{2}(y-\mu)^2}{\sigma^2}}.$$

For a QTL with different genotypes where $Q \sim N(\mu_Q, \sigma)$ and $q \sim N(\mu_q, \sigma)$, the likelihood function given the phenotype is (LYNCH and WALSH 1998),

$$L(\mu_Q, \mu_q, \sigma \mid y) = \prod_{i=1}^{n} \left[P(\mu_Q) \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\frac{1}{2}(y_i-\mu_Q)^2}{\sigma^2}} + P(\mu_q) \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\frac{1}{2}(y_i-\mu_q)^2}{\sigma^2}} \right].$$

Where n is the sample size and y_i is the observation of the quantitative trait for the *i*th individual. For most QTL analysis with unknown QTL genotype, $P(\mu_Q)$ and $P(\mu_q)$ are the probabilities that an individual with an observation belongs to the *Q*–mean or to the *q*–mean, respectively. Finally, the maximum likelihood (ML) estimates of the model parameters are calculated by setting the first

derivative $\ln(L) = 0$. The maximum can be found by using different maximization methods, e.g. expectation-maximization (EM), or Newton-Raphson. A test of significance is to compare the maximum likelihood with the likelihood of a null hypothesis model with the tested parameters omitted.

$$LR = -2\ln\left[\frac{L(\mu,\sigma^2)}{L(\hat{\mu}_Q,\hat{\mu}_q,\hat{\sigma}^2)}\right],$$

where $L(\mu,\sigma^2) = \prod_{i=1}^{n} f(y_i,\mu,\sigma^2)$ assumes no QTL is linked with the marker.

2.2.1.3. Variance Component Methods

Variance component methods have been widely used to detect QTL through obtaining the identity by descent (IBD) coefficients between relatives for the QTL based on marker and pedigree data (ALMASY and BLANGERO 1998; AMOS 1994; GEORGE *et al.* 2000; GRIGNOLA *et al.* 1996; VAN ARENDONK *et al.* 1998). The general genetic model is,

$$y = X\beta + Z_1 u + Z_2 q + e.$$

Where *y* is a vector of *n* phenotypic observations, β is a vector of fixed effects, u is a vector of random polygenetic effects, *q* is a vector of the random QTL effects, *e* is the residual vector, and *X*, Z_1, Z_2 are incidence matrices. The random effects (u,q,e) are assumed to follow $u \sim N(0, A\sigma_u^2)$, $q \sim N(0, G\sigma_q^2)$, $e \sim N(0, R\sigma_e^2)$, where *A* is the additive genetic relationship matrix based on

recorded pedigree, *G* is the genotype relationship matrix based on IBD probabilities at the QTL, and *R* is the residual covariance matrix. Then the total variance given the observed pedigree and marker genotype is equal to,

$$V = Z_1 A Z_1' \sigma_u^2 + Z_2 G Z_2' \sigma_q^2 + R .$$

The likelihood of phenotypic data given the IBD coefficients at each putative QTL position is estimated with ML or restricted maximum likelihood (REML), and the QTL is located at the position that has the highest likelihood value (ALMASY and BLANGERO 1998),

$$\log L(y \mid X\beta, \sigma_q^2, \sigma_u^2, \sigma_e^2) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|V| - \frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta) .$$

$$LR = -2\ln\left[\frac{L(\beta, \sigma_u^2, \sigma_e^2)}{L(\beta, \sigma_q^2, \sigma_u^2, \sigma_e^2)}\right].$$

In the variance-component model, the null hypothesis, that the additive genetic variance due to the QTL equals zero, is tested using the likelihood ratio test by comparing the likelihood of this restricted model with that of a model in which the variance due to the QTL is estimated (SELF and LIANG 1987).

With the increase of genomic marker densities, a variance component method called multi-marker linkage disequilibrium mapping has been extended to fine map QTL using LD generated from closely linked markers (MEUWISSEN and GODDARD 2000; MEUWISSEN and GODDARD 2001). In this method, the IBD probabilities between individuals with fully or partially unknown pedigree are

predicted from marker haplotype similarity instead of from pedigree information. It allows utilizing unknown relationships beyond the recorded pedigree as well as known relationships. The position estimates of QTL given data on a haplotype of markers spanning that position are more accurate than those from a single marker transmission disequilibrium test. The mapped QTL is usually located within a few centimorgans. Later this LD mapping method was extended to combine both linkage and linkage disequilibrium information for robust fine mapping QTL (MEUWISSEN *et al.* 2002). Meuwissen and Goddard (2004) further expanded the method for multi-locus QTL mapping using multi-trait data. The fitting of multiple QTL gives a much sharper indication of the QTL position than the single QTL model and can be used for disclosing multiple QTL. In addition, Han and Xu (2010) developed a multiple variance component model for genome-wide evaluation of QTL using either the ML method or the MCMC implemented Bayesian method. This model estimates multiple QTL variances and positions simultaneously, with the Bayesian method producing the optimal result and the ML method being computationally more efficient.

2.2.1.4. Bayesian Methods

In Bayesian methods, the parameters are treated as variables with their own probability distributions. The conditional distribution of parameters given the data is inferred. If appropriate prior distributions are used, this posterior distribution of parameters contains much more information than the point

estimates from maximum likelihood analysis. Let θ be a $m \times 1$ vector of parameters, the joint posterior distribution of the parameters is,

$$p(\theta \mid y) = \frac{p(y \mid \theta)\, p(\theta)}{p(y)} \propto p(y \mid \theta)\, p(\theta).$$

Where $p(\theta)$ is the prior distribution of parameters, and $p(y \mid \theta)$ is the likelihood function of the parameters. The marginal posterior distribution of the *k*th parameter θ_k is,

$$p(\theta_k \mid y) = \int \ldots \int p(\theta_k, \theta_{-k} \mid y) d\theta_{-k}.$$

Where θ_{-k} is a vector containing all parameters except the *k*th parameter.

The MCMC algorithm can be used to calculate the marginal posterior distribution for each parameter. It is a sampling based algorithm for repeatedly sampling a parameter from its fully conditional posterior distribution. Let $p(y \mid \theta_k, \theta_{-k})$ represent the likelihood of the parameters and $p(\theta_k, \theta_{-k})$ be the prior density of the parameters, the joint posterior distribution of the parameters is,

$$p(\theta_k, \theta_{-k} \mid y) \propto p(y \mid \theta_k, \theta_{-k})\, p(\theta_k, \theta_{-k}).$$

At the *t*th iteration, the fully conditional posterior distribution of θ_k is,

$$p(\theta_k \mid \theta_{-k}^{(t)}, y) = p(y \mid \theta_k, \theta_{-k}^{(t)})\, p(\theta_k, \theta_{-k}^{(t)}).$$

Based on if θ_k can be directly sampled from $p(\theta_k \mid \theta_{-k}^{(t)}, y)$ or if $p(\theta_k \mid \theta_{-k}^{(t)}, y)$ has an explicit form of distribution, the MCMC algorithms are classified into different methods, one called the Gibbs sampler and the other one called the Metropolis-Hastings algorithm. For the Gibbs sampler, the fully conditional posterior distribution has a simple form and θ_k can be directly sampled after the burn-in period and removal of the autocorrelation of samples (CASELLA and GEORGE 1992; GEMAN and GEMAN 1984). The Metropolis-Hastings algorithm is used when the distribution of $p(\theta_k \mid \theta_{-k}^{(t)}, y)$ and the way of drawing samples from it are unknown. It is an accept-reject algorithm for drawing samples of random variables. A variable is first drawn from a proposed distribution $q(\theta_k)$ which is similar to $p(\theta_k \mid \theta_{-k}^{(t)}, y)$ and then the sample is accepted or rejected based on an acceptance probability,

$$\alpha = \max\left[1, \frac{p(\theta_k^* \mid \theta_{-k}^{(t)}, y)}{p(\theta_k^{(t)} \mid \theta_{-k}^{(t)}, y)} \frac{q(\theta_k^{(t)})}{q(\theta_k^*)}\right].$$

Where $\theta_k^{(t)}$ is the old value of parameter θ_k and θ_k^* is the new value drawn from the proposed distribution $q(\theta_k)$. If θ_k^* is accepted, then $\theta_k^{(t+1)} = \theta_k^*$ or $\theta_k^{(t+1)} = \theta_k^{(t)}$ (HASTINGS 1970; METROPOLIS *et al.* 1953). Bayesian analysis provides the posterior distribution as the Bayesian estimate of a parameter. The most frequently used representatives of Bayesian estimate are the posterior sample size, posterior mean, posterior mode, posterior standard deviation, equal tail interval and highest posterior density interval (HU and XU 2009).

A Bayesian QTL mapping method was firstly introduced by Satagopan *et al.* (1996) using Bayes factors. Bayesian QTL mapping has also been used in outbred livestock populations with a granddaughter design, where single or bi-QTL models including both major gene and polygenic effects were fitted using a Gibbs sampler (UIMARI and HOESCHELE 1997; UIMARI *et al.* 1996). However, effects of major QTL on other chromosomes were not considered in these studies. Later, all chromosomes were simultaneously considered by Stephens (1998) using a Metropolis-Hastings scheme. The Bayesian method now works as a state-to-art method for dealing with multiple QTL and model selection. The two major Bayesian methods used for multiple QTL mapping are: (1) the reversible jump MCMC method (SILLANPAA and ARJAS 1998; YI and XU 2002a), in which the number of QTL is treated as a random variable and is estimated in the QTL analysis (YI *et al.* 2003), and (2) the Bayesian shrinkage method (WANG *et al.* 2005; XU 2003), which fits all markers into a single model and assigns a normal prior distribution to each marker coefficient, and assigns a scaled inverse chi-square distribution to the specific prior variance of each marker coefficient.

2.2.2. Linkage Mapping and Association Mapping

LD between a marker and a QTL is required for the QTL to be detected in QTL mapping studies. QTL mapping methods can be classified into two main approaches depending on the source of LD or recombination events used in the analysis: linkage mapping in families and population-based association mapping. Linkage mapping, also called family mapping, is a method for localizing genomic

regions that contain a gene related to the phenotype of interest based on the co-segregation of genetic markers and phenotypes in families with known relatedness over several generations (MYLES *et al.* 2009). Linkage mapping has been very successful in finding genes for rare, Mendelian, monogenic diseases, while it can only find loci that have the strongest influence for complex traits (SMITH and O'BRIEN 2005). Linkage mapping can only exploit the limited number of recombination events that occurred during the establishment of the mapping population, thus QTL are generally localized to large chromosomal regions (tens of centimorgans) unless the pedigree is very long or the family size is large (HERNANDEZ-SANCHEZ *et al.* 2009). These large intervals make it difficult to identify the causative variations underlying the QTL. In addition, QTL mapping results are not consistent across mapping populations since different QTL segregate in different populations with phenotypic diversity (HOLLAND 2007).

In contrast to linkage mapping using well designed populations, association mapping (also known as linkage disequilibrium mapping) uses a random sample of a natural population to perform QTL mapping. It searches for genotype-phenotype correlations across families and measures preferential segregation of a particular allele with a phenotype to assess the contribution of genetic variants to phenotypes. The major advantage of association mapping is that it can exploit all recombination events that have occurred in the evolutionary history of the population, thus providing much higher mapping resolution. In addition, this method can disclose more QTL without limitation to only the QTL segregating in the designed families with linkage mapping (MYLES *et al.* 2009).

This rapid and cost effective method works especially well for outbred populations in cattle without the laborious, lengthy and expensive process of constructing the mapping families. Furthermore, a direct analysis of extant populations is desirable for applying the results directly to the general population, as it can avoid the problem of different alleles segregating in different populations. In association mapping, many more markers are required to capture the short population-wise historical LD between markers and QTL, and population demography, e.g. population structure, can cause false positive results (MARCHINI *et al.* 2004).

In the past few years, joint linkage-association mapping, also called combined linkage and linkage-disequilibrium mapping, has been developed for the fine-scale mapping of genes affecting complex traits (BLOTT *et al.* 2003; MEUWISSEN *et al.* 2002; WU *et al.* 2002; WU and ZENG 2001). This method is optimal because it combines the power of association mapping and the robustness of linkage mapping (HERNANDEZ-SANCHEZ *et al.* 2009). Furthermore, false positives may be reduced since signals have to conform to both mapping methods' assumptions (MEUWISSEN *et al.* 2002). In cattle, this method had been used to refine previously reported QTL locations. For example, a QTL for twinning rate in dairy cattle was fine mapped to a region < 1 cM on chromosome 5 (MEUWISSEN *et al.* 2002) and a QTL affecting female fertility on BTA3 was refined to a set of narrow peaks (DRUET *et al.* 2008). This combined method was also applied to refine the position of a previously identified QTL for milk production traits on chromosome 6 to a 7.5 cM interval (OLSEN *et al.* 2004) and

from 7.5 cM to a 420 kb region (OLSEN *et al.* 2005) in Norwegian dairy cattle. In addition, some instances of important causative mutations were disclosed in the process of fine mapping QTL using this method. For example, both linkage and LD information were mined to improve the mapping resolution of a QTL on BTA20 with a major effect on milk yield and composition, and finally a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor (*GHR*) was identified (BLOTT *et al.* 2003). In another example, LD information was added to refine a QTL on BTA14 reported from linkage analysis to a 3 cM marker interval (FARNIR *et al.* 2002) and a non-conservative *K232A* substitution in the *DGAT1* gene with a major effect on milk fat content and other milk characteristics was identified (GRISART *et al.* 2002).

2.2.3. Reported QTL for Milk Production Traits in Dairy Cattle

To date, many studies involving the mapping of QTL affecting milk production traits have been carried out in dairy cattle and a total of 1,485 QTL are recorded in the CattleQTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/BT/summary) for milk traits, with most of them reported for milk yield and milk compositions. These reported QTL are from studies using different dairy cattle populations, marker panels and mapping methods. Some of these studies are described here. Georges *et al.* (1995) screened QTL controlling milk production in 1,518 elite Holstein cattle using 159 microsatellite markers covering ~1,645 cM of the bovine genome. Large QTL underlying the genetic variation of milk production were detected on chromosomes 1, 6, 9, 10 and 20 by using a

maximum likelihood multi-locus linkage analysis. Zhang *et al.* (1998) mapped QTL affecting milk production in a 1,794 Holstein population with granddaughter design using 206 markers flanking 2,497 cM on 29 BTA and two statistical methods: least squares and variance components methods. The average bootstrap confidence interval for the experiment-wise significant QTL was 48 cM and some chromosomes harbored QTL in coupling phase affecting several traits which is consist with the observed genetic correlations among traits. Heyen *et al.* (1999) conducted a genome scan for QTL influencing milk production in 1,068 North American Holstein-Friesian bulls using 174 markers covering 2,551 cM (85%) of the bovine genome with the ANOVA method. The results identified genome-wide significant marker effects on 11 chromosomes and a large effect QTL for fat percentage on BTA14 was also confirmed in another Israeli Holstein dairy herd. Riquet *et al.* (1999) fine mapped a previously identified QTL on BTA14 to a ~5 cM chromosome segment using seven selected heterozygous sires, where common haplotype associated with increased fat percentage was identified. Meuwissen and Goddard (2004) refined a previously identified QTL on BTA14 caused by segregation of *DGAT1* to a region of 0.04 cM using a multi-locus QTL mapping method, which combined linkage and LD information and used multi-trait data. The results showed that a much sharper QTL position was obtained by fitting multiple QTL in contrast to a single QTL model using multi-trait data; however, no indications for a second QTL affecting dairy traits were found on BTA14.

2.2.4. Reported Candidate Genes for Milk Production Traits in Dairy Cattle

Based on the reported QTL for milk production traits in conjunction with functional analysis and comparative mapping, a small number of functional candidate genes together with causative mutations have been identified, including a missense mutation in *DGAT1* (GRISART *et al.* 2002; WINTER *et al.* 2002) and a missense mutation in *ABCG2* (COHEN-ZINDER *et al.* 2005; OLSEN *et al.* 2007). Using the candidate gene approach, several other candidate genes with putative effects on milk production have been also identified in dairy cattle based on the functions of genes in fat, glucose and energy pathways, such as *OPN* (or *SPP1*) (KHATIB *et al.* 2007), *PPARGC1A* (KHATIB *et al.* 2007; WEIKARD *et al.* 2005), *CYP11B1* (KAUPE et al. 2007), *FASN* (MORRIS *et al.* 2007; ROY *et al.* 2006), *OLR1* (KHATIB *et al.* 2006), *SCD* (MELE *et al.* 2007; MOIOLI *et al.* 2007), *PRL* (DYBUS *et al.* 2005), *PRLR* and *GHR* (TURNER *et al.* 2010).

In the candidate gene approach, a gene is assumed to be involved in the physiology of the trait based on its known functions, often reported in other species. Sequence variations within or near the gene can be tested for association with variation in the phenotypic trait. Some important causative mutations have been successfully discovered using this approach, mainly for monogenic traits (ANDERSSON and GEORGES 2004). However, complex traits are affected by multiple genes with variable sizes of effects and the selection of candidate genes for sequencing and association analysis will miss genes that have not been regarded as an obvious candidate for a particular trait. For this reason, genome-

wide association studies (GWAS) are now often used to identify genomic regions of interest, and then candidates within the regions are identified for further study based on their known functions.

2.2.5. Reported QTL for Carcass Traits in Beef Cattle

QTL mapping studies for the identification of genetic markers and genes related to beef carcass traits have been carried out in several beef cattle populations, and a total of 1091 QTL have been recorded for meat traits in the CattleQTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/BT/summary), with most of them reported for carcass characters and meat quality traits. A few examples using different breeds of beef cattle are discussed here. Casas *et al.* (2003) mapped QTL for carcass composition in a half-sib family (n = 547) from a Brahman (*Bos indicus*) × Hereford (*Bos taurus*) sire. Mizoshita *et al.* (2004) detected QTL related to carcass traits in a half-sib family (n = 348) of purebred Japanese Black cattle using 342 microsatellite markers spanning 2,664 cM of 29 bovine autosomes. Takasuga *et al.* (2007) conducted a multiple QTL analysis for carcass traits in purebred Japanese Black cattle using 15 paternal half-sib families comprising 7,860 animals. McClure *et al.* (2010) reported a few hundred QTL for carcass traits using 390 microsatellites markers in two large commercial Angus populations. Two complementary approaches were used in this study: linkage analysis using the half-sib least squares and multipoint QTL interval analysis using Bayesian MCMC analyses of the entire pedigree. In addition, many other QTL mapping studies have been conducted for carcass traits (ABE *et al.* 2008;

CASAS *et al.* 1998; CASAS *et al.* 2004; CASAS *et al.* 2000; CASAS *et al.* 2001; GUTIERREZ-GIL *et al.* 2009; KIM *et al.* 2003; MACNEIL and GROSZ 2002; MOORE *et al.* 2003; NALAILA *et al.* 2011; STONE *et al.* 1999).

2.2.6. Reported Candidate Genes for Carcass Traits in Beef Cattle

A few candidate genes and polymorphisms within QTL regions have been reported for beef carcass traits. For example, the *NCAPG* gene is located in a narrow QTL region of 591 kb on BTA6 for beef carcass weight and an Ile-442-Met substitution in *NCAPG* was found to be significantly associated with carcass weight, longissimus muscle area and subcutaneous fat thickness in Japanese Black and Japanese Brown cattle (SETOGUCHI *et al.* 2009). *MSTN* (also called *GDF8*) within a QTL close to the centromere on BTA2 is thought to account for most of the variation of muscle mass and fat deposition in double-muscled cattle populations (CASAS *et al.* 1998). Variants in the *CRH* gene, located at a fat-related QTL region on BTA14 which has been reported by many studies, were found to be significantly associated with marbling and subcutaneous fat depth in a Wagyu × Limousin F2 population (WIBOWO *et al.* 2007).

In addition, many genes related to beef carcass traits have been identified using the candidate gene approach. For example, significant associations of SNPs in the gene *SCD1* were reported to be associated with fat deposition and composition in a Wagyu × Limousin population (JIANG *et al.* 2008). *IGF1* was found to be associated with fat deposition and carcass merit traits (ISLAM *et al.* 2009), *FABP4* with carcass weight and marbling (LEE *et al.* 2010a), *LEP* with

ultrasound ribeye area, carcass yield, backfat thickness and sensory traits (CORVA *et al.* 2009; GILL *et al.* 2009; NKRUMAH *et al.* 2005; SCHENKEL *et al.* 2005; WU *et al.* 2005), *GH* with beef carcass composition (TAYLOR *et al.* 1998; THOMAS *et al.* 2007; WU *et al.* 2005), *ADFP* with marbling score (CHEONG *et al.* 2009), *TG* with marbling score and subcutaneous fat thickness (DE *et al.* 2004; WOOD *et al.* 2006; WU *et al.* 2005), *POMC* with hot carcass weight (BUCHANAN *et al.* 2005), *CAPN1* and *CAST* with beef marbling and tenderness (CASAS *et al.* 2006; GILL *et al.* 2009; PAGE *et al.* 2002), *DGAT1* for sirloin weight, fat depth surrounding the sirloin, subcutaneous fat thickness and intramuscular fat deposition (GILL *et al.* 2009; THALLER *et al.* 2003; WU *et al.* 2005), *FABP3* with subcutaneous fat thickness (WU *et al.* 2005), and *GHR* with marbling score and odor (GILL *et al.* 2009; HALE *et al.* 2000).

2.2.7. Genome-wide Association Studies

With the completion of reference genome sequence assemblies, the discovery of millions of SNPs, and the development of cost-effective and high-throughput genotyping technologies, genome-wide association studies (GWAS) are increasingly being adopted in cattle as an approach for identifying trait-associated genetic variations (GODDARD and HAYES 2009). In GWAS, a dense set of SNPs across the genome is genotyped and used in the detection of statistical associations between a trait of interest and any of the markers. This method assumes that the functional alleles will likely be in LD with at least one of the genotyped markers.

2.2.7.1. Bayesian Methods Used in GWAS

Most commonly, the data for a GWAS are analyzed one SNP at a time using a simple linear model that includes the effect of a SNP, systematic environmental fixed effects, and the polygenic effect of the individual animal due to all the other genes affecting the trait (GODDARD and HAYES 2009). Recently, Bayesian methods have been developed to predict genomic breeding values (GEBVs) using high-density SNPs covering the whole genome. These methods allow simultaneous estimation of dense marker effects from a small number of phenotypic records and remove the limitation of insufficient of degrees of freedom in the least square methods, and thus can also be applied for GWAS studies aimed at identifying associated genetic variations.

Meuwissen *et al.* (2001) presented two Bayesian methods, termed BayesA and BayesB, with realistic assumptions on the variable variance explained by each locus from a prior distribution. A simulation study found that both methods resulted in much higher accuracy for prediction of marker effects in comparison with best linear unbiased prediction (BLUP), which assumes each marker has equal variance.

BayesA assumes that there are many markers of small effect and few of large effect on a phenotypic trait. Normal priors are first assumed for each marker effect. The variances of marker effects have a scaled inverted chi-square distribution $\sigma^2_{\alpha_j} \sim \chi^{-2}(v,s)$, with known degrees of freedom *v* and a scale parameter *s* derived from the assumed known additive genetic variance. The

posterior distribution of $\sigma^2_{\alpha_j}$ can be obtained from the combined prior distribution of $\sigma^2_{\alpha_j}$ and phenotypic data. Marker effects and variances can then be estimated through Gibbs sampling. An advantage of using an inverted chi-square distribution as a prior for the variances is that with normally distributed data, the posterior distribution of $\sigma^2_{\alpha_j}$ is also an inverted chi-squared.

In BayesB, the prior specification of a marker effect is zero with fixed probability π, and normally distributed with a locus specific variance with probability (1 – π). In comparison with BayesA, the prior distribution of variances of marker effects can be expressed as:

$$\begin{aligned} \sigma^2_{\alpha_j} &= 0 \qquad\quad (p = \pi) \\ \sigma^2_{\alpha_j} &\sim \chi^{-2}(v, s) \quad (p = 1 - \pi) \end{aligned}.$$

The assumption on variances of marker effects is consistent with many SNPs having a zero effect and a few SNPs having non-zero effects.

Other Bayesian methods have also been developed for genomic selection purposes, with different assumptions on the variance components of each marker (hyper-parameters of the normal priors for the regression coefficients) (FERNANDO and GARRICK 2009). The BayesC method is similar to BayesB except that a single common variance is used for all markers with non-zero effects, whereas locus-specific variance components are used in BayesB (KIZILKAYA *et al.* 2010). BayesC is more tolerant to prior genetic variance than BayesB and has

been applied in whole genome association studies of economically important traits in the pig (FAN *et al.* 2011; ONTERU *et al.* 2011).

Two additional Bayesian methods, BayesCπ and BayesDπ, have recently been developed, where the hyper-parameter π that a SNP has zero effect is set to be unknown and is estimated from the data (HABIER *et al.* 2011). The modification of unknown parameter π in both BayesCπ and BayesDπ is in order to overcome the drawbacks that arbitrary π values used in BayesA ($\pi = 0$) and BayesB ($\pi > 0$) and to remove the effects of arbitrary π on the shrinkage of SNP effects. The unknown π estimated from BayesCπ and BayesDπ will reflect the number of QTL for a trait of interest.

There are similarities and differences between these Bayes methods. BayesC and BayesCπ use the similar scale parameter *s* as in BayesA and BayesB, while unknown scale parameter *s* with Gamma prior is used in BayesDπ. A specific variance for each marker is used in BayesA, BayesB and BayesDπ, while BayesC and BayesCπ use a common variance for all SNP effects. The modification on common variance used in BayesC and BayesCπ is trying to solve the drawback in BayesA and BayesB that the full conditional posterior distribution of marker variance is dominated by the prior instead of the data (GIANOLA *et al.* 2009; HABIER *et al.* 2011).

Studies using simulated and real data found that a similar accuracy was obtained for BayesCπ, BayesDπ, BayesA and BayesB. BayesCπ has been suggested for routine applications since estimates of π in BayesCπ, in contrast to

BayesDπ, are sensitive to the number of simulated QTL and training data size, and can provide genetic architecture information for the trait (HABIER *et al.* 2011). Furthermore, Sun *et al.* (2011) conducted a simulation study using both BayesCπ and BayesB methods with windows of 10 consecutive SNPs. Results were found to be similar between BayesCπ and BayesB when π was set to 0.99; however, BayesCπ performed better than BayesB overall since the performance of BayesB dropped when π was decreased.

Overall, Bayesian methods can incorporate prior knowledge on the genetic control of complex traits to coerce negligible QTL effects towards zero. These methods are useful for GWAS QTL mapping by providing better inferences for real QTL based on the joint posterior distribution, which takes full account of all unknown parameters (HOESCHELE *et al.* 1997; ZOU and ZENG 2008).

2.2.7.2. GWAS for Milk Production Traits and Beef Carcass Traits in Cattle

In dairy cattle, whole genome association studies had been carried out for mapping QTL or genomic regions explaining variation in milk production. Kolbehdari *et al.* (2009) conducted a whole genome scan to identify QTL affecting milk production traits for 462 Canadian Holstein bulls using single marker LD regression and 1,536 SNP markers located in or nearby known genes. Genome-wise and chromosome-wise significant SNPs were identified and compared with the previously reported QTL regions in other dairy cattle populations. Several of the significantly associated SNP were located in genes known to encode components of the fat and protein metabolism pathways.

Jiang *et al.* (2010) carried out a GWAS for milk production traits in Chinese Holstein cattle using EBVs of 2,093 daughters from 14 paternal half-sib families and 54K SNPs. Associations were identified by using both a paternal transmission disequilibrium test approach and a mixed model regression analysis and in total 105 SNPs were detected to be significantly associated at the genome-wise level with one or multiple milk production traits.

Mai *et al.* (2010) mapped QTL for milk production traits in 1,039 Danish Jersey bulls by a GWAS using mixed model regression. Association tests for 33,090 SNPs resulted in 98 detected combinations of QTL and traits on 27 chromosomes.

Pryce *et al.* (2010) looked for associations for milk production traits using either single SNPs or haplotypes of SNP alleles (39,048 SNPs were genotyped) by first testing in a discovery Holstein population and then validating in both Holstein and Jersey cattle. The QTL intervals were narrowed down by the across-breed validation strategy. In comparison with single SNPs, the precision of QTL mapping increased with the haplotype length and the number of haplotypes discovered.

Bolormaa *et al.* (2010) carried out a multiple-trait GWAS for dairy traits using a principal component analysis and a series of bivariate analyses. Associations of 39,048 SNPs were tested in a discovery population with 767 Holstein bulls and validated using 386 Holstein bulls and 317 Jersey sires. The results from multiple-trait GWAS showed as good or better statistical power for

detecting associations than single trait GWAS. Multiple-trait GWAS reported additional associations without an increase in the false discovery rate; however, it did not increase the precision for the mapped QTL.

Cole *et al.* (2011) carried out a GWAS using predicted transmitting ability of milk production traits in contemporary U.S. Holstein cows and identified a number of chromosomal regions and candidate genes, such as *GNAS* for milk, fat and protein yields, *DGAT1-NIBP* for fat percentage, *FKBP2* for protein yield and percentage, and *MGMT* and *PDGFRA* for protein percentage.

Bouwman *et al.* (2011) conducted an association study using 50,000 SNPs for milk fat composition in Dutch dairy cattle using a two-step single SNP association analysis. The results identified a total of 54 regions on 29 chromosomes that were significantly associated with one or more fatty acids with many of them located on BTA14, 19 and 26. This study also disclosed several genes within the regions that either have functional evidence linking them to fat synthesis or are within previously identified QTL for fat yield or content, such as *ABCG2* and *PPARGC1A* on BTA6, *ACSS2* on BTA13, *DGAT1* on BTA14, *ACLY*, *SREBF1*, *STAT5A*, *GH*, and *FASN* on BTA19, *SCD1* on BTA26, and *AGPAT6* on BTA27.

Strucken *et al.* (2011) investigated the time-dependent genetic effects over different lactations for milk production traits with a GWAS using 44,962 SNPs in 152 divergent German Holstein-Friesian cows. The results showed that the variance explained by a particular locus changes from lactation to lactation, since

all significant effects were specific for a single lactation. The results also confirmed that most production traits vary in the degree of persistency after the peak of lactation as a result of genetic influence.

In beef cattle, GWAS have been carried out for detecting genetic variations associated with beef carcass traits. Lee *et al.* (2010b) detected associations for carcass quality traits using 32,756 SNPs in 289 Korean Hanwoo cattle. In total, 108 significant SNPs were identified using a simple linear regression model and a best set of 44 SNPs was selected from stepwise regression procedures. Kim *et al.* (2011) conducted a GWAS for carcass traits using 39,129 SNPs in 311 Korean beef cattle and disclosed a total of five SNPs, on BTA3, 6, 11, 13 and 16, that showed association with meat quantity or quality traits. Bolormaa *et al.* (2011) carried out a GWAS for meat and carcass traits in 940 Australian taurine and indicine cattle using 53,798 SNPs. The results were validated in another 1,338 animals genotyped for 335 SNPs and finally 27 significantly associated chromosomal regions were confirmed in both data sets.

2.2.8. Epistatic QTL Mapping

2.2.8.1 Methods for Epistatic QTL Mapping

The genetic variation of a quantitative trait is often controlled by the segregation of multiple interacting loci (CARLBORG and HALEY 2004; MOORE 2005) and epistasis is important in complex traits in animals, plants and human, as indicated by results from an increasingly number of recent studies (DONG *et al.*

2003; FERNANDEZ *et al.* 2000; FIJNEMAN *et al.* 1996; HOLLAND *et al.* 1997; LARK *et al.* 1995; NAGASE *et al.* 2001; SUGIYAMA *et al.* 2001; YU *et al.* 1997). Compared to the additive and dominance genetic effects, epistatic effects are hard to estimate and the discovery of epistatically interacting QTL is mainly hampered by the over-parameterized epistatic genetic models. Models that are able to decouple genetic interaction effects from the total genetic variance will be required for different genetic backgrounds (LE ROUZIC *et al.* 2008).

Early statistical models for single QTL mapping cannot be directly extended for epistatic QTL analysis. To date, several methods for simultaneously searching for multiple-QTL with epistasis in two or more dimensions have been developed. Epistatic QTL can be detected by first conducting a search for single main effects and then checking the interactions among those loci; however, this method will not work in the situation where no main effect is displayed (CULVERHOUSE *et al.* 2002). There are two-locus QTL detection models which fit two QTL at a time and their interactions together with the main effects (HALEY and KNOTT 1992; WANG *et al.* 1999). A stepwise model selection approach is also applied in mapping epistasis in backcross designs (KAO *et al.* 1999; ZENG *et al.* 1999). Carlborg *et al.* (2000) developed a genetic algorithm for simultaneous mapping of multiple interacting QTL to improve the computational efficiency in contrast to the previous step-by-step search (CARLBORG *et al.* 2000). Jannink and Jansen (2001) presented a one-dimensional search for epistatic QTL that involves detecting loci with high interaction between QTL and genetic background using a maximum likelihood method. This approach required large populations derived

from multiple related inbred-line crosses and gave epistatic QTL not only in pairwise but also higher-order interactions. This method was extended by Boer *et al.* (2002) to a penalized likelihood method for mapping epistatic QTL with one dimensional genome searches for high interactions between QTL and the genetic background. Overall, the above discussed methods use a variable selection technique to exclude those epistatic interactions with negligible effects. However, these methods may run a high risk of missing some important interaction effects by not fully exploring the large parameter space of models. Regarding this issue, Zhang and Xu (2005) developed a penalized maximum likelihood method for simultaneously estimating epistatic effects of QTL, which can accommodate a number of effects 15 times larger than the sample size. In this method, spurious QTL effects are shrunk towards zero, while QTL with large effects are estimated with virtually no shrinkage. Simulation studies showed comparable results between this method and the Bayesian shrinkage analysis, but with a much faster computational speed.

Bayesian methods have also been applied to mapping epistatic QTL. Sen and Churchill (2001) presented a Bayesian model selection method for QTL analysis which can accommodate multiple interacting QTL. This method asks for a pre-specified number of QTL which is actually unknown. Yi and Xu (2002b) developed a Bayesian method to map multiple QTL with epistatic effects for an unknown number of QTL, thus the dimension of the model becomes variable. By including all possible pairwise QTL interactions, this method results in a large number of parameters and the number of QTL is estimated by using the reversible

jump MCMC algorithm (SILLANPAA and ARJAS 1998). Yi *et al.* (2003) extended the Bayesian model by adding a variable selection procedure in which latent binary variables are used to indicate which main and epistatic effects of putative QTL are included in or excluded from the model. Multiple QTL with epistasis were identified where the reversible jump MCMC algorithm is used to determine both the number of QTL and to select main and epistatic effects. This method has been used to model epistasis in mice for obesity traits (YI *et al.* 2004). In order to greatly reduce the calculations in epistatic QTL mapping, Yi *et al.* (2005) improved the Bayesian model selection method by placing an upper bound on the number of QTL from prior knowledge, and by developing more efficient MCMC algorithms using the Gibbs sampler and Metropolis-Hastings algorithms. This improved method has detected novel epistatic QTL for obesity in mice. Yi *et al.* (2007) extended the Bayesian model selection framework for mapping epistatic QTL to include environmental effects and gene-environment interactions. This study also explored using the MCMC algorithm with new and fast sampling schemes and proper prior distributions incorporating prior knowledge about the genetic architecture of the complex traits. The updated method was applied in detecting new epistatic and gene-sex interactions for obesity-related traits in two populations of mice.

Xu (2007) proposed an empirical Bayes method (E-BAYES) under the mixed model framework that allows simultaneous estimation of main effects of all individual markers and epistatic effects of all pairs of markers in a single model. This method estimates prior variance components using marginal maximum-

likelihood and then estimates QTL effects using the Bayesian shrinkage method given the estimated prior variance components as if they were the true prior variances. This method without the MCMC samplings for inference of the parameter distribution is efficient in terms of computation ability. In comparison with other methods, such as the variable selection via stepwise regression or stochastic search variable selection (SSVS), a penalized likelihood (PENAL) method and the least absolute shrinkage and selection operator (LASSO) method, E-BAYES appeared to perform better in terms of minimizing the mean-squared error (MSE) and relatively short computing time for both simulation and real data. The method was used to map genome-wide interacting QTL for quantitative traits in barley where negligible epistasis was identified in contrast to the main effects, and the results showed that the appearance of epistasis between two loci did not depend on whether or not the loci had a significant main effects (XU and JIA 2007). The epistatic model using E-BAYES was also applied in the prediction of genomic values for quantitative traits in soybean where the squared correlation coefficient between the observed and predicted phenotype was 0.78 in comparison with 0.33 in the model including only the additive effects for prediction (HU *et al.* 2011).

2.2.8.2 Epistatic QTL Mapping in Animals

To date, several studies on mapping epistasis for complex traits in animals have been carried out using a variety of methods. Carlborg *et al.* (2003) conducted an epistatic QTL analysis for growth traits in an F2 cross chicken population

using 105 evenly distributed genetic markers. QTL were mapped using forward selection for loci with significant marginal effects and with a simultaneous search for epistatic QTL pairs. Results indicated that epistasis was mainly relevant for early growth, whereas additive genetic effects explained the major portion of the genetic variance later in life.

Carlborg *et al.* (2005) mapped epistatic QTL in an F2 cross mouse population using 93 microsatellite markers covering all chromosomes with an average marker spacing of 14.1 cM. This study was conducted by a simultaneous search for epistatic QTL pairs without assuming that the QTL had any effect individually. The results showed that the genetic model with epistatic QTL was able to increase by 8.8% – 128.3% the total explained genetic variance for several growth and body composition traits.

Yi *et al.* (2006) performed analyses of multiple epistatic QTL for body weight and body composition in mice using Bayesian model selection. This study revealed both strong main effects and epistatic interactions and an interacting network of multiple QTL for growth and body composition traits. However, both of the most main and epistatic effects had an opposite effect on early and late growth and the contribution of epistasis was more pronounced for body weights at older age.

Barendse *et al.* (2007) analyzed the epistasis between two candidate genes for beef tenderness in cattle: *CAPN1* on BTA29 and its inhibitor *CAST* on BTA7. Causative mutations within these genes were first identified and genotyped in

over 1500 animals of seven breeds. This study identified significant epistasis between SNPs at *CAPN1* and *CAST* in both taurine and zebu derived breeds and a larger additive × dominance component of epistasis than additive × additive and dominance × dominance components were observed.

Uemoto *et al.* (2009) mapped epistatic QTL for meat fatty acid composition in a Meishan × Duroc crossbred population using 180 microsatellite markers by fitting two loci and their interactions each time in the model. The analysis identified a total of 5 epistatic pairs located on chromosomes 4, 5, 9, and 16.

Significant epistatic QTL associated with meat quality and carcass composition traits were identified in a porcine Duroc × Pietrain population using 131 genetic markers spanning 18 autosomes and a two-step procedure implemented with a maximum likelihood method. This study distinguished 17 epistatic QTL pairs for carcass composition and 39 for meat quality traits, which explained up to 8% of the phenotypic variance. This study also revealed evidence for epistatic relationships between different chromosomal regions (GROSSE-BRINKHAUS *et al.* 2010).

Ankra-Badu *et al.* (2010) mapped epistatic QTL for body composition in a F2 reciprocal intercross between two chicken lines divergently selected for low or high growth rate using the Bayesian model selection method and 109 informative markers on 20 autosomes. The results discovered several QTL on different

chromosomes that interact with each other to affect body composition and abdominal fatness.

2.3. Literature Cited

ABE, T., J. SABURI, H. HASEBE, T. NAKAGAWA, T. KAWAMURA *et al.*, 2008 Bovine quantitative trait loci analysis for growth, carcass, and meat quality traits in an F-2 population from a cross between Japanese Black and Limousin. Journal of Animal Science **86:** 2821-2832.

ALMASY, L., and J. BLANGERO, 1998 Multipoint quantitative-trait linkage analysis in general pedigrees. American Journal of Human Genetics **62:** 1198-1211.

ALTSHULER, D., M. J. DALY and E. S. LANDER, 2008 Genetic Mapping in Human Disease. Science **322:** 881-888.

AMARAL, A. J., H. J. MEGENS, R. P. M. A. CROOIJMANS, H. C. M. HEUVEN and M. A. M. GROENEN, 2008 Linkage disequilibrium decay and haplotype block structure in the pig. Genetics **179:** 569-579.

AMOS, C. I., 1994 Robust variance-components approach for assessing genetic linkage in pedigrees. American Journal of Human Genetics **54:** 535-543.

ANDERSSON, L., and M. GEORGES, 2004 Domestic-animal genomics: deciphering the genetics of complex traits. Nature Reviews Genetics **5:** 202-212.

ANKRA-BADU, G. A., D. SHRINER, E. LE BIHAN-DUVAL, S. MIGNON-GRASTEAU, F. PITEL *et al.*, 2010 Mapping main, epistatic and sex-specific QTL for

body composition in a chicken population divergently selected for low or high growth rate. BMC Genomics **11:** 107.

ARDLIE, K. G., L. KRUGLYAK and M. SEIELSTAD, 2002 Patterns of linkage disequilibrium in the human genome. Nature Reviews Genetics **3:** 299-309.

BARENDSE, W., B. E. HARRISON, R. J. HAWKEN, D. M. FERGUSON, J. M. THOMPSON *et al.*, 2007 Epistasis between calpain 1 and its inhibitor calpastatin within breeds of cattle. Genetics **176:** 2601-2610.

BLOTT, S., J. J. KIM, S. MOISIO, A. SCHMIDT-KUNTZEL, A. CORNET *et al.*, 2003 Molecular dissection of a quantitative trait locus: A phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics **163:** 253-266.

BOER, M. P., C. J. F. TER BRAAK and R. C. JANSEN, 2002 A penalized likelihood method for mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics **162:** 951-960.

BOHMANOVA, J., M. SARGOLZAEI and F. S. SCHENKEL, 2010 Characteristics of linkage disequilibrium in North American Holsteins. BMC Genomics **11**.

BOLORMAA, S., L. R. P. NETO, Y. D. ZHANG, R. J. BUNCH, B. E. HARRISON *et al.*, 2011 A genome-wide association study of meat and carcass traits in Australian cattle. Journal of Animal Science **89:** 2297-2309.

BOLORMAA, S., J. E. PRYCE, B. J. HAYES and M. E. GODDARD, 2010 Multivariate analysis of a genome-wide association study in dairy cattle. Journal of Dairy Science **93:** 3818-3833.

BOUWMAN, A. C., H. BOVENHUIS, M. H. P. W. VISKER and J. A. M. VAN ARENDONK, 2011 Genome-wide association of milk fatty acids in Dutch dairy cattle. BMC Genetics **12**.

BUCHANAN, F. C., T. D. THUE, P. YU and D. C. WINKELMAN-SIM, 2005 Single nucleotide polymorphisms in the corticotrophin-releasing hormone and pro-opiomelancortin genes are associated with growth and carcass yield in beef cattle. Animal Genetics **36:** 127-131.

CARLBORG, O., L. ANDERSSON and B. KINGHORN, 2000 The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. Genetics **155:** 2003-2010.

CARLBORG, O., G. A. BROCKMANN and C. S. HALEY, 2005 Simultaneous mapping of epistatic QTL in DU6i x DBA/2 mice. Mammalian Genome **16:** 481-494.

CARLBORG, O., and C. S. HALEY, 2004 Epistasis: too often neglected in complex trait studies? Nature Reviews Genetics **5:** 618-U614.

CARLBORG, O., S. KERJE, K. SCHUTZ, L. JACOBSSON, P. JENSEN *et al.*, 2003 A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome Research **13:** 413-421.

CARLSON, C. S., M. A. EBERLE, M. J. RIEDER, J. D. SMITH, L. KRUGLYAK *et al.*, 2003 Additional SNPs and linkage-disequilibrium analyses are necessary

for whole-genome association studies in humans. Nature Genetics **33:** 518-521.

CASAS, E., J. W. KEELE, S. D. SHACKELFORD, M. KOOHMARAIE, T. S. SONSTEGARD *et al.*, 1998 Association of the muscle hypertrophy locus with carcass traits in beef cattle. Journal of Animal Science **76:** 468-473.

CASAS, E., J. W. KEELE, S. D. SHACKELFORD, M. KOOHMARAIE and R. T. STONE, 2004 Identification of quantitative trait loci for growth and carcass composition in cattle. Animal Genetics **35:** 2-6.

CASAS, E., S. D. SHACKELFORD, J. W. KEELE, M. KOOHMARAIE, T. P. L. SMITH *et al.*, 2003 Detection of quantitative trait loci for growth and carcass composition in cattle. Journal of Animal Science **81:** 2976-2983.

CASAS, E., S. D. SHACKLEFORD, J. W. KEELE, R. T. STONE, S. M. KAPPES *et al.*, 2000 Quantitative trait loci affecting growth and carcass composition of cattle segregating alternate forms of myostatin. Journal of Animal Science **78:** 560-569.

CASAS, E., R. T. STONE, J. W. KEELE, S. D. SHACKELFORD, S. M. KAPPES *et al.*, 2001 A comprehensive search for quantitative trait loci affecting growth and carcass composition of cattle segregating alternative forms of the myostatin gene. Journal of Animal Science **79:** 854-860.

CASAS, E., S. N. WHITE, T. L. WHEELER, S. D. SHACKELFORD, M. KOOHMARAIE *et al.*, 2006 Effects of calpastatin and micro-calpain markers in beef cattle on tenderness traits. Journal of Animal Science **84:** 520-525.

CASELLA, G., and E. I. GEORGE, 1992 Explaining the Gibbs Sampler. American Statistician **46:** 167-174.

CHEONG, H. S., D. H. YOON, J. S. BAE, L. H. KIM, E. M. KIM *et al.*, 2009 ADFP promoter polymorphism associated with marbling score in Korean cattle. BMB Reports **42:** 529-534.

COHEN-ZINDER, M., E. SEROUSSI, D. M. LARKIN, J. J. LOOR, A. EVERTS-VAN DER WIND *et al.*, 2005 Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Research **15:** 936-944.

COLE, J. B., G. R. WIGGANS, L. MA, T. S. SONSTEGARD, T. J. LAWLOR, JR. *et al.*, 2011 Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. BMC Genomics **12:** 408.

CORVA, P. M., G. V. F. MACEDO, L. A. SORIA, J. P. MAZZUCCO, M. MOTTER *et al.*, 2009 Effect of leptin gene polymorphisms on growth, slaughter and meat quality traits of grazing Brangus steers. Genetics and Molecular Research **8:** 105-116.

CULVERHOUSE, R., B. K. SUAREZ, J. LIN and T. REICH, 2002 A perspective on epistasis: Limits of models displaying no main effect. American Journal of Human Genetics **70:** 461-471.

DALY, M. J., J. D. RIOUX, S. E. SCHAFFNER, T. J. HUDSON and E. S. LANDER, 2001 High-resolution haplotype structure in the human genome. Nature Genetics **29:** 229-232.

DE ROOS, A. P., B. J. HAYES, R. J. SPELMAN and M. E. GODDARD, 2008 Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics **179:** 1503-1512.

DE, S., M. D. MACNEIL, X. L. WU, J. J. MICHAL, Q. X. XIAO *et al.*, 2004 Detection of quantitative trait loci for marbling and backfat in Wagyu x Limousin F2 crosses using a candidate gene approach. Journal of Animal Science **82:** 111-111.

DEKKERS, J. C. M., and F. HOSPITAL, 2002 The use of molecular genetics in the improvement of agricultural populations. Nature Reviews Genetics **3:** 22-32.

DONG, C. H., S. WANG, W. D. LI, D. LI, H. Y. ZHAO *et al.*, 2003 Interacting genetic loci on chromosomes 20 and 10 influence extreme human obesity. American Journal of Human Genetics **72:** 115-124.

DONNELLY, P., 2008 Progress and challenges in genome-wide association studies in humans. Nature **456:** 728-731.

DRUET, T., S. FRITZ, M. BOUSSAHA, S. BEN-JEMAA, F. GUILLAUME *et al.*, 2008 Fine mapping of quantitative trait loci affecting female fertility in dairy cattle on BTA03 using a dense single-nucleotide polymorphism map. Genetics **178:** 2227-2235.

DYBUS, A., W. GRZESIAK, H. KAMIENIECKI, I. SZATKOWSKA, Z. SOBEK *et al.*, 2005 Association of genetic variants of bovine prolactin with milk production traits of Black-and-White and Jersey cattle. Archives of Animal Breeding **48:** 149-156.

FAN, B., S. K. ONTERU, Z. Q. DU, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs. PLoS One **6:** e14726.

FARNIR, F., W. COPPIETERS, J. J. ARRANZ, P. BERZI, N. CAMBISANO *et al.*, 2000 Extensive genome-wide linkage disequilibrium in cattle. Genome Research **10:** 220-227.

FARNIR, F., B. GRISART, W. COPPIETERS, J. RIQUET, P. BERZI *et al.*, 2002 Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: Revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. Genetics **161:** 275-287.

FERNANDEZ, J. R., L. M. TARANTINO, S. M. HOFER, G. P. VOGLER and G. E. MCCLEARN, 2000 Epistatic quantitative trait loci for alcohol preference in mice. Behavior Genetics **30:** 431-437.

FERNANDO, R. L., and D. J. GARRICK, 2009 GenSel - User manual for a portfolio of genomic selection related analyses. Animal Breeding and Genetics, Iowa State University, Ames.

FIJNEMAN, R. J. A., S. S. DEVRIES, R. C. JANSEN and P. DEMANT, 1996 Complex interactions of new quantitative trait loci, Sluc1, Sluc2, Sluc3, and Sluc4,

that influence the susceptibility to lung cancer in the mouse. Nature Genetics **14:** 465-467.

GABRIEL, S. B., S. F. SCHAFFNER, H. NGUYEN, J. M. MOORE, J. ROY *et al.*, 2002a The structure of haplotype blocks in the human genome. Science **296:** 2225-2229.

GABRIEL, S. B., S. F. SCHAFFNER, D. E. REICH, P. SABETI, M. FREEDMAN *et al.*, 2002b An extended analysis of haplotype structure in the human genome. American Journal of Human Genetics **71:** 205-205.

GEMAN, S., and D. GEMAN, 1984 Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Machine Intelligence **6:** 721-741.

GEORGE, A. W., P. M. VISSCHER and C. S. HALEY, 2000 Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. Genetics **156:** 2081-2092.

GEORGES, M., D. NIELSEN, M. MACKINNON, A. MISHRA, R. OKIMOTO *et al.*, 1995 Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics **139:** 907-920.

GIANOLA, D., G. DE LOS CAMPOS, W. G. HILL, E. MANFREDI and R. FERNANDO, 2009 Additive genetic variability and the Bayesian alphabet. Genetics **183:** 347-363.

GILL, J. L., S. C. BISHOP, C. MCCORQUODALE, J. L. WILLIAMS and P. WIENER, 2009 Association of selected SNP with carcass and taste panel assessed

meat quality traits in a commercial population of Aberdeen Angus-sired beef cattle. Genetics Selection Evolution **41**.

GODDARD, M. E., 1991 Mapping Genes for Quantitative Traits Using Linkage Disequilibrium. Genetics Selection Evolution **23:** S131-S134.

GODDARD, M. E., and B. J. HAYES, 2009 Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics **10:** 381-391.

GOLDSTEIN, D. B., and M. E. WEALE, 2001 Population genomics: Linkage disequilibrium holds the key. Current Biology **11:** R576-R579.

GREENAWALT, D. M., X. F. CUI, Y. J. WU, Y. LIN, H. Y. WANG *et al.*, 2006 Strong correlation between meiotic crossovers and haplotype structure in a 2.5-Mb region on the long arm of chromosome 21. Genome Research **16:** 208-214.

GREENWOOD, T. A., B. K. RANA and N. J. SCHORK, 2004 Human haplotype block sizes are negatively correlated with recombination rates. Genome Research **14:** 1358-1361.

GRIGNOLA, F. E., I. HOESCHELE and B. TIER, 1996 Mapping quantitative trait loci in outcross populations via residual maximum likelihood .1. Methodology. Genetics Selection Evolution **28:** 479-490.

GRISART, B., W. COPPIETERS, F. FARNIR, L. KARIM, C. FORD *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research **12:** 222-231.

GROSSE-BRINKHAUS, C., E. JONAS, H. BUSCHBELL, C. PHATSARA, D. TESFAYE *et al.*, 2010 Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc x Pietrain population. Genetics Selection Evolution **42**.

GURYEV, V., B. M. G. SMITS, J. VAN DE BELT, M. VERHEUL, N. HUBNER *et al.*, 2006 Haplotype block structure is conserved across mammals. PLoS Genetics **2:** 1111-1118.

GUTIERREZ-GIL, B., J. L. WILLIAMS, D. HOMER, D. BURTON, C. S. HALEY *et al.*, 2009 Search for quantitative trait loci affecting growth and carcass traits in a cross population of beef and dairy cattle. Journal of Animal Science **87:** 24-36.

HABIER, D., R. L. FERNANDO, K. KIZILKAYA and D. J. GARRICK, 2011 Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics **12**.

HALE, C. S., W. O. HERRING, H. SHIBUYA, M. C. LUCY, D. B. LUBAHN *et al.*, 2000 Decreased growth in angus steers with a short TG-microsatellite allele in the P1 promoter of the growth hormone receptor gene. Journal of Animal Science **78:** 2099-2104.

HALEY, C. S., and S. A. KNOTT, 1992 A Simple Regression Method for Mapping Quantitative Trait Loci in Line Crosses Using Flanking Markers. Heredity **69:** 315-324.

HAN, L., and S. XU, 2010 Genome-wide evaluation for quantitative trait loci under the variance component model. Genetica **138:** 1099-1109.

HASTINGS, W. K., 1970 Monte-Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika **57:** 97-109.

HAYES, B. J., P. M. VISSCHER, H. C. MCPARTLAN and M. E. GODDARD, 2003 Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Research **13:** 635-643.

HERNANDEZ-SANCHEZ, J., J. A. GRUNCHEC and S. KNOTT, 2009 A web application to perform linkage disequilibrium and linkage analyses on a computational grid. Bioinformatics **25:** 1377-1383.

HEYEN, D. W., J. I. WELLER, M. RON, M. BAND, J. E. BEEVER *et al.*, 1999 A genome scan for QTL influencing milk production and health traits in dairy cattle. Physiological Genomics **1:** 165-175.

HILL, W. G., 1981 Estimation of Effective Population-Size from Data on Linkage Disequilibrium. Genetical Research **38:** 209-216.

HILL, W. G., and A. ROBERTSON, 1968 Linkage disequilibrium in finite populations. Theoretical and Applied Genetics **38:** 226--231.

HOESCHELE, I., P. UIMARI, F. E. GRIGNOLA, Q. ZHANG and K. M. GAGE, 1997 Advances in statistical methods to map quantitative trait loci in outbred populations. Genetics **147:** 1445-1457.

HOLLAND, J. B., 2007 Genetic architecture of complex traits in plants. Current Opinion in Plant Biology **10:** 156-161.

HOLLAND, J. B., H. S. MOSER, L. S. ODONOUGHUE and M. LEE, 1997 QTLs and epistasis associated with vernalization responses in oat. Crop Science **37:** 1306-1316.

HU, Z., and S. XU, 2009 PROC QTL-A SAS Procedure for Mapping Quantitative Trait Loci. International Journal of Plant Genomics **2009:** 141234.

HU, Z. Q., Y. G. LI, X. H. SONG, Y. P. HAN, X. D. CAI *et al.*, 2011 Genomic value prediction for quantitative traits under the epistatic model. BMC Genetics **12**.

ISLAM, K. K., M. VINSKY, R. E. CREWS, E. OKINE, S. S. MOORE *et al.*, 2009 Association analyses of a SNP in the promoter of IGF1 with fat deposition and carcass merit traits in hybrid, Angus and Charolais beef cattle. Animal Genetics **40:** 766-769.

JANNINK, J. L., and R. JANSEN, 2001 Mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics **157:** 445-454.

JANSEN, R. C., 1993 Interval Mapping of Multiple Quantitative Trait Loci. Genetics **135:** 205-211.

JANSEN, R. C., and P. STAM, 1994 High-Resolution of Quantitative Traits into Multiple Loci Via Interval Mapping. Genetics **136:** 1447-1455.

JEFFREYS, A. J., L. KAUPPI and R. NEUMANN, 2001 Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nature Genetics **29:** 217-222.

JEFFREYS, A. J., R. NEUMANN, M. PANAYI, S. MYERS and P. DONNELLY, 2005 Human recombination hot spots hidden in regions of strong marker association. Nature Genetics **37:** 601-606.

JIANG, C. J., and Z. B. ZENG, 1997 Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. Genetica **101:** 47-58.

JIANG, L., J. F. LIU, D. X. SUN, P. P. MA, X. D. DING *et al.*, 2010 Genome Wide Association Studies for Milk Production Traits in Chinese Holstein Population. PLoS One **5**.

JIANG, Z. H., J. J. MICHAL, D. J. TOBEY, T. F. DANIELS, D. C. RULE *et al.*, 2008 Significant associations of stearoyl-CoA desaturase (SCD1) gene with fat deposition and composition in skeletal muscle. International Journal of Biological Sciences **4:** 345-351.

KAO, C. H., Z. B. ZENG and R. D. TEASDALE, 1999 Multiple interval mapping for quantitative trait loci. Genetics **152:** 1203-1216.

KAUPE, B., H. BRANDT, E. M. PRINZENBERG and G. ERHARDT, 2007 Joint analysis of the influence of CYP11B1 and DGAT1 genetic variation on milk production, somatic cell score, conformation, reproduction, and productive lifespan in German Holstein cattle. Journal of Animal Science **85:** 11-21.

KEELE, J. W., S. D. SHACKELFORD, S. M. KAPPES, M. KOOHMARAIE and R. T. STONE, 1999 A region on bovine chromosome 15 influences beef longissimus tenderness in steers. Journal of Animal Science **77:** 1364-1371.

KHATIB, H., S. D. LEONARD, V. SCHUTZKUS, W. LUO and Y. M. CHANG, 2006 Association of the OLR1 gene with milk composition in holstein dairy cattle. Journal of Dairy Science **89:** 1753-1760.

KHATIB, H., I. ZAITOUN, J. WIEBELHAUS-FINGER, Y. M. CHANG and G. J. M. ROSA, 2007 The association of bovine PPARGC1A and OPN genes with milk composition in two independent holstein cattle populations. Journal of Dairy Science **90:** 2966-2970.

KHATKAR, M. S., A. COLLINS, J. A. L. CAVANAGH, R. J. HAWKEN, M. HOBBS *et al.*, 2006a A first-generation metric linkage disequilibrium map of bovine chromosome 6. Genetics **174:** 79-85.

KHATKAR, M. S., F. W. NICHOLAS, A. R. COLLINS, K. R. ZENGER, J. A. CAVANAGH *et al.*, 2008 Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics **9:** 187.

KHATKAR, M. S., P. C. THOMSON, I. TAMMEN, J. A. L. CAVANAGH, F. W. NICHOLAS *et al.*, 2006b Linkage disequilibrium on chromosome 6 in Australian Holstein-Friesian cattle. Genetics Selection Evolution **38:** 463-477.

KHATKAR, M. S., K. R. ZENGER, M. HOBBS, R. J. HAWKEN, J. A. L. CAVANAGH *et al.*, 2007 A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in Holstein-Friesian cattle. Genetics **176:** 763-772.

KIM, E. S., and B. W. KIRKPATRICK, 2009 Linkage disequilibrium in the North American Holstein population. Animal Genetics **40:** 279-288.

KIM, J. J., F. FARNIR, J. SAVELL and J. F. TAYLOR, 2003 Detection of quantitative trait loci for growth and beef carcass fatness traits in a cross between Bos

taurus (Angus) and Bos indicus (Brahman) cattle. Journal of Animal Science **81:** 1933-1942.

KIM, Y., J. RYU, J. WOO, J. B. KIM, C. Y. KIM *et al.*, 2011 Genome-wide association study reveals five nucleotide sequence variants for carcass traits in beef cattle. Animal Genetics **42:** 361-365.

KIZILKAYA, K., R. L. FERNANDO and D. J. GARRICK, 2010 Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. Journal of Animal Science **88:** 544-551.

KOLBEHDARI, D., Z. WANG, J. R. GRANT, B. MURDOCH, A. PRASAD *et al.*, 2009 A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. Journal of Animal Breeding and Genetics **126:** 216-227.

KUITTINEN, H., M. J. SILLANPAA and O. SAVOLAINEN, 1997 Genetic basis of adaptation: flowering time in Arabidopsis thaliana. Theoretical and Applied Genetics **95:** 573-583.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian Factors Underlying Quantitative Traits Using Rflp Linkage Maps. Genetics **121:** 185-199.

LANDER, E. S., and N. J. SCHORK, 1994 Genetic Dissection of Complex Traits (Vol 265, Pg 2037, 1994). Science **266:** 353-353.

LARK, K. G., K. CHASE, F. ADLER, L. M. MANSUR and J. H. ORF, 1995 Interactions between Quantitative Trait Loci in Soybean in Which Trait Variation at One Locus Is Conditional Upon a Specific Allele at Another.

Proceedings of the National Academy of Sciences of the United States of America **92:** 4656-4660.

LE ROUZIC, A., J. M. ALVAREZ-CASTRO and O. CARLBORG, 2008 Dissection of the genetic architecture of body weight in chicken reveals the impact of epistasis on domestication traits. Genetics **179:** 1591-1599.

LEE, S. H., J. H. J. VAN DER WERF, S. H. LEE, E. W. PARK, S. J. OH *et al.*, 2010a Genetic polymorphisms of the bovine Fatty acid binding protein 4 gene are significantly associated with marbling and carcass weight in Hanwoo (Korean Cattle). Animal Genetics **41:** 442-444.

LEE, Y. M., C. M. HAN, Y. LI, J. J. LEE, L. H. KIM *et al.*, 2010b A Whole Genome Association Study to Detect Single Nucleotide Polymorphisms for Carcass Traits in Hanwoo Populations. Asian-Australasian Journal of Animal Sciences **23:** 417-424.

LEWONTIN, R. C., 1964 The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics **49:** 49-67.

LEWONTIN, R. C., and K. KOJIMA, 1960 The Evolutionary Dynamics of Complex Polymorphisms. Evolution **14:** 458-472.

LINDBLAD-TOH, K., C. M. WADE, T. S. MIKKELSEN, E. K. KARLSSON, D. B. JAFFE *et al.*, 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature **438:** 803-819.

LYNCH, M., and B. WALSH, 1998 *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, Mass.

MACEACHERN, S., B. HAYES, J. MCEWAN and M. GODDARD, 2009 An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (Bos taurus) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle. BMC Genomics **10**.

MACNEIL, M. D., and M. D. GROSZ, 2002 Genome-wide scans for QTL affecting carcass traits in Hereford x composite double backcross populations. Journal of Animal Science **80:** 2316-2324.

MAI, M. D., G. SAHANA, F. B. CHRISTIANSEN and B. GULDBRANDTSEN, 2010 A genome-wide association study for milk production traits in Danish Jersey cattle using a 50K single nucleotide polymorphism chip. Journal of Animal Science **88:** 3522-3528.

MARCHINI, J., L. R. CARDON, M. S. PHILLIPS and P. DONNELLY, 2004 The effects of human population structure on large genetic association studies. Nature Genetics **36:** 512-517.

MARQUES, E., R. D. SCHNABEL, P. STOTHARD, D. KOLBEHDARI, Z. WANG *et al.*, 2008 High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle. BMC Genetics **9:** 45.

MCCLURE, M. C., N. S. MORSCI, R. D. SCHNABEL, J. W. KIM, P. YAO *et al.*, 2010 A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Animal Genetics **41:** 597-607.

MCKAY, S. D., R. D. SCHNABEL, B. M. MURDOCH, L. K. MATUKUMALLI, J. AERTS *et al.*, 2007 Whole genome linkage disequilibrium maps in cattle. BMC Genetics **8:** 74.

MCRAE, A. F., J. C. MCEWAN, K. G. DODDS, T. WILSON, A. M. CRAWFORD *et al.*, 2002 Linkage disequilibrium in domestic sheep. Genetics **160:** 1113-1122.

MELE, M., G. CONTE, B. CASTIGLIONI, S. CHESSA, N. P. P. MACCIOTTA *et al.*, 2007 Stearoyl-Coenzyme A desaturase gene polymorphism and milk fatty acid composition in Italian Holsteins. Journal of Dairy Science **90:** 4458-4465.

METROPOLIS, N., A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER and E. TELLER, 1953 Equation of State Calculations by Fast Computing Machines. Journal of Chemical Physics **21:** 1087-1092.

MEUWISSEN, T. H., and M. E. GODDARD, 2004 Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. Genetics Selection Evolution **36:** 261-279.

MEUWISSEN, T. H., A. KARLSEN, S. LIEN, I. OLSAKER and M. E. GODDARD, 2002 Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. Genetics **161:** 373-379.

MEUWISSEN, T. H. E., and M. E. GODDARD, 2000 Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. Genetics **155:** 421-430.

MEUWISSEN, T. H. E., and M. E. GODDARD, 2001 Prediction of identity by descent probabilities from marker-haplotypes. Genetics Selection Evolution **33:** 605-634.

MIZOSHITA, K., A. TAKANO, T. WATANABE, A. TAKASUGA and Y. SUGIMOTO, 2005 Identification of a 1.1-Mb region for a carcass weight QTL on bovine Chromosome 14. Mammalian Genome **16:** 532-537.

MIZOSHITA, K., T. WATANABE, H. HAYASHI, C. KUBOTA, H. YAMAKUCHI *et al.*, 2004 Quantitative trait loci analysis for growth and carcass traits in a half-sib family of purebred Japanese Black (Wagyu) cattle. Journal of Animal Science **82:** 3415-3420.

MOIOLI, B., G. CONTARINI, A. AVALLI, G. CATILLO, L. ORRU *et al.*, 2007 Short communication: Effect of stearoyl-coenzyme A desaturase polymorphism on fatty acid composition of milk. Journal of Dairy Science **90:** 3553-3558.

MOORE, J. H., 2005 A global view of epistasis. Nature Genetics **37:** 13-14.

MOORE, S. S., C. LI, J. BASARAB, W. M. SNELLING, J. KNEELAND *et al.*, 2003 Fine mapping of quantitative trait loci and assessment of positional candidate genes for backfat on bovine chromosome 14 in a commercial line of Bos taurus. Journal of Animal Science **81:** 1919-1925.

MORRIS, C. A., N. G. CULLEN, B. C. GLASS, D. L. HYNDMAN, T. R. MANLEY *et al.*, 2007 Fatty acid synthase effects on bovine adipose fat and milk fat. Mammalian Genome **18:** 64-74.

MYLES, S., J. PEIFFER, P. J. BROWN, E. S. ERSOZ, Z. W. ZHANG *et al.*, 2009 Association Mapping: Critical Considerations Shift from Genotyping to Experimental Design. Plant Cell **21:** 2194-2202.

NAGASE, H., J. H. MAO, J. P. DE KONING, T. MINAMI and A. BALMAIN, 2001 Epistatic interactions between skin tumor modifier loci in interspecific (spretus/musculus) backcross mice. Cancer Research **61:** 1305-1308.

NALAILA, S. M., P. STOTHARD, S. S. MOORE, Z. Q. WANG and C. X. LI, 2011 Whole genome fine mapping of quantitative trait loci for ultrasound and carcass merit traits in beef cattle. Canadian Journal of Animal Science **91:** 61-73.

NKRUMAH, J. D., C. LI, J. YU, C. HANSEN, D. H. KEISLER *et al.*, 2005 Polymorphisms in the bovine leptin promoter associated with serum leptin concentration, growth, feed intake, feeding behavior, and measures of carcass merit. Journal of Animal Science **83:** 20-28.

ODANI, M., A. NARITA, T. WATANABE, K. YOKOUCHI, Y. SUGIMOTO *et al.*, 2006 Genome-wide linkage disequilibrium in two Japanese beef cattle breeds. Animal Genetics **37:** 139-144.

OLSEN, H. G., S. LIEN, M. GAUTIER, H. NILSEN, A. ROSETH *et al.*, 2005 Mapping of a milk production quantitative trait locus to a 420-kb region on bovine chromosome 6. Genetics **169:** 275-283.

OLSEN, H. G., S. LIEN, M. SVENDSEN, H. NILSEN, A. ROSETH *et al.*, 2004 Fine mapping of milk production QTL on BTA6 by combined linkage and linkage disequilibrium analysis. Journal of Dairy Science **87:** 690-698.

OLSEN, H. G., H. NILSEN, B. HAYES, P. R. BERG, M. SVENDSEN *et al.*, 2007 Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. BMC Genetics **8**.

ONTERU, S. K., B. FAN, M. T. NIKKILA, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Whole-genome association analyses for lifetime reproductive traits in the pig. Journal of Animal Science **89:** 988-995.

PAGE, B. T., E. CASAS, M. P. HEATON, N. G. CULLEN, D. L. HYNDMAN *et al.*, 2002 Evaluation of single-nucleotide polymorphisms in CAPN1 for association with meat tenderness in cattle. Journal of Animal Science **80:** 3077-3085.

PATIL, N., A. J. BERNO, D. A. HINDS, W. A. BARRETT, J. M. DOSHI *et al.*, 2001 Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science **294:** 1719-1723.

PHILLIPS, M. S., R. LAWRENCE, R. SACHIDANANDAM, A. P. MORRIS, D. J. BALDING *et al.*, 2003 Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. Nature Genetics **33:** 382-387.

PRITCHARD, J. K., and M. PRZEWORSKI, 2001 Linkage disequilibrium in humans: Models and data. American Journal of Human Genetics **69:** 1-14.

PRYCE, J. E., S. BOLORMAA, A. J. CHAMBERLAIN, P. J. BOWMAN, K. SAVIN *et al.*, 2010 A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. Journal of Dairy Science **93:** 3331-3345.

REICH, D. E., M. CARGILL, S. BOLK, J. IRELAND, P. C. SABETI *et al.*, 2001 Linkage disequilibrium in the human genome. Nature **411:** 199-204.

REICH, D. E., S. F. SCHAFFNER, M. J. DALY, G. MCVEAN, J. C. MULLIKIN *et al.*, 2002 Human genome sequence variation and the influence of gene history, mutation and recombination. Nature Genetics **32:** 135-142.

RIQUET, J., W. COPPIETERS, N. CAMBISANO, J. J. ARRANZ, P. BERZI *et al.*, 1999 Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle. Proceedings of the National Academy of Sciences of the United States of America **96:** 9252-9257.

ROY, R., L. ORDOVAS, P. ZARAGOZA, A. ROMERO, C. MORENO *et al.*, 2006 Association of polymorphisms in the bovine FASN gene with milk-fat content. Animal Genetics **37:** 215-218.

SARGOLZAEI, M., F. S. SCHENKEL, G. B. JANSEN and L. R. SCHAEFFER, 2008 Extent of linkage disequilibrium in Holstein cattle in North America. Journal of Dairy Science **91:** 2106-2117.

SATAGOPAN, J. M., Y. S. YANDELL, M. A. NEWTON and T. C. OSBORN, 1996 A Bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo. Genetics **144:** 805-816.

SCHENKEL, F. S., S. P. MILLER, X. YE, S. S. MOORE, J. D. NKRUMAH *et al.*, 2005 Association of single nucleotide polymorphisms in the leptin gene with carcass and meat quality traits of beef cattle. Journal of Animal Science **83:** 2009-2020.

SELF, S. G., and K. Y. LIANG, 1987 Asymptotic Properties of Maximum-Likelihood Estimators and Likelihood Ratio Tests under Nonstandard Conditions. Journal of the American Statistical Association **82:** 605-610.

SELLNER, E. M., J. W. KIM, M. C. MCCLURE, K. H. TAYLOR, R. D. SCHNABEL *et al.*, 2007 Board-invited review: Applications of genomic information in livestock. Journal of Animal Science **85:** 3148-3158.

SEN, S., and G. A. CHURCHILL, 2001 A statistical framework for quantitative trait mapping. Genetics **159:** 371-387.

SETOGUCHI, K., M. FURUTA, T. HIRANO, T. NAGAO, T. WATANABE *et al.*, 2009 Cross-breed comparisons identified a critical 591-kb region for bovine carcass weight QTL (CW-2) on chromosome 6 and the Ile-442-Met substitution in NCAPG as a positional candidate. BMC Genetics **10**.

SILLANPAA, M. J., and E. ARJAS, 1998 Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. Genetics **148:** 1373-1388.

SLATKIN, M., 2008 Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics **9:** 477-485.

SMITH, M. W., and S. J. O'BRIEN, 2005 Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. Nature Reviews Genetics **6:** 623-632.

SOLLER, M., T. BRODY and A. GENIZI, 1976 Power of Experimental Designs for Detection of Linkage between Marker Loci and Quantitative Loci in Crosses between Inbred Lines. Theoretical and Applied Genetics **47:** 35-39.

STEPHENS, D. A., 1998 Bayesian analysis of quantitative trait locus data using reversible jump Markov chain Monte Carlo. Biometrics **54:** 1334-1347.

STONE, R. T., J. W. KEELE, S. D. SHACKELFORD, S. M. KAPPES and M. KOOHMARAIE, 1999 A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. Journal of Animal Science **77:** 1379-1384.

STRUCKEN, E. M., R. H. BORTFELDT, D. J. DE KONING and G. A. BROCKMANN, 2011 Genome-wide associations for investigating time-dependent genetic effects for milk production traits in dairy cattle. Animal Genetics**:** 1-8.

SUGIYAMA, F., G. A. CHURCHILL, D. C. HIGGINS, C. JOHNS, K. P. MAKARITSIS *et al.*, 2001 Concordance of murine quantitative trait loci for salt-induced hypertension with rat and human loci. Genomics **71:** 70-77.

SUN, X., D. HABIER, R. L. FERNANDO, D. J. GARRICK and J. C. DEKKERS, 2011 Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. BMC Proceedings **5 Suppl 3:** S13.

SUNYAEV, S., F. A. KONDRASHOV, P. BORK and V. RAMENSKY, 2003 Impact of selection, mutation rate and genetic drift on human genetic variation. Human Molecular Genetics **12:** 3325-3330.

TAKASUGA, A., T. WATANABE, Y. MIZOGUCHI, T. HIRANO, N. IHARA *et al.*, 2007 Identification of bovine QTL for growth and carcass traits in Japanese Black cattle by replication and identical-by-descent mapping. Mammalian Genome **18:** 125-136.

TANKSLEY, S. D., 1993 Mapping Polygenes. Annual Review of Genetics **27:** 205-233.

TAYLOR, J. F., L. L. COUTINHO, K. L. HERRING, D. S. GALLAGHER, R. A. BRENNEMAN *et al.*, 1998 Candidate gene analysis of GH1 for effects on growth and carcass composition of cattle. Animal Genetics **29:** 194-201.

TENESA, A., S. A. KNOTT, D. WARD, D. SMITH, J. L. WILLIAMS *et al.*, 2003 Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes. Journal of Animal Science **81:** 617-623.

TERWILLIGER, J. D., S. ZOLLNER, M. LAAN and S. PAABO, 1998 Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'Drift mapping' in small populations with no demographic expansion. Human Heredity **48:** 138-154.

THALLER, G., C. KUHN, A. WINTER, G. EWALD, O. BELLMANN *et al.*, 2003 DGAT1, a new positional and functional candidate gene for intramuscular fat deposition in cattle. Animal Genetics **34:** 354-357.

THOMAS, M. G., R. M. ENNS, K. L. SHIRLEY, M. D. GARCIA, A. J. GARRETT *et al.*, 2007 Associations of DNA polymorphisms in growth hormone and its transcriptional regulators with growth and carcass traits in two populations of Brangus bulls. Genetics and Molecular Research **6:** 222-237.

TURNER, L. B., B. E. HARRISON, R. J. BUNCH, L. R. P. NETO, Y. T. LI *et al.*, 2010 A genome-wide association study of tick burden and milk composition in cattle. Animal Production Science **50:** 235-245.

UEMOTO, Y., S. SATO, C. OHNISHI, S. TERAI, A. KOMATSUDA *et al.*, 2009 The effects of single and epistatic quantitative trait loci for fatty acid composition in a Meishan x Duroc crossbred population. Journal of Animal Science **87:** 3470-3476.

UIMARI, P., and I. HOESCHELE, 1997 Mapping-linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. Genetics **146:** 735-743.

UIMARI, P., G. THALLER and I. HOESCHELE, 1996 The use of multiple markers in a Bayesian method for mapping quantitative trait loci. Genetics **143:** 1831-1842.

VALLEJO, R. L., Y. L. LI, G. W. ROGERS and M. S. ASHWELL, 2003 Genetic diversity and background linkage disequilibrium in the North American Holstein cattle population. Journal of Dairy Science **86:** 4137-4147.

VAN ARENDONK, J. A. M., B. TIER, M. C. A. M. BINK and H. BOVENHUIS, 1998 Restricted maximum likelihood analysis of linkage between genetic markers and quantitative trait loci for a granddaughter design. Journal of Dairy Science **81:** 76-84.

VILLA-ANGULO, R., L. K. MATUKUMALLI, C. A. GILL, J. CHOI, C. P. VAN TASSELL *et al.*, 2009 High-resolution haplotype block structure in the cattle genome. BMC Genetics **10**.

WANG, D. L., J. ZHU, Z. K. LI and A. H. PATERSON, 1999 Mapping QTLs with epistatic effects and QTLxenvironment interactions by mixed linear model approaches. Theoretical and Applied Genetics **99:** 1255-1264.

WANG, H., Y. M. ZHANG, X. LI, G. L. MASINDE, S. MOHAN *et al.*, 2005 Bayesian shrinkage estimation of quantitative trait loci parameters. Genetics **170:** 465-480.

WANG, N., J. M. AKEY, K. ZHANG, R. CHAKRABORTY and L. JIN, 2002 Distribution of recombination crossovers and the origin of haplotype blocks: The interplay of population history, recombination, and mutation. American Journal of Human Genetics **71:** 1227-1234.

WEIKARD, R., C. KUHN, T. GOLDAMMER, G. FREYER and M. SCHWERIN, 2005 The bovine PPARGC1A gene: molecular characterization and association of an SNP with variation of milk fat synthesis. Physiological Genomics **21:** 1-13.

WIBOWO, T. A., J. J. MICHAL and Z. JIANG, 2007 Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. Genome **50:** 939-945.

WINTER, A., W. KRAMER, F. A. O. WERNER, S. KOLLERS, S. KATA *et al.*, 2002 Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA : diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proceedings of the National Academy of Sciences of the United States of America **99:** 9300-9305.

WOOD, I. A., G. MOSER, D. L. BURRELL, K. L. MENGERSEN and D. J. S. HETZEL, 2006 A meta-analytic assessment of a Thyroglobulin marker for marbling in beef cattle. Genetics Selection Evolution **38:** 479-494.

WU, R. L., C. X. MA and G. CASELLA, 2002 Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. Genetics **160:** 779-792.

WU, R. L., and Z. B. ZENG, 2001 Joint linkage and linkage disequilibrium mapping in natural populations. Genetics **157:** 899-909.

WU, X. L., M. D. MACNEIL, S. DE, Q. J. XIAO, J. J. MICHAL *et al.*, 2005 Evaluation of candidate gene effects for beef backfat via Bayesian model selection. Genetica **125:** 103-113.

XU, S. Z., 2003 Estimating polygenic effects using markers of the entire genome. Genetics **163:** 789-801.

XU, S. Z., 2007 An empirical Bayes method for estimating epistatic effects of quantitative trait loci. Biometrics **63:** 513-521.

XU, S. Z., and Z. Y. JIA, 2007 Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics **175:** 1955-1963.

YI, N. J., A. DIAMENT, S. CHIU, K. KIM, D. B. ALLISON *et al.*, 2004 Characterization of epistasis influencing complex spontaneous obesity in the BSB model. Genetics **167:** 399-409.

YI, N. J., D. SHRINER, S. BANERJEE, T. MEHTA, D. POMP *et al.*, 2007 An efficient Bayesian model selection approach for interacting quantitative trait loci models with many effects. Genetics **176:** 1865-1877.

YI, N. J., and S. Z. XU, 2002a Linkage analysis of quantitative trait loci in multiple line crosses. Genetica **114:** 217-230.

YI, N. J., and S. Z. XU, 2002b Mapping quantitative trait loci with epistatic effects. Genetical Research **79:** 185-198.

YI, N. J., S. Z. XU and D. B. ALLISON, 2003 Bayesian model choice and search strategies for mapping interacting quantitative trait loci. Genetics **165:** 867-883.

YI, N. J., B. S. YANDELL, G. A. CHURCHILL, D. B. ALLISON, E. J. EISEN *et al.*, 2005 Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. Genetics **170:** 1333-1344.

YI, N. J., D. K. ZINNIEL, K. KIM, E. J. EISEN, A. BARTOLUCCI *et al.*, 2006 Bayesian analyses of multiple epistatic QTL models for body weight and body composition in mice. Genetical Research **87:** 45-60.

YU, S. B., J. X. LI, Y. F. TAN, Y. J. GAO, X. H. LI *et al.*, 1997 Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proceedings of the National Academy of Sciences of the United States of America **94:** 9226-9231.

ZENG, Z. B., 1994 Precision mapping of quantitative trait loci. Genetics **136:** 1457-1468.

ZENG, Z. B., C. H. KAO and C. J. BASTEN, 1999 Estimating the genetic architecture of quantitative traits. Genetical Research **74:** 279-289.

ZHANG, K., J. M. AKEY, N. WANG, M. XIONG, R. CHAKRABORTY *et al.*, 2003 Randomly distributed crossovers may generate block-like pattern's of linkage disequilibrium: an act of genetic drift. Human Genetics **113:** 51-59.

ZHANG, Q., D. BOICHARD, I. HOESCHELE, C. ERNST, A. EGGEN *et al.*, 1998 Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics **149:** 1959-1973.

ZHANG, Y. M., and S. XU, 2005 A penalized maximum likelihood method for estimating epistatic effects of QTL. Heredity **95:** 96-104.

ZOU, W., and Z. B. ZENG, 2008 Statistical Methods for Mapping Multiple QTL. International Journal of Plant Genomics **2008:** 286561.

Chapter 3. Genome-wide Comparative Analysis of Haplotype Block Structure in Two Canadian Cattle Populations

3.1. Introduction

With the availability of high-density SNPs and cost-effective genotyping technologies, high resolution LD maps and haplotype block structure are being characterized in different cattle populations to aid investigations of cattle evolution and to facilitate association studies (MCKAY *et al.* 2007; VILLA-ANGULO *et al.* 2009). Haplotype blocks have been characterized in 1000 Holstein-Friesian cattle using 9 K SNPs with an average intermarker spacing of 251.8 kb (KHATKAR *et al.* 2007). Haplotype block structure has also been characterized for specific chromosomes in cattle. One study reported the block structure in 14 European and African cattle breeds using 1,536 SNPs mainly localized on BTA3 with an average intermarker distance of 311 kb (GAUTIER *et al.* 2007). Another study characterized the haplotype block structure in 101 regions on BTA6, 14, and 25 (spanning up to 7.6 Mb) with an average marker density of one SNP per 4 kb for 501 animals sampled from 19 worldwide taurine and indicine breeds (VILLA-ANGULO *et al.* 2009), in which non-differentiable haplotype block structure between dairy and beef breeds was reported. However, the genome-wide haplotype block structure remains unclear for beef cattle populations and the comparison of block structure between dairy and beef cattle breeds across the whole bovine genome has not been performed.

