University of Alberta

Application of Genomics-based Tools Leading to the Identification of Markers on Bovine Chromosome 14 Influencing Milk Production and Carcass Quality Traits

by

Elisa Ferreira Marques

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 2009
Edmonton, Alberta

Examining Committee

Dr. Stephen S. Moore, Animal Science

Dr. Zhiquan Wang, Animal Science

Dr. Walter Dixon, Animal Science

Dr. Michael Walter, Medical Genetics

Dr. Dominique Rocha, Animal Genetics, INRA-Université de Limoges

ABSTRACT

Genetic improvements in beef and dairy cattle can bring significant advances in satisfying the global food demand, which is expected to double by 2050. Several DNA markers have been identified on bovine chromosome 14 (BTA14), but low mapping resolution prevents their refinement for identification of causal mutations. The objective of this research was to apply radiation hybrid mapping technique to correctly map available high density markers, enabling the accurate assessment of linkage disequilibrium and the scanning of quantitative trait loci across the chromosome. The research also applied these techniques to identify candidate markers on BTA14 contributing to the genetic variation observed in milk production and carcass quality traits in Holstein and Angus cattle, respectively. The first study aimed at correctly ordering genetic markers along BTA14 and comparing the order to the bovine sequence assembly to aid collaborative efforts in improving the future versions of the assembly. A 12K radiation hybrid map of BTA14 was constructed using 843 single nucleotide polymorphism markers. The second study assessed the extent of linkage disequilibrium along the chromosome identifying specific regions in both Angus and Holstein cattle where non-random association between alleles of different loci occurred. For both breeds, results showed that average linkage disequilibrium extends to moderate levels up to 100 kilo base pairs and falls to background levels after 500 kilo base pairs. Correlation analysis for marker pairs common to these two breeds confirmed that the same marker phase is maintained only up to distances of 10 kilo base pairs. Linkage analysis studies for both breeds identified

markers on the basis of sire heterozygosity and linkage disequilibrium and reported quantitative trait loci affecting milk production and carcass quality traits. Finally, using marker function, association and linkage analysis results, several candidate markers demonstrating significant effects on these economically relevant traits were identified. The results from this study support the existence of considerable genetic variation for both milk production and carcass quality traits in Holstein and Angus cattle, respectively, demonstrating opportunities for genetic improvement.

ACKNOWLEGMENTS

This project and my success would not have been possible without my doctoral supervisor Dr. Stephen Moore, who not only gave me the opportunity to join his research group, but also trusted me to carry out this project. I would also like to acknowledge my committee members Dr. Zhiquan Wang and Dr. Walter Dixon for their comments and advice.

I want to acknowledge the Bovine Genomics Group harboring top notch personnel who have contributed to my success in this project and provided valuable comments and assistance throughout these years, especially: Dr. Graham Plastow, Dr. Paul Stothard, Dr. Jason Grant, Dr. Changxi Li, Yang Meng, Dr. Leluo Guan, Dr. Urmila Basu, Anna Szul, Jackson Mah, Fernando Soares, Brenda Murdoch and Xiu Zhaozi.

I must also acknowledge all fellow graduate students whom I have had the pleasure to work with, some of which have already left and continue to succeed in their endeavors: Dr Donald Nkrumah and Dr. Laura Sherman. They were both outstanding colleagues who shared and continue to share their knowledge in the field. I would also like to acknowledge the University of Alberta administrative personnel specially Jody Forslund and everyone who I had the pleasure to work with in the various committees and University events.

I would like to dedicate my thesis to my parents, Luiz Fernando Aarao Marques and Liliane Ferreira Marques and my siblings, Daniel Ferreira Marques and Laila Ferreira Marques who have supported me throughout my whole life. To my partner, Brian Howland, who provided his moral support throughout my graduate degree and for reassuring me the value of my skills, and to all my friends, specifically Daniela and Anna Villota for their support throughout these years. I also would like to thank the Department of Agricultural, Food and Nutritional Sciences, the University of Alberta and the Graduate Students' Association for their financial support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABCG2	ATP binding cassete, sub-family G, member 2 gene
ACTH	Adrenocorticotropic hormone
ADG	Average daily gain
ARMC1	Armadillo repeat containing 1 gene
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BES	Bacterial artificial chromosome end sequence
BF	Backfat
BFF	Weight of body fat trim
BLAST	Basic Local Alignment Search Tool
BTA	*Bos taurus* autosome
BWm_ABC	Birth weight maternal ABC
CBF	Carcass backfat
CBFA2T1	Core-binding factor, runt domain, alpha subunit 2; translocated to, 1;
CHORI-240	Children's Hospital Oakland Research Institute Bovine BAC library
cM	Centimorgan
CMAR	Carcass marbling
CMD	Congenital muscular dystonia gene
CRH	Corticotropin releasing hormone
CYP11B1	Cytochrome P450, family 11, subfamily B, polypeptide 1 gene

CYP7B1	Cytochrome P450, family 7, subfamily B, polypeptide 1 gene
DECR1	2,4−dienoyl CoA reductase 1, mitochondrial
DGAT1	Diacylglycerol acyltransferase1
DMI	Dry matter intake
DNA	Deoxyribonucleic acid
EHH	Extended haplotype homozygosity
Fat1	Fat depth (min fat in first quadrant, mm)
Fat2	Fat depth (min fat in second quadrant, mm)
Fat3	Fat depth (min fat in third quadrant, mm)
FDR	False discovery rate
FG	Feed / gain
FGF8	Fibroblast growth factor 8
GRFAT	Carcass gradefat
HCW	Hot carcass weight
HSA	*Homo sapiens* autosome
HSB	Homologous synteny block
IBISS	Interactive Bovine In Silico SNP
KCNB2	Potassium voltage-gated channel, shab-related subfamily, member 2 gene
LD	Linkage disequilibrium
LKH	Lin-Kernighan heuristic
LM7D	Shear force of longissimus dorsi aged 7 days
LMA	Longissimus muscle area