A previous study suggests that domestication and artificial selection have left detectable signatures of selection within the cattle genome (GIBBS *et al.* 2009). Dairy and beef cattle breeds under distinct selection for different economically important traits may have different haplotype block patterns. Comparative haplotype maps could enable us to explore the degree of diversity between cattle breeds and to detect genomic regions that have been subject to selective sweeps (GURYEV *et al.* 2006; KHATKAR *et al.* 2007; MCKAY *et al.* 2007). It is also necessary to understand the genomic haplotype structure in designing and interpreting association studies of phenotype and genotype.

The objective of this chapter was to comparatively assess the genome-wide LD and haplotype block structure in one Canadian Holstein population and one hybrid beef cattle population using about 40 K SNPs with an average intermarker distance of ~ 65 kb. First, breed specific block structure is characterized in terms of the genome coverage of blocks, average block size, and block haplotype diversity. Second, block boundary comparisons are performed and correlation between block size and average heterozygosity within blocks is examined.

3.2. Materials and Methods

3.2.1. DNA Collection

Semen samples from 647 proven Canadian Holstein bulls born in North America between 1985 and 2002 were obtained from the L'Alliance Boviteq Inc.

(Semex Alliance, Canada) for DNA extraction. The pedigree of the 647 bulls consisted of 71 sires and 492 dams with an average paternal family size of 9.1. Bulls were selected so that both family and population-wide information can be used.

Blood samples from 922 hybrid beef steers were obtained for DNA extraction. These animals were managed and tested for growth and feed efficiency under feedlot conditions at the University of Alberta's Kinsella beef cattle research station from 2003 to 2008. They were produced from a cross between Angus, Charolais, or University of Alberta hybrid bulls and the University of Alberta's experimental hybrid dam line. The hybrid bulls and the hybrid dam line were produced from crosses among three hybrid cattle lines, namely Beef Synthetic 1 (BS1), Beef Synthetic 2 (BS2), and Dairy × Beef Synthetic (DBS). BS1 was composed of approximately 33% Angus and Charolais, about 20% Galloway, and the remainder from other beef breeds. The BS2 hybrid was made up of approximately 60% Hereford and 40% other beef breeds. The DBS was composed of approximately 60% dairy breeds (Holstein, Brown Swiss, or Simmental) and approximately 40% beef breeds, mainly Angus and Charolais (GOONEWARDENE *et al.* 2003).

3.2.2. Genotyping Platform and Marker Selection

High-throughput genotyping was carried out for both populations using the Illumina BovineSNP50 BeadChip. Missing values were imputed using fastPHASE (SCHEET and STEPHENS 2006). Chromosomal coordinates for SNPs

were obtained by aligning approximately 250 bp flanking each SNP by BLAST to the bovine genome sequence assembly Btau4.0. Unmapped SNPs on Btau_4.0 were excluded. SNPs located on chromosome X were also eliminated, as they are hemizygous in bulls and steers. This resulted in 49,312 and 49,128 SNPs on 29 autosomes for Canadian Holstein and hybrid beef cattle, respectively. SNPs with minor allele frequency (MAF) < 0.05 or exhibiting deviations from Hardy-Weinberg equilibrium (HWE) with $P < 0.001$ were also excluded from the analysis. Finally, 37,986 and 40,472 SNPs were considered for pairwise LD and haplotype block analysis for the Canadian Holstein and hybrid beef cattle, respectively. Among the total SNPs, there were 34,846 markers shared by both cattle populations.

3.2.3. LD Analysis

Whole-chromosome haplotypes were constructed from the unphased SNP genotype data using fastPHASE (SCHEET and STEPHENS 2006). LD was assessed by generating D' and r^2 values among all pairs of syntenic markers using Haploview (BARRETT *et al.* 2005). Tag SNPs were selected based on the pairwise tagging method of the Tagger program (DE BAKKER *et al.* 2005) implemented in Haploview, using a threshold of $r^2 \geq 0.8$.

3.2.4. Haplotype Block Analysis

Haplotype blocks were identified for all autosomes using the Haploview software (BARRETT *et al.* 2005). Haplotype blocks are chromosome regions that

are inherited from the ancestors of the current population without substantial recombination. They were identified based on the estimates of the normalized measure of allelic association, *D'*, for all pair-wise combinations of SNPs within each chromosome (GABRIEL *et al.* 2002). Confidence bounds of *D'* values were used instead of the point estimates of *D'*, since the latter can fluctuate upward with a small number of samples or with rare alleles. A pair of markers was defined to be in strong LD if the confidence interval minima upper bound on *D'* was 0.98 and lower bound was 0.7. Other default values used for block identification were: (1) the upper confidence interval maximum for strong historical recombination was set to 0.9; (2) markers with MAF below 0.05 were excluded; (3) the fraction of strong LD in informative comparisons was at least 0.95; (4) the 4^{th} gamete was observed at frequency > 0.01; (5) the strong LD spine was extended if $D' > 0.8$. Finally, a haplotype block was identified as a region over which 95% of marker pairs showed strong LD (GABRIEL *et al.* 2002). Genome-wide haplotype block maps were then built, and the block size, haplotype diversity within blocks, and genome coverage were characterized for the Canadian Holstein and hybrid beef populations. Finally, the relationship between the block size and average heterozygosity of SNPs within block was analyzed.

3.3. Results

The distribution of the intermarker distances of the SNPs used in this study is shown in Figure 3 – 1. The total number of SNPs on each chromosome varied from 2,454 on BTA1 to 716 on BTA28 for Canadian Holstein cattle and

from 2,616 on BTA1 to 747 on BTA28 for hybrid beef cattle. The average interval between SNPs, average heterozygosity and average MAF were 66.86 ± 64.57 kb, 0.376 ± 0.122, 0.281 ± 0.129 for Canadian Holstein cattle and 62.79 ± 59.66 kb, 0.375 ± 0.121, 0.281 ± 0.129 for hybrid beef cattle (Table 3 – 1).

3.3.1. General LD and Tag SNPs

LD was assessed using D' and r^2 among all pairs of syntenic markers for both the Canadian Holstein and hybrid beef cattle populations based on the inferred linkage phase of SNPs. Pairwise LD displayed a negative exponential distribution in relation to pairwise marker physical distance. The Canadian Holstein cattle, on average, displayed stronger LD than the hybrid beef cattle across all chromosomes, particularly for closely spaced markers. For instance, the average r^2 at 0.1 Mb distance was 0.21 in Canadian Holstein and 0.14 in hybrid beef; while the average D' at 0.1 Mb distance was 0.78 in Canadian Holstein and 0.66 in hybrid beef on BTA14 (Figure 3 – 2). The longer-range LDs were comparable between the two cattle populations. The percentage of identified tag SNPs was 85.36% for the Canadian Holstein cattle and 92.48% for the hybrid beef cattle (Table 3 – 2 and Table 3 – 3).

3.3.2. Genome-wide Comparison of Haplotype Blocks

Haplotype blocks were identified for all 29 autosomes. This resulted in a total of 1,716 and 950 haplotype blocks consisting of two or more SNPs for the Canadian Holstein and hybrid beef cattle, respectively. The number of haplotype

blocks was generally proportional to the chromosome length, gradually decreasing from BTA1 to BTA 29 (Figure 3 – 3). These blocks comprised a total of 8,249 and 4,140 SNPs and covered 366.78 Mb (14.41%) and 146.78 Mb (5.77%) of the whole bovine autosomal sequence map for the Canadian Holstein and hybrid beef, respectively. The locations of haplotype blocks in the bovine genome are presented in Figure 3 – 4 and Figure 3 – 5. Detailed information on the number and size of haplotype blocks for each chromosome is summarized in Table 3 – 2 and Table 3 – 3.

3.3.3. Comparative Diversities of Haplotype Blocks

The average haplotype block size was 213.74 ± 153.03 kb in Canadian Holstein and 154.50 ± 110.62 kb in hybrid beef. However, the size of each block varied dramatically, from ~1 kb to 1,937 kb (BTA29) in the Canadian Holstein cattle and from ~1 kb to 1,263 kb (BTA7) in the hybrid beef cattle. Most blocks were in 100 – 400 kb in Canadian Holstein and 50 – 300 kb in hybrid beef (Figure 3 – 6), while the genomic sequences spanned by haplotype blocks were in 200 – 500 kb in Canadian Holstein and 100 – 400 kb in hybrid beef (Figure 3 – 7). The Kolmogorov-Smirnov (K-S) test indicates that a significant difference (*P*-value < 2.2×10^{-16}) exists between the distributions of block size in the two cattle populations (Figure 3 – 8).

In terms of marker number, most of the blocks were composed of 4 to 5 SNPs in both cattle populations (Figure 3 – 9 and Table 3 – 4). The average block size gradually increased with the addition of markers within a block (Figure 3 –

10). The maximum number of SNPs in a block was 20 with a block length of 1,172 kb on BTA7 in Canadian Holstein, while the maximum number of SNPs in a block was 17 with a block length of 694 kb on BTA5 in hybrid beef.

The average number of haplotypes in each haplotype block category (classified in terms of marker number in a block) was similar between two cattle populations (Table 3 – 4). In general, there was a slight increase for the mean number of haplotypes with the increase of markers in a block (Figure 3 – 11). For example, the average number of haplotypes was 4.34, 5.13, 5.62, 6.01, and 6.10 for blocks consisting of 4, 5, 6, 7, and 8 SNPs, respectively, in Holstein cattle. The observed number of haplotypes existing within a block was far less than the expected number of haplotypes in a pure random combination case (Figure 3 – 12).

3.3.4. Regional Comparison of Haplotype Blocks

Different haplotype block structures were observed in genomic regions that contain genes associated with economically important traits in cattle. Haplotype blocks close to *DGAT1* on BTA14 and *PPARGC1A* on BTA6 were only observed in the Holstein cattle (Figure 3 – 13 and Figure 3 – 14); *DGAT1* (GRISART *et al.* 2002; RIQUET *et al.* 1999) and *PPARGC1A* (ESTALL *et al.* 2009; WEIKARD *et al.* 2005) show strong associations with milk production traits in dairy cattle. One haplotype block close to *LEP* on BTA4 was only identified in the hybrid beef cattle (Figure 3 – 15); *LEP* controls food intake, energy balance and body composition in mammals (GEARY *et al.* 2003; HOUSEKNECHT *et al.*

1998) and there is a significant correlation between serum leptin and carcass quality (backfat thickness, marbling and rib fat percentage) in beef cattle (BUCHANAN *et al.* 2002; GEARY *et al.* 2003; KONONOFF *et al.* 2005).

Although different block distributions and block boundary discordances existed between the Holstein and hybrid beef cattle, there were haplotype blocks shared (2,177 SNPs within haplotype blocks) by the two cattle populations (Figure 3 – 16). For example, one haplotype block close to *GH1* on BTA19 was identified in both cattle populations (Figure 3 – 17), and *GH1* was found to be associated with growth and carcass composition in beef cattle (SCHLEE *et al.* 1994a; SCHLEE *et al.* 1994b; TAYLOR *et al.* 1998) and milk production in dairy cattle (LAGZIEL *et al.* 1996).

3.3.5. Relationship between Haplotype Block Size and the Average within-Block Heterozygosity

The mean block heterozygosity, measured by averaging the heterozygosities of each marker involved in a block, was 0.357 ± 0.072 (0.154 – 0.552) in Holstein and 0.351 ± 0.081 (0.117 – 0.529) in hybrid beef. A significant negative correlation ($P < 0.05$) between haplotype block size and average within-block heterozygosity was found on BTA5, 7, 13, 14, and 23 in Holstein and on BTA3, 5, 7, 13, 22, 24, and 27 in hybrid beef (Table 3 – 5). One example of the relationship between haplotype block size and the average heterozygosity on BTA5 is shown in Figure 3 – 18.

3.4. Discussion

In this work we present a comparative assessment of genome-wide LD and haplotype block structure in Holstein and hybrid beef cattle populations using about 40,000 SNPs. The two cattle populations used in this study are the result of different breeding processes. The Holstein population has been under strong directional selection for superior milking ability and conformation, whereas the hybrid beef population has been subjected to continuous crossbreeding with moderate selection for performance and growth rate.

The results of the LD analysis in Holstein cattle are in agreement with those of previous studies, in that significant LD in Holstein extended to 40 – 60 kb (BOHMANOVA *et al.* 2010; KHATKAR *et al.* 2008). Much less extensive LD exists in the hybrid beef cattle in comparison with the previously reported LD in Japanese Black and Japanese Brown beef cattle, in that significant LD was observed for most syntenic marker pairs < 40 cM apart using only 246 autosomal microsatellite markers (ODANI *et al.* 2006). Given that a previous study suggests that ~30,000 uniformly distributed SNPs should be used to construct a complete LD map (VILLA-ANGULO *et al.* 2009), our study using over 40,000 SNPs provides a better estimation of LD in beef cattle. However, LD in the hybrid beef cattle population cannot fully represent LD in other pure breed beef populations since a recently admixed population could have larger LD compared to the purebred beef cattle, especially if the two parental populations have large differences in allele frequencies (GREENWOOD *et al.* 2004). On the other hand, the larger LD in an

admixed population could elevate the type I error rate for detecting genes underlying complex traits from LD based association studies (DENG *et al.* 2001). The intensive directional selection for high milk production ability through artificial insemination and progeny testing since 1950 could be one possible reason for the stronger LD in the Holstein cattle in contrast to that in the hybrid beef cattle (BROTHERSTONE and GODDARD 2005). Denser SNPs are necessary for the hybrid beef cattle with the purpose of association studies than the Canadian Holstein cattle.

Marker density affects the results of haplotype block analyses. The average block size was 69.7 kb and the genome coverage of blocks was about 2.18% in Holstein-Friesian Cattle, based on markers with median spacing of 93.9 kb (9,195 SNPs) (KHATKAR *et al.* 2007); whereas the average block size was 10.3 kb and the genome coverage of blocks was about 34.7% in 19 worldwide taurine and indicine breeds using markers with an average intermarker distance of 4 kb (VILLA-ANGULO *et al.* 2009). The smaller blocks identified in the latter study may be the result of the increased marker density which may uncover more ancestral recombination events, break up larger blocks and refine block boundaries (PHILLIPS *et al.* 2003). However, the haplotype blocks exhibited an overall mean size of 213.7 kb in Holstein and 154.5 kb in hybrid beef with median marker spacing of 45.26 kb, which is about 3 times larger than that reported by Khatkar *et al.* (2007) and about 15 – 20 times larger than that reported by Villa-Angulo *et al.* (2009). The block size identified in this study was larger than previous studies with either higher or lower marker densities (KHATKAR *et al.* 2007; VILLA-

ANGULO *et al.* 2009), indicating that further studies on assessment of the effect of marker density on haplotype block partition should be further investigated. In addition, dense marker sets may be necessary for further stable views of fine-scale haplotype block patterns in the bovine genome (KHATKAR *et al.* 2007; VILLA-ANGULO *et al.* 2009).

Our study on genome-wide comparison of haplotype blocks revealed different block structure between dairy and beef cattle in terms of number, size, distribution and genome coverage of blocks. However, non-differentiable haplotype block structure between dairy and beef breeds was reported in a previous study based on comparison of haplotype block structure in 101 genomic regions on BTA6, 14, and 25 spanning up to ~ 7.6 Mb (VILLA-ANGULO *et al.* 2009). This disagreement could be drawn from the different density of markers and number of animals used, different methods for characterizing the haplotype blocks, and different measures used to quantify block similarities and block boundary consistency between populations. The correlation of the numbers of haplotype blocks in a genomic region, the approach used by Villa-Angulo *et al.* (2009), may not be satisfactory for comparing block structures across breeds. Instead our study provided more detailed comparisons for haplotype blocks between the dairy and beef cattle populations, where the location, marker composition, haplotype diversity, boundary overlap, and common blocks were fully identified across the 29 bovine autosomes. In addition, our study on whole genome comparison of haplotype blocks may provide a better view on the block

pattern between dairy and beef cattle in contrast to the limited comparisons within a few mega-base pairs.

Haplotype blocks are shaped by multiple evolutionary factors and the relative contribution of each factor on forming specific haplotype pattern can differ not only among populations, but also throughout the genome (GOLDSTEIN and WEALE 2001). Haplotype blocks could be created by selection of beneficial or deleterious alleles together with physically linked loci (selective sweeps or background selection) (PHILLIPS *et al.* 2003). Previous studies have found that selection governs the conservation of haplotype blocks in orthologous genomic regions in different species, especially in genic regions (GURYEV *et al.* 2006), and gene history is quantitatively the main force for local patterns of genome sequence variation (REICH *et al.* 2002). The observed different block structures at functional genomic regions between two populations may in this study due to the distinct selection of genes associated with different economically important traits in cattle breeds. Regions of similar block structure could be the result of similar selection processes acting on genomic regions with genes affecting multiple traits (GURYEV *et al.* 2006). These results indicate that comparative analysis of local haplotype block structure could be used in the future searches for functional candidate genes. In addition, the significant negative correlation between block sizes and the average marker heterozygosity within blocks on some chromosomes probably further relates to the effect of directional artificial selection on reducing the genetic diversity through selective sweeps or background selection. Furthermore, the strong selection intensity experienced in the Holstein cattle

could be one possible reason for the observed larger block size in comparison with block size in hybrid beef. Since selection is localized around specific genes and is not averaged throughout the genome, it can be the major cause of local haplotype block structure in functional genomic regions.

The population history is another reason affecting haplotype block size (REICH *et al.* 2002). The haplotype block pattern in the hybrid beef cattle could be partially created through crossbreeding (gene flow), especially if subpopulations with haplotypes occur in different frequencies (GODDARD 1991; GREENWOOD *et al.* 2004). In addition, finite population size can be an alternative mechanism for block structure. It was found that humans, with a relatively heterogeneous founder population, had small blocks with a median size of 45 kb, and the inbred populations of laboratory mice, which experienced a recent genetic bottleneck during domestication, had large blocks spanning hundreds of kilo-base pairs (GURYEV *et al.* 2006). In cattle, the effective population size was over 50,000 at 10,000 generations ago and decreased to ~100 over the last 50 generations by breed formation and artificial breeding techniques (DE ROOS *et al.* 2008; MACEACHERN *et al.* 2009), which could result in loss of some haplotypes and can create haplotype blocks (TERWILLIGER *et al.* 1998). Furthermore, genetic drift alone can lead to block-like patterns of LD (ZHANG *et al.* 2003). The Holstein population with long-term inbreeding (genetic drift) could increase the haplotype sharing, which provides another possibility for the observed larger haplotype blocks. Finally, mutation was found to jointly dictate haplotype block characteristics with population demographic history and recombination (WANG *et*

al. 2002), in which striking negative correlation exist between block structure, in terms of length and genome coverage of blocks, and recombination rate (GREENWOOD *et al.* 2004).

Overall, the genomic patterns of haplotype blocks in different cattle populations can be the results from selection together with other demographic factors. The structure of blocks throughout the genome may reflect the population history and breeding system, while blocks in functional genomic regions may mainly reflect the history of selection (SLATKIN 2008). Assessing the effect of a specific factor on the haplotype block structure will require further in depth studies in other cattle populations with well-known demographic and breeding history.

3.5. Conclusions

Genome-wide LD, haplotype block partitioning and haplotype diversity in one Canadian Holstein and one hybrid beef population were characterized and compared using about 40,000 SNPs. Consistent with previous analyses in cattle, larger LD was observed in the Holstein cattle, such that r^2 averaged ~0.21 at 100 kb in the Holstein and ~0.14 in the hybrid beef. Haplotype blocks exhibited an overall mean size of 213.7 kb and 154.5 kb with an average of 4.8 and 4.6 haplotypes per block in the Holstein and the hybrid beef, respectively. Denser and larger haplotype blocks were identified for the Holstein, whereas limited haplotype diversities were existed for both populations. Analyses of genome-wide block pattern exhibited a clear differentiation between the Holstein and the hybrid

beef cattle. Regional comparison of block structures, as well as the negative correlation between the block size and average within-block marker heterozygosity, revealed that distinct selection in cattle breeds may have a role in shaping the block pattern in bovine genome.

Comparisons of LD and haplotype blocks may help us to understand the breeding history and to pinpoint functionally important genomic segments and genes showing significant evidence of positive selection in cattle breeds. Our study provides the first whole-genome comparison of the location, marker composition, haplotype diversity and boundary overlap for haplotype blocks between the dairy and beef cattle populations. In addition, this study also provides the first high-density reference LD map and haplotype block map for beef cattle, which could be used in designing and interpreting association studies.

3.6. Literature Cited

BARRETT, J. C., B. FRY, J. MALLER and M. J. DALY, 2005 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics **21:** 263-265.

BOHMANOVA, J., M. SARGOLZAEI and F. S. SCHENKEL, 2010 Characteristics of linkage disequilibrium in North American Holsteins. BMC Genomics **11**.

BROTHERSTONE, S., and M. GODDARD, 2005 Artificial selection and maintenance of genetic variance in the global dairy cow population. Philosophical Transactions of the Royal Society B-Biological Sciences **360:** 1479-1488.

BUCHANAN, F. C., C. J. FITZSIMMONS, A. G. VAN KESSEL, T. D. THUE, D. C. WINKELMAN-SIM *et al.*, 2002 Association of a missense mutation in the

bovine leptin gene with carcass fat content and leptin mRNA levels. Genetics Selection Evolution **34:** 105-116.

DE BAKKER, P. I. W., R. YELENSKY, I. PE'ER, S. B. GABRIEL, M. J. DALY *et al.*, 2005 Efficiency and power in genetic association studies. Nature Genetics **37:** 1217-1223.

DE ROOS, A. P. W., B. J. HAYES, R. J. SPELMAN and M. E. GODDARD, 2008 Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics **179:** 1503-1512.

DENG, H. W., W. M. CHEN and R. R. RECKER, 2001 Population admixture: detection by Hardy-Weinberg test and its quantitative effects on linkage-disequilibrium methods for localizing genes underlying complex traits. Genetics **157:** 885-897.

ESTALL, J. L., M. KAHN, M. P. COOPER, F. M. FISHER, M. K. WU *et al.*, 2009 Sensitivity of Lipid Metabolism and Insulin Signaling to Genetic Alterations in Hepatic Peroxisome Proliferator-Activated Receptor-gamma Coactivator-1 alpha Expression. Diabetes **58:** 1499-1508.

GABRIEL, S. B., S. F. SCHAFFNER, H. NGUYEN, J. M. MOORE, J. ROY *et al.*, 2002 The structure of haplotype blocks in the human genome. Science **296:** 2225-2229.

GAUTIER, M., T. FARAUT, K. MOAZAMI-GOUDARZI, V. NAVRATIL, M. FOGFIO *et al.*, 2007 Genetic and haplotypic structure in 14 European and African cattle breeds. Genetics **177:** 1059-1070.

GEARY, T. W., E. L. MCFADIN, M. D. MACNEIL, E. E. GRINGS, R. E. SHORT *et al.*, 2003 Leptin as a predictor of carcass composition in beef cattle. Journal of Animal Science **81:** 1-8.

GIBBS, R. A., J. F. TAYLOR, C. P. VAN TASSELL, W. BARENDSE, K. A. EVERSOIE *et al.*, 2009 Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. Science **324:** 528-532.

GODDARD, M. E., 1991 Mapping Genes for Quantitative Traits Using Linkage Disequilibrium. Genetics Selection Evolution **23:** S131-S134.

GOLDSTEIN, D. B., and M. E. WEALE, 2001 Population genomics: Linkage disequilibrium holds the key. Current Biology **11:** R576-R579.

GOONEWARDENE, L. A., Z. WANG, M. A. PRICE, R. C. YANG, R. T. BERG *et al.*, 2003 Effect of udder type and calving assistance on weaning traits of beef and dairy x beef calves. Livestock Production Science **81:** 47-56.

GREENWOOD, T. A., B. K. RANA and N. J. SCHORK, 2004 Human haplotype block sizes are negatively correlated with recombination rates. Genome Research **14:** 1358-1361.

GRISART, B., W. COPPIETERS, F. FARNIR, L. KARIM, C. FORD *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research **12:** 222-231.

GURYEV, V., B. M. G. SMITS, J. VAN DE BELT, M. VERHEUL, N. HUBNER *et al.*, 2006 Haplotype block structure is conserved across mammals. PLoS Genetics **2:** 1111-1118.

HOUSEKNECHT, K. L., C. A. BAILE, R. L. MATTERI and M. E. SPURLOCK, 1998 The biology of leptin: A review. Journal of Animal Science **76:** 1405-1420.

KHATKAR, M. S., F. W. NICHOLAS, A. R. COLLINS, K. R. ZENGER, J. AL CAVANAGH *et al.*, 2008 Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics **9**.

KHATKAR, M. S., K. R. ZENGER, M. HOBBS, R. J. HAWKEN, J. A. CAVANAGH *et al.*, 2007 A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in holstein-friesian cattle. Genetics **176:** 763-772.

KONONOFF, P. J., H. M. DEOBALD, E. L. STEWART, A. D. LAYCOCK and F. L. S. MARQUESS, 2005 The effect of a leptin single nucleotide polymorphism on quality grade, yield grade, and carcass weight of beef cattle. Journal of Animal Science **83:** 927-932.

LAGZIEL, A., E. LIPKIN and M. SOLLER, 1996 Association between SSCP haplotypes at the bovine growth hormone gene and milk protein percentage. Genetics **142:** 945-951.

MACEACHERN, S., B. HAYES, J. MCEWAN and M. GODDARD, 2009 An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (Bos taurus) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle. BMC Genomics **10**.

MCKAY, S. D., R. D. SCHNABEL, B. M. MURDOCH, L. K. MATUKUMALLI, J. AERTS *et al.*, 2007 Whole genome linkage disequilibrium maps in cattle. BMC Genetics **8**.

ODANI, M., A. NARITA, T. WATANABE, K. YOKOUCHI, Y. SUGIMOTO *et al.*, 2006 Genome-wide linkage disequilibrium in two Japanese beef cattle breeds. Animal Genetics **37:** 139-144.

PATIL, N., A. J. BERNO, D. A. HINDS, W. A. BARRETT, J. M. DOSHI *et al.*, 2001 Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science **294:** 1719-1723.

PHILLIPS, M. S., R. LAWRENCE, R. SACHIDANANDAM, A. P. MORRIS, D. J. BALDING *et al.*, 2003 Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. Nature Genetics **33:** 382-387.

REICH, D. E., S. F. SCHAFFNER, M. J. DALY, G. MCVEAN, J. C. MULLIKIN *et al.*, 2002 Human genome sequence variation and the influence of gene history, mutation and recombination. Nature Genetics **32:** 135-142.

RIQUET, J., W. COPPIETERS, N. CAMBISANO, J. J. ARRANZ, P. BERZI *et al.*, 1999 Fine-mapping of quantitative trait loci by identity by descent in outbred populations: Application to milk production in dairy cattle. Proceedings of the National Academy of Sciences of the United States of America **96:** 9252-9257.

SCHEET, P., and M. STEPHENS, 2006 A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing

genotypes and haplotypic phase. American Journal of Human Genetics **78:** 629-644.

SCHLEE, P., R. GRAML, O. ROTTMANN and F. PIRCHNER, 1994a Influence of Growth-Hormone Genotypes on Breeding Values of Simmental Bulls. Journal of Animal Breeding and Genetics **111:** 253-256.

SCHLEE, P., R. GRAML, E. SCHALLENBERGER, D. SCHAMS, O. ROTTMANN *et al.*, 1994b Growth-Hormone and Insulin-Like Growth-Factor-I Concentrations in Bulls of Various Growth-Hormone Genotypes. Theoretical and Applied Genetics **88:** 497-500.

SLATKIN, M., 2008 Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics **9:** 477-485.

TAYLOR, J. F., L. L. COUTINHO, K. L. HERRING, D. S. GALLAGHER, R. A. BRENNEMAN *et al.*, 1998 Candidate gene analysis of GH1 for effects on growth and carcass composition of cattle. Animal Genetics **29:** 194-201.

TERWILLIGER, J. D., S. ZOLLNER, M. LAAN and S. PAABO, 1998 Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'Drift mapping' in small populations with no demographic expansion. Human Heredity **48:** 138-154.

VILLA-ANGULO, R., L. K. MATUKUMALLI, C. A. GILL, J. CHOI, C. P. VAN TASSELL *et al.*, 2009 High-resolution haplotype block structure in the cattle genome. BMC Genetics **10**.

WANG, N., J. M. AKEY, K. ZHANG, R. CHAKRABORTY and L. JIN, 2002 Distribution of recombination crossovers and the origin of haplotype

blocks: The interplay of population history, recombination, and mutation. American Journal of Human Genetics **71:** 1227-1234.

WEIKARD, R., C. KUHN, T. GOLDAMMER, G. FREYER and M. SCHWERIN, 2005 The bovine PPARGC1A gene: molecular characterization and association of an SNP with variation of milk fat synthesis. Physiological Genomics **21:** 1-13.

ZHANG, K., J. M. AKEY, N. WANG, M. XIONG, R. CHAKRABORTY *et al.*, 2003 Randomly distributed crossovers may generate block-like pattern's of linkage disequilibrium: an act of genetic drift. Human Genetics **113:** 51-59.

Tables

Table 3 – 1. Number of SNPs, average intermarker spacing (kb), MAF and average heterozygosity for 29 BTA in the Canadian Holstein and hybrid beef cattle populations. The BTA length (Mb) is also presented.

BTA	BTA Length (Mb)	Canadian Holstein cattle				Hybrid beef cattle			
		N	Spacing (kb) ± SD	MAF ± SD	He ± SD	N	Spacing (kb) ± SD	MAF ± SD	He ± SD
1	161.11	2454	65.6±59.1	0.279±0.128	0.374±0.121	2616	61.53±52.64	0.276±0.13	0.368±0.123
2	140.80	2021	69.6±70.31	0.282±0.129	0.376±0.122	2179	64.5±64.84	0.283±0.131	0.377±0.122
3	127.92	1912	66.52±66.63	0.286±0.129	0.381±0.122	2055	62.27±62.05	0.275±0.128	0.372±0.121
4	124.45	1840	67.43±61.07	0.28±0.13	0.375±0.123	2008	61.81±50.75	0.278±0.128	0.373±0.121
5	125.85	1599	78.72±83.66	0.275±0.13	0.369±0.126	1684	74.72±81.9	0.281±0.131	0.374±0.122
6	122.56	1899	64.51±62.54	0.276±0.129	0.374±0.125	2037	60.14±57.16	0.282±0.13	0.377±0.122
7	112.08	1610	69.39±68.63	0.28±0.13	0.376±0.124	1754	63.72±61.98	0.281±0.129	0.376±0.121
8	116.94	1743	67.12±60.12	0.277±0.127	0.375±0.12	1911	61.22±53.07	0.274±0.129	0.369±0.122
9	108.15	1496	72.27±70.49	0.271±0.129	0.369±0.125	1643	65.8±62.24	0.275±0.133	0.368±0.125
10	106.38	1595	66.28±68.48	0.271±0.128	0.367±0.124	1735	60.93±71.78	0.277±0.128	0.372±0.121
11	110.17	1688	65.2±61.33	0.278±0.127	0.371±0.118	1756	62.68±58.01	0.281±0.13	0.374±0.121
12	85.36	1223	69.69±72.58	0.279±0.13	0.374±0.124	1304	65.39±66.03	0.278±0.131	0.371±0.123
13	84.42	1296	64.72±56.61	0.287±0.129	0.382±0.122	1415	59.46±50.8	0.282±0.131	0.376±0.124
14	81.35	1283	63.4±54.67	0.298±0.125	0.397±0.116	1356	59.98±49.99	0.29±0.124	0.386±0.114
15	84.63	1278	66±59.99	0.283±0.128	0.382±0.123	1318	64.18±59.22	0.279±0.128	0.373±0.119
16	77.91	1145	67.95±72.52	0.28±0.128	0.379±0.124	1234	63.16±64.8	0.28±0.13	0.373±0.12
17	76.51	1183	64.64±60.27	0.279±0.129	0.373±0.122	1256	60.84±53.3	0.274±0.129	0.367±0.121
18	66.14	1027	64.36±65.68	0.291±0.126	0.386±0.119	1055	62.55±60.5	0.291±0.128	0.384±0.118
19	65.31	1041	62.56±53.01	0.285±0.129	0.377±0.119	1053	61.82±52.43	0.295±0.124	0.387±0.113
20	75.80	1158	65.17±63.82	0.27±0.131	0.364±0.123	1227	61.58±55.62	0.281±0.132	0.374±0.125
21	69.17	1020	67.85±66.32	0.284±0.129	0.378±0.12	1057	65.48±64.02	0.281±0.128	0.377±0.12
22	61.85	944	65.35±53.1	0.276±0.132	0.37±0.126	1001	61.62±46.95	0.275±0.128	0.369±0.121
23	53.38	811	65.63±58.67	0.291±0.129	0.386±0.122	860	62.02±54.64	0.292±0.126	0.387±0.118
24	65.02	925	70.17±62	0.281±0.129	0.374±0.121	1011	64.29±56.11	0.289±0.131	0.382±0.12
25	44.06	766	56.76±44.58	0.292±0.13	0.386±0.12	783	55.53±44.34	0.302±0.128	0.392±0.115
26	51.75	797	64.01±49.73	0.275±0.131	0.371±0.125	834	61.23±47.31	0.269±0.127	0.367±0.123
27	48.75	730	66.81±87.54	0.279±0.131	0.372±0.122	767	63.58±85.98	0.282±0.131	0.374±0.12
28	46.08	716	64.27±52.45	0.288±0.128	0.384±0.12	747	61.63±50.24	0.285±0.128	0.379±0.119
29	52.00	786	65.96±74.45	0.287±0.129	0.381±0.121	816	63.44±61.55	0.283±0.129	0.375±0.121
Overall	2545.90	37986	66.86±64.57	0.281±0.129	0.376±0.122	40472	62.79±59.66	0.281±0.129	0.375±0.121

Table 3 – 2. Number of haplotype blocks, average block size, chromosome coverage of blocks (in kb and %) and number of SNPs in blocks for 29 BTA in the Canadian Holstein cattle. The number and percentage of tag SNPs are also presented.

BTA	N-SNP	N-blocks	Total block length (kb)	Block mean size ± SD (kb)	Block size (min)	Block size (max)	BTA Length (Mb)	% coverage	No. of SNPs in blocks	Tag SNPs (pairwise)	% pairwise tagSNP
1	2454	129	28079.14	217.67±127.92	0.131	737.952	161.11	17.43	646	2008	81.83
2	2021	102	21162.74	207.48±136.07	1.641	805.229	140.80	15.03	494	1691	83.67
3	1912	96	21087.41	219.66±152.49	0.108	908.567	127.92	16.48	461	1624	84.94
4	1840	82	16495.48	201.16±110.64	11.574	510.578	124.45	13.25	387	1585	86.14
5	1599	71	19074.34	268.65±167.16	0.148	1031.62	125.85	15.16	356	1348	84.30
6	1899	107	21560.48	201.5±163.06	1.36	1224.84	122.56	17.59	523	1592	83.83
7	1610	82	19243.5	234.68±194.93	0.003	1172.31	112.08	17.17	426	1299	80.68
8	1743	89	21954.75	246.68±163.68	3.416	1006.62	116.94	18.77	473	1431	82.10
9	1496	60	13547.24	225.79±129.83	0.449	611.368	108.15	12.53	297	1309	87.50
10	1595	79	16683.51	211.18±131.01	21.467	789.47	106.38	15.68	396	1310	82.13
11	1688	71	13396.11	188.68±146.71	5.467	820.199	110.17	12.16	322	1468	86.97
12	1223	54	10827.46	200.51±150.03	0.237	862.154	85.36	12.68	249	1079	88.23
13	1296	71	15905.5	224.02±198.78	0.382	1244.62	84.42	18.84	330	1075	82.95
14	1283	63	15715.46	249.45±135.89	0.211	600.714	81.35	19.32	320	1023	79.73
15	1278	52	9659.268	185.76±100.81	14.413	443.014	84.63	11.41	239	1134	88.73
16	1145	75	17340.05	231.2±196.1	3.727	1107.42	77.91	22.26	369	926	80.87
17	1183	46	9332.599	202.88±119.93	0.333	577.403	76.51	12.20	219	1034	87.40
18	1027	33	6220.297	188.49±116.77	10.468	538.168	66.14	9.40	153	930	90.56
19	1041	47	9321.983	198.34±99.84	3.937	487.839	65.31	14.27	221	922	88.57
20	1158	51	11620.2	227.85±165.4	15.752	1045.49	75.80	15.33	239	954	82.38
21	1020	38	7063.58	185.88±89.8	13.428	490.568	69.17	10.21	176	895	87.75
22	944	41	7673.629	187.16±128.89	6.394	712.044	61.85	12.41	183	826	87.50
23	811	23	4003.912	174.08±121.43	1.568	473.552	53.38	7.50	102	724	89.27
24	925	34	6077.87	178.76±128.98	0.095	597.6	65.02	9.35	144	811	87.68
25	766	27	4019.837	148.88±99.48	7.69	367.33	44.06	9.12	110	697	90.99
26	797	29	6430.431	221.74±159.85	0.281	842.903	51.75	12.43	133	695	87.20
27	730	18	4271.452	237.3±154.51	0.151	542.63	48.75	8.76	86	670	91.78
28	716	16	2892.055	180.75±86.7	87.22	405.2	46.08	6.28	66	666	93.02
29	786	30	6121.304	204.04±338.01	1.86	1936.68	52.00	11.77	129	700	89.06
All	37986	1716	366781.6	213.74±153.03	0.003	1936.68	2545.90	14.41	8249	32426	85.36

Table 3 – 3. Number of haplotype blocks, average block size, chromosome coverage of blocks (in kb and %) and number of SNPs in blocks for 29 BTA in the hybrid beef cattle. The number and percentage of tag SNPs are also presented.

BTA	N-SNP	N-blocks	Total block length (kb)	Block mean size ± SD (kb)	Block size (min)	Block size (max)	BTA Length (Mb)	% coverage	No. of SNPs in blocks	Tag SNPs (pairwise)	% pairwise tagSNP
1	2616	79	11074.03	140.18 ±83.87	0.131	515.309	161.11	6.87	340	2378	90.90
2	2179	48	7159.122	149.15 ±98.13	1.641	517.084	140.80	5.08	207	1991	91.37
3	2055	58	8901.178	153.47 ±107.25	0.108	494.486	127.92	6.96	250	1866	90.80
4	2008	68	11600.96	170.6 ±109.83	11.574	708.195	124.45	9.32	314	1819	90.59
5	1684	39	6876.229	176.31 ±130.06	2.444	694.634	125.85	5.46	179	1560	92.64
6	2037	49	7257.434	148.11 ±104.51	2.66	465.823	122.56	5.92	216	1863	91.46
7	1754	45	8057.654	179.06 ±200.97	0.003	1263.42	112.08	7.19	200	1587	90.48
8	1911	62	10364.24	167.17 ±79.57	28.215	380.395	116.94	8.86	301	1736	90.84
9	1643	42	7443.768	177.23 ±117.78	0.449	602.077	108.15	6.88	197	1526	92.88
10	1735	40	6562.707	164.07 ±120.53	0.284	656.932	106.38	6.17	181	1588	91.53
11	1756	41	6307.962	153.85 ±83.48	0.886	398.343	110.17	5.73	188	1598	91.00
12	1304	27	4270.301	158.16 ±111.04	0.237	427.384	85.36	5.00	118	1221	93.63
13	1415	39	5296.792	135.82 ±92.41	0.382	399.273	84.42	6.27	158	1322	93.43
14	1356	28	4823.41	172.26 ±107.44	0.211	411.231	81.35	5.93	126	1246	91.89
15	1318	21	3025.87	144.09 ±83.18	3.701	336.039	84.63	3.58	88	1230	93.32
16	1234	32	4084.634	127.64 ±88.29	3.727	383.237	77.91	5.24	129	1131	91.65
17	1256	31	4278.714	138.02 ±76.79	7.015	292.901	76.51	5.59	126	1175	93.55
18	1055	26	4850.358	186.55 ±125.6	10.468	449.016	66.14	7.33	116	986	93.46
19	1053	22	2991.782	135.99 ±88.09	3.937	324.377	65.31	4.58	88	994	94.40
20	1227	25	3167.615	126.7 ±56.71	18.595	233.242	75.80	4.18	105	1141	92.99
21	1057	16	2053.651	128.35 ±81.16	13.428	361.617	69.17	2.97	67	998	94.42
22	1001	19	2602.692	136.98 ±75.66	6.394	253.715	61.85	4.21	75	948	94.71
23	860	9	1400.497	155.61 ±207.39	1.568	677.221	53.38	2.62	35	828	96.28
24	1011	21	3273.713	155.89 ±94.98	0.095	351.198	65.02	5.03	89	939	92.88
25	783	14	1341.415	95.82 ±69.81	7.69	208.64	44.06	3.04	50	741	94.64
26	834	15	2127.487	141.83 ±76.15	0.281	327.541	51.75	4.11	64	787	94.36
27	767	19	2917.757	153.57 ±105.45	0.151	449.325	48.75	5.99	76	732	95.44
28	747	4	564.273	141.07 ±50.09	95.084	196.067	46.08	1.22	17	725	97.05
29	816	11	2103.018	191.18 ±294.23	13.808	1036.87	52.00	4.04	40	771	94.49
All	40472	950	146779.3	154.50 ±110.62	0.003	1263.42	2545.90	5.77	4140	37427	92.48

Table 3 – 4. Number of haplotype blocks, average block size and average number of haplotypes observed within blocks ranging in size from 2 SNPs to ≥ 10 SNPs in the Canadian Holstein and hybrid beef cattle populations.

	Canadian Holstein cattle					Hybrid beef cattle				
	N-blocks	Mean block size ± SD (kb)	Min (kb)	Max (kb)	Mean N-haplotypes	N-blocks	Mean block size ± SD (kb)	Min (kb)	Max (kb)	Mean N-haplotypes
2-SNP blocks	79	10.07 ±6.89	0.003	19.978	2.76 (2-3)	92	9.94 ±6.59	0.003	19.978	2.83 (2-4)
3-SNP blocks	9	59.93 ±103.29	20.688	335.254	3.44 (3-4)	13	24.97 ±4.03	16.507	28.717	3.46 (2-4)
4-SNP blocks	818	171.35 ±104.65	55.243	1045.49	4.34 (2-6)	512	137.61 ±65.61	54.979	505.145	4.36 (2-5)
5-SNP blocks	468	226.95 ±118.47	71.584	1131.54	5.13 (2-8)	231	191.92 ±92.54	71.584	1036.87	5.12 (2-8)
6-SNP blocks	159	287.68 ±125.88	119.291	789.47	5.62 (3-9)	53	261.76 ±170.84	121.343	1263.42	5.57 (3-9)
7-SNP blocks	90	328.65 ±206.83	160.023	1936.68	6.01 (3-9)	24	307.39 ±123.1	164.489	602.077	5.88 (3-8)
8-SNP blocks	51	373.82 ±147.53	198.365	842.903	6.10 (3-11)	14	294.02 ±58.6	202.765	398.343	6.86 (4-8)
9-SNP blocks	10	509.74 ±311.98	243.307	1244.62	6.40 (5-9)	7	418.3 ±130.7	262.428	656.932	6.71 (4-9)
≥ 10-SNP blocks	32	612.11 ±249.89	324.277	1224.84	6.90 (3-10)	4	614.21 ±116.24	456.199	708.195	9.50 (7-11)
All	1716	213.74 ±153.03	0.003	1936.68	4.80 (2-11)	950	154.5 ±110.62	0.003	1263.42	4.56 (2-11)

Table 3 – 5. Chromosomes with significant correlation between haplotype block sizes and average heterozygosities within blocks in the Canadian Holstein and hybrid beef cattle populations.

Canadian Holstein cattle			Hybrid beef cattle		
BTA	Correlation coefficient (r)	*P*-value	BTA	Correlation coefficient (r)	*P*-value
–	–	–	3	-0.3344	0.0103
5	-0.3131	0.0078	5	-0.4170	0.0083
7	-0.3677	0.0007	7	-0.4031	0.0060
13	-0.3404	0.0037	13	-0.3617	0.0236
14	-0.3583	0.0039	-	–	–
–	–	–	22	-0.7681	0.0001
23	-0.5507	0.0065	-	–	–
–	–	–	24	-0.4844	0.0260
–	–	–	27	-0.4852	0.0352

Figures

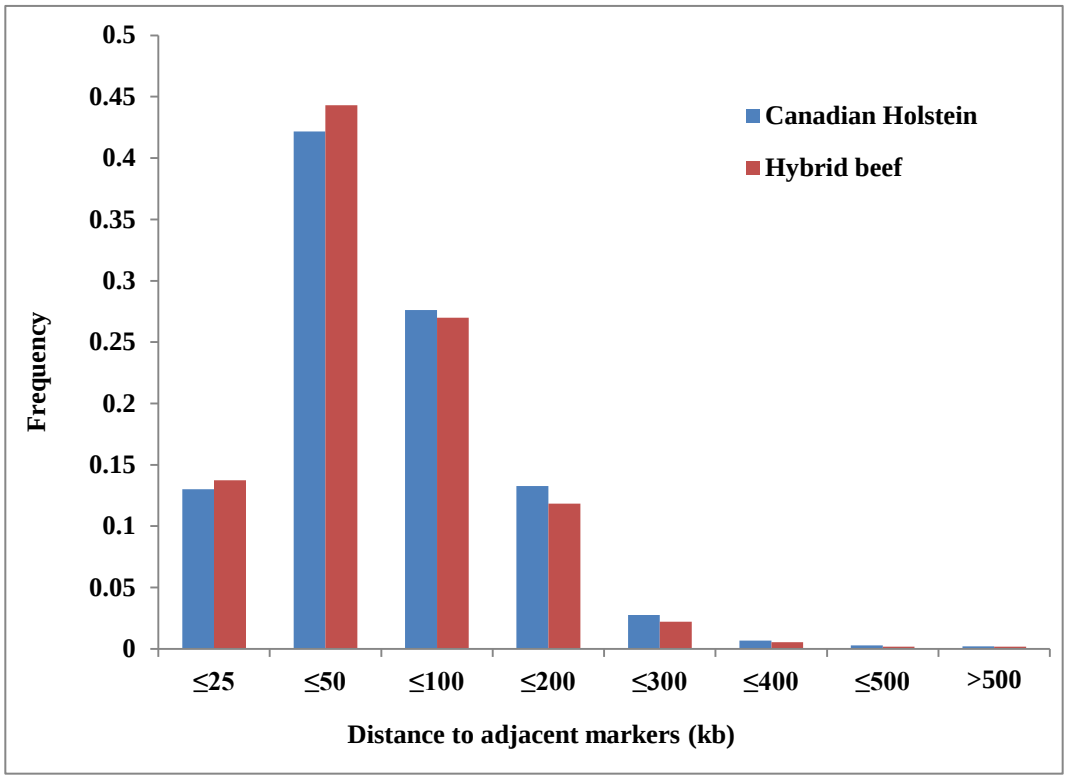

Figure 3 – 1. Distribution of adjacent marker distances at 8 distance intervals (0 – 25, 25 – 50, 50 – 100, 100 – 200, 200 – 300, 300 – 400, 400 – 500 and > 500 kb) for 37,986 SNPs in the Canadian Holstein cattle and 40,472 SNPs in the hybrid beef cattle.

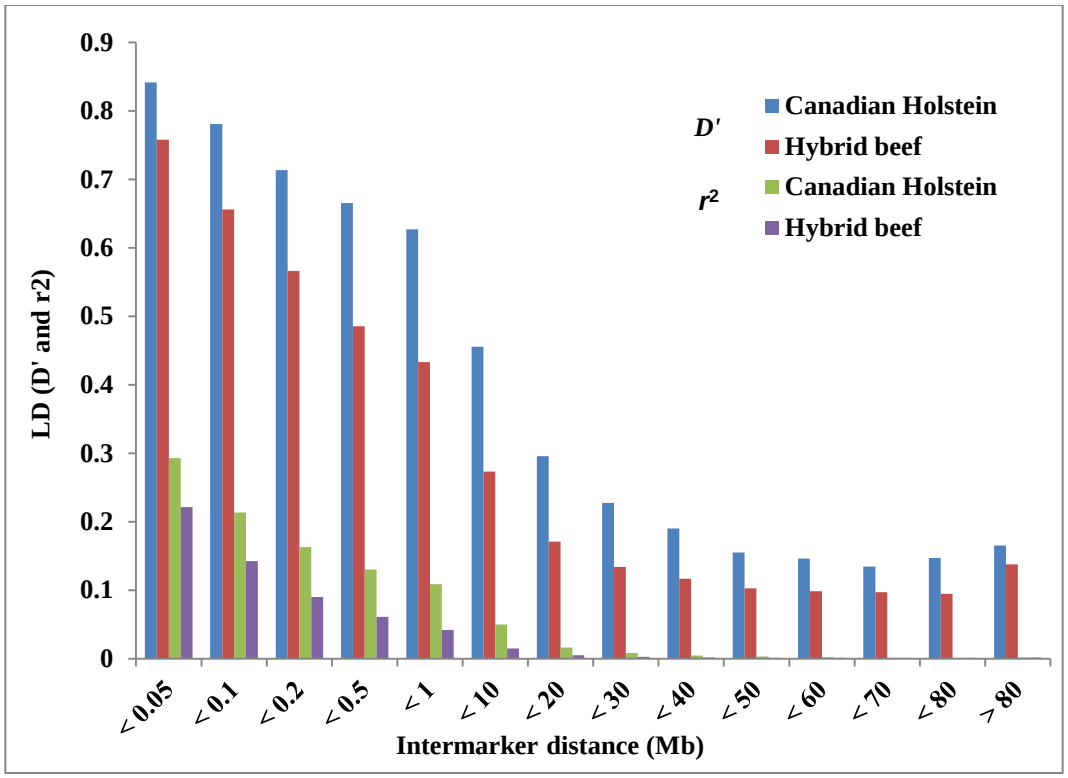

Figure 3 – 2. Distribution of *D'* and r^2 estimates for the Canadian Holstein cattle and the hybrid beef cattle on BTA14. Average pairwise LD is depicted for each cattle breed in each bin of intermarker distance.

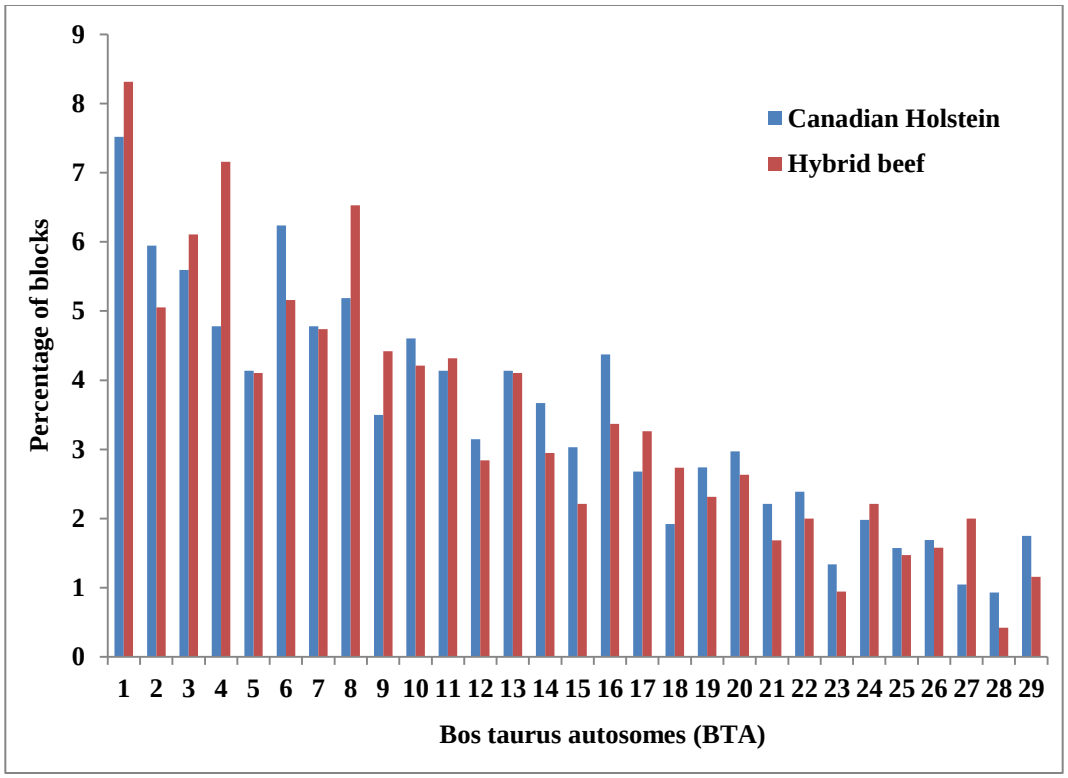

Figure 3 – 3. The percentage distribution of hyplotype blocks on 29 bovine autosomes for Canadian Holstein cattle and hybrid beef cattle.

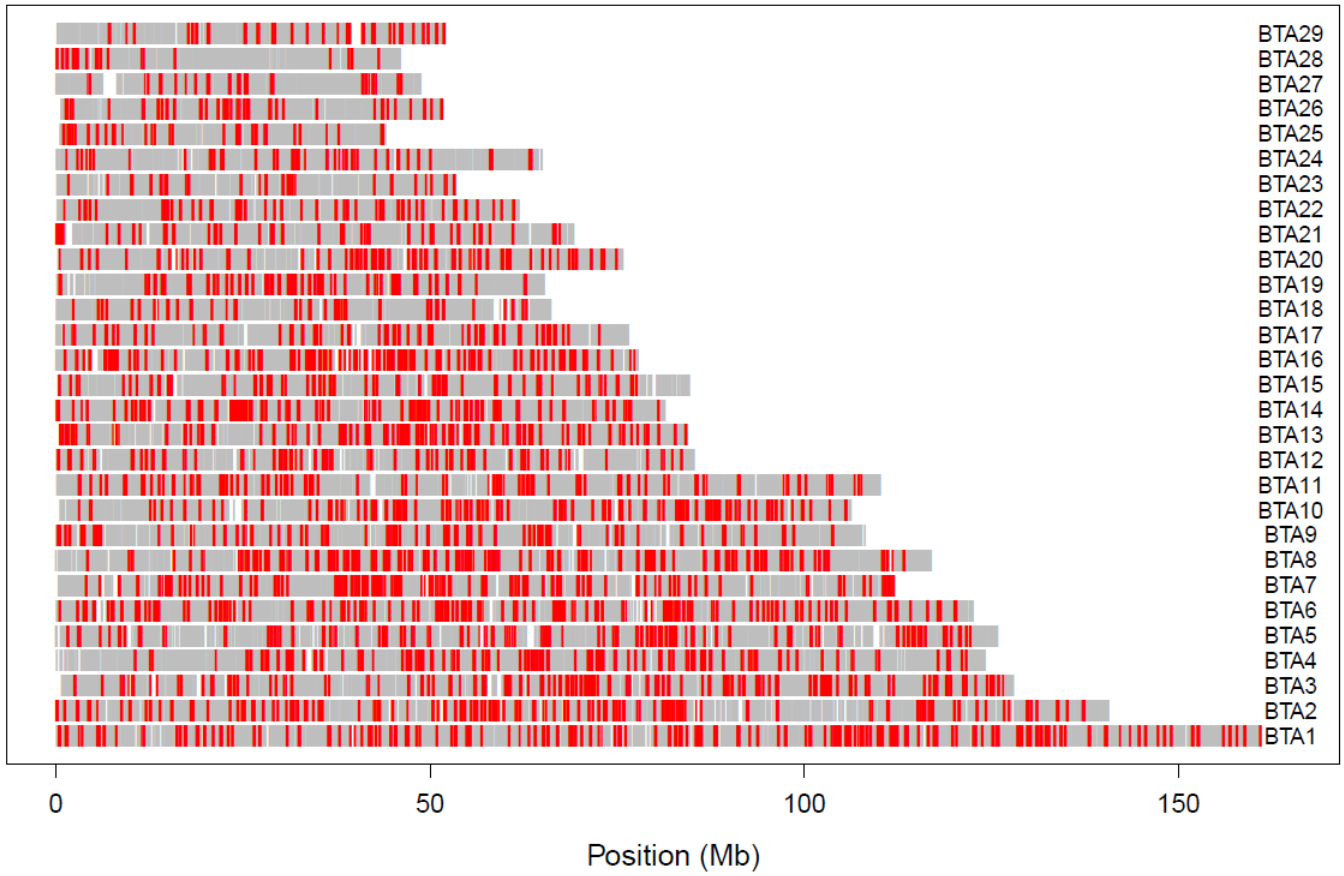

Figure 3 – 4. The physical distribution of haplotype blocks (red colour) on 29 bovine autosomes in Canadian Holstein cattle. The light grey colour represents the distribution of all 37,986 SNPs.

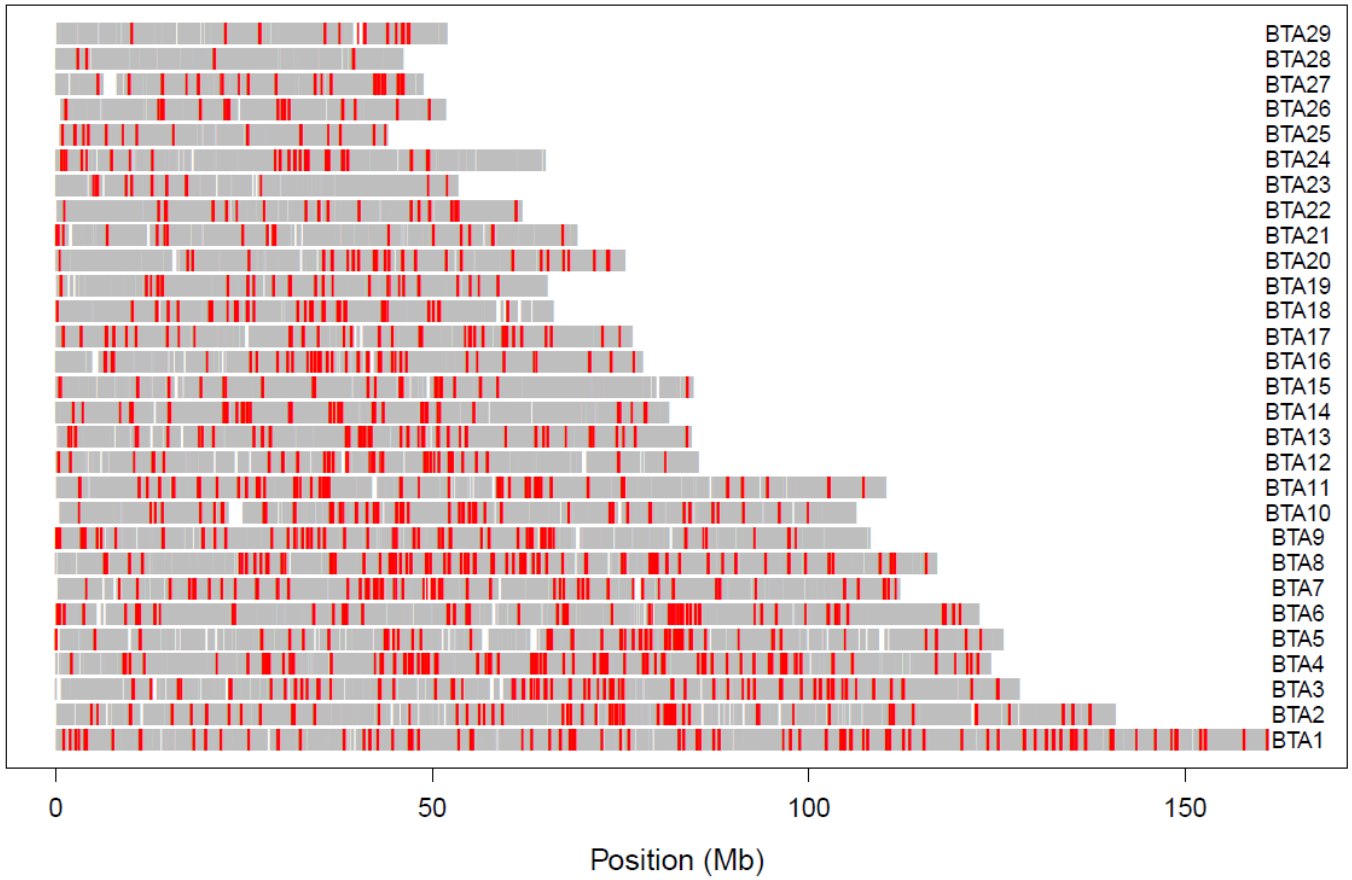

Figure 3 – 5. The physical distribution of haplotype blocks (red colour) on 29 bovine autosomes in hybrid beef cattle. The light grey colour represents the distribution of all 40,472 SNPs.

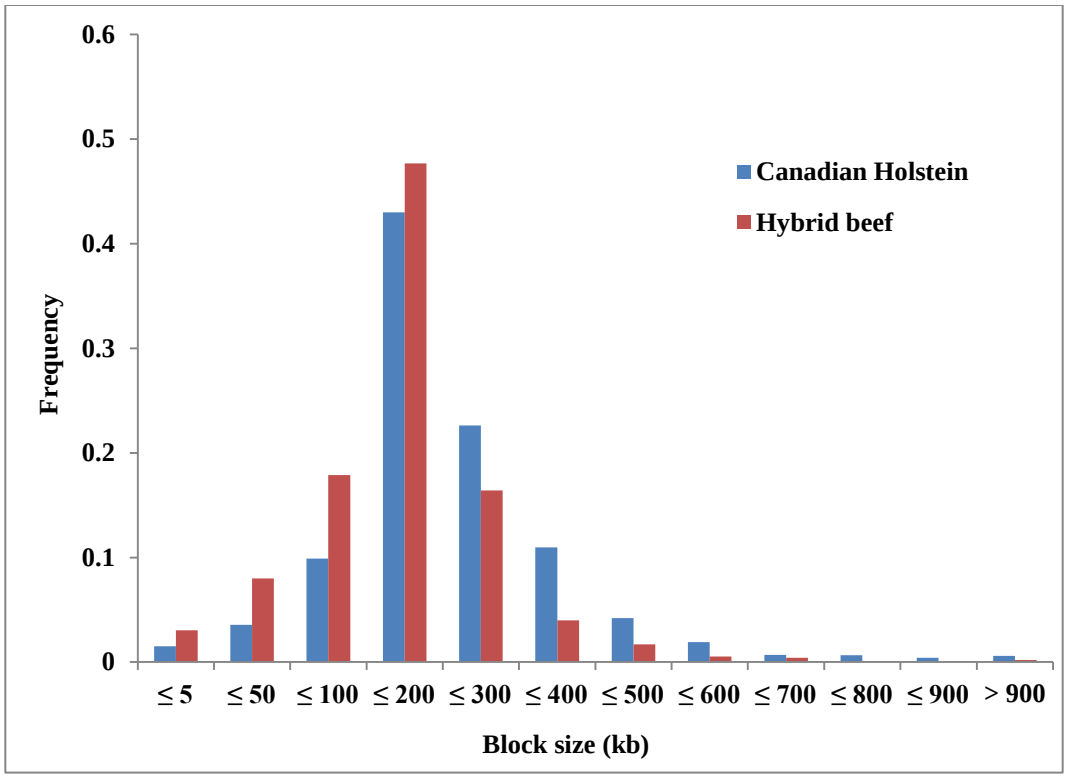

Figure 3 – 6. Frequency distribution of haplotype blocks in 12 block size categories for 1,716 blocks in the Canadian Holstein cattle and 950 blocks in the hybrid beef cattle.

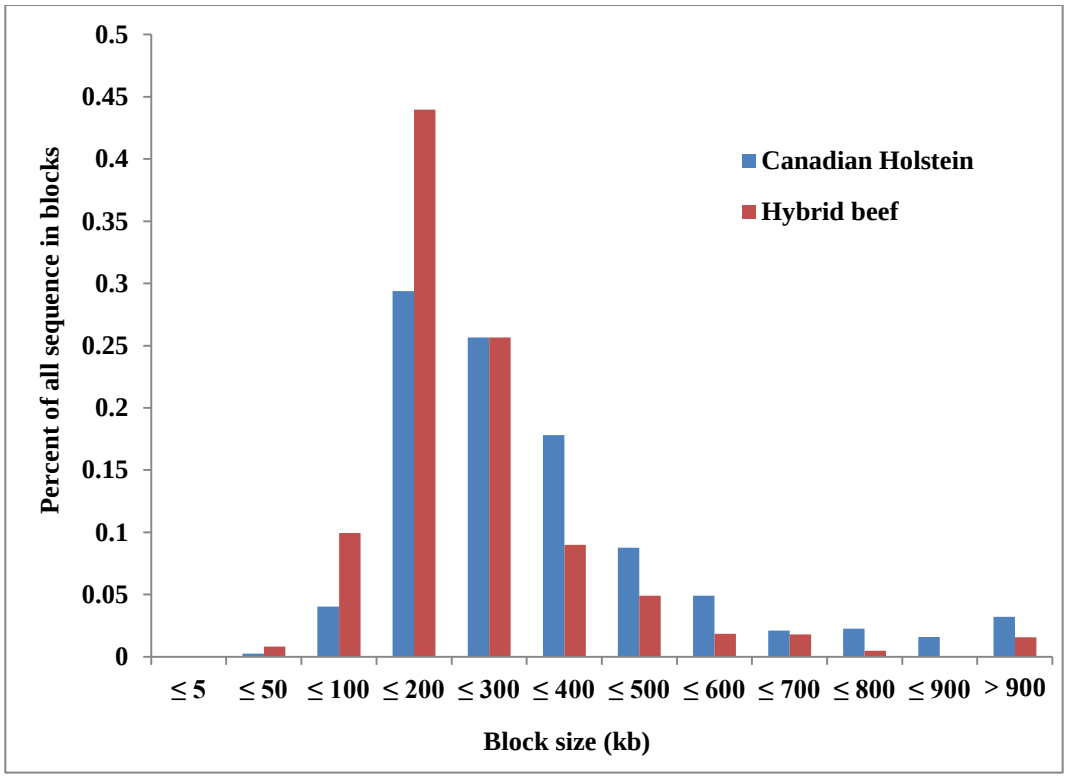

Figure 3 – 7. Proportion of genome sequence spanned by all blocks in 12 block size categories for the Canadian Holstein cattle and the hybrid beef cattle.

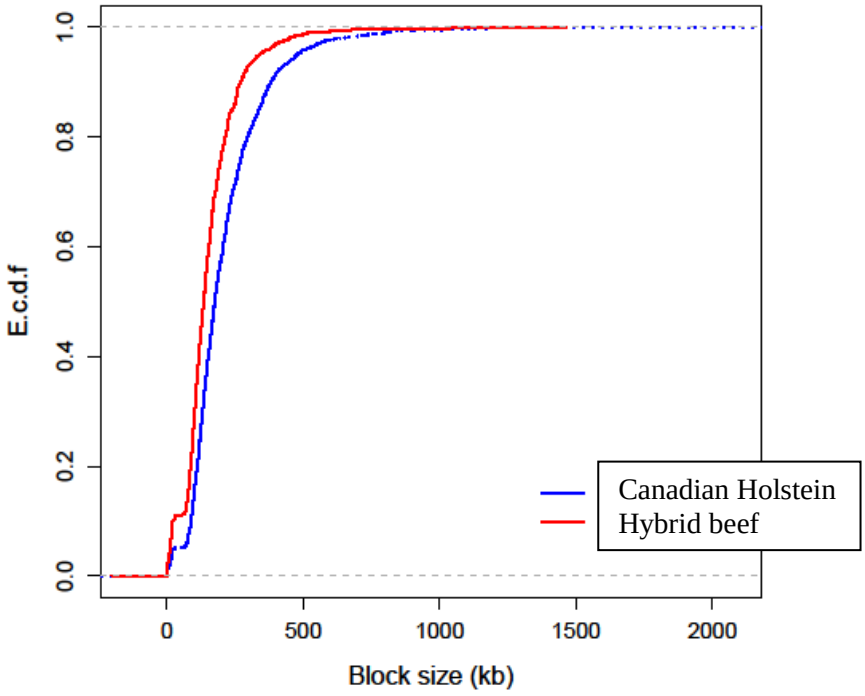

Figure 3 – 8. The patterns of the empirical cumulative distribution function (e.c.d.f.) of haplotype block size for the Canadian Holstein cattle and the hybrid beef cattle.

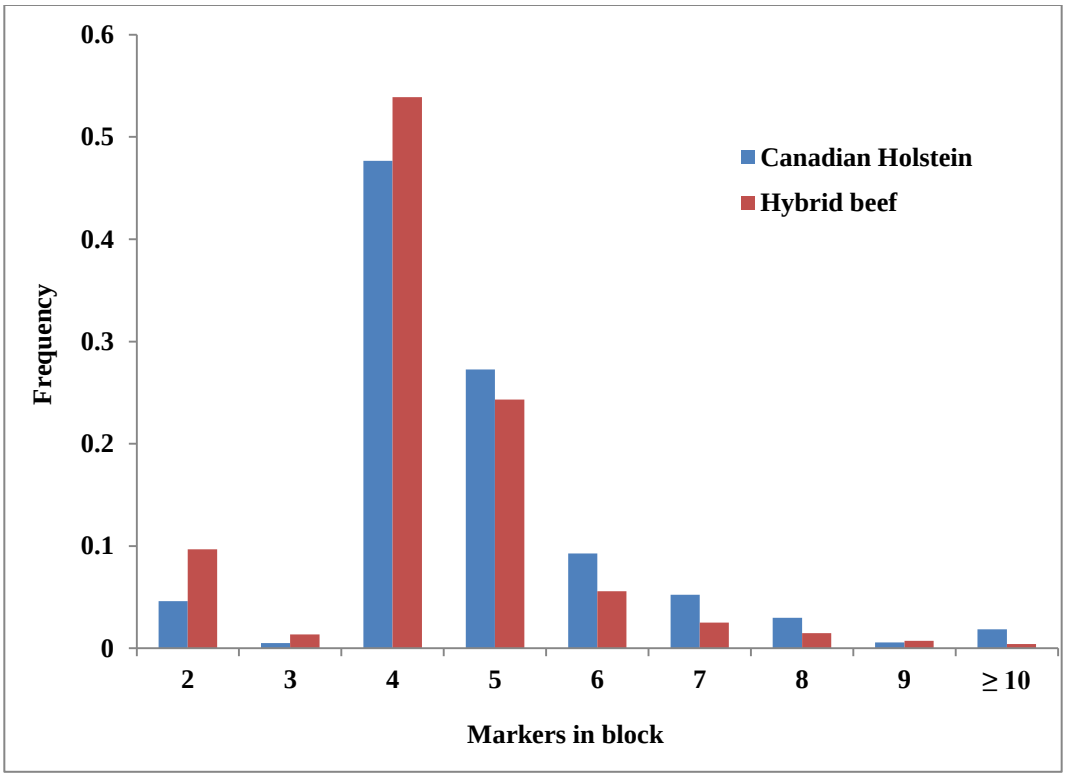

Figure 3 – 9. Frequency distribution of haplotype blocks in 9 block categories in terms of the number of SNP markers within a block (2, 3, 4, 5, 6, 7, 8, 9, and ≥ 10) for blocks in the Canadian Holstein cattle and the hybrid beef cattle.

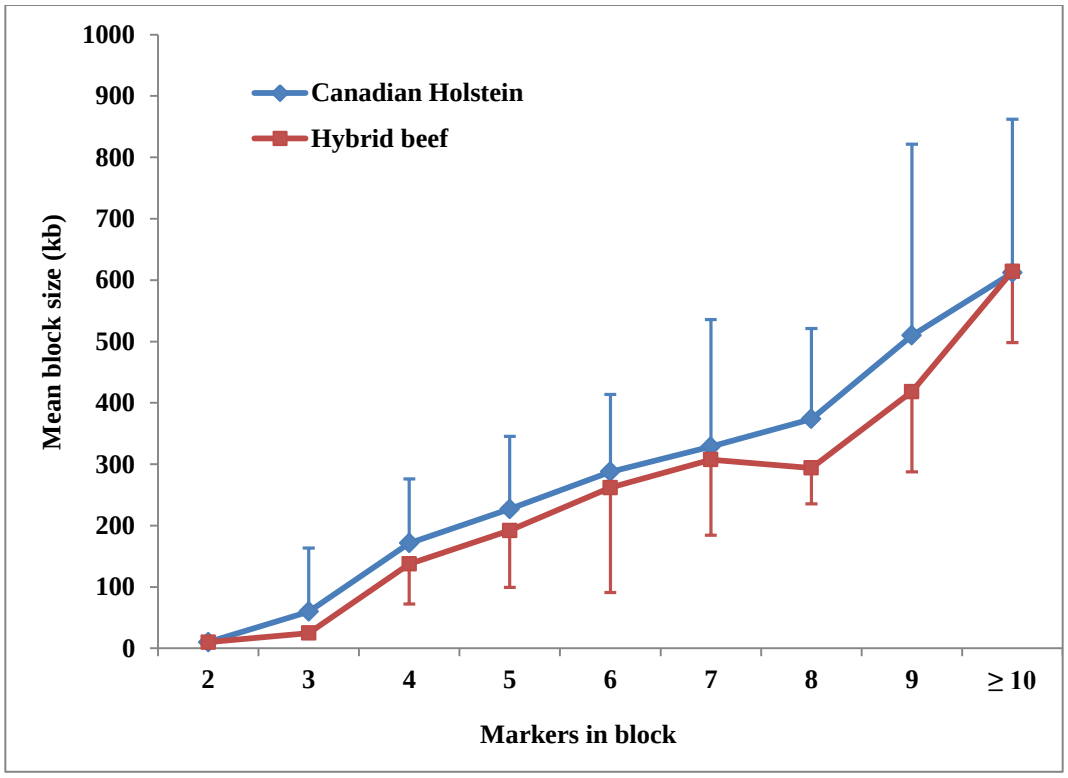

Figure 3 – 10. Haplotype block size distribution as a function of the number of markers in a block. The vertical bars represent the positive or negative standard deviation of the mean block size in different cattle populations.

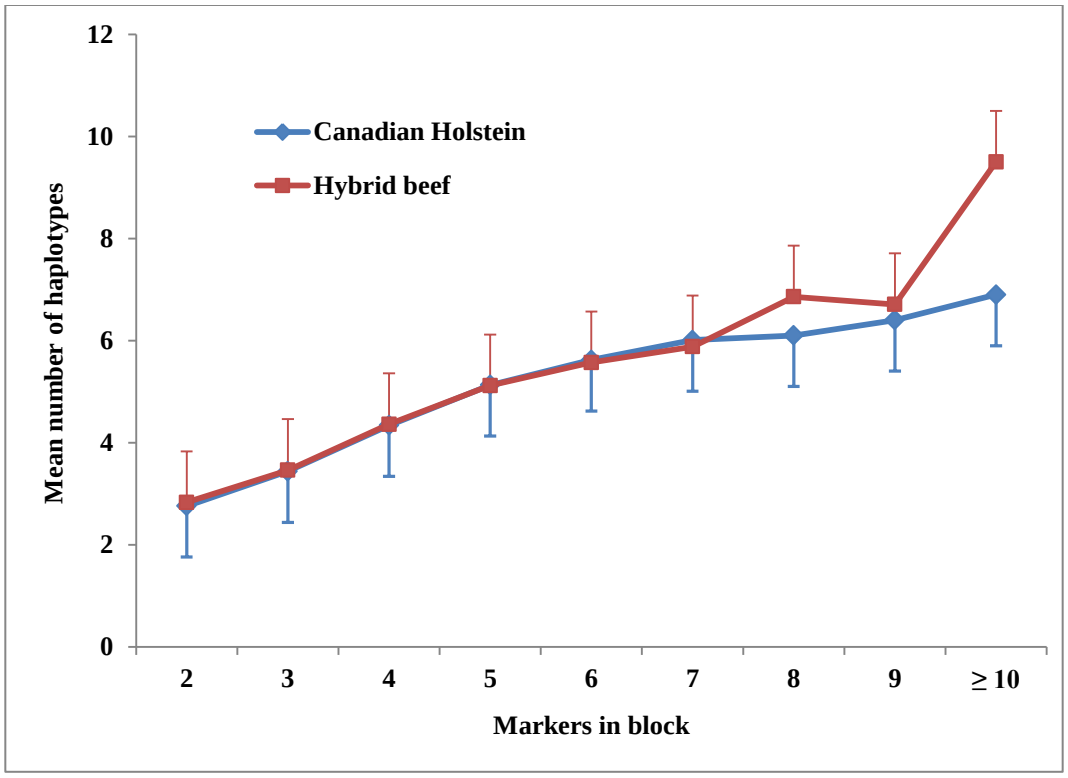

Figure 3 – 11. Distribution of the mean number of haplotypes as a function of the number of markers in a block. The vertical bars represent the positive or negative standard deviation of the mean number of haplotypes in different cattle populations.

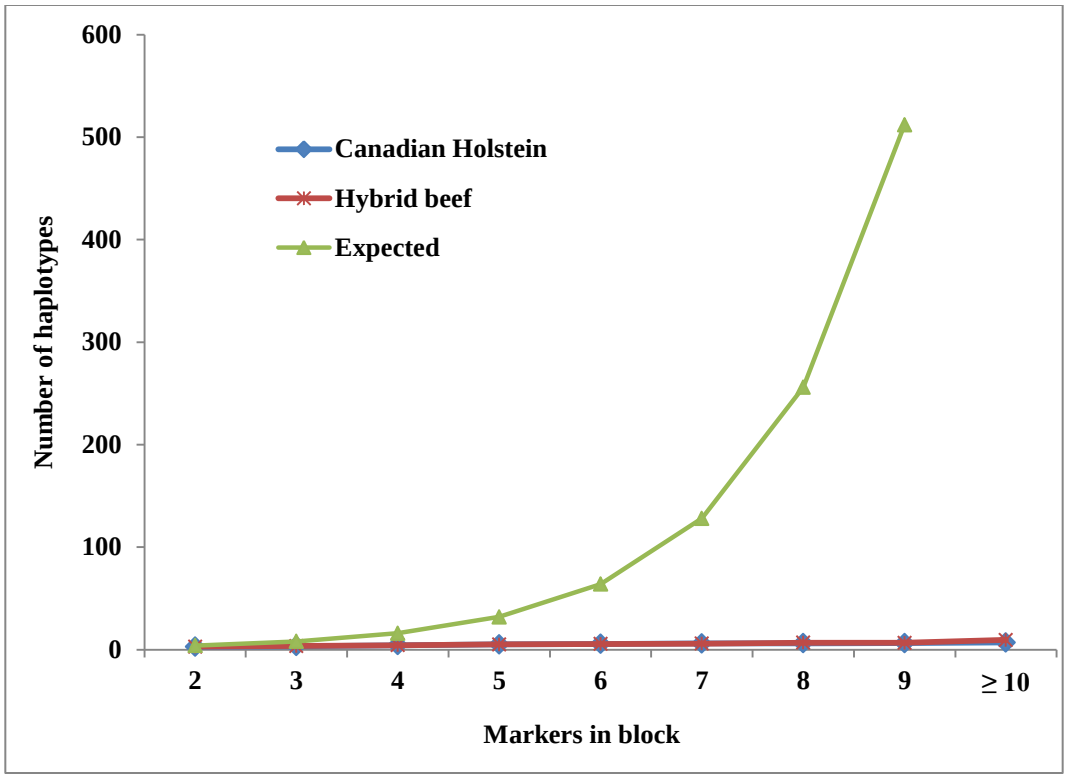

Figure 3 – 12. Distribution of the mean number of haplotypes as a function of the number of markers in block. The maximum expected number of haplotypes in a block is also presented. In a pure random combination case, a block of *N* independent biallelic SNPs could in theory generate 2^N different haplotypes (PATIL *et al.* 2001).

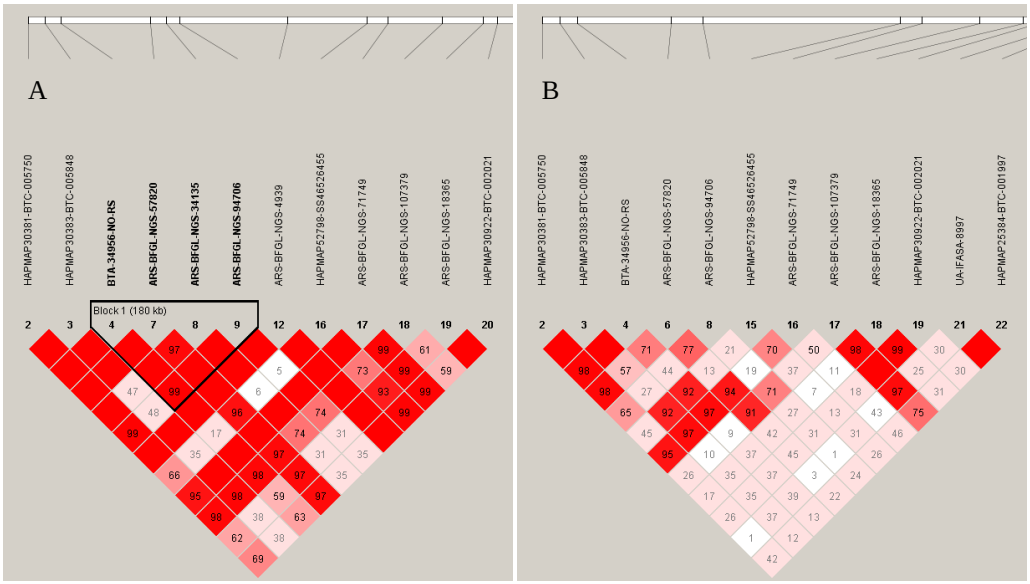

Figure 3 – 13. Comparison of haplotype block maps for a portion of BTA14 (0 – 763,331 bp) in the Canadian Holstein cattle (A) and the hybrid beef cattle (B). One block (black triangle at 101,473 – 443,937 bp) closely flanking gene *DGAT1* (444,097 – 446,810 bp) was identified only in the Canadian Holstein cattle.

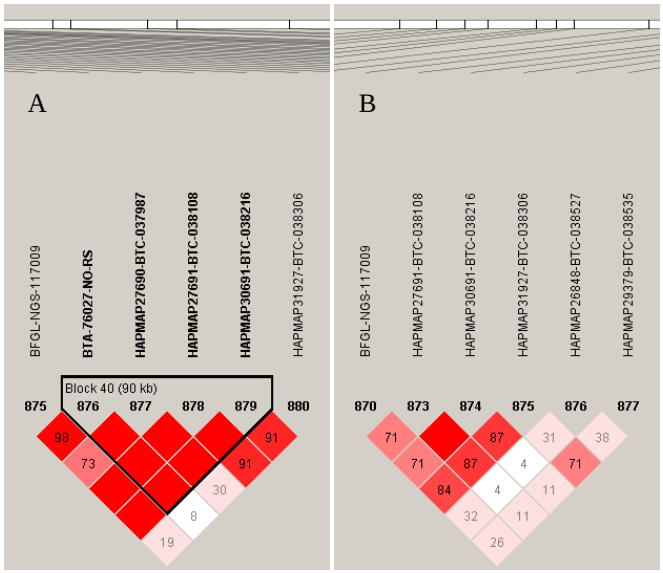

Figure 3 – 14. Comparison of haplotype block maps for a portion of BTA6 (44,553,262 – 44,729,116 bp) in the Canadian Holstein cattle (A) and the hybrid beef cattle (B). One block (black triangle at 44,553,262 – 44,643,940 bp) closely flanking gene *PPARGC1A* (44,813,186 – 44,919,653 bp) was identified only in the Canadian Holstein cattle.

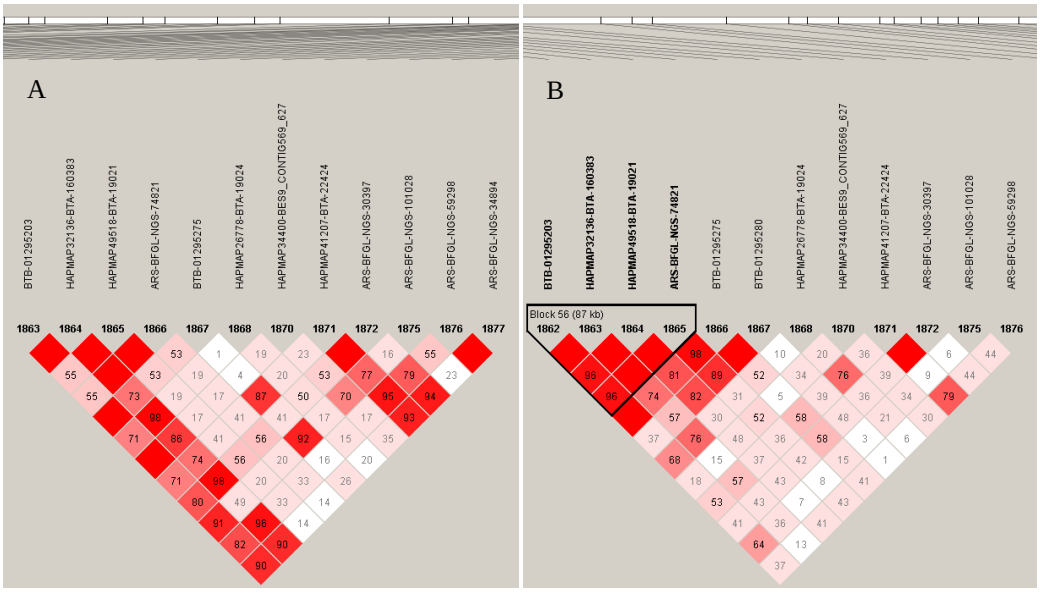

Figure 3 – 15. Comparison of haplotype block maps for a portion of BTA4 (95,097,951 – 95,715,499 bp) in the Canadian Holstein cattle (A) and the hybrid beef cattle (B). One block (black triangle at 95,097,951 – 95,185,083 bp) flanking gene *LEP* (95,655,925 – 95,672,659 bp) was identified only in the hybrid beef cattle.

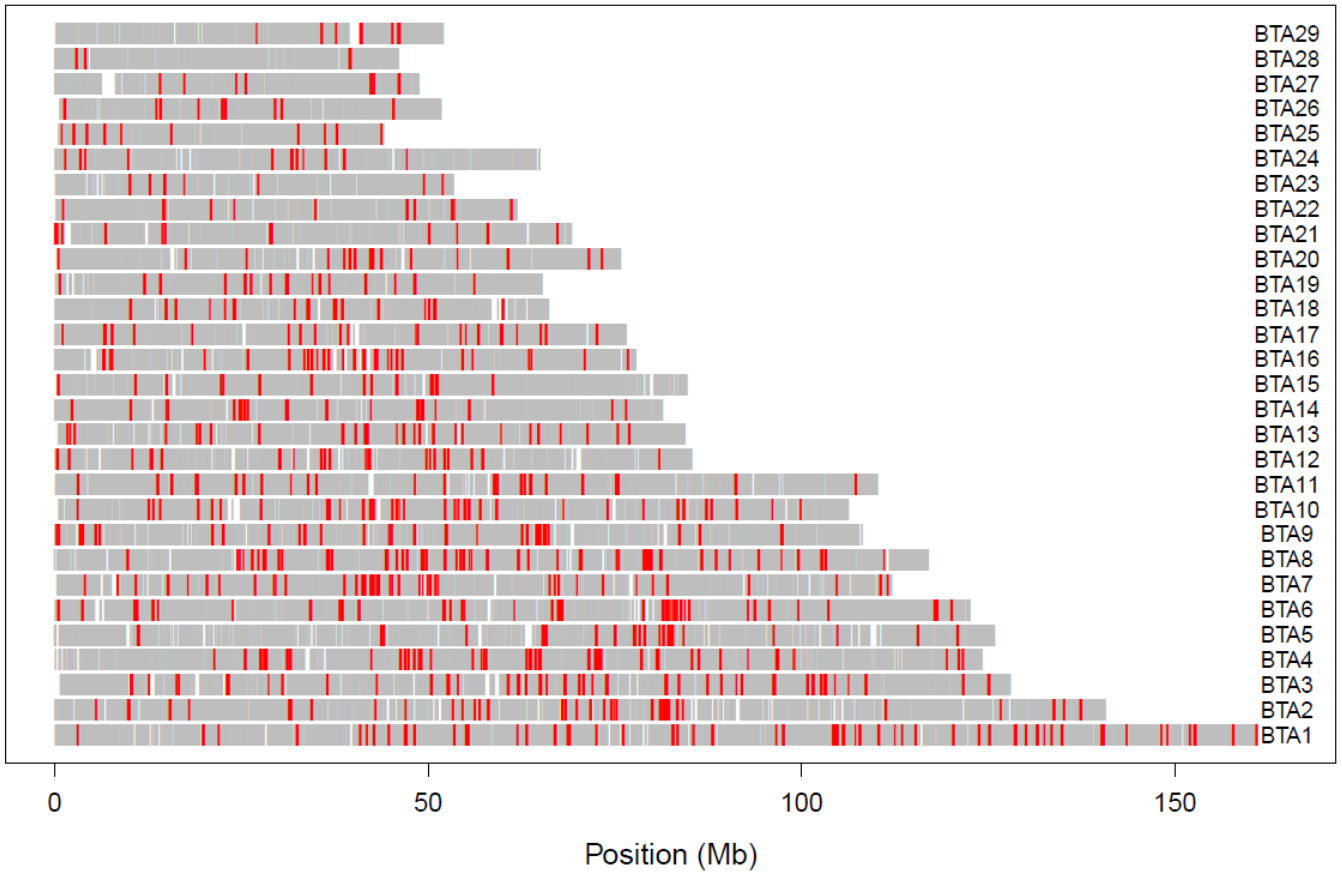

Figure 3 – 16. The physical distribution of overlapped haplotype blocks (red colour) between the Canadian Holstein cattle and the hybrid beef cattle on 29 bovine autosomes. The light grey colour represents the distribution of all 37,986 SNPs in the Canadian Holstein cattle.

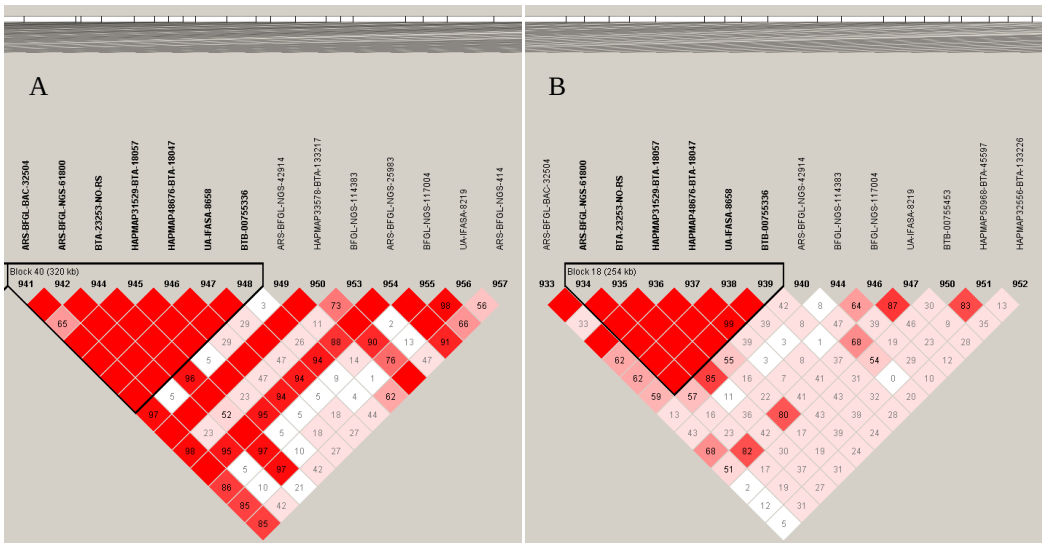

Figure 3 – 17. Comparison of haplotype block maps for a portion of BTA19 (48,045,874 – 49,028,970 bp) in the Canadian Holstein cattle (A) and the hybrid beef cattle (B). One block (black triangle) at 48,045,874 – 48,366,536 bp in the Canadian Holstein and at 48,111,550 – 48,366,536 bp in the hybrid beef) flanking gene *GH1* (48,768,617 – 48,772,013 bp) was identified in both cattle populations.

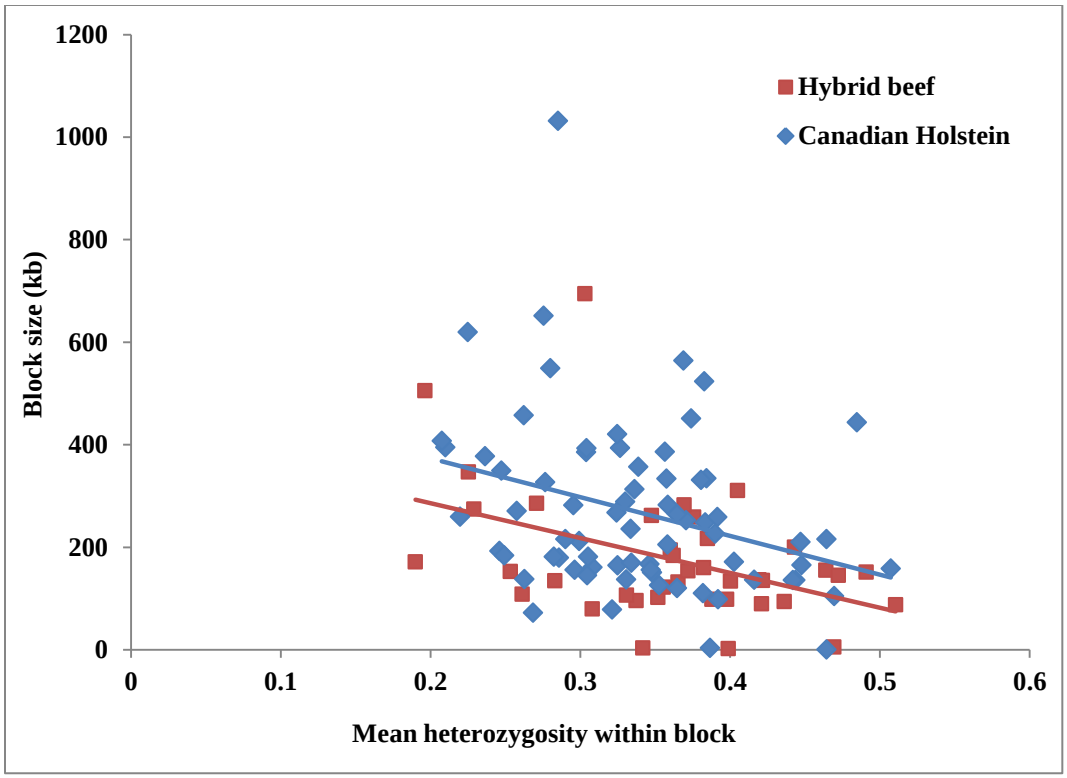

Figure 3 – 18. Distribution of block size in terms of the average heterozygosity within blocks for haplotype blocks on BTA5 in the Canadian Holstein cattle and the hybrid beef cattle. Significant negative correlations between the haplotype block size and average heterozygosity within blocks on BTA5 were observed in both populations.

Chapter 4. Whole-Genome Association Study for Milk Production Traits in Canadian Holstein Cattle

4.1. Introduction

Genetic improvement of milk production in dairy cattle based on phenotypes of pedigree herds has resulted in significant performance gains (DEKKERS and HOSPITAL 2002). In order to incorporate genetic marker information into selection of animals with superior genetic merit, many QTL mapping studies for the identification of genetic markers and genes have been carried out in different dairy cattle breeds (ASHWELL *et al.* 2001; GEORGES *et al.* 1995; OLSEN *et al.* 2002; VIITALA *et al.* 2003; ZHANG *et al.* 1998). In conjunction with functional analysis and comparative mapping, a few candidate genes and causative mutations for milk yield and composition underlying the reported QTL have been identified, such as a missense mutation in *DGAT* (GRISART *et al.* 2002; WINTER *et al.* 2002) and a missense mutation in *ABCG2* (COHEN-ZINDER *et al.* 2005; OLSEN *et al.* 2007). However, most of the previously reported QTL from linkage analysis were placed in large confidence intervals making it difficult to identify the causative genes or mutations underlying the QTL by positional cloning. Genes that account for variation in milk synthesis and secretion in dairy cattle are still largely unknown, as are the gene networks and pathways.

With the completion of the bovine genome sequence assembly and the availability of high density SNPs, genome-wide association studies (GWAS) are now available as a powerful and efficient method to detect genetic loci and causal

genes (ALTSHULER *et al.* 2008). Using GWAS, QTL can be fine mapped within regions of 1 – 2 Mb (SELLNER *et al.* 2007), thus facilitating the subsequent search for causative mutations. In this way, hundreds of genetic variants associated with complex human diseases and traits had been identified (MANOLIO *et al.* 2009). GWAS has also been applied in livestock (CHARLIER *et al.* 2008; HAYES *et al.* 2009). In dairy cattle, GWAS had been carried out for mapping QTL or genomic regions explaining variation in milk production using single marker LD regression (JIANG *et al.* 2010; KOLBEHDARI *et al.* 2009; MAI *et al.* 2010) or haplotypes of SNPs (PRYCE *et al.* 2010) or multiple-trait GWAS (BOLORMAA *et al.* 2010). However, GWAS using the BovineSNP50 BeadChip had not been conducted in the Canadian Holstein cattle for the discovery of genes underlying milk production traits. In addition, GWAS using genomic selection methods has not been carried out in this population. These methods have proven useful in genome association studies in swine (FAN *et al.* 2011; ONTERU *et al.* 2011).

The objective of this study was to identify genome regions and potential functional candidate genes affecting milk production traits in Canadian Holstein cattle. Associations of high density genome-wide SNPs with targeted traits were investigated using single marker LD regression and Bayesian regression. Functional candidate genes for milk production were identified within the associated regions, and potential gene networks influencing the traits were defined.

4.2. Materials and Methods

4.2.1. Animals and Traits

The Canadian Holstein bulls used for this study are described in Chapter 3. Five milk production traits including milk yield (MY), fat yield (FY), protein yield (PY), fat percentage (FP) and protein percentage (PP) were considered. The estimated breeding values (EBVs) of these traits were obtained from the online genetic evaluations files of the Canadian Dairy Network released in April 2008 (www.cdn.ca). Finally, de-regressed EBVs were used as the response variable in this study based on the previous studies showing that de-regressed EBVs can produce more reliable results in genomic regression analysis than EBVs (GARRICK *et al.* 2009; OSTERSEN *et al.* 2011). De-regression was done according to the following mixed model equation (SCHAEFFER 1994):

$$\left[D^{-1}+A^{-1}\alpha\right]\left[u\right]=\left[D^{-1}y\right].$$

Where D is a diagonal matrix containing the inverse of the number of effective daughters on the diagonal, A is the numerator relationship matrix among bulls with EBVs, $\alpha=(4-h^2)/h^2$, h^2 is heritability, u is the EBV vector and the y vector contains de-regressed EBVs. The *A* matrix for the 647 bulls was obtained from a pedigree tracing back to 1960. Here the y vector was unknown. Since the numbers of effective daughters were not available, the numbers of daughters (n) were calculated from the reliabilities (r) as $r=\frac{n}{n+\alpha}\Rightarrow n=\frac{r\alpha}{1-r}$. The

descriptive statistics of these traits are given in Table 4 – 1.

4.2.2. SNP Markers

The DNA collection and genotyping for these animals are described in Chapter 3. SNPs not mapping to the Btau4.0 reference assembly and SNPs located on chromosome X were excluded, as were SNPs with missing genotype calls. In addition, SNPs with MAF < 0.1 were discarded. This filtering yielded 29,552 SNPs for use in the whole genome association analysis. The summarized chromosome-wide SNP information is given in Table 4 – 2.

4.2.3. Genome-wide Association Studies

4.2.3.1. Single Marker LD Regression

4.2.3.1.1. Statistical Analysis

The phenotypic values of individuals were regressed on their SNP genotypes. SNPs were assumed to be in LD with the QTL over the entire genome. The simple single locus LD regression model was shown to have good power and accuracy for QTL fine mapping (ZHAO *et al.* 2007). The association between the phenotypic values and marker genotypes was implemented by successively fitting single SNPs in a linear mixed model (YU *et al.* 2006) as follows:

$$y = Xb + Z\alpha + e\,.$$

Where *y* is the vector of phenotypes; *X* is the design matrix; *b* is the vector of coefficients of the regression on recoded SNP genotypes; *Z* is the incidence matrix for animal effects; $\alpha \sim N(0, A\sigma_a^2)$ is a vector of the polygenic animal effects and $e \sim N(0, I\sigma_e^2)$ is the vector of residuals, in which *A* is an additive genetic relationship matrix of animals and *I* is an identity matrix, and σ_a^2 and σ_e^2 are the animal's additive polygenic variance and residual error variance, respectively. SNP allele substitution fixed effects (*b*) and random background polygenic effects (α) were evaluated in this model. Values in the design matrix, *X*, were coded as 0, 1, 2 for the SNP genotypes, representing the number of copies of the minor allele carried by the individual. The F-statistic, type I error (*P*-value) and allele substitution effects were estimated for all SNPs. The analysis was performed using a R script to call univariate analysis using the animal model in the ASReml package (GILMOUR 2009).

4.2.3.1.2. Significance Testing using False Discovery Rate

To control the type I error in multiple testing, a false discovery rate (FDR) correction, which is the expected proportion of falsely detected QTL, was used to establish the statistical significance critical value in this study (BENJAMINI and HOCHBERG 1995). Significant control based on genome-wise type I error was used in this study. The genome-wise threshold was adjusted based on the total *n* number of SNPs on the bovine genome used in this study and represents a very conservative approach for large number of markers in a whole genome association study. Here, *n* numbers of tests were performed on a genome, and the

P-values were ranked from lowest to highest. FDR was calculated as:

$FDR = \frac{n \times P(k)}{k}$, where *k* is the individual relative test position in the rank. Correction was performed for each trait separately. Significant SNPs were detected at three genome-wise FDR thresholds (5%, 1% and 0.1%).

4.2.3.2. Bayesian Regression

4.2.3.2.1. Statistical Analysis

Statistical analyses were conducted using Bayesian methods implemented in the GenSel software (http://bigs.ansci.iastate.edu) for each trait separately. The statistical model is briefly introduced here and a more detailed explanation of the method is given by Kizilkaya et al. (2010). Let *n* be the number of animals and *K* be the number of SNPs. The vector of phenotypic values for a trait can be described by the following linear model,

$$y_i = Xb + \sum_{j=1}^{K} Z_j u_j \delta_j + e_i$$

where y is a n $\times 1$ vector of phenotypes of the analyzed trait, X is the incidence matrix for fixed effects, *b* is the vector of fixed effects, Z_j is a $n \times 1$ vector of the genotype covariate indicators for locus $j(\forall j = 1,...,K)$, Z_j takes one of three values $\{10, 0, -10\}$ depending on the genotype of animal i for locus j, u_j is the random substitution effect for locus *j*, which is assumed to be normally

distributed $N(0,\sigma_u^2)$ when δ_j = 1 but u_j = 0 when δ_j = 0. δ_j is a random 0 / 1 variable indicating the absence (with probability π) or presence (with probability $1-\pi$) of locus j in the model, and e is the residual error vector with an assumed normal distribution $N(0,\sigma_e^2)$ (KIZILKAYA *et al.* 2010). The Bayesian mixture models assumed $\pi = 0.99$ or $1-\pi = 0.01$, corresponding to about 300 non-zero SNPs fitted per iteration of each Markov chain. The analyses were implemented with 1,000 burn-in iterations and 40,000 MCMC iterations. Since de-regressed EBVs were used for milk production traits, weighted Bayesian analysis was carried out in this study and the appropriate weights for analyzing the de-regressed data with heterogeneous variance were:

$$w_i = \frac{1-h^2}{[c+(1-r_i^2)/r_i^2]h^2},$$

where *c* is the part of the genetic variance not explained by markers, h^2 is the heritability of the trait, and r_i^2 is the reliability of the de-regressed EBV of the *i*th animal (GARRICK *et al.* 2009). In this study, *c* = 0.5 and h^2 = 0.37 were used. Results were obtained in the form of a post burn-in posterior distribution for the effect of every SNP fitted simultaneously with other informative SNPs.

In this study, we performed association analyses through a combination of Bayesian models with different assumptions on the variance components of each marker (hyper-parameters of the normal priors for the regression coefficients) (FERNANDO and GARRICK 2009; KIZILKAYA *et al.* 2010). The total genetic

variance and residual variance components were first estimated using the BayesC method. The BayesC method is less sensitive to the prior genetic variance compared to the BayesB method and is useful for the estimation of prior genetic and error variances (KIZILKAYA *et al.* 2010). The BayesC method assumes that SNP markers have a common variance and therefore the same variance ratio in the mixed model equations is used to sample effects. The variance components estimated from BayesC were used in BayesB to estimate all SNP marker effects simultaneously. The BayesB method not only assumes a prior distribution of marker effects where many SNPs are likely to have no effect and only a few will have a moderate to large effect, but also allows each marker to have its own variance and degree of shrinkage in the model (MEUWISSEN *et al.* 2001).

In order to remove the effects of many SNPs in high LD with a particular QTL, genomic merit was predicted for 1 Mb sliding windows based on the posterior means of the effects of the SNPs within the window. In total, there were 2,552 unique non-overlapping SNP windows in the genome. The most informative genomic regions were selected based on the proportion of variance explained by each 1 Mb window and the SNP within this window that explained the largest proportion of genetic variance was used to denote the variance explained by the window. The estimated proportion of genetic variance contributed by the 1Mb sliding windows was plotted against genomic location using the R software.

4.2.3.2.2. Hypothesis Testing

The SNP windows contributing the most genetic variance were considered as putative QTL for each trait and their significance values were calculated based on the posterior distributions of the test statistic. In this study, the posterior distributions of the genetic variance explained by each 1Mb SNP window were generated and compared with the null hypothesis of no QTL in the identified SNP window. SNP windows with $P < 0.2$ were selected as suggestive QTL for the purpose of functional gene searches.

4.2.3.3. Comparisons of Identified QTL

The significant SNPs and QTL regions identified from single marker LD regression and Bayesian regression were compared with previously reported QTL for the same trait retrieved from the CattleQTLdb database (http://www.animalgenome.org/cgi-bin/gbrowse/cattle/). Results from the two statistical methods used in this study were also compared.

4.2.3.4. QTL Annotation

4.2.3.4.1. Gene Network Analysis among Positional Candidate Genes

All positional candidate genes located within 0.5 Mb windows on each side of the significant SNPs in single marker LD regression were identified using the Ensembl Genome Browser (www.biomart.org). Possible gene networks among all positional candidate genes were identified and those related to milk production traits were selected based on known relationships and functions

available in the Ingenuity Pathway Analysis (IPA) database (Ingenuity Systems, www.ingenuity.com).

4.2.3.4.2. Gene Search and Functional Annotation

Gene identification was performed for suggestive QTL regions from Bayesian regression using Ensembl release 63. Gene annotations were obtained from the National Center for Biotechnology Information (NCBI) as Gene Ontology terms and as information from associated articles in PubMed and from Gene References into Functions (GeneRIFs). A list of potential functional candidate genes with evidence linking them to lipid and protein metabolism and intracellular molecular transport was compiled.

4.3. Results and Discussion

Association analyses for milk production traits were performed using single marker LD regression and Bayesian regression. The average marker interval, heterozygosity and MAF for the SNPs used in this study were 85.91 ± 87.71 kb, 0.402 ± 0.098, and 0.303 ± 0.115, respectively (Table 4 – 2).

4.3.1. Association Analyses using Single Marker Regression

Single marker regression identified 94, 97, 17, 206, and 136 significant SNPs at the genome-wise FDR $P < 0.05$ for MY, FY, PY, FP, and PP, respectively. Of these markers, 49, 54, 6, 150, and 72 were significant at the genome-wise FDR $P < 0.01$ and 24, 31, 0, 119, and 39 were significant at the

genome-wise FDR $P < 0.001$ for MY, FY, PY, FP, and PP, respectively (Figure 4 – 1). No genome-wise significant SNPs at FDR $P < 0.001$ was identified for PY. The physical position, estimated allele substitution fixed effect, heterozygosity and MAF of the identified significant SNP, are presented for each trait in Tables 4 – 3 to 4 – 7. There is considerable overlap among the SNPs identified for some of the traits, for example, 59 SNPs are shared between FY and FP, and 103 SNPs are shared between FP and PP (Appendix 1). Most of these common SNPs are located on BTA14.

In total, 316 SNPs on 28 chromosomes were identified for five milk production traits using a genome-wise FDR of 5%. A large proportion of the identified SNPs are clustered on BTA14: 60 out of 94, 61 out of 97, 4 out of 17, 151 out of 206, and 106 out of 136 for MY, FY, PY, FP, and PP, respectively. These BTA14 SNPs are located within the 0 – 16 Mb region with only one exception, for PP (Table 4 – 7). Additional clusters of significant SNPs were observed (Table 4 – 8 and Figure 4 – 2). For example, a cluster of SNPs on BTA5 was observed for FY and FP, and clusters of SNPs on BTA6 were found for PP (Table 4 – 8, Figure 4 – 3). BTA1, 5, 11 and 14 are chromosomes having identified SNPs affecting at least four out of the five analyzed milk production traits.

A detected QTL was deemed to be supported by a previously reported QTL if it overlapped with QTL reported in the cattle QTL database (www.animalgenome.org/cgi-bin/QTLdb/BT/index). Generally, results from this

study agree very well with previously reported QTL, although many new large-effect SNPs were identified (Table 4 – 9).

4.3.1.1. Milk Yield

For MY, we reported novel associations on BTA9, 10, 11, 12, 17, and 28, and confirmed previously reported QTL on BTA1, 5, 7, 14, 18, 19, 21, and 27. SNPs on BTA14 for MY were located within 0.07 – 15.52 Mb and overlap with a QTL reported in several different cattle breeds (BAGNATO *et al.* 2008; BOICHARD *et al.* 2003; DAETWYLER *et al.* 2008; THALLER *et al.* 2003). The gene *DGAT1* within this region was reported to have a large effect on milk yield and composition (GRISART *et al.* 2002; SUN *et al.* 2009). Similar results on BTA14 were observed for other milk production traits. BTA1 and BTA5 may harbor important variants for milk yield since several significant SNPs were reported in this study and in previous studies, for example Viitala *et al.* (2003).

4.3.1.2. Fat Yield

Significant SNPs for FY detected at a genome-wise FDR of 5% were placed on 15 chromosomes. BTA5 was found to be an important chromosome for FY with 13 significant SNPs. Eleven of these SNPs, which clustered together, overlap with a previously described QTL at 80.145 – 111.5 cM (Olsen et al. 2002) and partially overlap with other reported QTL in this region for fat yield (LILLEHAMMER *et al.* 2007; LUND *et al.* 2008). In this study, we identified only one SNP on BTA6 for FY, however, this SNP overlaps with several previously

reported QTL and the candidate gene *CSN3* (BOVENHUIS and WELLER 1994; HECK *et al.* 2009). We also identified several novel associated SNPs on BTA1, 3, 5, 7, 9, 15, 24, 25, and 29, for example, two SNPs at 111.26 – 112.16 Mb on BAT3 and two SNPs at 14.42 – 14.50 Mb on BTA5.

4.3.1.3. Protein Yield

For PY, 17 significant SNPs at a genome-wise FDR of 5% threshold were distributed on 9 chromosomes. SNPs identified on BTA5, 9, 16, and 21 were novel findings from this study. We found that the significant SNPs on BTA1, 10, 11, and 24 are located close (within 2 – 5 cM) to several previously reported QTL for protein yield in Canadian Holstein cattle (DAETWYLER *et al.* 2008), however, they do not overlap. This dis-concordance may result from the way that the SNP centimorgan positions were interpolated from base-pair positions in the previous study. The different maker panels and methods used could also affect the results. The QTL reported for PY in Daetwyler *et al.* (2008) were identified either using variance component linkage analysis or single marker LD regression and 9,919 SNPs from the Affymetrix MegAllele GeneChip Bovine Mapping 10K SNP array.

4.3.1.4. Fat Percentage

In comparison with the number of significant SNPs identified for yield traits (MY, FY, and PY), larger numbers of SNPs were found for the percentage traits (FP and PP). For FP, significant SNPs were placed on 15 chromosomes. Across all five milk production traits, significant SNPs on BTA4 were only

identified for FP and 4 out of the five SNPs are located within a QTL at 52.49 – 112.7 cM (LINDERSSON *et al.* 1998). Similar to FY, BTA5 was shown to be an important chromosome for FP with 18 identified significant SNPs. Sixteen out of these 18 SNPs are supported by previous studies (BENNEWITZ *et al.* 2003; HEYEN *et al.* 1999), and a candidate gene, *OLR1*, has been identified (SCHENNINK *et al.* 2009). In addition, the SNP on BTA20, located within many previously identified QTL (ARRANZ *et al.* 1998; ZHANG *et al.* 1998), is close to the reported candidate gene *GHR* (SUN *et al.* 2009; WATERS *et al.* 2011). This study has also revealed novel SNPs for FP on BTA1, 3, 4, 8, 10, 11, 12, 15, 16, 17, and 18 including 7 SNPs on BTA1 within 116.61 – 151.29 Mb and 4 SNPs on BTA12 within 49.50 – 50.88 Mb.

4.3.1.5. Protein Percentage

Significant SNPs for PP are distributed on 10 chromosomes. Ten significant SNPs on BTA6 were identified, several of which coincide with previously reported QTL in this region (BENNEWITZ *et al.* 2004; CHEN *et al.* 2006; OLSEN *et al.* 2004). All but one SNP of the PP SNPs on BTA14 are clustered within a region at 0.05 – 14.79 Mb and the significant SNP not in this region at 59.77 Mb overlaps with a QTL peak at 60.05 cM (SCHNABEL *et al.* 2005). As with FP, significant SNPs on BTA20 are identified for PP. These SNPs on BTA20 were located within a few previously reported QTL (ARRANZ *et al.* 1998; BENNEWITZ *et al.* 2004; BOICHARD *et al.* 2003) and one of these SNPs is located close to the previously reported candidate gene *GHR* (REARDON *et al.* 2010; SUN

et al. 2009; WATERS *et al.* 2011). For PP, we also identified several novel significant SNPs on BTA6, 11, 13, 15, 17 and 23.

Following the association analyses, we identified positional candidate genes within 0.5 Mb windows on each side of the significant SNPs and identified one potential gene network among these positional candidate genes for MY, PY and PP each through functional clustering analysis (Table 4 – 10). We did not find clear networks related to FY and FP based on current functional annotations available in the IPA database. The gene networks identified are mainly involved in carbohydrate and protein metabolism, protein post-translational modification, and molecular transport, and may be important for milk production in dairy cattle. The physical positions and cellular locations of the genes involved in the gene networks are presented in the Appendix 2 – 4. In comparison with the other recent GWAS for milk production traits using high density SNP chips (JIANG *et al.* 2010; MAI *et al.* 2010; PRYCE *et al.* 2010), our study includes a relatively systematical search for candidate genes or gene networks following the GWAS. Overall, our study using single marker LD regression identified a larger number of significant SNPs as well as potential gene networks involving the positional candidate genes. These SNPs, genes and networks may provide future directions in genetic improvement and candidate gene studies for milk production traits.

4.3.2. Association Analyses using Bayesian Regression

Posterior variance components and the genetic contribution of each SNP window were estimated from the BayesB method for five milk production traits.

Within the total variance of the de-regressed EBVs, the proportion explained by SNPs was 0.41 for MY, 0.48 for FY, 0.38 for PY, 0.72 for FP and 0.48 for PP (Table 4 – 11). In comparison with PY and MY, a larger genetic variance was accounted for by the SNPs for FP, with about 300 SNPs likely accounting for 72% total variance of the de-regressed EBVs (π = 0.99).

Based on the significance level determined by the posterior distribution of each SNP window under the null hypothesis that the window did not harbor QTL, SNP windows having high genetic variance on the trait were identified as candidate chromosomal regions. In total, 1, 2, 1, 3 and 2 chromosome regions (1 Mb window) were identified for MY, FY, PY, FP and PP, respectively, with a *P*-value less than 0.2 (Table 4 – 12). The detailed results for all traits, including positions of the associated chromosomal regions, number of SNPs within each region, *P*-values, accounted variances of the de-regressed EBVs, genes and previously reported QTL located in the regions are presented in Table 4 – 13. There is extensive overlap between the associated regions from this study and previously reported QTL regions for similar traits; moreover there was one novel QTL newly reported for PP on BTA10 (Table 4 – 12, Table 4 – 13).

In this study, we only identified one highly significant QTL region ($P < 0.001$) associated with MY on BTA14 at 0.05 – 1.00 Mb, which explained 10.5% total variance based on the de-regressed EBVs of MY (Table 4 – 12, Table 4 – 13, and Figure 4 – 4A). This QTL region was also identified for FY ($P < 0.001$), PY ($P < 0.10$), FP ($P < 0.001$) and PP ($P < 0.001$), and explained 18.08%, 2.32%,

38.94% and 10.66% of variance among the de-regressed EBVs for each trait, respectively. This region contains the *DGAT1* gene, and has been reported many times in previous QTL mapping studies (GRISART *et al.* 2002; SUN *et al.* 2009; WELLER *et al.* 2003). DGAT catalyzes the final step in triglyceride synthesis and functional polymorphisms have been identified for *DGAT1* in multiple cattle populations (GRISART *et al.* 2002; GRISART *et al.* 2004; WINTER *et al.* 2002). Three additional genes (*RPL8*, *GPT* and *MAPK15*) out of the 40 genes within this region, were identified in our study as being functionally relevant for milk production traits based on previous functional studies. *MAPK15* may affect milk production through down-regulating transactivation of the glucocorticoid receptor since glucocorticoid is an important hormone in maintaining milking (SAELZLER *et al.* 2006). *RPL8* plays an important role in protein biosynthesis and has a strong effect on the translational activity of ribosomes (UHLEIN *et al.* 1998). *GPT* was found to be associated with insulin resistance, obesity and diabetes and is considered to be an indicator for diabetes and metabolic syndrome in adults (DUBERN *et al.* 2006; KIM *et al.* 2009; MOJIMINIYI *et al.* 2010).

In addition to the QTL on BTA14, another genomic region at 95.07 – 95.88 Mb on BTA5 was identified for FY which explains 2.6% of the variance among the de-regressed EBVs for FY (Figure 4 – 4B). This region coincides with a QTL peak at 96 cM (80.145 – 111.5 cM) for milk fat yield in Norwegian dairy cattle (OLSEN *et al.* 2002). Within this region, two genes, *GOLT1B* and *IAPP*, were selected as functional candidates for FY in our study. GOLT1B, a component for Golgi apparatus and endoplasmic reticulum, is involved in vesicle-

mediated endosome-Golgi transport (CONCHON *et al.* 1999). IAPP is thought to be involved in glucose homeostasis (OTTO-BUCZKOWSKA *et al.* 2008) and can inhibit glucose-stimulated insulin secretion (SOTY *et al.* 2011).

For FP, QTL regions were identified on BTA14 and BTA20 (Figure 4 – 4D). On BTA14, in addition to the QTL at 0.50 – 1.00 Mb, another QTL at 3.02 – 3.99 Mb was identified which explains 2.61% of the genetic variance of the trait. This QTL overlaps with many previously reported QTL (BOICHARD *et al.* 2003; KUHN *et al.* 2004; THALLER *et al.* 2003). However, no candidate gene was identified in this region and further investigations should be carried out. A QTL on BTA20 at 35.15 – 35.70 Mb was identified, which explains 1.28% of the variance of the de-regressed EBVs for FP. This region overlaps with several previously reported QTL for milk fat percentage (ARRANZ *et al.* 1998; ZHANG *et al.* 1998), and one candidate gene, *GHR*, was described (WATERS *et al.* 2011).

In this study, one novel QTL was identified for PP on BTA10 at 46.00 – 46.95 Mb, which explained 1.78% of the variance among the de-regressed EBVs of PP (Figure 4 – 4E). Three genes, *PPIB*, *SNX1* and *HERC1*, within this newly reported QTL with functions related to protein transport and secretory pathway, were selected as candidate genes. SNX1 was found to be involved in several stages of intracellular protein transport (CARLTON *et al.* 2004; HAFT *et al.* 1998). The endoplasmic reticulum protein PPIB functions in accelerating the protein folding and is an important part of the protein secretary pathway (CARONI *et al.* 1991). HERC1 is a large protein involved in intracellular membrane traffic

(GARCIA-GONZALO *et al.* 2003) and may function in protein modification and transport. The functionally relevant genes in associated regions for all analyzed traits are given in Table 4 – 12 and Table 4 – 13. The concordance of QTL regions among five milk production traits may result from the high genetic correlations among the traits.

In the present study, 1 – 3 associated SNP windows with significant genetic contributions were discovered for each trait using the Bayesian approach. The small number of regions suggests that larger sample sizes will be required for a better power to detect the associated genomic regions for these traits. The results agreed in most cases with QTL locations for similar traits from previous independent studies (Table 4 – 13). One novel QTL region on BTA10 for milk protein percentage was identified. Following the association study, nine potential functional candidate genes on three chromosomes were identified from the suggestive QTL regions for milk production traits, based on currently available annotations, primarily based on studies conducted in other species (Table 4 – 12 and Table 4 – 13). This study rediscovered one previously reported candidate gene *DGAT1*, and reported several novel functional candidate genes, including *RPL8*, *GPT* and *MAPK15* for MY, FY, PY, FP and PP, *GOLT1B* and *IAPP* for FY, *PPIB*, *SNX1* and *HERC1* for PP. These candidate genes have functions related to lipid and protein metabolism and intracellular molecular transport. Further studies of these novel candidate genes should be conducted to establish these associations and to identify the causal mutations.

4.3.3. Association Analyses using Different Methods

For dairy milk production traits, genome-wide significant SNPs were first identified using single marker LD regression and then genomic regions harboring variations affecting these traits were identified using the BayesB method (MEUWISSEN *et al.* 2001). The BayesB method was originally developed for predicting the genetic merits of animals, typically with the aim of accelerating the genetic improvement of animals through genomic selection. Single marker regression fits one SNP at a time as a fixed effect, while the BayesB method analyzes all SNPs simultaneously as random effects and allows different degrees of shrinkage for each marker. Theoretically, the Bayesian model has advantages over the single marker approach as it can address the problem of too many highly correlated SNPs in the model, analyze all SNPs simultaneously, and avoid the problems of model selection and multiple testing (BALDING 2006; BEAUMONT and RANNALA 2004). In addition, Bayesian methods can incorporate our prior knowledge on genetic control of phenotypic traits to coerce negligible QTL effects towards zero. However, a larger number of animals is required for Bayesian analysis than for single marker analysis since all markers are be fitted simultaneously in a Bayesian model to partition the total phenotypic variance of targeted traits. The parameters that need to be estimated are usually several times larger than the number of available phenotypic records. A significant level of $P < 0.2$ was used in our study to declare a QTL region as significantly associated with a trait. Future studies using larger sample sizes will be required for a better power

to detect significant QTL regions at higher significance levels with Bayesian analyses.

Regarding the problem of multiple testing with single marker analysis, FDR, which is the proportion of false positive test results among all positives, was applied to adjust the genome-wise significance levels. This adjustment may help to reduce the type-I error rate and provide appropriately conservative results in this large scale SNP association study. However, this correction can only solve the problem with multiple testing and does not account for the effects of other SNPs in the bovine genome on the estimation of the effect of a particular SNP. In our study, this correction resulted in 316 SNPs at the genome-wise 5% FDR; however, several highly correlated SNPs were still obtained, for example, clusters of large numbers of SNPs in 0 – 16 Mb regions on BTA14 across analyzed traits. Although these regions are much smaller than the previously reported QTL from interval mapping which span over 20 – 40 cM, these SNPs still result in large genomic regions which could potentially contain hundreds of genes.

In comparison, the non-overlapping 1Mb SNP sliding windows used in the Bayesian analysis could account for the high LD among neighboring SNPs in identifying the associated regions. Previous whole genome association studies using sliding windows showed better performance in detecting chromosomal regions than using single SNPs in the Bayesian analyses (FAN *et al.* 2011; ONTERU *et al.* 2011; SUN *et al.* 2011). In addition, the resulting 1 Mb SNP windows will facilitate the subsequent search for functional candidate genes. In

this study, we also applied a recently developed genomic selection method called BayesC (FERNANDO and GARRICK 2009) to derive the variance components before the association analyses using the BayesB method. The BayesC method was derived from the BayesB method and was found to be less sensitive to the given priors of genetic variance in comparison with the BayesB method (KIZILKAYA *et al.* 2010).

Compared with previous QTL mapping studies in Canadian Holstein cattle using single marker analysis, our study using the BovineSNP50 BeadChip (50K) assay gave a larger number of significant SNPs, even after adjusting for a 5% FDR. For instance, 31 SNPs for MY, 7 for FY, 22 for PY, were identified in a whole genome association study for milk production traits using 9,919 SNPs and 484 Canadian Holstein sires (DAETWYLER *et al.* 2008); while only a few genome-wise significant SNPs (< 10) were reported for milk production traits using 1,536 SNPs and 462 Canadian Holstein bulls (KOLBEHDARI *et al.* 2009). In addition, this study represents the first GWAS using Bayesian regression for milk production traits in Canadian Holstein cattle. The SNPs or regions from both methods overlapped with or were in close proximity to QTL cited in the literature (Table 4 – 9 and Table 4 – 13), and there was good agreement between the two methods (Table 4 – 14). For example, the highly significant QTL region on BTA14 at 0.05 – 1.00 Mb across all analyzed traits in the Bayesian analysis was also identified through all traits in the single marker analysis, thus the known candidate gene *DGAT1* was reconfirmed by both methods. Overlapping results from different methods may increase our confidence to declare new QTL or

candidate genes for target traits. Furthermore, both methods provided some novel associations for milk production traits in Canadian Holstein cattle, for example, the novel SNPs on BTA1 for FP and novel QTL region on BTA10 for PP. Novel findings together with the reproduced positive associations should facilitate population-specific genetic improvement programs in the Canadian Holstein cattle population.

4.4. Conclusions

Emerging genomic resources and bioinformatics tools enable systematic identification of genome-wide sequence variants and candidate genes associated with complex traits in cattle. The present analyses, using the BovineSNP50 BeadChip, identified several QTL regions and genes associated with dairy milk production traits in Canadian Holstein cattle. Single marker analyses presented strong evidence for the presence of 316 genome-wise significant (FDR $P < 0.05$) SNPs on 28 chromosomes with a large proportion of the associations on BTA 1, 5, 11 and 14; whereas the Bayesian analyses identified 5 QTL on four chromosomes. In addition to many associations coinciding with previously reported QTL, several novel QTL regions were identified in both the single marker and Bayesian analyses. Furthermore, many concordances of QTL regions between the two methods were found.

Following the association analyses, three gene networks among positional candidate genes were identified from single marker analyses, whereas nine functional candidate genes on three chromosomes were identified in the Bayesian

analyses. Both the gene networks and candidate genes are linked through studies primarily in other species to lipid metabolism, small molecular biochemistry and molecular transport, all of which might be related to phenotypic variation of milk yield and composition. Overall, novel SNPs, QTL regions and functional candidate genes identified in this study will contribute to a better understanding of the molecular mechanisms of milk synthesis and secretion in cattle and could be applied to improve the accuracy of genetic evaluation of milk production traits after verification studies in other populations.

4.5. Literature Cited

ALAIN, K., N. A. KARROW, C. THIBAULT, J. ST-PIERRE, M. LESSARD *et al.*, 2009 Osteopontin: an early innate immune marker of Escherichia coli mastitis harbors genetic polymorphisms with possible links with resistance to mastitis. BMC Genomics **10**.

ALTSHULER, D., M. J. DALY and E. S. LANDER, 2008 Genetic Mapping in Human Disease. Science **322:** 881-888.

ARRANZ, J. J., W. COPPIETERS, P. BERZI, N. CAMBISANO, B. GRISART *et al.*, 1998 A QTL affecting milk yield and composition maps to bovine chromosome 20: a confirmation. Animal Genetics **29:** 107-115.

ASHWELL, M. S., Y. DA, C. P. VAN TASSELL, P. M. VANRADEN, R. H. MILLER *et al.*, 1998 Detection of putative loci affecting milk production and composition, health, and type traits in a United States Holstein population. Journal of Dairy Science **81:** 3309-3314.

ASHWELL, M. S., D. W. HEYEN, T. S. SONSTEGARD, C. P. VAN TASSELL, Y. DA *et al.*, 2004 Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. Journal of Dairy Science **87:** 468-475.

ASHWELL, M. S., C. P. VAN TASSELL and T. S. SONSTEGARD, 2001 A genome scan to identify quantitative trait loci affecting economically important traits in a US Holstein population. Journal of Dairy Science **84:** 2535-2542.

BAGNATO, A., F. SCHIAVINI, A. ROSSONI, C. MALTECCA, M. DOLEZAL *et al.*, 2008 Quantitative trait loci affecting milk yield and protein percentage in a three-country Brown Swiss population. Journal of Dairy Science **91:** 767-783.

BALDING, D. J., 2006 A tutorial on statistical methods for population association studies. Nature Reviews Genetics **7:** 781-791.

BEAUMONT, M. A., and B. RANNALA, 2004 The Bayesian revolution in genetics. Nature Reviews Genetics **5:** 251-261.

BENJAMINI, Y., and Y. HOCHBERG, 1995 Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological **57:** 289-300.

BENNEWITZ, J., N. REINSCH, C. GROHS, H. LEVEZIEL, A. MALAFOSSE *et al.*, 2003 Combined analysis of data from two granddaughter designs: A simple strategy for QTL confirmation and increasing experimental power in dairy cattle. Genetics Selection Evolution **35:** 319-338.

BENNEWITZ, J., N. REINSCH, V. GUIARD, S. FRITZ, H. THOMSEN *et al.*, 2004 Multiple quantitative trait loci mapping with cofactors and application of alternative variants of the false discovery rate in an enlarged granddaughter design. Genetics **168:** 1019-1027.

BLOTT, S., J. J. KIM, S. MOISIO, A. SCHMIDT-KUNTZEL, A. CORNET *et al.*, 2003 Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics **163:** 253-266.

BOICHARD, D., C. GROHS, F. BOURGEOIS, F. CERQUEIRA, R. FAUGERAS *et al.*, 2003 Detection of genes influencing economic traits in three French dairy cattle breeds. Genetics Selection Evolution **35:** 77-101.

BOLORMAA, S., J. E. PRYCE, B. J. HAYES and M. E. GODDARD, 2010 Multivariate analysis of a genome-wide association study in dairy cattle. Journal of Dairy Science **93:** 3818-3833.

BONAKDAR, E., H. R. RAHMANI, M. A. EDRISS and B. E. S. TABATABAEI, 2010 IGF-I gene polymorphism, but not its blood concentration, is associated with milk fat and protein in Holstein dairy cows. Genetics and Molecular Research **9:** 1726-1734.

BOVENHUIS, H., and J. I. WELLER, 1994 Mapping and Analysis of Dairy-Cattle Quantitative Trait Loci by Maximum-Likelihood Methodology Using Milk Protein Genes as Genetic-Markers. Genetics **137:** 267-280.

CARLTON, J., M. BUJNY, B. J. PETER, V. M. J. OORSCHOT, A. RUTHERFORD *et al.*, 2004 Sorting nexin-1 mediates tubular endosome-to-TGN transport through coincidence sensing of high-curvature membranes and 3-phosphoinositides. Current Biology **14:** 1791-1800.

CARONI, P., A. ROTHENFLUH, E. MCGLYNN and C. SCHNEIDER, 1991 S-Cyclophilin - New Member of the Cyclophilin Family Associated with the Secretory Pathway. Journal of Biological Chemistry **266:** 10739-10742.

CHARLIER, C., W. COPPIETERS, F. ROLLIN, D. DESMECHT, J. S. AGERHOLM *et al.*, 2008 Highly effective SNP-based association mapping and management of recessive defects in livestock. Nature Genetics **40:** 449-454.

CHEN, H. Y., Q. ZHANG, C. C. YIN, C. K. WANG, W. J. GONG *et al.*, 2006 Detection of quantitative trait loci affecting milk production traits on bovine chromosome 6 in a Chinese Holstein population by the daughter design. Journal of Dairy Science **89:** 782-790.

COHEN-ZINDER, M., E. SEROUSSI, D. M. LARKIN, J. J. LOOR, A. EVERTS-VAN DER WIND *et al.*, 2005 Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Research **15:** 936-944.

CONCHON, S., X. C. CAO, C. BARLOW and H. R. B. PELHAM, 1999 Got1p and Sft2p: membrane proteins involved in traffic to the Golgi complex. EMBO Journal **18:** 3934-3946.

DAETWYLER, H. D., F. S. SCHENKEL, M. SARGOLZAEI and J. A. B. ROBINSON, 2008 A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. Journal of Dairy Science **91:** 3225-3236.

DEKKERS, J. C. M., and F. HOSPITAL, 2002 The use of molecular genetics in the improvement of agricultural populations. Nature Reviews Genetics **3:** 22-32.

DUBERN, B., J. P. GIRARDET and P. TOUNIAN, 2006 Insulin resistance and ferritin as major determinants of abnormal serum aminotransferase in severely obese children. International Journal of Pediatric Obesity **1:** 77-82.

FAN, B., S. K. ONTERU, Z. Q. DU, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Genome-Wide Association Study Identifies Loci for Body Composition and Structural Soundness Traits in Pigs. PLoS One **6**.

FERNANDO, R. L., and D. J. GARRICK, 2009 GenSel - User manual for a portfolio of genomic selection related analyses. Animal Breeding and Genetics, Iowa State University, Ames.

FREYER, G., C. KUHN, R. WEIKARD, Q. ZHANG, M. MAYER *et al.*, 2002 Multiple QTL on chromosome six in dairy cattle affecting yield and content traits. Journal of Animal Breeding and Genetics **119:** 200-200.

GAO, H., M. FANG, J. LIU and Q. ZHANG, 2009 Bayesian shrinkage mapping for multiple QTL in half-sib families. Heredity **103:** 368-376.

GARCIA-GONZALO, F. R., C. CRUZ, P. MUNOZ, S. MAZUREK, E. EIGENBRODT *et al.*, 2003 Interaction between HERC1 and M2-type pyruvate kinase. FEBS Letters **539:** 78-84.

GARRICK, D. J., J. F. TAYLOR and R. L. FERNANDO, 2009 Deregressing estimated breeding values and weighting information for genomic regression analyses. Genetics Selection Evolution **41**.

GEORGES, M., D. NIELSEN, M. MACKINNON, A. MISHRA, R. OKIMOTO *et al.*, 1995 Mapping Quantitative Trait Loci Controlling Milk-Production in Dairy-Cattle by Exploiting Progeny Testing. Genetics **139:** 907-920.

GIBLIN, L., S. T. BUTLER, B. M. KEARNEY, S. M. WATERS, M. J. CALLANAN *et al.*, 2010 Association of bovine leptin polymorphisms with energy output and energy storage traits in progeny tested Holstein-Friesian dairy cattle sires. BMC Genetics **11**.

GILMOUR, A. R., GOGEL, B.J., CULLIS, B.R., AND THOMPSON, R., 2009 ASReml User Guide Release 3.0. VSN International Ltd **Hemel Hempstead, HP1 1ES, UK**.

GRISART, B., W. COPPIETERS, F. FARNIR, L. KARIM, C. FORD *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research **12:** 222-231.

GRISART, B., F. FARNIR, L. KARIM, N. CAMBISANO, J. J. KIM *et al.*, 2004 Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition.

Proceedings of the National Academy of Sciences of the United States of America **101:** 2398-2403.

HAFT, C. R., M. D. SIERRA, V. A. BARR, D. H. HAFT and S. I. TAYLOR, 1998 Identification of a family of sorting nexin molecules and characterization of their association with receptors. Molecular and Cellular Biology **18:** 7278-7287.

HARDER, B., J. BENNEWITZ, N. REINSCH, G. THALLER, H. THOMSEN *et al.*, 2006 Mapping of quantitative trait loci for lactation persistency traits in German Holstein dairy cattle. Journal of Animal Breeding and Genetics **123:** 89-96.

HAYES, B. J., P. J. BOWMAN, A. J. CHAMBERLAIN, K. SAVIN, C. P. VAN TASSELL *et al.*, 2009 A Validated Genome Wide Association Study to Breed Cattle Adapted to an Environment Altered by Climate Change. PLoS One **4**.

HECK, J. M. L., A. SCHENNINK, H. J. F. VAN VALENBERG, H. BOVENHUIS, M. H. P. W. VISKER *et al.*, 2009 Effects of milk protein variants on the protein composition of bovine milk. Journal of Dairy Science **92:** 1192-1202.

HEYEN, D. W., J. I. WELLER, M. RON, M. BAND, J. E. BEEVER *et al.*, 1999 A genome scan for QTL influencing milk production and health traits in dairy cattle. Physiological Genomics **1:** 165-175.

HOGLUND, J. K., A. J. BUITENHUIS, B. GULDBRANDTSEN, G. SU, B. THOMSEN *et al.*, 2009 Overlapping chromosomal regions for fertility traits and production traits in the Danish Holstein population. Journal of Dairy Science **92:** 5712-5719.

JIANG, L., J. F. LIU, D. X. SUN, P. P. MA, X. D. DING *et al.*, 2010 Genome Wide Association Studies for Milk Production Traits in Chinese Holstein Population. PLoS One **5**.

KIM, C. H., J. Y. PARK, K. U. LEE, J. H. KIM and H. K. KIM, 2009 Association of serum gamma-glutamyltransferase and alanine aminotransferase activities with risk of type 2 diabetes mellitus independent of fatty liver. Diabetes-Metabolism Research and Reviews **25:** 64-69.

KIZILKAYA, K., R. L. FERNANDO and D. J. GARRICK, 2010 Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. Journal of Animal Science **88:** 544-551.

KOLBEHDARI, D., Z. WANG, J. R. GRANT, B. MURDOCH, A. PRASAD *et al.*, 2009 A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. Journal of Animal Breeding and Genetics **126:** 216-227.

KUHN, C., G. THALLER, A. WINTER, O. R. P. BININDA-EMONDS, B. KAUPE *et al.*, 2004 Evidence for multiple alleles at the DGAT1 locus better explains a quantitative tip trait locus with major effect on milk fat content in cattle. Genetics **167:** 1873-1881.

LEYVA-BACA, I., F. SCHENKEL, B. S. SHARMA, G. B. JANSEN and N. A. KARROW, 2007 Identification of single nucleotide polymorphisms in the bovine CCL2, IL8, CCR2 and IL8RA genes and their association with health and production in Canadian Holsteins. Animal Genetics **38:** 198-202.

LILLEHAMMER, M., M. ARNYASI, S. LIEN, H. G. OLSEN, E. SEHESTED *et al.*, 2007 A genome scan for quantitative trait locus by environment interactions for production traits. Journal of Dairy Science **90:** 3482-3489.

LINDERSSON, M., L. ANDERSSON-EKLUND, D. J. DE KONING, A. LUNDEN, A. MAKI-TANILA *et al.*, 1998 Mapping of serum amylase-1 and quantitative trait loci for milk production traits to cattle chromosome 4. Journal of Dairy Science **81:** 1454-1461.

LIPKIN, E., M. O. MOSIG, A. DARVASI, E. EZRA, A. SHALOM *et al.*, 1998 Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: Analysis of milk protein percentage. Genetics **149:** 1557-1567.

LOOFT, C., N. REINSCH, C. KARALL-ALBRECHT, S. PAUL, M. BRINK *et al.*, 2001 A mammary gland EST showing linkage disequilibrium to a milk production QTL on bovine Chromosome 14. Mammalian Genome **12:** 646-650.

LUND, M. S., B. GULDBRANDTSEN, A. J. BUITENHUIS, B. THOMSEN and C. BENDIXEN, 2008 Detection of quantitative trait loci in Danish Holstein cattle affecting clinical mastitis, somatic cell score, udder conformation traits, and assessment of associated effects on milk yield. Journal of Dairy Science **91:** 4028-4036.

MAI, M. D., G. SAHANA, F. B. CHRISTIANSEN and B. GULDBRANDTSEN, 2010 A genome-wide association study for milk production traits in Danish Jersey cattle using a 50K single nucleotide polymorphism chip. Journal of Animal Science **88:** 3522-3528.

MANOLIO, T. A., F. S. COLLINS, N. J. COX, D. B. GOLDSTEIN, L. A. HINDORFF *et al.*, 2009 Finding the missing heritability of complex diseases. Nature **461:** 747-753.

MEI, G., C. C. YIN, X. D. DING and Q. ZHANG, 2009 Fine mapping quantitative trait loci affecting milk production traits on bovine chromosome 6 in a Chinese Holstein population. Journal of Genetics and Genomics **36:** 653-660.

MEUWISSEN, T. H. E., B. J. HAYES and M. E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157:** 1819-1829.

MOJIMINIYI, O. A., N. A. ABDELLA and H. AL MOHAMMEDI, 2010 Higher Levels of Alanine Aminotransferase Within the Reference Range Predict Unhealthy Metabolic Phenotypes of Obesity in Normoglycemic First-Degree Relatives of Patients With Type 2 Diabetes Mellitus. Journal of Clinical Hypertension **12:** 301-308.

MOSIG, M. O., E. LIPKIN, G. KHUTORESKAYA, E. TCHOURZYNA, M. SOLLER *et al.*, 2001 A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion. Genetics **157:** 1683-1698.

NADESALINGAM, J., Y. PLANTE and J. P. GIBSON, 2001 Detection of QTL for milk production on Chromosomes 1 and 6 of Holstein cattle. Mammalian Genome **12:** 27-31.

NASLUND, J., W. F. FIKSE, G. R. PIELBERG and A. LUNDEN, 2008 Frequency and effect of the bovine acyl-CoA : diacylglycerol acyltransferase 1 (DGAT1) K232A polymorphism in Swedish dairy cattle. Journal of Dairy Science **91:** 2127-2134.

OLSEN, H. G., L. GOMEZ-RAYA, D. I. VAGE, I. OLSAKER, H. KLUNGLAND *et al.*, 2002 A genome scan for quantitative trait loci affecting milk production in Norwegian dairy cattle. Journal of Dairy Science **85:** 3124-3130.

OLSEN, H. G., S. LIEN, M. SVENDSEN, H. NILSEN, A. ROSETH *et al.*, 2004 Fine mapping of milk production QTL on BTA6 by combined linkage and linkage disequilibrium analysis. Journal of Dairy Science **87:** 690-698.

OLSEN, H. G., H. NILSEN, B. HAYES, P. R. BERG, M. SVENDSEN *et al.*, 2007 Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. BMC Genetics **8**.

ONTERU, S. K., B. FAN, M. T. NIKKILA, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Whole-genome association analyses for lifetime reproductive traits in the pig. Journal of Animal Science **89:** 988-995.

OSTERSEN, T., O. F. CHRISTENSEN, M. HENRYON, B. NIELSEN, G. S. SU *et al.*, 2011 Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in pure-bred pigs. Genetics Selection Evolution **43**.

OTTO-BUCZKOWSKA, E., U. MAZUR-DWORZECKA and T. DWORZECKI, 2008 Role of amylin in glucose homeostasis and its perspective use in diabetes management. Przeglad Lekarski **65:** 135-139.

PLANTE, Y., J. P. GIBSON, J. NADESALINGAM, H. MEHRABANI-YEGANEH, S. LEFEBVRE *et al.*, 2001 Detection of quantitative trait loci affecting milk production traits on 10 chromosomes in Holstein cattle. Journal of Dairy Science **84:** 1516-1524.

PRYCE, J. E., S. BOLORMAA, A. J. CHAMBERLAIN, P. J. BOWMAN, K. SAVIN *et al.*, 2010 A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. Journal of Dairy Science **93:** 3331-3345.

REARDON, W., A. M. MULLEN, T. SWEENEY and R. M. HAMILL, 2010 Association of polymorphisms in candidate genes with colour, water-holding capacity, and composition traits in bovine M. longissimus and M. semimembranosus. Meat Science **86:** 270-275.

RODRIGUEZ-ZAS, S. L., B. R. SOUTHEY, D. W. HEYEN and H. A. LEWIN, 2002 Detection of quantitative trait loci influencing dairy traits using a model for longitudinal data. Journal of Dairy Science **85:** 2681-2691.

SAELZLER, M. P., C. C. SPACKMAN, Y. LIU, L. C. MARTINEZ, J. P. HARRIS *et al.*, 2006 ERK8 down-regulates transactivation of the glucocorticoid receptor through Hic-5. Journal of Biological Chemistry **281:** 16821-16832.

SCHAEFFER, L. R., 1994 Multiple-Country Comparison of Dairy Sires. Journal of Dairy Science **77:** 2671-2678.

SCHENNINK, A., H. BOVENHUIS, K. M. LEON-KLOOSTERZIEL, J. A. M. VAN ARENDONK and M. H. P. W. VISKER, 2009 Effect of polymorphisms in the

FASN, OLR1, PPARGC1A, PRL and STAT5A genes on bovine milk-fat composition. Animal Genetics **40:** 909-916.

SCHNABEL, R. D., T. S. SONSTEGARD, J. F. TAYLOR and M. S. ASHWELL, 2005 Whole-genome scan to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families. Animal Genetics **36:** 408-416.

SCHROOTEN, C., M. C. A. M. BINK and H. BOVENHUIS, 2004 Whole genome scan to detect chromosomal regions affecting multiple traits in dairy cattle. Journal of Dairy Science **87:** 3550-3560.

SCHULMAN, N. F., G. SAHANA, M. S. LUND, S. M. VIITALA and J. H. VILKKI, 2008 Quantitative trait loci for fertility traits in Finnish Ayrshire cattle. Genetics Selection Evolution **40:** 195-214.

SELLNER, E. M., J. W. KIM, M. C. MCCLURE, K. H. TAYLOR, R. D. SCHNABEL *et al.*, 2007 Board-invited review: Applications of genomic information in livestock. Journal of Animal Science **85:** 3148-3158.

SOTY, M., M. VISA, S. SORIANO, C. CARMONA MDEL, A. NADAL *et al.*, 2011 Involvement of ATP-sensitive potassium (K(ATP)) channels in the loss of beta-cell function induced by human islet amyloid polypeptide. Journal of Biological Chemistry **286:** 40857-40866.

SPELMAN, R. J., W. COPPIETERS, L. KARIM, J. A. M. VANARENDONK and H. BOVENHUIS, 1996 Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population. Genetics **144:** 1799-1807.

SUN, D., J. JIA, Y. MA, Y. ZHANG, Y. WANG *et al.*, 2009 Effects of DGAT1 and GHR on milk yield and milk composition in the Chinese dairy population. Animal Genetics **40:** 997-1000.

SUN, X., D. HABIER, R. L. FERNANDO, D. J. GARRICK and J. C. DEKKERS, 2011 Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. BMC Proceedings **5 Suppl 3:** S13.

THALLER, G., W. KRAMER, A. WINTER, B. KAUPE, G. ERHARDT *et al.*, 2003 Effects of DGAT1 variants on milk production traits in German cattle breeds. Journal of Animal Science **81:** 1911-1918.

UHLEIN, M., W. WEGLOHNER, H. URLAUB and B. WITTMANN-LIEBOLD, 1998 Functional implications of ribosomal protein L2 in protein biosynthesis as shown by in vivo replacement studies. Biochemical Journal **331:** 423-430.

VELMALA, R. J., H. J. VILKKI, K. T. ELO, D. J. DE KONING and A. V. MAKI-TANILA, 1999 A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle. Animal Genetics **30:** 136-143.

VIITALA, S. M., N. F. SCHULMAN, D. J. DE KONING, K. ELO, R. KINOS *et al.*, 2003 Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle. Journal of Dairy Science **86:** 1828-1836.

WANG, Z. Y., G. L. WANG, J. M. HUANG, Q. L. LI, C. F. WANG *et al.*, 2011 Novel SNPs in the ATP1B2 gene and their associations with milk yield, milk composition and heat-resistance traits in Chinese Holstein cows. Molecular Biology Reports **38:** 1749-1755.

WATERS, S. M., M. S. MCCABE, D. J. HOWARD, L. GIBLIN, D. A. MAGEE *et al.*, 2011 Associations between newly discovered polymorphisms in the Bos taurusgrowth hormone receptor gene and performance traits in Holstein-Friesian dairy cattle. Animal Genetics **42:** 39-49.

WELLER, J. I., M. GOLIK, E. SEROUSSI, E. EZRA and M. RON, 2003 Population-wide analysis of a QTL affecting milk-fat production in the Israeli Holstein population. Journal of Dairy Science **86:** 2219-2227.

WINTER, A., W. KRAMER, F. A. O. WERNER, S. KOLLERS, S. KATA *et al.*, 2002 Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA : diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proceedings of the National Academy of Sciences of the United States of America **99:** 9300-9305.

YU, J. M., G. PRESSOIR, W. H. BRIGGS, I. V. BI, M. YAMASAKI *et al.*, 2006 A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics **38:** 203-208.

ZHANG, Q., D. BOICHARD, I. HOESCHELE, C. ERNST, A. EGGEN *et al.*, 1998 Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics **149:** 1959-1973.

ZHAO, H. H., R. L. FERNANDO and J. C. M. DEKKERS, 2007 Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. Genetics **175:** 1975-1986.

Tables

Table 4 – 1. Descriptive statistics of de-regressed EBVs for five milk production traits examined in the Canadian Holstein cattle population.

Trait[1]	N	Mean	SD	Min	Max	Heritability[2]	Reliability[3]
MY	647	588.64	839.14	-1609.30	2962.11	0.41	91.62
FY	647	15.57	31.11	-85.29	101.13	0.34	91.62
PY	647	20.96	23.14	-48.96	87.92	0.37	91.62
FP	647	-0.052	0.32	-0.73	0.91	0.50	91.62
PP	647	0.019	0.13	-0.37	0.52	0.50	91.62

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); PP = protein percentage (%). [2]Heritability values used by the Canadian Dairy Network (CDN) in the genetic evaluation. The values were literature averages for FP and PP. [3]The CDN used PY reliability for all milk production traits.

Table 4 – 2. Number of SNPs, average intermarker distance (kb), average heterozygosity and average MAF of SNPs on 29 autosomes in Canadian Holstein cattle.

BTA	N-SNPs	Intermarker spacing ± SD	Heterozygosity ± SD	MAF ± SD
1	1,921	83.81±78.64	0.397±0.099	0.298±0.116
2	1,554	90.4±97.72	0.401±0.098	0.304±0.115
3	1,496	85±89.92	0.406±0.096	0.307±0.115
4	1,430	86.71±86.51	0.403±0.098	0.304±0.115
5	1,209	104.14±113.35	0.398±0.099	0.299±0.115
6	1,475	83.02±83.13	0.4±0.1	0.297±0.115
7	1,283	87.09±89.17	0.4±0.103	0.302±0.117
8	1,373	85.21±81.98	0.402±0.096	0.301±0.113
9	1,122	96.03±96.22	0.402±0.096	0.3±0.112
10	1,253	84.31±92.61	0.395±0.1	0.293±0.113
11	1,309	84.09±83.24	0.395±0.095	0.299±0.114
12	945	90.22±93.15	0.402±0.097	0.304±0.115
13	1,000	83.89±82.7	0.409±0.096	0.31±0.114
14	1,033	78.75±74.1	0.416±0.095	0.314±0.112
15	984	85.68±80.55	0.409±0.097	0.306±0.113
16	887	87.74±95.75	0.406±0.099	0.303±0.113
17	907	84.27±83.57	0.4±0.097	0.3±0.114
18	825	80.14±90.05	0.408±0.095	0.308±0.112
19	829	78.54±67.39	0.399±0.098	0.303±0.117
20	860	87.79±98.61	0.394±0.097	0.296±0.117
21	791	87.52±89.62	0.402±0.096	0.306±0.115
22	730	84.46±77.07	0.401±0.099	0.302±0.116
23	645	82.55±83.32	0.413±0.096	0.314±0.114
24	713	91.07±86.73	0.404±0.096	0.307±0.115
25	622	69.79±64.46	0.41±0.096	0.314±0.116
26	608	83.81±71.36	0.402±0.098	0.302±0.116
27	550	88.63±109.39	0.396±0.099	0.3±0.118
28	579	79.5±70.23	0.406±0.099	0.307±0.116
29	619	83.59±95.75	0.408±0.095	0.31±0.113
Overall	29,552	85.91±87.71	0.402±0.098	0.303±0.115

Table 4 – 3. Genome-wise significant SNPs for milk yield (MY) using single marker LD regression in the Canadian Holstein cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-NGS-38890	1	57,143,016	0.20	0.32	-253.3 ±57.38	19.48	1.19E-05**
HAPMAP44594-BTA-25389	1	57,702,386	0.22	0.36	-231.2 ±58.18	15.79	7.88E-05*
ARS-BFGL-BAC-7205	1	120,983,739	0.40	0.52	222.4 ±50.06	19.73	1.05E-05**
BFGL-NGS-110032	1	121,119,261	0.20	0.33	-241.7 ±60.56	15.93	7.33E-05*
ARS-BFGL-NGS-553	1	125,411,899	0.35	0.47	-239.3 ±49.81	23.08	1.94E-06**
ARS-BFGL-NGS-37759	1	132,886,245	0.41	0.48	183.5 ±47.76	14.75	0.000135*
HAPMAP48798-BTA-51401	1	133,390,970	0.32	0.43	217.4 ±51.15	18.06	2.46E-05*
ARS-BFGL-NGS-93090	1	138,007,561	0.26	0.37	228.3 ±55.56	16.89	4.48E-05*
ARS-BFGL-NGS-105044	1	147,597,144	0.13	0.23	-322.4 ±70.63	20.83	6.01E-06**
ARS-BFGL-NGS-6826	5	14,793,744	0.47	0.50	-183.4 ±48.01	14.59	0.000147*
ARS-BFGL-NGS-57681	5	114,710,762	0.31	0.43	-190.7 ±49.76	14.69	0.000139*
BTA-74965-NO-RS	5	114,799,035	0.13	0.21	-305.4 ±68.18	20.06	8.88E-06**
HAPMAP25014-BTA-123017	5	119,734,177	0.11	0.19	-331.9 ±75.93	19.11	1.44E-05**
ARS-BFGL-NGS-109212	5	120,674,018	0.12	0.20	-331.5 ±72.15	21.11	5.22E-06**
ARS-BFGL-NGS-104684	7	17,903,848	0.20	0.35	238.4 ±60.17	15.7	8.25E-05*
BTA-78475-NO-RS	7	17,929,919	0.16	0.29	255.4 ±64.88	15.5	9.15E-05*
BTB-01895465	9	93,903,479	0.26	0.39	-210.3 ±53.89	15.24	0.000105*
HAPMAP47338-BTA-59326	10	15,172,465	0.44	0.49	209.5 ±48.4	18.73	1.75E-05*
ARS-BFGL-NGS-43296	10	34,965,551	0.29	0.43	-203.9 ±50.76	16.14	6.58E-05*
BFGL-NGS-115446	11	12,763,718	0.25	0.36	-234.7 ±55.54	17.86	2.75E-05*
ARS-BFGL-NGS-40307	11	12,784,299	0.25	0.36	-234.5 ±55.5	17.86	2.74E-05*
ARS-BFGL-NGS-40347	11	13,334,124	0.21	0.31	-239.6 ±58.34	16.87	4.55E-05*
HAPMAP15326-RS29013300	11	41,385,158	0.49	0.53	189.7 ±48.84	15.08	0.000114*
ARS-BFGL-NGS-39065	11	105,482,391	0.21	0.32	-230.2 ±55.31	17.33	3.57E-05*
ARS-BFGL-NGS-80488	12	14,768,766	0.25	0.41	217 ±54.53	15.84	7.68E-05*
ARS-BFGL-NGS-105816	12	14,796,693	0.17	0.30	273.1 ±62.57	19.05	1.48E-05**
HAPMAP30383-BTC-005848	14	76,703	0.35	0.50	-319.2 ±49.61	41.4	2.42E-10***
BTA-34956-NO-RS	14	101,473	0.48	0.55	221.1 ±47.62	21.56	4.16E-06**
ARS-BFGL-NGS-57820	14	236,532	0.24	0.38	-462 ±52.4	77.73	1.10E-17***
ARS-BFGL-NGS-34135	14	260,341	0.41	0.51	-274.2 ±46.89	34.19	7.97E-09***
ARS-BFGL-NGS-94706	14	281,533	0.41	0.51	-284.6 ±46.62	37.26	1.79E-09***
ARS-BFGL-NGS-4939	14	443,937	0.24	0.38	-458.8 ±53.2	74.35	5.06E-17***
ARS-BFGL-NGS-71749	14	596,341	0.28	0.43	220.7 ±50.47	19.12	1.45E-05**
ARS-BFGL-NGS-107379	14	679,600	0.28	0.44	-375.9 ±51.53	53.22	8.80E-13***
HAPMAP25384-BTC-001997	14	835,054	0.44	0.51	-201.7 ±47.71	17.88	2.69E-05*
HAPMAP24715-BTC-001973	14	856,889	0.44	0.51	-201.1 ±47.67	17.8	2.81E-05*
BTA-35941-NO-RS	14	894,252	0.42	0.53	-222 ±47.65	21.7	3.88E-06**
ARS-BFGL-NGS-26520	14	996,982	0.36	0.48	257.9 ±46.95	30.17	5.75E-08***
ARS-BFGL-NGS-22866	14	1,131,952	0.48	0.51	182.8 ±46.52	15.44	9.44E-05*
HAPMAP29758-BTC-003619	14	1,339,276	0.49	0.53	-204.5 ±45.75	19.99	9.20E-06**
HAPMAP30646-BTC-002054	14	1,461,085	0.44	0.53	263.9 ±46.24	32.58	1.76E-08***
HAPMAP30086-BTC-002066	14	1,490,178	0.50	0.54	273.6 ±45.22	36.6	2.49E-09***
HAPMAP30374-BTC-002159	14	1,546,591	0.42	0.53	-214.5 ±47.09	20.76	6.24E-06**
HAPMAP32970-BTC-064990	14	2,288,510	0.33	0.46	-207.5 ±51.55	16.2	6.38E-05*
HAPMAP24986-BTC-065021	14	2,313,595	0.33	0.46	-207.5 ±51.55	16.2	6.38E-05*
ARS-BFGL-NGS-22111	14	2,347,219	0.38	0.49	227.6 ±48.52	22	3.34E-06**
UA-IFASA-7269	14	2,370,256	0.38	0.49	227.6 ±48.52	22	3.34E-06**
HAPMAP26072-BTC-065132	14	2,391,826	0.38	0.51	193.9 ±48.36	16.08	6.8E-05*
ARS-BFGL-NGS-100480	14	2,607,583	0.39	0.51	-266.7 ±48.86	29.79	6.88E-08***
HAPMAP27703-BTC-053907	14	2,826,073	0.38	0.50	250.3 ±48.97	26.12	4.25E-07***
HAPMAP23302-BTC-052123	14	3,099,635	0.27	0.41	-211.2 ±54.8	14.85	0.000128*
HAPMAP25217-BTC-067767	14	3,189,312	0.40	0.51	282.6 ±48.23	34.34	7.40E-09***
UA-IFASA-6329	14	3,465,237	0.50	0.54	198.2 ±47.91	17.12	3.97E-05*
ARS-BFGL-NGS-3571	14	3,587,018	0.44	0.51	-277.7 ±46.87	35.1	5.10E-09***
BFGL-NGS-110563	14	3,799,228	0.46	0.52	-285.3 ±47.12	36.65	2.40E-09***
HAPMAP32262-BTC-066621	14	3,834,069	0.35	0.48	235.8 ±49.66	22.55	2.53E-06**
BFGL-NGS-115947	14	3,865,962	0.39	0.50	-283 ±48.93	33.46	1.14E-08***

HAPMAP30091-BTC-005211	14	3,940,998	0.37	0.50	259 ±49.43	27.46	2.18E-07***
ARS-BFGL-BAC-24804	14	4,157,675	0.27	0.40	-221.9 ±52.9	17.6	3.11E-05*
HAPMAP23454-BTC-046932	14	4,182,816	0.27	0.39	-217.9 ±52.77	17.05	4.12E-05*
HAPMAP51646-BTA-86764	14	4,302,229	0.45	0.55	213.9 ±49.18	18.92	1.58E-05**
HAPMAP26591-BTC-056596	14	4,477,036	0.47	0.52	217.5 ±47.54	20.93	5.72E-06**
HAPMAP30988-BTC-056315	14	4,693,901	0.38	0.50	212.3 ±49.62	18.31	2.16E-05*
BFGL-NGS-112858	14	4,956,375	0.27	0.40	-238.3 ±53.06	20.17	8.40E-06**
ARS-BFGL-NGS-55227	14	5,085,416	0.27	0.41	-206.3 ±53.52	14.85	0.000128*
BFGL-NGS-113706	14	5,117,434	0.40	0.51	-239.8 ±49.9	23.09	1.93E-06**
HAPMAP32236-BTC-049785	14	5,139,498	0.41	0.50	-245.8 ±48.8	25.37	6.16E-07***
UA-IFASA-6228	14	5,204,594	0.37	0.50	-267.6 ±50.53	28.05	1.63E-07***
BFGL-NGS-110894	14	5,282,438	0.34	0.45	-288 ±50.24	32.85	1.53E-08***
HAPMAP32234-BTC-048199	14	5,640,338	0.35	0.48	-273.9 ±48.38	32.06	2.25E-08***
HAPMAP26283-BTC-048098	14	5,696,729	0.30	0.43	-225.9 ±51.85	18.98	1.54E-05**
UA-IFASA-6647	14	5,808,644	0.38	0.49	-233.9 ±47.89	23.86	1.31E-06**
ARS-BFGL-NGS-102953	14	5,867,266	0.44	0.51	-261.2 ±47	30.89	4.00E-08***
ARS-BFGL-BAC-20850	14	7,928,144	0.46	0.49	-253.8 ±48.69	27.18	2.51E-07***
HAPMAP30097-BTC-007678	14	7,969,429	0.48	0.48	232.5 ±47.22	24.25	1.08E-06**
HAPMAP31564-BTC-007633	14	7,998,736	0.45	0.48	-244.1 ±48.08	25.77	5.07E-07***
UA-IFASA-5356	14	8,132,747	0.42	0.49	-255.8 ±47.83	28.6	1.24E-07***
ARS-BFGL-BAC-20261	14	8,656,676	0.40	0.48	-189.5 ±48.83	15.06	0.000115*
HAPMAP57409-RS29021898	14	11,524,613	0.30	0.42	-211.2 ±51.47	16.83	4.61E-05*
ARS-BFGL-NGS-8221	14	11,720,820	0.47	0.50	224.3 ±47.25	22.52	2.56E-06**
ARS-BFGL-NGS-24160	14	11,963,066	0.40	0.48	219.3 ±47.89	20.96	5.63E-06**
ARS-BFGL-NGS-105600	14	11,985,275	0.50	0.49	197.4 ±47.15	17.52	3.24E-05*
HAPMAP46741-BTA-00625	14	12,063,432	0.49	0.50	-195.9 ±47.2	17.23	3.76E-05*
UA-IFASA-7696	14	12,380,364	0.38	0.48	193.6 ±48.81	15.74	8.08E-05*
ARS-BFGL-NGS-63270	14	14,365,665	0.30	0.45	193.3 ±49.71	15.12	0.000111*
HAPMAP38314-BTA-42171	14	15,515,272	0.47	0.49	-195.3 ±47.54	16.88	4.5E-05*
ARS-BFGL-NGS-105690	17	18,690,385	0.19	0.31	-244.4 ±58.59	17.4	3.45E-05*
ARS-BFGL-NGS-8684	18	53,442,280	0.21	0.33	-216.8 ±56.86	14.53	0.000151*
ARS-BFGL-NGS-104611	19	37,377,576	0.21	0.33	-234.1 ±57.78	16.42	5.69E-05*
HAPMAP59410-SS46526044	21	30,726,309	0.20	0.30	243.1 ±57.38	17.95	2.6E-05*
BFGL-NGS-111976	27	19,052,634	0.12	0.20	-264.3 ±69.37	14.52	0.000152*
HAPMAP40395-BTA-113989	28	103,000	0.11	0.19	284.1 ±72.98	15.16	0.000109*
BFGL-NGS-118787	28	317,989	0.18	0.28	252.8 ±62.97	16.12	6.66E-05*
ARS-BFGL-NGS-72206	28	1,970,800	0.10	0.18	299.7 ±74.64	16.13	6.61E-05*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip. *Significant at genome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.01$; ***Significant at genome-wise FDR $P < 0.001$.