LMY	Lean meat yield
LPL	Lipoprotein lipase
MAF	Minor allele frequency
MARC	Meat Animal Research Center
MCM4	Minichromosome maintenance complex component 4
MGC138052	Mitogen-activated protein kinase 15 gene
miRNAs	Micro RNAs
MMWT	Metabolic mid test weight
mRNA	Messenger RNA
NCBI	National Center for Biotechnology Information
OPN	Osteopontin gene
PCR	Polymerase chain reaction
QTL	Quantitative trait loci
QTN	Quantitative trait nucleotide
RFLP	Restriction fragment length polymorphism
RGS22	Regulator of G-protein signaling 22
RH	Radiation hybrid mapping
RW	Rib weight
SF	Subcutaneous fat trim
SNP	Single nucleotide polymorphism
TRIM55	Tripartite motif-containing 55
TSP	Traveling salesman problem
UBF	Ultrasound backfat

UGL	Estimated lean meat yield
ULMA	Ultrasound longissimus muslce area
UMAR	Ultrasound marbling
USDA	United States Department of Agriculture
WADG	Average daily gain from birth to weaning
WG	ABC for average daily gain from birth to weaning
WGm_ABC	Weaning gain maternal ABC
WWT	Weaning weight
YGRADE	Yield Grade

CHAPTER 1

General Introduction

1.1 Introduction

Breeders are faced with the challenge of using diverse resources to produce cattle that are profitable to all segments of the industry and meat products that are in demand by consumers. Health issues concerning consumers drive the increase in product quality standards. In some breeding programs, selection is directed against backfat thickness, because reduced carcass fatness benefits carcass quality and production. In dairy systems, there is an increased importance of protein concentration on milk pricing, with genetic manipulation being one of the strategies for increase in milk protein concentration. Milk production traits such as milk volume, fat content (%) and protein content (%) are also among the most highly important traits for dairy producers. Milk fat and protein content contribute to the quality of dairy products, with milk protein directly affecting cheese yield. These traits are quantitative in nature, that is, their phenotypic expression is genetically and environmentally determined. Moreover, the genetic aspect encompasses more than one gene, making the identification of all the genetic variation very challenging.

Quantitative Trait Loci (**QTL**) are regions on a chromosome where gene(s) affecting a quantitative trait exist. Several QTL affecting milk production traits have been reported for BTA14 (Georges *et al.* 1995; Coppieters *et al.* 1998; Ron *et al.* 1998; Zhang *et al.* 1998). In beef cattle, a number of studies have

shown BTA14 to harbor QTL for fat deposition traits (Stone *et al.* 1999; Casas *et al.* 2003; Moore *et al.* 2003) and longissimus muscle area (Stone *et al.* 1999). In 1994, Andersson *et al.* (1994) suggested that much of the genetic variation in quantitative traits was a result of the interaction of the alleles of a few genes with the environment. Those genes could either be single genes with major effects (Boehnke & Moll 1989) or be a small cluster of genes producing a major effect (de Vries *et al.* 1996). To date, only a few genes with conclusive effects have been identified (Van Laere *et al.* 2003; Grisart *et al.* 2004; Cohen-Zinder *et al.* 2005; Clop *et al.* 2006). The lack of more successful cases is perhaps due to the existence of several genes with an effect, as well as gene-gene interactions.

The identification of markers associated with traits of interest becomes possible with the growing availability of genome wide single nucleotide polymorphisms (**SNP**) and their abundance in cattle makes them very attractive markers. Several public databases such as the Interactive Bovine In Silico SNP (IBISS) database, the National Center for Biotechnology Information (NCBI) and the Bovine Sequencing Initiative offer hundreds of thousands of SNPs to researchers interested in the genetic variation underlying economically relevant traits in cattle. These SNPs can be selected to represent regions in chromosomes where known QTL exist.

Another method of increasing the current knowledge of meat or milk production genes is through comparative genomics. Knowledge of genes affecting lipid metabolism in other species can be used to find similar genes in other less studied species. This positional candidate approach especially in farm animals

takes advantage of comparative information and it has become more useful with the completion of the human map and the creation of gene functions and gene expression patterns databases (Andersson *et al.* 1994). Human and bovine sequence alignment information has reported that bovine chromosome 14 contains regions homologous to human chromosome 8 (Gallagher & Womack 1992). Genes mapped on human chromosome 8 known to affect lipid metabolism are considered start points for candidate genes affecting, for instance, marbling traits on bovine chromosome 14.

It is undeniable the amount of improvement that can be obtained through the genetic knowledge underlying both meat and dairy traits. In the past 40 years, dairy systems have seen an almost 100% increase in milk yield (Georges 2007). Linkage disequilibrium studies on human chromosomes (Dawson *et al.* 2002; De La Vega *et al.* 2005) have given great insight for such studies in livestock (Farnir *et al.* 2000; Andersson 2001). Understanding the patterns of linkage disequilibrium in humans (Dawson *et al.* 2002; Cardon & Abecasis 2003; De La Vega *et al.* 2005) can bring astonishing results to help unravel and predict the variations that underlie complex traits in cattle. Together, knowledge of marker-marker relationship (i.e linkage disequilibrium) coupled to linkage analysis and function of a gene on lipid metabolism can result in the fine mapping and identification of the causative mutations spanning the QTL regions affecting meat and milk production traits. In addition, it is anticipated to contribute towards enhancing the understanding of the genetic, biochemical and physiological pathways that regulate mammalian adipogenesis. The characterization of the

genetics influencing the phenotypic variation in meat quality and milk production traits will provide a means for efficient livestock management, greatly accelerating the rate of genetic progress. Besides this, the verification and validation of associations of SNP markers in different cattle populations will build on the development and commercialization of DNA marker tests.

1.2 Research Hypotheses

Several markers have been identified on bovine chromosome 14, but difficulty in mapping resolution has prevented the refinement of their order for application of future genetic work necessary for identification of candidate markers. The objective of this research was to apply radiation hybrid mapping technique to correctly map these available markers, enabling the accurate assessment of linkage disequilibrium and the scanning of quantitative trait loci across the chromosome. The research also aimed at applying these techniques to identify candidate markers on bovine chromosome 14 contributing to the genetic variation observed in milk production and meat quality traits in Holstein and Angus cattle. We hypothesize that there is considerable genetic variation in milk production and carcass quality traits in Holstein and Angus, respectively, which can be identified by quantitative trait loci and by genetic marker associations.