Table 4 – 4. Genome-wise significant SNPs for fat yield (FY) using single marker LD regression in the Canadian Holstein cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
HAPMAP41782-BTA-16216	1	16,276,347	0.28	0.41	-9.77 ±2.03	23.24	1.81E-06**
ARS-BFGL-NGS-43783	1	54,006,178	0.38	0.47	-6.97 ±1.82	14.62	0.000144*
BTA-97386-NO-RS	2	5,156,155	0.13	0.23	9.91 ±2.59	14.68	0.00014*
HAPMAP51953-BTA-48787	2	101,230,933	0.30	0.40	-7.5 ±1.9	15.59	8.74E-05*
BFGL-NGS-118243	3	111,260,250	0.34	0.47	7.92 ±1.92	17.05	4.12E-05*
BFGL-NGS-117787	3	112,164,376	0.38	0.50	7.63 ±1.85	17.05	4.12E-05*
HAPMAP46581-BTA-75566	5	14,416,892	0.22	0.34	-8.8 ±2.13	17.11	4.02E-05*
ARS-BFGL-NGS-1161	5	14,500,783	0.22	0.34	-8.65 ±2.12	16.6	5.21E-05*
HAPMAP43345-BTA-74211	5	86,720,671	0.30	0.43	-7.44 ±1.95	14.61	0.000145*
ARS-BFGL-NGS-17920	5	91,062,497	0.40	0.49	8.57 ±1.87	20.9	5.84E-06**
BFGL-NGS-111850	5	92,502,082	0.30	0.43	7.7 ±2	14.75	0.000135*
ARS-BFGL-NGS-15506	5	94,284,625	0.45	0.51	6.84 ±1.73	15.61	8.64E-05*
BTA-98453-NO-RS	5	94,386,404	0.41	0.48	-7.03 ±1.78	15.5	9.17E-05*
BTB-01267042	5	95,595,198	0.32	0.43	9.06 ±1.97	21.23	4.96E-06**
HAPMAP23365-BTA-156277	5	97,370,232	0.21	0.34	-10.14 ±2.15	22.23	2.98E-06**
HAPMAP60862-RS29018508	5	97,572,284	0.18	0.29	-10.78 ±2.32	21.61	4.09E-06**
BTA-15560-NO-RS	5	98,456,165	0.28	0.40	-8.83 ±1.98	19.93	9.50E-06**
HAPMAP33512-BTA-158274	5	98,624,100	0.27	0.38	-7.59 ±1.99	14.58	0.000148*
ARS-BFGL-NGS-106674	5	112,671,018	0.18	0.30	8.92 ±2.29	15.15	0.00011*
HAPMAP51409-BTA-122717	6	90,356,013	0.30	0.42	-7.67 ±1.93	15.82	7.79E-05*
BTB-01880776	7	64,095,705	0.26	0.41	-7.91 ±2.01	15.44	9.44E-05*
BTA-12616-NO-RS	7	64,712,172	0.49	0.47	-6.43 ±1.69	14.41	0.000161*
ARS-BFGL-NGS-107513	7	72,719,276	0.34	0.47	-8.19 ±1.85	19.72	1.06E-05**
BTB-01843749	9	36,578,501	0.35	0.49	7.87 ±1.98	15.84	7.69E-05*
BTB-01346626	9	58,105,834	0.19	0.33	9.44 ±2.4	15.42	9.58E-05*
BTB-01350179	12	57,912,194	0.35	0.44	7.26 ±1.87	15.09	0.000114*
BFGL-NGS-113128	12	58,312,167	0.48	0.53	7.32 ±1.8	16.48	5.53E-05*
HAPMAP30381-BTC-005750	14	50,872	0.35	0.48	-10.75 ±1.72	39.15	7.26E-10***
HAPMAP30383-BTC-005848	14	76,703	0.35	0.50	12.09 ±1.9	40.68	3.43E-10***
BTA-34956-NO-RS	14	101,473	0.48	0.55	-8.15 ±1.82	19.94	9.44E-06**
ARS-BFGL-NGS-57820	14	236,532	0.24	0.38	17.62 ±2	77.82	1.07E-17***
ARS-BFGL-NGS-34135	14	260,341	0.41	0.51	11.57 ±1.78	42.09	1.74E-10***
ARS-BFGL-NGS-94706	14	281,533	0.41	0.51	10.32 ±1.79	33.25	1.26E-08***
ARS-BFGL-NGS-4939	14	443,937	0.24	0.38	19.49 ±2	95.16	4.76E-21***
ARS-BFGL-NGS-71749	14	596,341	0.28	0.43	-8.81 ±1.95	20.49	7.16E-06**
ARS-BFGL-NGS-107379	14	679,600	0.28	0.44	16.27 ±1.94	70.05	3.62E-16***
HAPMAP25384-BTC-001997	14	835,054	0.44	0.51	10.13 ±1.8	31.68	2.72E-08***
HAPMAP24715-BTC-001973	14	856,889	0.44	0.51	10.16 ±1.8	31.9	2.44E-08***
BTA-35941-NO-RS	14	894,252	0.42	0.53	11.32 ±1.8	39.65	5.62E-10***
ARS-BFGL-NGS-26520	14	996,982	0.36	0.48	-10.18 ±1.8	31.99	2.33E-08***
ARS-BFGL-NGS-22866	14	1,131,952	0.48	0.51	-7.36 ±1.78	17.15	3.92E-05*
ARS-BFGL-NGS-3122	14	1,264,233	0.48	0.53	-7.92 ±1.82	19.02	1.51E-05**
HAPMAP29758-BTC-003619	14	1,339,276	0.49	0.53	7.68 ±1.75	19.26	1.33E-05**
HAPMAP30646-BTC-002054	14	1,461,085	0.44	0.53	-11.63 ±1.76	43.6	8.43E-11***
HAPMAP30086-BTC-002066	14	1,490,178	0.50	0.54	-13.04 ±1.71	58.22	8.48E-14***
HAPMAP30374-BTC-002159	14	1,546,591	0.42	0.53	11.67 ±1.77	43.29	9.77E-11***
ARS-BFGL-NGS-74378	14	1,889,210	0.26	0.40	9.34 ±2.02	21.31	4.72E-06**
UA-IFASA-9288	14	2,201,870	0.25	0.40	11.29 ±2.1	28.93	1.06E-07***
HAPMAP32970-BTC-064990	14	2,288,510	0.33	0.46	9.77 ±1.95	25.2	6.74E-07***
HAPMAP24986-BTC-065021	14	2,313,595	0.33	0.46	9.77 ±1.95	25.2	6.74E-07***
ARS-BFGL-NGS-22111	14	2,347,219	0.38	0.49	-9.66 ±1.85	27.33	2.33E-07***
UA-IFASA-7269	14	2,370,256	0.38	0.49	-9.66 ±1.85	27.33	2.33E-07***
HAPMAP26072-BTC-065132	14	2,391,826	0.38	0.51	-8.84 ±1.85	22.9	2.12E-06**
BFGL-NGS-118081	14	2,511,265	0.40	0.51	7.17 ±1.84	15.2	0.000107*
ARS-BFGL-NGS-56327	14	2,580,414	0.31	0.45	10.51 ±1.97	28.55	1.27E-07***
ARS-BFGL-NGS-100480	14	2,607,583	0.39	0.51	12.04 ±1.85	42.46	1.46E-10***
UA-IFASA-5306	14	2,711,615	0.25	0.40	11.33 ±2.1	29.08	9.80E-08***

HAPMAP27703-BTC-053907	14	2,826,073	0.38	0.50	-9.47±1.88	25.53	5.68E-07***
HAPMAP22692-BTC-068210	14	3,018,726	0.27	0.41	10.86±2.05	27.92	1.74E-07***
BFGL-NGS-110993	14	3,059,045	0.39	0.50	8.03±1.9	17.86	2.73E-05*
HAPMAP23302-BTC-052123	14	3,099,635	0.27	0.41	10.43±2.07	25.36	6.21E-07***
HAPMAP25217-BTC-067767	14	3,189,312	0.40	0.51	-8.11±1.87	18.92	1.58E-05**
UA-IFASA-6329	14	3,465,237	0.50	0.54	-9.89±1.81	29.88	6.60E-08***
ARS-BFGL-NGS-56339	14	3,498,807	0.37	0.51	8.14±1.89	18.58	1.89E-05*
ARS-BFGL-NGS-3571	14	3,587,018	0.44	0.51	11.06±1.79	38.35	1.06E-09***
UA-IFASA-8927	14	3,640,094	0.46	0.51	7.23±1.76	16.94	4.36E-05*
BFGL-NGS-110563	14	3,799,228	0.46	0.52	10.14±1.81	31.54	2.91E-08***
HAPMAP32262-BTC-066621	14	3,834,069	0.35	0.48	-9.99±1.89	27.84	1.80E-07***
BFGL-NGS-115947	14	3,865,962	0.39	0.50	13.72±1.83	56.19	2.25E-13***
HAPMAP30091-BTC-005211	14	3,940,998	0.37	0.50	-7.44±1.91	15.19	0.000107*
HAPMAP51646-BTA-86764	14	4,302,229	0.45	0.55	-8.39±1.87	20.07	8.85E-06**
HAPMAP26591-BTC-056596	14	4,477,036	0.47	0.52	-8.13±1.81	20.11	8.67E-06**
HAPMAP30988-BTC-056315	14	4,693,901	0.38	0.50	-8.45±1.89	19.94	9.44E-06**
HAPMAP51078-BTA-87682	14	5,064,063	0.41	0.52	-7.6±1.9	16	7.08E-05*
BFGL-NGS-113706	14	5,117,434	0.40	0.51	7.78±1.91	16.58	5.26E-05*
HAPMAP32236-BTC-049785	14	5,139,498	0.41	0.50	8.35±1.87	20.03	9.04E-06**
UA-IFASA-6228	14	5,204,594	0.37	0.50	7.91±1.94	16.58	5.26E-05*
BFGL-NGS-110894	14	5,282,438	0.34	0.45	9.45±1.93	24.06	1.19E-06**
HAPMAP33635-BTC-049051	14	5,318,261	0.39	0.50	-8.22±1.83	20.22	8.19E-06**
HAPMAP32234-BTC-048199	14	5,640,338	0.35	0.48	7.71±1.87	17.02	4.18E-05*
UA-IFASA-6647	14	5,808,644	0.38	0.49	8.61±1.83	22.15	3.09E-06**
ARS-BFGL-NGS-102953	14	5,867,266	0.44	0.51	6.94±1.81	14.65	0.000142*
ARS-BFGL-BAC-8730	14	6,252,101	0.49	0.53	8.88±1.77	25.24	6.56E-07***
HAPMAP24518-BTC-062393	14	6,457,257	0.41	0.53	7.61±1.95	15.3	0.000102*
HAPMAP22779-BTC-061888	14	6,847,768	0.46	0.50	6.94±1.8	14.93	0.000123*
HAPMAP30986-BTC-056068	14	8,731,323	0.20	0.32	10.05±2.23	20.35	7.68E-06**
HAPMAP26301-BTC-055949	14	8,772,916	0.18	0.30	10.96±2.28	23.01	2.01E-06**
HAPMAP31247-BTC-009373	14	9,332,841	0.16	0.28	10.07±2.57	15.4	9.65E-05*
HAPMAP39958-BTA-96258	15	56,889,859	0.40	0.46	6.94±1.8	14.85	0.000128*
ARS-BFGL-NGS-25325	23	15,596,936	0.13	0.23	10.5±2.69	15.19	0.000107*
ARS-BFGL-BAC-29830	24	6,110,902	0.27	0.39	8.23±2.07	15.76	8.02E-05*
ARS-BFGL-NGS-38218	24	7,631,249	0.38	0.47	-7.13±1.8	15.72	8.17E-05*
ARS-BFGL-NGS-42400	24	48,636,081	0.35	0.48	-7.84±1.92	16.64	5.1E-05*
BTA-60128-NO-RS	25	38,544,241	0.41	0.47	6.81±1.78	14.7	0.000139*
ARS-BFGL-NGS-16336	26	34,575,102	0.39	0.49	-7.29±1.86	15.29	0.000102*
ARS-BFGL-NGS-43058	26	37,536,046	0.13	0.23	-10.7±2.7	15.74	8.1E-05*
ARS-BFGL-NGS-16904	29	50,395,795	0.38	0.48	-7.61±1.83	17.37	3.52E-05*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip. *Significant at genome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.01$; ***Significant at genome-wise FDR $P < 0.001$.

Table 4 – 5. Genome-wise significant SNPs for protein yield (PY) using single marker LD regression in the Canadian Holstein cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-BAC-7205	1	120,983,739	0.40	0.52	7.33 ±1.39	27.95	1.71E-07**
BTA-114286-NO-RS	1	121,075,154	0.40	0.50	-6.05 ±1.41	18.43	2.04467E-05*
ARS-BFGL-NGS-105044	1	147,597,144	0.13	0.23	-9.58 ±1.96	23.88	1.30E-06**
ARS-BFGL-NGS-109212	5	120,674,018	0.12	0.20	-8.74 ±2.01	18.87	1.62655E-05*
BTB-01843749	9	36,578,501	0.35	0.49	6.23 ±1.44	18.7	0.000017761*
BTB-01513309	9	85,200,611	0.15	0.25	8.27 ±1.87	19.6	1.13031E-05*
HAPMAP47338-BTA-59326	10	15,172,465	0.44	0.49	7.12 ±1.34	28.32	1.43E-07**
ARS-BFGL-NGS-37154	10	102,563,532	0.37	0.47	-5.58 ±1.31	18.28	2.19585E-05*
HAPMAP33734-BTA-162283	11	40,627,655	0.33	0.46	6.69 ±1.45	21.21	5.01E-06*
HAPMAP15326-RS29013300	11	41,385,158	0.49	0.53	7.15 ±1.34	28.33	1.43E-07**
ARS-BFGL-NGS-57820	14	236,532	0.24	0.38	-7.95 ±1.51	27.58	2.05E-07**
ARS-BFGL-NGS-4939	14	443,937	0.24	0.38	-7.68 ±1.54	25.04	7.26E-07**
ARS-BFGL-NGS-107379	14	679,600	0.28	0.44	-6.42 ±1.47	19.03	1.49828E-05*
UA-IFASA-5356	14	8,132,747	0.42	0.49	-5.9 ±1.34	19.43	1.22702E-05*
BFGL-NGS-117892	16	35,581,157	0.24	0.38	-6.38 ±1.48	18.48	0.000019824*
HAPMAP59410-SS46526044	21	30,726,309	0.20	0.30	6.82 ±1.59	18.33	2.14764E-05*
BTB-00891813	24	58,257,735	0.42	0.47	-5.65 ±1.26	20.16	8.45E-06*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip. *Significant at genome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.01$.

Table 4 – 6. Genome-wise significant SNPs for fat percentage (FP) using single marker LD regression in the Canadian Holstein cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
HAPMAP51685-BTA-113867	1	15,157,455	0.50	0.55	0.075 ±0.019	14.95	0.000122413*
ARS-BFGL-NGS-11580	1	116,611,632	0.33	0.43	0.07 ±0.019	13.47	0.000262519*
UA-IFASA-1795	1	116,820,290	0.41	0.47	0.067 ±0.019	13.2	0.000302283*
ARS-BFGL-NGS-31534	1	120,277,406	0.19	0.31	0.094 ±0.024	15.32	0.000100853*
HAPMAP46305-BTA-53011	1	120,372,217	0.19	0.31	0.094 ±0.024	15.32	0.000100853*
ARS-BFGL-NGS-553	1	125,411,899	0.35	0.47	0.073 ±0.019	14.32	0.000168675*
ARS-BFGL-NGS-106616	1	136,401,656	0.24	0.38	-0.08 ±0.022	13.69	0.000234027*
BFGL-NGS-117639	1	151,286,556	0.25	0.40	0.078 ±0.022	13.08	0.000321823*
ARS-BFGL-NGS-26099	3	14,418,594	0.17	0.30	-0.093 ±0.025	13.9	0.000209772*
ARS-BFGL-NGS-3374	3	14,779,899	0.16	0.28	-0.1 ±0.026	14.77	0.000133529*
ARS-BFGL-NGS-42701	3	20,130,429	0.40	0.53	0.082 ±0.02	17.53	3.22438E-05**
ARS-BFGL-NGS-65128	3	119,705,760	0.12	0.20	-0.099 ±0.027	13.42	0.000269429*
HAPMAP59011-RS29027498	4	27,732,338	0.39	0.50	-0.073 ±0.02	13.41	0.000272542*
BTB-01261088	4	85,681,605	0.44	0.48	-0.067 ±0.018	13.41	0.000271038*
BTB-01652664	4	88,536,202	0.37	0.47	0.071 ±0.018	15.82	7.75661E-05*
ARS-BFGL-NGS-16675	4	88,562,392	0.37	0.47	0.071 ±0.018	15.82	7.75661E-05*
BFGL-NGS-110883	4	96,110,843	0.38	0.47	0.078 ±0.019	17.76	2.86763E-05**
BTA-73613-NO-RS	5	60,366,994	0.15	0.26	-0.113 ±0.026	18.92	1.59131E-05**
BTB-01267080	5	95,538,815	0.21	0.35	0.091 ±0.024	14.65	0.000142371*
HAPMAP36106-SCAFFOLD22038_1672	5	102,836,545	0.27	0.42	0.1 ±0.021	22.19	3.03E-06***
BFGL-NGS-119645	5	103,296,590	0.45	0.54	0.082 ±0.019	18.64	1.83207E-05**
ARS-BFGL-NGS-61040	5	103,492,379	0.12	0.22	0.107 ±0.028	14.88	0.000126377*
HAPMAP28380-BTA-74649	5	103,598,357	0.49	0.56	0.075 ±0.018	17.2	3.81594E-05**
ARS-BFGL-NGS-15498	5	103,709,355	0.43	0.55	0.1 ±0.019	29.41	8.29E-08***
BFGL-NGS-115294	5	103,967,265	0.49	0.53	-0.095 ±0.019	26.13	4.21E-07***
HAPMAP53773-SS46526912	5	104,577,460	0.39	0.49	-0.088 ±0.019	20.68	6.50E-06**
HAPMAP59520-RS29021624	5	105,028,275	0.39	0.49	-0.091 ±0.019	22.21	3.01E-06***
HAPMAP47511-BTA-114200	5	105,350,648	0.28	0.42	0.087 ±0.021	17.57	3.16554E-05**
HAPMAP47185-BTA-114173	5	105,874,101	0.43	0.48	0.068 ±0.019	13	0.000336413*
UA-IFASA-6670	5	107,086,994	0.37	0.48	0.085 ±0.02	18.24	2.25365E-05**
ARS-BFGL-NGS-37981	5	108,087,764	0.48	0.50	-0.095 ±0.019	24.78	8.38E-07***
HAPMAP48257-BTA-24329	5	108,117,322	0.37	0.49	0.083 ±0.019	20.32	7.78E-06**
BTA-74798-NO-RS	5	108,214,402	0.47	0.51	0.082 ±0.019	19.49	1.18841E-05**
ARS-USMARC-226	5	108,288,993	0.33	0.48	0.072 ±0.019	13.59	0.000246555*
BFGL-NGS-118038	5	115,509,870	0.18	0.30	-0.088 ±0.024	13.88	0.000211996*
ARS-BFGL-NGS-21043	8	355,812	0.19	0.31	-0.088 ±0.023	14.47	0.0001561*
BTA-23646-NO-RS	8	94,732,857	0.18	0.33	-0.087 ±0.024	13.26	0.00029295*
ARS-BFGL-NGS-103752	10	103,736,758	0.26	0.42	-0.078 ±0.022	13.06	0.000325245*
ARS-BFGL-NGS-40347	11	13,334,124	0.21	0.31	0.082 ±0.023	13.21	0.000302845*
BTA-94053-NO-RS	11	19,321,756	0.19	0.30	0.086 ±0.023	14.33	0.000168157*
HAPMAP26194-BTA-158111	11	20,901,523	0.36	0.48	-0.076 ±0.02	14.39	0.000163516*
BTB-01256624	11	36,712,545	0.11	0.20	-0.107 ±0.029	13.78	0.000223475*
BTB-01406729	12	49,499,513	0.21	0.35	-0.09 ±0.022	16.78	4.73269E-05**
BTB-01236909	12	50,265,419	0.17	0.28	0.09 ±0.025	13.06	0.000325245*
BTB-01980482	12	50,852,178	0.17	0.28	0.095 ±0.025	14.35	0.000166021*
BTB-02092928	12	50,880,300	0.17	0.28	0.095 ±0.025	14.35	0.000166021*
BTB-00502017	12	66,468,572	0.41	0.49	0.075 ±0.019	15.95	7.26231E-05*
HAPMAP30381-BTC-005750	14	50,872	0.35	0.48	-0.167 ±0.017	101.62	3.27E-22***
HAPMAP30383-BTC-005848	14	76,703	0.35	0.50	0.235 ±0.018	180.37	1.97E-36***
BTA-34956-NO-RS	14	101,473	0.48	0.55	-0.162 ±0.018	84.32	5.70E-19***
ARS-BFGL-NGS-57820	14	236,532	0.24	0.38	0.341 ±0.017	415.62	2.09E-71***
ARS-BFGL-NGS-34135	14	260,341	0.41	0.51	0.214 ±0.017	165.73	6.45E-34***
ARS-BFGL-NGS-94706	14	281,533	0.41	0.51	0.205 ±0.017	149.68	4.29E-31***
ARS-BFGL-NGS-4939	14	443,937	0.24	0.38	0.359 ±0.017	474.33	7.45E-79***
HAPMAP52798-SS46526455	14	565,311	0.45	0.49	-0.092 ±0.017	27.77	1.87E-07***
ARS-BFGL-NGS-71749	14	596,341	0.28	0.43	-0.165 ±0.019	77.2	1.55E-17***

ARS-BFGL-NGS-107379	14	679,600	0.28	0.44	0.298±0.017	301.98	9.99E-56***
ARS-BFGL-NGS-18365	14	741,867	0.28	0.40	-0.115±0.02	32.82	1.56E-08***
HAPMAP30922-BTC-002021	14	763,331	0.27	0.39	-0.111±0.02	29.94	6.42E-08***
UA-IFASA-8997	14	812,103	0.22	0.33	-0.082±0.021	14.82	0.000130075*
HAPMAP25384-BTC-001997	14	835,054	0.44	0.51	0.174±0.017	100.32	4.89E-22***
HAPMAP24715-BTC-001973	14	856,889	0.44	0.51	0.174±0.017	100.28	4.95E-22***
BTA-35941-NO-RS	14	894,252	0.42	0.53	0.193±0.017	125.77	8.78E-27***
ARS-BFGL-NGS-26520	14	996,982	0.36	0.48	-0.192±0.017	127.02	5.31E-27***
UA-IFASA-6878	14	1,044,041	0.47	0.50	-0.108±0.018	37.61	1.51E-09***
ARS-BFGL-NGS-22866	14	1,131,952	0.48	0.51	-0.139±0.017	63.52	7.24E-15***
ARS-BFGL-NGS-3122	14	1,264,233	0.48	0.53	-0.134±0.018	55.71	2.74E-13***
HAPMAP29758-BTC-003619	14	1,339,276	0.49	0.53	0.151±0.017	79.04	6.08E-18***
HAPMAP30646-BTC-002054	14	1,461,085	0.44	0.53	-0.209±0.017	161.32	4.21E-33***
HAPMAP30086-BTC-002066	14	1,490,178	0.50	0.54	-0.226±0.016	207.28	6.97E-41***
HAPMAP30374-BTC-002159	14	1,546,591	0.42	0.53	0.193±0.017	130.49	1.23E-27***
ARS-BFGL-NGS-74378	14	1,889,210	0.26	0.40	0.164±0.02	68.45	7.50E-16***
UA-IFASA-9288	14	2,201,870	0.25	0.40	0.181±0.021	77.31	1.34E-17***
HAPMAP24777-BTC-064977	14	2,261,623	0.42	0.54	-0.111±0.019	35.62	3.96E-09***
HAPMAP32970-BTC-064990	14	2,288,510	0.33	0.46	0.172±0.019	81.63	1.91E-18***
HAPMAP24986-BTC-065021	14	2,313,595	0.33	0.46	0.172±0.019	81.63	1.91E-18***
ARS-BFGL-NGS-22111	14	2,347,219	0.38	0.49	-0.175±0.018	95.86	3.40E-21***
UA-IFASA-7269	14	2,370,256	0.38	0.49	-0.175±0.018	95.86	3.40E-21***
HAPMAP26072-BTC-065132	14	2,391,826	0.38	0.51	-0.154±0.018	72.48	1.20E-16***
BFGL-NGS-118081	14	2,511,265	0.40	0.51	0.13±0.018	50.98	2.53E-12***
ARS-BFGL-NGS-56327	14	2,580,414	0.31	0.45	0.176±0.019	83.28	9.04E-19***
ARS-BFGL-NGS-100480	14	2,607,583	0.39	0.51	0.215±0.017	152.04	1.64E-31***
UA-IFASA-5306	14	2,711,615	0.25	0.40	0.182±0.021	77.56	1.19E-17***
ARS-BFGL-NGS-54400	14	2,736,947	0.40	0.49	-0.07±0.019	13.91	0.000208656*
ARS-BFGL-NGS-103321	14	2,785,213	0.40	0.51	-0.089±0.019	21.76	3.76E-06***
ARS-BFGL-BAC-25166	14	2,805,785	0.12	0.22	0.111±0.028	15.71	8.22304E-05*
HAPMAP27703-BTC-053907	14	2,826,073	0.38	0.50	-0.182±0.018	101.99	2.34E-22***
HAPMAP22783-BTC-068255	14	2,989,275	0.49	0.55	0.104±0.019	31.49	2.98E-08***
HAPMAP22692-BTC-068210	14	3,018,726	0.27	0.41	0.177±0.02	77.42	1.27E-17***
BFGL-NGS-110993	14	3,059,045	0.39	0.50	0.147±0.019	61.39	1.94E-14***
HAPMAP23302-BTC-052123	14	3,099,635	0.27	0.41	0.182±0.02	81.09	2.43E-18***
HAPMAP25217-BTC-067767	14	3,189,312	0.40	0.51	-0.182±0.018	105.08	6.13E-23***
UA-IFASA-6329	14	3,465,237	0.50	0.54	-0.168±0.018	91.58	2.25E-20***
ARS-BFGL-NGS-56339	14	3,498,807	0.37	0.51	0.106±0.019	30.9	3.98E-08***
ARS-BFGL-NGS-3571	14	3,587,018	0.44	0.51	0.209±0.017	156.55	2.63E-32***
UA-IFASA-8927	14	3,640,094	0.46	0.51	0.111±0.018	40.39	3.94E-10***
BFGL-NGS-118478	14	3,660,264	0.11	0.20	0.138±0.03	20.76	6.26E-06**
BFGL-NGS-110563	14	3,799,228	0.46	0.52	0.204±0.017	143.46	5.49E-30***
HAPMAP32262-BTC-066621	14	3,834,069	0.35	0.48	-0.184±0.018	102.87	1.62E-22***
BFGL-NGS-115947	14	3,865,962	0.39	0.50	0.237±0.017	194.06	1.02E-38***
HAPMAP30091-BTC-005211	14	3,940,998	0.37	0.50	-0.168±0.018	83.27	9.18E-19***
ARS-BFGL-BAC-24839	14	3,993,200	0.48	0.53	-0.096±0.018	27.92	1.73E-07***
ARS-BFGL-BAC-24804	14	4,157,675	0.27	0.40	0.145±0.02	52.9	1.02E-12***
HAPMAP23454-BTC-046932	14	4,182,816	0.27	0.39	0.146±0.02	53.93	6.30E-13***
HAPMAP51646-BTA-86764	14	4,302,229	0.45	0.55	-0.16±0.018	76.55	1.87E-17***
UA-IFASA-9107	14	4,356,232	0.41	0.49	0.112±0.018	39.42	6.28E-10***
HAPMAP31968-BTC-056754	14	4,399,113	0.42	0.49	0.077±0.018	17.51	3.25496E-05**
HAPMAP26591-BTC-056596	14	4,477,036	0.47	0.52	-0.161±0.018	83.52	8.09E-19***
HAPMAP23618-BTC-056528	14	4,518,666	0.33	0.47	0.109±0.019	32.85	1.53E-08***
UA-IFASA-6370	14	4,619,625	0.31	0.42	0.077±0.02	15.54	8.99195E-05*
HAPMAP30988-BTC-056315	14	4,693,901	0.38	0.50	-0.162±0.018	76.78	1.69E-17***
ARS-BFGL-NGS-43719	14	4,781,194	0.30	0.43	0.095±0.02	23.13	1.89E-06***
UA-IFASA-4560	14	4,922,757	0.35	0.45	0.127±0.019	46.22	2.43E-11***
BFGL-NGS-112858	14	4,956,375	0.27	0.40	0.139±0.02	47.2	1.52E-11***
HAPMAP51078-BTA-87682	14	5,064,063	0.41	0.52	-0.128±0.019	45.51	3.39E-11***
ARS-BFGL-NGS-55227	14	5,085,416	0.27	0.41	0.131±0.02	41.54	2.27E-10***
BFGL-NGS-113706	14	5,117,434	0.40	0.51	0.165±0.019	78.51	7.72E-18***
HAPMAP32236-BTC-049785	14	5,139,498	0.41	0.50	0.173±0.018	91.33	2.51E-20***

UA-IFASA-6228	14	5,204,594	0.37	0.50	0.176±0.019	88.23	9.93E-20***
ARS-BFGL-BAC-20965	14	5,225,004	0.36	0.51	-0.125±0.019	42.41	1.49E-10***
BFGL-NGS-110894	14	5,282,438	0.34	0.45	0.197±0.018	114.97	8.61E-25***
HAPMAP33635-BTC-049051	14	5,318,261	0.39	0.50	-0.133±0.018	54.36	5.15E-13***
HAPMAP27091-BTC-048823	14	5,356,988	0.41	0.52	-0.135±0.019	52.23	1.40E-12***
HAPMAP23851-BTC-048718	14	5,387,836	0.33	0.49	-0.121±0.019	38.49	9.86E-10***
HAPMAP32234-BTC-048199	14	5,640,338	0.35	0.48	0.175±0.018	95.49	4.00E-21***
HAPMAP26283-BTC-048098	14	5,696,729	0.30	0.43	0.101±0.02	25.74	5.13E-07***
UA-IFASA-6647	14	5,808,644	0.38	0.49	0.169±0.018	91.13	2.73E-20***
HAPMAP32948-BTC-047992	14	5,839,290	0.29	0.42	0.141±0.02	49.37	5.44E-12***
ARS-BFGL-NGS-102953	14	5,867,266	0.44	0.51	0.163±0.018	86.83	1.85E-19***
HAPMAP25716-BTC-047850	14	5,937,550	0.44	0.51	-0.094±0.018	27.49	2.15E-07***
HAPMAP23799-BTC-047701	14	6,044,246	0.42	0.51	-0.076±0.019	16.74	4.83004E-05**
ARS-BFGL-NGS-73096	14	6,214,126	0.49	0.50	0.078±0.019	17.21	3.80768E-05**
ARS-BFGL-BAC-8730	14	6,252,101	0.49	0.53	0.129±0.018	54.28	5.35E-13***
HAPMAP41162-BTA-115830	14	6,538,124	0.35	0.49	0.078±0.02	15.22	0.000105875*
HAPMAP26598-BTC-062212	14	6,567,155	0.33	0.46	0.074±0.02	14.48	0.000155183*
UA-IFASA-8561	14	6,779,090	0.17	0.29	0.103±0.024	18.01	2.52101E-05**
HAPMAP22779-BTC-061888	14	6,847,768	0.46	0.50	0.119±0.018	44.43	5.68E-11***
HAPMAP22964-BTC-001504	14	6,932,267	0.17	0.29	0.104±0.024	18.62	1.84623E-05**
ARS-BFGL-NGS-85289	14	7,062,935	0.49	0.49	0.073±0.019	15.3	0.000101749*
HAPMAP22724-BTC-001541	14	7,134,175	0.42	0.47	0.084±0.018	21.83	3.63E-06***
ARS-BFGL-NGS-29470	14	7,353,500	0.29	0.41	0.074±0.02	13.15	0.000310472*
BFGL-NGS-110918	14	7,408,578	0.33	0.45	0.071±0.019	13.55	0.000251879*
HAPMAP41241-BTA-35529	14	7,751,463	0.45	0.51	0.119±0.018	42.21	1.65E-10***
ARS-BFGL-NGS-101419	14	7,790,462	0.27	0.43	0.081±0.021	15.21	0.000106308*
ARS-BFGL-BAC-18370	14	7,812,530	0.42	0.49	-0.101±0.019	29.58	7.65E-08***
UA-IFASA-7842	14	7,857,977	0.48	0.53	0.103±0.018	32.35	1.95E-08***
HAPMAP31565-BTC-007750	14	7,882,481	0.45	0.49	0.072±0.018	15.5	9.16431E-05*
ARS-BFGL-BAC-20850	14	7,928,144	0.46	0.49	0.103±0.019	30.32	5.34E-08***
HAPMAP30097-BTC-007678	14	7,969,429	0.48	0.48	-0.112±0.018	38.32	1.08E-09***
HAPMAP31564-BTC-007633	14	7,998,736	0.45	0.48	0.079±0.019	17.93	2.63809E-05**
UA-IFASA-5356	14	8,132,747	0.42	0.49	0.089±0.019	22.74	2.30E-06***
ARS-BFGL-NGS-1857	14	8,323,182	0.48	0.52	0.093±0.019	23.24	1.79E-06***
HAPMAP24767-BTC-058058	14	8,353,977	0.44	0.53	0.091±0.02	21.6	4.08E-06**
HAPMAP22733-BTC-058037	14	8,376,617	0.48	0.53	-0.095±0.019	25.14	6.90E-07***
HAPMAP23517-BTC-058008	14	8,403,522	0.37	0.47	0.083±0.02	18.01	2.52639E-05**
ARS-BFGL-NGS-41494	14	8,508,157	0.45	0.51	0.088±0.019	22.2	3.01E-06***
HAPMAP30653-BTC-008335	14	8,543,256	0.43	0.53	-0.092±0.019	24.06	1.19E-06***
ARS-USMARC-PARENT-DQ846690-NO-RS	14	8,561,778	0.50	0.53	0.087±0.019	21.35	4.62E-06**
HAPMAP24065-BTC-072634	14	8,629,153	0.42	0.51	-0.105±0.019	31.93	2.41E-08***
ARS-BFGL-BAC-20261	14	8,656,676	0.40	0.48	0.125±0.019	46.09	2.60E-11***
ARS-BFGL-NGS-28580	14	8,691,664	0.27	0.41	0.074±0.021	13.09	0.000320181*
HAPMAP30986-BTC-056068	14	8,731,323	0.20	0.32	0.131±0.022	34.14	8.17E-09***
HAPMAP26301-BTC-055949	14	8,772,916	0.18	0.30	0.134±0.023	34.02	8.66E-09***
HAPMAP25450-BTC-055819	14	8,810,386	0.28	0.40	0.092±0.02	20.85	5.95E-06**
HAPMAP30158-BTC-057468	14	8,850,016	0.44	0.47	0.075±0.018	17.41	3.42869E-05**
HAPMAP26890-BTC-057492	14	8,875,202	0.49	0.51	-0.113±0.018	39.4	6.36E-10***
HAPMAP24111-BTC-009033	14	9,098,067	0.42	0.47	0.106±0.019	30.28	5.43E-08***
HAPMAP27677-BTC-009441	14	9,498,494	0.39	0.49	-0.102±0.019	28.76	1.15E-07***
HAPMAP53853-RS29015157	14	9,670,932	0.40	0.50	-0.105±0.019	31.58	2.86E-08***
UA-IFASA-8705	14	9,809,560	0.45	0.50	0.07±0.018	14.9	0.000124849*
BFGL-NGS-117252	14	9,990,357	0.44	0.51	0.079±0.018	18.67	1.80058E-05**
HAPMAP30171-BTC-063976	14	10,347,750	0.49	0.55	0.08±0.019	18.59	0.000018751**
HAPMAP23251-BTC-063743	14	10,390,402	0.34	0.48	-0.093±0.019	23.42	1.63E-06***
HAPMAP25440-BTC-053771	14	10,464,996	0.43	0.51	-0.082±0.019	19.3	1.30631E-05**
HAPMAP26328-BTC-063507	14	10,534,560	0.40	0.53	-0.078±0.019	16.74	4.83092E-05**
HAPMAP25770-BTC-068957	14	10,563,449	0.33	0.48	-0.106±0.019	30.58	4.66E-08***
HAPMAP31038-BTC-069141	14	10,632,341	0.47	0.52	-0.095±0.018	27.41	2.23E-07***
UA-IFASA-6074	14	10,701,474	0.47	0.52	-0.099±0.018	29.78	6.91E-08***
HAPMAP23304-BTC-069412	14	10,768,341	0.46	0.51	0.104±0.018	33.8	9.62E-09***

UA-IFASA-6621	14	10,875,943	0.46	0.51	0.101 ±0.018	31.6	2.82E-08***
HAPMAP38378-BTA-114219	14	10,953,660	0.46	0.51	-0.112 ±0.018	38.78	8.56E-10***
HAPMAP24056-BTC-060048	14	10,975,727	0.35	0.47	-0.111 ±0.019	33.2	1.29E-08***
HAPMAP33448-BTC-060018	14	11,022,829	0.37	0.48	-0.105 ±0.019	31.18	3.47E-08***
UA-IFASA-5274	14	11,136,304	0.45	0.50	0.088 ±0.018	23.81	1.34E-06***
ARS-BFGL-NGS-104836	14	11,377,913	0.33	0.45	-0.111 ±0.019	33.19	1.30E-08***
BTA-35971-NO-RS	14	11,504,490	0.33	0.45	-0.107 ±0.019	31.48	3.00E-08***
HAPMAP57409-RS29021898	14	11,524,613	0.30	0.42	0.085 ±0.02	18.25	0.000022341**
BFGL-NGS-119442	14	11,545,868	0.39	0.49	0.097 ±0.019	26.38	3.74E-07***
ARS-BFGL-NGS-8221	14	11,720,820	0.47	0.50	-0.092 ±0.018	25.45	5.93E-07***
HAPMAP23784-BTC-010226	14	11,748,068	0.49	0.50	0.072 ±0.019	15.02	0.00011747*
ARS-BFGL-NGS-24160	14	11,963,066	0.40	0.48	-0.098 ±0.018	28.26	1.47E-07***
ARS-BFGL-NGS-105600	14	11,985,275	0.50	0.49	-0.082 ±0.018	20.35	7.68E-06**
HAPMAP46741-BTA-00625	14	12,063,432	0.49	0.50	0.08 ±0.018	19.06	1.47731E-05**
UA-IFASA-7696	14	12,380,364	0.38	0.48	-0.085 ±0.019	20.19	8.32E-06**
ARS-BFGL-NGS-33755	14	14,072,516	0.30	0.43	0.073 ±0.02	13.94	0.000205422*
BFGL-NGS-117354	14	14,132,142	0.30	0.43	0.073 ±0.02	13.94	0.000205422*
ARS-BFGL-NGS-63270	14	14,365,665	0.30	0.45	-0.071 ±0.019	13.66	0.000237747*
HAPMAP58921-RS29010046	14	14,784,728	0.24	0.36	-0.075 ±0.021	13.11	0.000316812*
BTA-13384-RS29018344	15	57,429,381	0.27	0.38	0.086 ±0.021	16.65	5.11034E-05*
BTB-01090150	16	60,113,461	0.21	0.33	0.082 ±0.022	14.28	0.000172123*
ARS-BFGL-NGS-95458	17	41,789,183	0.49	0.54	0.068 ±0.019	13.01	0.000334138*
HAPMAP41532-BTA-40872	17	41,810,695	0.37	0.49	-0.093 ±0.019	22.68	2.38E-06***
HAPMAP3803-BTA-21924	17	51,840,285	0.36	0.46	0.068 ±0.018	13.66	0.000237747*
ARS-BFGL-BAC-28014	18	6,688,498	0.11	0.20	-0.128 ±0.033	15	0.000119269*
ARS-BFGL-NGS-76756	20	35,610,597	0.26	0.35	0.072 ±0.02	13.23	0.000297721*
ARS-BFGL-NGS-19519	26	48,822,071	0.42	0.51	-0.073 ±0.019	14.98	0.000119909*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip. *Significant at genome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.01$; ***Significant at genome-wise FDR $P < 0.001$.

Table 4 – 7. Genome-wise significant SNPs for protein percentage (PP) using single marker LD regression in the Canadian Holstein cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-NGS-42701	3	20,130,429	0.40	0.53	0.034±0.008	17.92	2.64E-05*
HAPMAP43907-BTA-67629	3	38,712,900	0.32	0.44	-0.031±0.008	14.23	0.000177*
ARS-BFGL-NGS-19374	6	19,635,119	0.12	0.20	0.043±0.011	15.47	9.29E-05*
HAPMAP41636-BTA-77903	6	21,058,743	0.27	0.40	0.033±0.008	15.74	8.08E-05*
ARS-BFGL-NGS-56572	6	47,292,121	0.14	0.25	0.043±0.01	17.44	3.37E-05*
HAPMAP44302-BTA-86279	6	50,416,046	0.41	0.47	0.03±0.007	18	2.53E-05*
ARS-BFGL-NGS-41652	6	50,645,902	0.49	0.47	-0.03±0.007	18.4	2.06E-05**
HAPMAP55170-RS29017690	6	60,904,764	0.47	0.55	-0.03±0.008	14.65	0.000142*
ARS-BFGL-NGS-62682	6	61,252,856	0.48	0.51	-0.028±0.008	13.87	0.000213*
BTB-01534149	6	66,150,019	0.36	0.49	0.031±0.008	15.16	0.000109*
ARS-BFGL-NGS-33719	6	69,739,267	0.16	0.28	0.04±0.01	15.01	0.000118*
HAPMAP60836-RS29027147	6	71,189,704	0.29	0.43	-0.033±0.008	16.44	5.64E-05*
ARS-BFGL-NGS-89119	11	9,430,780	0.22	0.34	0.035±0.009	14.08	0.000191*
HAPMAP57340-RS29010501	11	46,636,484	0.42	0.48	0.03±0.008	15.37	9.78E-05*
BFGL-NGS-111691	11	46,688,491	0.37	0.46	0.031±0.008	15.07	0.000114*
ARS-BFGL-NGS-91269	11	47,313,224	0.28	0.40	0.033±0.008	16.51	5.44E-05*
HAPMAP43116-BTA-97124	11	47,534,009	0.18	0.29	0.039±0.009	17.46	3.34E-05*
HAPMAP40956-BTA-34026	13	11,829,385	0.21	0.32	-0.036±0.01	14.02	0.000197*
ARS-BFGL-NGS-23971	13	45,326,653	0.11	0.20	-0.049±0.012	17.3	3.62E-05*
HAPMAP30381-BTC-005750	14	50,872	0.35	0.48	-0.037±0.007	27.54	2.09E-07***
HAPMAP30383-BTC-005848	14	76,703	0.35	0.50	0.047±0.008	35.17	4.93E-09***
BTA-34956-NO-RS	14	101,473	0.48	0.55	-0.032±0.008	18.61	1.86E-05**
ARS-BFGL-NGS-57820	14	236,532	0.24	0.38	0.069±0.008	68.01	9.14E-16***
ARS-BFGL-NGS-34135	14	260,341	0.41	0.51	0.042±0.007	32.42	1.89E-08***
ARS-BFGL-NGS-94706	14	281,533	0.41	0.51	0.042±0.007	32.47	1.84E-08***
ARS-BFGL-NGS-4939	14	443,937	0.24	0.38	0.07±0.008	69.76	4.10E-16***
ARS-BFGL-NGS-71749	14	596,341	0.28	0.43	-0.039±0.008	24.7	8.59E-07***
ARS-BFGL-NGS-107379	14	679,600	0.28	0.44	0.057±0.008	48.98	6.50E-12***
HAPMAP25384-BTC-001997	14	835,054	0.44	0.51	0.032±0.008	17.54	3.21E-05*
HAPMAP24715-BTC-001973	14	856,889	0.44	0.51	0.031±0.008	16.78	4.73E-05*
BTA-35941-NO-RS	14	894,252	0.42	0.53	0.047±0.007	39.58	5.81E-10***
ARS-BFGL-NGS-26520	14	996,982	0.36	0.48	-0.041±0.007	31.27	3.32E-08***
ARS-BFGL-NGS-22866	14	1,131,952	0.48	0.51	-0.032±0.007	19.65	1.09E-05**
HAPMAP29758-BTC-003619	14	1,339,276	0.49	0.53	0.033±0.007	21.17	5.06E-06**
HAPMAP30646-BTC-002054	14	1,461,085	0.44	0.53	-0.043±0.007	34.4	7.17E-09***
HAPMAP30086-BTC-002066	14	1,490,178	0.50	0.54	-0.049±0.007	48.48	8.24E-12***
HAPMAP30374-BTC-002159	14	1,546,591	0.42	0.53	0.046±0.007	39.43	6.25E-10***
ARS-BFGL-NGS-74378	14	1,889,210	0.26	0.40	0.035±0.008	17.05	4.12E-05*
UA-IFASA-9288	14	2,201,870	0.25	0.40	0.039±0.009	19.91	9.58E-06**
HAPMAP24777-BTC-064977	14	2,261,623	0.42	0.54	-0.032±0.008	17.66	3.01E-05*
HAPMAP32970-BTC-064990	14	2,288,510	0.33	0.46	0.038±0.008	22.2	3.01E-06**
HAPMAP24986-BTC-065021	14	2,313,595	0.33	0.46	0.038±0.008	22.2	3.01E-06**
ARS-BFGL-NGS-22111	14	2,347,219	0.38	0.49	-0.038±0.008	25.34	6.24E-07***
UA-IFASA-7269	14	2,370,256	0.38	0.49	-0.038±0.008	25.34	6.24E-07***
HAPMAP26072-BTC-065132	14	2,391,826	0.38	0.51	-0.034±0.008	20.45	7.28E-06**
BFGL-NGS-118081	14	2,511,265	0.40	0.51	0.034±0.008	20.55	6.93E-06**
ARS-BFGL-NGS-56327	14	2,580,414	0.31	0.45	0.041±0.008	24.6	9.03E-07***
ARS-BFGL-NGS-100480	14	2,607,583	0.39	0.51	0.045±0.008	34.75	6.05E-09***
UA-IFASA-5306	14	2,711,615	0.25	0.40	0.039±0.009	19.78	1.02E-05**
HAPMAP27703-BTC-053907	14	2,826,073	0.38	0.50	-0.04±0.008	26.66	3.24E-07***
HAPMAP22783-BTC-068255	14	2,989,275	0.49	0.55	0.031±0.008	16.15	6.54E-05*
HAPMAP22692-BTC-068210	14	3,018,726	0.27	0.41	0.041±0.009	23.24	1.79E-06**
BFGL-NGS-110993	14	3,059,045	0.39	0.50	0.038±0.008	23.12	1.90E-06**
HAPMAP23302-BTC-052123	14	3,099,635	0.27	0.41	0.044±0.009	26.85	2.95E-07***
HAPMAP25217-BTC-067767	14	3,189,312	0.40	0.51	-0.043±0.008	31.92	2.41E-08***
UA-IFASA-6329	14	3,465,237	0.50	0.54	-0.041±0.007	30.4	5.09E-08***
ARS-BFGL-NGS-3571	14	3,587,018	0.44	0.51	0.045±0.007	37.52	1.57E-09***

UA-IFASA-8927	14	3,640,094	0.46	0.51	0.03 ±0.007	17.5	3.27E-05*
BFGL-NGS-110563	14	3,799,228	0.46	0.52	0.046 ±0.007	37.56	1.54E-09***
HAPMAP32262-BTC-066621	14	3,834,069	0.35	0.48	-0.051 ±0.008	43.28	9.82E-11***
BFGL-NGS-115947	14	3,865,962	0.39	0.50	0.05 ±0.008	42.01	1.80E-10***
HAPMAP30091-BTC-005211	14	3,940,998	0.37	0.50	-0.042 ±0.008	29.49	7.97E-08***
ARS-BFGL-BAC-24839	14	3,993,200	0.48	0.53	-0.028 ±0.007	13.85	0.000215*
ARS-BFGL-BAC-24804	14	4,157,675	0.27	0.40	0.038 ±0.008	20.53	7.00E-06**
HAPMAP23454-BTC-046932	14	4,182,816	0.27	0.39	0.037 ±0.008	19.85	9.87E-06**
HAPMAP51646-BTA-86764	14	4,302,229	0.45	0.55	-0.042 ±0.008	29.76	6.98E-08***
UA-IFASA-9107	14	4,356,232	0.41	0.49	0.031 ±0.007	17.04	4.14E-05*
HAPMAP26591-BTC-056596	14	4,477,036	0.47	0.52	-0.042 ±0.007	31.8	2.56E-08***
HAPMAP23618-BTC-056528	14	4,518,666	0.33	0.47	0.029 ±0.008	14.02	0.000197*
HAPMAP30988-BTC-056315	14	4,693,901	0.38	0.50	-0.041 ±0.008	27.97	1.69E-07***
UA-IFASA-4560	14	4,922,757	0.35	0.45	0.036 ±0.008	21.46	4.37E-06**
BFGL-NGS-112858	14	4,956,375	0.27	0.40	0.039 ±0.008	21.77	3.74E-06**
ARS-BFGL-NGS-55227	14	5,085,416	0.27	0.41	0.035 ±0.008	17.59	3.12E-05*
BFGL-NGS-113706	14	5,117,434	0.40	0.51	0.042 ±0.008	28.21	1.50E-07***
HAPMAP32236-BTC-049785	14	5,139,498	0.41	0.50	0.043 ±0.008	31.68	2.71E-08***
UA-IFASA-6228	14	5,204,594	0.37	0.50	0.04 ±0.008	25.31	6.34E-07***
ARS-BFGL-BAC-20965	14	5,225,004	0.36	0.51	-0.031 ±0.008	15.09	0.000113*
BFGL-NGS-110894	14	5,282,438	0.34	0.45	0.038 ±0.008	22.52	2.56E-06**
HAPMAP27091-BTC-048823	14	5,356,988	0.41	0.52	-0.032 ±0.008	16.92	4.4E-05*
HAPMAP23851-BTC-048718	14	5,387,836	0.33	0.49	-0.035 ±0.008	18.59	1.87E-05**
HAPMAP32234-BTC-048199	14	5,640,338	0.35	0.48	0.035 ±0.008	20.49	7.14E-06**
HAPMAP26283-BTC-048098	14	5,696,729	0.30	0.43	0.035 ±0.008	18.56	1.9E-05**
UA-IFASA-6647	14	5,808,644	0.38	0.49	0.04 ±0.008	27.82	1.82E-07***
HAPMAP32948-BTC-047992	14	5,839,290	0.29	0.42	0.041 ±0.008	23.72	1.40E-06**
ARS-BFGL-NGS-102953	14	5,867,266	0.44	0.51	0.047 ±0.007	40.55	3.64E-10***
HAPMAP25716-BTC-047850	14	5,937,550	0.44	0.51	-0.03 ±0.007	17.19	3.83E-05*
ARS-BFGL-NGS-73096	14	6,214,126	0.49	0.50	0.031 ±0.008	16.11	6.68E-05*
HAPMAP22779-BTC-061888	14	6,847,768	0.46	0.50	0.036 ±0.007	23.16	1.86E-06**
HAPMAP22724-BTC-001541	14	7,134,175	0.42	0.47	0.027 ±0.007	13.74	0.000228*
BFGL-NGS-118604	14	7,446,141	0.46	0.52	-0.028 ±0.007	14.29	0.000171*
HAPMAP41241-BTA-35529	14	7,751,463	0.45	0.51	0.035 ±0.008	21.49	4.31E-06**
ARS-BFGL-BAC-18370	14	7,812,530	0.42	0.49	-0.036 ±0.008	22.23	2.97E-06**
UA-IFASA-7842	14	7,857,977	0.48	0.53	0.035 ±0.007	22.17	3.06E-06**
ARS-BFGL-BAC-20850	14	7,928,144	0.46	0.49	0.034 ±0.008	19.69	1.07E-05**
HAPMAP30097-BTC-007678	14	7,969,429	0.48	0.48	-0.034 ±0.007	20.54	6.96E-06**
ARS-BFGL-NGS-1857	14	8,323,182	0.48	0.52	0.033 ±0.008	17.58	3.14E-05*
HAPMAP24767-BTC-058058	14	8,353,977	0.44	0.53	0.03 ±0.008	13.79	0.000222*
HAPMAP22733-BTC-058037	14	8,376,617	0.48	0.53	-0.03 ±0.008	14.52	0.000152*
HAPMAP30653-BTC-008335	14	8,543,256	0.43	0.53	-0.029 ±0.008	13.82	0.000219*
ARS-USMARC-PARENT-DQ846690-NO-RS	14	8,561,778	0.50	0.53	0.029 ±0.008	14.23	0.000177*
HAPMAP24065-BTC-072634	14	8,629,153	0.42	0.51	-0.032 ±0.008	17.72	2.92E-05*
ARS-BFGL-BAC-20261	14	8,656,676	0.40	0.48	0.035 ±0.008	21.05	5.38E-06**
HAPMAP30158-BTC-057468	14	8,850,016	0.44	0.47	0.029 ±0.007	15.7	8.25E-05*
HAPMAP26890-BTC-057492	14	8,875,202	0.49	0.51	-0.037 ±0.007	25.39	6.09E-07***
HAPMAP27677-BTC-009441	14	9,498,494	0.39	0.49	-0.037 ±0.008	22.41	2.71E-06**
HAPMAP53853-RS29015157	14	9,670,932	0.40	0.50	-0.036 ±0.008	21.69	3.89E-06**
UA-IFASA-8705	14	9,809,560	0.45	0.50	0.031 ±0.007	17.36	3.51E-05*
HAPMAP36733-SCAFFOLD230838_1182	14	10,106,755	0.43	0.52	0.03 ±0.008	14.49	0.000154*
HAPMAP25468-BTC-063927	14	10,369,144	0.40	0.50	-0.03 ±0.008	15.46	9.34E-05*
HAPMAP25440-BTC-053771	14	10,464,996	0.43	0.51	-0.037 ±0.008	24.26	1.07E-06***
HAPMAP25770-BTC-068957	14	10,563,449	0.33	0.48	-0.036 ±0.008	20.92	5.74E-06**
HAPMAP38378-BTA-114219	14	10,953,660	0.46	0.51	-0.031 ±0.007	16.85	4.56E-05*
HAPMAP24056-BTC-060048	14	10,975,727	0.35	0.47	-0.039 ±0.008	24.66	8.77E-07***
HAPMAP33448-BTC-060018	14	11,022,829	0.37	0.48	-0.036 ±0.008	22.09	3.18E-06**
ARS-BFGL-NGS-104836	14	11,377,913	0.33	0.45	-0.039 ±0.008	24.2	1.10E-06***
BTA-35971-NO-RS	14	11,504,490	0.33	0.45	-0.039 ±0.008	24.44	9.79E-07***
BFGL-NGS-119442	14	11,545,868	0.39	0.49	0.029 ±0.008	13.98	0.000201*

ARS-BFGL-NGS-8221	14	11,720,820	0.47	0.50	-0.029 ±0.008	15.11	0.000112*
HAPMAP23784-BTC-010226	14	11,748,068	0.49	0.50	0.029 ±0.008	14.52	0.000152*
ARS-BFGL-NGS-24160	14	11,963,066	0.40	0.48	-0.029 ±0.008	14.41	0.000161*
UA-IFASA-7696	14	12,380,364	0.38	0.48	-0.03 ±0.008	14.62	0.000144*
ARS-BFGL-NGS-63270	14	14,365,665	0.30	0.45	-0.035 ±0.008	20.26	8.02E-06**
ARS-BFGL-NGS-549	14	14,409,359	0.28	0.40	0.031 ±0.008	14.5	0.000154*
HAPMAP58921-RS29010046	14	14,784,728	0.24	0.36	-0.034 ±0.008	16.98	4.27E-05*
HAPMAP41515-BTA-35210	14	59,770,694	0.47	0.49	-0.028 ±0.007	15.25	0.000104*
BFGL-NGS-119538	15	26,541,932	0.31	0.44	0.03 ±0.008	13.98	0.000201*
HAPMAP60722-RS29027621	15	26,564,544	0.31	0.44	0.03 ±0.008	13.98	0.000201*
BTB-00666435	17	473,998	0.37	0.45	0.03 ±0.007	16.94	4.36E-05*
BTA-42041-NO-RS	17	7,263,715	0.30	0.42	-0.032 ±0.008	14.72	0.000137*
HAPMAP41532-BTA-40872	17	41,810,695	0.37	0.49	-0.032 ±0.008	16.48	5.52E-05*
BFGL-NGS-117951	19	53,162,863	0.40	0.49	0.027 ±0.007	13.74	0.000228*
ARS-BFGL-NGS-18978	20	26,555,886	0.16	0.29	-0.038 ±0.01	14.46	0.000157*
ARS-BFGL-NGS-20700	20	27,669,913	0.15	0.27	-0.041 ±0.01	15.34	9.94E-05*
HAPMAP50988-BTA-50138	20	29,240,614	0.21	0.33	-0.034 ±0.009	15.01	0.000118*
HAPMAP39811-BTA-122745	20	35,432,863	0.28	0.39	-0.034 ±0.009	15.28	0.000102*
BTA-56731-NO-RS	23	46,011,251	0.27	0.39	-0.03 ±0.008	13.79	0.000222*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip. *Significant at genome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.01$; ***Significant at genome-wise FDR $P < 0.001$.

Table 4 – 8. Chromosomal distribution of significant SNPs for milk production traits in the Canadian Holstein cattle using single marker LD regression.

Trait	N*	BTA(N)	N**	BTA(N)	N***	BTA(N)
MY	94	1(9), 5(5), 7(2), 9(1), 10(2), 11(5), 12(2), 14(60), 17(1), 18(1), 19(1), 21(1), 27(1), 28(3)	49	1(4), 5(3), 12(1), 14(41)	24	14(24)
FY	97	1(2), 2(2), 3(2), 5(13), 6(1), 7(3), 9(2), 12(2), 14(61), 15(1), 23(1), 24(3), 25(1), 26(2), 29(1)	54	1(1), 5(5), 7(1), 14(47)	31	14(31)
PY	17	1(3), 5(1), 9(2), 10(2), 11(2), 14(4), 16(1), 21(1), 24(1)	6	1(2), 10(1), 11(1), 14(2)	0	-
FP	206	1(8), 3(4), 4(5), 5(18), 8(2), 10(1), 11(4), 12(5), 14(151), 15(1), 16(1), 17(3), 18(1), 20(1), 26(1)	150	3(1), 4(1), 5(13), 12(1), 14(133), 17(1)	119	5(5), 14(113), 17(1)
PP	136	3(2), 6(10), 11(5), 13(2), 14(106), 15(2), 17(3), 19(1), 20(4), 23(1)	72	6(1), 14(71)	39	14(39)

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); and PP = protein percentage (%).*Significant at genome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.01$; ***Significant at genome-wise FDR $P < 0.001$

Table 4 – 9. Comparison of the identified significant SNPs from single marker LD regression in the Canadian Holstein cattle with previously reported QTL for milk production traits in different dairy cattle breeds.

Trait[1]	BTA	SNP position (Mb)	N-SNPs	Previously reported QTL peak (cM) [2]	Previously reported QTL interval (cM) [2] or gene[3]	Type[4]	Reference
MY	1	57.14-57.70	2	49.6	38.438-60.26	Overlap	(NADESALINGAM *et al.* 2001)
		120.98-133.39	5	110.1	99.155-122.0	Overlap	(VIITALA *et al.* 2003)
				122.0	-	Overlap	(BAGNATO *et al.* 2008)
		138.00, 147.60	2	142.2	-	Close	(BAGNATO *et al.* 2008)
				142	-	Close	(DAETWYLER *et al.* 2008)
	5	14.79	1	89.6	6.9-102.6	Overlap	(BENNEWITZ *et al.* 2004)
		114.71-120.67	4	90.5	80.145-121.7	Overlap	(VIITALA *et al.* 2003)
				106.9	103.1-112.4	Close	(BENNEWITZ *et al.* 2003)
				107.4	103.1-112.4	Close	(BENNEWITZ *et al.* 2003)
	7	17.90, 17.93	2	-	16.75-36.94	Overlap	(LILLEHAMMER *et al.* 2007)
	9	93.90	1	-	-	-	-
	10	15.17	1	-	-	-	-
		34.97	1	-	-	-	-
	11	12.76-13.33	3	-	-	-	-
		41.39	1	-	-	-	-
		105.48	1	112.3	-	Close	(BAGNATO *et al.* 2008)
	12	14.77, 14.80	2	-	-	-	-
	14	0.08-15.52	60	0.5	DGAT1	Overlap	(GRISART *et al.* 2002)
				0.5	DGAT1	Overlap	(SUN *et al.* 2009)
				1	0-5.125	Overlap	(HARDER *et al.* 2006)
				2	0-5.125	Overlap	(THALLER *et al.* 2003)
				3.99	0-5.125	Overlap	(BENNEWITZ *et al.* 2003)
				-	0-5.125	Overlap	(LOOFT *et al.* 2001)
				4	-	Overlap	(DAETWYLER *et al.* 2008)
				5	-	Overlap	(DAETWYLER *et al.* 2008)
				5.1	-	Overlap	(BAGNATO *et al.* 2008)
				5.125	0-25	Overlap	(BOICHARD *et al.* 2003)
				5.125	-	Overlap	(RODRIGUEZ-ZAS *et al.* 2002)
				6	-	Overlap	(DAETWYLER *et al.* 2008)
				10.5	-	Overlap	(BAGNATO *et al.* 2008)
				12	-	Overlap	(DAETWYLER *et al.* 2008)
	17	18.69	1	-	-	-	-
	18	53.44	1	54.7	-	Overlap	(BAGNATO *et al.* 2008)
				65.6	54.713-76.57	Close	(HARDER *et al.* 2006)
	19	37.38	1	36	-	Overlap	(DAETWYLER *et al.* 2008)
				42.5	ATP1B2	Close	(WANG *et al.* 2011)
	21	30.73	1	26.6	12.601-35.89	Overlap	(VIITALA *et al.* 2003)
	27	19.05	1	44.7	14.63-45.253	Overlap	(VIITALA *et al.* 2003)
	28	0.10-1.97	3	-	-	-	-
FY	1	16.28	1	18.6	5.113-23.94	Overlap	(NADESALINGAM *et al.* 2001)
		54.01	1	-	-	-	-
	2	5.16	1	34.1	3.856-38.015	Overlap	(SCHNABEL *et al.* 2005)
		101.23	1	-	0-128.8	Overlap	(LEYVA-BACA *et al.* 2007)
	3	111.26, 112.16	2	-	-	-	-
	5	14.42, 14.50	2	19.1	17.287-41.69	Close	(PLANTE *et al.* 2001)
				-	17.28-32.25	Close	(SCHROOTEN *et al.* 2004)
		86.72-112.67	11	96	80.145-111.5	Overlap	(OLSEN *et al.* 2002)
				103.9	103.1-112.4	Overlap	(BENNEWITZ *et al.* 2003)
				79.9	73.44-91.19	Overlap	(LUND *et al.* 2008)
				-	52.40-90.84	Overlap	(LILLEHAMMER *et al.* 2007)
	6	90.36	1	90.7	CSN3	Overlap	(BOVENHUIS and WELLER 1994)
				90.5	CSN3	Overlap	(BOVENHUIS and WELLER 1994)
				90.7	CSN3	Overlap	(HECK *et al.* 2009)

				86	79.192-96.98	Overlap	(VELMALA *et al.* 1999)
				94.1	IL8	Close	(LEYVA-BACA *et al.* 2007)
				67.40	35.4-90.1	Close	(CHEN *et al.* 2006)
	7	64.10-72.72	3	-	-	-	-
	9	36.58	1	-	-	-	-
		58.11	1	-	-	-	-
	12	57.91, 58.31	2	57.05	20.845-101.9	Overlap	(VIITALA *et al.* 2003)
				62	-	Close	(SCHULMAN *et al.* 2008)
	14	0.05-9.33	61	-10.1	-	Overlap	(SCHULMAN *et al.* 2008)
				0.5	DGAT1	Overlap	(SUN *et al.* 2009)
				0.5	DGAT1	Overlap	(WELLER *et al.* 2003)
				0.5	DGAT1	Overlap	(GRISART *et al.* 2002)
				0.5	DGAT1	Overlap	(NASLUND *et al.* 2008)
				2	0-5.125	Overlap	(THALLER *et al.* 2003)
				3	-	Overlap	(DAETWYLER *et al.* 2008)
				3.3	0-10.501	Overlap	(VIITALA *et al.* 2003)
				3.99	0-5.125	Overlap	(BENNEWITZ *et al.* 2003)
				4.56	0-5.125	Overlap	(BENNEWITZ *et al.* 2003)
				-	0-5.125	Overlap	(LOOFT *et al.* 2001)
				5.125	-	Overlap	(HEYEN *et al.* 1999)
				5.125	0-42	Overlap	(BOICHARD *et al.* 2003)
				6.2	-	Overlap	(ASHWELL *et al.* 2001)
				7.0	0-14.011	Overlap	(ASHWELL *et al.* 2004)
				8.9	0-17.846	Overlap	(WINTER *et al.* 2002)
	15	56.89	1	-	-	-	-
	23	15.60	1	49.2	8.5-51.8	Overlap	(BENNEWITZ *et al.* 2004)
	24	6.11, 7.63	2	-	-	-	-
		48.64	1	-	-	-	-
	25	38.54	1	-	-	-	-
	26	34.58, 37.54	2	-	31.65-37.63	Overlap	(HOGLUND *et al.* 2009)
				23.7	22.862-41.64	Overlap	(PLANTE *et al.* 2001)
				39.4	37.63-53.09	Close	(LUND *et al.* 2008)
				28.6	15.1-35.8	Overlap	(BENNEWITZ *et al.* 2004)
	29	50.40	1	-	-	-	-
PY	1	120.98, 121.08	2	93.5	77.686-122.3	Overlap	(NADESALINGAM *et al.* 2001)
		147.60	1	142.2	-	Close	(HEYEN *et al.* 1999)
				142	-	Close	(DAETWYLER *et al.* 2008)
	5	120.67	1	-	-	-	-
	9	36.58	1	-	-	-	-
		85.20	1	-	-	-	-
	10	15.17	1	-	0-118.8	Overlap	(LUND *et al.* 2008)
				21	-	Close	(DAETWYLER *et al.* 2008)
		102.56	1	-	24.7-127.2	Overlap	(SCHROOTEN *et al.* 2004)
				99	-	Close	(DAETWYLER *et al.* 2008)
	11	40.63, 41.39	2	38	-	Close	(DAETWYLER *et al.* 2008)
	14	0.24-8.13	4	0.5	DGAT1	Overlap	(SUN *et al.* 2009)
				0.5	DGAT1	Overlap	(WELLER *et al.* 2003)
				0.5	DGAT1	Overlap	(GRISART *et al.* 2002)
				2	0-5.125	Overlap	(THALLER *et al.* 2003)
				-	0-5.125	Overlap	(LOOFT *et al.* 2001)
				4	-	Overlap	(DAETWYLER *et al.* 2008)
	16	35.58	1	-	-	-	-
	21	30.73	1	-	-	-	-
	24	58.26	1	53	-	Close	(DAETWYLER *et al.* 2008)
FP	1	15.16	1	19.8	5.113-23.94	Overlap	(NADESALINGAM *et al.* 2001)
				7.6	5.113-23.94	Overlap	(NADESALINGAM *et al.* 2001)
		116.61-151.29	7	-	-	-	-
	3	14.42, 14.78	2	-	-	-	-
		20.13	1	17.36	17.088-32.09	Overlap	(VIITALA *et al.* 2003)
		119.71	1	-	-	-	-
	4	27.73	1	-	-	-	-
		85.68-96.11	4	81.6	72.31-91.19	Overlap	(LINDERSSON *et al.* 1998)

				85.3	52.49-112.7	Overlap	(LINDERSSON *et al.* 1998)
				78	LEP	Close	(GIBLIN *et al.* 2010)
	5	60.37	1	89.6	5.3-104.9	Overlap	(BENNEWITZ *et al.* 2004)
				69	IGF1	Close	(BONAKDAR *et al.* 2010)
		95.54	1	85.5	80.145-90.84	Close	(ASHWELL *et al.* 2004)
				89.6	5.3-104.9	Overlap	(BENNEWITZ *et al.* 2004)
		102.84-115.51	16	103.1	-	Overlap	(HEYEN *et al.* 1999)
				104.1	OLR1	Overlap	(SCHENNINK *et al.* 2009)
				105.4	103.1-112.4	Overlap	(BENNEWITZ *et al.* 2003)
				120	111.59-121.7	Overlap	(OLSEN *et al.* 2002)
	8	0.36	1	-	-	-	-
		94.73	1	-	-	-	-
	10	103.74	1	-	-	-	-
	11	13.33-20.90	3	-	-	-	-
		36.71	1	33.59	-	Close	(ASHWELL *et al.* 1998)
	12	49.50-50.88	4	-	-	-	-
		66.47	1	-	-	-	-
	14	0.05-14.78	151	-0.7	(-0.9)-51.5	Overlap	(BENNEWITZ *et al.* 2004)
				0.5	DGAT1	Overlap	(GRISART *et al.* 2002)
				0.5	DGAT1	Overlap	(WELLER *et al.* 2003)
				0.66	0-10.501	Overlap	(VIITALA *et al.* 2003)
				2	0-5.125	Overlap	(THALLER *et al.* 2003)
				2.5	0-5.125	Overlap	(KUHN *et al.* 2004)
				5.125	0-9	Overlap	(BOICHARD *et al.* 2003)
				5.125	0-33.31	Overlap	(BENNEWITZ *et al.* 2003)
				5.125	-	Overlap	(HEYEN *et al.* 1999)
				6.2	-	Overlap	(ASHWELL *et al.* 2001)
				7.0	0-14.011	Overlap	(ASHWELL *et al.* 2004)
	15	57.43	1	-	-	-	-
	16	60.11	1	-	-	-	-
	17	41.79, 41.81	2	-	-	-	-
		51.84	1	62.3	54.709-92.06	Close	(PLANTE *et al.* 2001)
	18	6.69	1	-	-	-	-
	20	35.61	1	37.7	31.866-43.54	Overlap	(ZHANG *et al.* 1998)
				34	GHR	Close	(WATERS *et al.* 2011)
				42.9	GHR	Close	(SUN *et al.* 2009)
				31.85	20.165-43.54	Overlap	(ARRANZ *et al.* 1998)
				43.39	37.496-42.29	Close	(BLOTT *et al.* 2003)
	26	48.82	1	58.7	42.48-60.476	Overlap	(PLANTE *et al.* 2001)
				51.5	22.862-52.45	Overlap	(VIITALA *et al.* 2003)
				-	43.22-53.09	Overlap	(HOGLUND *et al.* 2009)
PP	3	20.13	1	27.41	6-32	Overlap	(BOICHARD *et al.* 2003)
				25.0	22.617-27.41	Close	(ASHWELL *et al.* 2004)
				17.36	17.088-32.09	Overlap	(VIITALA *et al.* 2003)
				-	21.7-71.6	Close	(HEYEN *et al.* 1999)
		38.71	1	53.5	34.629-54.2	Overlap	(PLANTE *et al.* 2001)
				-	21.7-71.6	Overlap	(HEYEN *et al.* 1999)
				43.29	-	Close	(HEYEN *et al.* 1999)
				43.29	-	Close	(RODRIGUEZ-ZAS *et al.* 2002)
	6	19.64, 21.06	2	15.36	-	Close	(SCHROOTEN *et al.* 2004)
		47.29-50.65, 60.90-71.19	8	49.4	17.4-83.4	Overlap	(BENNEWITZ *et al.* 2004)
				53.72	35.39-59.73	Overlap	(SPELMAN *et al.* 1996)
				42.2	29.8-50.9	Overlap	(FREYER *et al.* 2002)
				39.4	36-53.8	Overlap	(FREYER *et al.* 2002)
				69.12	-	Overlap	(MEI *et al.* 2009)
				61.4	53.724-70.74	Overlap	(NADESALINGAM *et al.* 2001)
				71	70.741-72.43	Overlap	(VELMALA *et al.* 1999)
				50.9	43.936-57.9	Overlap	(OLSEN *et al.* 2004)
				93.1	47.3-98	Overlap	(CHEN *et al.* 2006)
				67.4	-	Overlap	(BAGNATO *et al.* 2008)
				63.9	-	Overlap	(BAGNATO *et al.* 2008)

			66.5	53.724-59.73	Overlap	(ZHANG *et al.* 1998)
			50.8	47.82-53.724	Overlap	(ASHWELL *et al.* 2004)
			52.4	35.398-53.72	Overlap	(VIITALA *et al.* 2003)
			72.4	-	Close	(BAGNATO *et al.* 2008)
			53.72	-	Close	(LIPKIN *et al.* 1998)
			53.72	-	Close	(COHEN-ZINDER *et al.* 2005)
			53.72	-	Close	(MOSIG *et al.* 2001)
			53.72	-	Close	(RODRIGUEZ-ZAS *et al.* 2002)
			73	-	Close	(FREYER *et al.* 2002)
			43.93	-	Close	(GAO *et al.* 2009)
			38.9	MED28	Close	(COHEN-ZINDER *et al.* 2005)
			38.9	LAP3	Close	(COHEN-ZINDER *et al.* 2005)
			38.5	SPP1	Close	(ALAIN *et al.* 2009)
			37.7	FAM13A1	Close	(COHEN-ZINDER *et al.* 2005)
			38.1	HERC6	Close	(COHEN-ZINDER *et al.* 2005)
			38.3	ABCG2	Close	(COHEN-ZINDER *et al.* 2005)
			38.4	PKD2	Close	(COHEN-ZINDER *et al.* 2005)
			38.4	SPP1	Close	(COHEN-ZINDER *et al.* 2005)
11	9.43	1	-	-	-	-
	46.64-47.53	4	-	-	-	-
13	11.83	1	8.993	-	Close	(SCHROOTEN *et al.* 2004)
	45.33	1	57.5	41.728-73.29	Overlap	(ASHWELL *et al.* 2004)
14	0.05-14.78	105	-0.7	(-0.9)-59.7	Overlap	(BENNEWITZ *et al.* 2004)
			0.0	-	Overlap	(BAGNATO *et al.* 2008)
			0.5	DGAT1	Overlap	(WELLER *et al.* 2003)
			0.5	DGAT1	Overlap	(GRISART *et al.* 2002)
			0.5	DGAT1	Overlap	(NASLUND *et al.* 2008)
			2	0-5.125	Overlap	(THALLER *et al.* 2003)
			5.1	-	Overlap	(BAGNATO *et al.* 2008)
			5.125	0-70	Overlap	(BOICHARD *et al.* 2003)
			8.1	-	Overlap	(BAGNATO *et al.* 2008)
			21.8	5.125-33.31	Overlap	(BENNEWITZ *et al.* 2003)
	59.77	1	60.5	51.941-69.01	Overlap	(SCHNABEL *et al.* 2005)
			43.63	10.501-60.68	Overlap	(VIITALA *et al.* 2003)
15	26.54, 26.56	2	-	-	-	-
17	0.47	1	-	-	-	-
	7.26	1	-	-	-	-
	41.81	1	-	-	-	-
19	53.16	1	62.1	4.1-102.6	Overlap	(BENNEWITZ *et al.* 2004)
20	26.56-29.24, 35.43	4	31.85	20.165-43.54	Overlap	(ARRANZ *et al.* 1998)
			31.86	21-50	Overlap	(BOICHARD *et al.* 2003)
			42.7	15.5-68	Overlap	(BENNEWITZ *et al.* 2004)
			34	GHR	Overlap	(REARDON *et al.* 2010)
			34	GHR	Overlap	(WATERS *et al.* 2011)
			37.7	31.866-43.54	Overlap	(ZHANG *et al.* 1998)
			42.9	GHR	Close	(SUN *et al.* 2009)
23	46.01	1	42.9	-	Close	(BAGNATO *et al.* 2008)
			52.3	-	Close	(BAGNATO *et al.* 2008)

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); and PP = protein percentage (%). [2]Previous QTL information was obtained from CattleQTLdb (www.animalgenome.org/cgi-bin/QTLdb/BT/index). [3]Genes reported within the QTL region: *ABCG2*: ATP-binding cassette, sub-family G, member 2; *ATP1B2*: ATPase, Na+/K+ transporting, beta 2 polypeptide; *CSN3*: casein kappa; *DGAT1*: diacylglycerol O-acyltransferase homolog 1; *FAM13A1*: family with sequence similarity 13, member A1; *GHR*: growth hormone receptor; *HERC6*: HECT and RLD domain containing E3 ubiquitin protein ligase family member 6; *IGF1*: insulin-like growth factor 1; *IL8*: interleukin 8; *LAP3*: leucine aminopeptidase 3; *LEP*: leptin; *MED28*: mediator complex subunit 28; *OLR1*: oxidized low density lipoprotein receptor 1; *PKD2*: polycystic kidney disease 2; *SPP1*: secreted phosphoprotein 1. [4]The type of concordance with QTL in literature. For SNPs without overlapping QTL, the closest previously identified QTL was presented.

Table 4 – 10. Gene networks among positional candidate genes related to milk production traits in the Canadian Holstein cattle. Positional candidate genes within 0.5 Mb windows of significant SNPs (at genome-wise FDR $P < 0.05$) from the single marker LD regression for milk production traits were considered in the functional clustering analyses using the Ingenuity Pathway Analysis (IPA) database.

Trait	Functional network	Moleculars in network	N
MY	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry	ABCB10, ABCC3, ABHD10, ACSBG1, AP2S1, ARMC8, C22orf32, CALML4, CCDC77, CCDC94, CHRAC1, COMMD5, CPSF1, ERC1, FSIP1, GPT, HIF3A, HLTF, HPS3, KDM4B, KIAA0196, LRRC6, LRRC14, MPND, MRPL27, MTRF1L, NAGA, NPAS1, NUP214, OPA3, OPLAH, PLIN5, POLR3H, PWP2, SLC1A5, SLC25A17, SLC35A5, SLC9A10, SNRPD2, STRN4, SYMPK, TCF20, TM4SF4, TOB2, TONSL, TRMT12, TSPO, URB2, XPNPEP3, ZC3H4	50
PY	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry	C22orf32, CALML4, COMMD5,FEM1B, FOXH1, GPT, HPS3, LRRC3, LRRC6, NAGA, PWP2, TM4SF4, TXNL1	13
PP	Small Molecule Biochemistry, Carbohydrate Metabolism, Post-Translational Modification	AATK, AZI1, AZIN1, C17orf70, CHMP6, CHRNA9, DKK2, ETFDH, FIP1L1, FOXH1, FSCN2, FXYD2, GCC2, GCM2, GRINA, KHDRBS3, LNX1, LY6E, N4BP2, NPTX1, NSMCE2, OC90, OPLAH, OXCT1, PCSK7, PITRM1, PRMT6, RECQL4, RNF139, RPL37, SCFD2, SH3D19, SLC5A7, SNX18, ST3GAL1, Sult1c2, SULT1C4*, TGFBRAP1, Tssk5, TSTA3	40

[1]MY = milk yield (kg); PY = protein yield (kg); and PP = protein percentage (%).

Table 4 – 11. Posterior means of variance components explained by whole genome SNPs for milk production traits in the Canadian Holstein cattle using the BayesB method.

Trait[1]	Genetic variance	Residual variance	Total variance	Proportion of phenotypic variance explained by all SNPs
MY	460237	658634	1118871	0.41
FY	684.966	753.189	1438.15	0.48
PY	334.962	535.81	870.773	0.38
FP	0.08910	0.03534	0.1244	0.72
PP	0.01246	0.01337	0.02583	0.48

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); and PP = protein percentage (%).

Table 4 – 12. Summary of significantly ($P < 0.2$) associated QTL regions and candidate genes for milk production traits in the Canadian Holstein cattle identified using the BayesB method.

Trait[1]	N-associated regions	BTA	Functionally relevant genes in associated regions (BTA)[2]	Previously identified overlapping QTL (BTA)[3]
MY	1	14	RPL8(14), GPT(14), DGAT1(14), MAPK15(14)	MY(14), FY(14), PY(14), FP(14), PP(14)
FY	2	5, 14	RPL8(14), GPT(14), DGAT1(14), MAPK15(14), GOLT1B(5), IAPP(5)	MY(14), FY(14), PY(14), FP(14), PP(14), FY(5)
PY	1	14	RPL8(14), GPT(14), DGAT1(14), MAPK15(14)	MY(14), FY(14), PY(14), FP(14), PP(14)
FP	3	14, 20	RPL8(14), GPT(14), DGAT1(14), MAPK15(14)	MY(14), FY(14), PY(14), FP(14), PP(14), FP(20)
PP	2	10, 14	RPL8(14), GPT(14), DGAT1(14), MAPK15(14), PPIB(10), SNX1(10), HERC1(10)	MY(14), FY(14), PY(14), FP(14), PP(14)

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); PP = protein percentage (%). [2] *DGAT1* = diacylglycerol O-acyltransferase 1; *GOLT1B* = golgi transport 1B; *GPT* = glutamic-pyruvate transaminase (alanine aminotransferase); *HERC1* = HECT and RLD domain containing E3 ubiquitin protein ligase family member 1; *IAPP* = islet amyloid polypeptide; *MAPK15* = mitogen-activated protein kinase 15; *PPIB* = peptidylprolyl isomerase B (cyclophilin B); *RPL8* = ribosomal protein L8; *SNX1* = sorting nexin 1. [3]The QTL information was obtained from CattleQTLdb (http://www.animalgenome.org/cgi-bin/gbrowse/cattle/).

Table 4 - 13. Detailed information for QTL regions associated ($P < 0.2$) with milk production traits in the Canadian Holstein cattle identified using the BayesB method.

Trait[1]	BTA	Start[2]	End[2]	N-SNPs[3]	%Var[4]	*P*-value	Genes at SNP window[5]	Previously reported QTL at the SNP window (Reference)[6]	Previously related traits
MY	14	50,872	996,982	17	10.5	<0.001	C14H8orf33, ZNF34, RPL8*, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT*, PPP1R16A, FOXH1, KIFC2, CYHR1, CYHR1, TONSL, VPS28, CPSF1, ADCK5, GPR172A, SCRT1, DGAT1*, HSF1, HEATR7A, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, NRBP2, PUF60, MAPK15*, CCDC166, TSTA3, PYCRL, EEF1D, NAPRT1, GSDMD	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (SUN *et al.* 2009)	Milk yield
FY	14	50,872	996,982	17	18.08	<0.001	C14H8orf33, ZNF34, RPL8*, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT*, PPP1R16A, FOXH1, KIFC2, CYHR1, CYHR1, TONSL, VPS28, CPSF1, ADCK5, GPR172A, SCRT1, DGAT1*, HSF1, HEATR7A, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, NRBP2, PUF60, MAPK15*, CCDC166, TSTA3, PYCRL, EEF1D, NAPRT1, GSDMD	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (SUN *et al.* 2009) 0.5 cM (DGAT1) (NASLUND *et al.* 2008)	Milk fat yield
	5	95,068,492	95,878,365	10	2.6	0.16	GYS2, C5H12orf39, GOLT1B*, RECQL, PYROXD1, IAPP*, SLCO1A2, SLCO1B3, SLCO1C1	96 (80.145-111.5) cM (OLSEN *et al.* 2002)	Milk fat yield
PY	14	50,872	996,982	17	2.32	0.10	C14H8orf33, ZNF34, RPL8*, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT*, PPP1R16A, FOXH1, KIFC2, CYHR1, CYHR1, TONSL, VPS28, CPSF1, ADCK5, GPR172A, SCRT1, DGAT1*, HSF1, HEATR7A, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, NRBP2, PUF60, MAPK15*, CCDC166, TSTA3, PYCRL, EEF1D, NAPRT1, GSDMD	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (SUN *et al.* 2009)	Milk protein yield
FP	14	50,872	996,982	17	38.94	<0.001	C14H8orf33, ZNF34, RPL8*, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT*, PPP1R16A, FOXH1, KIFC2, CYHR1, CYHR1, TONSL, VPS28, CPSF1, ADCK5, GPR172A, SCRT1, DGAT1*, HSF1, HEATR7A, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, NRBP2, PUF60, MAPK15*, CCDC166, TSTA3, PYCRL, EEF1D, NAPRT1, GSDMD	0.66 (0-10.501) cM (VIITALA *et al.* 2003) 0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003)	Milk fat percentage

	14	3,018,726	3,993,200	21	2.61	0.04	LOC618755	2 (0-5.125) cM (THALLER *et al.* 2003) 2.5 (0-5.125) cM (KUHN *et al.* 2004) 5.125 (0-9) cM (BOICHARD *et al.* 2003)	Milk fat percentage
	20	35,153,927	35,702,448	8	1.28	0.20	PLCXD3, C6, C7	37.7 (31.866-43.54) cM (ZHANG *et al.* 1998) 34 cM (GHR) (WATERS *et al.* 2011) 31.85 (20.165-43.54) cM (ARRANZ *et al.* 1998)	Milk fat percentage
PP	14	50,872	996,982	17	10.66	<0.001	C14H8orf33, ZNF34, RPL8*, COMMD5, ARHGAP39, C14H8orf82, LRRC24, LRRC14, RECQL4, MFSD3, GPT*, PPP1R16A, FOXH1, KIFC2, CYHR1, CYHR1, TONSL, VPS28, CPSF1, ADCK5, GPR172A, SCRT1, DGAT1*, HSF1, HEATR7A, MAF1, SHARPIN, CYC1, GPAA1, EXOSC4, OPLAH, NRBP2, PUF60, MAPK15*, CCDC166, TSTA3, PYCRL, EEF1D, NAPRT1, GSDMD	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (NASLUND *et al.* 2008)	Milk protein percentage
	10	46,001,270	46,950,341	16	1.78	0.19	TRIP4, PAF, CSNK1G1, PPIB*, SNX22, SNX1*, FAM96A, HERC1*, FBXL22, USP3	--	--

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); and PP = protein percentage (%). [2]Start and end position of the associated SNP windows (bp) based on the Btau4.0 reference assembly (http://www.hgsc.bcm.tmc.edu/ftp-archive/Btaurus/fasta/Btau20070913-freeze/). [3]The number of SNPs located within each associated SNP window. [4]The proportion of variance among the de-regressed EBVs explained by the SNP window. [5]Genes located within the SNP window and the genes with potential functions related to milk production traits were indicated with an asterisk (*). [6]Previously identified QTL at the associated QTL region with QTL peak location, QTL span and reference. Previously reported candidate genes within these regions are also presented. The QTL information was obtained from CattleQTLdb (www.animalgenome.org/cgi-bin/QTLdb/BT/index). *DGAT1*: diacylglycerol O-acyltransferase 1; *GHR*: growth hormone receptor.