1.3 Literature Cited

Andersson L. (2001) Genetic dissection of phenotypic diversity in farm animals. *Nat Rev Genet* **2**, 130-8.

Andersson L., Haley C.S., Ellegren H., Knott S.A., Johansson M., Andersson K., Andersson-Eklund L., Edfors-Lilja I., Fredholm M., Hansson I. & et al. (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**, 1771-4.

Boehnke M. & Moll P.P. (1989) Identifying pedigrees segregating at a major locus for a quantitative trait: an efficient strategy for linkage analysis. *Am J Hum Genet* **44**, 216-24.

Cardon L.R. & Abecasis G.R. (2003) Using haplotype blocks to map human complex trait loci. *Trends Genet* **19**, 135-40.

Casas E., Shackelford S.D., Keele J.W., Koohmaraie M., Smith T.P. & Stone R.T. (2003) Detection of quantitative trait loci for growth and carcass composition in cattle. *J Anim Sci* **81**, 2976-83.

Clop A., Marcq F., Takeda H., Pirottin D., Tordoir X., Bibe B., Bouix J., Caiment F., Elsen J.M., Eychenne F., Larzul C., Laville E., Meish F., Milenkovic D., Tobin J., Charlier C. & Georges M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet* **38**, 813-8.

Cohen-Zinder M., Seroussi E., Larkin D.M., Loor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I. & Ron M. (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 Res* **15**, 936-44.

Coppieters W., Riquet J., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Nezer C., Simon P., Vanmanshoven P., Wagenaar D. & Georges M. (1998) A QTL with major effect on milk yield and composition maps to bovine chromosome 14. *Mamm Genome* **9**, 540-4.

Dawson E., Abecasis G.R., Bumpstead S., Chen Y., Hunt S., Beare D.M., Pabial J., Dibling T., Tinsley E., Kirby S., Carter D., Papaspyridonos M., Livingstone S., Ganske R., Lohmussaar E., Zernant J., Tonisson N., Remm M., Magi R., Puurand T., Vilo J., Kurg A., Rice K., Deloukas P., Mott R., Metspalu A., Bentley D.R., Cardon L.R. & Dunham I. (2002) A first-generation linkage disequilibrium map of human chromosome 22. *Nature* **418**, 544-8.

De La Vega F.M., Isaac H., Collins A., Scafe C.R., Halldorsson B.V., Su X., Lippert R.A., Wang Y., Laig-Webster M., Koehler R.T., Ziegle J.S., Wogan L.T., Stevens J.F., Leinen K.M., Olson S.J., Guegler K.J., You X., Xu L.H., Hemken H.G., Kalush F., Itakura M., Zheng Y., de The G., O'Brien S.J., Clark A.G., Istrail S., Hunkapiller M.W., Spier E.G. & Gilbert D.A. (2005) The linkage disequilibrium maps of three human chromosomes across four populations reflect their demographic history and a common underlying recombination pattern. *Genome Res* **15**, 454-62.

de Vries H.G., van der Meulen M.A., Rozen R., Halley D.J., Scheffer H., ten Kate L.P., Buys C.H. & te Meerman G.J. (1996) Haplotype identity between individuals who share a CFTR mutation allele "identical by descent": demonstration of the usefulness of the haplotype-sharing concept for gene mapping in real populations. *Hum Genet* **98**, 304-9.

Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D. & Georges M. (2000) Extensive genome-wide linkage disequilibrium in cattle. *Genome Res* **10**, 220-7.

Gallagher D.S., Jr. & Womack J.E. (1992) Chromosome conservation in the Bovidae. *J Hered* **83**, 287-98.

Georges M. (2007) Mapping, fine mapping, and molecular dissection of quantitative trait Loci in domestic animals. *Annu Rev Genomics Hum Genet* **8**, 131-62.

Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., Pasquino A.T., Sargeant L.S., Sorensen A., Steele M.R., Zhao X. & et al. (1995) Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* **139**, 907-20.

Grisart B., Farnir F., Karim L., Cambisano N., Kim J.J., Kvasz A., Mni M., Simon P., Frere J.M., Coppieters W. & Georges M. (2004) Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc Natl Acad Sci U S A* **101**, 2398-403.

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C. & Benkel B. (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. *J Anim Sci* **81**, 1919-25.

Ron M., Heyen D.W., Weller J.I., Band M., Feldmesser E., Pasternak H., Da Y., Wiggans G.R., Vanraden P., Ezra E. & Lewin H.A. (1998) *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, 6WCCALP Congress, Armidale, Australia*, 422-5.

Stone R.T., Keele J.W., Shackelford S.D., Kappes S.M. & Koohmaraie M. (1999) A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. *J Anim Sci* **77**, 1379-84.

Van Laere A.S., Nguyen M., Braunschweig M., Nezer C., Collette C., Moreau L., Archibald A.L., Haley C.S., Buys N., Tally M., Andersson G., Georges M. & Andersson L. (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **425**, 832-6.

Zhang Q., Boichard D., Hoeschele I., Ernst C., Eggen A., Murkve B., Pfister-Genskow M., Witte L.A., Grignola F.E., Uimari P., Thaller G. & Bishop M.D. (1998) Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics* **149**, 1959-73.

CHAPTER 2

Literature Review

2.1 Chromosomal Maps

There are several types of maps used to identify the order of genes or other markers within the chromosome. The oldest of all, the genetic map, uses meiotic information to infer the location of the genes within the chromosomes. Genetic linkage mapping takes advantage of heterozygosity of molecular markers among parents. The most hypervariable of the polymorphisms are the microsatellites (repeat unit of 1-5 base pairs) and the minisatellite (repeat unit of tens of base pairs). Because of their hypervariable feature, they have been widely used for building genetic linkage maps. However, compared to microsatellites and minisatellites, single nucleotide polymorphisms (SNPs) are the most abundant (Steele & Georges 1991) types of polymorphism, and with the decrease in genotyping costs associated with them, more cattle genetic maps are incorporating these single point mutations (Ihara *et al.* 2004).