Table 4 - 14. Comparison of the identified associated SNPs and QTL regions between single marker LD regression and Bayesian regression for milk production traits in the Canadian Holstein cattle.

Trait[1]	BTA	BayesB regression		Single marker LD regression		Previously reported QTL at the SNP window (Reference)[4]	Previously related traits
		SNP window (Mb)[2]	N-SNPs[2]	SNP position (Mb)[3]	N-SNPs[3]		
MY	14	0.05-1.00	17	0.08-15.52	60	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (SUN *et al.* 2009)	Milk yield
FY	14	0.05-1.00	17	0.05-9.33	61	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (SUN *et al.* 2009) 0.5 cM (DGAT1) (NASLUND *et al.* 2008)	Milk fat yield
	5	95.07-95.88	10	86.72-112.67	11	96 (80.145-111.5) cM (OLSEN *et al.* 2002)	Milk fat yield
PY	14	0.05-1.00	17	0.24-8.13	4	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (SUN *et al.* 2009)	Milk protein yield
FP	14	0.05-1.00	17	0.05-14.78	151	0.66 (0-10.501) cM (VIITALA *et al.* 2003) 0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003)	Milk fat percentage
	14	3.02-3.99	21	0.05-14.78	151	2 (0-5.125) cM (THALLER *et al.* 2003) 2.5 (0-5.125) cM (KUHN *et al.* 2004) 5.125 (0-9) cM (BOICHARD *et al.* 2003)	Milk fat percentage
	20	35.15-35.70	8	35.61	1	37.7 (31.866-43.54) cM (ZHANG *et al.* 1998) 34 cM (GHR) (WATERS *et al.* 2011) 31.85 (20.165-43.54) cM (ARRANZ *et al.* 1998)	Milk fat percentage
PP	14	0.05-1.00	17	0.05-14.78	105	0.5 cM (DGAT1) (GRISART *et al.* 2002) 0.5 cM (DGAT1) (WELLER *et al.* 2003) 0.5 cM (DGAT1) (NASLUND *et al.* 2008)	Milk protein percentage
	10	46.00-46.95	16	-	-	--	--

[1]MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); and PP = protein percentage (%). [2]Physical position of the associated SNP windows (Mb) using the BayesB regression and the number of SNPs located within the SNP window. [3]Physical position of the significantly associated SNP (Mb) using single marker LD regression and the number of SNPs within the region. [4]Previously identified QTL at the associated QTL region with QTL peak, QTL span and reference. Previously reported candidate genes within these regions are also presented. The QTL information was obtained from CattleQTLdb (www.animalgenome.org/cgi-bin/QTLdb/BT/index). *DGAT1*: diacylglycerol O-acyltransferase 1; *GHR*: growth hormone receptor.

Figures

Figure 4 – 1A

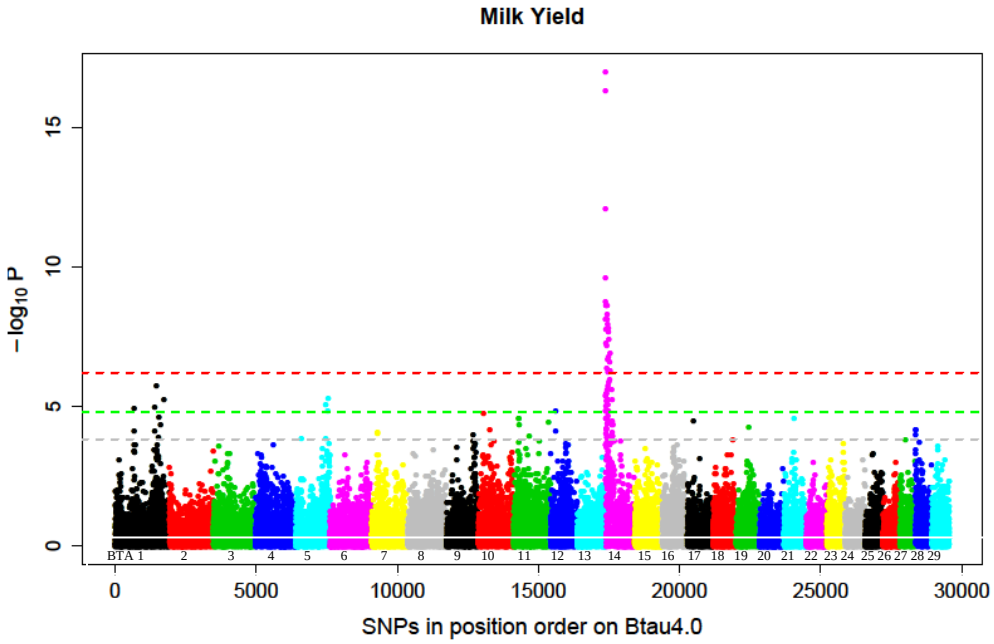

Figure 4 – 1B

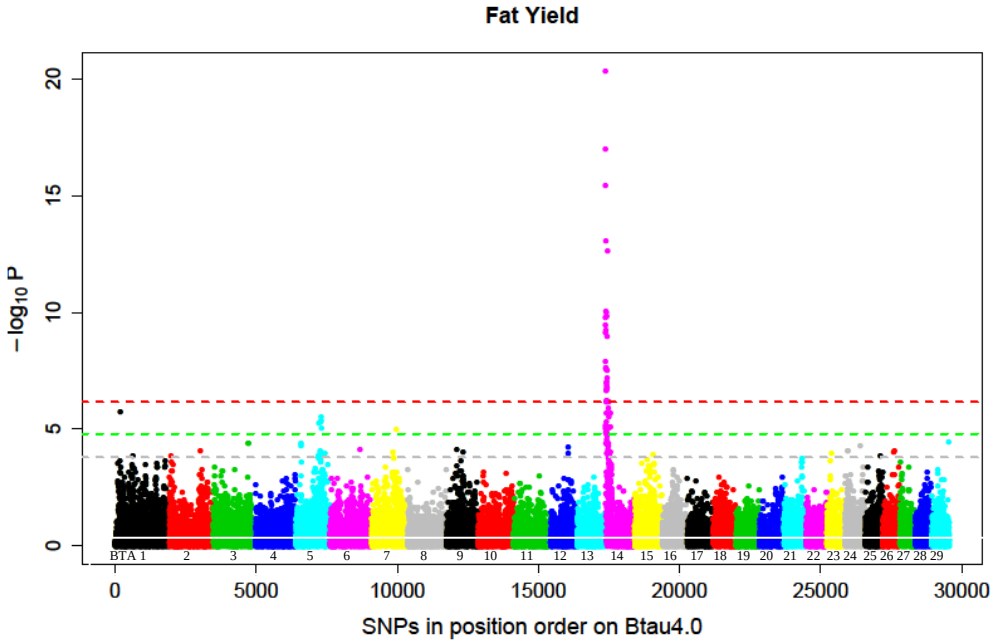

Figure 4 – 1C

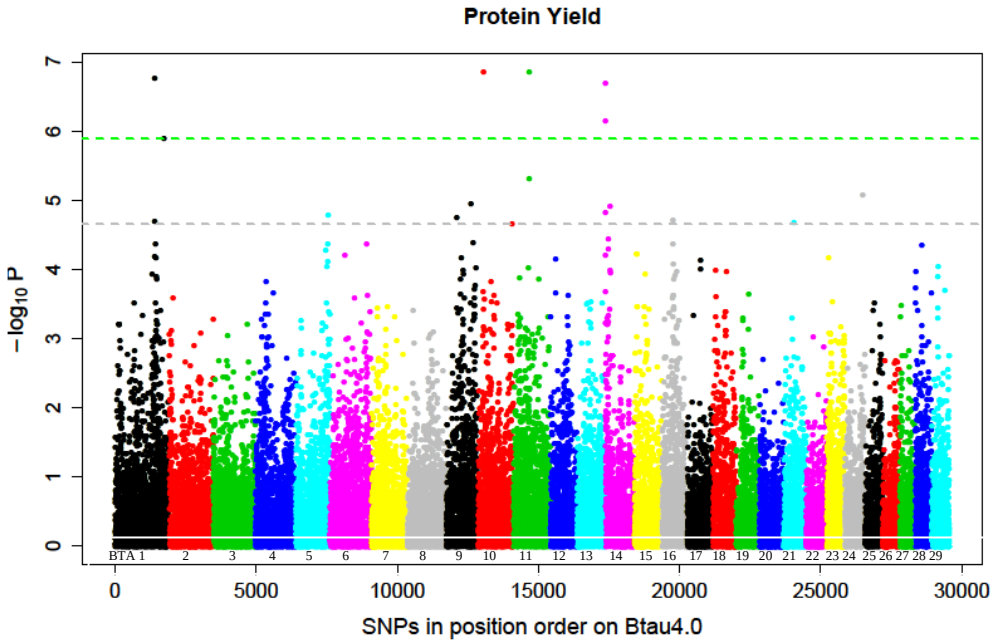

Figure 4 – 1D

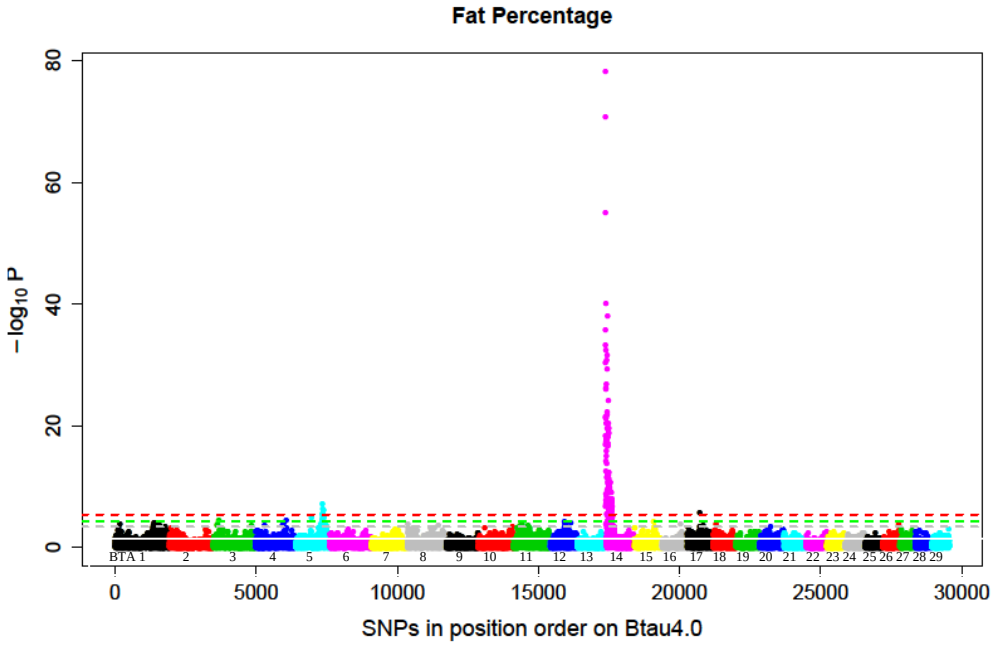

Figure 4 – 1E

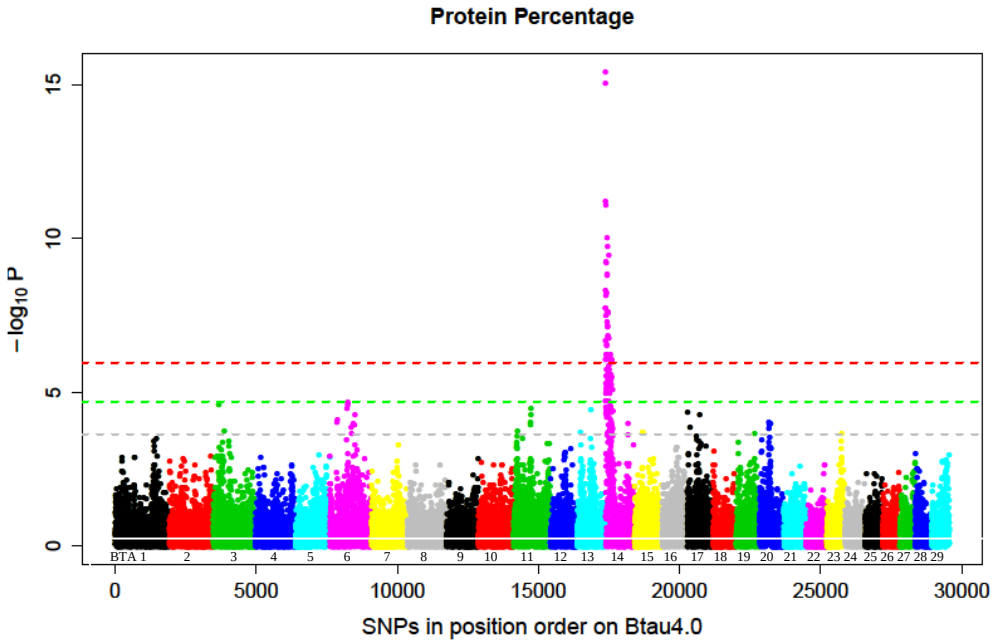

Figure 4 – 1. Whole-genome association analyses for milk production traits in Canadian Holstein cattle using single marker LD regression: A) milk yield (MY); B) fat yield (FY); C) protein yield (PY); D) fat percentage (FP); E) protein percentage (PP). The X-axis is the genomic location of the SNP on Btau4.0 (Mb). The Y-axis sows the negative logarithm (base 10) of the *P*-values for the SNP allele substitution effects. Different shades represent SNPs on different chromosomes from BTA1 (left) to BTA29 (right). The dashed lines represent the genome-wise adjusted thresholds for FDR $P < 0.05$ (grey line), $P < 0.01$ (green line) and $P < 0.001$ (red line).

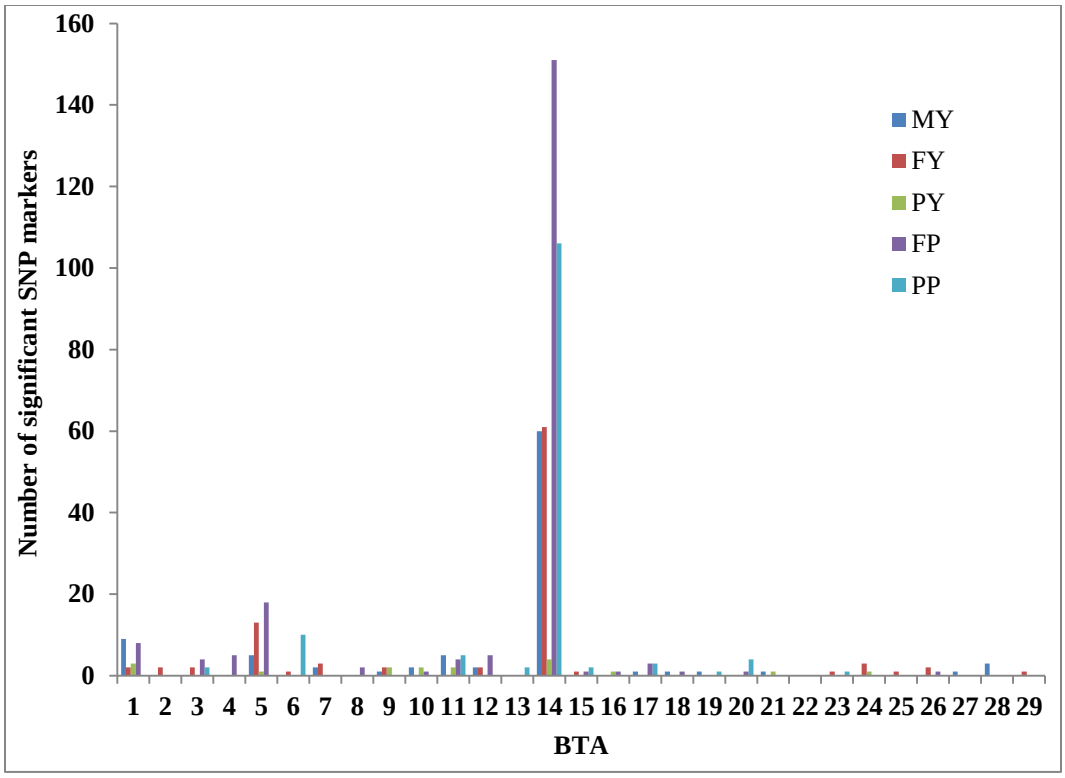

Figure 4 – 2. Chromosomal distribution of significant SNPs at genome-wise 5% FDR threshold for milk production traits in the Canadian Holstein cattle using single marker LD regression. MY = milk yield; FY = fat yield; PY = protein yield; FP = fat percentage; PP = protein percentage.

Figure 4 – 3A

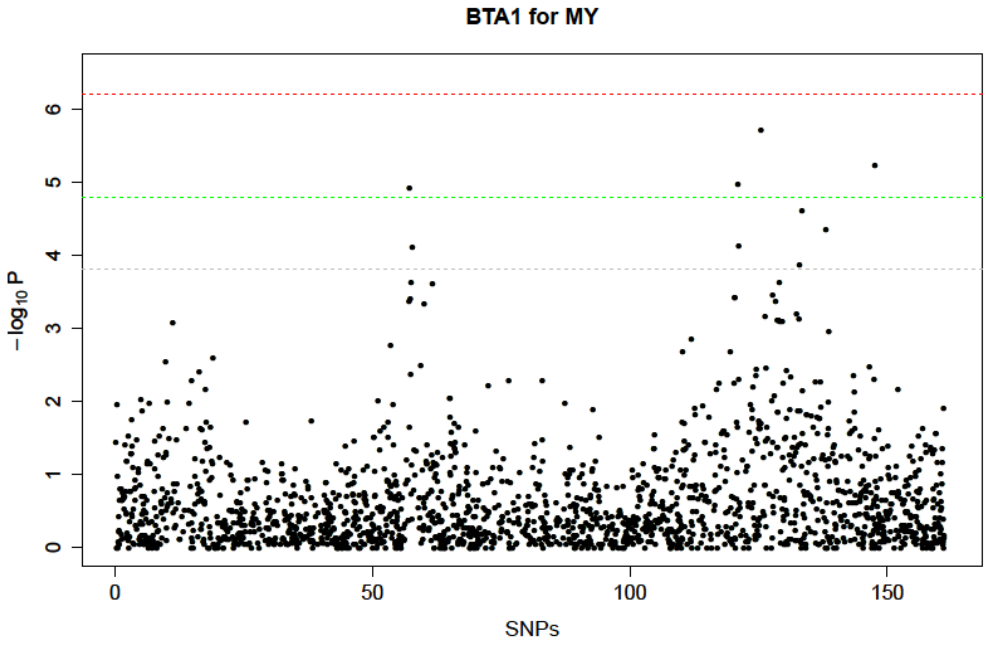

Figure 4 – 3B

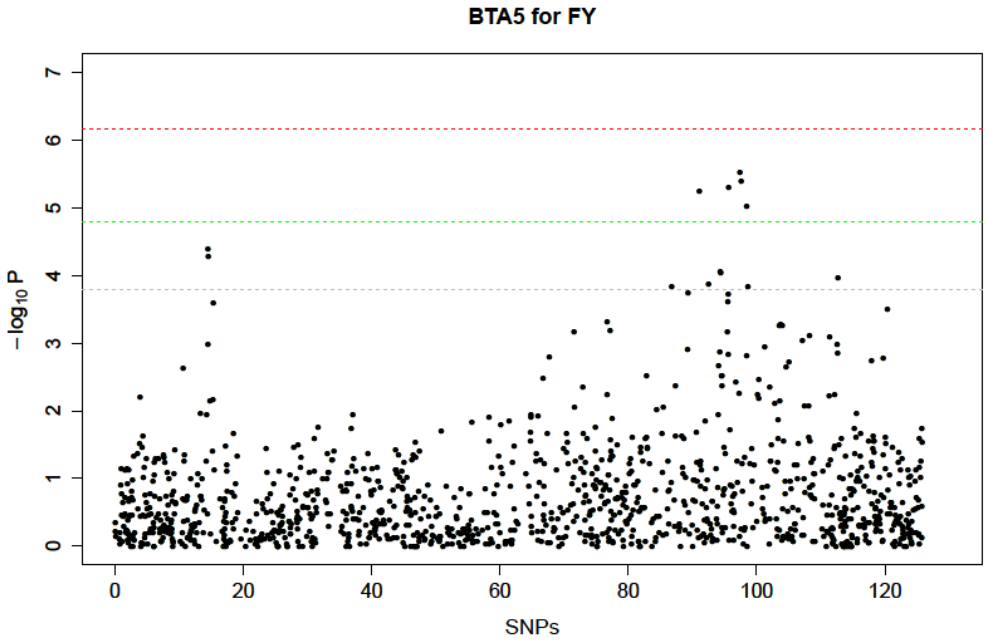

Figure 4 – 3C

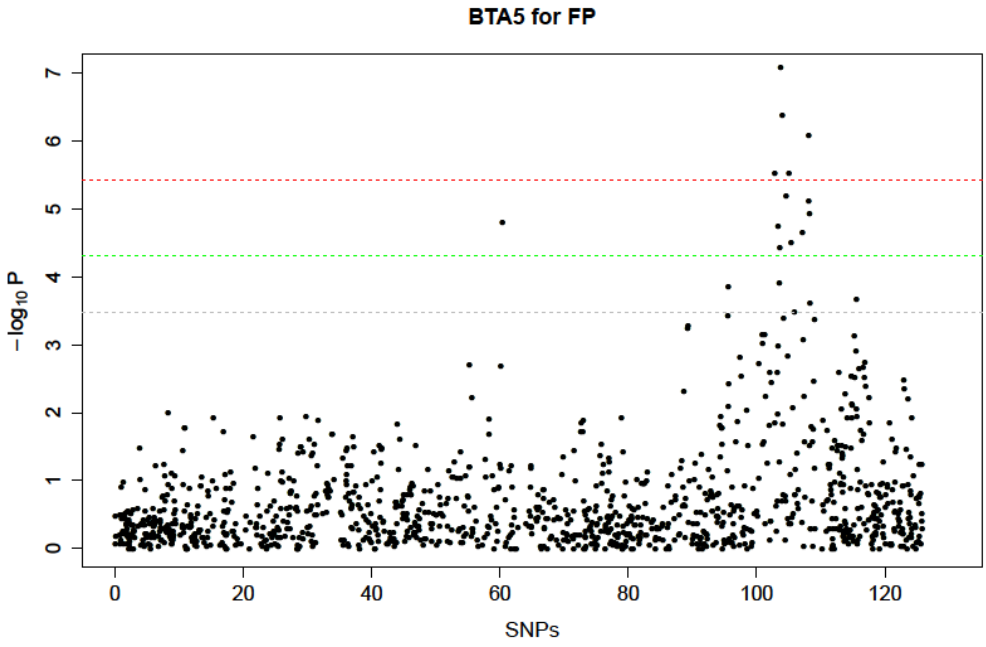

Figure 4 – 3D

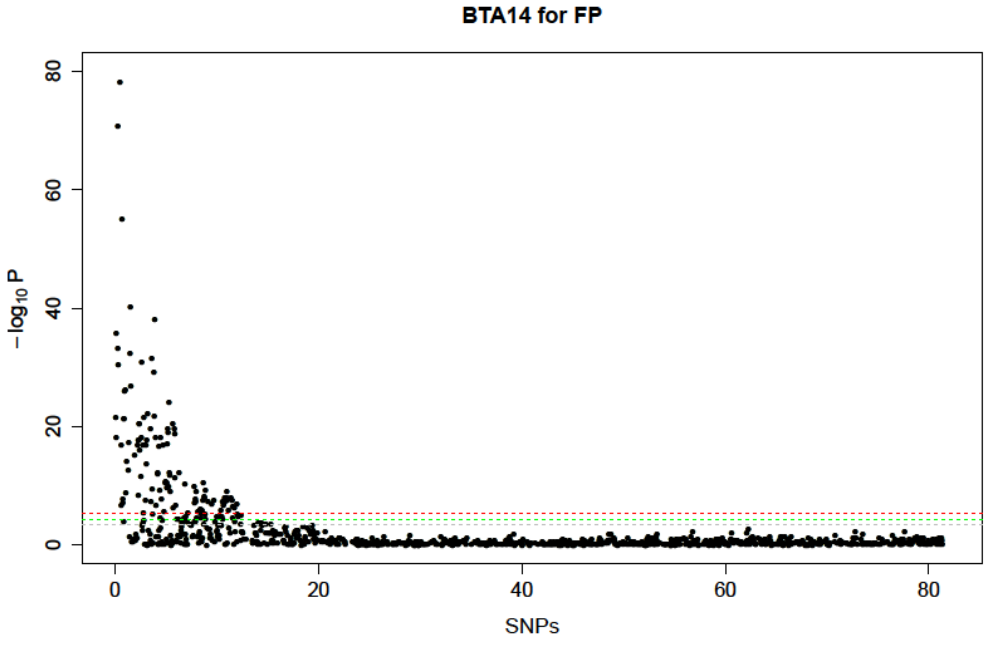

Figure 4 – 3E

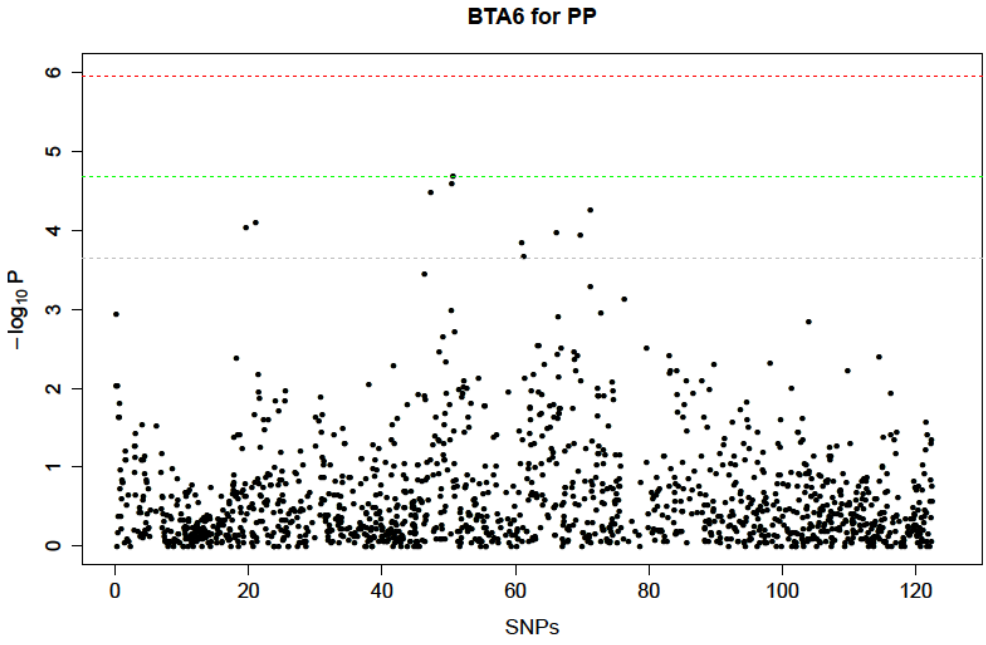

Figure 4 – 3. Examples of chromosomes with clustered significant SNPs identified for milk production traits in the Canadian Holstein cattle using single marker LD regression: A) BTA1 for milk yield (MY); B) BTA5 for fat yield (FY); C) BTA5 for fat percentage (FP); D) BTA14 for FP; E) BTA6 for protein percentage (PP). The X-axis shows the physical location of the SNP on each chromosome (Mb). The Y-axis shows the negative logarithm (base 10) of the *P*-values for the SNP allele substitution effects. The dashed lines represent the genome-wise adjusted thresholds for FDR $P < 0.05$ (grey line), $P < 0.01$ (green line) and $P < 0.001$ (red line).

Figure 4 – 4A

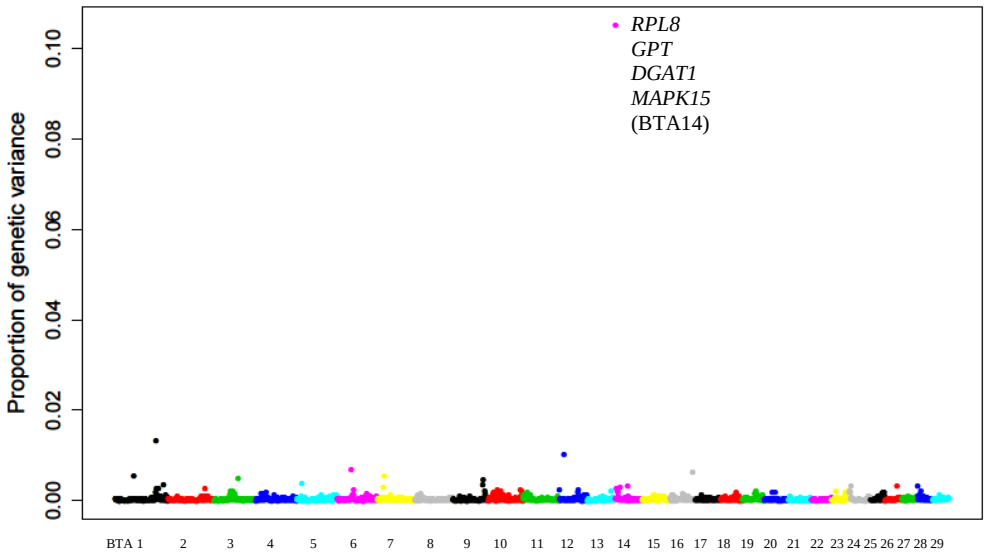

Figure 4 – 4B

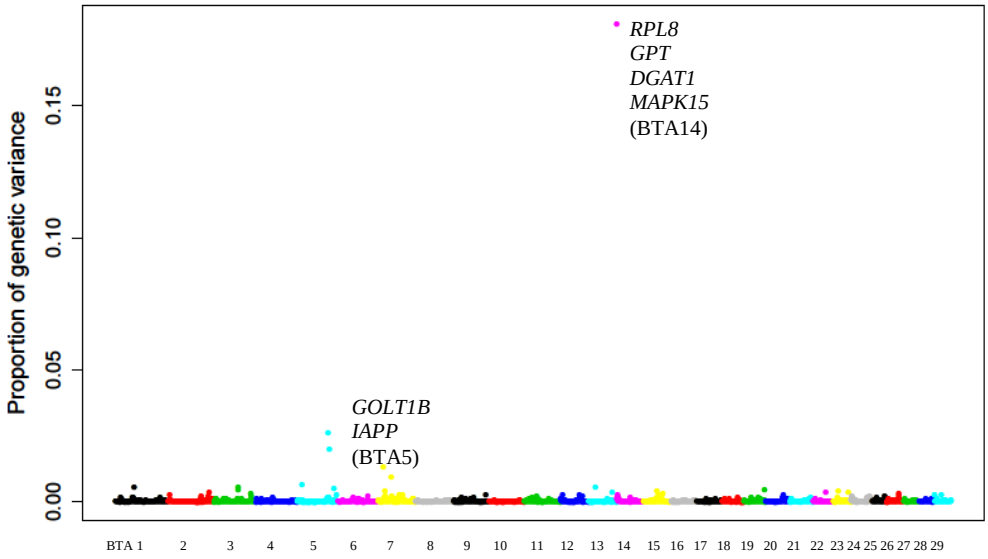

Figure 4 – 4C

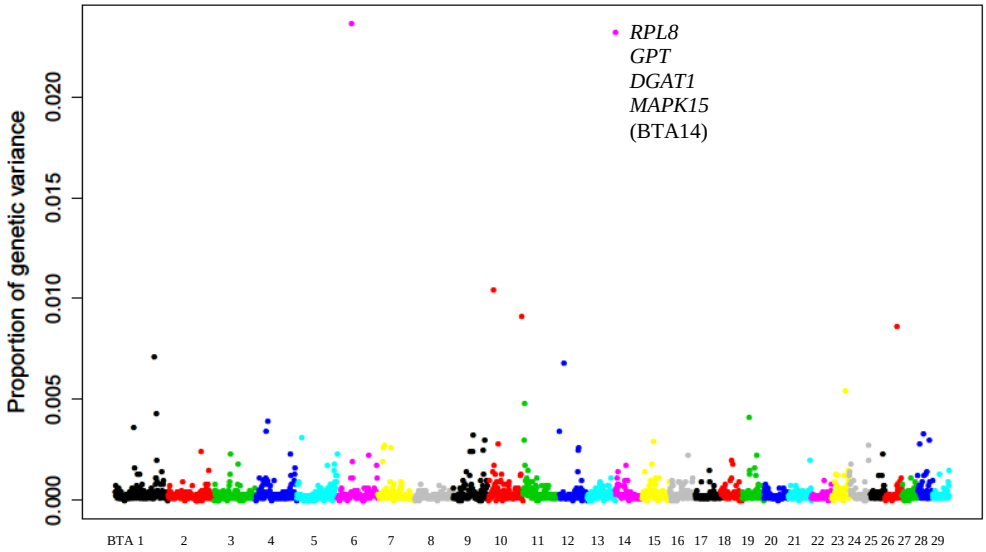

Figure 4 – 4D

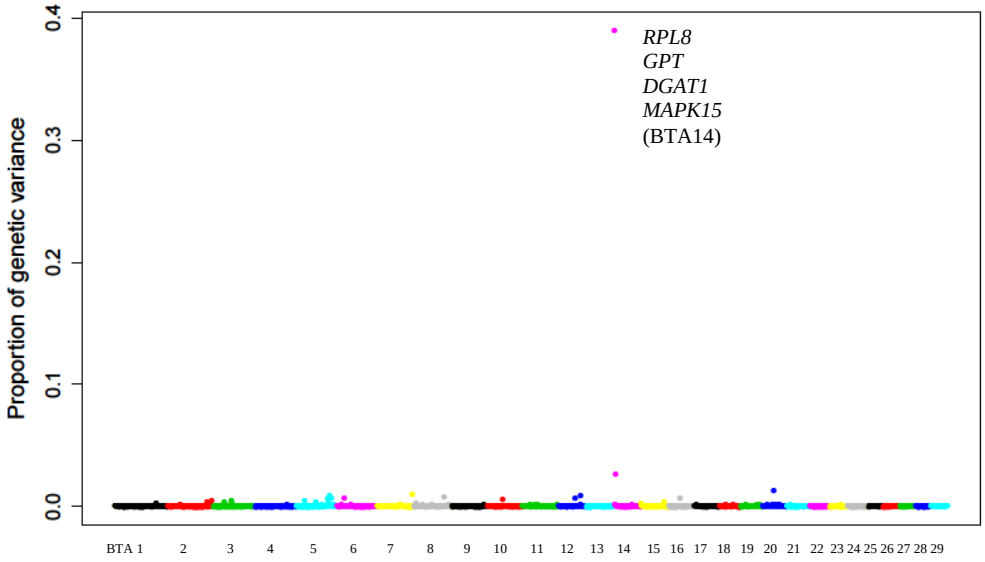

Figure 4 – 4E

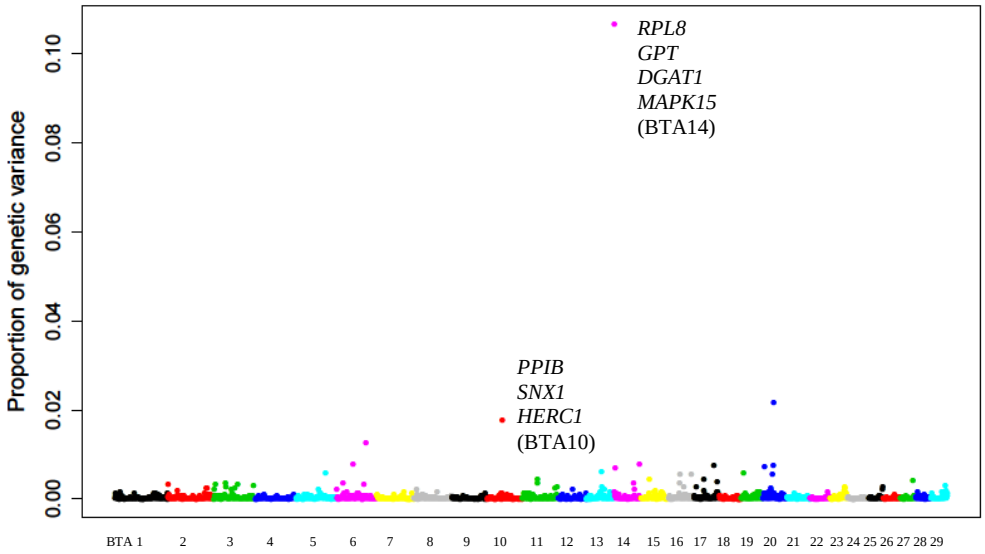

Figure 4 – 4. Proportion of genetic variance explained by each window of 1Mb consecutive SNP markers across the genome for milk production traits in Canadian Holstein cattle using the BayesB method: A) milk yield (MY); B) fat yield (FY); C) protein yield (PY); D) fat percentage (FP); E) protein percentage (PP). The X-axis shows the SNP marker position in genome order according to Btau4.0 (Mb), and the Y-axis shows the proportion of genetic variance contributed by the 1 Mb SNP window (the exact candidate regions). Different colours represent SNPs on different chromosomes from BTA1 (left) to BTA29 (right). *DGAT1* = diacylglycerol O-acyltransferase 1; *GOLT1B* = golgi transport 1B; *GPT* = glutamic-pyruvate transaminase (alanine aminotransferase); *HERC1* = HECT and RLD domain containing E3 ubiquitin protein ligase family member 1; *IAPP* = islet amyloid polypeptide; *MAPK15* = mitogen-activated protein kinase 15; *PPIB* = peptidylprolyl isomerase B (cyclophilin B); *RPL8* = ribosomal protein L8; *SNX1* = sorting nexin 1.

Chapter 5. Whole-Genome Association Study for Ultrasound and Carcass Traits in Hybrid Beef Cattle

5.1. Introduction

Carcass quantity and quality determine much of the economic value of beef cattle and thus these traits are of great importance to the beef industry. Carcass-related traits in cattle are moderately to highly heritable but can only be measured post-slaughter. These traits are among those that would benefit most from the use of genetic marker or gene information in the prediction of individual genetic merit (DAVIS and DENISE 1998). Consequently, it is important to identify genetic markers linked to, or genes that influence, beef carcass quantity and quality.

QTL mapping is the first step towards identification of the genes and causal polymorphisms (SEATON *et al.* 2002), and QTL associated with beef carcass traits have been described (ABE *et al.* 2008; BOLORMAA *et al.* 2011; CASAS *et al.* 2003; GUTIERREZ-GIL *et al.* 2009; MCCLURE *et al.* 2010; TAKASUGA *et al.* 2007). However, few genes have been specifically linked to carcass traits. For example, a QTL for beef carcass weight was fine mapped on BTA6 and a non-synonymous mutation was identified, in the gene *NCAPG*, in a Japanese Brown × Black population (SETOGUCHI *et al.* 2009). Variation in the *MSTN* gene, located within a QTL close to the centromere on BTA2, accounts for most of the observed variation in muscle mass and fat deposition in double-muscled cattle populations (CASAS *et al.* 1998).

High-throughput genotyping of SNPs allows genome-wide evaluation of QTL with a resolution of 1 – 2 Mb (SELLNER *et al.* 2007), and this will assist in the identification of additional functional candidate genes. A few genome-wide association studies (GWAS) have been carried out to identify SNPs associated with beef carcass traits, using single marker regression and the BovineSNP50 BeadChip (BOLORMAA *et al.* 2011; KIM *et al.* 2011; LEE *et al.* 2010b). However, whole genome association studies using high density SNP chips have not been reported for beef carcass traits in the hybrid beef cattle used in the conventional Canadian commercial beef production system. Recently, genomic selection methodologies (FERNANDO and GARRICK 2009; MEUWISSEN *et al.* 2001) for predicting the genetic merit of animals were successfully applied to GWAS for the discovery of candidate genes in swine (FAN *et al.* 2011; ONTERU *et al.* 2011). These methods can also be used in beef cattle to identify QTL regions or genes related to carcass traits.

The objective of this study was to identify candidate SNPs or genes strongly associated with beef carcass traits in a Canadian hybrid beef population. Genetic associations of SNPs on the BovineSNP50 BeadChip with carcass traits were investigated using two different methodologies: single marker LD regression and Bayesian regression. The results from the two methods were compared with previously reported QTL and potential candidate genes within the associated genomic regions were identified based on previous association and functional studies.

5.2. Materials and Methods

5.2.1. Animals and Traits

The hybrid beef steers used for this study are described in Chapter 3. Five carcass merit traits including carcass weight (CWT), carcass average backfat (CABF), carcass ribeye area (CREA), carcass grade fat (CGF), carcass lean meat yield (CLMY) and three ultrasound measurement traits including ultrasound marbling (UMAR), ultrasound backfat (UBF) at the 12th to 13th ribs and ultrasound ribeye area (UREA) were considered in this study. Briefly, UREA, UBF, and UMAR were recorded during feeding tests at 28-day intervals for a period of approximately 90 days using an Aloka 500V realtime ultrasound with a 17-cm, 3.5-MHz linear array transducer (Overseas Monitor Corporation Ltd., Richmond, BC). CWT was calculated as the sum of the weights of the left and right sides of each split carcass. CABF was measured over the ribeye muscle at 12th rib. CREA was measured on the cross section of the *longissimus dorsi* muscle between the 12th and 13th ribs. CGF was measured at the 12th – 13th rib. CLMY, an estimate of the saleable meat, was estimated using the following equation: lean meat yield (%) = 57.96 + [0.202 × *L. thoracis* area (cm^2)] – [0.027 × warm carcass weight (kg)] – [0.703 × average backfat thickness (mm)] (BASARAB *et al.* 2003). The descriptive statistics of these traits are given in Table 5 – 1.

5.2.2. SNP Markers

The DNA collection and genotyping for these animals are described in Chapter 3. SNPs not mapping to the Btau4.0 reference assembly and SNPs located on chromosome X were excluded. To avoid false associations and genotyping errors, SNPs with MAF < 0.05 or heterozygosity < 0.05 were discarded, as were SNPs with a chi-square (χ^2) value > 600 for the Hardy-Weinberg disequilibrium (HWD) test. This filtering yielded 40,809 SNPs for the whole genome association analysis. The summarized chromosome-wide SNP information is given in Table 5 – 2.

5.2.3. Genome-wide Association Studies

5.2.3.1. Single Marker LD Regression

The statistical model and significance test used in single marker LD regression are described in Chapter 4. For carcass merit traits, the fixed effects considered in the model were sire breed, slaughter age, and contemporary group. For ultrasound measurement traits, sire breed, measurement age, and contemporary group were considered.

In comparison with the analysis described in Chapter 4, two levels of significance controls were used in this study based on both the genome-wise and chromosome-wise type I error rate. The chromosome-wise threshold was adjusted for the number of SNPs tested for each chromosome, while the genome-wise

threshold was adjusted for the total number of SNPs tested. Significant SNPs are reported for both significance levels, using a 5% FDR threshold.

5.2.3.2. Bayesian Regression

The Bayesian statistical model and hypothesis testing approach are described in Chapter 4. Similar to the single marker LD regression, three systematic environmental effects (sire breed, contemporary group, and age of the animal) were considered in the model. For beef carcass traits, the Bayesian mixture models assumed $\pi = 0.95$ or $1-\pi = 0.05$, corresponding to about 2,000 non-zero SNP markers fitted per iteration of each Markov chain. In total, 2,556 unique 1 Mb SNP windows in the whole genome were fitted in this study.

5.2.3.3. Comparisons of Identified QTL

As described in Chapter 4, the significant SNPs and QTL regions identified from single marker LD regression and Bayesian regression were compared with previously reported QTL from the CattleQTLdb database.

5.2.3.4. QTL Annotation

Following the association analyses, candidate genes and gene network searches were carried out as described in Chapter 4. Potential functional candidate genes related to whole body lipid homeostasis, skeletal and muscular tissue growth, development and function were identified.

5.3. Results and Discussion

Whole genome association of 40,809 SNPs with eight beef carcass traits was performed by using both single marker LD regression and Bayesian regression in Canadian commercial hybrid beef cattle. The average intermarker spacing, heterozygosity and MAF for SNPs used in the association analyses were 62.28 $\pm$ 58.91 kb, 0.38 $\pm$ 0.12, and 0.28 $\pm$ 0.13, respectively (Table 5 – 2).

5.3.1. Association Analyses using Single Marker Regression

Single marker regression identified 69, 14, 146, 26, 13, 11, 24 and 25 significant SNPs at the chromosome-wise level for CWT, CABF, CREA, CGF, CLMY, UMAR, UBF and UREA, respectively. Using the more stringent genome-wise significance cut-off, there were 31, 29, 1 and 1 SNPs identified for CWT, CREA, CLMY and UMAR, respectively (Figure 5 – 1). Significant SNPs at the genome-wise threshold were not identified for CABF, CGF, UBF and UREA. The physical position, estimated allele substitution fixed effect, heterozygosity and MAF of the significant SNPs are presented for each trait in Tables 5 – 3 to 5 – 10. Several significant SNPs were shared among the analyzed beef carcass traits (Appendix 5). Most of the shared SNPs involve the carcass merit traits (for example, 25 SNPs overlap between CWT and CREA). Few common SNPs were found between the carcass merit traits and the ultrasound measurement traits.

The 328 significant SNPs at the chromosome-wise 5% FDR are distributed among 24 chromosomes (Figure 5 – 2, Table 5 – 11). No significant

SNPs were found on BTA9, 10, 16, 18 and 27. Different chromosomal distributions of significant SNPs were observed for different traits. For instance, most significant SNPs, 50 out of 69 for CWT, 10 out of 14 for CABF, 84 out of 146 for CREA, were clustered on BTA6, whereas most of the significant SNPs were found on BTA8 for both CGF (21 out of 26) and UBF (12 out of 24) (Figure 5 – 3). In addition, there were some other chromosomes having a higher density of significant SNPs, BTA4 for CREA, BTA2 for CREA and CLMY, and BTA5 for UREA (Figure 5 – 3). Overall, for beef ultrasound measurement and carcass merit traits, significant SNPs had a higher frequency on BTA2, 6, 8 and 13 (Figure 5 – 2 and Table 5 – 11).

5.3.1.1. Carcass Weight

Results from this study coincide with previously reported QTL for similar traits, and identify several novel regions of interest (Table 5 – 12). For CWT, it was observed that 43 significant SNPs cluster at 36.09 – 66.46 Mb on BTA6 (Figure 5 – 3A). This QTL region has been reported for carcass weight in different beef cattle populations (CASAS *et al.* 2000; SETOGUCHI *et al.* 2009; TAKASUGA *et al.* 2007) and one functional candidate gene *NCAPG* has been identified (SETOGUCHI *et al.* 2009). Many significant SNPs overlap with previously reported QTL in commercial Angus cattle (MCCLURE *et al.* 2010), for example, the five significant SNPs on BTA14 at 22.63 – 22.84 Mb (Figure 5 – 3B), five SNPs on BTA17 at 61.60 – 61.97 Mb (Figure 5 – 3C), and two genome-wise significant SNPs at 26.67 – 26.69 Mb on BTA24. Overall, most of the

significant SNPs identified for CWT are located within previously reported QTL regions, while a few SNPs are newly reported in this study on chromosomes 2, 6, 7, and 26.

5.3.1.2. Carcass Average Backfat

For CABF, 14 significant SNPs are distributed on three chromosomes. The SNPs on BTA2 are newly reported for CABF in this study. On BTA6 (Figure 5 – 3D), 9 out of 10 significant SNPs are located within a previously identified QTL at 17.00 – 43.93 cM in Angus cattle (MCCLURE *et al.* 2010). The remaining BTA6 SNP is close to a QTL at 89.35 – 118.0 cM (MCCLURE *et al.* 2010). The significant SNP on BTA13 overlaps with a QTL peak at 67 cM (MCCLURE *et al.* 2010).

5.3.1.3. Carcass Ribeye Area

For CREA, seven SNPs located at 2.01 – 10.09 Mb on BTA2 have been reported in many previous studies (ALEXANDER *et al.* 2007; CASAS *et al.* 1998; MORRIS *et al.* 2009), and the *MSTN* gene located within this region at 5.6 cM was identified as a functional candidate in double-muscled cattle (ALLAIS *et al.* 2010; CASAS *et al.* 1998; ESMAILIZADEH *et al.* 2008; MARTINEZ *et al.* 2010) (Figure 5 – 3E). On BTA4, 22 out of the 23 SNPs are newly identified for CREA in this study, while only one of them confirms a previous QTL peak in Japanese Black cattle (Takasuga *et al.* 2007) (Figure 5 – 3F). A total of 84 SNPs on BTA6 were found (Figure 5 – 3G), and partially overlap with previously reported QTL (Casas

et al. 2000; Casas *et al.* 2003; McClure *et al.* 2010). The reported candidate gene *NCAPG* at 39.1 cM on BTA6 overlaps with some of the significant SNPs (SETOGUCHI *et al.* 2009). The genome-wise significant SNP on BTA11 is located in the peak of a QTL at 103 cM in commercial US Angus cattle (MCCLURE *et al.* 2010), and eight SNPs on BTA19 at 43.67 – 55.30 Mb overlap with a QTL at 43.81 – 50 cM (TAYLOR *et al.* 1998) (Figure 5 – 3H). In addition, many novel SNPs were identified for CREA in this study, on BTA1, 2, 4, 6, 7, 8, 15, 19 and 25.

5.3.1.4. Carcass Grade Fat

Results for CGF were compared with previously reported QTL for fat thickness at the 12th rib since CGF in this study was measured at the 12th – 13th rib. Most (21 out of 26) significant SNPs for CGF are distributed on BTA8 (Figure 5 – 3I), with 10 SNPs newly reported in this study and 11 SNPs located within a previously described QTL at 11 – 47 cM (CASAS *et al.* 2001). The significant SNPs on BTA13 and BTA29 overlap with a known QTL peak at 67 cM and 27 cM, respectively (MCCLURE *et al.* 2010). The SNPs on BTA2 represent novel associations for CGF.

5.3.1.5. Carcass Lean Meat Yield

The results for CLMY were compared with previous QTL mapping results for yield grade since yield grade refers to the proportion of lean meat and has been classified in this study as follows: 1 ≥ 59%; 2 = 54 to 58%; and 3 ≤ 54%. A

small number of QTL have been reported for yield grade in the CattleQTLdb. Most SNPs on BTA2 (Figure 5 – 3J) are consistent with previous results, for example, one SNP is within a QTL at 3.86 – 10.77 cM (CASAS *et al.* 1998) and another four SNPs are within a QTL at 38 – 79 cM (CASAS *et al.* 2003). The SNPs on BTA6, 8, 13 and 21 do not overlap with known yield grade QTL.

5.3.1.6. Ultrasound Marbling

For UMAR, the SNPs on BTA2, 5, 7 and 12 identified in this study have not been linked to this trait before. The SNP on BTA4 is located within reported QTL in Japanese black cattle (MIZOSHITA *et al.* 2004; YOKOUCHI *et al.* 2009). The two SNPs reported on BTA14 are located within a QTL at 30 – 87 cM (CASAS *et al.* 2003). Two reported functional candidate genes for beef marbling score, *CRH* and *FABP4* (LEE *et al.* 2010a; WIBOWO *et al.* 2007), are located close to the only identified genome-wide significant SNP for UMAR on BTA14 at 43.74 Mb. The SNP on BTA22 overlaps with a QTL peak at 47 cM (MCCLURE *et al.* 2010).

5.3.1.7. Ultrasound Backfat

Although half of the significant SNPs for UBF are located on BTA8 (Figure 5 – 3K), no UBF QTL have been previously reported for this region. The SNPs reported on BTA20, 22 and 28 are also novel results from this study. One SNP on BTA13 is located within a reported QTL at 8.99 – 38.65 cM (MCCLURE *et al.* 2010) and another SNP on BTA13 overlaps with a QTL at 62.80 – 73.63 cM

(MCCLURE *et al.* 2010). The SNP on BTA21 is located within a QTL at 0 – 10.96 cM (MCCLURE *et al.* 2010), and the SNP on BTA23 is within one reported QTL at 52.29 – 67.93 cM (MCCLURE *et al.* 2010) and another QTL at 52.29 – 58.19 cM (LI *et al.* 2004).

5.3.1.8. Ultrasound Ribeye Area

For UREA, all significant SNPs on BTA5 (Figure 5 – 3M) represent novel associations except for one SNP at 45.35 Mb, which is located within a previously reported QTL (CASAS *et al.* 2003; MCCLURE *et al.* 2010). Novel associations were also identified on BTA2, 8, 13, 15, 20 and 22. The SNP on BTA6 overlaps a QTL peak at 41 cM (MCCLURE *et al.* 2010), where the functional candidate gene *NCAPG* is located (SETOGUCHI *et al.* 2009). Two SNPs on BTA8 overlap with two known QTL (MCCLURE *et al.* 2010), and SNPs on BTA12 and 26 are also consistent with previous findings (ALEXANDER *et al.* 2007; STONE *et al.* 1999).

Gene identifiers within 0.5 Mb windows on each side of the significant SNPs were obtained and gene networks among these genes were identified through the IPA functional clustering analysis, which is based on functional annotations in human and mice. Specifically, 2, 3, 1, 1, 1 and 2 gene networks with potential functions related to beef carcass traits were identified for CWT, CREA, CGF, CLMY, UBF and UREA, respectively (Table 5 – 13). This study did not find clear networks related to CABF and UMAR. These gene networks are mainly involved in carbohydrate metabolism, lipid metabolism, energy production, cellular growth and proliferation, and skeletal and muscular system

development and function. The physical positions and cellular locations of the gene products involved in the gene networks are presented in Appendix 6 – 11. Recently, whole genome association studies for beef carcass traits with high density SNP chips have been carried out in Korean and Australian beef populations (BOLORMAA *et al.* 2011; KIM *et al.* 2011; LEE *et al.* 2010b), however, systematically studies of candidate genes or gene networks were not carried out following the GWAS.

Overall, our study using single marker regression identified a large number of significant SNPs and potential gene networks associated with beef carcass traits. In addition to the novel associations identified, e.g. SNPs on BTA28 for UBF (Figure 5 – 3L), many of the significant SNPs are supported by previously reported QTL. However, the QTL span was greatly narrowed down in this study, for example, SNPs on BTA14 clustered at 22.63 – 22.84 Mb and SNPs on BTA17 clustered at 61.60 – 61.97 Mb; these SNPs are located within much smaller intervals compared to the overlapping previously reported QTL which extend over 20 cM (MCCLURE *et al.* 2010) (Table 5 – 12, Figure 5 – 3B and Figure 5 – 3C). These narrower and novel reported SNP regions and gene networks may provide future directions in association studies of beef carcass traits.

5.3.2. Association Analyses using Bayesian Regression

In the Bayesian analyses, the posterior genetic and residual variances and the genetic contribution of each SNP window for target traits were estimated

using the BayesB method. The proportion of phenotypic variance explained by SNPs was 0.51 for CWT, 0.25 for CABF, 0.37 for CREA, 0.27 for CGF, 0.24 for CLMY, 0.65 for UMAR, 0.36 for UBF, and 0.31 for UREA (Table 5 – 14). The genetic contribution to variation was relatively high for some traits. For example, UMAR with about 2,000 SNPs likely accounting for 65% of total genetic variance (π = 0.95), and CWT with the same number of SNPs accounting for 51%. Other traits, such as CLMY and CABF, had relatively lower heritability estimates. Overall, the proportion of phenotypic variance explained by all markers was between 0.24 – 0.65 for the beef carcass traits, reflecting that these traits in cattle had moderate to high genetic variance proportions. These variance proportions were similar to the heritability estimates previously reported in other beef populations (BERTRAND *et al.* 2001).

Candidate chromosomal regions were obtained by finding those genomic locations comprising SNP windows with the highest genetic variance on the trait. Significance was determined by the posterior distribution of each SNP window under the null hypothesis that the window did not harbor QTL. In total, 3, 4, 3, 2, 3, 7, 3 and 2 chromosome regions were identified with a *P*-value less than 0.2 for CWT, CABF, CREA, CGF, CLMY, UBF, UMAR and UREA, respectively. Several suggestive QTL regions are shared among the analyzed beef carcass traits. For example, the QTL region on BTA6 at 38.01 – 38.98 Mb is significant for CWT, CABF, CGF and CLMY, while the QTL region on BTA2 positioned at 31.02 – 32.00 Mb is associated with CABF, CREA, CLMY, UMAR, and UBF. In addition, one QTL on BTA14 at 5.01 – 5.97 Mb was found for all traits analyzed.

Following the association study, 10 functional candidate genes on five chromosomes were identified within the suggestive QTL regions. Information on the associated QTL regions together with the functionally relevant genes in the regions are given in Table 5 – 15 and Figures 5 – 4A to 5 – 4H. Many overlaps between the associated regions from this study with previously reported QTL regions related to similar beef carcass traits were observed (Table 5 – 15). However, there are novel candidate regions reported in this study. The detailed results for all traits, including positions of the associated chromosomal regions, number of SNPs within each region, *P*-values, accounted phenotypic variances, genes and previously reported QTL located in the regions are presented in Table 5 – 16.

Three QTL regions (1 Mb windows) associated with CWT were found on BTA6 and BTA14. Among those, the QTL region on BTA6 (38.01 – 38.98 Mb) was highly significantly ($P < 0.004$) associated with CWT and explained 12.47% of the phenotypic variance of CWT. This region overlaps with a previously reported QTL for carcass weight in Japanese Black cattle and harbors the functional candidate gene *NCAPG* (SETOGUCHI *et al.* 2009; TAKASUGA *et al.* 2007). In this study, two functional candidate genes, *NCAPG* and *LCORL*, were identified in this region. Previous association studies and metabolomic profiles both indicated that *NCAPG* has a major effect on *Bos taurus* growth (EBERLEIN *et al.* 2009; WEIKARD *et al.* 2010). The gene *LCORL* in the same QTL region was found to have a function in controlling stature in cattle (PRYCE *et al.* 2011) and in humans has effects on human trunk length and skeletal frame size (SORANZO *et*

al. 2009) and adult height (WEEDON *et al.* 2008). The QTL region on BTA14 at 22.02 – 22.97 Mb was also highly significant ($P < 0.006$) with 5.86% explained phenotypic variance for CWT and it is located within a previously reported QTL for carcass weight (MCCLURE *et al.* 2010). The *TGS1* gene within this region was reported to affect adult human height (GUDBJARTSSON *et al.* 2008), determine pediatric stature (ZHAO *et al.* 2010) and regulate cellular lipid metabolic processes and adipogenesis (QI *et al.* 2003). Another QTL region on BTA14 at 5.01 – 5.97 Mb was newly reported in this study and no functional candidate genes were identified.

For CABF, four SNP windows were found on three chromosomes, BTA2, 6 and 14. The QTL region on BTA2 positioned at 31.02 – 32.00 Mb is located within one reported QTL for fat depth (CASAS *et al.* 2003). The *GALNT3* gene within this region was identified as a functional candidate since *GALNT3* was found to be associated with mouse weight (GHAZALPOUR *et al.* 2006) and *GALNT3*-deficient males showed growth retardation and significantly increased bone mineral density (ICHIKAWA *et al.* 2009). The other QTL on BTA2 at 139.00 – 139.99 Mb is a novel result from this study. The QTL at 38.01 – 38.98 Mb on BTA6 overlaps with one reported QTL associated with fat thickness in a commercial Angus cattle population (MCCLURE *et al.* 2010), while the QTL at 5.01 – 5.97 Mb on BTA14 was previously found to be associated with backfat thickness in a commercial line of *Bos taurus* (MOORE *et al.* 2003).

Suggestive QTL regions for CREA are distributed on BTA2, 7 and 14 and none of these regions have been reported in previous QTL mapping studies. Two QTL regions for CGF were found on BTA6 and BTA14 and were also identified for CABF. Of the three QTL regions identified for CLMY on BTA2, 6 and 14, only the QTL on BTA14 at 5.01 – 5.97 Mb overlaps with a previously reported QTL, for USDA yield grade (CASAS *et al.* 2003). The regions on BTA2 and BTA6 are newly reported for CLMY in the current study.

Seven QTL regions were identified for UMAR on BTA2, 4, 8, 14 and 25. The QTL on BTA8 (1.02 – 1.96 Mb) containing the newly proposed candidate gene *CLCN3* explained 3.33% of the phenotypic variance for UMAR and is located within or close to a previously reported QTL for beef marbling score in other beef cattle populations (MCCLURE *et al.* 2010; TAKASUGA *et al.* 2007). CLCN3 is localized in insulin granules and plays a role in insulin processing and insulin secretion (DERIY *et al.* 2009; FU *et al.* 2010) and was found to participate in the fibroblast-to-myofibroblast transition (YIN *et al.* 2008). The QTL at 43.05 – 43.98 Mb on BTA14, explaining 2.18% of the phenotypic variance for UMAR, was previously reported for marbling and subcutaneous fat depth in different populations (CASAS *et al.* 2003; LEE *et al.* 2010a; WIBOWO *et al.* 2007). The gene *EXT1* within this region was selected as a functional candidate since EXT1 is involved in maintaining the phenotype and function of joint-forming cells and coordinating local signaling pathways in long bone growth (MUNDY *et al.* 2011). An important QTL on BTA4 (59.05 – 59.97 Mb) explained 2.23% of the phenotypic variance for UMAR and is located within or close to a reported QTL

for beef marbling score in Japanese Black cattle (MIZOSHITA *et al.* 2004; TAKASUGA *et al.* 2007). The SNP window on BTA14 at 5.01 – 5.97 Mb coincides with a QTL peak at 5 cM (MCCLURE *et al.* 2010). The two QTL on BTA2 and one QTL on BTA25 are novel findings from the current study.

The three identified QTL regions for UBF were found on BTA2, 14 and 28. The two QTL regions on BTA2 and BTA14 are described in the CABF results section. The QTL region on BTA28 (26.02 – 26.99 Mb), which explained 1.5% of the phenotypic variance, is close to a reported QTL at 23 cM for beef fat thickness (MCCLURE *et al.* 2010). Three genes including *SGPL1*, *PCBD1*, and *SLC29A3* within this region were selected in the current study as potential functional candidate genes for UBF based on their functional annotations related to lipid metabolism. *SGPL1* is involved in lipid metabolic pathways and deficiency of *SGPL1* disrupts lipid homeostasis (BEKTAS *et al.* 2010). In humans, *PCBD1* may determine the degree of adiposity (NAUKKARINEN *et al.* 2010). Inactivating mutations in *SLC29A3* can cause insulin-dependent diabetes and can profoundly affect cell size and number through effects on the insulin signaling pathway (CLIFFE *et al.* 2009). The *ADAMTS14* gene within this region was also selected in the present study as a functional candidate since this gene functions in the maturation of collagen fibers and a nsSNP was found to be associated with some osteoarthritis phenotypes (RODRIGUEZ-LOPEZ *et al.* 2009).

For UREA, two QTL regions on BTA13 and BTA14 were identified. The QTL at 80.05 – 80.96 Mb on BTA13 overlaps with a reported QTL for ribeye

muscle area in a commercial Angus cattle population (MCCLURE *et al.* 2010), while the QTL region on BTA14 was newly reported for UREA in this study. No functional candidate genes were identified for UREA from the two QTL regions based on currently available functional studies.

The present analyses, using the BovineSNP50 BeadChip, identified 2 – 7 QTL regions associated with each beef carcass trait. The results support several QTL locations for similar traits from previously independent studies, and reveal some novel QTL regions. The analyzed beef carcass traits shared a number of suggestive QTL regions. The concordance of QTL regions may result from the high genetic correlation among these traits (PARIACOTE *et al.* 1998). The number of detected QTL regions per trait was small in the Bayesian analyses, suggesting that a larger population size will be required for greater power to detect the associated genomic regions for these traits. The results support the previously documented candidate gene *NCAPG* and identified several novel functional candidate genes for beef carcass traits, such as *TGS1* for carcass weight; *GALNT3* for carcass backfat, ribeye area, lean meat yield, and marbling; *CLCN3* and *EXT1* for marbling (Tables 5 – 15 and 5 – 16). The selected functional candidate genes have previously determined functions related to lipid metabolism, cellular growth and proliferation and skeletal and muscular system development and function.

5.3.3. Association Analyses using Different Methods

Significant SNPs and suggestive chromosomal regions harboring variation affecting beef carcass traits were identified in this study using the BovineSNP50

BeadChip assay and two different association methods: single marker LD regression and the BayesB method (MEUWISSEN *et al.* 2001) in the Canadian commercial hybrid beef cattle. Both methods revealed several novel associations for beef carcass traits in addition to many that are supported by previously reported QTL.

Generally, larger number of significant associations was identified by using the single marker LD regression compared to the Bayesian regression. Regarding to the problem of multiple testing in single marker analysis, FDR correction on the reported significance of SNPs was used to reduce the type I error. Overall, we examined 62 genome-wise and 328 chromosome-wise significant SNPs (on 24 chromosomes) at the 5% FDR for beef carcass traits. The QTL span was greatly reduced in this study compared to many previous reported QTL that extended for tens of centimorgans (Table 5 – 12). However, significant SNPs are still spread out over a few centimorgans, which makes the identification of compelling functional genes difficult. Due to large number of candidates found, only potential gene networks among the positional candidate genes are described in the text.

Compared to single marker regression which fits one marker at a time in the model, the Bayesian methods avoided the problems with model selection and multiple testing and fitted all genome-wide dense markers simultaneously into the model. In this study, we used a recently developed Bayesian method, BayesC (FERNANDO and GARRICK 2009) to derive the variance components. These

components were used for the association analyses using the BayesB approach, since BayesC is less sensitive to the given priors of genetic and residual variances compared to BayesB (KIZILKAYA *et al.* 2010). We also used non-overlapping 1 Mb SNP sliding windows in the Bayesian analysis to remove the highly correlated SNPs in the model based on results from previous studies (FAN *et al.* 2011; ONTERU *et al.* 2011; SUN *et al.* 2011). In this way, the Bayesian regression will remove many redundant QTL regions caused by the high LD and the resulting QTL regions, actually 1 Mb SNP windows, will facilitate the subsequent search for functional candidate genes. However, there are also challenges associated with the Bayesian analysis. Since the Bayesian methods analyze dense markers simultaneously in the model, a larger number of phenotypes are required for identifying the associated regions. In this study, only a few suggestive QTL regions ($P < 0.2$) were found to be associated with traits based on the posterior distribution of the test statistic. One major reason could be the small population size used in this study where 922 animals were used in fitting all 40,809 SNPs simultaneously in the Bayesian model. Future studies using larger sample sizes will be required for a better power to detect QTL regions at higher significance levels.

For the typical Canadian commercial beef production system using hybrid beef cattle, identification of population-specific QTL regions and genetic markers could lead to more efficient genetic improvement programs. In this study, we identified many QTL regions consistent with those from previous studies, and found many overlapping associations between the two methods used (Table 5 –

17). Both methods were able to detect some important QTL regions reported by many previous studies, for example, one QTL for CWT on BTA6 was identified at 38.01 – 38.98 Mb using Bayesian regression, while 39 SNPs located within 36.09 – 46.52 Mb were found in the single marker regression. In addition, there were novel QTL identified for this specific cattle population from both methods. For instance, both methods reported a novel QTL at 31.02 – 32.00 Mb on BTA2 for UMAR. The results from different methods may increase the confidence of declaration of new QTL regions or candidate genes for beef carcass traits. In addition, following the association analyses, functional candidate genes and possible gene networks were identified in this study for beef carcass traits based on previous association and functional studies. These provide new directions for future studies since most previous studies only report the gene closest to the identified associations. The novel QTL and candidate gene information from this study should be validated in other populations before being used in specific breeding programs.

5.4. Conclusions

GWAS using the BovineSNP50 BeadChip (50K) assay led to the identification of many QTL and candidate genes underlying beef carcass traits in the Canadian commercial hybrid beef cattle. Association using single marker LD regression presented strong evidence for the presence of 62 genome-wise and 328 chromosome-wise significant (5% FDR) SNPs associated with beef carcass traits on 24 chromosomes with a large proportion of the significant associations

clustered on BTA 2, 6, 8 and 13. Functional clustering analyses for positional candidate genes disclosed several gene networks which might be related to the phenotypic variation of beef carcass yield and composition. Using Bayesian regression, 12 QTL regions on 9 chromosomes with significant genetic contribution were discovered for beef carcass traits and 10 functional candidate genes on five chromosomes were identified. Both methods disclosed several novel associations in addition to many that coincide with previously reported QTL in other beef populations, and many concordances of associations were observed between the two methods. The networks and genes highlighted in this work may lead to a more detailed understanding of the molecular mechanisms controlling beef yield and composition in cattle. The information on associated QTL regions, SNPs and genes could be used in further identification of causal mutations and in MAS for genetic improvement of beef quality and quantity, following validation in other populations.

5.5. Literature Cited

ABE, T., J. SABURI, H. HASEBE, T. NAKAGAWA, T. KAWAMURA *et al.*, 2008 Bovine quantitative trait loci analysis for growth, carcass, and meat quality traits in an F-2 population from a cross between Japanese Black and Limousin. Journal of Animal Science **86:** 2821-2832.

ALEXANDER, L. J., T. W. GEARY, W. M. SNELLING and M. D. MACNEIL, 2007 Quantitative trait loci with additive effects on growth and carcass traits in a Wagyu-Limousin F2 population. Animal Genetics **38:** 413-416.

ALLAIS, S., H. LEVEZIEL, N. PAYET-DUPRAT, J. F. HOCQUETTE, J. LEPETIT *et al.*, 2010 The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds. Journal of Animal Science **88:** 446-454.

BASARAB, J. A., M. A. PRICE, J. L. AALHUS, E. K. OKINE, W. M. SNELLING *et al.*, 2003 Residual feed intake and body composition in young growing cattle. Canadian Journal of Animal Science **83:** 189-204.

BEKTAS, M., M. L. ALLENDE, B. G. LEE, W. P. CHEN, M. J. AMAR *et al.*, 2010 Sphingosine 1-Phosphate Lyase Deficiency Disrupts Lipid Homeostasis in Liver. Journal of Biological Chemistry **285:** 10880-10889.

BERTRAND, J. K., R. D. GREEN, W. O. HERRING and D. W. MOSER, 2001 Genetic evaluation for beef carcass traits. Journal of Animal Science **79:** E190–E200.

BOLORMAA, S., L. R. P. NETO, Y. D. ZHANG, R. J. BUNCH, B. E. HARRISON *et al.*, 2011 A genome-wide association study of meat and carcass traits in Australian cattle. Journal of Animal Science **89:** 2297-2309.

CASAS, E., J. W. KEELE, S. D. SHACKELFORD, M. KOOHMARAIE, T. S. SONSTEGARD *et al.*, 1998 Association of the muscle hypertrophy locus with carcass traits in beef cattle. Journal of Animal Science **76:** 468-473.

CASAS, E., S. D. SHACKELFORD, J. W. KEELE, M. KOOHMARAIE, T. P. L. SMITH *et al.*, 2003 Detection of quantitative trait loci for growth and carcass composition in cattle. Journal of Animal Science **81:** 2976-2983.

CASAS, E., S. D. SHACKLEFORD, J. W. KEELE, R. T. STONE, S. M. KAPPES *et al.*, 2000 Quantitative trait loci affecting growth and carcass composition of cattle segregating alternate forms of myostatin. Journal of Animal Science **78:** 560-569.

CASAS, E., R. T. STONE, J. W. KEELE, S. D. SHACKELFORD, S. M. KAPPES *et al.*, 2001 A comprehensive search for quantitative trait loci affecting growth and carcass composition of cattle segregating alternative forms of the myostatin gene. Journal of Animal Science **79:** 854-860.

CLIFFE, S. T., J. M. KRAMER, K. HUSSAIN, J. H. ROBBEN, E. K. DE JONG *et al.*, 2009 SLC29A3 gene is mutated in pigmented hypertrichosis with insulin-dependent diabetes mellitus syndrome and interacts with the insulin signaling pathway. Human Molecular Genetics **18:** 2257-2265.

DAVIS, G. P., and S. K. DENISE, 1998 The impact of genetic markers on selection. Journal of Animal Science **76:** 2331-2339.

DERIY, L. V., E. A. GOMEZ, D. A. JACOBSON, X. Q. WANG, J. A. HOPSON *et al.*, 2009 The Granular Chloride Channel ClC-3 is Permissive for Insulin Secretion. Cell Metabolism **10:** 316-323.

EBERLEIN, A., A. TAKASUGA, K. SETOGUCHI, R. PFUHL, K. FLISIKOWSKI *et al.*, 2009 Dissection of Genetic Factors Modulating Fetal Growth in Cattle Indicates a Substantial Role of the Non-SMC Condensin I Complex, Subunit G (NCAPG) Gene. Genetics **183:** 951-964.

ESMAILIZADEH, A. K., C. D. BOTTEMA, G. S. SELLICK, A. P. VERBYLA, C. A. MORRIS *et al.*, 2008 Effects of the myostatin F94L substitution on beef traits. Journal of Animal Science **86:** 1038-1046.

FAN, B., S. K. ONTERU, Z. Q. DU, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Genome-Wide Association Study Identifies Loci for Body Composition and Structural Soundness Traits in Pigs. PLoS One **6**.

FERNANDO, R. L., and D. J. GARRICK, 2009 GenSel - User manual for a portfolio of genomic selection related analyses. Animal Breeding and Genetics, Iowa State University, Ames.

FU, J. F., J. F. FU, B. GAO, L. WANG, N. Y. ZHANG *et al.*, 2010 Expression patterns of ClC-3 mRNA and protein in aortic smooth muscle, kidney and brain in diabetic rats. Clinical and Investigative Medicine **33:** E146-E154.

GHAZALPOUR, A., S. DOSS, B. ZHANG, S. WANG, C. PLAISIER *et al.*, 2006 Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genetics **2:** 1182-1192.

GUDBJARTSSON, D. F., G. B. WALTERS, G. THORLEIFSSON, H. STEFANSSON, B. V. HALLDORSSON *et al.*, 2008 Many sequence variants affecting diversity of adult human height. Nature Genetics **40:** 609-615.

GUTIERREZ-GIL, B., J. L. WILLIAMS, D. HOMER, D. BURTON, C. S. HALEY *et al.*, 2009 Search for quantitative trait loci affecting growth and carcass traits in a cross population of beef and dairy cattle. Journal of Animal Science **87:** 24-36.

ICHIKAWA, S., A. H. SORENSON, A. M. AUSTIN, D. S. MACKENZIE, T. A. FRITZ *et al.*, 2009 Ablation of the Galnt3 Gene Leads to Low-Circulating Intact Fibroblast Growth Factor 23 (Fgf23) Concentrations and Hyperphosphatemia Despite Increased Fgf23 Expression. Endocrinology **150:** 2543-2550.

KIM, J. J., F. FARNIR, J. SAVELL and J. F. TAYLOR, 2003 Detection of quantitative trait loci for growth and beef carcass fatness traits in a cross between Bos taurus (Angus) and Bos indicus (Brahman) cattle. Journal of Animal Science **81:** 1933-1942.

KIM, Y., J. RYU, J. WOO, J. B. KIM, C. Y. KIM *et al.*, 2011 Genome-wide association study reveals five nucleotide sequence variants for carcass traits in beef cattle. Animal Genetics **42:** 361-365.

KIZILKAYA, K., R. L. FERNANDO and D. J. GARRICK, 2010 Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. Journal of Animal Science **88:** 544-551.

LEE, S. H., J. H. J. VAN DER WERF, S. H. LEE, E. W. PARK, S. J. OH *et al.*, 2010a Genetic polymorphisms of the bovine Fatty acid binding protein 4 gene are significantly associated with marbling and carcass weight in Hanwoo (Korean Cattle). Animal Genetics **41:** 442-444.

LEE, Y. M., C. M. HAN, Y. LI, J. J. LEE, L. H. KIM *et al.*, 2010b A Whole Genome Association Study to Detect Single Nucleotide Polymorphisms

for Carcass Traits in Hanwoo Populations. Asian-Australasian Journal of Animal Sciences **23:** 417-424.

LI, C., J. BASARAB, W. M. SNELLING, B. BENKEL, J. KNEELAND *et al.*, 2004 Identification and fine mapping of quantitative trait loci for backfat on bovine chromosomes 2, 5, 6, 19, 21, and 23 in a commercial line of Bos taurus. Journal of Animal Science **82:** 967-972.

MARTINEZ, A., N. ALDAI, R. CELAYA and K. OSORO, 2010 Effect of breed body size and the muscular hypertrophy gene in the production and carcass traits of concentrate-finished yearling bulls. Journal of Animal Science **88:** 1229-1239.

MCCLURE, M. C., N. S. MORSCI, R. D. SCHNABEL, J. W. KIM, P. YAO *et al.*, 2010 A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Animal Genetics **41:** 597-607.

MEUWISSEN, T. H. E., B. J. HAYES and M. E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157:** 1819-1829.

MIZOSHITA, K., T. WATANABE, H. HAYASHI, C. KUBOTA, H. YAMAKUCHI *et al.*, 2004 Quantitative trait loci analysis for growth and carcass traits in a half-sib family of purebred Japanese Black (Wagyu) cattle. Journal of Animal Science **82:** 3415-3420.

MOORE, S. S., C. LI, J. BASARAB, W. M. SNELLING, J. KNEELAND *et al.*, 2003 Fine mapping of quantitative trait loci and assessment of positional

candidate genes for backfat on bovine chromosome 14 in a commercial line of Bos taurus. Journal of Animal Science **81:** 1919-1925.

MORRIS, C. A., W. S. PITCHFORD, N. G. CULLEN, A. K. ESMAILIZADEH, S. M. HICKEY *et al.*, 2009 Quantitative trait loci for live animal and carcass composition traits in Jersey and Limousin back-cross cattle finished on pasture or feedlot. Animal Genetics **40:** 648-654.

MUNDY, C., T. YASUDA, T. KINUMATSU, Y. YAMAGUCHI, M. IWAMOTO *et al.*, 2011 Synovial joint formation requires local Ext1 expression and heparan sulfate production in developing mouse embryo limbs and spine. Developmental Biology **351:** 70-81.

NAUKKARINEN, J., I. SURAKKA, K. H. PIETILAINEN, A. RISSANEN, V. SALOMAA *et al.*, 2010 Use of Genome-Wide Expression Data to Mine the "Gray Zone" of GWA Studies Leads to Novel Candidate Obesity Genes. PLoS Genetics **6**.

ONTERU, S. K., B. FAN, M. T. NIKKILA, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Whole-genome association analyses for lifetime reproductive traits in the pig. Journal of Animal Science **89:** 988-995.

PARIACOTE, F., L. D. VAN VLECK and R. E. HUNSLEY, 1998 Genetic and of phenotypic parameters for carcass traits American Shorthorn beef cattle. Journal of Animal Science **76:** 2584-2588.

PRYCE, J. E., B. HAYES, S. BOLORMAA and M. E. GODDARD, 2011 Polymorphic Regions Affecting Human Height Also Control Stature in Cattle. Genetics **187:** 981-984.

QI, C., S. SURAPUREDDI, Y. J. ZHU, S. T. YU, P. KASHIREDDY *et al.*, 2003 Transcriptional coactivator PRIP, the peroxisome proliferator-activated receptor gamma (PPAR gamma)-interacting protein, is required for PPAR gamma- mediated adipogenesis. Journal of Biological Chemistry **278:** 25281-25284.

RODRIGUEZ-LOPEZ, J., M. POMBO-SUAREZ, J. LOUGHLIN, A. TSEZOU, F. J. BLANCO *et al.*, 2009 Association of a nsSNP in ADAMTS14 to some osteoarthritis phenotypes. Osteoarthritis and Cartilage **17:** 321-327.

SEATON, G., C. S. HALEY, S. A. KNOTT, M. KEARSEY and P. M. VISSCHER, 2002 QTL Express: mapping quantitative trait loci in of simple and complex pedigrees. Bioinformatics **18:** 339-340.

SELLNER, E. M., J. W. KIM, M. C. MCCLURE, K. H. TAYLOR, R. D. SCHNABEL *et al.*, 2007 Board-invited review: Applications of genomic information in livestock. Journal of Animal Science **85:** 3148-3158.

SETOGUCHI, K., M. FURUTA, T. HIRANO, T. NAGAO, T. WATANABE *et al.*, 2009 Cross-breed comparisons identified a critical 591-kb region for bovine carcass weight QTL (CW-2) on chromosome 6 and the Ile-442-Met substitution in NCAPG as a positional candidate. BMC Genetics **10**.

SORANZO, N., F. RIVADENEIRA, U. CHINAPPEN-HORSLEY, I. MALKINA, J. B. RICHARDS *et al.*, 2009 Meta-Analysis of Genome-Wide Scans for Human Adult Stature Identifies Novel Loci and Associations with Measures of Skeletal Frame Size. PLoS Genetics **5**.

STONE, R. T., J. W. KEELE, S. D. SHACKELFORD, S. M. KAPPES and M. KOOHMARAIE, 1999 A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. Journal of Animal Science **77:** 1379-1384.

SUN, X., D. HABIER, R. L. FERNANDO, D. J. GARRICK and J. C. DEKKERS, 2011 Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. BMC Proceedings **5 Suppl 3:** S13.

TAKASUGA, A., T. WATANABE, Y. MIZOGUCHI, T. HIRANO, N. IHARA *et al.*, 2007 Identification of bovine QTL for growth and carcass traits in Japanese Black cattle by replication and identical-by-descent mapping. Mammalian Genome **18:** 125-136.

TAYLOR, J. F., L. L. COUTINHO, K. L. HERRING, D. S. GALLAGHER, R. A. BRENNEMAN *et al.*, 1998 Candidate gene analysis of GH1 for effects on growth and carcass composition of cattle. Animal Genetics **29:** 194-201.

WEEDON, M. N., H. LANGO, C. M. LINDGREN, C. WALLACE, D. M. EVANS *et al.*, 2008 Genome-wide association analysis identifies 20 loci that influence adult height. Nature Genetics **40:** 575-583.