Genetic maps can be very useful only when the makers used are informative, meaning, there is a high enough number of recombinants. The first cattle genetic maps only included a few hundred markers (Barendse *et al.* 1994; Bishop *et al.* 1994). Several chromosomes were unrepresented and it was clear that many more markers were needed. Over the next several years, a number of independent studies had included an additional 1600 loci to the later versions of the cattle linkage map (Ma *et al.* 1996; Barendse *et al.* 1997; Kappes *et al.* 1997). Comparison of these maps yielded that a number of gaps still needed to be filled.

These gaps showed a low density of markers. They eventually became the stepping stone for other maps able to handle polymorphic and non-polymorphic markers.

The objective of physical mapping lies in the identification of a set of overlapping cloned fragments. These fragments could comprise a whole chromosome or genome. When clones have been ordered, they now represent libraries of DNA sequences which can then be used in various genetic analyses. Cloned DNA fragments are first generated by breaking a number of identical chromosomes into fragments (inserts) which are then joined to other DNA molecules (vectors). The resulting vector with a foreign piece of DNA is incorporated into organisms such as yeast or Escherichia coli, creating either an yeast artificial chromosomes (YACs) or a bacterial artificial chromosomes (BACs). Even though this approach can be useful for localizing markers to a small region, it still presents challenges in areas of the chromosomes that are difficult to clone (i.e near the centromere) due to the presence of repetitive sequences.

2.1.1 Radiation Hybrid Mapping

Other mapping methods such as radiation hybrid maps (Womack *et al.* 1997; Rexroad *et al.* 2000; Williams *et al.* 2002; Itoh *et al.* 2005) have enabled the community to create high resolution maps because they work with both polymorphic and non-polymorphic markers. Radiation hybrid mapping technique was first proposed in 1975 by Goss and Harris (1975) as a new method for mapping genes on human chromosomes. They subjected human chromosomes to

large doses of radiation and then fused them with hamster cells. In order to determine the linear order of groups of genes and to estimate the distance between them, the frequency with which pairs of linked genes were co-transferred after irradiation was measured. This method was resurrected and systematically used as a human gene mapping tool by 1990 by Cox and associates (Cox *et al.* 1990). It was also employed for mapping bovine genes, when the first bovine whole genome radiation hybrid panel was constructed in 1997 (Womack *et al*., 1997). In this case, X-ray treated bovine cells were fused to rodent cells forming a panel of different hybrids, each containing a unique representation of cattle chromosome fragments. Bovine cells carried the thymidine kinase gene (TK^+), whereas hamster cells lacked the thymidine kinase gene (TK^-). Once all cells were plated onto a HAT medium (hypoxanthine, aminopterin and thymidine), the hybrid (fused) cells were the only one which survived in the presence of this selective medium (Womack *et al.* 1997). Closely linked markers would be incorporated at high frequencies, because of a low probability of an X-irradiation break to occur between closely linked loci. So, the typing of these closely linked markers for either presence or absence in a particular fragment will look very similar, meaning that, nearby loci tend to show similar retention patterns.

One of the disadvantages of radiation hybrid mapping is that the denser the marker map gets, the harder it is to correctly estimate the inter-marker distance. When markers are placed not every Mb, but several markers per Mb, there is no way to reliably estimate their inter-marker order and distance. These distances can only be accurately determined by the ultimate whole genome

sequence map. The ultimate map is still a whole genome sequencing map. The first draft of the bovine sequence assembly became publicly available in 2004; and since its first draft, several other drafts have been released. However, a number of studies have cast doubt on some of the sequence order produced by these assembly versions and these versions need more scrutiny before the final assembly can be trusted (Everts-van der Wind *et al.* 2005; Jann *et al.* 2006; McKay *et al.* 2007a).

2.2 Quantitative Trait Loci (QTL) Mapping

Quantitative trait loci (QTL) mapping is recognized as the first step in identifying the gene or genes affecting any multigenic trait of interest. This analysis is also known as linkage analysis and it can be thought of as the process of determining the approximate chromosomal location of a gene.

Most QTL analysis results to date have low resolution, with intervals that extend several centimorgans (Sonstegard *et al.* 2001). The success of QTL mapping lies in not only the phenotypic and genotypic information, but also the family size and informativeness of the families. The frequency of informative sibships is also increased when the level of informativeness of markers increases. According to Lynch and Walsh (1998), in outbred populations such as livestock, only a fraction of the parents are informative, in contrast to inbred-line crosses. In order for the marker-trait association to be picked up, the parent must be heterozygous at both the QTL and the linked marker (Lynch and Walsh 1998). Lynch and Walsh (1998) also examined the trade off between increasing the number of families versus increasing the family size. According to their study,

increasing the number of sibs per family is more efficient than increasing the number of families. They explain that the number of families should be relatively large to ensure that at least one will be informative and more importantly large numbers of animals in each family to have power within the informative family.

Since the first dairy QTL mapping study (Georges *et al.* 1995), several other studies have either confirmed or identified major genes responsible for the QTL peaks detected in cattle populations. Grisart *et al.* (2002 and Winter *et al.* (2002) independently identified that a lysine to alanine substitution in diacyl glycerol acyltransferase 1 gene (*DGAT1*) was the causative mutation affecting milk fat percentage in Holstein cattle. Later, Grisart *et al.* (2004) confirmed through QTL cloning that in fact the K232A mutation was the causative mutation. Another QTL peak observed on BTA6 had several genes as candidates for harboring the causative mutation: *OPN* (Schnabel *et al.* 2005) and *ABCG2* (Cohen-Zinder *et al.* 2005; Olsen *et al.* 2007), with *ABCG2* being conclusively linked to the detected QTL peak (Cohen-Zinder et al, 2005). According to Georges (2007), in most cases the identified QTL only explains a small portion of the genetic variance. These estimates are in most cases overestimated as he mentions, due to the Beavis effect or the winner's curse (Georges *et al.* 1995; Beavis 1998); demonstrating that much of the genetic variation is yet to be identified.

The use of several markers to identify and estimate the number of QTL in a chromosome raises the issue of multiple testing. For example, if a significance threshold of 5% is applied, after 100 points along the chromosome were analyzed,

one would expect 5 (100*0.05) false positives. This problem can be circumvented by applying statistical techniques such as permutations (Churchill and Doerge 1994) and False Discovery Rates (FDR) (Benjamini & Hochberg 1995). Churchill and Doerge (1994) explained that permutation testing is performed by first creating data sets by randomly shuffling the phenotypes across the genotypes, removing any relationship between genotype and phenotype. FDR on the other hand extracts the information from the distribution of p-values over all performed tests.

2.3 Linkage Disequilibrium (LD)

Linkage disequilibrium (LD) is the non-random association between alleles at two or more loci, in other words, alleles from different loci are not segregating independently. Traditionally, linkage analysis has been used as an important tool to find the gene(s) responsible for such phenotypic variations. As traits under study become more complex, linkage analysis has become limited (Talbot *et al.* 1999). The difficulties in obtaining large and informative samples contribute to less accurate estimates of the location of genes underlying certain traits when using linkage analysis. One distinct difference between linkage analysis and linkage disequilibrium mapping is that the former evaluates LD within families, while the latter looks at LD between the marker and the QTL for the entire population (Hayes 2007).

One measure of LD is r^2, which was proposed in 1968 by Hill and Robertson (1968). This measure is described by:

$$r^2 = \frac{(freq(A1\_B1) * freq(A2\_B2) - freq(A1\_B2) * freq(A2\_B1))^2}{freq(A1) * freq(A2) * freq(B1) * freq(B2)},$$

Where freq (A1_B1) is the frequency of the A1_B1 haplotype in the population is the frequency of A1 in the population and, likewise for the other alleles.

Compared to D (Hill 1981), r^2 is less dependent of allele frequency and therefore it is more suitable to compare the extent of LD among pairs of loci. D is defined by:

$$D = freq(A1\_B1) * freq(A2\_B2) - freq(A1\_B2) * freq(A2\_B1)$$

D' is another measure of LD proposed by Lewontin (1964). As explained by McRae *et al.* (2002), this measure is affected by small sample sizes, hence the preference of using r^2 over D', which is described by:

$$D' = \frac{|D|}{D\max},$$

If D>0: Dmax = min[freq(A1)*freq(B2), -1*freq(A2)*freq(B1)],

if D<0: Dmax = max[-freq(A1)*freq(B1), -1*freq(A2)*freq(B2)].

Successes in linkage disequilibrium-based mapping of Mendelian disorders (Hastbabka et al. 1994), have led investigators to use the same

procedure in search of loci underlying complex traits in cattle (Grisart *et al.* 2002). The growing availability of genome-wide molecular markers such as SNPs and microsatellites have provided the possibility of applying association studies to genomes, therefore uncovering the variations underlying the quantitative traits.

Because of its multiple alleles, one microsatellite usually provides more information for linkage analysis than does one SNP, but the situation is more complex for LD (Jorde 2000). Genetic estimation and comparison among populations require many more SNPs relative to microsatellites. This is because microsatellites have more alleles (~5-20) versus two for SNPs (Mariette *et al.* 2002). Morin *et al.* (2004) suggests that two to six times more SNPs will be needed in order to reach the same resolution as microsatellite loci. On the other hand, the high mutation rate of microsatellites which generally decreases LD more rapidly might yield an unreliable understanding of the true variation underlying the trait under study, as well as give inaccurate divergent times and gene flow among populations (Kalinowski 2002). Alternatively, SNPs make attractive markers because they are abundant in cattle (Heaton *et al.* 2001) and relatively stable in mammals (Thomson 2001).

Dekkers (2004) suggests that due to the extensive genome-wide LD observed in livestock populations, it is possible that informative markers every 1 or 2 cM might be sufficient to detect most QTLs. In haplotype block detection, Gabriel *et al.* (2002) demonstrated that a density of one marker per 7.8 Kb was sufficient to reflect 51 haplotype block patterns throughout the human genome.

LD can be a result of migration, mutation, selection, small finite population size or other genetic events which the population experiences (Lander and Schork 1994). Finite population size is generally implicated as the major cause of LD in livestock populations, since effective population sizes for most such populations are relatively small (Andersson 2001). LD due to migration is significant when crossing inbred lines but small when crossing breeds that do not differ as markedly in gene frequencies (Stephens et al., 1994). Selection is an important cause of LD (Bulmer 1971), and it will preferentially generate disequilibrium between loci influencing the selected phenotype (Farnir *et al.* 2000).

Assessing the extent of LD across the genome can provide insights into the causes of LD. More specifically, calculation of chromosome segment homozygosity (CSH) as detailed by Hayes *et al.* (2003), a multi-locus type of LD assessment, can be useful in estimating the effective population size (Tenesa *et al.* 2007). In livestock, LD comparison among breeds can also uncover relationship between breeds, where LD decline is expected to be similar for similar breeds, for example, Zenger *et al.* (2007) showed that both the Dutch and Australian Holstein populations show similar LD decline, congruent with their relatedness.

According to Farnir *et al.* (2000), selection is not the major contributor of genome-wide LD levels found in cattle, because of the uniform distribution of LD across the genome and not only at the QTL undergoing selection. Alternatively, migration is likely to have influenced the levels of LD observed. For example, some breeds of African cattle are the products of a progressive, male-driven

admixture between the indigenous taurine breeds (*Bos taurus*) and more recently introduced zebu cattle (*Bos indicus*) (Brudford et al., 2003).

Whole genome LD studies continue to emerge, as the number of polymorphisms available increases and the costs associated with genotyping them decreases. McKay *et al.* (2007b) assessed the extent of LD across both Holstein and Angus breeds using a total of approximately 3,000 markers genome wide. The measure of LD used was r^2 which has previously been shown to be more robust than D' (McRae *et al.* 2002). The results showed that a total of 30,000 markers would be ideal to capture the genetic variation in future whole genome association analysis.

2.4 Association Analysis

In general terms, an association exists between any two characteristics if they occur more often than would be expected by chance in any individual in a population. Association does not immediately indicate causality, since many markers can be in linkage disequilibrium and in this case their effects are confounded.

Association analysis is a simple regression analysis where a particular phenotype is regressed on the genotype of interest. If there are other factors that could be influencing the association between an allele and the phenotype, those factors need to be modeled in the analysis. For example, there are often breed and contemporary group differences between animals tested, which could cause confounding effects with the genetic factors that are being tested. This

information needs to be included in the statistical analysis in order to prevent incorrect conclusions.