WEIKARD, R., E. ALTMAIER, K. SUHRE, K. M. WEINBERGER, H. M. HAMMON *et al.*, 2010 Metabolomic profiles indicate distinct physiological pathways affected by two loci with major divergent effect on Bos taurus growth and lipid deposition. Physiological Genomics **42A:** 79-88.

WIBOWO, T. A., J. J. MICHAL and Z. JIANG, 2007 Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. Genome **50:** 939-945.

YIN, Z. H., Y. A. TONG, H. Q. ZHU and M. A. WATSKY, 2008 ClC-3 is required for LPA-activated Cl- current activity and fibroblast-to-myofibroblast differentiation. American Journal of Physiology-Cell Physiology **294:** C535-C542.

YOKOUCHI, K., Y. MIZOGUCHI, T. WATANABE, E. IWAMOTO, Y. SUGIMOTO *et al.*, 2009 Identification of a 3.7-Mb region for a marbling QTL on bovine chromosome 4 by identical-by-descent and association analysis. Animal Genetics **40:** 945-951.

ZHAO, J. H., M. Y. LI, J. P. BRADFIELD, H. T. ZHANG, F. D. MENTCH *et al.*, 2010 The role of height-associated loci identified in genome wide association studies in the determination of pediatric stature. BMC Medical Genetics **11**.

Tables

Table 5 – 1. Number of steers and descriptive statistics for beef ultrasound and carcass merit traits in the commercial hybrid beef cattle population.

Trait	Steers, N	Mean	SD	Min	Max
Carcass weight (CWT), kg	836	317.98	29.76	207.20	453.19
Carcass average backfat (CABF), mm	836	12.45	3.98	2.67	26.67
Carcass ribeye area (CREA), cm^2	836	82.94	8.39	53.00	113.00
Carcass grade fat (CGF), mm	836	11.03	3.95	2.00	26.00
Carcass lean meat yield (CLMY), %	836	57.38	3.66	44.67	66.18
Ultrasound marbling (UMAR)	922	4.35	1.07	2.82	31.15
Ultrasound backfat (UBF), mm	922	13.46	5.33	1.02	29.97
Ultrasound ribeye area (UREA), cm^2	922	61.47	8.43	31.50	88.17

Table 5 – 2. Number of SNPs, average intermarker distance (kb), average heterozygosity and average MAF of SNPs on 29 autosomes in commercial hybrid beef cattle.

BTA	N-SNPs	Intermarker spacing ± SD	Heterozygosity ± SD	MAF ± SD
1	2640	60.97 ±51.61	0.369 ±0.123	0.276 ±0.13
2	2201	63.86 ±64.08	0.377 ±0.121	0.283 ±0.131
3	2064	62 ±61.72	0.373 ±0.121	0.276 ±0.128
4	2016	61.56 ±50.6	0.373 ±0.121	0.279 ±0.128
5	1706	73.78 ±79.3	0.375 ±0.122	0.282 ±0.13
6	2057	59.55 ±56.41	0.378 ±0.121	0.282 ±0.129
7	1774	63 ±60.99	0.377 ±0.121	0.282 ±0.128
8	1933	60.52 ±51.83	0.37 ±0.122	0.275 ±0.129
9	1649	65.56 ±62.17	0.368 ±0.125	0.275 ±0.133
10	1747	60.51 ±71.46	0.373 ±0.12	0.279 ±0.128
11	1771	62.15 ±57.12	0.376 ±0.121	0.282 ±0.13
12	1314	64.89 ±65.08	0.372 ±0.123	0.28 ±0.131
13	1424	59.08 ±50.55	0.377 ±0.123	0.283 ±0.13
14	1364	59.63 ±49.89	0.386 ±0.114	0.289 ±0.124
15	1329	63.68 ±58.65	0.374 ±0.119	0.28 ±0.128
16	1242	62.75 ±64.47	0.372 ±0.12	0.279 ±0.13
17	1264	60.45 ±52.84	0.367 ±0.121	0.274 ±0.129
18	1068	61.89 ±59.87	0.384 ±0.118	0.291 ±0.129
19	1061	61.35 ±51.55	0.388 ±0.112	0.295 ±0.124
20	1236	61.13 ±54.92	0.375 ±0.124	0.282 ±0.131
21	1067	64.86 ±63.38	0.377 ±0.12	0.281 ±0.127
22	1010	61.07 ±46.76	0.369 ±0.121	0.275 ±0.128
23	868	61.44 ±53.44	0.387 ±0.118	0.292 ±0.126
24	1024	63.47 ±54.78	0.383 ±0.12	0.289 ±0.131
25	785	55.39 ±44.01	0.392 ±0.115	0.302 ±0.128
26	844	60.5 ±45.03	0.367 ±0.123	0.27 ±0.127
27	772	63.17 ±85.61	0.374 ±0.12	0.282 ±0.131
28	755	60.98 ±49.37	0.379 ±0.119	0.285 ±0.128
29	824	62.82 ±61.08	0.376 ±0.121	0.283 ±0.129
Overall	40809	62.28 ±58.91	0.375 ±0.121	0.281 ±0.129

Table 5 – 3. Chromosome-wise significant (FDR $P < 0.05$) SNPs for carcass weight (CWT) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-NGS-14875	2	25,174,430	0.14	0.24	8.659 ±2.056	17.74	2.94E-05*
ARS-BFGL-NGS-104900	6	6,995,395	0.26	0.41	-6.077 ±1.729	12.35	0.000469
ARS-BFGL-NGS-18384	6	24,251,269	0.18	0.3	6.247 ±1.838	11.55	0.000711
BTA-75645-no-rs	6	29,675,685	0.15	0.24	6.396 ±1.898	11.36	0.000787
ARS-BFGL-NGS-12563	6	31,783,985	0.44	0.51	5.194 ±1.522	11.65	0.000674
Hapmap59443-rs29009843	6	31,937,563	0.32	0.43	-5.147 ±1.538	11.2	0.000857
ARS-BFGL-NGS-28684	6	32,765,343	0.17	0.29	7.758 ±1.907	16.55	5.21E-05
BFGL-NGS-114855	6	36,088,413	0.24	0.37	-5.605 ±1.664	11.34	0.000795
ARS-BFGL-NGS-49600	6	36,280,417	0.36	0.46	-5.798 ±1.459	15.78	7.74E-05
Hapmap26264-BTC-037159	6	36,969,462	0.26	0.39	-8.663 ±1.584	29.92	6.03E-08*
Hapmap33644-BTC-062485	6	37,045,881	0.19	0.3	-7.552 ±1.775	18.1	2.35E-05*
Hapmap30134-BTC-034283	6	37,852,400	0.32	0.46	-7.411 ±1.604	21.34	4.50E-06*
Hapmap26308-BTC-057761	6	37,963,147	0.41	0.51	8.207 ±1.48	30.75	4.00E-08*
BFGL-NGS-112812	6	38,014,254	0.21	0.35	-7.837 ±1.801	18.93	1.53E-05*
ARS-BFGL-NGS-45457	6	38,102,327	0.49	0.53	7.086 ±1.488	22.67	2.29E-06*
Hapmap27083-BTC-041166	6	38,212,941	0.49	0.53	-6.338 ±1.488	18.14	2.3E-05*
Hapmap23507-BTC-041133	6	38,233,088	0.47	0.54	-6.228 ±1.529	16.59	5.12E-05
Hapmap31285-BTC-041097	6	38,256,889	0.5	0.53	6.665 ±1.482	20.23	7.90E-06*
Hapmap33630-BTC-041044	6	38,301,284	0.18	0.29	-8.197 ±1.863	19.36	1.23E-05*
Hapmap33628-BTC-041023	6	38,326,147	0.49	0.53	7.086 ±1.488	22.67	2.29E-06*
Hapmap27537-BTC-060891	6	38,638,962	0.44	0.53	-4.973 ±1.505	10.91	0.000998
BTB-01709638	6	38,658,980	0.11	0.2	-8.264 ±2.274	13.21	0.000296
Hapmap31044-BTC-071337	6	38,729,866	0.45	0.53	-6.063 ±1.495	16.44	5.50E-05
Hapmap33170-BTC-071249	6	38,756,335	0.27	0.4	-9.226 ±1.635	31.82	2.37E-08*
BTB-01326707	6	38,824,038	0.17	0.28	-7.371 ±1.975	13.92	0.000205
Hapmap33339-BTC-071052	6	38,914,556	0.48	0.52	5.991 ±1.432	17.51	3.17E-05*
Hapmap26618-BTC-070864	6	38,982,338	0.34	0.47	-6.549 ±1.496	19.17	1.35E-05*
Hapmap23923-BTC-066021	6	39,108,078	0.14	0.25	-8.436 ±2.15	15.4	9.48E-05
Hapmap32513-BTC-066089	6	39,139,502	0.47	0.5	4.824 ±1.418	11.58	0.0007
Hapmap27298-BTC-035654	6	39,159,587	0.19	0.3	-7.094 ±1.847	14.75	0.000134
Hapmap32210-BTC-035534	6	39,223,437	0.31	0.46	6.352 ±1.598	15.8	7.71E-05
Hapmap43932-BTA-75850	6	39,486,004	0.39	0.48	-5.973 ±1.467	16.58	5.16E-05
BTB-00260450	6	39,509,020	0.37	0.5	-7.841 ±1.514	26.82	2.85E-07*
Hapmap23907-BTC-036006	6	39,770,115	0.33	0.44	5.902 ±1.501	15.47	9.13E-05
Hapmap27299-BTC-035816	6	39,794,334	0.42	0.5	8.416 ±1.489	31.93	2.24E-08*
Hapmap33079-BTA-163567	6	40,096,368	0.42	0.49	-6.168 ±1.532	16.22	6.25E-05
Hapmap32714-BTC-037559	6	41,028,973	0.28	0.42	5.809 ±1.604	13.12	0.000311
Hapmap23186-BTC-046762	6	41,208,356	0.29	0.41	-6.568 ±1.527	18.5	1.90E-05*
ARS-BFGL-NGS-72188	6	41,541,414	0.33	0.45	5.084 ±1.529	11.06	0.000924
ARS-BFGL-NGS-70160	6	41,625,126	0.49	0.54	6.283 ±1.43	19.32	1.25E-05*
BTA-75905-no-rs	6	41,680,774	0.21	0.36	8.062 ±1.887	18.25	2.19E-05*
Hapmap47224-BTA-24614	6	43,071,076	0.43	0.49	5.166 ±1.467	12.4	0.000455
BTB-00825210	6	43,296,255	0.43	0.49	5.24 ±1.438	13.28	0.000285
BFGL-NGS-112781	6	45,334,394	0.33	0.42	5.812 ±1.458	15.89	7.33E-05
Hapmap55150-rs29025709	6	46,068,197	0.32	0.44	5.726 ±1.517	14.25	0.000172
Hapmap52362-ss46526804	6	46,522,536	0.44	0.52	5.439 ±1.469	13.71	0.000228
Hapmap49465-BTA-11618	6	51,685,676	0.3	0.41	-6.139 ±1.592	14.86	0.000127
Hapmap54653-rs29025767	6	60,652,061	0.4	0.48	-5.262 ±1.424	13.66	0.000234
ARS-BFGL-NGS-62682	6	61,252,856	0.21	0.33	7.65 ±1.751	19.09	1.43E-05*
BFGL-NGS-115746	6	66,459,822	0.25	0.38	6.615 ±1.617	16.74	4.71E-05
BTB-00987135	6	91,388,558	0.11	0.2	7.374 ±2.239	10.85	0.001033
ARS-BFGL-NGS-18900	7	92,033,645	0.49	0.5	5.997 ±1.437	17.41	3.36E-05*
BTB-01532239	14	22,634,364	0.22	0.36	-6.506 ±1.724	14.24	0.000172
BTB-01530788	14	22,720,374	0.21	0.35	-10.07 ±1.702	35	4.84E-09*
BTB-01530836	14	22,768,981	0.2	0.33	-10.28 ±1.749	34.56	6.01E-09*
BTB-00557585	14	22,803,367	0.14	0.25	-10.54 ±2.042	26.64	3.08E-07*
BTB-00557532	14	22,838,802	0.14	0.26	-9.844 ±1.988	24.53	8.88E-07*

Hapmap43265-BTA-42290	14	30,005,787	0.2	0.31	6.8 ±1.727	15.5	8.97E-05
ARS-BFGL-NGS-37733	14	35,349,697	0.18	0.3	-7.943 ±1.913	17.24	3.68E-05*
BTB-01327818	17	61,598,441	0.11	0.21	8.652 ±2.225	15.12	0.000109
BTB-00681880	17	61,849,917	0.12	0.23	9.467 ±2.198	18.56	1.87E-05*
BTB-00681858	17	61,879,045	0.14	0.24	8.744 ±2.117	17.06	4.03E-05
BTB-00681839	17	61,905,115	0.14	0.25	8.425 ±2.085	16.33	5.87E-05
BTB-00681799	17	61,968,093	0.14	0.24	8.744 ±2.117	17.06	4.03E-05
ARS-BFGL-BAC-30721	24	3,829,304	0.13	0.23	8.387 ±2.153	15.17	0.000107
Hapmap54266-rs29023063	24	9,303,210	0.27	0.42	7.081 ±1.672	17.93	2.56E-05*
Hapmap42436-BTA-114985	24	26,665,511	0.08	0.14	12.05 ±2.653	20.62	6.49E-06*
ARS-BFGL-NGS-1585	24	26,691,760	0.08	0.14	12.05 ±2.653	20.62	6.49E-06*
ARS-BFGL-NGS-46579	26	39,777,201	0.09	0.17	10.63 ±2.511	17.92	2.59E-05*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip; *Significant at genome-wise FDR $P < 0.05$.

Table 5 – 4. Chromosome-wise significant (FDR $P < 0.05$) SNPs for carcass average backfat (CABF) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-NGS-12099	2	65,109,733	0.49	0.48	0.8301 ±0.1965	17.84	2.71E-05
BFGL-NGS-110609	2	68,908,330	0.09	0.17	-1.497 ±0.3497	18.33	2.13E-05
Hapmap47231-BTA-27375	2	71,545,226	0.13	0.25	1.311 ±0.2992	19.2	1.34E-05
ARS-BFGL-NGS-45457	6	38,102,327	0.49	0.53	-0.8918 ±0.2089	18.22	2.21E-05
Hapmap27083-BTC-041166	6	38,212,941	0.49	0.53	0.7963 ±0.2088	14.55	0.000147218
Hapmap23507-BTC-041133	6	38,233,088	0.47	0.54	0.8082 ±0.214	14.26	0.000171792
Hapmap31285-BTC-041097	6	38,256,889	0.5	0.53	-0.8844 ±0.2077	18.13	2.31E-05
Hapmap33628-BTC-041023	6	38,326,147	0.49	0.53	-0.8918 ±0.2089	18.22	2.21E-05
Hapmap27529-BTC-050639	6	38,420,476	0.28	0.43	0.9219 ±0.2341	15.51	8.95E-05
Hapmap27537-BTC-060891	6	38,638,962	0.44	0.53	0.8322 ±0.2101	15.69	8.11E-05
Hapmap31044-BTC-071337	6	38,729,866	0.45	0.53	0.7743 ±0.2098	13.63	0.000237293
BTB-00250665	6	41,506,939	0.42	0.54	0.8186 ±0.215	14.5	0.000152173
ARS-BFGL-NGS-72838	6	118,790,441	0.46	0.51	-0.8447 ±0.2055	16.9	4.39E-05
Hapmap53271-rs29022375	13	65,631,645	0.35	0.48	-1.028 ±0.2234	21.18	4.95E-06

[1]SNP identification number on the Illumina BovineSNP50 BeadChip.

Table 5 – 5. Chromosome-wise significant (FDR $P < 0.05$) SNPs for carcass ribeye area (CREA) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
Hapmap52536-rs29011297	1	43,082,466	0.20	0.31	2.325 ±0.5248	19.63	1.07E-05*
BFGL-NGS-115117	2	2,008,108	0.21	0.33	-2.167 ±0.5482	15.63	8.51E-05
Hapmap43973-BTA-93569	2	2,874,374	0.39	0.49	1.644 ±0.4543	13.09	0.000317
ARS-BFGL-NGS-102353	2	3,983,056	0.39	0.51	-1.734 ±0.4591	14.27	0.00017
Hapmap57611-rs29021061	2	5,735,608	0.41	0.51	1.663 ±0.4445	14	0.000196
ARS-BFGL-BAC-2576	2	6,095,949	0.25	0.38	1.833 ±0.4931	13.82	0.000216
ARS-BFGL-NGS-17147	2	6,502,336	0.27	0.40	2.13 ±0.4871	19.11	1.41E-05*
Hapmap23397-BTA-133598	2	10,092,609	0.30	0.43	-1.7 ±0.4758	12.76	0.000377
BTB-00084737	2	22,987,199	0.11	0.20	-2.607 ±0.7291	12.79	0.000372
BTB-00088008	2	26,936,940	0.39	0.49	-1.705 ±0.4477	14.5	0.000151
BTA-46926-no-rs	2	26,979,349	0.12	0.21	2.587 ±0.6558	15.56	8.68E-05
BTB-00091527	2	27,205,349	0.17	0.27	-2.471 ±0.6015	16.88	4.45E-05
BTB-00089278	2	27,711,955	0.38	0.47	1.718 ±0.4409	15.19	0.000106
ARS-BFGL-BAC-33842	2	31,996,368	0.24	0.39	-1.877 ±0.5179	13.13	0.000309
BTB-00088688	2	35,323,343	0.29	0.40	1.92 ±0.4574	17.62	3.01E-05*
ARS-BFGL-BAC-31338	2	40,343,813	0.25	0.40	2.428 ±0.4945	24.11	1.10E-06*
BTA-49947-no-rs	2	45,421,998	0.11	0.21	-2.596 ±0.6846	14.38	0.000161
Hapmap41882-BTA-47592	2	48,587,809	0.19	0.31	2.453 ±0.5768	18.08	2.43E-05*
Hapmap52345-rs29024293	2	60,249,495	0.24	0.36	-1.924 ±0.5265	13.36	0.000276
ARS-BFGL-NGS-5970	2	117,446,980	0.08	0.16	2.809 ±0.7894	12.66	0.000396
BFGL-NGS-112252	3	1,480,364	0.06	0.12	-4.19 ±0.9453	19.64	1.08E-05*
BTA-25432-no-rs	3	6,553,148	0.18	0.31	2.286 ±0.5473	17.44	3.29E-05*
BFGL-NGS-114514	4	1,815,714	0.21	0.32	1.711 ±0.4929	12.04	0.000549
BTA-65897-no-rs	4	25,976,226	0.28	0.41	-2.035 ±0.4982	16.69	4.9E-05
BTB-00169916	4	26,216,280	0.16	0.26	2.186 ±0.609	12.89	0.000353
BTB-01186320	4	27,002,372	0.27	0.39	-1.702 ±0.4831	12.41	0.000452
BTA-23043-no-rs	4	29,463,119	0.50	0.49	1.82 ±0.4525	16.17	6.45E-05
BTB-00171649	4	29,840,416	0.31	0.45	-1.931 ±0.4944	15.25	0.000104
ARS-BFGL-NGS-26477	4	29,863,125	0.47	0.52	1.586 ±0.4312	13.53	0.000251
BTA-96367-no-rs	4	33,692,196	0.13	0.22	2.519 ±0.6292	16.03	6.83E-05
BTB-00174592	4	34,362,493	0.12	0.21	2.541 ±0.6644	14.63	0.000142
ARS-BFGL-NGS-24321	4	63,394,829	0.25	0.41	-1.896 ±0.5381	12.42	0.000452
Hapmap40188-BTA-71028	4	69,497,399	0.36	0.47	-1.64 ±0.4624	12.58	0.000414
BTA-106667-no-rs	4	70,711,368	0.19	0.31	2.002 ±0.5575	12.89	0.000351
BTA-106668-no-rs	4	70,732,467	0.18	0.30	2.085 ±0.5485	14.45	0.000155
BTB-01162076	4	82,920,744	0.19	0.32	1.888 ±0.5425	12.12	0.000525
BTB-01943785	4	86,777,167	0.35	0.48	1.773 ±0.462	14.72	0.000135
Hapmap50629-BTA-25069	4	89,420,236	0.14	0.25	2.16 ±0.6067	12.68	0.000391
BTA-25084-no-rs	4	89,484,089	0.44	0.50	-1.557 ±0.4233	13.53	0.00025
Hapmap60639-rs29020598	4	102,242,253	0.36	0.48	1.719 ±0.4624	13.82	0.000216
BFGL-NGS-115811	4	106,034,362	0.24	0.37	2.198 ±0.5103	18.56	1.86E-05*
BTA-72175-no-rs	4	106,127,725	0.10	0.19	2.493 ±0.6877	13.14	0.000307
Hapmap32072-BTA-142491	4	106,961,853	0.35	0.44	-1.653 ±0.4403	14.09	0.000187
BFGL-NGS-112115	4	110,748,286	0.30	0.41	-1.818 ±0.4682	15.08	0.000112
ARS-BFGL-NGS-100194	4	118,243,167	0.37	0.47	2.181 ±0.4495	23.54	1.49E-06*
Hapmap54435-rs29022514	6	18,153,463	0.28	0.41	-1.803 ±0.4774	14.26	0.000172
BFGL-NGS-114096	6	23,878,186	0.43	0.50	-1.463 ±0.4529	10.44	0.001289
ARS-BFGL-NGS-55441	6	24,182,983	0.12	0.21	2.243 ±0.6301	12.68	0.000391
ARS-BFGL-NGS-63778	6	24,220,361	0.32	0.44	-1.533 ±0.4562	11.29	0.000817
ARS-BFGL-NGS-18384	6	24,251,269	0.18	0.30	2.019 ±0.5485	13.55	0.000248
Hapmap27407-BTA-143867	6	26,537,285	0.22	0.34	1.618 ±0.5098	10.07	0.001564
Hapmap50091-BTA-75608	6	27,831,792	0.49	0.52	1.391 ±0.4395	10.01	0.001616
Hapmap55575-rs29016266	6	31,006,415	0.43	0.48	1.363 ±0.4338	9.87	0.001743
ARS-BFGL-NGS-105606	6	31,285,197	0.25	0.40	-1.631 ±0.5039	10.47	0.001263
ARS-BFGL-NGS-12563	6	31,783,985	0.44	0.51	1.413 ±0.4544	9.67	0.001939
Hapmap49740-BTA-75691	6	31,807,490	0.22	0.33	1.59 ±0.5068	9.84	0.001769
Hapmap59443-rs29009843	6	31,937,563	0.32	0.43	-1.778 ±0.4571	15.14	0.000108

ARS-BFGL-NGS-28684	6	32,765,343	0.17	0.29	2.193 ±0.5706	14.77	0.000131
BFGL-NGS-114855	6	36,088,413	0.24	0.37	-1.635 ±0.496	10.87	0.00102
ARS-BFGL-NGS-49600	6	36,280,417	0.36	0.46	-1.378 ±0.4365	9.97	0.001649
Hapmap26264-BTC-037159	6	36,969,462	0.26	0.39	-2.17 ±0.4755	20.83	5.80E-06*
Hapmap33644-BTC-062485	6	37,045,881	0.19	0.30	-2.096 ±0.5308	15.59	8.56E-05
Hapmap43675-BTA-75814	6	37,231,101	0.49	0.49	-1.308 ±0.4209	9.66	0.001951
Hapmap54103-rs29010895	6	37,433,107	0.43	0.51	1.831 ±0.447	16.77	4.67E-05
Hapmap30134-BTC-034283	6	37,852,400	0.32	0.46	-1.801 ±0.4814	14	0.000196
Hapmap26308-BTC-057761	6	37,963,147	0.41	0.51	1.69 ±0.4463	14.34	0.000164
ARS-BFGL-NGS-45457	6	38,102,327	0.49	0.53	1.408 ±0.4476	9.9	0.001714
Hapmap33630-BTC-041044	6	38,301,284	0.18	0.29	-1.913 ±0.5595	11.69	0.00066
Hapmap33628-BTC-041023	6	38,326,147	0.49	0.53	1.408 ±0.4476	9.9	0.001714
Hapmap27529-BTC-050639	6	38,420,476	0.28	0.43	-1.854 ±0.4999	13.75	0.000223
Hapmap31044-BTC-071337	6	38,729,866	0.45	0.53	-1.662 ±0.4466	13.85	0.000211
Hapmap26618-BTC-070864	6	38,982,338	0.34	0.47	-1.762 ±0.4467	15.57	8.63E-05
Hapmap23923-BTC-066021	6	39,108,078	0.14	0.25	-2.232 ±0.6426	12.07	0.000541
BTB-00260450	6	39,509,020	0.37	0.50	-2.03 ±0.4547	19.94	9.13E-06*
Hapmap27299-BTC-035816	6	39,794,334	0.42	0.50	1.4 ±0.4509	9.63	0.001982
Hapmap33744-BTC-050901	6	40,542,918	0.50	0.51	-1.669 ±0.43	15.05	0.000113
BTB-00250403	6	41,568,481	0.18	0.30	-1.912 ±0.5795	10.89	0.001011
BTA-75905-no-rs	6	41,680,774	0.21	0.36	2.603 ±0.5621	21.44	4.30E-06*
BTB-00251835	6	42,157,836	0.18	0.31	-2.089 ±0.5946	12.34	0.00047
BTA-95818-no-rs	6	42,786,669	0.41	0.47	1.428 ±0.4207	11.53	0.000718
Hapmap47224-BTA-24614	6	43,071,076	0.43	0.49	1.498 ±0.4377	11.71	0.000655
ARS-BFGL-NGS-1611	6	43,253,609	0.37	0.46	-1.69 ±0.4473	14.27	0.00017
BTB-00825210	6	43,296,255	0.43	0.49	1.611 ±0.4292	14.08	0.000188
ARS-BFGL-NGS-23937	6	43,378,454	0.20	0.35	1.861 ±0.5392	11.91	0.000587
BTB-01893222	6	43,660,166	0.25	0.36	1.564 ±0.4737	10.9	0.001003
ARS-BFGL-NGS-2525	6	45,627,476	0.31	0.46	-2.06 ±0.4855	17.99	2.49E-05*
Hapmap55150-rs29025709	6	46,068,197	0.32	0.44	1.964 ±0.4511	18.96	1.51E-05*
ARS-BFGL-NGS-39570	6	46,320,086	0.07	0.13	2.63 ±0.812	10.49	0.00125
ARS-BFGL-NGS-21182	6	46,429,903	0.31	0.45	-2.091 ±0.4786	19.09	1.41E-05*
Hapmap52362-ss46526804	6	46,522,536	0.44	0.52	2.035 ±0.436	21.78	3.58E-06*
ARS-BFGL-NGS-41037	6	48,639,804	0.34	0.46	-1.672 ±0.4533	13.61	0.00024
BTB-01794972	6	51,709,445	0.15	0.26	-1.88 ±0.6014	9.77	0.001838
ARS-BFGL-NGS-12583	6	51,749,549	0.08	0.14	-2.556 ±0.8147	9.85	0.001763
Hapmap31996-BTC-066011	6	53,329,685	0.21	0.38	2.037 ±0.5312	14.71	0.000135
Hapmap23860-BTC-065677	6	53,445,430	0.26	0.36	1.524 ±0.4706	10.48	0.001256
ARS-BFGL-NGS-60567	6	53,852,962	0.22	0.37	1.899 ±0.5187	13.41	0.000266
Hapmap32218-BTC-040966	6	53,921,383	0.41	0.48	-1.415 ±0.4435	10.18	0.001476
ARS-BFGL-NGS-2697	6	55,121,390	0.40	0.51	-1.878 ±0.4585	16.78	4.63E-05
Hapmap54653-rs29025767	6	60,652,061	0.40	0.48	-1.7 ±0.4241	16.06	6.7E-05
ARS-BFGL-NGS-62682	6	61,252,856	0.21	0.33	1.687 ±0.5269	10.25	0.001426
Hapmap51312-BTA-76569	6	61,532,962	0.17	0.29	1.856 ±0.5889	9.93	0.001686
Hapmap51763-BTA-76571	6	61,569,918	0.21	0.33	2.451 ±0.5071	23.37	1.60E-06*
ARS-BFGL-NGS-97839	6	61,802,199	0.37	0.48	1.495 ±0.453	10.9	0.001005
ARS-BFGL-NGS-106015	6	61,938,805	0.38	0.48	1.895 ±0.4486	17.85	2.67E-05*
Hapmap38694-BTA-76566	6	62,334,030	0.44	0.52	2.089 ±0.4544	21.12	5.02E-06*
BFGL-NGS-114148	6	62,985,094	0.29	0.40	-1.48 ±0.4615	10.28	0.001398
Hapmap31190-BTA-161167	6	64,741,834	0.35	0.48	-1.532 ±0.4624	10.98	0.000963
BTA-76599-no-rs	6	65,862,811	0.13	0.23	-2.215 ±0.689	10.34	0.001364
BTA-122855-no-rs	6	67,190,573	0.44	0.51	-1.66 ±0.4573	13.17	0.000304
Hapmap60836-rs29027147	6	71,189,704	0.37	0.53	1.614 ±0.4734	11.62	0.000685
Hapmap54879-rs29017018	6	71,217,153	0.40	0.54	1.755 ±0.4676	14.08	0.000188
ARS-BFGL-NGS-69314	6	74,871,315	0.23	0.35	1.608 ±0.4982	10.42	0.001298
Hapmap50775-BTA-76767	6	74,911,168	0.41	0.50	1.399 ±0.4357	10.32	0.001368
BTB-00263321	6	75,030,032	0.31	0.46	1.468 ±0.4697	9.77	0.001837
ARS-BFGL-NGS-93711	6	75,063,474	0.23	0.35	1.761 ±0.4981	12.5	0.00043
ARS-BFGL-NGS-14715	6	75,359,134	0.19	0.32	1.8 ±0.5608	10.31	0.001375
ARS-BFGL-NGS-27643	6	80,157,015	0.13	0.23	2.498 ±0.6382	15.32	9.86E-05
Hapmap41083-BTA-76098	6	80,715,299	0.41	0.53	1.701 ±0.4775	12.69	0.000391
BTB-00264414	6	81,296,477	0.21	0.36	1.734 ±0.5343	10.54	0.001216

ARS-BFGL-NGS-2935	6	92,835,501	0.49	0.50	1.505 ±0.4374	11.84	0.000613
ARS-BFGL-NGS-21859	6	115,332,735	0.14	0.25	-2.62 ±0.6752	15.06	0.000117
ARS-BFGL-NGS-1155	6	115,635,484	0.20	0.34	-1.94 ±0.5884	10.88	0.001042
ARS-BFGL-NGS-43597	6	118,232,278	0.07	0.14	-2.932 ±0.9175	10.21	0.001482
ARS-BFGL-NGS-88797	6	118,267,986	0.12	0.22	-2.187 ±0.6933	9.95	0.001678
ARS-BFGL-NGS-56487	6	119,339,222	0.20	0.32	-2.033 ±0.5651	12.94	0.000344
BFGL-NGS-113112	6	120,374,438	0.21	0.33	-1.854 ±0.5369	11.92	0.000591
ARS-BFGL-NGS-17086	6	120,450,992	0.28	0.42	-1.719 ±0.5109	11.32	0.000806
ARS-BFGL-NGS-54737	6	120,570,412	0.39	0.50	-1.546 ±0.47	10.82	0.001053
ARS-BFGL-NGS-41730	6	122,509,741	0.21	0.34	-1.801 ±0.5703	9.98	0.001654
ARS-BFGL-NGS-101886	7	91,903,228	0.44	0.49	-1.845 ±0.426	18.75	1.68E-05*
ARS-BFGL-NGS-104947	8	29,550,270	0.27	0.39	2.3 ±0.5041	20.82	6.09E-06*
ARS-BFGL-NGS-90334	11	103,403,245	0.05	0.11	4.053 ±0.9252	19.19	1.34E-05*
BTB-00620537	15	73,686,200	0.49	0.54	-1.981 ±0.4619	18.4	2.05E-05*
BFGL-NGS-113600	15	73,715,150	0.35	0.45	2.087 ±0.4565	20.91	5.65E-06*
ARS-BFGL-NGS-3016	19	5,395,507	0.06	0.12	4.374 ±1.166	14.07	0.000228
ARS-BFGL-NGS-14867	19	6,847,978	0.22	0.34	2.333 ±0.5116	20.8	5.93E-06*
Hapmap60292-rs29022330	19	19,536,785	0.23	0.35	1.924 ±0.5313	13.11	0.000314
ARS-BFGL-NGS-28901	19	43,671,889	0.48	0.50	1.609 ±0.4379	13.51	0.000256
BTB-00758263	19	50,638,953	0.33	0.40	1.649 ±0.4446	13.75	0.000224
Hapmap31094-BTA-133271	19	50,781,403	0.44	0.48	1.829 ±0.4216	18.82	1.63E-05*
BFGL-NGS-112313	19	52,647,017	0.48	0.49	1.754 ±0.4265	16.91	4.35E-05
ARS-BFGL-NGS-47995	19	52,695,896	0.47	0.48	1.783 ±0.4245	17.63	3E-05*
ARS-BFGL-NGS-88163	19	52,721,943	0.46	0.50	2.172 ±0.4312	25.38	5.99E-07*
ARS-BFGL-NGS-68655	19	53,020,578	0.30	0.40	1.833 ±0.4636	15.63	8.51E-05
ARS-BFGL-NGS-28507	19	55,298,722	0.33	0.45	-1.614 ±0.4569	12.48	0.000436
Hapmap58450-rs29021989	25	33,719,124	0.44	0.49	-2.368 ±0.4479	27.96	1.68E-07*

[1]SNP identification number on the Illumina BovineSNP50 BeadChip; *Significant at genome-wise FDR $P < 0.05$.

Table 5 – 6. Chromosome-wise significant (FDR $P < 0.05$) SNPs for carcass grade fat (CGF) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-NGS-12099	2	65,109,733	0.49	0.48	0.832 ±0.196	18.12	2.34E-05
BFGL-NGS-110609	2	68,908,330	0.09	0.17	-1.442 ±0.347	17.27	3.67E-05
Hapmap47231-BTA-27375	2	71,545,226	0.13	0.25	1.418 ±0.297	22.85	2.09E-06
BTA-28730-no-rs	8	12,699,662	0.34	0.45	0.748 ±0.21	12.76	0.000376
BTA-18947-no-rs	8	12,720,834	0.40	0.49	0.713 ±0.205	12.09	0.000535
BTA-92138-no-rs	8	14,684,960	0.28	0.40	0.9 ±0.224	16.2	6.26E-05
BTB-00225031	8	15,494,877	0.12	0.22	1.153 ±0.315	13.39	0.000271
BTB-00095463	8	15,697,935	0.08	0.16	1.543 ±0.367	17.68	2.95E-05
ARS-BFGL-NGS-15251	8	15,913,957	0.47	0.44	0.747 ±0.192	15.06	0.000114
BTB-01925657	8	15,940,103	0.48	0.53	-0.884 ±0.208	18.15	2.31E-05
BFGL-NGS-117504	8	30,247,817	0.16	0.27	1.143 ±0.263	18.83	1.61E-05
BTB-00339941	8	30,276,201	0.16	0.27	1.143 ±0.263	18.83	1.61E-05
BTA-80950-no-rs	8	38,510,343	0.41	0.48	-0.699 ±0.2	12.28	0.000484
BTB-01669383	8	42,929,958	0.17	0.28	0.923 ±0.262	12.4	0.000455
BTA-107830-no-rs	8	61,733,553	0.08	0.14	1.82 ±0.379	23.07	1.87E-06
BTB-00349926	8	61,756,588	0.21	0.33	0.943 ±0.25	14.19	0.000178
BTB-00350133	8	62,123,766	0.08	0.14	1.769 ±0.377	22.06	3.11E-06
BTA-81286-no-rs	8	62,178,362	0.08	0.14	1.769 ±0.377	22.06	3.11E-06
BTB-00350198	8	62,251,314	0.08	0.14	1.769 ±0.377	22.06	3.11E-06
BTB-00350366	8	62,315,972	0.08	0.14	1.769 ±0.377	22.06	3.11E-06
BTB-01552322	8	63,114,323	0.08	0.14	1.629 ±0.359	20.57	6.64E-06
BTB-00351461	8	64,164,842	0.08	0.15	1.644 ±0.377	19.04	1.45E-05
ARS-BFGL-NGS-13571	8	89,506,070	0.25	0.39	0.938 ±0.236	15.79	7.72E-05
BTA-82770-no-rs	8	112,255,430	0.36	0.47	0.794 ±0.214	13.83	0.000215
Hapmap53271-rs29022375	13	65,631,645	0.35	0.48	-1.073 ±0.222	23.44	1.58E-06
BTB-01014755	29	27,607,708	0.05	0.10	1.762 ±0.436	16.38	5.7E-05

[1]SNP identification number on the Illumina BovineSNP50 BeadChip.

Table 5 – 7. Chromosome-wise significant (FDR $P < 0.05$) SNPs for carcass lean meat yield (CLMY) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-NGS-52961	2	7,879,761	0.14	0.26	1.279 ±0.296	18.68	1.86E-05
ARS-BFGL-BAC-29585	2	31,256,175	0.42	0.50	0.735 ±0.187	15.41	9.42E-05
BTB-01361440	2	61,919,313	0.28	0.40	0.822 ±0.206	15.95	7.21E-05
BFGL-NGS-115485	2	65,854,004	0.46	0.49	0.729 ±0.181	16.15	6.43E-05
BFGL-NGS-110609	2	68,908,330	0.09	0.17	1.338 ±0.321	17.41	3.38E-05
Hapmap47231-BTA-27375	2	71,545,226	0.13	0.25	-1.171 ±0.274	18.3	2.12E-05
Hapmap27529-BTC-050639	6	38,420,476	0.28	0.43	-0.882 ±0.214	17.01	4.11E-05
ARS-BFGL-NGS-2935	6	92,835,501	0.49	0.50	0.801 ±0.187	18.45	1.99E-05
BFGL-NGS-117504	8	30,247,817	0.16	0.27	-1.028 ±0.242	18.04	2.41E-05
BTB-00339941	8	30,276,201	0.16	0.27	-1.028 ±0.242	18.04	2.41E-05
Hapmap40032-BTA-93874	13	45,627,032	0.33	0.43	-0.848 ±0.189	20.21	7.95E-06
Hapmap53271-rs29022375	13	65,631,645	0.35	0.48	1.066 ±0.204	27.38	2.18E-07*
ARS-BFGL-NGS-4716	21	47,992,701	0.13	0.21	-1.171 ±0.282	17.3	3.56E-05

[1]SNP identification number on the Illumina BovineSNP50 BeadChip; *Significant at genome-wise FDR $P < 0.05$.

Table 5 – 8. Chromosome-wise significant (FDR $P < 0.05$) SNPs for ultrasound marbling (UMAR) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
ARS-BFGL-BAC-30229	2	31,607,614	0.07	0.13	0.532 ±0.127	17.68	2.87006E-05
ARS-BFGL-BAC-30252	2	31,652,458	0.07	0.13	0.532 ±0.127	17.68	2.87006E-05
ARS-BFGL-NGS-64127	2	35,515,940	0.11	0.22	0.405 ±0.086	22.19	2.85E-06
Hapmap53694-rs29015972	4	59,971,948	0.07	0.14	0.447 ±0.104	18.42	1.96027E-05
BTB-02020184	5	93,494,954	0.05	0.10	0.444 ±0.105	17.83	2.65618E-05
ARS-BFGL-NGS-25628	7	103,339,927	0.07	0.13	0.439 ±0.097	20.44	6.96E-06
BTA-85910-no-rs	7	106,943,605	0.07	0.13	0.48 ±0.101	22.48	2.46E-06
BTA-90709-no-rs	12	38,564,181	0.06	0.10	0.459 ±0.106	18.67	0.00001724
ARS-BFGL-NGS-24568	14	43,740,541	0.10	0.18	0.466 ±0.089	27.56	1.89E-07*
Hapmap51129-BTA-104919	14	57,333,346	0.18	0.32	0.293 ±0.071	17.25	0.000035836
ARS-BFGL-NGS-26419	22	47,085,996	0.09	0.15	0.373 ±0.091	16.75	4.63993E-05

[1]SNP identification number on the Illumina BovineSNP50 BeadChip; *Significant at genome-wise FDR $P < 0.05$.

Table 5 – 9. Chromosome-wise significant (FDR $P < 0.05$) SNPs for ultrasound backfat (UBF) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
BTB-01089722	8	58,476,229	0.20	0.32	0.043 ±0.011	15.53	8.74E-05
ARS-BFGL-NGS-108096	8	59,780,226	0.50	0.52	0.033 ±0.009	14.84	0.000125
ARS-BFGL-NGS-96805	8	69,820,394	0.40	0.49	-0.041 ±0.009	21.72	3.62E-06
ARS-BFGL-NGS-29663	8	69,840,958	0.40	0.49	-0.042 ±0.009	22.45	2.50E-06
ARS-BFGL-NGS-48897	8	72,136,388	0.35	0.45	0.037 ±0.009	16.24	6.04E-05
Hapmap38700-BTA-81760	8	72,160,479	0.35	0.45	0.035 ±0.009	14.18	0.000177
ARS-BFGL-NGS-11101	8	83,907,555	0.08	0.15	-0.065 ±0.017	14.83	0.000126
Hapmap53638-rs29019444	8	88,011,068	0.09	0.16	-0.058 ±0.016	13.26	0.000287
ARS-BFGL-NGS-66366	8	88,574,678	0.07	0.14	0.062 ±0.017	13.35	0.000273
BTB-01235513	8	88,644,805	0.11	0.20	0.063 ±0.014	19.32	1.23E-05
ARS-BFGL-NGS-13571	8	89,506,070	0.25	0.39	0.045 ±0.01	18.63	1.76E-05
BTB-01266056	8	98,656,165	0.15	0.25	0.049 ±0.012	16.09	6.54E-05
Hapmap46799-BTA-32094	13	35,739,205	0.14	0.26	-0.058 ±0.014	18.09	2.32E-05
ARS-BFGL-NGS-12144	13	68,440,299	0.33	0.44	-0.037 ±0.009	16.42	5.5E-05
Hapmap39724-BTA-122305	20	34,953,908	0.40	0.50	-0.039 ±0.009	17.91	2.55E-05
ARS-BFGL-NGS-2398	21	4,034,246	0.23	0.36	-0.048 ±0.011	18.8	1.62E-05
BFGL-NGS-112796	22	2,700,147	0.31	0.42	0.044 ±0.01	20.93	5.42E-06
ARS-BFGL-NGS-40920	22	51,422,755	0.23	0.36	0.043 ±0.011	15.53	8.74E-05
ARS-BFGL-NGS-35019	23	52,726,475	0.39	0.49	0.041 ±0.009	20.15	8.07E-06
ARS-BFGL-NGS-102613	28	26,194,404	0.44	0.51	0.044 ±0.009	23.82	1.25E-06
ARS-BFGL-NGS-30132	28	27,173,775	0.28	0.44	0.04 ±0.01	15.31	9.8E-05
BFGL-NGS-111957	28	27,438,070	0.28	0.41	0.038 ±0.01	15.67	8.13E-05
Hapmap48728-BTA-36414	28	29,682,393	0.33	0.42	0.037 ±0.009	16.52	5.23E-05
BTB-00987935	28	34,653,168	0.49	0.50	0.033 ±0.009	13.38	0.000269

[1]SNP identification number on the Illumina BovineSNP50 BeadChip.

Table 5 – 10. Chromosome-wise significant (FDR $P < 0.05$) SNPs for ultrasound ribeye area (UREA) using single marker LD regression in commercial hybrid beef cattle.

SNP_ID[1]	BTA	Position(bp)	MAF	Heterozygosity	Effect ± SD	F-test	*P*-value
Hapmap51692-BTA-119034	2	84,096,221	0.33	0.44	-1.745 ±0.405	18.54	0.00001851
ARS-BFGL-NGS-18347	2	94,256,195	0.16	0.27	-2.005 ±0.484	17.13	3.82E-05
BTB-00219600	5	11,276,200	0.29	0.41	-1.61 ±0.426	14.31	0.000165915
BTB-00219662	5	11,304,815	0.29	0.41	-1.61 ±0.426	14.31	0.000165915
Hapmap42658-BTA-74674	5	11,407,908	0.29	0.41	-1.61 ±0.426	14.31	0.000165915
BTA-105507-no-rs	5	45,346,211	0.35	0.44	-1.473 ±0.367	16.08	6.57E-05
ARS-BFGL-NGS-2133	5	68,059,571	0.42	0.5	-1.548 ±0.39	15.79	7.65E-05
BTA-73797-no-rs	5	68,762,563	0.4	0.51	-1.398 ±0.361	14.99	0.000115758
ARS-BFGL-NGS-57509	5	108,835,957	0.22	0.34	2.042 ±0.433	22.23	2.79E-06
ARS-BFGL-NGS-25950	5	108,980,836	0.48	0.54	-1.615 ±0.368	19.26	1.27E-05
ARS-BFGL-NGS-72188	6	41,541,414	0.33	0.45	1.737 ±0.385	20.36	7.26E-06
BFGL-NGS-118172	8	56,621,847	0.38	0.47	-1.507 ±0.377	15.99	6.89E-05
Hapmap53638-rs29019444	8	88,011,068	0.09	0.16	-3.002 ±0.644	21.72	3.64E-06
BTB-00374945	8	112,756,815	0.2	0.33	1.976 ±0.472	17.55	3.07E-05
BTB-01113508	12	1,911,991	0.07	0.14	-2.822 ±0.674	17.51	3.14E-05
BTA-06771-rs29021128	13	29,247,337	0.12	0.22	-2.56 ±0.564	20.58	6.53E-06
Hapmap51588-BTA-36751	15	37,907,640	0.42	0.48	1.57 ±0.366	18.42	1.96E-05
Hapmap39724-BTA-122305	20	34,953,908	0.4	0.5	-1.521 ±0.378	16.22	6.11E-05
Hapmap51600-BTA-50467	20	38,936,262	0.29	0.43	-1.667 ±0.403	17.08	3.91E-05
Hapmap49835-BTA-104494	20	40,167,239	0.13	0.2	-2.509 ±0.527	22.65	2.27E-06
BTA-89598-no-rs	20	57,735,717	0.09	0.17	-2.299 ±0.593	15.03	0.000113369
ARS-BFGL-NGS-19917	22	48,595,383	0.45	0.5	-1.525 ±0.369	17.1	3.88E-05
ARS-BFGL-NGS-16330	22	50,674,396	0.22	0.36	1.785 ±0.44	16.43	5.48E-05
ARS-BFGL-NGS-5577	26	22,648,915	0.44	0.51	1.558 ±0.388	16.12	6.47E-05
ARS-BFGL-NGS-36064	26	22,691,335	0.48	0.49	1.524 ±0.37	16.97	4.16E-05

[1]SNP identification number on the Illumina BovineSNP50 BeadChip.

Table 5 – 11. Chromosomal distribution of significant SNPs for beef ultrasound and carcass merit traits in commercial hybrid beef cattle using single marker LD regression.

Trait[1]	N*	BTA(N)	N**	BTA(N)
CWT	31	2(1), 6(19), 7(1), 14(5), 17(1), 24(3), 26(1)	67	2(1), 6(50), 7(1), 14(7), 17(5), 24(4), 26(1)
CABF	0	--	14	2(3), 6(10), 13(1)
CREA	29	1(1), 2(4), 3(2), 4(2), 6(10), 7(1), 8(1), 11(1), 15(2), 19(4), 25(1)	146	1(1), 2(19), 3(2), 4(23), 6(84), 7(1), 8(1), 11(1), 15(2), 19(11), 25(1)
CGF	0	--	26	2(3), 8(21), 13(1), 29(1)
CLMY	1	13(1)	13	2(6), 6(2), 8(2), 13(2), 21(1)
UMAR	1	14(1)	11	2(3), 4(1), 5(1), 7(2), 12(1), 14(2), 22(1)
UBF	0	--	24	8(12), 13(2), 20(1), 21(1), 22(2), 23(1), 28(5)
UREA	0	--	25	2(2), 5(8), 6(1), 8(3), 12(1), 13(1), 15(1), 20(4), 22(2), 26(2)

[1]CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2). *Significant at chromosome-wise FDR $P < 0.05$; **Significant at genome-wise FDR $P < 0.05$.

Table 5 – 12. Comparison of the significant SNPs from single marker LD regression in commercial hybrid beef cattle with previously reported QTL for beef carcass traits in different beef cattle breeds.

Trait[1]	BTA	SNP position (Mb)	N-SNPs	Previously reported QTL peak (cM) [2]	Previously reported QTL interval (cM) [2] or gene[3]	Type[4]	Reference
CWT	2	25.17	1	19.71	11.912-23.11	Close	(KIM *et al.* 2003)
	6	6.99 – 32.77	6	-	-	-	-
		36.09 – 66.46	43	38	-	Overlap	(TAKASUGA *et al.* 2007)
				39.1	NCAPG	Overlap	(SETOGUCHI *et al.* 2009)
				42	37-55	Overlap	(SETOGUCHI *et al.* 2009)
				51	48-58	Overlap	(CASAS *et al.* 2000)
				58	43-75	Overlap	(SETOGUCHI *et al.* 2009)
		91.39	1	94	89.35-101.4	Overlap	(MCCLURE *et al.* 2010)
	7	92.03	1	85	77.19-90.70	Close	(MCCLURE *et al.* 2010)
	14	22.63 – 22.84	5	26	-	Close	(TAKASUGA *et al.* 2007)
				39	17.84-43.63	Overlap	(MCCLURE *et al.* 2010)
		30.01, 35.35	2	34	-	Close	(TAKASUGA *et al.* 2007)
				36	-	Overlap	(TAKASUGA *et al.* 2007)
				39	17.84-43.63	Overlap	(MCCLURE *et al.* 2010)
	17	61.60 – 61.97	5	63	57.09-80.85	Overlap	(MCCLURE *et al.* 2010)
	24	3.83, 9.30	2	14	1-16.33	Overlap	(MCCLURE *et al.* 2010)
		26.67, 26.69	2	27	23.68-30.53	Overlap	(MCCLURE *et al.* 2010)
	26	39.78	1	46	42.48-52.45	Close	(MCCLURE *et al.* 2010)
CABF	2	65.11 – 71.55	3	54	21-60	Close	(CASAS *et al.* 2003)
	6	38.10 – 41.51	9	36	17.00-43.93	Overlap	(MCCLURE *et al.* 2010)
		118.79	1	113	89.35-118.0	Overlap	(MCCLURE *et al.* 2010)
	13	65.63	1	67	62.80-73.63	Overlap	(MCCLURE *et al.* 2010)
CREA	1	43.08	1	-	-	-	-
	2	2.01 – 10.09	7	1.1	-	Close	(ABE *et al.* 2008)
				5.6	MSTN	Overlap	(MARTINEZ *et al.* 2010)
				5.6	MSTN	Overlap	(ESMAILIZADEH *et al.* 2008)
				5.6	MSTN	Overlap	(ALLAIS *et al.* 2010)
				6.3	3.856-10.772	Overlap	(CASAS *et al.* 1998)
				8.6	0-10.8	Overlap	(ALEXANDER *et al.* 2007)
				10	-	Overlap	(MORRIS *et al.* 2009)
		22.99 – 60.25	11	52	-	Close	(TAKASUGA *et al.* 2007)
		117.45	1	-	-	-	-
	3	1.48, 6.55	2	2	0-9.342	Overlap	(MCCLURE *et al.* 2010)
	4	1.82	1	-	-	-	-
		25.98 – 34.36	8	-	-	-	-
		63.39	1	64	-	Overlap	(TAKASUGA *et al.* 2007)
				60	42.49-67.471	Overlap	(MIZOSHITA *et al.* 2004)
		69.50 – 89.48	7	-	-	-	-
		102.24 – 118.24	6	-	-	-	-
	6	18.15 – 32.77	13	9	0-26	Close	(CASAS *et al.* 2003)
		36.09 – 81.30	61	39.1	NCAPG	Overlap	(SETOGUCHI *et al.* 2009)
				41	34.45-43.93	Overlap	(MCCLURE *et al.* 2010)
				51	47-58	Overlap	(CASAS *et al.* 2000)
		92.84 – 122.51	10	103	89.35-118.0	Overlap	(MCCLURE *et al.* 2010)
	7	91.90	1	81	65.30-90.70	Close	(MCCLURE *et al.* 2010)
	8	29.55	1	-	-	-	-
	11	103.40	1	103	97.57-122.3	Overlap	(MCCLURE *et al.* 2010)
	15	73.69, 73.72	2	-	-	-	-
	19	5.40 – 19.54	3	-	-	-	-
		43.67 – 55.30	8	44	43.814-50	Overlap	(TAYLOR *et al.* 1998)
	25	33.72	1	-	-	-	-
CGF	2	65.11 – 71.55	3	54	21-60	Close	(CASAS *et al.* 2003)
	8	12.70 – 42.93	11	25	6-30	Overlap	(CASAS *et al.* 2001)
				30	11-47	Overlap	(CASAS *et al.* 2001)

		61.73 – 112.26	10	-	-	-	-
	13	65.63	1	67	62.80-73.63	Overlap	(McClure *et al.* 2010)
	29	27.61	1	27	19.58-29.20	Overlap	(McClure *et al.* 2010)
CLMY	2	7.88	1	6.3	3.856-10.77	Overlap	(Casas *et al.* 1998)
				5.3	-	Close	(Abe *et al.* 2008)
		31.26	1	-	-	-	-
		61.92 – 71.55	4	52	38-79	Overlap	(Casas *et al.* 2003)
	6	38.42, 92.84	2	-	-	-	-
	8	30.25, 30.28	2	-	-	-	-
	13	45.63, 65.63	2	-	-	-	-
	21	47.99	1	-	-	-	-
UMAR	2	31.61 – 35.52	3	-	-	-	-
	4	59.97	1	55	52.49-67.471	Overlap	(Mizoshita *et al.* 2004)
				56	-	Close	(Takasuga *et al.* 2007)
				66	-	Close	(Takasuga *et al.* 2007)
				44	30-88	Overlap	(Yokouchi *et al.* 2009)
	5	93.49	1	82	81.91-90.84	Close	(McClure *et al.* 2010)
	7	103.34, 106.94	2	-	-	-	-
	12	38.56	1	-	-	-	-
	14	43.74, 57.33	2	47	30-87	Overlap	(Casas *et al.* 2003)
				46.3	CRH	Close	(Wibowo *et al.* 2007)
				44.2	FABP4	Overlap	(Lee *et al.* 2010a)
	22	47.09	1	34.6	30.56-47.06	Overlap	(Gutierrez-Gil *et al.* 2009)
				47	42.37-47.06	Overlap	(McClure *et al.* 2010)
UBF	8	58.48 – 98.66	12	-	-	-	-
	13	35.74	1	25	8.993-38.65	Overlap	(McClure *et al.* 2010)
		68.44	1	67	62.80-73.63	Overlap	(McClure *et al.* 2010)
	20	34.95	1	-	-	-	-
	21	4.03	1	2	0-10.96	Overlap	(McClure *et al.* 2010)
	22	2.70, 51.42	2	-	-	-	-
	23	52.73	1	55.24	52.29-58.19	Overlap	(Li *et al.* 2004)
				59	52.29-67.93	Overlap	(McClure *et al.* 2010)
				44	37.72-52.29	Close	(McClure *et al.* 2010)
	28	26.19 – 34.65	5	23	16.06-24.77	Close	(McClure *et al.* 2010)
UREA	2	84.10, 94.26	2	-	-	-	-
	5	11.28 – 11.41	3	-	-	-	-
		45.35	1	53	38-66	Overlap	(Casas *et al.* 2003)
				40	29.42-52.09	Overlap	(McClure *et al.* 2010)
		68.06, 68.76	2	-	-	-	-
		108.84, 108.98	2	-	-	-	-
	6	41.54	1	41	34.45-43.93	Overlap	(McClure *et al.* 2010)
				39.1	NCAPG	Close	(Setoguchi *et al.* 2009)
	8	56.62	1	58	50.11-66.03	Overlap	(McClure *et al.* 2010)
		88.01	1	-	-	-	-
		112.76	1	104	92.72-118.7	Overlap	(McClure *et al.* 2010)
	12	1.91	1	9.6	0-15.11	Overlap	(Alexander *et al.* 2007)
	13	29.25	1	12	1-27.60	Close	(McClure *et al.* 2010)
	15	37.91	1	54	37.96-54.29	Close	(McClure *et al.* 2010)
	20	34.95 – 57.74	4	-	-	-	-
	22	48.60, 50.67	2	-	-	-	-
	26	22.65, 22.69	2	18	15.459-22.86	Overlap	(Stone *et al.* 1999)

[1]CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2). [2]Previous QTL information was obtained from CattleQTLdb (www.animalgenome.org/cgi-bin/QTLdb/BT/index). Significant SNPs identified for CABF, CGF, and UBF were compared with the QTL reported for fat thickness at the 12th rib; SNPs identified for CWT, CREA, CLMY, UMAR and UREA were compared with the QTL reported for carcass weight, ribeye area, yield grade, marbling score and ribeye area, respectively. [3]Genes reported within the QTL region: *CRH*: corticotropin releasing hormone; *FABP4*: fatty acid binding protein 4, adipocyte; *MSTN*: myostatin; *NCAPG*: non-SMC condensin I complex, subunit G. [4]The type of concordance with QTL in literature. For SNPs without overlapping QTL, the closest previously identified QTL was presented.

Table 5 – 13. Gene networks among positional candidate genes related to beef carcass traits in commercial hybrid beef cattle. Positional candidate genes within 0.5 Mb windows of significant SNPs (at genome-wise FDR $P < 0.05$) from the single marker LD regression for beef carcass traits were considered in the functional clustering analyses using the Ingenuity Pathway Analysis (IPA) database.

Trait[1]	Functional network	Positional candidate genes in network	N
CWT	Small Molecule Biochemistry, Cellular Development, Cell Cycle	ANAPC4, ANKRD17, ARMC1, ATOH1, B4GALT6, C18orf55, C4orf34, CHRNA9, CNDP2, COX7B, DLX1, DLX2, DNAJC5B, FABP2, FAM184B, FBXO15, FBXW8, GABRG1, GRID2, GUF1, HNF4G, KCNIP4, KLHL5, LAP3, LCORL, MANBA, MED13L, MMRN1, MYOZ2, N4BP2, NAP1L5, NCAPG, PACRGL, PDS5A, PI4K2B, PIGY, PLAG1, RGS10, RP1, SEC24D, SEPSECS, SLC39A8, SMARCAD1, SYNPO2, TGS1, TIAL1, TMEM68, TRIM55, TTR	49
	Cellular Function and Maintenance, Skeletal and Muscular System Development and Function	AFP, ALB, ARRDC3, BANK1, BMPR1B, CXCL5, DHX15, DSC1, DSC2, DSG1, DSG2, DSG3, EIF3A, EYA1, GPR98, GPR125, GRK5, HAT1, HERC3, HERC5, HERC6, HPGDS, IBSP, IL8, ITGA6, LYN, MOS, NFKB1, PDK1, PKD2, PPARGC1A, PRDX3, RAPGEF4, RHOH, RNF125, RNF138, SLC25A12, SLIT2, SNCA, SOD3, SPP1, TLR10, UBE2K, UGDH	44
CREA	Carbohydrate Metabolism, Small Molecule Biochemistry, Digestive System Development and Function	AGAP3, AKAP1, BMPR1B, C2orf88, CANT1, Chn2, CLEC2L, DDIT4L, EZH1, FEZF1, FIP1L1, FOXN1, GALNT5, HCRT, HIBADH, HS6ST1, IER5L, IGF2BP3, IGFBP7, JHDM1D, LAMTOR3, LAP3, LRP2, MMRN1, MPZL1, MXD4, NHEDC2, NOSTRIN, NPB, PACRGL, PCYT2, PDHA2, PGS1, RAB34, RFNG, RHEB, RPTOR, SCN3A, SCPEP1, SH3GLB2, SLC4A2, SLC4A10, SMARCD3, SOSTDC1, TMEM100, TNRC6C, TOR1B, UBN2, WDR91, ZC3HAV1L, ZER1	51
	Molecular Transport, Energy Production, Nucleic Acid Metabolism	ACVR1C, API5, ATP5I, ATP6V0A4, BBS5, BPTF, CCDC137, COL25A1, DCAF6, DHRS9, DHX15, DOCK10, EPHA1, FAM114A1, GABRA2, GABRG1, GABRR3, GPNMB, GPS1, HIBCH, INPP1, JAZF1, KCNH7, KIAA1549, KLHL11, KRT35, LRRC8A, MAFG, MYO3B, NAB1, NCAPG, PCTP, PDE6B, PI4K2B, PIP, PPM1K, PRKAG2, PROCA1, PSMC3IP, RAB5C, SAP130, SCFD2, SLC34A2, SMURF2, TIPRL, TMEM139, TSPAN5, TWISTNB, UNC5C, WDR19	50
	Carbohydrate Metabolism, Lipid Metabolism, Small Molecule Biochemistry	AATK, ABCG2, ACVR1, AGR2, AKR1B10, ALDOC, ANAPC4, AOAH, APBB2, ASPSCR1, AZI1, CHRNA1, CHRNA9, DMTF1, DNAJB14, DOLPP1, EPHA6, FGFRL1, G6PC2, GALNT3, GRK4, HSPB9, IDUA, LCORL, LEPREL4, LOC285141, MACC1, METTL11A, N4BP2, OCIAD1, PARM1, PIGY, PTGES, PYCR1, RGS5, SLC13A2, SLC39A8, SMARCAD1, SOD3, SORCS2, TAPT1, THOC4, TRIM24, TSPAN13, UGGT1, WHSC1	46
CGF	Amino Acid Metabolism, Small Molecule Biochemistry, Cellular Development	ACMSD, C20orf4, GDF5, GLIPR2, GRHPR, HPS5, OR8D4, PIGO, R3HDM1, RBM39, RECK, RUSC2, SCAND1, SIT1, SNAPC3, SPAG4, STOML2, UEVLD, UQCC, VWA5A	20

CLMY	Cellular Growth and Proliferation	ACMSD, ACTR3, BTC, C20orf4, COL3A1, CPNE1, DPP10, EIF6, EPB41L1, FAM184B, GALNT3, GDF5, HNMT, LAP3, LCORL, MGAT5, NCAPG, NFS1, NKX2-1, NKX2-8, PARM1, PAX9, PFKP, PHF20, PITRM1, PSIP1, RBM39, ROMO1, SCAND1, SCN1A, SCN9A, SNAPC3, SPAG4, UQCC	34
UBF	Skeletal and Muscular Disorders, Carbohydrate Metabolism	ANXA11, C6, C9orf89, CCDC71, DAG1, DDIT4, DOK2, EIF4EBP2, GFRA2, GNAI2, GPX1, HYAL1, HYAL2, HYAL3, IARS, IPPK, KIAA1274, LPL, MAP3K8, MAT1A, MST1, MST1R, NOL8, OMD, P4HTM, PLAU, PPIF, PRF1, PSAP, QARS, REEP4, RHOA, SELS, SEMA3F, SFTPA1, SFTPD, SLC38A3, SMC2, UBA7, UNC5B, USP19, VCL, WDR6, ZNF484	44
UREA	Gene Expression, Cellular Development, Skeletal and Muscular System Development and Function	A2ML1, ACRBP, ACTR6, BTBD10, CDH18, CEP78, CLEC4A, CLEC4E, CLEC6A, DCLRE1C, DPCD, ECM2, ELOVL3, FAM171A1, FAR1, FGF8, FOXJ2, HPS6, HSPA14, IFRD2, KCNIP4, MANF, MEIG1, MFAP5, MYO1B, NIF3L1, NUP155, OBFC2A, OGN, ORC2, OXCT1, PACRGL, PITX3, Rassf1, RBM5, SDPR, SEMA3B, SLC1A3, SUV39H2, TKT, TRAK2, WDR70, ZMYND10, ZNF484	44
	Amino Acid Metabolism, Molecular Transport, Small Molecule Biochemistry	ALS2CR4, ALS2CR12, APOBEC1, ASPN, BZW1, CACNA1D, CACNA2D2, CHDH, CPNE8, CYB561D2, DCP1A, FAM107B, FAM171A1, FBXW4, HYAL1, HYAL2, HYAL3, NDUFB3, NECAP1, NIPBL, NPRL2, POLL, PPRC1, RFT1, SCYL2, SFXN3, SLC17A8, SPON1, STRADB, TMEM115	30

[1]CWT = carcass weight (kg); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2).

Table 5 – 14. Posterior means of variance components explained by whole genome SNPs for beef carcass traits in the commercial hybrid beef cattle using the BayesB method.

Trait[1]	Genetic variance	Residual variance	Total variance	Proportion of phenotypic variance explained by all SNPs
CWT	349.83	337.65	687.48	0.51
CABF	3.49	10.31	13.81	0.25
CREA	21.59	36.06	57.65	0.37
CGF	3.77	9.94	13.71	0.27
CLMY	2.86	9.03	11.89	0.24
UMAR	0.63	0.35	0.98	0.65
UBF	0.01	0.02	0.03	0.36
UREA	12.01	26.39	38.41	0.31

[1]CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2).

Table 5 – 15. Summary of significantly ($P < 0.2$) associated QTL regions and functional candidate genes within the regions for beef ultrasound and carcass merit traits in hybrid beef cattle identified using the BayesB method.

Trait[1]	N-associated regions	BTA	Functionally relevant genes in associated regions (BTA)[2]	Previously identified overlapping QTL (BTA)[3]
CWT	3	6, 14	NCAPG(6), LCORL(6), TGS1(14)	Carcass weight (6, 14)
CABF	4	2, 6, 14	GALNT3(2), NCAPG(6), LCORL(6)	Backfat depth or thickness (2, 6, 14)
CREA	3	2, 7, 14	GALNT3(2)	--
CGF	2	6, 14	NCAPG(6), LCORL(6)	Backfat thickness(6, 14)
CLMY	3	2, 6, 14	GALNT3(2), NCAPG(6), LCORL(6)	USDA yield grade (14)
UMAR	7	2, 4, 8, 14, 25	GALNT3(2), CLCN3(8), EXT1(14)	Beef marbling score (4, 8, 14); ribeye muscle area (8); weaning weight (14); carcass weight (14); subcutaneous fat depth(14)
UBF	3	2, 14, 28	GALNT3(2), ADAMTS14(28), SGPL1(28), PCBD1(28), SLC29A3(28)	Backfat depth or thickness (2, 14, 28)
UREA	2	13, 14	--	Ribeye muscle area (13)

[1]CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2). [2]*ADAMTS14* = ADAM metallopeptidase with thrombospondin type 1 motif, 14; *CLCN3* = chloride channel 3; *EXT1* = exostosin 1; *GALNT3* = UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3; *LCORL* = ligand dependent nuclear receptor corepressor-like; *NCAPG* = non-SMC condensin I complex, subunit G; *PCBD1* = pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha; *SGPL1* = sphingosine-1-phosphate lyase 1; *SLC29A3* = solute carrier family 29 (nucleoside transporters), member 3; *TGS1* = trimethylguanosine synthase 1. [3]The QTL information was obtained from CattleQTLdb (http://www.animalgenome.org/cgi-bin/gbrowse/cattle/).

Table 5 – 16. Detailed information for QTL regions associated ($P < 0.2$) with beef carcass traits in commercial hybrid beef cattle using the BayesB method.

Trait[1]	BTA	Start[2]	End[2]	N-SNPs[3]	%Var[4]	*P*-value	Genes at SNP window[5]	Previously reported QTL at the SNP window (Reference)[6]	Previously related traits
CWT	6	38,014,254	38,982,338	24	12.47	0.004	DCAF16, NCAPG,* LCORL*	38 cM (TAKASUGA *et al.* 2007)	Carcass weight
								39.1 cM (NCAPG) (SETOGUCHI *et al.* 2009)	Carcass weight
	14	22,019,957	22,967,675	17	5.86	0.006	SOX17, RP1, TMEM68, TGS1*	39 (17.84-43.63) cM (MCCLURE *et al.* 2010)	Carcass weight
	14	5,010,064	5,971,945	33	0.07	0.18	KHDRBS3	--	--
CABF	6	38,014,254	38,982,338	24	1.1	0.13	DCAF16, NCAPG,* LCORL*	36 (17.00-43.93) cM (MCCLURE *et al.* 2010)	Fat thickness
	14	5,010,064	5,971,945	33	0.08	0.19	KHDRBS3	5.125 cM (MOORE *et al.* 2003)	Backfat
	2	31,016,869	31,996,368	30	0.09	0.20	SCN9A, SCN1A, TTC21B, GALNT3,* CSRNP3, SCN2A	54 (21-60) cM (CASAS *et al.* 2003)	Fat depth
	2	139,004,086	139,987,690	30	0.09	0.20	ARHGEF10L, RCC2, PADI4	--	--
CREA	7	91,027,447	91,903,228	11	5.76	0.056	CETN3, MBLAC2, POLR3G, LYSMD3	81 (65.30-90.70) cM (MCCLURE *et al.* 2010)	Ribeye muscle area
	2	31,016,869	31,996,368	30	0.19	0.17	SCN9A, SCN1A, TTC21B, GALNT3,* CSRNP3, SCN2A	--	--
	14	5,010,064	5,971,945	33	0.09	0.18	KHDRBS3	--	--
CGF	6	38,014,254	38,982,338	24	0.83	0.14	DCAF16, NCAPG,* LCORL*	36 (17.00-43.93) cM (MCCLURE *et al.* 2010)	Fat thickness
	14	5,010,064	5,971,945	33	0.08	0.19	KHDRBS3	5.125cM (MOORE *et al.* 2003)	Backfat
CLMY	14	5,010,064	5,971,945	33	0.07	0.18	KHDRBS3	19 (0-24) cM (CASAS *et al.* 2003)	USDA yield grade
	2	31,016,869	31,996,368	30	0.27	0.18	SCN9A, SCN1A, TTC21B, GALNT3,* CSRNP3, SCN2A	--	--
	6	38,014,254	38,982,338	24	0.33	0.20	DCAF16, NCAPG,* LCORL*	--	--
UMAR	8	1,016,022	1,957,751	14	3.33	0.089	SH3RF1, CLCN3,* C8H4orf27, LOC615521, MFAP3L, AADAT	4 (TAKASUGA *et al.* 2007)	Beef marbling score
								7 (1-11.34) cM (MCCLURE *et al.* 2010)	Marbling score and ribeye muscle area
	14	43,053,810	43,984,234	21	2.18	0.12	SAMD12, EXT1*	44.2 cM (FABP4) (LEE *et al.* 2010a)	Marbling and carcass weight
								46.3 cM (CRH) (WIBOWO *et al.* 2007)	Marbling and subcutaneous fat depth
								47 (30-87) cM (CASAS *et al.* 2003)	Marbling
	4	59,049,049	59,971,948	21	2.23	0.13	IMMP2L	55 (52.49-67.47) cM (MIZOSHITA *et al.* 2004)	Beef marbling score
								56 cM (TAKASUGA *et al.* 2007)	Beef marbling score
	2	31,016,869	31,996,368	30	0.24	0.16	SCN9A, SCN1A, TTC21B, GALNT3,* CSRNP3, SCN2A	--	--

	14	5,010,064	5,971,945	33	0.06	0.19	KHDRBS3	5 (0.00-5.13) cM (McClure *et al.* 2010)	Marbling score and weaning weight
	25	12,019,320	12,986,220	27	0.15	0.19	SNX29, CPPED1, SHISA9	--	--
	2	139,004,086	139,987,690	30	0.1	0.19	ARHGEF10L, RCC2, PADI4	--	--
UBF	28	26,021,026	26,993,363	21	1.5	0.15	PRF1, ADAMTS14,* C28H10orf27, SGPL1,* PCBD1,* UNC5B, SLC29A3,* CDH23	23 (16.06-24.77) cM (McClure *et al.* 2010)	Fat thickness
	14	5,010,064	5,971,945	33	0.07	0.19	KHDRBS3	5.125cM (Moore *et al.* 2003)	Backfat
	2	31,016,869	31,996,368	30	0.08	0.20	SCN9A, SCN1A, TTC21B, GALNT3,* CSRNP3, SCN2A	54 (21-60) cM (Casas *et al.* 2003)	Fat depth
UREA	14	5,010,064	5,971,945	33	0.12	0.17	KHDRBS3	--	--
	13	80,053,094	80,955,946	27	0.44	0.19	SALL4, ZFP64	82 (73.63-91.37) cM (McClure *et al.* 2010)	Ribeye muscle area

[1]CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2). [2]Start and end position of the associated SNP windows (bp) based on the Btau4.0 reference assembly (http://www.hgsc.bcm.tmc.edu/ftp-archive/Btaurus/fasta/Btau20070913-freeze/). [3]The number of SNPs located within each associated SNP window. [4]The proportion of phenotypic variance explained by each SNP window. [5]Genes located within the SNP window and the genes with potential functions related to beef carcass traits were indicated with an asterisk (*).[6]Previously identified QTL at the associated QTL region with QTL peak location, QTL span and reference. Previously reported candidate genes within these regions are also presented. The QTL information was obtained from CattleQTLdb (www.animalgenome.org/cgi-bin/QTLdb/BT/index). Associated regions identified for CABF, CGF, and UBF were compared with the QTL reported for fat thickness at the 12th rib; QTL regions identified for CWT, CREA, CLMY, UMAR and UREA were compared with the QTL reported for carcass weight, ribeye area, yield grade, marbling score and ribeye area, respectively. *CRH*: corticotropin releasing hormone; *FABP4*: fatty acid binding protein 4, adipocyte; *NCAPG*: non-SMC condensin I complex, subunit G.

Table 5 – 17. Comparison of the associated SNPs and QTL regions between single marker LD regression and Bayesian regression for beef carcass traits in commercial hybrid beef cattle.

Trait[1]	BTA	BayesB regression		Single marker LD regression		Previously reported QTL at the SNP window (Reference)[4]	Previously related traits
		SNP window (Mb)[2]	N-SNPs[2]	SNP position (Mb)[3]	N-SNPs[3]		
CWT	6	38.01-38.98	24	36.09-46.52	39	38 cM (TAKASUGA *et al.* 2007)	Carcass weight
						39.1 cM (NCAPG) (SETOGUCHI *et al.* 2009)	Carcass weight
						42 (37-55) cM (SETOGUCHI *et al.* 2009)	Carcass weight
	14	22.02-22.97	17	22.63-22.84	5	39 (17.84-43.63) cM (MCCLURE *et al.* 2010)	Carcass weight
	14	5.01-5.97	33	--	--	--	--
CABF	6	38.01-38.98	24	38.10-38.73	8	36 (17.00-43.93) cM (MCCLURE *et al.* 2010)	Fat thickness
	14	5.01-5.97	33	--	--	5.125 cM (MOORE *et al.* 2003)	Backfat
	2	31.02-32.00	30	--	--	54 (21-60) cM (CASAS *et al.* 2003)	Fat depth
	2	139.00-139.99	30	--	--	--	--
CREA	7	91.03-91.90	11	91.90	1	81 (65.30-90.70) cM (MCCLURE *et al.* 2010)	Ribeye muscle area
	2	31.02-32.00	30	32.00	1	--	--
	14	5.01-5.97	33	--	--	--	--
CGF	6	38.01-38.98	24	--	--	36 (17.00-43.93) cM (MCCLURE *et al.* 2010)	Fat thickness
	14	5.01-5.97	33	--	--	5.125cM (MOORE *et al.* 2003)	Backfat
CLMY	14	5.01-5.97	33	--	--	19 (0-24) cM (CASAS *et al.* 2003)	USDA yield grade
	2	31.02-32.00	30	31.26	1	--	--
	6	38.01-38.98	24	38.42	1	--	--
UMAR	8	1.02-1.96	14	--	--	4 (TAKASUGA *et al.* 2007)	Beef marbling score
						7 (1-11.34) cM (MCCLURE *et al.* 2010)	Marbling score and ribeye muscle area
	14	43.05-43.98	21	43.74	1	44.2 cM (FABP4) (LEE *et al.* 2010a)	Marbling and carcass weight
						46.3 cM (CRH) (WIBOWO *et al.* 2007)	Marbling and subcutaneous fat depth
						47 (30-87) cM (CASAS *et al.* 2003)	Marbling
	4	59.05-59.97	21	59.97	1	55 (52.49-67.47) cM (MIZOSHITA *et al.* 2004)	Beef marbling score
						56 cM (TAKASUGA *et al.* 2007)	Beef marbling score
	2	31.02-32.00	30	31.61, 31.65	2	--	--
	14	5.01-5.97	33	--	--	5 (0.00-5.13) cM (MCCLURE *et al.* 2010)	Marbling score and weaning weight
	25	12,02-12.99	27	--	--	--	--
	2	139.00-139.99	30	--	--	--	--
UBF	28	26.02-26.99	21	26.19-34.65	5	23 (16.06-24.77) cM (MCCLURE *et al.* 2010)	Fat thickness
	14	5.01-5.97	33	--	--	5.125cM (MOORE *et al.* 2003)	Backfat
	2	31.02-32.00	30	--	--	54 (21-60) cM (CASAS *et al.* 2003)	Fat depth
UREA	14	5.01-5.97	33	--	--	--	--

13	80.05-80.96	27	--	--	82 (73.63-91.37) cM (MCCLURE *et al.* 2010)	Ribeye muscle area

[1]CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2). [2]Physical position of the associated SNP windows (Mb) using BayesB regression and the number of SNPs located within the SNP window. [3]Physical position of the significantly associated SNP (Mb) using the single marker LD regression and the number of SNPs within the region. [4]Previously identified QTL at the associated QTL region with QTL peak, QTL span and reference. Previously reported candidate genes within these regions are also presented. The QTL information was obtained from CattleQTLdb (www.animalgenome.org/cgi-bin/QTLdb/BT/index). *CRH*: corticotropin releasing hormone; *FABP4*: fatty acid binding protein 4, adipocyte; *NCAPG*: non-SMC condensin I complex, subunit G.

Figures

Figure 5 – 1A

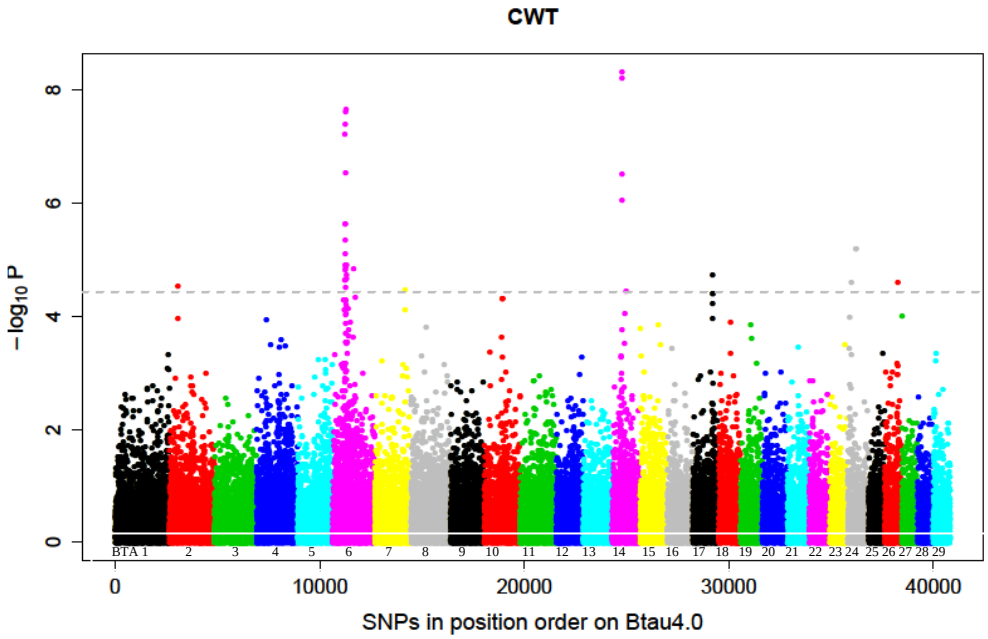

Figure 5 – 1B

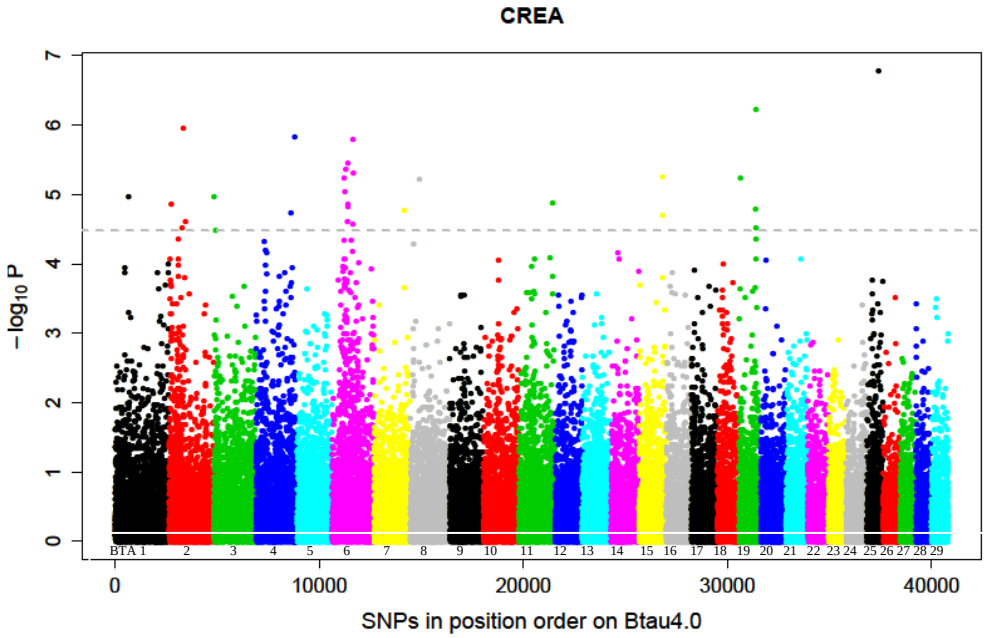

Figure 5 – 1. Whole-genome association analyses for beef carcass merit traits in commercial hybrid beef cattle using single marker LD regression: A) carcass weight (CWT); B) carcass ribeye area (CREA). The X-axis is the genomic location of the SNP on Btau4.0 (Mb). The Y-axis represents the negative logarithm (base 10) of the *P*-values for the SNP allele substitution effects. Different shades represented SNPs on different chromosomes from BTA1 (left) to BTA29 (right). The dashed line in grey represents the genome-wise threshold at FDR $P < 0.05$.