QTL mapping tests for the presence of linkage between traits and loci. The next logical step after QTL mapping is association analysis. After attempting to narrow down the region of a QTL, markers underlying these regions should be tested for associations with the traits. This is also known as the positional candidate gene approach, where a gene with known function under a QTL is selected to be further analyzed. The first success story in candidate gene association and later cloning in dairy cattle is the *DGAT1* story (Grisart et al., 2004). *DGAT1* is from the family of enzymes that catalyzes the last step in triacylglycerol biosynthesis. Grisart *et al.* (2004) successfully reported, through QTL cloning, that a non-conservative lysine to alanine substitution (K232A) in this gene was the quantitative trait nucleotide (QTN) affecting milk fat composition. This work followed Riquet *et al.* (1999) who successfully fine mapped a QTL in the proximal region of BTA14 to 0.5 cM. Since then, work on other candidate genes have shown strong support for being quantitative trait nucleotides (QTN): *ABCG2* Y581S (Cohen-Zinder *et al.* 2005) influencing milk composition in cattle, *IGF2* intron 3-3072 (G-A) (Van Laere *et al.* 2003) affecting muscle mass in pigs and *MSTN* 3' untranslated region (UTR) g+6723 (G-A) (Clop *et al.* 2006) affecting muscle mass in sheep.

For other populations, for example beef cattle, a few polymorphisms in candidate genes have shown association with several of the meat quality traits. Barendse (1999) reported that a polymorphism in the thyroglobulin gene was

associated with increased marbling in beef cattle, but reports were contradictory when other populations were used to assess this polymorphism (Moore *et al.* 2003; Casas *et al.* 2007). Other associations continue to arise in different cattle populations, but need scrutiny before they can be considered quantitative trait nucleotide (QTN) (Barendse; White *et al.* 2005; Michal *et al.* 2006; Zhang *et al.* 2008).

The genomics-based tools reviewed here are applied in the following chapters described in this thesis. Overall, the creation of the radiation hybrid map of the bovine chromosome 14 aided in the determining the location of the markers along the chromosome. Once the location was determined, linkage disequilibrium was assessed for all marker-pairs for both Holstein and Angus cattle. Together, location of the markers and linkage disequilibrium maps enabled the identification of quantitative trait loci. Once the regions were established, further characterization took place providing the means for the identification of markers affecting milk production and meat quality in Holstein and Angus cattle, respectively.

2.5 Literature Cited

Andersson L. (2001) Genetic dissection of phenotypic diversity in farm animals. *Nat Rev Genet* **2**, 130-8.

Barendse W. DNA markers for meat tenderness. *WO02064820.2002*, pp. 1–88.

Barendse W., Armitage S.M., Kossarek L.M., Shalom A., Kirkpatrick B.W., Ryan A.M., Clayton D., Li L., Neibergs H.L., Zhang N. & et al. (1994) A genetic linkage map of the bovine genome. *Nat Genet* **6**, 227-35.

Barendse W., Vaiman D., Kemp S.J., Sugimoto Y., Armitage S.M., Williams J.L., Sun H.S., Eggen A., Agaba M., Aleyasin S.A., Band M., Bishop M.D., Buitkamp J., Byrne K., Collins F., Cooper L., Coppettiers W., Denys B., Drinkwater R.D., Easterday K., Elduque C., Ennis S., Erhardt G., Li L. & et al. (1997) A medium-density genetic linkage map of the bovine genome. *Mamm Genome* **8**, 21-8.

Barendse W.J. (1999) Assessing lipid metabolism. *Patent, International Publication Number: WO 99/23248. World International Property Organization.*

Beavis W.D. (1998) QTL Analysis: power, precision and accuracy. *In Molecular Dissection of Complex traitsed, ed AH Patterson*, pp. 145-62.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B-Methodological*, 57:289-300.

Bishop M.D., Kappes S.M., Keele J.W., Stone R.T., Sunden S.L., Hawkins G.A., Toldo S.S., Fries R., Grosz M.D., Yoo J. & et al. (1994) A genetic linkage map for cattle. *Genetics* **136**, 619-39.

Bulmer M.G. (1971) The effect of selection on genetic variability. *Am. Nat* **105**, 201-11.

Casas E., White S.N., Shackelford S.D., Wheeler T.L., Koohmaraie M., Bennett G.L. & Smith T.P. (2007) Assessing the association of single nucleotide polymorphisms at the thyroglobulin gene with carcass traits in beef cattle. *J Anim Sci* **85**, 2807-14.

Churchill G.A. & Doerge R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963-71.

Clop A., Marcq F., Takeda H., Pirottin D., Tordoir X., Bibe B., Bouix J., Caiment F., Elsen J.M., Eychenne F., Larzul C., Laville E., Meish F., Milenkovic D., Tobin J., Charlier C. & Georges M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet* **38**, 813-8.

Cohen-Zinder M., Seroussi E., Larkin D.M., Loor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I. & Ron M. (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 Res* **15**, 936-44.

Cox D.R., Burmeister M., Price E.R., Kim S. & Myers R.M. (1990) Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250**, 245-50.

Dekkers J.C. (2004) Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *J Anim Sci* **82 E-Suppl**, E313-28.

Everts-van der Wind A., Larkin D.M., Green C.A., Elliott J.S., Olmstead C.A., Chiu R., Schein J.E., Marra M.A., Womack J.E. & Lewin H.A. (2005) A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. *Proc Natl Acad Sci U S A* **102**, 18526-31.

Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D. & Georges M. (2000) Extensive genome-wide linkage disequilibrium in cattle. *Genome Res* **10**, 220-7.

Gabriel S.B., Schaffner S.F., Nguyen H., Moore J.M., Roy J., Blumenstiel B., Higgins J., DeFelice M., Lochner A., Faggart M., Liu-Cordero S.N., Rotimi C., Adeyemo A., Cooper R., Ward R., Lander E.S., Daly M.J. & Altshuler D. (2002) The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9.

Georges M. (2007) Mapping, fine mapping, and molecular dissection of quantitative trait Loci in domestic animals. *Annu Rev Genomics Hum Genet* **8**, 131-62.

Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., Pasquino A.T., Sargeant L.S., Sorensen A., Steele M.R., Zhao X. & et al. (1995) Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* **139**, 907-20.

Goss S.J. & Harris H. (1975) New method for mapping genes in human chromosomes. *Nature* **255**, 680-4.

Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M. & Snell R. (2002) Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res* **12**, 222-31.

Grisart B., Farnir F., Karim L., Cambisano N., Kim J.J., Kvasz A., Mni M., Simon P., Frere J.M., Coppieters W. & Georges M. (2004) Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc Natl Acad Sci U S A* **101**, 2398-403.

Hayes B. (2007) QTL Mapping, MAS and Genomic Selection. *Animal Breeding & Genetics Short Courses (Iowa State University)*, [http://www.ans.iastate.edu/stud/courses/short/].

Hayes B.J., Visscher P.M., McPartlan H.C. & Goddard M.E. (2003) Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res* **13**, 635-43.

Heaton M.P., Chitko-McKnown C.G., Grosse W.M., Keele J.W., Keen J.E. & Laegreid W.W. (2001) Interleukin-8 haplotype structure from nucleotide sequence variation in commercial populations of U.S. beef cattle. *Mamm Genome* **12**, 219-26.

Hill W.G. (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* **38**, 209 - 16.

Hill W.G. & Robertson A. (1968) The effects of inbreeding at loci with heterozygote advantage. *Genetics* **60**, 615-28.

Ihara N., Takasuga A., Mizoshita K., Takeda H., Sugimoto M., Mizoguchi Y., Hirano T., Itoh T., Watanabe T., Reed K.M., Snelling W.M., Kappes S.M., Beattie C.W., Bennett G.L. & Sugimoto Y. (2004) A comprehensive genetic map of the cattle genome based on 3802 microsatellites. *Genome Res* **14**, 1987-98.

Itoh T., Watanabe T., Ihara N., Mariani P., Beattie C.W., Sugimoto Y. & Takasuga A. (2005) A comprehensive radiation hybrid map of the bovine genome comprising 5593 loci. *Genomics* **85**, 413-24.

Jann O.C., Aerts J., Jones M., Hastings N., Law A., McKay S., Marques E., Prasad A., Yu J., Moore S.S., Floriot S., Mahe M.F., Eggen A., Silveri L., Negrini R., Milanesi E., Ajmone-Marsan P., Valentini A., Marchitelli C., Savarese M.C., Janitz M., Herwig R., Hennig S., Gorni C., Connor E.E., Sonstegard T.S., Smith T., Drogemuller C. & Williams J.L. (2006) A second generation radiation hybrid map to aid the assembly of the bovine genome sequence. *BMC Genomics* **7**, 283.

Jorde L.B. (2000) Linkage disequilibrium and the search for complex disease genes. *Genome Res* **10**, 1435-44.

Kalinowski S.T. (2002) How many alleles per locus should be used to estimate genetic distances? *Heredity* **88**, 62-5.

Kappes S.M., Keele J.W., Stone R.T., McGraw R.A., Sonstegard T.S., Smith T.P., Lopez-Corrales N.L. & Beattie C.W. (1997) A second-generation linkage map of the bovine genome. *Genome Res* **7**, 235-49.

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

Lynch M. & Walsh B. (1998) Genetics and Analysis of Quantitative Traits. Sunderland, MA: Sinauer.

Ma R.Z., Beever J.E., Da Y., Green C.A., Russ I., Park C., Heyen D.W., Everts R.E., Fisher S.R., Overton K.M., Teale A.J., Kemp S.J., Hines H.C., Guerin G. & Lewin H.A. (1996) A male linkage map of the cattle (Bos taurus) genome. *J Hered* **87**, 261-71.

Mariette S., Le Corre V., Austerlitz F. & Kremer A. (2002) Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Mol Ecol* **11**, 1145-56.

McKay S.D., Schnabel R.D., Murdoch B.M., Aerts J., Gill C.A., Gao C., Li C., Matukumalli L.K., Stothard P., Wang Z., Van Tassell C.P., Williams J.L., Taylor J.F. & Moore S.S. (2007a) Construction of bovine whole-genome radiation hybrid and linkage maps using high-throughput genotyping. *Anim Genet* **38**, 120-5.

McKay S.D., Schnabel R.D., Murdoch B.M., Matukumalli L.K., Aerts J., Coppieters W., Crews D., Dias Neto E., Gill C.A., Gao C., Mannen H., Stothard P., Wang Z., Van Tassell C.P., Williams J.L., Taylor J.F. & Moore S.S. (2007b) Whole genome linkage disequilibrium maps in cattle. *BMC Genet* **8**, 74.

McRae A.F., McEwan J.C., Dodds K.G., Wilson T., Crawford A.M. & Slate J. (2002) Linkage disequilibrium in domestic sheep. *Genetics* **160**, 1113-22.

Michal J.J., Zhang Z.W., Gaskins C.T. & Jiang Z. (2006) The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. *Anim Genet* **37**, 400-2.

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C. & Benkel B. (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. *J Anim Sci* **81**, 1919-25.

Morin P.A., Luikart G. & Robertson K. (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* **19**, 208-16.

Olsen H.G., Nilsen H., Hayes B., Berg P.R., Svendsen M., Lien S. & Meuwissen T. (2007) Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. *BMC Genet* **8**, 32.

Rexroad C.E., 3rd, Owens E.K., Johnson J.S. & Womack J.E. (2000) A 12,000 rad whole genome radiation hybrid panel for high resolution mapping in

cattle: characterization of the centromeric end of chromosome 1. *Anim Genet* **31**, 262-5.

Riquet J., Coppieters W., Cambisano N., Arranz J.J., Berzi P., Davis S.K., Grisart B., Farnir F., Karim L., Mni M., Simon P., Taylor J.F., Vanmanshoven P., Wagenaar D., Womack J.E. & Georges M. (1999) Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle. *Proc Natl Acad Sci U S A* **96**, 9252-7.

Schnabel R.D., Kim J.J., Ashwell M.S., Sonstegard T.S., Van Tassell C.P., Connor E.E. & Taylor J.F. (2005) Fine-mapping milk production quantitative trait loci on BTA6: analysis of the bovine osteopontin gene. *Proc Natl Acad Sci U S A* **102**, 6896-901.