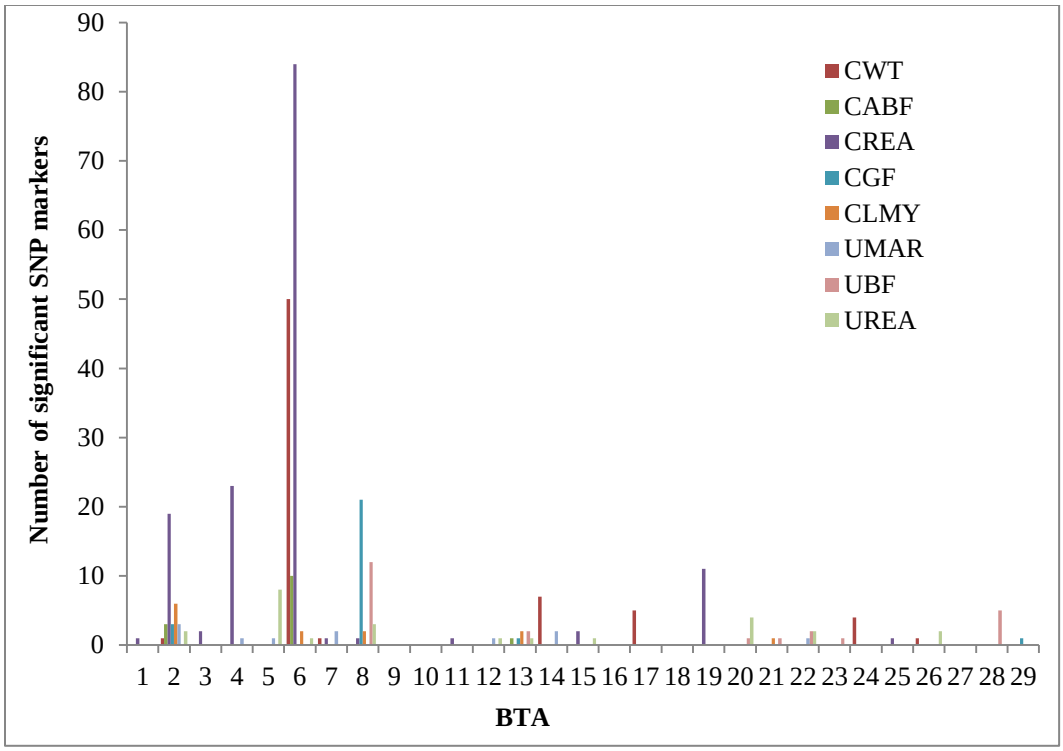

Figure 5 – 2. Chromosomal distribution of significant SNPs at chromosome-wise 5% FDR threshold for beef carcass traits in commercial hybrid beef cattle using single marker LD regression. CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2).

Figure 5 – 3A

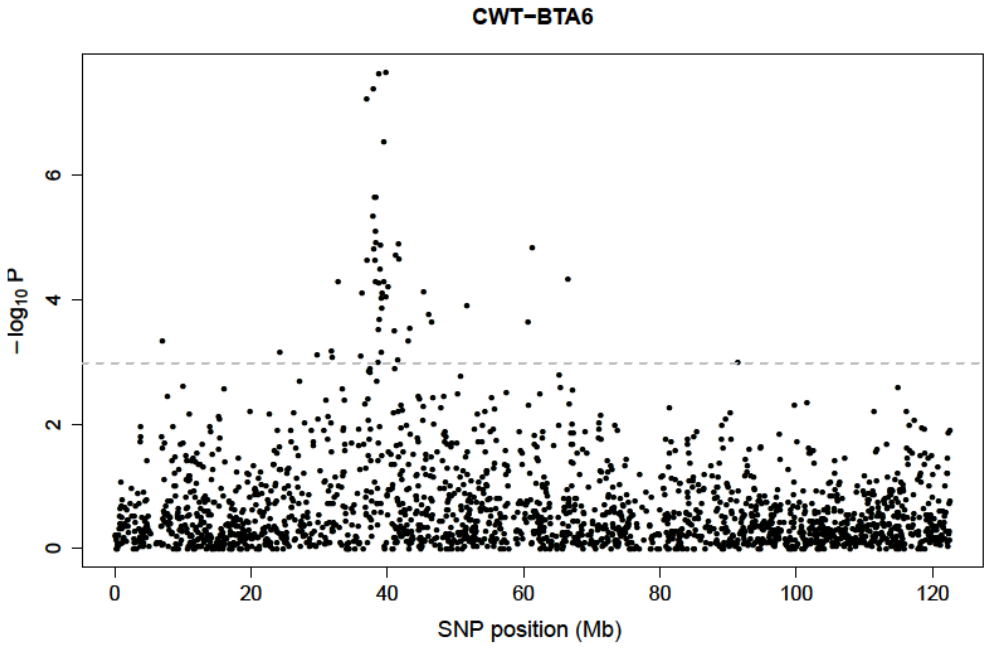

Figure 5 – 3B

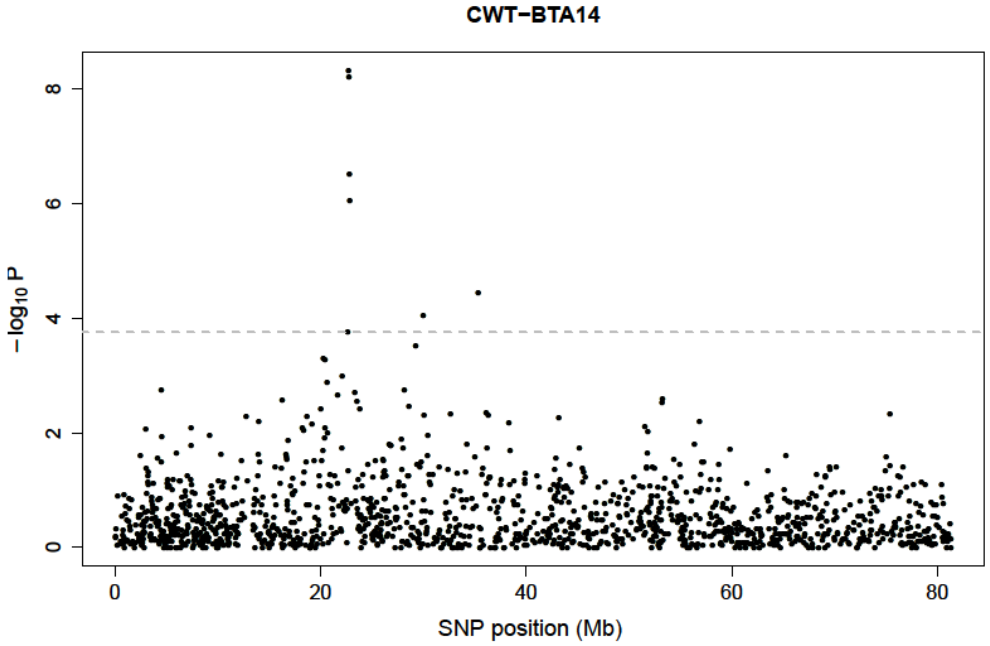

Figure 5 – 3C

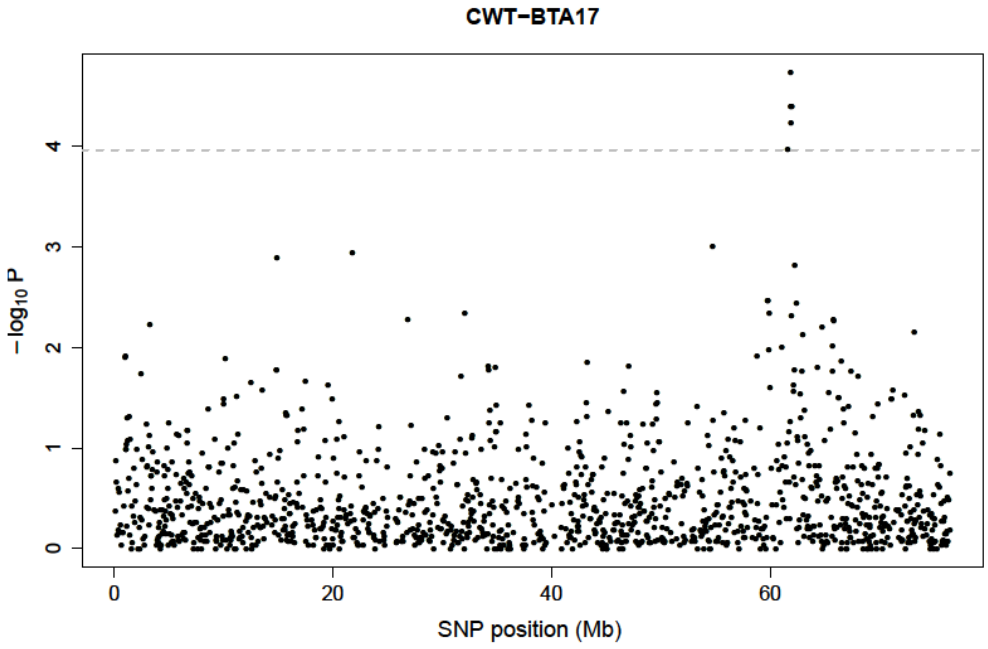

Figure 5 – 3D

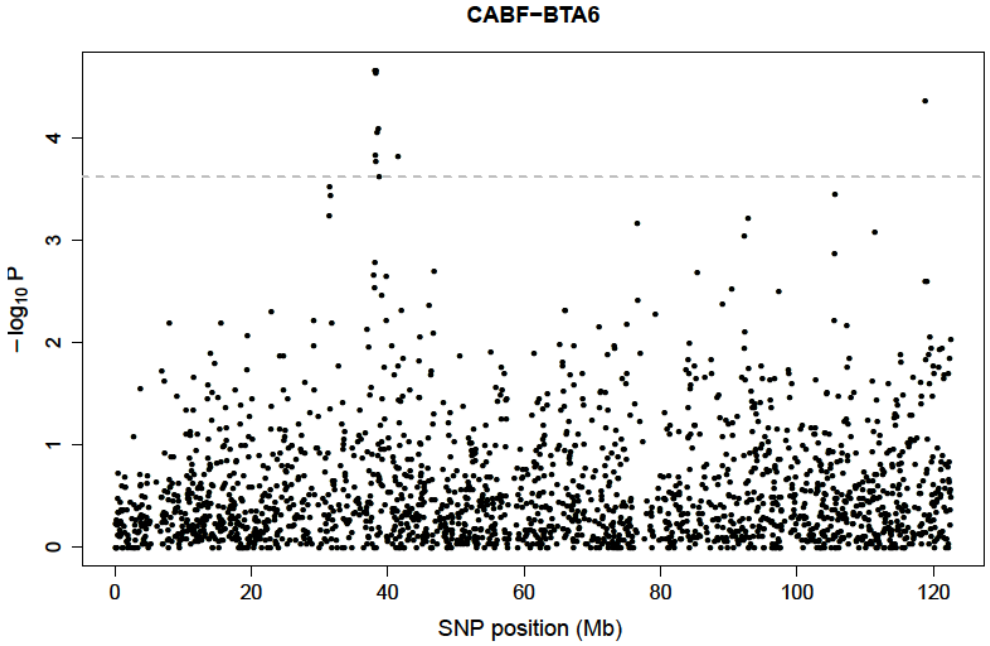

Figure 5 – 3E

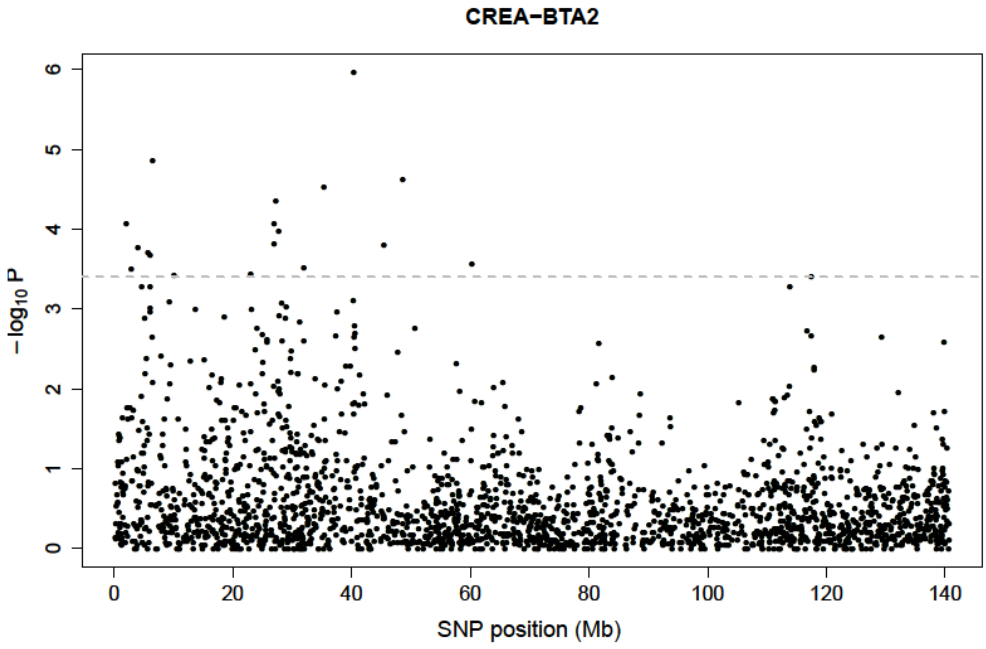

Figure 5 – 3F

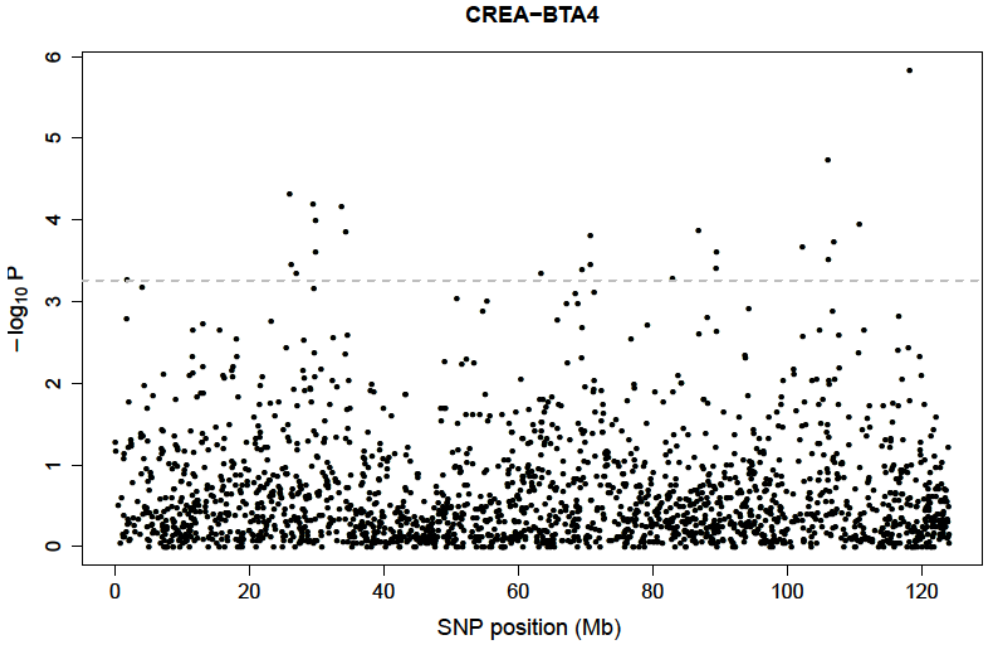

Figure 5 – 3G

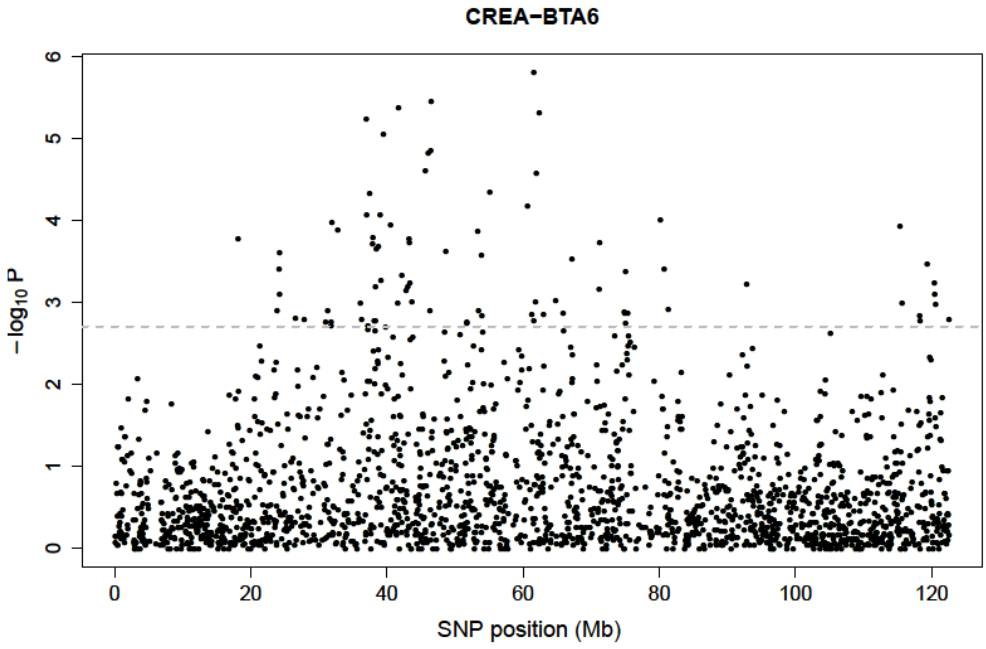

Figure 5 – 3H

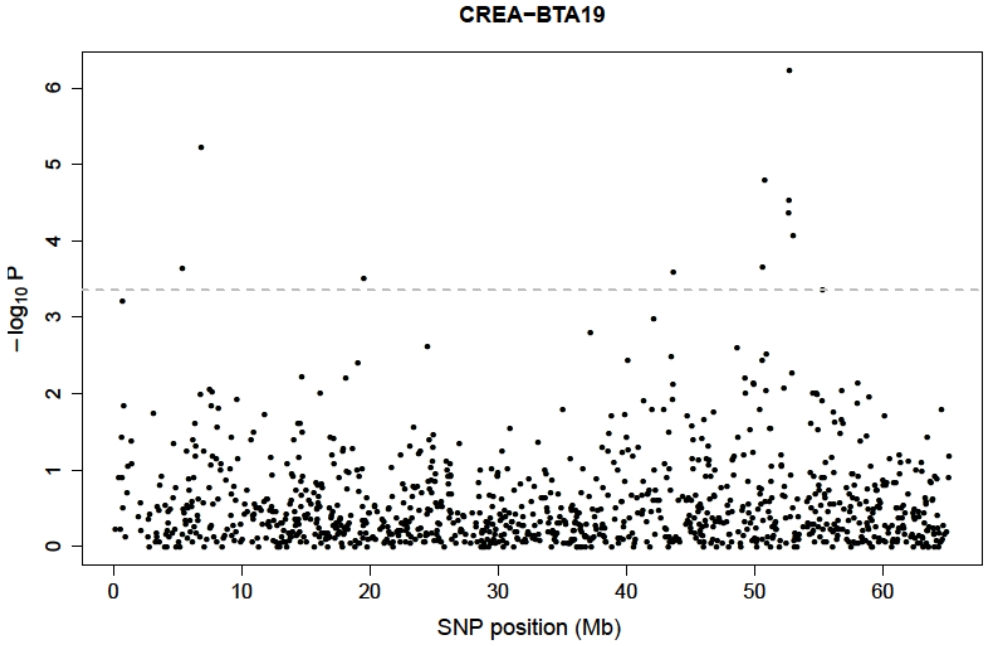

Figure 5 – 3I

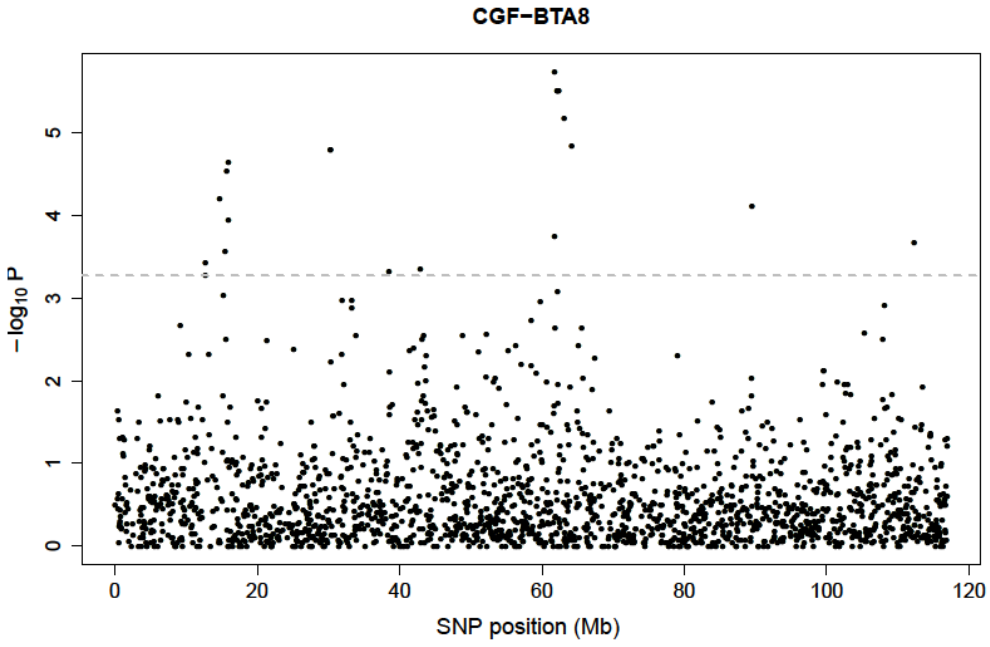

Figure 5 – 3J

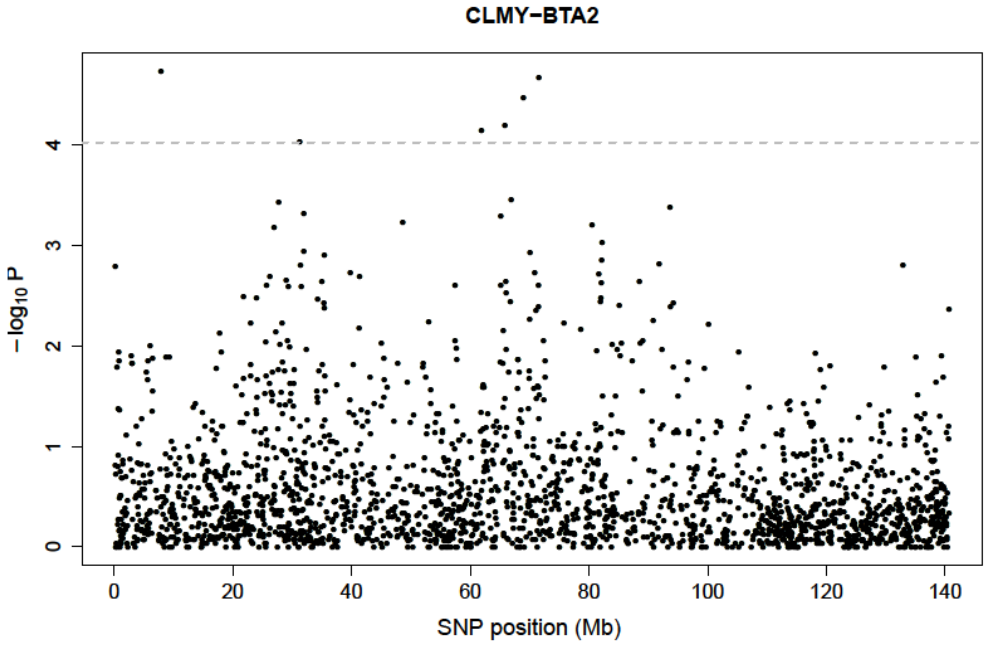

Figure 5 – 3K

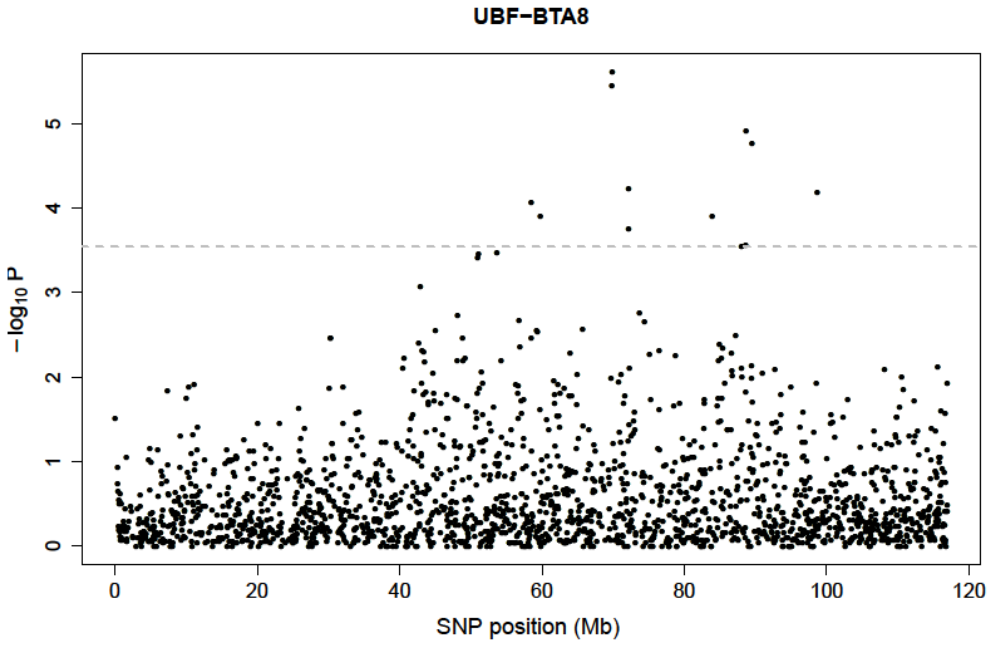

Figure 5 – 3L

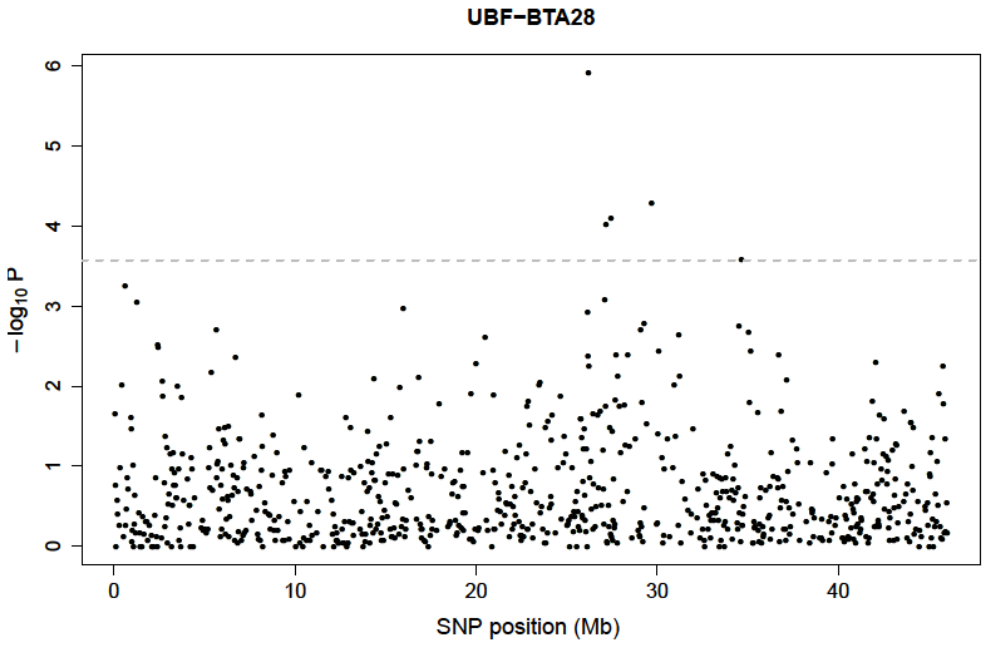

Figure 5 – 3M

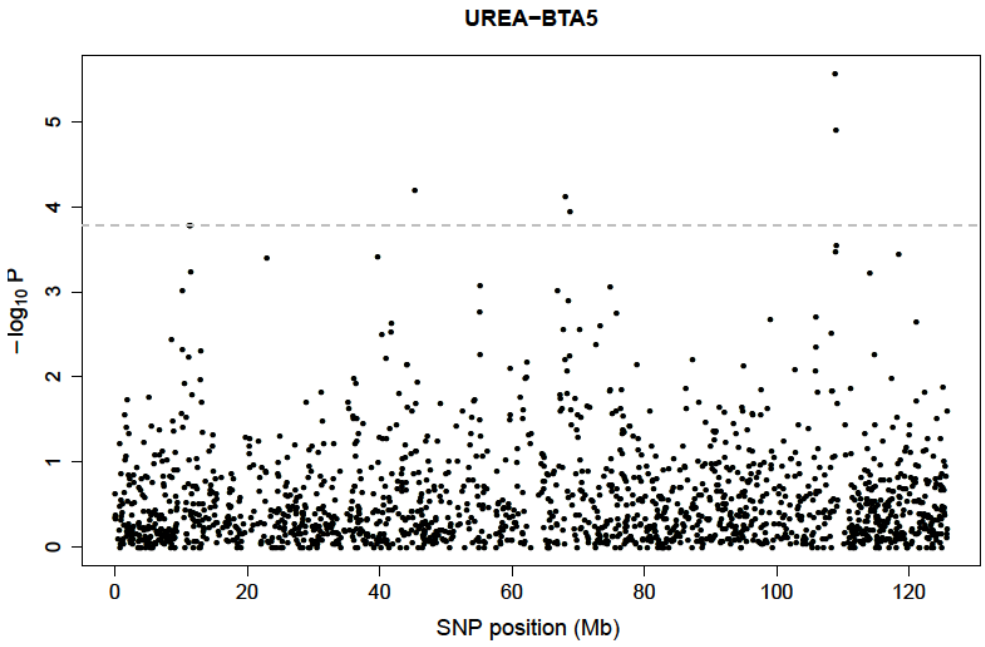

Figure 5 – 3. Examples of chromosomes with clustered significant SNPs identified for beef carcass traits in the commercial hybrid beef cattle using single marker LD regression: A) BTA6 for carcass weight (CWT); B) BTA14 for CWT; C) BTA17 for CWT; D) BTA6 for carcass average backfat (CABF); E) BTA2 for carcass ribeye area (CREA); F) BTA4 for CREA; G) BTA6 for CREA; H) BTA19 for CREA;. I) BTA8 for carcass grade fat (CGF); J) BTA2 for carcass lean meat yield (CLMY); K) BTA8 for ultrasound backfat (UBF); L) BTA28 for UBF; M) BTA5 for ultrasound ribeye area (UREA). The X-axis is the physical location of the SNP on each chromosome (Mb). The Y-axis represents the negative logarithm (base 10) of the *P*-values for the SNP allele substitution effects. The dashed line in grey represents the chromosome-wise FDR adjusted thresholds at $P < 0.05$.

Figure 5 – 4A

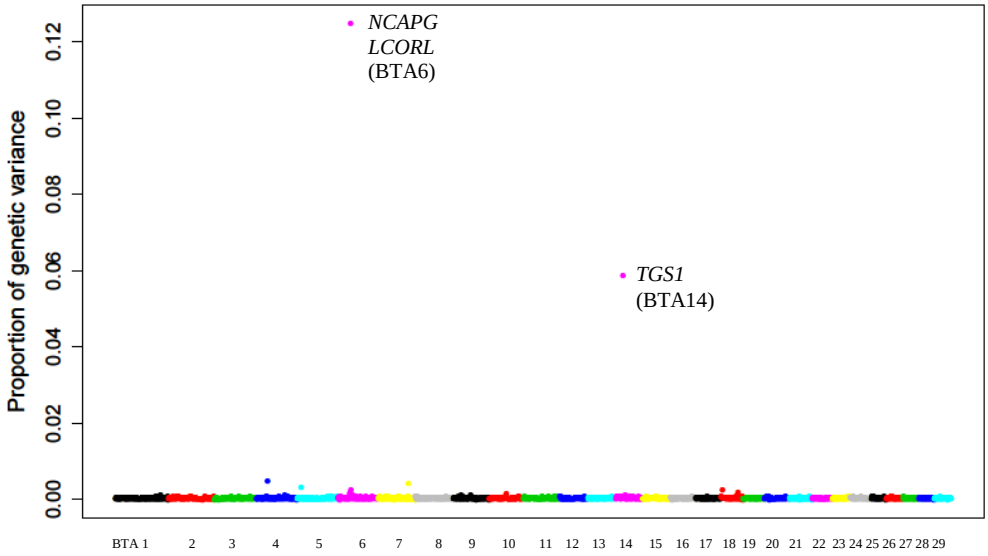

Figure 5 – 4B

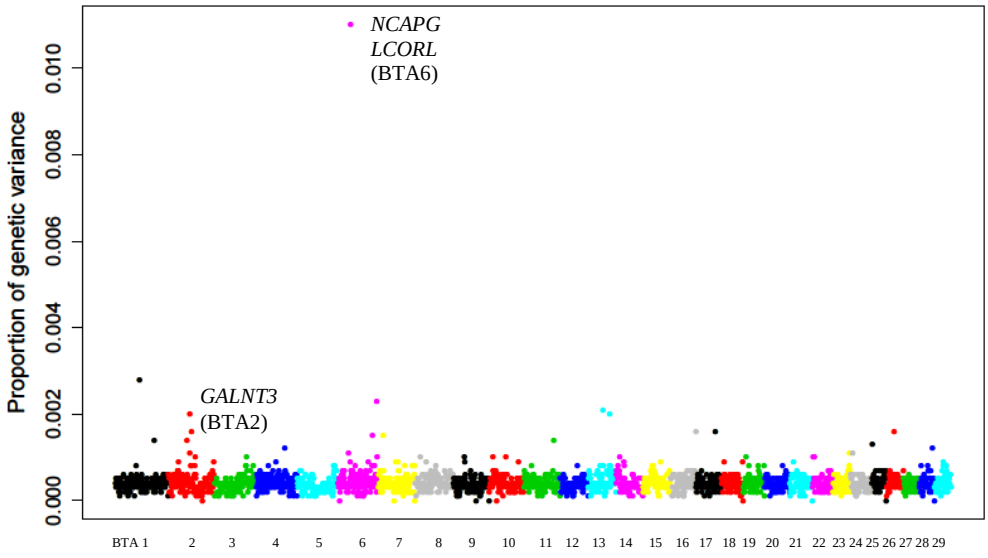

Figure 5 – 4C

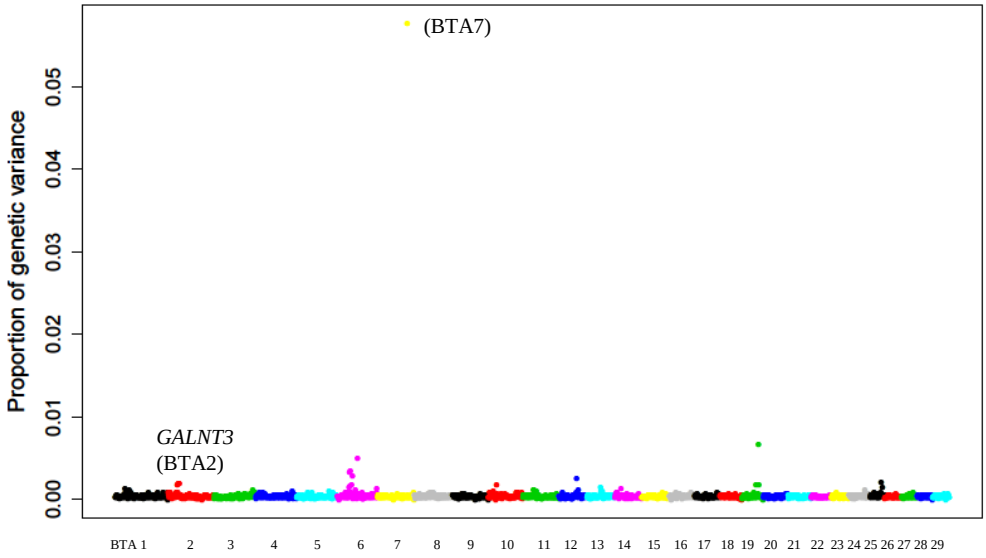

Figure 5 – 4D

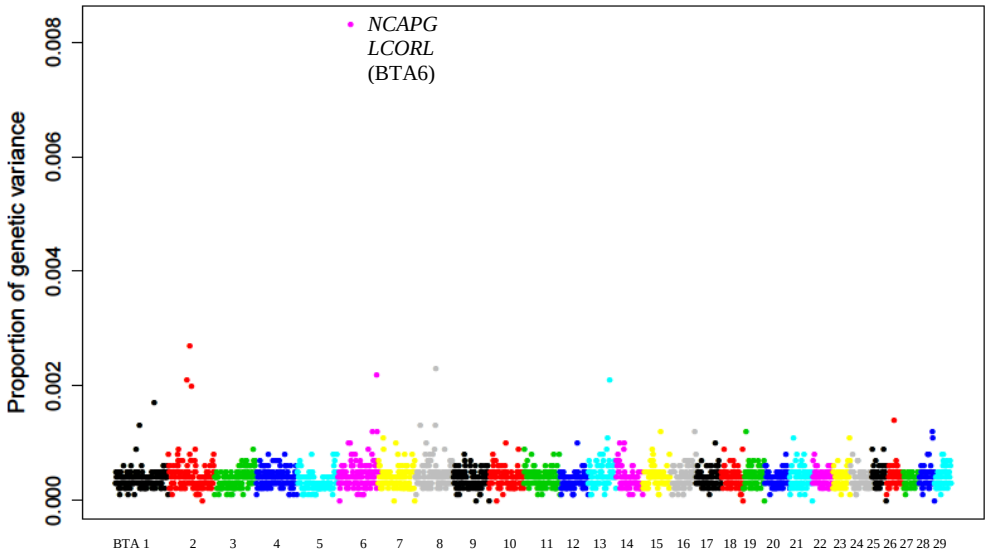

Figure 5 – 4E

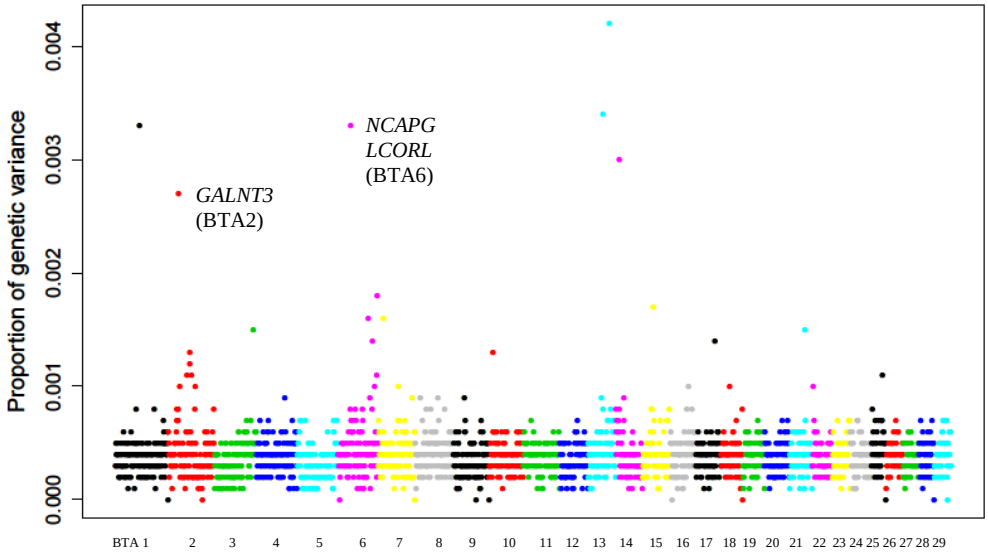

Figure 5 – 4F

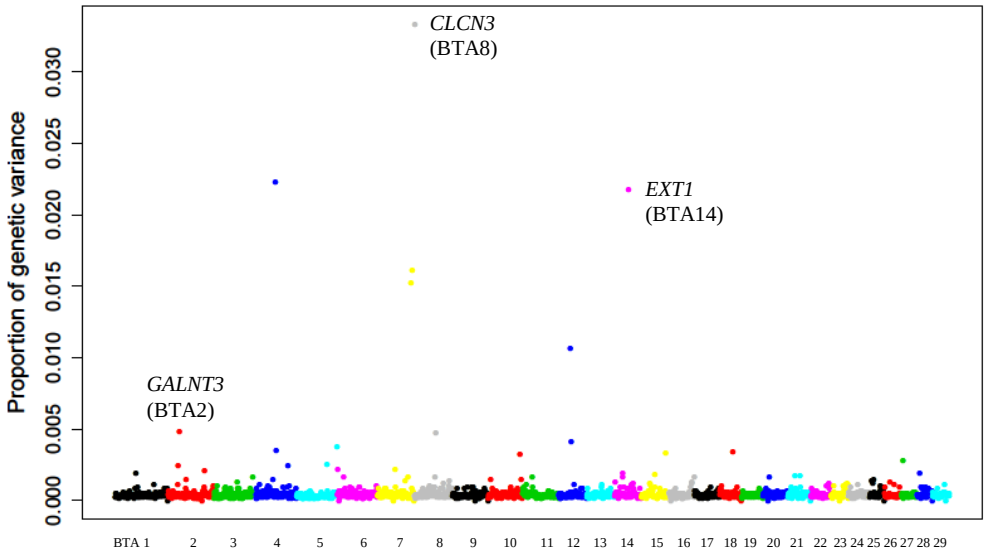

Figure 5 – 4G

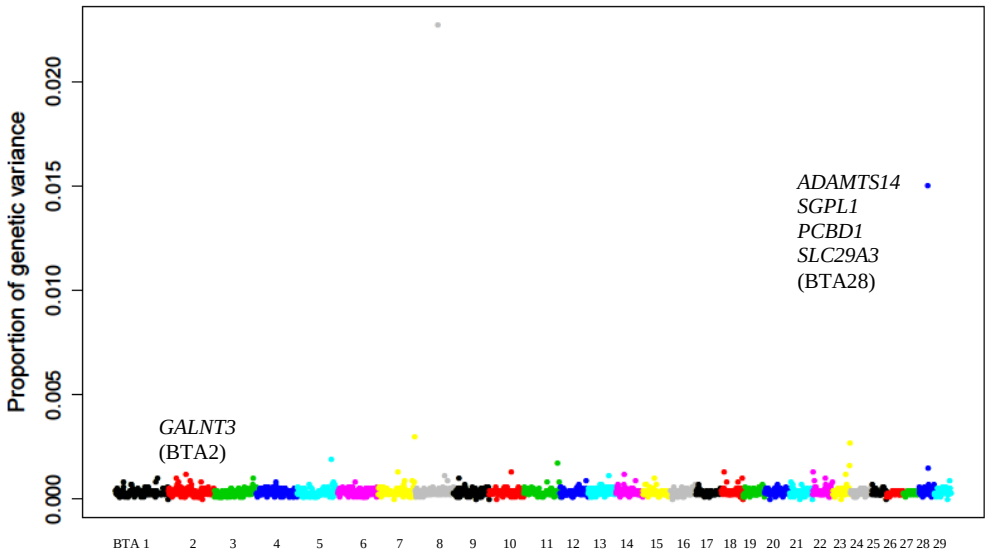

Figure 5 – 4H

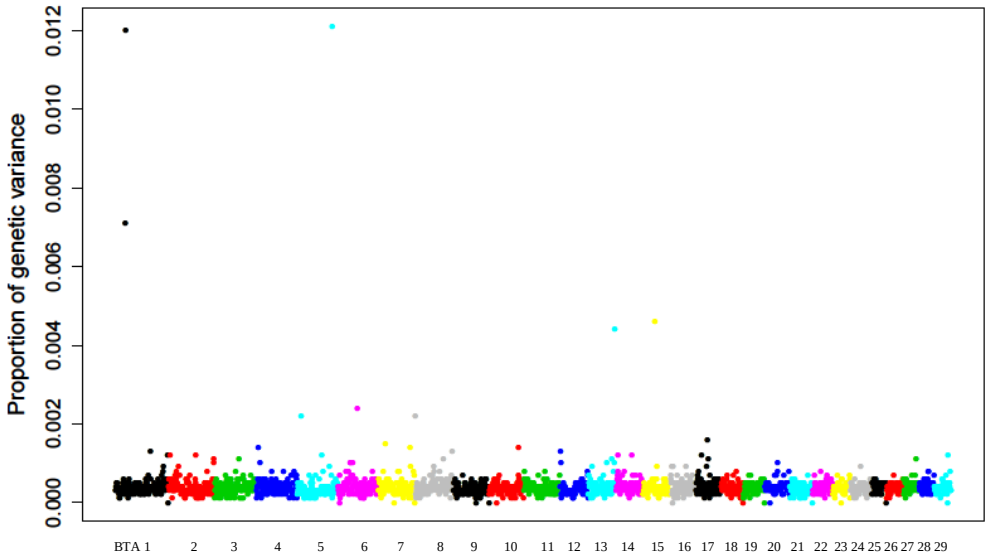

Figure 5 – 4. Proportion of genetic variance explained by each window of 1Mb consecutive SNP markers across the genome for beef ultrasound and carcass merit traits in commercial hybrid beef cattle: A) carcass weight (CWT); B) carcass average backfat (CABF); C) carcass ribeye area (CREA); D) carcass grade fat (CGF); E) carcass lean meat yield (CLMY); F) ultrasound marbling (UMAR); G) ultrasound backfat (UBF); H) ultrasound ribeye area (UREA). The X-axis is SNP marker position in genome order according to Btau4.0 (Mb), and the Y-axis represents the proportion of genetic variance contributed by the 1 Mb SNP window (the exact candidate regions). Different colours represent SNPs on different chromosomes from BTA1 (left) to BTA29 (right). *ADAMTS14* = ADAM metallopeptidase with thrombospondin type 1 motif, 14; *CLCN3* = chloride channel 3; *EXT1* = exostosin 1; *GALNT3* = UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3; *LCORL* = ligand dependent nuclear receptor corepressor-like; *NCAPG* = non-SMC condensin I complex, subunit G; *PCBD1* = pterin-4 alpha-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha; *SGPL1* = sphingosine-1-phosphate lyase 1; *SLC29A3* = solute carrier family 29 (nucleoside transporters), member 3; *TGS1* = trimethylguanosine synthase 1.

Chapter 6. Genome-wide Analysis of Epistasis for Quantitative Traits in Canadian Holstein Cattle Using an Empirical Bayes Method

6.1. Introduction

Genetic variation of quantitative traits is often controlled by the segregation of multiple interacting loci (CARLBORG and HALEY 2004; MOORE 2005). Previous studies found that interactions among QTL (epistasis) play an important role in the expression of phenotypes (ANKRA-BADU *et al.* 2010; CARLBORG *et al.* 2005; CARLBORG *et al.* 2003; GROSSE-BRINKHAUS *et al.* 2010; UEMOTO *et al.* 2009). The role of epistasis in the genetic architecture of complex traits can be explored by epistatic QTL mapping (CARLBORG and HALEY 2004) and the statistical framework to incorporate epistasis was initiated half a century ago (COCKERHAM 1954; KEMPTHORNE 1954). Recently, several methods have been developed for the estimation of epistasis controlling quantitative variation (BOER *et al.* 2002; CARLBORG *et al.* 2000; JANNINK and JANSEN 2001; WANG *et al.* 1999; YI *et al.* 2007; YI *et al.* 2005; YI *et al.* 2003; ZENG *et al.* 1999). However, epistatic effect analysis through single epistatic effect or model selection for multiple epistatic effects may result in the loss of some important interaction effects (XU and JIA 2007; YI and XU 2002; ZHANG and XU 2005). A recently developed empirical Bayesian method allows simultaneous estimation of marginal effects (additive and dominance) of all individual markers and epistatic effects of all pairs of markers in a single model (XU 2007). This new approach is different from other full-model shrinkage methods either by using a Bayesian

framework (ZHANG and XU 2005) or with a higher computational efficiency (YI and XU 2002). Epistasis studies (HU *et al.* 2011; XU and JIA 2007) using this method are true multiple-effect analyses that require no variable selection, thus no important epistatic QTL can be missed if the genome coverage of markers is sufficiently high.

Dairy cattle are highly specialized agricultural species for milk production and have experienced intensive artificial selection in the last 50 years (BROTHERSTONE and GODDARD 2005). Most of the economically important traits under selection in dairy cattle are polygenetic traits controlled by multiple genes. Many studies on mapping QTL for these traits have been carried out in the past two decades and most of them ignored the QTL interactions (GEORGES *et al.* 1995; KHATKAR *et al.* 2004; ZHANG *et al.* 1998). So far, only a small proportion of the total genetic variance has been explained by the reported additive and dominance effects of QTL (HAYES and GODDARD 2010). Little knowledge of the true genetic architecture (*i.e.*, the number of genes and alleles, as well as the nature of interactions among them) that underlies dairy traits of interest is available. With the availability of the bovine genome sequence assembly and a large number of SNP markers (ELSIK *et al.* 2009; MATUKUMALLI *et al.* 2009), genome-wide epistatic QTL effects can be studied in cattle to test their genetic contribution. The purpose of this study was to dissect the additive-by-additive epistatic effects, in addition to the additive effects, in the genetic architectures of economically important traits in dairy cattle. A genome-wide epistatic QTL

analysis was carried out in the Canadian Holstein cattle using the empirical Bayes method (XU 2007).

6.2. Materials and Methods

6.2.1. Animal Resource and Phenotypes

The Canadian Holstein cattle used for this study are described in Chapter 3. Four economically important traits in dairy cattle were considered in this study: two milk production related traits, milk yield (MY) and somatic cell score (SCS); and two functional traits, daughter calving ability (DCA) and milking temperament (MT). The descriptive statistics of these traits are given in Table 6 – 1.

6.2.2. Genotyping Platform and Marker Selection

Following the marker filtering described in Chapter 4, a set of 316 evenly spaced SNPs covering the autosomal bovine genome with an average intermarker distance of 8.80 Mb was selected using the differential evolutionary algorithm (KINGHORN 1998). A small number of markers was selected for this study in order to fit all pairwise epistatic effects together with all marginal effects in the oversaturated epistatic QTL model. The descriptive statistics for the selected markers are given in Table 6 – 2. The three genotypes at each locus were coded as +1, 0 and –1, respectively.

6.2.3. QTL Mapping

The statistical analyses were conducted by using the empirical Bayes method (XU 2007). In summary, this method estimates prior variance components using marginal maximum-likelihood and then estimates QTL effects using the Bayesian shrinkage method given the estimated prior variance components as if they were the true prior variances. This method, without the MCMC sampling for inference of the parameter distribution, is computationally efficient. The statistical models used in this study are briefly introduced here: (1) the additive effect QTL model which includes only the additive QTL effects; and (2) the epistatic effect QTL model, constructed by adding the additive-by-additive (A × A) QTL effects into the additive effect QTL model. The additive QTL effects of all individual markers and the A × A epistatic effects of all pairs of markers were estimated simultaneously in the epistatic effect QTL model. Let n be the number of bulls and m be the number of markers, the phenotypic value for a trait is described by the following linear models,

$$y = 1\mu + \sum_{l=1}^{m} Z_l \gamma_l + \varepsilon \tag{1}$$

$$y = 1\mu + \sum_{l=1}^{m} Z_l \gamma_l + \sum_{l'>l}^{m} \left(Z_l \otimes Z_{l'} \right) \gamma_{ll'} + \varepsilon \tag{2}$$

where y is a $n \times 1$ vector, μ is the population mean, $Z_l = \left(Z_{1l} ... Z_{nl} \right)^T$ is a $n \times 1$ vector of the genotype indicators for locus $l \left(\forall l = 1, ..., m \right)$, Z_{il} takes one of three

values $\{1, 0, -1\}$ depending on the genotype of bull i for locus l, γ_l is the additive effect for locus l and $\gamma_{ll'}$ is the epistatic effect between loci l and l', and ε is the residual error vector with an assumed normal distribution $N\left(0, I\sigma^2\right)$. The notation $Z_l \otimes Z_{l'}$ represents a direct product of vectors Z_l and $Z_{l'}$. Excluding μ, the total number of QTL effects was m = 316 additive effects for the additive effect QTL model (1), while the total number of QTL effects for the epistatic effect QTL model (2) was $p = m \times (m+1)/2$ = 50,086, including m = 316 additive effects and $m \times (m-1)/2$ = 49,770 A × A epistatic effects.

The data were analyzed using a SAS IML program downloaded from http://www.statgen.ucr.edu. The hyper-parameters were chosen as $(\tau, \omega) = (-1, \omega)$, where τ was set as –1 as suggested in previous studies (TER BRAAK *et al.* 2005; XU 2003; XU 2007). As the value of ω in general should be small, we varied ω from $5e^{-14}$ to 0 incremented by 10-fold each time and chose the value that minimized the prediction error for each trait (TIBSHIRANI 1996).

Based on the fact that the sample size was not sufficiently large to shrink all small effects to zero and many spurious QTL effects still occurred, genome-wise critical values for significant QTL effects at the 5% level $(\alpha = 0.05)$ were determined by permutation tests (1,000 reshuffled samples for the additive effect QTL model and 50 reshuffled samples for the epistatic effect QTL model) (XU and JIA 2007).

6.2.4. Estimation of Multi-locus Genetic Variance

The total additive variance $V_A = \sum_{l=1}^{m} \gamma_l^2$, the total epistatic variance $V_{AA} = \sum_{l<l'}^{m} \gamma_{ll'}^2$ (neglecting the covariance caused by linkage), the overall genetic variance $V_G = V_A + V_{AA}$, the corresponding proportions of the phenotypic variance contributed by the additive, the epistatic, and the total genetic variances $H_A = V_A / V_P$, $H_{AA} = V_{AA} / V_P$, and $H_G = V_G / V_P$, where V_P is the observed phenotypic variance, were calculated, respectively, for each trait by using QTL effects only declared as significant.

Allele frequency was considered in calculating the QTL genetic contributions since the outbreed cattle population here was not a well-designed population for QTL mapping. The epistatic effect QTL model (2) was rewritten for an individual animal j as:

$$y_j = b_0 + \sum_{l=1}^{m} x_{jl} b_l + \sum_{l'>l}^{m} \left(x_{jl} \otimes x_{jl'} \right) b_{ll'} + \varepsilon_j .$$

The phenotypic variance was then calculated using:

$$\mathrm{var}(y_j) = \sum_{l=1}^{m} \mathrm{var}(x_{jl}) b_l^2 + \sum_{ll'=1}^{\frac{1}{2}l(l-1)} \mathrm{var}(x_{jll'}) b_{ll'}^2 + \mathrm{var}(\varepsilon_j) .$$

Let $\sigma_p^2 = \mathrm{var}(y_j)$, $\sigma_0^2 = \mathrm{var}(\varepsilon_j)$, $v_l^2 = \mathrm{var}(x_{jl})$ and $v_{ll'}^2 = \mathrm{var}(x_{jll'})$, then the phenotypic variance was rewritten as:

$$\sigma_p^2 = \sum_{l=1}^{m} v_l^2 b_l^2 + \sum_{ll'=1}^{\frac{1}{2}l(l-1)} v_{ll'}^2 b_{ll'}^2 + \sigma_0^2 .$$

The variances for the additive coefficients v_l^2 and the epistatic coefficient $v_{ll'}^2$ were calculated as:

$$v_l^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{jl} - \bar{x}_{jl})^2$$
$$v_{ll'}^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{jll'} - \bar{x}_{jll'})^2$$

where n is the sample size, x_{jl} is the genotype indicator variable for the additive effect of the j th animal at locus l, $x_{jll'}$ is the genotype indicator variable for the l th additive by l' th additive epistatic effect of the j th animal, which is the direct product of x_{jl} and $x_{jl'}$.

6.3. Results

In this study, 316 evenly distributed SNPs covering the entire autosomal bovine genome were used. The average intermarker spacing was 8.80 Mb (5.53 ~ 13.05 Mb) and the genome coverage of this marker set was 2,546 Mb. The average heterozygosity, average polymorphic information content (PIC) and average MAF were 0.41, 0.32 and 0.31, respectively. Both the additive effect QTL model and epistatic effect QTL model were analyzed to assess the importance of epistasis relative to additivity for traits in dairy cattle. Critical values at an experimental type I error rate of $\alpha = 0.05$ were obtained from the

average of 1000 (additive effect QTL model) and 50 (epistatic effect QTL model) reshuffled samples. These critical values together with the selected ω-values for each trait are listed in Table 6 – 3 and Table 6 – 4.

In the additive effect QTL model, the total number of significant additive QTL effects detected (N_A) ranged from 21 (DCA and MT) to 63 (SCS) with an average of 37. The sum of all significant QTL effects accounted for, on average, 12.87% of the phenotypic variance, with DCA having the highest (22.43%) and MT the lowest (9.24%) proportions (Table 6 – 3). Additive QTL with large effects were found for all traits, with the largest QTL explaining ~1.87% (MY) to 3.36% (SCS) of the phenotypic variance. Different distributions of additive QTL effects were observed for different traits, with MY, SCS and MT having more large effect QTL and DCA having more small to medium effect QTL (Figure 6 – 1, blue lines).

For the epistatic effect QTL model, the total number of additive QTL effects detected (N_A) ranged from 20 (MT) to 81 (SCS) with an average of 50, whereas the total number of detected A × A epistatic QTL effects (N_{AA}) ranged from 3,112 (MT) to 7,996 (SCS) out of the total 49,770 pairwise interaction effects with an average of 5,627 (Table 6 – 4). These estimated QTL effects were plotted against the genome locations (3D) shown in Figure 6 – 2. Different architectures of QTL effects were observed for different traits. The sum of all the significant QTL effects contributed on average 19.91% of the phenotypic variance, with DCA having the highest (27.35%) and MT the lowest (13.64%)

proportions. For the 19.91% total genetic variance, 4.25% and 15.66% were accounted by the additive and epistatic variances, respectively. The largest additive QTL explained ~0.98% (MT) to 3.09% (DCA) of the phenotypic variance, whereas the largest epistatic QTL explained ~1.73% (MY) to 3.93% (MT) of the phenotypic variance (Table 6 – 4). Epistatic QTL with large effects existed for all analyzed traits and the A × A epistatic variance accounted for a large proportion of the total phenotypic variation. Within the total estimated genetic variance $V_G = V_A + V_{AA}$, the relative proportion of epistatic variance V_{AA}/V_G was 65.66%, 75.43%, 82.43%, and 92.85% for MY, SCS, DCA, and MT, respectively. Although the cumulative contribution from significant epistatic effects was generally larger than that from the additive effects, QTL interactions may play more important roles for DCA and MT than for MY and SCS in dairy cattle.

In comparison with the additive model, the epistatic model increased the explained phenotypic variance for all traits, on average, from 12.87% to 19.91%, with SCS having the highest (11.15%) and MT the lowest (4.40%) proportions. With the increase of the total explained phenotypic variance, however, the cumulative contribution from the significant additive effects in the epistatic model decreased by 3.82% (MY), 4.77% (SCS), 17.63% (DCA) and 8.26% (MT) compared to that in the additive model. The architecture changes of additive QTL effects between two models are also shown in Figure 6 – 1 (blue lines vs. red lines). For MY and SCS, most of the additive QTL detected in the additive model were still present with only small changes in the magnitudes of the effects when

the A × A epistatic effects were added into the model. However, many large additive QTL detected for DCA and MT from the additive model were not confirmed using the epistatic model. The architecture changes of the additive QTL effects between the two models were larger for DCA and MT than for MY and SCS.

6.4. Discussion

Genome-wide epistatic QTL mapping was carried out for four dairy traits using 316 evenly spaced markers in Canadian Holstein cattle. Results showed that the total phenotypic variance explained by markers was increased when A × A epistatic effects were considered. The cumulative contribution from significant epistatic effects accounted for a considerably larger proportion of genetic variance than the additive effects. Different genetic architectures of the additive and epistatic effects were observed for different traits.

6.4.1. Model

The epistatic effect QTL model used in this study is an oversaturated model without variable selection. In the empirical Bayes method, the variance component of each marker (hyper-parameter of the normal prior for each regression coefficient) is estimated from the variance component analysis and then used in the Bayesian analysis for estimation of the QTL effect (CARLIN and LOUIS 1996). Phenotypic data are used twice in this method, once for estimation of the prior hyper-parameters and then for estimation of the QTL effects (HU and

XU 2009). The empirical Bayes method without the MCMC sampling for inference of the parameter distribution is computationally efficient, thus is especially suitable for epistatic QTL analysis because of the large number of variables included in the epistatic genetic models. Although there is no hierarchical model used in this method, the scaled inverse chi-square distribution assigned to the variance components for the regression coefficients has two hyper-parameters (τ, ω), where τ is the degree of freedom and ω is the scale parameter (XU 2007). Here, τ was set at –1 in this study since previous studies found other values of τ did not shrink the parameters properly (TER BRAAK *et al.* 2005; XU 2003; XU 2007). Our study found that different ω values can result in different distributions of QTL effects, and the scale parameter ω was finally selected from the range of 15 fitted values that minimized the prediction error for each trait (Table 6 – 3 and Table 6 – 4).

The presence of an epistatic effect between two loci in the empirical Bayes method does not depend on whether or not the two loci both have significant additive effects (XU 2007), whereas some previous methods on epistasis analysis firstly identify loci with significant additive effects and then examine the epistatic effects only among those loci (KAO and ZENG 2002). In this study, 13.64% – 27.35% of the genetic variance was identified by the 316 markers using the epistatic effect QTL model. The remaining unexplained genetic variance may be caused by ignoring the allelic dominance effects at individual loci and the non-additive epistasis among loci since only additive and A × A epistatic effects were considered in the model. Compared to the study of epistasis in barley using the

same method, the absence of dominance effects in double haploid (DH) lines might be one possible reason for the larger explained genetic variance (about 34.93% – 47.00%) by additive and A × A epistatic effects (XU and JIA 2007). In addition, the low density marker panel used in this study might be another reason for the lower identified genetic variances and future studies using higher density of markers may capture more genetic variance.

6.4.2. Sample Size

This study considered only the pair-wise epistatic effects. Higher order interactions, e.g., three-locus interactions, could also be included in the model. However, such models require large sample sizes. A previous study found that population size can dictate the types of identifiable genetic interactions (CARLBORG *et al.* 2006). Compared to the MCMC-based fully Bayesian methods, the empirical Bayes method is more robust to small sample sizes by involving a smaller number of parameters (XU and JIA 2007), in which variance components are estimated separately using a marginal maximum-likelihood method before the Bayes analysis. However, the results from our study were still limited by decoupling 316 additive effects and 49,770 A × A interaction effects simultaneously from the total phenotypic variance using phenotypic records from 647 Holstein cattle. Future studies with much larger sample sizes need to be carried out to validate the results from this study.

6.4.3. Significance Test

In the empirical Bayes method, a large number of effects were shrunk to zero, but many effects deviating from zero still remained. These deviations, very small individually, collectively can contribute to a large proportion of the trait variance, because of the extremely large number of epistatic effects included in the model. In this study, the epistatic QTL model contained a total of 50,086 effects, which was about 77 times as large as the sample size. Xu (2007) suggests using empirical critical values from permutation analysis (CHURCHILL and DOERGE 1994) for separation of the statistically significant QTL from the non-significant ones (XU 2007). Without such a statistical test, the total genetic variance from the large number of non-significant epistatic effects may dominate over the additive variance. In the permutation tests, 1000 reshuffled samples were used for the additive model with a small number of parameters. For the epistatic model, 50 reshuffled samples were used and we found that the variance of the critical values obtained from these samples was small (Table 6 – 5). When taking the average of the 50 reshuffled samples, the coefficient of variation (CV) was about 0.43% – 1.22%. This shows that 50 reshuffled samples should be sufficient for estimating the true critical value in the epistatic model. Finally, each QTL effect deemed significant had a chance of $\alpha = 0.05$ to be a false positive based on the thresholds drawn from the permutation tests.

6.4.4. Genetic Background of Traits

In this study, different distributions of QTL effects were observed for different traits in both the additive model and the epistatic model (Figure 6 – 1 and Figure 6 – 2), and larger architecture changes of the additive QTL between the two models were observed for DCA and MT than for MY and SCS (Figure 6 – 1). It was found that many large additive QTL shown in the additive model for functional traits shrunk dramatically to many small epistatic QTL in the epistatic model, indicating these additive QTL found in the additive model were actually not caused by an individual genetic effect but rather by the combination of beneficial alleles at different loci. These results showed that polygenic traits may have different genetic architectures that are controlled by a few major genes or by many minor gene interactions. Further studies of the gene networks and pathways may help to understand more about the nature of the epistatic effects in the analyzed traits.

Although the A × A epistatic effects appear to be an important component of all dairy traits, the relative importance was larger for the functional traits (DCA and MT) than for the milk related traits (MY and SCS). This may explain why functional traits have relatively low heritability and slow progress on genetic improvement from progeny selection compared to the milk production traits. The low narrow sense heritability in functional traits could result from the large unaccounted epistasis components. A strategy for genetic improvement of these types of traits should capitalize on both the additive and A × A epistatic

interactions. Recently one study in soybean found that the accuracy of genomic value prediction for a quantitative trait was increased from 0.33 when only additive effects were used for prediction, to 0.78 when the epistatic effects were also included in the model (HU *et al.* 2011). This study on soybean showed that using the epistatic model in genomic value prediction for a quantitative trait can help to achieve the maximum efficiency for genetic improvement.

6.4.5. Population History and Selection Response

Here we characterized the role of epistasis on quantitative traits in an outbred cattle population, whereas most previous studies on epistasis used model systems or selected lines derived from the same base population or crosses from extreme strains or breeds (ANKRA-BADU *et al.* 2010; CARLBORG *et al.* 2005; CARLBORG *et al.* 2003; GROSSE-BRINKHAUS *et al.* 2010; XU and JIA 2007; YI *et al.* 2006). As a certain combination of alleles can adapt to a specific environment or a special purpose of production, epistasis was found to be an important source of variation contributing to speciation and breed formation (WRIGHT 1931). The long-term intensive selection for genetic improvement of milk production in dairy cattle is not only the selection of the individual genes, but rather the epistatic selection of combinations of genes or networks affecting the traits (OTTO and WHITLOCK 2009). This might be one possible way to explain the results that epistasis accounted a large proportion of phenotypic variation for dairy traits in the Holstein cattle population in this study. In comparison, epistasis was found to

be small in magnitude relative to the additive effects in less diversified varieties of crops or breeds of animals (XU and JIA 2007; YI *et al.* 2006).

Overall, this study provided the first empirical Bayes analysis of epistasis for quantitative traits in Canadian Holstein cattle. The new statistical model including both the additive and the epistatic genetic effects provided new insights into how genetic interactions in a large set of loci jointly contribute to phenotypic variation. Simultaneous estimation of all genetic effects was an effective method to avoid missing some important epistatic effects. This study suggests that we need a research strategy to identify QTL for economically important traits that embraces, rather than ignores, the complexity of the genotype to phenotype relationship. Including the ignored epistatic effects in conventional QTL mapping studies could result in a dramatic increase of genetic variance explained by markers. The results may explain why there is still a large amount of genetic variation for the selected traits after long term selection in agricultural species with remarkable genetic changes (CARLBORG *et al.* 2006). Further studies on epistasis in dairy cattle should be carried out in populations with larger sample sizes and denser genetic markers before applying these epistatic QTL in future breeding programs. The gene networks underlying these significant epistatic QTL should also be explored to understand the molecular mechanisms mediating the epistasis.

6.5. Conclusions

Genome-wide epistatic QTL mapping was carried out for four quantitative traits in Canadian Holstein cattle using the empirical Bayes method where all additive effects and pair-wise A × A epistatic effects are simultaneously estimated from a single model. For both the milk production related traits and functional traits, epistatic QTL with large effects were identified and the A × A epistatic variance made a substantial contribution to the total phenotypic variation. The cumulative contribution from significant epistatic effects was even larger than that from the additive effects. Different genetic architectures existed for different traits with QTL interactions playing more important roles for functional traits than for the milk production traits. The identified A × A epistatic QTL should be considered in future breeding programs after further validation studies using larger population sizes and denser genome coverage of markers. Studies should also be carried out to explore the relationship between statistical epistasis and biological gene interactions.

6.6. Literature Cited

ANKRA-BADU, G. A., D. SHRINER, E. LE BIHAN-DUVAL, S. MIGNON-GRASTEAU, F. PITEL *et al.*, 2010 Mapping main, epistatic and sex-specific QTL for body composition in a chicken population divergently selected for low or high growth rate. BMC Genomics **11:** 107.

BOER, M. P., C. J. F. TER BRAAK and R. C. JANSEN, 2002 A penalized likelihood method for mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics **162:** 951-960.

BROTHERSTONE, S., and M. GODDARD, 2005 Artificial selection and maintenance of genetic variance in the global dairy cow population. Philosophical Transactions of the Royal Society B-Biological Sciences **360:** 1479-1488.

CARLBORG, O., L. ANDERSSON and B. KINGHORN, 2000 The use of a genetic algorithm for simultaneous mapping of multiple interacting quantitative trait loci. Genetics **155:** 2003-2010.

CARLBORG, O., G. A. BROCKMANN and C. S. HALEY, 2005 Simultaneous mapping of epistatic QTL in DU6i x DBA/2 mice. Mamm Genome **16:** 481-494.

CARLBORG, O., and C. S. HALEY, 2004 Epistasis: too often neglected in complex trait studies? Nature Reviews Genetics **5:** 618-625.

CARLBORG, O., L. JACOBSSON, P. AHGREN, P. SIEGEL and L. ANDERSSON, 2006 Epistasis and the release of genetic variation during long-term selection. Nature Genetics **38:** 418-420.

CARLBORG, O., S. KERJE, K. SCHUTZ, L. JACOBSSON, P. JENSEN *et al.*, 2003 A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome Research **13:** 413-421.

CARLIN, B. P., and T. A. LOUIS, 1996 Bayes and Empirical Bayes Methods for Data Analysis. Chapman & Hall/CRC, London.

CHURCHILL, G. A., and R. W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963-971.

COCKERHAM, C. C., 1954 An Extension of the Concept of Partitioning Hereditary Variance for Analysis of Covariances among Relatives When Epistasis Is Present. Genetics **39:** 859-882.

ELSIK, C. G., R. L. TELLAM, K. C. WORLEY, R. A. GIBBS, A. R. R. ABATEPAULO *et al.*, 2009 The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution. Science **324:** 522-528.

GEORGES, M., D. NIELSEN, M. MACKINNON, A. MISHRA, R. OKIMOTO *et al.*, 1995 Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics **139:** 907-920.

GROSSE-BRINKHAUS, C., E. JONAS, H. BUSCHBELL, C. PHATSARA, D. TESFAYE *et al.*, 2010 Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc x Pietrain population. Genetics Selection Evolution **42:** 39.

HAYES, B., and M. GODDARD, 2010 Genome-wide association and genomic selection in animal breeding. Genome **53:** 876-883.

HU, Z., and S. XU, 2009 PROC QTL-A SAS Procedure for Mapping Quantitative Trait Loci. International Journal of Plant Genomics **2009:** 141234.

HU, Z. Q., Y. G. LI, X. H. SONG, Y. P. HAN, X. D. CAI *et al.*, 2011 Genomic value prediction for quantitative traits under the epistatic model. BMC Genetics **12**.

JANNINK, J. L., and R. JANSEN, 2001 Mapping epistatic quantitative trait loci with one-dimensional genome searches. Genetics **157:** 445-454.

KAO, C. H., and Z. B. ZENG, 2002 Modeling epistasis of quantitative trait loci using Cockerham's model. Genetics **160:** 1243-1261.

KEMPTHORNE, O., 1954 The correlation between relatives in a random mating population. Proceedings of the Royal Society B-Biological Sciences **143:** 102-113.

KHATKAR, M. S., P. C. THOMSON, I. TAMMEN and H. W. RAADSMA, 2004 Quantitative trait loci mapping in dairy cattle: review and meta-analysis. Genetics Selection Evolution **36:** 163-190.

KINGHORN, B. P., 1998 Managing genetic change under operational and cost constraints. 36th National Congress of the South African Association of Animal Science. University of Stellenbosch 5-8 April**:** 9-16.

MATUKUMALLI, L. K., C. T. LAWLEY, R. D. SCHNABEL, J. F. TAYLOR, M. F. ALLAN *et al.*, 2009 Development and Characterization of a High Density SNP Genotyping Assay for Cattle. PLoS One **4**.

MOORE, J. H., 2005 A global view of epistasis. Nature Genetics **37:** 13-14.

OTTO, S. P., and M. C. WHITLOCK, 2009 The impact of epistatic selection on the genomic traces of selection. Molecular Ecology **18:** 4985-4987.

TER BRAAK, C. J. F., M. P. BOER and M. C. A. M. BINK, 2005 Extending Xu's Bayesian model for estimating polygenic effects using markers of the entire genome. Genetics **170:** 1435-1438.

TIBSHIRANI, R., 1996 Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society Series B-Methodological **58:** 267-288.

UEMOTO, Y., S. SATO, C. OHNISHI, S. TERAI, A. KOMATSUDA *et al.*, 2009 The effects of single and epistatic quantitative trait loci for fatty acid composition in a Meishan x Duroc crossbred population. Journal of Animal Science **87:** 3470-3476.

WANG, D. L., J. ZHU, Z. K. LI and A. H. PATERSON, 1999 Mapping QTLs with epistatic effects and QTLxenvironment interactions by mixed linear model approaches. Theoretical and Applied Genetics **99:** 1255-1264.

WRIGHT, S., 1931 Evolution in Mendelian Populations. Genetics **16:** 97-159.

XU, S., and Z. JIA, 2007 Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics **175:** 1955-1963.

XU, S. Z., 2003 Estimating polygenic effects using markers of the entire genome. Genetics **163:** 789-801.

XU, S. Z., 2007 An empirical Bayes method for estimating epistatic effects of quantitative trait loci. Biometrics **63:** 513-521.

YI, N., D. SHRINER, S. BANERJEE, T. MEHTA, D. POMP *et al.*, 2007 An efficient Bayesian model selection approach for interacting quantitative trait loci models with many effects. Genetics **176:** 1865-1877.

YI, N., and S. XU, 2002 Mapping quantitative trait loci with epistatic effects. Genetics Research **79:** 185-198.

YI, N., B. S. YANDELL, G. A. CHURCHILL, D. B. ALLISON, E. J. EISEN *et al.*, 2005 Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. Genetics **170:** 1333-1344.

YI, N., D. K. ZINNIEL, K. KIM, E. J. EISEN, A. BARTOLUCCI *et al.*, 2006 Bayesian analyses of multiple epistatic QTL models for body weight and body composition in mice. Genetics Research **87:** 45-60.

YI, N. J., S. Z. XU and D. B. ALLISON, 2003 Bayesian model choice and search strategies for mapping interacting quantitative trait loci. Genetics **165:** 867-883.

ZENG, Z. B., C. H. KAO and C. J. BASTEN, 1999 Estimating the genetic architecture of quantitative traits. Genetics Research **74:** 279-289.

ZHANG, Q., D. BOICHARD, I. HOESCHELE, C. ERNST, A. EGGEN *et al.*, 1998 Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. Genetics **149:** 1959-1973.

ZHANG, Y. M., and S. XU, 2005 A penalized maximum likelihood method for estimating epistatic effects of QTL. Heredity **95:** 96-104.

Tables

Table 6 – 1. Descriptive statistics for four economically important traits in Canadian Holstein cattle.

Trait	N	Mean	SD	Min	Max
MY	647	588.64	839.14	−1609.30	2962.11
SCS	647	3.06	0.32	2.17	4.31
DCA	524	99.65	7.40	69.73	117.82
MT	501	100.27	7.39	71.92	124.33

MY = milk yield; SCS = somatic cell score; DCA = daughter calving ability; MT = milking temperament.

Table 6 – 2. Number of SNPs, intermarker distance (Mb), average heterozygosity (He), average polymorphic information content (PIC) and average MAF of SNPs on 29 autosomes in Canadian Holstein cattle.

BTA	N	BTA length	Max. spacing	Min. spacing	Average spacing	Average He	Average PIC	Average MAF
1	20	161.11	9.40	7.82	8.46	0.43	0.33	0.34
2	16	140.80	10.49	8.66	9.33	0.39	0.31	0.28
3	16	127.92	12.44	5.61	8.39	0.40	0.32	0.32
4	15	124.45	9.53	8.27	8.82	0.40	0.31	0.30
5	13	125.85	12.27	9.49	10.44	0.44	0.33	0.34
6	15	122.56	9.43	8.07	8.73	0.35	0.28	0.24
7	14	112.08	11.56	5.53	8.50	0.41	0.32	0.32
8	14	116.94	9.35	8.72	8.97	0.38	0.30	0.27
9	12	108.15	10.50	9.25	9.75	0.43	0.33	0.33
10	13	106.38	13.05	5.97	8.72	0.46	0.35	0.38
11	14	110.17	9.55	7.90	8.46	0.40	0.32	0.31
12	10	85.36	9.82	8.86	9.43	0.40	0.31	0.29
13	11	84.42	8.74	8.04	8.37	0.38	0.30	0.29
14	11	81.35	8.68	7.40	8.13	0.47	0.35	0.38
15	10	84.63	9.66	9.04	9.33	0.39	0.31	0.27
16	10	77.91	9.10	8.23	8.63	0.40	0.31	0.28
17	10	76.51	9.09	7.57	8.48	0.38	0.30	0.29
18	9	66.14	9.73	7.38	8.18	0.41	0.32	0.31
19	9	65.31	9.01	7.48	8.12	0.41	0.32	0.30
20	9	75.80	10.72	7.43	9.34	0.45	0.35	0.35
21	9	69.17	8.78	8.26	8.63	0.40	0.31	0.29
22	8	61.85	9.17	8.43	8.79	0.41	0.31	0.31
23	7	53.38	9.13	8.52	8.82	0.43	0.32	0.32
24	8	65.02	9.77	8.73	9.17	0.40	0.31	0.31
25	7	44.06	7.40	7.03	7.22	0.36	0.28	0.27
26	7	51.75	8.69	8.25	8.44	0.40	0.31	0.29
27	6	48.75	9.82	9.41	9.72	0.42	0.34	0.33
28	6	46.08	10.15	8.62	9.15	0.48	0.37	0.44
29	7	52.00	9.45	8.10	8.60	0.40	0.31	0.29
All	316	2545.90	13.05	5.53	8.80	0.41	0.32	0.31

Table 6 – 3. Genetic contribution of significant QTL effects in the additive effect QTL model at an experimental type I error rate of $\alpha = 0.05$.

Trait	MY	SCS	DCA	MT
V_P	703056	0.0994	54.7556	54.6586
ω-value	5.00E-01	5.00E-06	5.00E-01	5.00E-02
Critical value	0.0189	0.0005	0.9034	0.2332
N_A	44	63	21	21
V_A	69172.1	0.0099	12.2808	5.0523
H_A	0.0984	0.0996	0.2243	0.0924
$V_{A\,(MAX)}$	13144.2	0.0033	1.0557	1.0264
$H_{A\,(MAX)}$	0.0187	0.0336	0.0193	0.0188

MY = milk yield; SCS = somatic cell score; DCA = daughter calving ability; MT = milking temperament; V_P: phenotypic variance; ω-value: value of hyper-parameter ω that minimized the prediction error; Critical value: genome-wise critical value for declaration of significant QTL effects at 5% level from permutation test using 1,000 reshuffled samples; N_A: number of significant additive effects; V_A: summed variance of additive effects; H_A: proportion of additive variance; $V_{A\,(MAX)}$: variance of the largest additive effect; $H_{A\,(MAX)}$: H_A of the largest additive effect.

Table 6 – 4. Genetic contribution of significant QTL effects in the epistatic effect QTL model at an experimental type I error rate of $\alpha = 0.05$.

Trait	MY	SCS	DCA	MT
V_P	703056.2	0.0994	54.7556	54.6586
ω-value	5.00E-01	5.00E-09	5.00E-05	5.00E-04
Critical value	1.64E-02	3.93E-07	1.47E-04	1.53E-03
N_A	49	81	48	20
N_{AA}	5303	7996	6098	3112
V_A	42292.41	0.0052	2.6309	0.5330
V_{AA}	80866.48	0.0158	12.3464	6.9206
V_G	123158.9	0.0210	14.9773	7.4536
H_A	0.0602	0.0519	0.0480	0.0098
H_{AA}	0.1150	0.1593	0.2255	0.1266
H_G	0.1752	0.2111	0.2735	0.1364
$H_{A\ (MAX)}$	0.0130	0.0207	0.0309	0.0097
$H_{AA\ (MAX)}$	0.0173	0.0225	0.0335	0.0393

MY = milk yield; SCS = somatic cell score; DCA = daughter calving ability; MT = milking temperament; V_P: phenotypic variance; ω-value: value of hyper-parameter ω that minimized the prediction error; Critical value: genome-wise critical value for declaration of significant QTL effects at 5% level from permutation test using 50 reshuffled samples; N_A: number of significant additive effects; N_{AA}: number of significant A × A epistatic effects; V_A: summed variance of additive effects; V_{AA}: summed variance of A × A epistatic effects; V_G: total genetic variance; H_A: proportion of additive variance; H_{AA}: proportion of A × A epistatic variance; H_G: proportion of total genetic variance; $H_{A(MAX)}$: H_A of the largest additive effect; $H_{AA(MAX)}$: H_{AA} of the largest A × A epistatic effect.

Table 6 – 5. Standard deviation (SD) and coefficient of variation (CV) of the critical values at 5% level from permutation test in the epistatic effect QTL model.

Trait	Critical value	SD	CV	$CV / \sqrt{50}$
MY	0.0164	5.02E-04	0.0305	0.0043
SCS	3.93E-07	2.06E-08	0.0524	0.0074
DCA	1.47E-04	1.26E-05	0.0860	0.0122
MT	1.53E-03	4.88E-05	0.0320	0.0045

MY = milk yield; SCS = somatic cell score; DCA = daughter calving ability; MT = milking temperament; Critical value: genome-wise critical value for declaration of significant QTL effects at 5% level from permutation test using 50 reshuffled samples.

Figures

Figure 6 – 1A

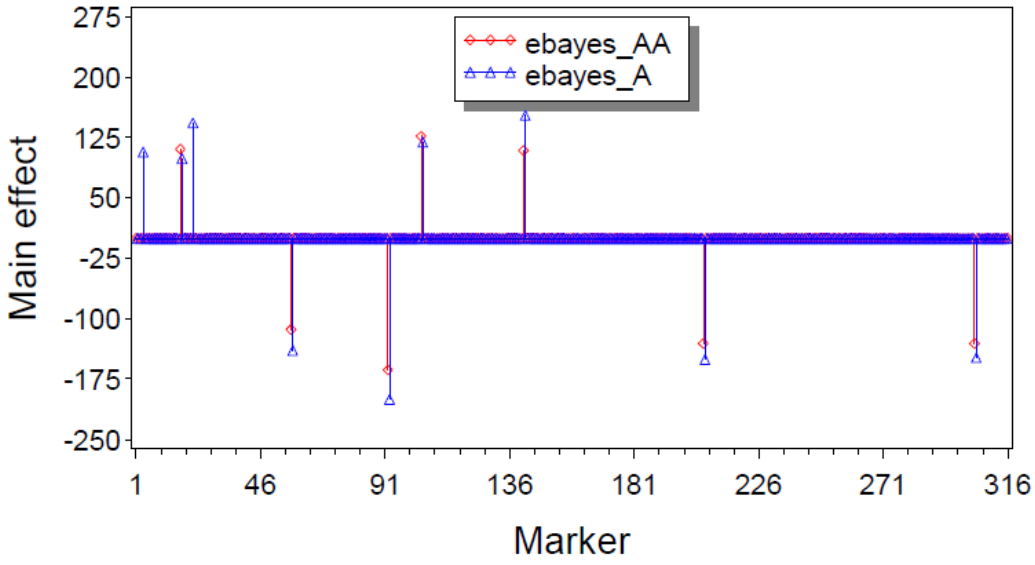

Figure 6 – 1B

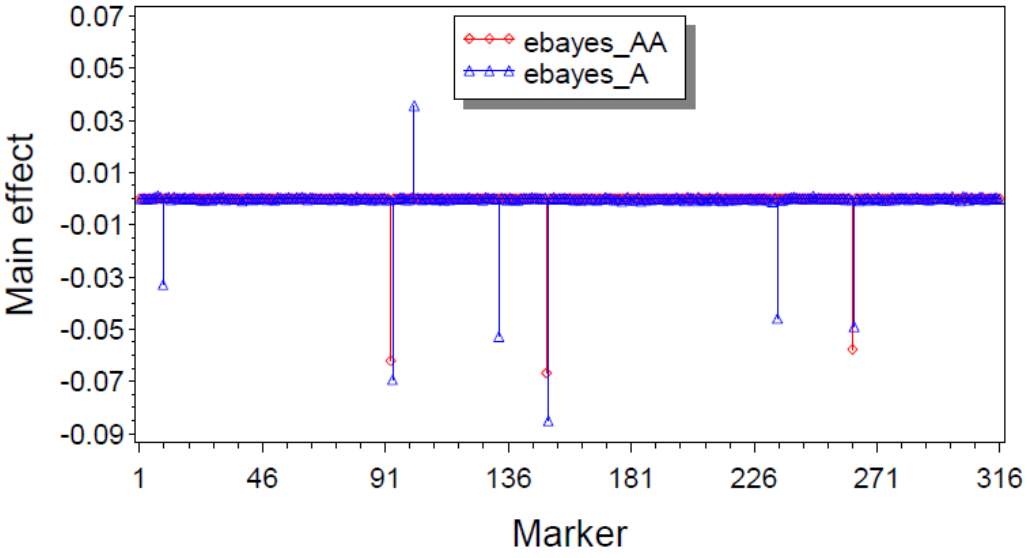

Figure 6 – 1C

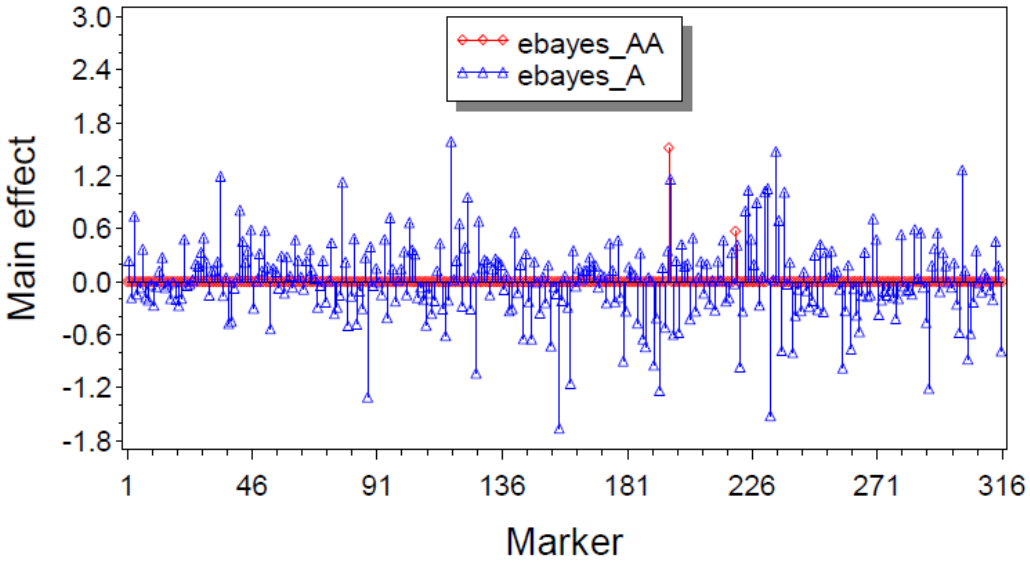

Figure 6 – 1D

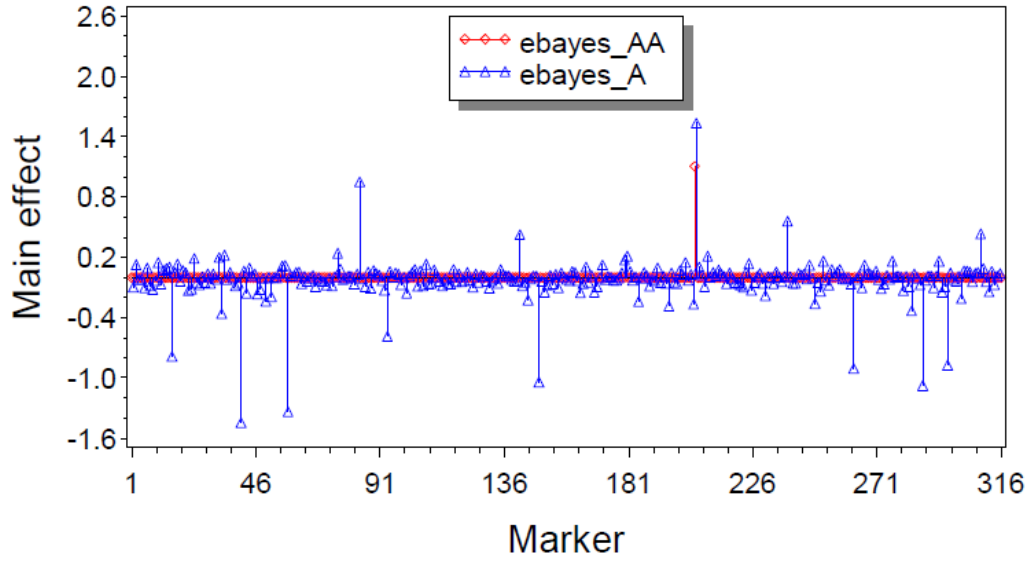

Figure 6 – 1. Additive QTL effects for four agronomic traits in the Canadian Holstein cattle from both the additive effect QTL model (ebayes_A with blue lines) and the epistatic effect QTL model (ebayes_AA with red lines): A) milk yield; B) somatic cell score; C) daughter calving ability; D) milking temperament.

Figure 6 – 2A

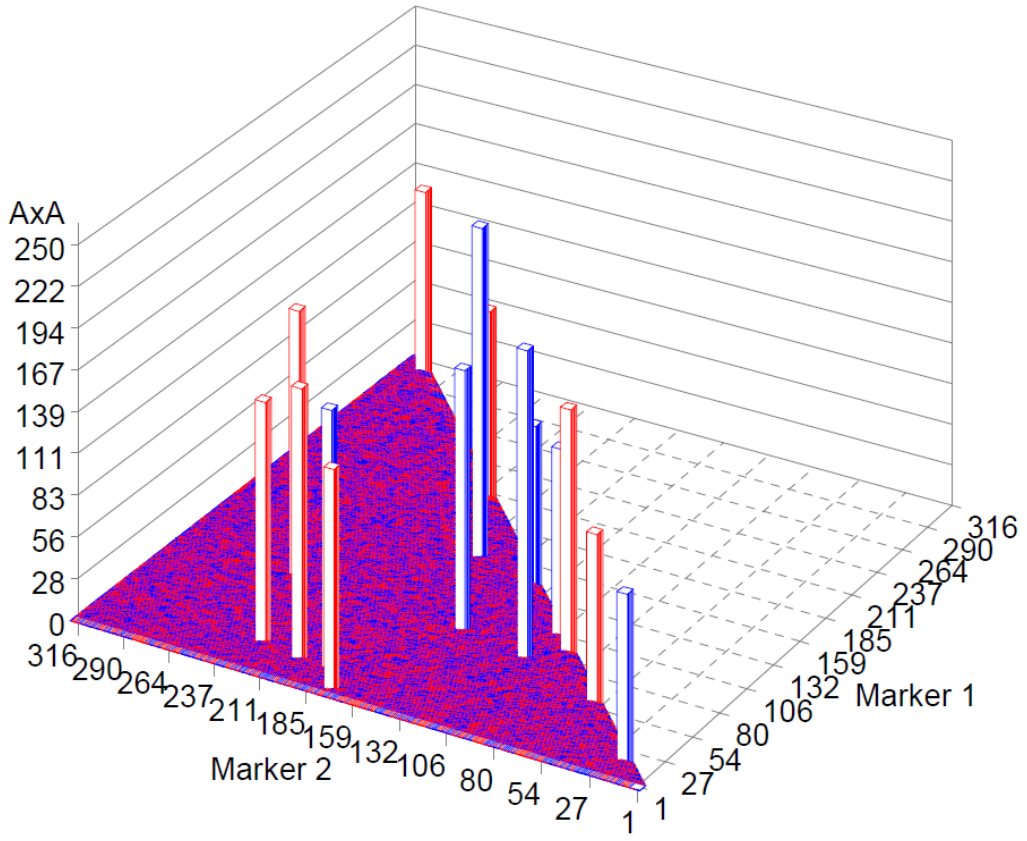
AxA
250
222
194
167
139
111
83
56
28
0
316
290
264
237
211
185
159
132
106
80
54
27
1
Marker 2
316
290
264
237
211
185
159
132
106
80
54
27
1
Marker 1

Figure 6 – 2B

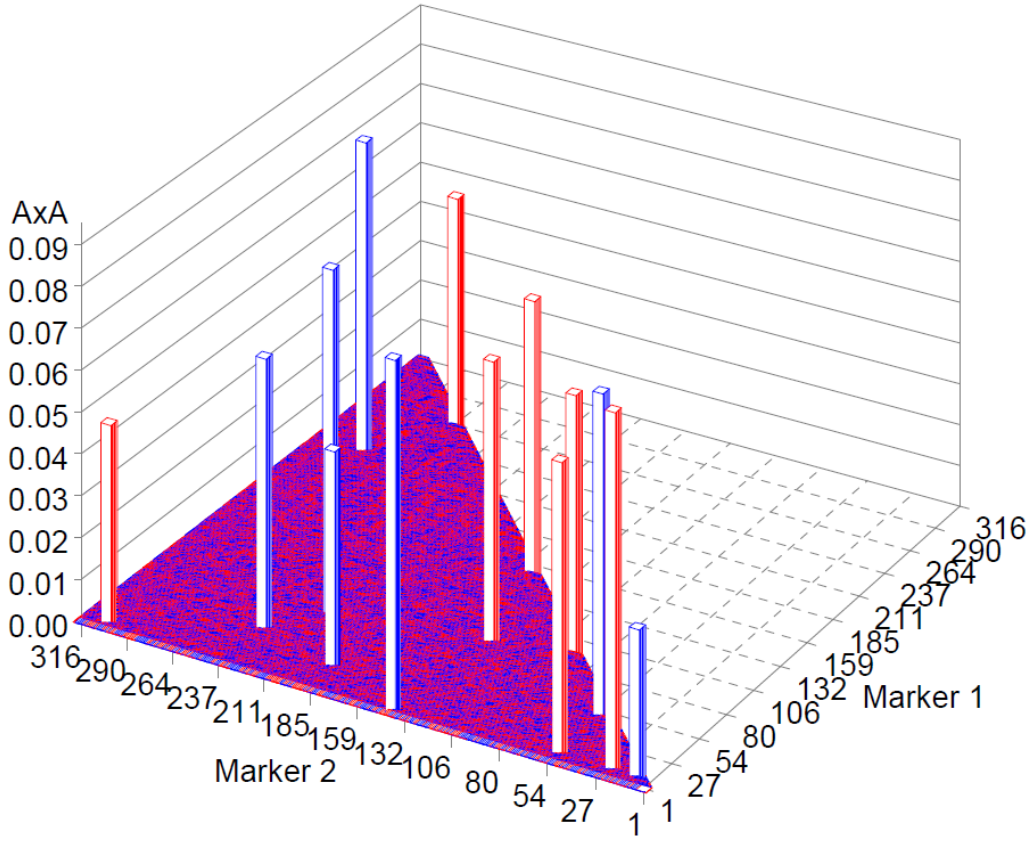
AxA
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
316
290
264
237
211
185
159
132
106
80
54
27
1
Marker 2
Marker 1

Figure 6 – 2C

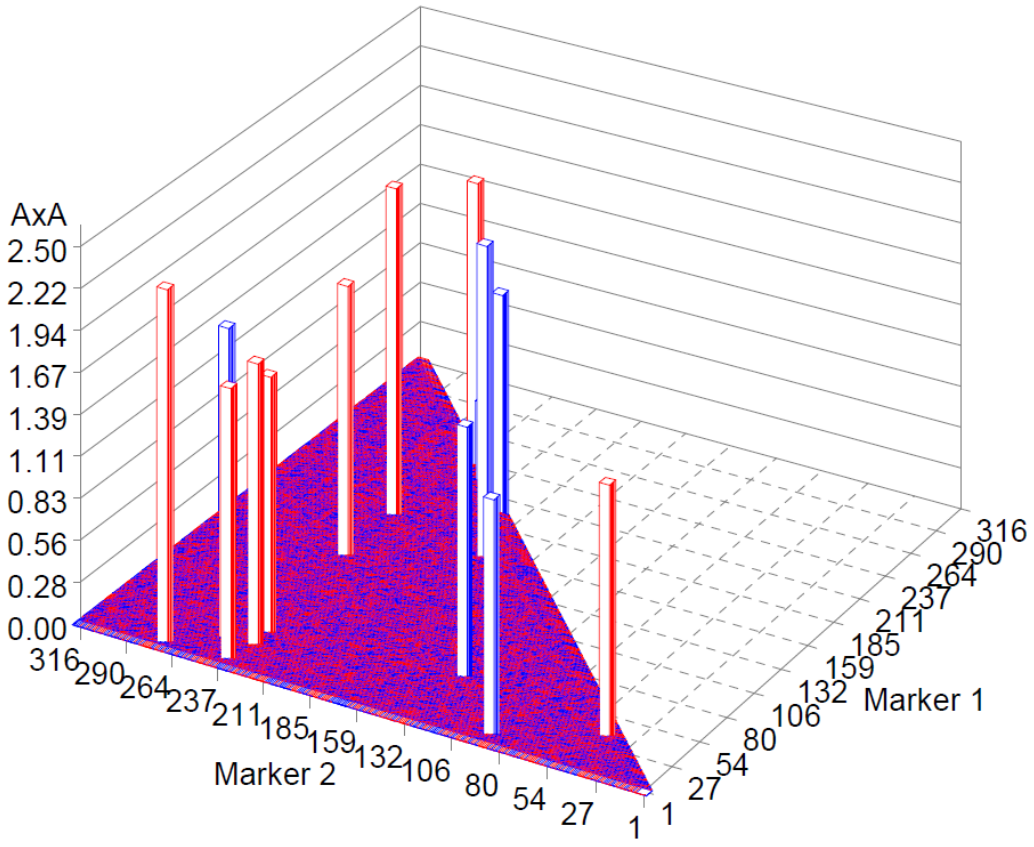
AxA
2.50
2.22
1.94
1.67
1.39
1.11
0.83
0.56
0.28
0.00
316
290
264
237
211
185
159
132
106
80
54
27
1
Marker 2
Marker 1

Figure 6 – 2D

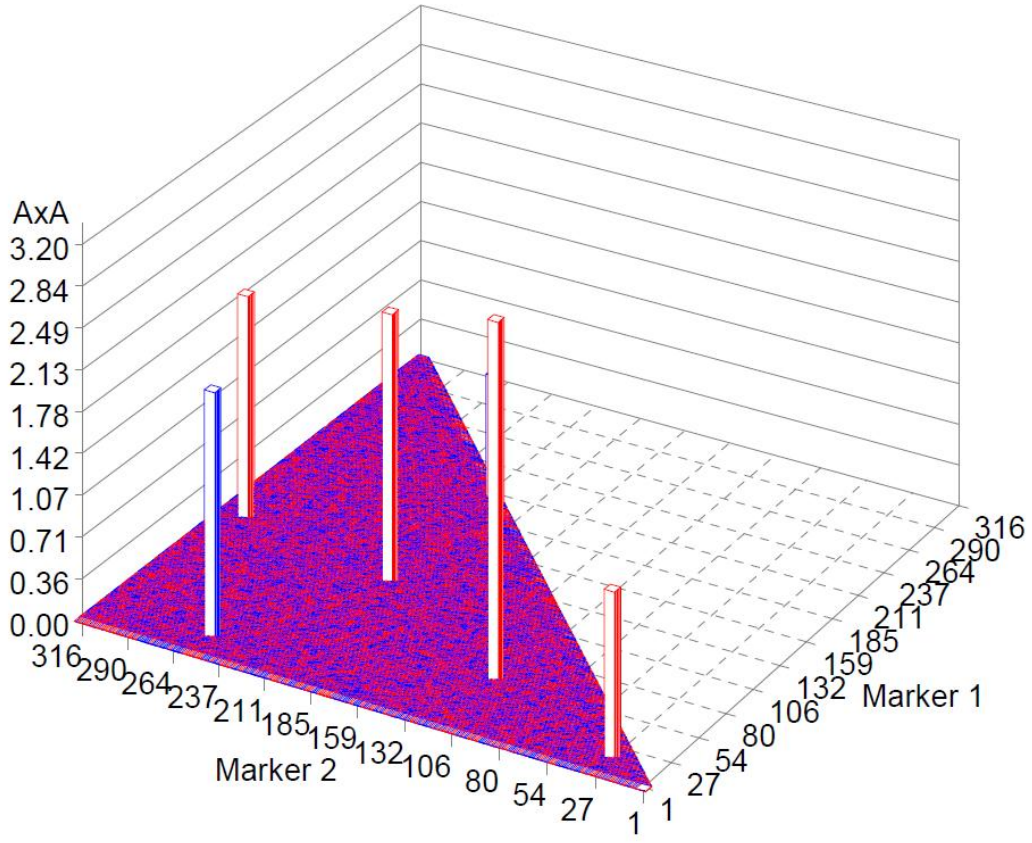

Figure 6 – 2. Additive and additive-by-additive (A × A) epistatic QTL effects for four traits in the Canadian Holstein from the epistatic effect QTL model: A) milk yield; B) somatic cell score; C) daughter calving ability; D) milking temperament. The additive effects are shown on the diagonals and the A × A epistatic effects are on the left triangle of the 3D plots. Blue prisms represent positive effects and red prisms represent negative effects.

Chapter 7. General Discussion and Future Directions

7.1. General Discussion

Milk production and beef carcass yield and quality determine much of the economic profits of the cattle industry and provide important sources of human food and nutrition. Genetic improvement of these economically important traits has the potential to better feed the world population, which has now surpassed 7 billion and is expected to reach 9.1 billion in 2050 (http://www.prb.org/). There have been increasing successes worldwide by incorporating genetic marker information into the genetic breeding programs of livestock, especially in cattle (DEKKERS 2004; HAYES *et al.* 2009). The major objective of this thesis was to identify genomic regions or markers contributing to the phenotypic variation of milk and beef production traits in cattle using a 50 K SNP marker panel with efforts to disclose major genes determining these traits.

Chapter 3 presents a comparative assessment of genome-wide LD and haplotype block structure in one Canadian Holstein population and one hybrid beef cattle population using ~40, 000 SNPs with an average marker spacing of 65 kb. Larger LD was observed in the Holstein cattle with r^2 averaging ~0.21 at 100 kb in contrast to ~0.14 in the hybrid beef. Results in Holstein cattle were consistent with previous LD studies in dairy cattle (BOHMANOVA *et al.* 2010; KHATKAR *et al.* 2008), while much less extensive LD in hybrid beef cattle was found compared with the LD in the Japanese Black and Brown beef cattle using sparse markers (ODANI *et al.* 2006). Our study of LD in beef cattle using high

density markers is more reliable since the characterization of a complete LD map should use ~30,000 uniformly distributed SNPs (VILLA-ANGULO *et al.* 2009). The strong LD observed in Holstein cattle could result from the intensive artificial selection of dairy bulls with genetics for strong milking ability through artificial insemination and progeny testing in the past half century (BROTHERSTONE and GODDARD 2005). Characterization of LD before the association studies in our specific research cattle populations is of great importance because it will help us to determine the feasibility of a QTL mapping method and the required marker density. In addition, the LD phase between a genetic marker and a QTL in one population can be different or even reversed in another population. Our results on the empirical LD indicated that denser SNPs were necessary for the hybrid beef cattle for the purpose of association studies.

The genome-wide comparison of the haplotype block pattern exhibited clear differentiation between the Holstein and the hybrid beef cattle in terms of block numbers, genome coverage of blocks, average block size and block boundary discordances, whereas limited haplotype diversities existed for both populations. Interestingly, the block size identified in this study was larger than previous studies with either higher or lower marker densities (KHATKAR *et al.* 2007; VILLA-ANGULO *et al.* 2009), indicating that further studies should be carried out to better understand the effect of marker density on the disclosure of haplotype blocks. The differences of block structure between the two populations revealed in this study could be mainly from the distinct population demographic histories. However, our results were inconsistent with a previous study on small

genomic regions with a total of 7.6 Mb, where non differentiable haplotype block structure between dairy and beef breeds were reported (VILLA-ANGULO *et al.* 2009). This disagreement could be due to the different density of markers and number of animals used, different methods for characterizing the haplotype blocks, and different measures used to quantify block similarities and block boundary consistency between populations. A regional comparison of block structures revealed that distinct selection of economically important traits in cattle breeds may have a role in shaping the block pattern and that the observed common blocks in both breeds could result from similar selection processes acting on genomic regions affecting multiple traits (GURYEV *et al.* 2006). The negative correlation between the block size and average marker heterozygosity within blocks may also provide some hints on the effects of selection on haplotype block formation through selective sweeps. Overall, comparative analyses of haplotype block structure could be employed for the future detection of genomic regions that have been subject to selective sweeps, where most often the functionally important genes are located (GURYEV *et al.* 2006; KHATKAR *et al.* 2007; MCKAY *et al.* 2007).

After characterization of the LD and haplotype block regions, high-density SNP markers were utilized to perform whole genome-wide association studies for milk production traits in Canadian Holstein bulls (n = 647) (**Chapter 4**). We tested two statistical methods: single marker LD regression and Bayesian regression using the BayesB method (MEUWISSEN *et al.* 2001). Previous studies had reported that the simple single locus LD regression model has good power

and accuracy for QTL fine mapping (GRAPES *et al.* 2004; ZHAO *et al.* 2007) and in this study this method identified a total of 316 SNPs on 28 chromosomes at a genome-wise FDR of 5% for five milk production traits. Compared with previous studies in Canadian Holstein cattle using the same method (DAETWYLER *et al.* 2008; KOLBEHDARI *et al.* 2009), our analyses with the BovineSNP50 BeadChip (50K) assay identified a larger number of significant SNPs. We also conducted the first GWAS using the BayesB method for milk production traits in Canadian Holstein cattle; however, only 1 – 3 suggestive QTL regions with $P < 0.2$ were discovered for each trait. The low number of detected QTL in the Bayesian analyses suggests that a larger population size will be required for better power to detect the associated genomic regions for these traits.

In **Chapter 5**, genome-wide SNPs were tested for their associations with beef carcass traits in hybrid beef cattle (n = 922). In comparison with the ~30 K SNPs selected for association studies in dairy cattle, we selected ~40 K SNPs in the association analyses for beef cattle since the extent of LD in beef was found to be smaller than in dairy. Similar to Chapter 4, larger number of associations were identified by using the single marker analyses, with 62 genome-wise and 328 chromosome-wise significant (FDR $P < 0.05$) SNPs detected on 24 chromosomes for eight beef carcass traits. A total of 12 suggestive QTL regions ($P < 0.2$) on 9 chromosomes were detected for the beef carcass traits using the BayesB method.

Both the association studies for milk production traits and beef carcass traits disclosed different distributions of significant SNPs on 29 chromosomes,

with a large proportion of the associations clustered on BTA1, 5, 11 and 14 for milk production traits and on BTA2, 6, 8 and 13 for beef carcass traits. This may to some extent indicate the different genetic control for economically important traits in cattle. In Chapter 4 and 5, both methods were shown to be able to confirm several important QTL regions reported by many previous studies for similar traits. Several novel QTL regions were also identified for the analyzed traits. In contrast to most previously reported QTL using linkage analysis, the QTL span in this study was greatly shrunk by using the LD regression mapping methods exploiting population-wised LD with high-density markers (MEUWISSEN and GODDARD 2000). In addition, we observed considerable overlap of significant associations between the two methods used. Notably, several of the novel associations were detected by both methods. The consistent results from different methods can increase our confidence in these novel QTL, which may be specific to these cattle populations. For the typical Canadian cattle production system using specific cattle breeds, identification of population-specific QTL regions and markers is an important step towards enhancing the genetic improvement programs for these specific cattle populations.

In comparison with other GWAS analyses with high density SNP chips recently carried out for milk production traits (JIANG *et al.* 2010; MAI *et al.* 2010; PRYCE *et al.* 2010) and for beef carcass traits (BOLORMAA *et al.* 2011; KIM *et al.* 2011; LEE *et al.* 2010), our studies included systematic searches for candidate genes or gene networks following the GWAS. Both of the studies in Chapter 4 and Chapter 5 identified several gene networks among positional candidate genes

and functional candidate genes. The gene networks identified from single marker analyses and the candidate genes screened from Bayesian regression showed similar functions, with those found for milk production traits showing functions related to lipid metabolism, small molecular biochemistry and intracellular molecular transport and those associated with beef carcass traits having functions in lipid metabolism, cellular growth and proliferation, and skeletal and muscular system development and function. In addition, both studies in Chapter 4 and 5 rediscovered some previously documented candidate genes, for example, *DGAT1* with milk production traits (GRISART *et al.* 2002; GRISART *et al.* 2004; WINTER *et al.* 2002) and *NCAPG* with muscle mass and growth regulation (EBERLEIN *et al.* 2009; WEIKARD *et al.* 2010). The novel functional candidate genes and potential networks will together contribute to a better understanding of the molecular mechanisms of milk synthesis and secretion and beef growth and composition in cattle and could be applied to improve accuracy of genetic evaluation of these traits after verification studies in other populations. Many candidate genes previously reported for analyzed traits through the candidate gene approach were not re-identified in our association studies. This is in agreement with human studies where the candidate gene approach, in which associations are tested for a few genes based on known functions, has been declared woefully inadequate since most declared disease genes were not detected using GWAS (ALTSHULER *et al.* 2008). Compared to the candidate gene approaches with guesswork in choice of candidate genes and variations within them, more insights on the genetic control of the phenotype of interest can be gained from GWAS.

In both Chapter 4 and 5, larger numbers of significant associations were identified from the single marker LD regression compared to the Bayesian regression. The Bayesian method, which fits all markers simultaneously in a model, requires more phenotypic data than single marker analysis to have the same power in detecting the associated QTL regions. However, results from the Bayesian regression using the non-overlapping 1Mb SNP sliding windows will facilitate future work on gene identification by removing many redundant QTL regions caused by the high LD (FAN *et al.* 2011; ONTERU *et al.* 2011; SUN *et al.* 2011).

Following the one-dimensional genome scans for associations between single markers and phenotype, we had also conducted a genome-wide epistatic QTL mapping study in **Chapter 6**, in which both the additive (A) and additive-by-additive (A × A) epistatic effects were considered for quantitative traits in Canadian Holstein cattle. We applied a multi-effect analysis using the empirical Bayes method (XU 2007), that requires no variable selection, and 316 evenly spaced markers. Our results showed that the total phenotypic variance explained by markers was increased, on average, from 12.87% to 19.91% when A × A epistatic effects were considered. This trend was consistent with previously reported epistasis in other species (CARLBORG *et al.* 2005; GROSSE-BRINKHAUS *et al.* 2010). The results may explain why only small proportions of genetic variance were captured by marginal effects (HAYES and GODDARD 2010) and why there was still a large amount of genetic variation for the selected traits after long term selection (CARLBORG *et al.* 2006). The observed large magnitude of epistasis

relative to the additive effects in dairy cattle could result from the long term epistatic selection for certain combinations of alleles that can adapt to a particular type of production during speciation and breed formation (OTTO and WHITLOCK 2009; WRIGHT 1931).

The epistasis study also set a good example for exploring the genetic architecture of important agronomic traits. Several additive QTL found in a search for only marginal additive effects, were not detected in the epistatic model, because the epistasis cancelled out the individual effects of the QTL. This suggests that the additive QTL found in the additive model are actually not caused by an individual genetic effect but rather by the combination of beneficial alleles at different loci. These large additive effects were split into many small interaction effects in the epistatic model and caused a decrease in the genetic contribution from the additive effects. Our results suggest that QTL interactions may be of differing importance for different types of traits in dairy cattle and that those traits with low narrow sense heritability could actually have large unaccounted epistasis components. Breeding programs with consideration of both the additive and A × A epistatic interactions should be carried out for faster genetic improvement of these type of traits.

The empirical Bayes method used in this study for epistasis analysis has advantages over other previously used variable selection methods (KAO and ZENG 2002; YI *et al.* 2005; ZENG *et al.* 1999) for simultaneous estimation of all effects, and is more computationally efficient than previous shrinkage methods (YI and

XU 2002; ZHANG and XU 2005). Our study confirmed that the empirical Bayes method without the MCMC sampling for inference of the parameter distribution was computationally efficient and suitable for epistatic QTL analysis with a large number of parameters (HU and XU 2009; XU 2007). In addition, the presence of epistatic effects between two loci in the empirical Bayes method did not depend on whether or not the two loci both have significant additive effects (XU 2007) and this method will work especially in the situation where no main effect is displayed (CULVERHOUSE *et al.* 2002).

7.2. Future Directions

Although larger numbers of SNPs were applied in this thesis in comparison with previous ones (KHATKAR *et al.* 2007; KIM and KIRKPATRICK 2009), future studies with even denser markers might be necessary for complete views of fine-scale haplotype block patterns in the bovine genome and may provide a better comparison of block structures between dairy and beef cattle (KHATKAR *et al.* 2007; VILLA-ANGULO *et al.* 2009). In addition, assessing the effects of selection on haplotype block patterns will require further in depth studies in other beef populations with much clearer breeding and selection history. Furthermore, studies of LD in other purebred beef cattle populations should also be carried out since the reported LD for our hybrid beef cattle population as a recently admixed population cannot well represent LD in other purebred beef cattle populations.

The novel QTL reported in this thesis, especially those overlapping across different methods, should be further validated in other cattle populations to establish unbiased association before using in MAS for genetic improvement of dairy and beef production traits. Future association studies using Bayesian regression should be mainly explored since this method can avoid the problems of model selection, multiple testing and marker correlation, and estimates effects of dense markers simultaneously (BALDING 2006; BEAUMONT and RANNALA 2004). In addition to the BayesB (MEUWISSEN *et al.* 2001) and BayesC (KIZILKAYA *et al.* 2010) methods performed in Chapter 4 and 5, other newly developed Bayesian methods, *i.e.*, BayesCπ and BayesDπ (HABIER *et al.* 2011), can further aid in identifying associations for traits of interest. Regarding the power of Bayesian analysis in this thesis, future studies using larger sample sizes of phenotypic records should be carried out for better power to detect significant QTL regions at higher significance levels. Another direction to follow emerging from this work is the functional analysis of genes reported in this thesis in further identification of causal mutations. More specifically, the genes should be sequenced and variation within them should be further tested for their association with the traits of interest.

Future QTL mapping studies that embrace epistatic effects should be carried out to recognize the effect of epistasis on quantitative traits in outbreed cattle populations. Although the empirical Bayes method was more robust to small sample sizes compared to the MCMC-based fully Bayesian method (XU and JIA 2007), the epistatic QTL reported for dairy traits in this thesis should be further validated in other dairy populations with much larger sample sizes and

denser genetic markers before applying these epistatic QTL in future breeding programs. Higher order interactions, e.g., three-locus interactions could also be considered in future studies. In addition to the additive epistasis considered in this thesis, the allelic dominance effects and non-additive epistasis among loci should be examined in future studies. Furthermore, a more in-depth understanding of the biological and genetic mechanisms underlying these significant epistatic QTL should also be explored beyond the statistical estimates of these QTL effects.

7.3. Conclusions

The present study was carried out to characterize the structure of bovine genomic variation and to map important genomic regions or genes via GWAS for different economically important traits using the BovineSNP50 BeadChip (50K). By characterizing LD, we presented the first high-density genome-wide reference haplotype block map for beef cattle and the first genome-wide comparative haplotype block maps between dairy and beef cattle. The patterns of LD and haplotype blocks could be used in designing and interpreting association studies and could help in the detection of functionally important genomic regions and genes which show significant evidence of positive selection. GWAS was shown to be a powerful and efficient method in association studies by which we detected several novel QTL regions in both dairy and beef cattle and highlighted functional candidate genes and networks for further investigation. This study also describes the first GWAS analyses using the Bayesian regression method in cattle. The information on associated markers and genes has the potential for being utilized in

future breeding programs for selection of genetically superior animals. The genome-wide epistatic QTL analyses in Canadian Holstein cattle identified strong epistasis with considerable contribution to the phenotypic variation of quantitative traits in cattle. Different genetic architectures existed for different traits with QTL interactions playing more important roles for traits with lower heritability. The identified A × A epistatic QTL may be considered in future breeding programs after further validation studies using larger population sizes and denser genome coverage of markers.

7.4. Literature Cited

ALTSHULER, D., M. J. DALY and E. S. LANDER, 2008 Genetic Mapping in Human Disease. Science **322:** 881-888.

BALDING, D. J., 2006 A tutorial on statistical methods for population association studies. Nature Reviews Genetics **7:** 781-791.

BEAUMONT, M. A., and B. RANNALA, 2004 The Bayesian revolution in genetics. Nature Reviews Genetics **5:** 251-261.

BOHMANOVA, J., M. SARGOLZAEI and F. S. SCHENKEL, 2010 Characteristics of linkage disequilibrium in North American Holsteins. BMC Genomics **11**.

BOLORMAA, S., L. R. P. NETO, Y. D. ZHANG, R. J. BUNCH, B. E. HARRISON *et al.*, 2011 A genome-wide association study of meat and carcass traits in Australian cattle. Journal of Animal Science **89:** 2297-2309.

BROTHERSTONE, S., and M. GODDARD, 2005 Artificial selection and maintenance of genetic variance in the global dairy cow population. Philosophical Transactions of the Royal Society B-Biological Sciences **360:** 1479-1488.

CARLBORG, O., G. A. BROCKMANN and C. S. HALEY, 2005 Simultaneous mapping of epistatic QTL in DU6i x DBA/2 mice. Mammalian Genome **16:** 481-494.

CARLBORG, O., L. JACOBSSON, P. AHGREN, P. SIEGEL and L. ANDERSSON, 2006 Epistasis and the release of genetic variation during long-term selection. Nature Genetics **38:** 418-420.

CULVERHOUSE, R., B. K. SUAREZ, J. LIN and T. REICH, 2002 A perspective on epistasis: Limits of models displaying no main effect. American Journal of Human Genetics **70:** 461-471.

DAETWYLER, H. D., F. S. SCHENKEL, M. SARGOLZAEI and J. A. B. ROBINSON, 2008 A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. Journal of Dairy Science **91:** 3225-3236.

DEKKERS, J. C., 2004 Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. Journal of Animal Science **82 E-Suppl:** E313-328.

EBERLEIN, A., A. TAKASUGA, K. SETOGUCHI, R. PFUHL, K. FLISIKOWSKI *et al.*, 2009 Dissection of Genetic Factors Modulating Fetal Growth in Cattle Indicates a Substantial Role of the Non-SMC Condensin I Complex, Subunit G (NCAPG) Gene. Genetics **183:** 951-964.

FAN, B., S. K. ONTERU, Z. Q. DU, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Genome-Wide Association Study Identifies Loci for Body Composition and Structural Soundness Traits in Pigs. PLoS One **6**.

GRAPES, L., J. C. M. DEKKERS, M. F. ROTHSCHILD and R. L. FERNANDO, 2004 Comparing linkage disequilibrium-based methods for fine mapping quantitative trait loci. Genetics **166:** 1561-1570.

GRISART, B., W. COPPIETERS, F. FARNIR, L. KARIM, C. FORD *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: Identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Research **12:** 222-231.

GRISART, B., F. FARNIR, L. KARIM, N. CAMBISANO, J. J. KIM *et al.*, 2004 Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proceedings of the National Academy of Sciences of the United States of America **101:** 2398-2403.

GROSSE-BRINKHAUS, C., E. JONAS, H. BUSCHBELL, C. PHATSARA, D. TESFAYE *et al.*, 2010 Epistatic QTL pairs associated with meat quality and carcass composition traits in a porcine Duroc x Pietrain population. Genetics Selection Evolution **42**.

GURYEV, V., B. M. SMITS, J. VAN DE BELT, M. VERHEUL, N. HUBNER *et al.*, 2006 Haplotype block structure is conserved across mammals. PLoS Genetics **2:** e121.

HABIER, D., R. L. FERNANDO, K. KIZILKAYA and D. J. GARRICK, 2011 Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics **12**.

HAYES, B., and M. GODDARD, 2010 Genome-wide association and genomic selection in animal breeding. Genome **53:** 876-883.

HAYES, B. J., P. J. BOWMAN, A. J. CHAMBERLAIN and M. E. GODDARD, 2009 Invited review: Genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science **92:** 433-443.

HU, Z., and S. XU, 2009 PROC QTL-A SAS Procedure for Mapping Quantitative Trait Loci. International Journal of Plant Genomics **2009:** 141234.

JIANG, L., J. LIU, D. SUN, P. MA, X. DING *et al.*, 2010 Genome wide association studies for milk production traits in Chinese Holstein population. PLoS One **5**.

KAO, C. H., and Z. B. ZENG, 2002 Modeling epistasis of quantitative trait loci using Cockerham's model. Genetics **160:** 1243-1261.

KHATKAR, M. S., F. W. NICHOLAS, A. R. COLLINS, K. R. ZENGER, J. AL CAVANAGH *et al.*, 2008 Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics **9**.

KHATKAR, M. S., K. R. ZENGER, M. HOBBS, R. J. HAWKEN, J. A. L. CAVANAGH *et al.*, 2007 A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in Holstein-Friesian cattle. Genetics **176:** 763-772.

KIM, E. S., and B. W. KIRKPATRICK, 2009 Linkage disequilibrium in the North American Holstein population. Animal Genetics **40:** 279-288.

KIM, Y., J. RYU, J. WOO, J. B. KIM, C. Y. KIM *et al.*, 2011 Genome-wide association study reveals five nucleotide sequence variants for carcass traits in beef cattle. Animal Genetics **42:** 361-365.

KIZILKAYA, K., R. L. FERNANDO and D. J. GARRICK, 2010 Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. Journal of Animal Science **88:** 544-551.

KOLBEHDARI, D., Z. WANG, J. R. GRANT, B. MURDOCH, A. PRASAD *et al.*, 2009 A whole genome scan to map QTL for milk production traits and somatic cell score in Canadian Holstein bulls. Journal of Animal Breeding and Genetics **126:** 216-227.

LEE, Y. M., C. M. HAN, Y. LI, J. J. LEE, L. H. KIM *et al.*, 2010 A Whole Genome Association Study to Detect Single Nucleotide Polymorphisms for Carcass Traits in Hanwoo Populations. Asian-Australasian Journal of Animal Sciences **23:** 417-424.

MAI, M. D., G. SAHANA, F. B. CHRISTIANSEN and B. GULDBRANDTSEN, 2010 A genome-wide association study for milk production traits in Danish Jersey cattle using a 50K single nucleotide polymorphism chip. Journal of Animal Science **88:** 3522-3528.

MCKAY, S. D., R. D. SCHNABEL, B. M. MURDOCH, L. K. MATUKUMALLI, J. AERTS *et al.*, 2007 Whole genome linkage disequilibrium maps in cattle. BMC Genetics **8:** 74.

MEUWISSEN, T. H. E., and M. E. GODDARD, 2000 Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. Genetics **155:** 421-430.

MEUWISSEN, T. H. E., B. J. HAYES and M. E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157:** 1819-1829.

ODANI, M., A. NARITA, T. WATANABE, K. YOKOUCHI, Y. SUGIMOTO *et al.*, 2006 Genome-wide linkage disequilibrium in two Japanese beef cattle breeds. Animal Genetics **37:** 139-144.

ONTERU, S. K., B. FAN, M. T. NIKKILA, D. J. GARRICK, K. J. STALDER *et al.*, 2011 Whole-genome association analyses for lifetime reproductive traits in the pig. Journal of Animal Science **89:** 988-995.

OTTO, S. P., and M. C. WHITLOCK, 2009 The impact of epistatic selection on the genomic traces of selection. Molecular Ecology **18:** 4985-4987.

PRYCE, J. E., S. BOLORMAA, A. J. CHAMBERLAIN, P. J. BOWMAN, K. SAVIN *et al.*, 2010 A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. Journal of Dairy Science **93:** 3331-3345.

SUN, X., D. HABIER, R. L. FERNANDO, D. J. GARRICK and J. C. DEKKERS, 2011 Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. BMC Proceedings **5 Suppl 3:** S13.

VILLA-ANGULO, R., L. K. MATUKUMALLI, C. A. GILL, J. CHOI, C. P. VAN TASSELL *et al.*, 2009 High-resolution haplotype block structure in the cattle genome. BMC Genetics **10**.

WEIKARD, R., E. ALTMAIER, K. SUHRE, K. M. WEINBERGER, H. M. HAMMON *et al.*, 2010 Metabolomic profiles indicate distinct physiological pathways affected by two loci with major divergent effect on Bos taurus growth and lipid deposition. Physiological Genomics **42A:** 79-88.

WINTER, A., W. KRAMER, F. A. O. WERNER, S. KOLLERS, S. KATA *et al.*, 2002 Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA : diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proceedings of the National Academy of Sciences of the United States of America **99:** 9300-9305.

WRIGHT, S., 1931 Evolution in Mendelian populations. Genetics **16:** 0097-0159.

XU, S., and Z. JIA, 2007 Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics **175:** 1955-1963.

XU, S. Z., 2007 An empirical Bayes method for estimating epistatic effects of quantitative trait loci. Biometrics **63:** 513-521.

YI, N. J., and S. Z. XU, 2002 Mapping quantitative trait loci with epistatic effects. Genetics Research **79:** 185-198.

YI, N. J., B. S. YANDELL, G. A. CHURCHILL, D. B. ALLISON, E. J. EISEN *et al.*, 2005 Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. Genetics **170:** 1333-1344.

ZENG, Z. B., C. H. KAO and C. J. BASTEN, 1999 Estimating the genetic architecture of quantitative traits. Genetics Research **74:** 279-289.

ZHANG, Y. M., and S. XU, 2005 A penalized maximum likelihood method for estimating epistatic effects of QTL. Heredity **95:** 96-104.

ZHAO, H. H., R. L. FERNANDO and J. C. M. DEKKERS, 2007 Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. Genetics **175:** 1975-1986.

Appendices

Appendix 1. Overlapping genome-wise significant SNPs (FDR $P < 0.05$) among milk production traits from whole genome association study using single marker LD regression in Canadian Holstein cattle.

Trait	MY	FY	PY	FP	PP	Total
MY	94	42	10	61	54	-
FY	-	97	4	59	52	-
PY	-	-	17	4	3	-
FP	-	-	-	206	103	-
PP	-	-	-	-	136	-
Total	-	-	-	-	-	3

MY = milk yield (kg); FY = fat yield (kg); PY = protein yield (kg); FP = fat percentage (%); PP = protein percentage (%).

Appendix 2. Summary of positional candidate genes in potential gene network for milk yield from whole genome association study using single marker LD regression in Canadian Holstein cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
ABHD10	1	57,435,818	57,448,955	Cytoplasm	other
SLC9A10	1	57,635,019	57,731,162	unknown	other
SLC35A5	1	58,077,010	58,097,614	unknown	transporter
TM4SF4	1	120,600,100	120,627,687	Plasma Membrane	other
HPS3	1	120,950,144	120,992,566	Cytoplasm	other
HLTF	1	121,020,645	121,084,579	Nucleus	transcription regulator
ARMC8	1	132,863,648	132,970,686	unknown	other
PWP2	1	147,388,099	147,401,301	Nucleus	other
CCDC77	5	114,201,495	114,228,221	unknown	other
ERC1	5	114,717,781	114,958,856	Cytoplasm	other
SLC25A17	5	119,197,333	119,239,897	Cytoplasm	transporter
XPNPEP3	5	119,270,170	119,311,327	Cytoplasm	peptidase
TOB2	5	119,695,117	119,705,164	Nucleus	other
POLR3H	5	119,772,236	119,784,777	Nucleus	enzyme
NAGA	5	120,182,058	120,190,242	Cytoplasm	enzyme
C22orf32	5	120,198,466	120,202,165	Extracellular Space	other
TCF20	5	120,260,419	120,309,324	Nucleus	transcription regulator
TSPO	5	121,103,387	121,116,024	Cytoplasm	transmembrane receptor
KDM4B	7	17,649,227	17,728,367	unknown	other
PLIN5	7	18,197,116	18,208,271	unknown	other
MPND	7	18,330,624	18,340,107	unknown	other
CCDC94	7	18,394,674	18,408,673	unknown	other
MTRF1L	9	93,437,385	93,443,266	Cytoplasm	translation regulator
CALML4	10	14,851,879	14,861,294	unknown	other
FSIP1	10	35,272,363	35,332,774	unknown	other
NUP214	11	105,059,653	105,150,009	Nucleus	transporter
COMMD5	14	118,906	120,941	Nucleus	other
LRRC14	14	195,644	199,106	unknown	other
GPT	14	209,124	212,128	Cytoplasm	enzyme
TONSL	14	266,557	277,561	Cytoplasm	transcription regulator
CPSF1	14	396,495	404,637	Nucleus	other
OPLAH	14	599,578	608,670	unknown	enzyme
CHRAC1	14	2,424,524	2,427,765	Nucleus	enzyme
LRRC6	14	8,049,136	8,120,135	Cytoplasm	other
KIAA0196	14	14,554,115	14,601,800	Cytoplasm	other
TRMT12	14	15,108,499	15,110,329	unknown	other
OPA3	18	52,900,437	52,966,902	Cytoplasm	other
SNRPD2	18	53,038,139	53,040,997	Nucleus	other
SYMPK	18	53,136,461	53,165,702	Cytoplasm	other
HIF3A	18	53,354,238	53,385,788	Nucleus	transcription regulator
STRN4	18	53,619,561	53,643,096	Cytoplasm	other
SLC1A5	18	53,657,678	53,669,150	Plasma Membrane	transporter
AP2S1	18	53,724,455	53,734,090	Cytoplasm	transporter
NPAS1	18	53,876,672	53,892,486	Nucleus	transcription regulator
ZC3H4	18	53,912,204	53,947,958	unknown	other
ABCC3	19	37,220,837	37,268,446	Plasma Membrane	transporter
MRPL27	19	37,482,259	37,487,088	Cytoplasm	other
ACSBG1	21	30,613,012	30,672,217	Cytoplasm	enzyme
ABCB10	28	499,875	532,292	Cytoplasm	transporter
URB2	28	1,732,053	1,758,192	Nucleus	other

Appendix 3. Summary of positional candidate genes in potential gene network for protein yield from whole genome association study using single marker LD regression in Canadian Holstein cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
TM4SF4	1	120,600,100	120,627,687	Plasma Membrane	other
HPS3	1	120,950,144	120,992,566	Cytoplasm	other
LRRC3	1	147,112,789	147,115,122	unknown	other
PWP2	1	147,388,099	147,401,301	Nucleus	other
NAGA	5	120,182,058	120,190,242	Cytoplasm	enzyme
C22orf32	5	120,198,466	120,202,165	Extracellular Space	other
CALML4	10	14,851,879	14,861,294	unknown	other
FEM1B	10	14,922,884	14,934,497	Nucleus	transcription regulator
COMMD5	14	118,906	120,941	Nucleus	other
GPT	14	209,124	212,128	Cytoplasm	enzyme
FOXH1	14	239,960	241,362	Nucleus	transcription regulator
LRRC6	14	8,049,136	8,120,135	Cytoplasm	other
TXNL1	24	58,173,961	58,216,690	Cytoplasm	enzyme

Appendix 4. Summary of positional candidate genes in potential gene network for protein percentage from whole genome association study using single marker LD regression in Canadian Holstein cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
PRMT6	3	39,208,742	39,210,100	Nucleus	enzyme
DKK2	6	19,591,210	19,721,448	Extracellular Space	other
N4BP2	6	61,484,389	61,538,984	Cytoplasm	kinase
CHRNA9	6	61,702,088	61,715,473	Plasma Membrane	transmembrane receptor
SCFD2	6	71,168,458	71,566,216	unknown	transporter
FIP1L1	6	71,571,106	71,635,861	Nucleus	other
LNX1	6	71,636,847	71,763,175	Cytoplasm	enzyme
TGFBRAP1	11	9,473,152	9,506,953	Cytoplasm	other
GCC2	11	46,531,790	46,565,122	Cytoplasm	other
SULT1C4	11	46,569,792	46,577,456	Cytoplasm	enzyme
Sult1c2	11	46,592,931	46,598,171	Cytoplasm	enzyme
SLC5A7	11	46,721,311	46,744,137	Plasma Membrane	transporter
PITRM1	13	45,247,547	45,272,018	Cytoplasm	peptidase
RECQL4	14	199,404	410,175	Nucleus	enzyme
FOXH1	14	239,960	241,362	Nucleus	transcription regulator
Tssk5	14	543,582	545,871	unknown	kinase
OPLAH	14	599,578	608,670	unknown	enzyme
TSTA3	14	906,366	911,206	Plasma Membrane	enzyme
GRINA	14	1,059,703	1,060,760	unknown	ion channel
LY6E	14	1,385,601	1,389,305	Plasma Membrane	other
KHDRBS3	14	5,764,816	5,908,681	Nucleus	other
ST3GAL1	14	7,372,446	7,387,594	Cytoplasm	enzyme
OC90	14	8,532,296	8,556,535	Extracellular Space	enzyme
NSMCE2	14	14,310,935	14,541,975	Nucleus	other
RNF139	14	15,052,567	15,099,405	Cytoplasm	enzyme
AZIN1	14	59,765,276	59,795,942	Cytoplasm	enzyme
PCSK7	15	26,327,640	26,353,523	Cytoplasm	peptidase
FXYD2	15	26,912,132	26,919,643	Plasma Membrane	ion channel
SH3D19	17	7,154,107	7,201,974	Plasma Membrane	other
ETFDH	17	42,272,291	42,328,446	Cytoplasm	enzyme
C17orf70	19	52,714,691	52,722,710	Nucleus	other
FSCN2	19	52,725,152	52,730,938	Cytoplasm	other
AZI1	19	52,991,344	53,004,927	Cytoplasm	other
AATK	19	53,049,375	53,064,869	Cytoplasm	kinase
CHMP6	19	53,157,453	53,164,741	Cytoplasm	other
NPTX1	19	53,556,091	53,561,995	Extracellular Space	other
SNX18	20	26,210,150	26,238,612	Cytoplasm	transporter
OXCT1	20	34,833,498	34,953,986	Cytoplasm	enzyme
RPL37	20	35,785,937	35,788,537	Cytoplasm	other
GCM2	23	46,089,932	46,096,024	Nucleus	transcription regulator

Appendix 5. Overlapping chromosome-wise significant SNPs (FDR $P < 0.05$) among beef carcass traits from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Trait	CWT	CABF	CREA	CGF	CLMY	UMAR	UBF	UREA	Overall
CWT	69	7	25	0	0	0	0	1	-
CABF	-	14	4	4	4	0	0	0	-
CREA	-	-	146	0	2	0	0	0	-
CGF	-	-	-	26	5	0	1	0	-
CLMY	-	-	-	-	13	0	0	0	-
UMAR	-	-	-	-	-	11	0	0	-
UBF	-	-	-	-	-	-	24	2	-
UREA	-	-	-	-	-	-	-	25	-
Overall	-	-	-	-	-	-	-	-	0

CWT = carcass weight (kg); CABF = carcass average backfat (mm); CREA = carcass ribeye area (cm^2); CGF = carcass grade fat (mm); CLMY = carcass lean meat yield (%); UMAR = ultrasound marbling; UBF = ultrasound backfat (mm); UREA = ultrasound ribeye area (cm^2).

Appendix 6. Summary of positional candidate genes in potential gene networks for carcass weight from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
RAPGEF4	2	24,377,850	24,732,350	Cytoplasm	other
PDK1	2	24,822,252	24,866,940	Cytoplasm	kinase
ITGA6	2	24,966,035	25,051,812	Plasma Membrane	other
DLX2	2	25,287,355	25,289,363	Nucleus	transcription regulator
DLX1	2	25,300,700	25,304,443	Nucleus	transcription regulator
HAT1	2	25,403,525	25,443,422	Nucleus	enzyme
SLC25A12	2	25,477,147	25,575,796	Cytoplasm	transporter
FABP2	6	6,817,374	6,820,334	Cytoplasm	transporter
MYOZ2	6	6,929,164	6,995,820	Cytoplasm	other
SYNPO2	6	7,093,095	7,306,171	Cytoplasm	other
SEC24D	6	7,364,958	7,476,797	Cytoplasm	transporter
MANBA	6	23,674,703	23,802,364	Cytoplasm	enzyme
NFKB1	6	23,818,264	23,940,305	Nucleus	transcription regulator
SLC39A8	6	24,031,738	24,113,066	Extracellular Space	transporter
BANK1	6	24,492,639	24,639,895	unknown	other
BMPR1B	6	31,251,194	31,303,130	Plasma Membrane	kinase
HPGDS	6	32,201,690	32,214,073	Cytoplasm	enzyme
SMARCAD1	6	32,220,382	32,308,866	Nucleus	enzyme
ATOH1	6	32,560,789	32,562,969	Nucleus	transcription regulator
GRID2	6	33,050,484	33,114,326	Plasma Membrane	ion channel
MMRN1	6	36,344,037	36,417,274	Extracellular Space	other
SNCA	6	36,504,250	36,651,060	Cytoplasm	other
NAP1L5	6	36,894,197	36,896,330	unknown	other
HERC3	6	36,907,814	36,958,912	Cytoplasm	enzyme
PIGY	6	37,069,642	37,072,440	Plasma Membrane	other
HERC5	6	37,076,253	37,120,870	Cytoplasm	enzyme
HERC6	6	37,128,540	37,185,776	Cytoplasm	enzyme
PKD2	6	37,431,967	37,490,645	Plasma Membrane	ion channel
SPP1	6	37,511,674	37,518,672	Extracellular Space	cytokine
IBSP	6	37,668,427	37,682,179	Extracellular Space	other
LAP3	6	37,961,794	37,987,164	Cytoplasm	peptidase
FAM184B	6	38,001,547	38,059,383	unknown	other
NCAPG	6	38,153,047	38,199,149	Nucleus	other
LCORL	6	38,231,086	38,327,423	Nucleus	transcription regulator
SLIT2	6	41,199,406	41,353,160	Extracellular Space	other
PACRGL	6	41,397,759	41,418,385	unknown	other
KCNIP4	6	41,420,001	41,596,754	Plasma Membrane	ion channel
GPR125	6	43,200,819	43,296,065	Plasma Membrane	G-protein coupled receptor
PPARGC1A	6	44,732,239	44,838,841	Nucleus	transcription regulator
DHX15	6	45,484,488	45,538,459	Nucleus	enzyme
SOD3	6	45,741,186	45,744,454	Extracellular Space	enzyme
SEPSECS	6	46,059,735	46,094,483	Cytoplasm	other
PI4K2B	6	46,163,517	46,199,651	Cytoplasm	kinase
ANAPC4	6	46,279,716	46,312,880	Nucleus	enzyme
TLR10	6	60,338,715	60,345,705	Plasma Membrane	transmembrane receptor
KLHL5	6	60,633,501	60,706,518	unknown	other
UGDH	6	60,965,528	61,000,293	Nucleus	enzyme
C4orf34	6	61,012,429	61,071,031	unknown	other
UBE2K	6	61,155,723	61,230,738	Cytoplasm	transcription regulator
PDS5A	6	61,266,410	61,394,604	Nucleus	other
N4BP2	6	61,484,389	61,538,984	Cytoplasm	kinase
RHOH	6	61,579,571	61,627,670	Plasma Membrane	enzyme
CHRNA9	6	61,702,088	61,715,473	Plasma Membrane	transmembrane receptor
GUF1	6	66,059,818	66,084,364	Cytoplasm	other
GABRG1	6	66,951,630	67,047,315	Plasma Membrane	ion channel
ANKRD17	6	91,166,355	91,254,785	unknown	other

ALB	6	91,461,165	91,479,536	Extracellular Space	transporter
AFP	6	91,487,384	91,508,932	Extracellular Space	transporter
IL8	6	91,790,453	91,794,218	Extracellular Space	cytokine
CXCL5	6	91,877,879	91,880,039	Extracellular Space	cytokine
GPR98	7	91,336,981	91,557,203	Plasma Membrane	G-protein coupled receptor
ARRDC3	7	92,055,611	92,063,674	unknown	other
RP1	14	22,193,803	22,202,968	Cytoplasm	other
TMEM68	14	22,891,757	22,927,546	unknown	other
TGS1	14	22,927,647	22,953,141	Nucleus	enzyme
LYN	14	23,085,065	23,133,139	Cytoplasm	kinase
MOS	14	23,188,641	23,189,213	Cytoplasm	kinase
PLAG1	14	23,219,718	23,221,723	Nucleus	transcription regulator
ARMC1	14	29,948,965	29,973,726	Cytoplasm	other
DNAJC5B	14	30,300,948	30,393,961	unknown	transporter
TRIM55	14	30,414,892	30,461,308	Cytoplasm	other
EYA1	14	35,074,955	35,233,609	Nucleus	phosphatase
HNF4G	14	35,500,695	35,635,567	Nucleus	transcription regulator
FBXW8	17	61,148,228	61,262,700	unknown	other
COX7B	17	61,163,457	61,163,876	Cytoplasm	enzyme
MED13L	17	61,935,480	61,995,820	Nucleus	other
CNDP2	24	3,910,194	3,920,980	Cytoplasm	peptidase
C18orf55	24	4,154,843	4,160,558	Cytoplasm	other
FBXO15	24	4,161,227	4,199,112	unknown	other
RNF138	24	26,159,987	26,199,831	unknown	other
RNF125	24	26,224,547	26,267,764	unknown	other
B4GALT6	24	26,541,313	26,606,363	Cytoplasm	enzyme
TTR	24	26,625,148	26,634,024	Extracellular Space	transporter
DSG2	24	26,685,789	26,716,184	Plasma Membrane	other
DSG3	24	26,748,956	26,779,570	Plasma Membrane	other
DSG1	24	26,842,426	26,886,351	Plasma Membrane	other
DSC1	24	27,036,237	27,074,363	Plasma Membrane	other
DSC2	24	27,100,875	27,130,987	Plasma Membrane	other
EIF3A	26	39,705,112	39,735,630	Cytoplasm	translation regulator
PRDX3	26	39,793,508	39,802,856	Cytoplasm	enzyme
GRK5	26	40,039,082	40,091,265	Plasma Membrane	kinase
RGS10	26	40,136,209	40,177,089	Cytoplasm	other
TIAL1	26	40,201,143	40,220,306	Nucleus	transcription regulator

Appendix 7. Summary of positional candidate genes in potential gene networks for carcass ribeye area from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
EPHA6	1	42,348,344	42,921,576	Plasma Membrane	kinase
GABRR3	1	43,172,901	43,209,533	Plasma Membrane	transmembrane receptor
HS6ST1	2	4,193,997	4,231,176	Plasma Membrane	enzyme
UGGT1	2	4,309,971	4,409,788	Cytoplasm	enzyme
SAP130	2	4,461,692	4,557,199	Nucleus	transcription regulator
NAB1	2	5,839,634	5,876,300	Nucleus	transcription regulator
INPP1	2	6,140,073	6,173,067	Cytoplasm	phosphatase
HIBCH	2	6,192,073	6,348,621	Cytoplasm	enzyme
C2orf88	2	6,354,242	6,432,420	unknown	other
CHRNA1	2	22,722,693	22,742,289	Plasma Membrane	transmembrane receptor
LOC285141	2	26,523,660	26,538,811	unknown	other
MYO3B	2	26,762,397	27,004,945	unknown	kinase
BBS5	2	27,547,152	27,568,469	Cytoplasm	other
LRP2	2	27,705,743	27,923,795	Plasma Membrane	transporter
DHRS9	2	27,961,841	27,985,729	Cytoplasm	enzyme
G6PC2	2	28,152,299	28,159,020	Cytoplasm	phosphatase
NOSTRIN	2	28,195,356	28,258,237	Cytoplasm	transcription regulator
GALNT3	2	31,590,413	31,614,133	Cytoplasm	enzyme
SCN3A	2	32,142,020	32,229,198	Plasma Membrane	ion channel
KCNH7	2	35,071,518	35,228,855	Plasma Membrane	ion channel
SLC4A10	2	35,593,894	35,858,015	Plasma Membrane	transporter
ACVR1	2	39,926,385	40,003,195	Plasma Membrane	kinase
ACVR1C	2	40,185,872	40,233,876	Plasma Membrane	kinase
GALNT5	2	40,431,059	40,472,517	Cytoplasm	enzyme
DOCK10	2	117,416,542	117,654,254	unknown	other
TIPRL	3	938,838	1,008,061	unknown	other
DCAF6	3	1,123,224	1,254,298	Nucleus	transcription regulator
MPZL1	3	1,478,302	1,560,160	Plasma Membrane	other
RGS5	3	6,961,395	7,014,145	Plasma Membrane	other
SOSTDC1	4	26,097,541	26,101,785	Extracellular Space	other
TSPAN13	4	26,380,963	26,419,028	Plasma Membrane	other
AGR2	4	26,426,406	26,438,905	Extracellular Space	other
TWISTNB	4	29,683,187	29,694,034	Nucleus	other
MACC1	4	30,049,715	30,075,448	Nucleus	other
GPNMB	4	33,202,335	33,228,580	Plasma Membrane	enzyme
IGF2BP3	4	33,269,266	33,308,105	Cytoplasm	translation regulator
DMTF1	4	34,323,197	34,373,852	Nucleus	transcription regulator
AOAH	4	63,021,465	63,207,269	unknown	enzyme
Chn2	4	69,341,355	69,678,883	Plasma Membrane	other
JAZF1	4	70,724,244	71,054,807	Nucleus	transcription regulator
HIBADH	4	71,206,571	71,314,100	Cytoplasm	enzyme
FEZF1	4	89,766,920	89,769,219	unknown	other
AKR1B10	4	101,949,003	101,966,623	Cytoplasm	enzyme
WDR91	4	102,522,670	102,548,004	unknown	other
TRIM24	4	105,819,538	105,932,400	Nucleus	transcription regulator
ATP6V0A4	4	106,040,038	106,086,748	Cytoplasm	transporter
ZC3HAV1L	4	106,295,893	106,306,710	unknown	other
KIAA1549	4	106,313,359	106,361,757	unknown	other
UBN2	4	106,433,379	106,502,620	unknown	other
CLEC2L	4	106,705,627	106,708,876	unknown	other
JHDM1D	4	107,214,741	107,251,383	Nucleus	enzyme
PIP	4	110,485,785	110,495,960	Extracellular Space	other
TMEM139	4	110,727,627	110,729,506	unknown	other
EPHA1	4	110,841,565	110,855,892	Plasma Membrane	kinase
SLC4A2	4	117,917,252	117,929,812	Plasma Membrane	transporter
AGAP3	4	117,940,377	117,995,737	Nucleus	transcription regulator

SMARCD3	4	118,105,517	118,138,732	Nucleus	transcription regulator
RHEB	4	118,301,814	118,353,690	Cytoplasm	other
PRKAG2	4	118,383,509	118,681,815	Cytoplasm	kinase
COL25A1	6	17,789,311	17,915,021	Cytoplasm	other
NHEDC2	6	23,408,864	23,453,698	Plasma Membrane	other
SLC39A8	6	24,031,738	24,113,066	Extracellular Space	transporter
DDIT4L	6	26,201,468	26,206,424	Cytoplasm	other
DNAJB14	6	26,437,809	26,475,619	unknown	enzyme
LAMTOR3	6	26,484,326	26,495,726	Cytoplasm	other
TSPAN5	6	27,734,168	27,911,471	Plasma Membrane	other
PDHA2	6	30,516,350	30,517,678	Cytoplasm	enzyme
UNC5C	6	31,048,643	31,239,605	Plasma Membrane	transmembrane receptor
BMPR1B	6	31,251,194	31,303,130	Plasma Membrane	kinase
SMARCAD1	6	32,220,382	32,308,866	Nucleus	enzyme
MMRN1	6	36,344,037	36,417,274	Extracellular Space	other
PIGY	6	37,069,642	37,072,440	Plasma Membrane	other
PPM1K	6	37,268,109	37,290,151	Cytoplasm	phosphatase
ABCG2	6	37,304,738	37,421,681	Plasma Membrane	transporter
LAP3	6	37,961,794	37,987,164	Cytoplasm	peptidase
NCAPG	6	38,153,047	38,199,149	Nucleus	other
LCORL	6	38,231,086	38,327,423	Nucleus	transcription regulator
PACRGL	6	41,397,759	41,418,385	unknown	other
DHX15	6	45,484,488	45,538,459	Nucleus	enzyme
SOD3	6	45,741,186	45,744,454	Extracellular Space	enzyme
PI4K2B	6	46,163,517	46,199,651	Cytoplasm	kinase
ANAPC4	6	46,279,716	46,312,880	Nucleus	enzyme
SLC34A2	6	46,496,138	46,521,367	Plasma Membrane	transporter
FAM114A1	6	60,403,409	60,459,346	unknown	other
WDR19	6	60,720,183	60,781,600	Extracellular Space	other
N4BP2	6	61,484,389	61,538,984	Cytoplasm	kinase
CHRNA9	6	61,702,088	61,715,473	Plasma Membrane	transmembrane receptor
APBB2	6	62,105,047	62,486,330	Cytoplasm	other
GABRG1	6	66,951,630	67,047,315	Plasma Membrane	ion channel
GABRA2	6	67,198,924	67,340,085	Plasma Membrane	ion channel
OCIAD1	6	70,775,179	70,798,072	Cytoplasm	other
SCFD2	6	71,168,458	71,566,216	unknown	transporter
FIP1L1	6	71,571,106	71,635,861	Nucleus	other
IGFBP7	6	75,049,708	75,130,502	Extracellular Space	transporter
PARM1	6	92,968,112	93,096,535	Extracellular Space	other
TAPT1	6	115,982,733	116,032,080	Plasma Membrane	G-protein coupled receptor
PDE6B	6	117,868,393	117,899,321	Cytoplasm	enzyme
ATP5I	6	117,901,248	117,902,790	Cytoplasm	transporter
IDUA	6	118,177,504	118,191,626	Cytoplasm	enzyme
FGFRL1	6	118,216,357	118,219,180	Plasma Membrane	transmembrane receptor
WHSC1	6	118,987,389	119,029,770	Nucleus	other
MXD4	6	119,341,676	119,349,171	Nucleus	transcription regulator
GRK4	6	120,029,916	120,072,154	Plasma Membrane	kinase
SORCS2	6	122,213,300	122,260,420	Plasma Membrane	transporter
ZER1	11	102,886,812	102,915,280	unknown	enzyme
LRRC8A	11	103,076,757	103,103,223	unknown	other
SH3GLB2	11	103,172,379	103,189,424	Cytoplasm	other
DOLPP1	11	103,229,242	103,238,569	Cytoplasm	enzyme
IER5L	11	103,310,300	103,311,508	unknown	other
METTL11A	11	103,701,167	103,708,828	Nucleus	enzyme
PTGES	11	103,800,389	103,811,856	Cytoplasm	enzyme
TOR1B	11	103,898,682	103,903,458	Cytoplasm	other
API5	15	73,430,189	73,456,673	Cytoplasm	other
TMEM100	19	5,189,449	5,192,339	unknown	other
PCTP	19	5,359,190	5,384,018	Cytoplasm	transporter
SCPEP1	19	7,033,105	7,063,997	Cytoplasm	peptidase
AKAP1	19	7,171,111	7,183,414	Cytoplasm	other
SLC13A2	19	19,830,979	19,855,104	Plasma Membrane	transporter

FOXN1	19	19,878,022	19,891,364	Nucleus	transcription regulator
ALDOC	19	19,919,314	19,922,828	Cytoplasm	enzyme
PROCA1	19	20,020,529	20,027,212	unknown	other
RAB34	19	20,029,283	20,030,118	Cytoplasm	enzyme
KRT35	19	43,036,132	43,186,658	Cytoplasm	other
LEPREL4	19	43,356,719	43,363,114	Nucleus	other
KLHL11	19	43,399,075	43,405,783	unknown	other
HSPB9	19	43,591,543	43,592,056	Cytoplasm	other
RAB5C	19	43,592,835	43,614,711	Cytoplasm	enzyme
HCRT	19	43,636,053	43,637,289	Extracellular Space	other
PSMC3IP	19	43,987,969	43,992,075	Nucleus	other
EZH1	19	44,114,127	44,146,413	Nucleus	enzyme
SMURF2	19	50,270,228	50,314,755	Cytoplasm	enzyme
BPTF	19	50,499,978	50,590,280	Nucleus	transcription regulator
GPS1	19	52,294,882	52,299,161	Nucleus	other
RFNG	19	52,301,004	52,303,533	Cytoplasm	enzyme
ASPSCR1	19	52,377,561	52,405,315	Cytoplasm	other
PYCR1	19	52,442,617	52,447,280	Cytoplasm	enzyme
MAFG	19	52,451,905	52,457,467	Nucleus	transcription regulator
PCYT2	19	52,467,494	52,474,648	Cytoplasm	enzyme
NPB	19	52,476,278	52,477,335	Extracellular Space	other
THOC4	19	52,485,066	52,488,850	Nucleus	transcription regulator
CCDC137	19	52,637,445	52,643,095	unknown	other
AZI1	19	52,991,344	53,004,927	Cytoplasm	other
AATK	19	53,049,375	53,064,869	Cytoplasm	kinase
RPTOR	19	53,175,424	53,363,915	Cytoplasm	other
CANT1	19	54,951,431	54,955,155	Extracellular Space	enzyme
PGS1	19	55,364,881	55,402,115	Cytoplasm	enzyme
TNRC6C	19	55,629,561	55,669,178	unknown	other

Appendix 8. Summary of positional candidate genes in potential gene network for carcass grade fat from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
R3HDM1	2	64,601,841	64,764,417	unknown	other
ACMSD	2	65,355,856	65,409,437	Cytoplasm	enzyme
SNAPC3	8	30,580,555	30,650,215	Nucleus	other
PIGO	8	62,000,491	62,006,902	Cytoplasm	enzyme
STOML2	8	62,013,177	62,016,436	Plasma Membrane	other
RUSC2	8	62,379,329	62,441,169	unknown	other
SIT1	8	62,483,210	62,484,760	Plasma Membrane	other
RECK	8	63,082,768	63,162,315	Plasma Membrane	other
GLIPR2	8	63,175,694	63,194,547	Cytoplasm	other
GRHPR	8	64,330,017	64,338,931	Cytoplasm	enzyme
UQCC	13	65,157,010	65,251,937	Cytoplasm	other
GDF5	13	65,264,319	65,267,468	Extracellular Space	growth factor
SPAG4	13	65,424,637	65,429,342	Cytoplasm	other
RBM39	13	65,491,558	65,518,140	Nucleus	transcription regulator
SCAND1	13	65,669,436	65,670,267	Nucleus	transcription regulator
C20orf4	13	65,945,722	65,965,913	unknown	other
UEVLD	29	27,407,252	27,461,742	Cytoplasm	enzyme
HPS5	29	27,632,837	27,667,895	Cytoplasm	other
OR8D4	29	27,841,740	27,843,245	Plasma Membrane	G-protein coupled receptor
VWA5A	29	28,099,374	28,113,887	unknown	other

Appendix 9. Summary of positional candidate genes in potential gene network for carcass lean meat yield from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
COL3A1	2	7,741,002	7,779,700	Extracellular Space	other
SCN9A	2	31,042,176	31,133,329	Plasma Membrane	ion channel
SCN1A	2	31,270,013	31,363,759	Plasma Membrane	ion channel
GALNT3	2	31,590,413	31,614,133	Cytoplasm	enzyme
HNMT	2	61,972,450	62,017,036	Cytoplasm	enzyme
ACMSD	2	65,355,856	65,409,437	Cytoplasm	enzyme
MGAT5	2	65,791,198	66,046,862	Cytoplasm	enzyme
ACTR3	2	68,831,148	68,852,297	Plasma Membrane	other
DPP10	2	70,927,242	71,099,935	unknown	peptidase
LAP3	6	37,961,794	37,987,164	Cytoplasm	peptidase
FAM184B	6	38,001,547	38,059,383	unknown	other
NCAPG	6	38,153,047	38,199,149	Nucleus	other
LCORL	6	38,231,086	38,327,423	Nucleus	transcription regulator
BTC	6	92,785,266	92,835,094	Extracellular Space	growth factor
PARM1	6	92,968,112	93,096,535	Extracellular Space	other
PSIP1	8	30,535,430	30,574,350	Nucleus	other
SNAPC3	8	30,580,555	30,650,215	Nucleus	other
PITRM1	13	45,247,547	45,272,018	Cytoplasm	peptidase
PFKP	13	45,272,887	45,325,518	Cytoplasm	kinase
EIF6	13	65,130,220	65,136,106	Cytoplasm	translation regulator
UQCC	13	65,157,010	65,251,937	Cytoplasm	other
GDF5	13	65,264,319	65,267,468	Extracellular Space	growth factor
SPAG4	13	65,424,637	65,429,342	Cytoplasm	other
CPNE1	13	65,432,559	65,468,475	unknown	transporter
NFS1	13	65,471,267	65,488,244	Cytoplasm	enzyme
ROMO1	13	65,488,292	65,489,823	Cytoplasm	other
RBM39	13	65,491,558	65,518,140	Nucleus	transcription regulator
PHF20	13	65,539,433	65,665,747	Nucleus	other
SCAND1	13	65,669,436	65,670,267	Nucleus	transcription regulator
EPB41L1	13	65,885,041	65,939,277	Plasma Membrane	other
C20orf4	13	65,945,722	65,965,913	unknown	other
NKX2-1	21	47,547,391	47,550,216	Nucleus	transcription regulator
NKX2-8	21	47,613,497	47,615,115	Nucleus	transcription regulator
PAX9	21	47,709,915	47,722,932	Nucleus	transcription regulator

Appendix 10. Summary of positional candidate genes in potential gene network for ultrasound backfat from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
LPL	8	70,187,161	70,214,332	Cytoplasm	enzyme
GFRA2	8	72,205,731	72,306,991	Plasma Membrane	transmembrane receptor
DOK2	8	72,427,648	72,432,271	Plasma Membrane	other
REEP4	8	72,518,587	72,522,572	unknown	other
ZNF484	8	87,963,413	87,970,747	Nucleus	other
IARS	8	88,005,770	88,087,293	Cytoplasm	enzyme
NOL8	8	88,094,398	88,119,707	Nucleus	other
OMD	8	88,206,467	88,221,378	Extracellular Space	other
IPPK	8	88,357,724	88,416,932	Cytoplasm	kinase
C9orf89	8	88,654,335	88,669,779	Cytoplasm	other
SMC2	8	98,622,768	98,667,171	Nucleus	transporter
MAP3K8	13	35,222,366	35,248,667	Cytoplasm	kinase
C6	20	35,449,496	35,525,973	Extracellular Space	other
SELS	21	3,877,736	3,886,203	Cytoplasm	other
HYAL2	22	50,927,751	50,932,890	Cytoplasm	enzyme
HYAL1	22	50,935,916	50,939,017	Cytoplasm	enzyme
HYAL3	22	50,943,394	50,945,096	Cytoplasm	enzyme
GNAI2	22	51,006,126	51,026,284	Plasma Membrane	enzyme
SLC38A3	22	51,038,046	51,052,801	Plasma Membrane	transporter
SEMA3F	22	51,068,558	51,078,300	Extracellular Space	other
MST1R	22	51,263,376	51,276,476	Plasma Membrane	kinase
UBA7	22	51,327,194	51,336,258	Cytoplasm	enzyme
MST1	22	51,416,063	51,420,964	Extracellular Space	growth factor
DAG1	22	51,521,576	51,534,748	Plasma Membrane	transmembrane receptor
RHOA	22	51,637,653	51,683,018	Cytoplasm	enzyme
GPX1	22	51,684,530	51,685,408	Cytoplasm	enzyme
CCDC71	22	51,795,111	51,798,797	Nucleus	other
USP19	22	51,829,926	51,841,292	Cytoplasm	peptidase
QARS	22	51,843,519	51,850,823	Cytoplasm	enzyme
WDR6	22	51,906,867	51,915,248	Cytoplasm	other
P4HTM	22	51,915,530	51,920,984	Cytoplasm	enzyme
EIF4EBP2	28	25,846,337	25,867,588	Cytoplasm	other
KIAA1274	28	25,959,581	26,009,613	Cytoplasm	phosphatase
PRF1	28	26,034,422	26,038,537	Cytoplasm	other
UNC5B	28	26,746,382	26,835,042	Plasma Membrane	transmembrane receptor
PSAP	28	27,353,596	27,387,074	Extracellular Space	other
DDIT4	28	27,716,022	27,718,047	Cytoplasm	other
PLAU	28	29,177,892	29,183,938	Extracellular Space	peptidase
VCL	28	29,261,195	29,371,804	Plasma Membrane	enzyme
PPIF	28	34,527,625	34,533,722	Cytoplasm	enzyme
ANXA11	28	34,678,315	34,719,037	Nucleus	other
SFTPD	28	34,956,688	34,961,597	Extracellular Space	other
SFTPA1	28	35,127,184	35,131,284	Extracellular Space	transporter
MAT1A	28	35,147,081	35,164,031	Cytoplasm	enzyme

Appendix 11. Summary of positional candidate genes in potential gene networks for ultrasound ribeye area from whole genome association study using single marker LD regression in commercial hybrid beef cattle.

Gene	BTA	Start (bp)	End (bp)	Location	Type
MYO1B	2	83,714,680	83,914,421	Cytoplasm	other
OBFC2A	2	84,190,581	84,199,369	Nucleus	other
SDPR	2	84,374,348	84,387,379	Plasma Membrane	other
BZW1	2	93,764,164	93,775,089	Cytoplasm	translation regulator
NIF3L1	2	93,826,598	93,841,566	Cytoplasm	other
ORC2	2	93,845,592	93,883,458	Nucleus	other
NDUFB3	2	93,956,131	93,966,641	Cytoplasm	enzyme
ALS2CR12	2	94,177,823	94,215,667	Cytoplasm	other
TRAK2	2	94,230,106	94,272,120	Plasma Membrane	transporter
STRADB	2	94,307,242	94,331,055	Cytoplasm	kinase
ALS2CR4	2	94,463,237	94,485,609	unknown	other
ACRBP	5	10,770,945	10,778,132	Extracellular Space	other
CPNE8	5	45,665,007	45,944,391	unknown	other
SCYL2	5	69,021,889	69,203,803	unknown	other
ACTR6	5	69,094,816	69,114,346	unknown	transporter
SLC17A8	5	69,222,568	69,261,920	Plasma Membrane	transporter
A2ML1	5	108,438,453	108,487,655	Cytoplasm	other
MFAP5	5	108,589,246	108,600,991	Extracellular Space	other
APOBEC1	5	108,681,942	108,687,553	Cytoplasm	enzyme
FOXJ2	5	108,934,285	108,952,320	Nucleus	transcription regulator
NECAP1	5	108,980,942	108,992,705	Plasma Membrane	other
CLEC4A	5	109,006,517	109,018,347	Plasma Membrane	transmembrane receptor
CLEC6A	5	109,032,003	109,048,605	Plasma Membrane	other
CLEC4E	5	109,088,137	109,098,725	unknown	other
PACRGL	6	41,397,759	41,418,385	unknown	other
KCNIP4	6	41,420,001	41,596,754	Plasma Membrane	ion channel
CEP78	8	56,548,037	56,582,897	Cytoplasm	other
ZNF484	8	87,963,413	87,970,747	Nucleus	other
OGN	8	88,188,034	88,203,623	Extracellular Space	growth factor
ASPN	8	88,236,061	88,261,403	Extracellular Space	other
ECM2	8	88,277,170	88,316,411	Extracellular Space	other
FAM107B	13	28,844,020	28,920,442	unknown	other
HSPA14	13	29,128,108	29,152,706	Cytoplasm	peptidase
SUV39H2	13	29,162,737	29,185,252	Nucleus	transcription regulator
DCLRE1C	13	29,189,765	29,223,777	Nucleus	enzyme
MEIG1	13	29,231,640	29,248,248	Nucleus	other
FAM171A1	13	29,454,243	29,644,562	unknown	other
FAM171A1	13	29,454,243	29,644,562	unknown	other
SPON1	15	37,242,121	37,557,628	Extracellular Space	other
FAR1	15	37,779,626	37,856,711	Cytoplasm	enzyme
BTBD10	15	38,093,079	38,138,651	unknown	ion channel
OXCT1	20	34,833,498	34,953,986	Cytoplasm	enzyme
WDR70	20	38,932,437	39,213,167	unknown	other
NUP155	20	39,218,706	39,268,111	Nucleus	transporter
NIPBL	20	39,411,530	39,620,673	Nucleus	transcription regulator
SLC1A3	20	39,826,428	39,907,716	Plasma Membrane	transporter
CDH18	20	56,901,150	57,306,268	Plasma Membrane	other
CHDH	22	48,177,391	48,200,809	Cytoplasm	enzyme
CACNA1D	22	48,206,364	48,549,487	Plasma Membrane	ion channel
DCP1A	22	48,689,528	48,733,764	Nucleus	other
TKT	22	48,754,843	48,778,401	Cytoplasm	enzyme
RFT1	22	48,860,495	48,898,958	unknown	other
MANF	22	50,406,244	50,410,083	Extracellular Space	other
CACNA2D2	22	50,752,556	50,890,955	Plasma Membrane	ion channel
TMEM115	22	50,895,096	50,899,834	Plasma Membrane	other
CYB561D2	22	50,900,676	50,903,590	unknown	enzyme

NPRL2	22	50,903,710	50,907,037	unknown	kinase
ZMYND10	22	50,908,522	50,912,716	Cytoplasm	other
Rassf1	22	50,916,048	50,922,052	Cytoplasm	other
HYAL2	22	50,927,751	50,932,890	Cytoplasm	enzyme
HYAL1	22	50,935,916	50,939,017	Cytoplasm	enzyme
HYAL3	22	50,943,394	50,945,096	Cytoplasm	enzyme
IFRD2	22	50,945,927	50,972,912	unknown	other
SEMA3B	22	50,990,922	50,997,263	Extracellular Space	other
RBM5	22	51,124,702	51,148,827	Nucleus	other
SFXN3	26	22,165,941	22,175,398	Cytoplasm	transporter
POLL	26	22,631,355	22,639,249	Nucleus	enzyme
DPCD	26	22,639,305	22,661,392	unknown	other
FBXW4	26	22,661,880	22,744,754	unknown	other
FGF8	26	22,808,438	22,813,158	Extracellular Space	growth factor
PPRC1	26	23,008,798	23,022,702	Extracellular Space	other
ELOVL3	26	23,074,153	23,076,897	Cytoplasm	enzyme
PITX3	26	23,078,139	23,089,423	Nucleus	transcription regulator
HPS6	26	23,166,497	23,169,118	Cytoplasm	other