Sonstegard T.S., Van Tassell C.P. & Ashwell M.S. (2001) Dairy Cattle Genomics: Tools to accelerate genetic improvement? *J Anim Sci* **79**, E:307-E15.

Steele M.R. & Georges M. (1991) Generation of bovine multisite haplotypes using random cosmid clones. *Genomics* **10**, 889-904.

Talbot C.J., Nicod A., Cherny S.S., Fulker D.W., Collins A.C. & Flint J. (1999) High-resolution mapping of quantitative trait loci in outbred mice. *Nat Genet* **21**, 305-8.

Tenesa A., Navarro P., Hayes B.J., Duffy D.L., Clarke G.M., Goddard M.E. & Visscher P.M. (2007) Recent human effective population size estimated from linkage disequilibrium. *Genome Res* **17**, 520-6.

Thomson G. (2001) Significance levels in genome scans. *Adv Genet* **42**, 475-86.

Van Laere A.S., Nguyen M., Braunschweig M., Nezer C., Collette C., Moreau L., Archibald A.L., Haley C.S., Buys N., Tally M., Andersson G., Georges M. & Andersson L. (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **425**, 832-6.

White S.N., Casas E., Wheeler T.L., Shackelford S.D., Koohmaraie M., Riley D.G., Chase C.C., Jr., Johnson D.D., Keele J.W. & Smith T.P. (2005) A new single nucleotide polymorphism in CAPN1 extends the current tenderness marker test to include cattle of Bos indicus, Bos taurus, and crossbred descent. *J Anim Sci* **83**, 2001-8.

Williams J.L., Eggen A., Ferretti L., Farr C.J., Gautier M., Amati G., Ball G., Caramorr T., Critcher R., Costa S., Hextall P., Hills D., Jeulin A., Kiguwa S.L., Ross O., Smith A.L., Saunier K., Urquhart B. & Waddington D. (2002) A bovine whole-genome radiation hybrid panel and outline map. *Mamm Genome* **13**, 469-74.

Winter A., Kramer W., Werner F.A., Kollers S., Kata S., Durstewitz G., Buitkamp J., Womack J.E., Thaller G. & Fries R. (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. *Proc Natl Acad Sci U S A* **99**, 9300-5.

Womack J.E., Johnson J.S., Owens E.K., Rexroad C.E., 3rd, Schlapfer J. & Yang Y.P. (1997) A whole-genome radiation hybrid panel for bovine gene mapping. *Mamm Genome* **8**, 854-6.

Zenger K.R., Khatkar M.S., Cavanagh J.A., Hawken R.J. & Raadsma H.W. (2007) Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. *Anim Genet* **38**, 7-14.

Zhang S., Knight T.J., Reecy J.M. & Beitz D.C. (2008) DNA polymorphisms in bovine fatty acid synthase are associated with beef fatty acid composition. *Anim Genet* **39**, 62-70.

CHAPTER 3

A high resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly[1]

3.1 Introduction

Radiation hybrid (RH) mapping is a powerful tool for establishing marker order across a number of species (Womack *et al.* 1997; Yerle *et al.* 1998; Williams *et al.* 2002). The advantage of RH mapping over other mapping approaches such as linkage maps is that RH mapping does not require polymorphic markers or large families, therefore increasing the number of loci potentially mapped.

In 2005, Everts-van der Wind *et al.* (2005) published the most comprehensive bovine whole genome radiation hybrid map including a total of 3000 markers on 29 chromosomes. Two other genome wide RH maps (Jann *et al.* 2006; McKay *et al.* 2007) with additional markers have been released since then. Considering that linkage maps are only useful when there is recombination between markers, a higher resolution RH panel provides the means to order closely linked markers. In addition, RH maps can be used either as scaffolds for correct genome assembly or for identifying and resolving misassembled regions of the genome sequence (Pitel *et al.* 2004). Currently, there are four whole genome radiation hybrid panels available in cattle (Womack *et al.* 1997; Rexroad

[1] A version of this chapter has been published. Marques et al., 2007. BMC Genomics: 8:254

et al. 2000; Williams *et al.* 2002; Itoh *et al.* 2005), with the highest resolution (12K rad) developed by Rexroad *et al.* (2000).

The increased availability of markers has led to the development of new methods for RH data analysis and map construction. The comparative mapping approach, a newly incorporated algorithm in CarthaGene (de Givry *et al.* 2005), takes advantage of the information already available for a particular genome sequence assembly, building more robust maps than the traditional approach. Using simulated data, less than 10% of the markers were wrongly positioned using the comparative mapping approach, while 33% of incorrectly positioned markers were observed using the traditional RH approach (Faraut *et al.* 2007). The traditional RH approach relies on heuristic methods resulting in framework maps that include only a small portion of all the markers (20% to 50%) (Faraut *et al.* 2007). On the other hand, the comparative mapping approach extends the usual statistical model describing the RH data (Boehnke *et al.* 1991) by adding a non-uniform prior distribution on the possible orders. Overall, the comparative mapping approach exploits the knowledge of a completely sequenced genome containing markers that have orthologous relationships with markers genotyped through the RH panel (de Givry *et al.* 2005; Faraut *et al.* 2007).

Our study uses this new mapping algorithm to build a high resolution radiation hybrid map of bovine chromosome 14 (BTA14) comparing specific discrepancies between our map and the latest sequence assembly. The identification of the correct order of markers on a specific chromosome is essential to the research community. Specifically, the large number of carcass

fatness quantitative trait loci (QTL) on BTA14 (Moore *et al.* 2003; Casas *et al.* 2004) makes it a prime target for fine scale mapping.

3.2 **Materials and Methods**

3.2.1 Compilation and development of SNP markers on BTA 14

SNPs included in the construction of the RH map were compiled and selected from the Baylor College of Medicine bovine database (ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/). At the time the experiment was conducted Btau_2.0 was the latest version available and all markers selected and genotyped were thought to belong to chromosome 14. Additional SNPs were derived from BAC end sequences (BES) of the CHORI-240 library (http://bacpac.chori.org/bovine240.htm) and IBISS (Interactive Bovine In Silico SNP) database (www.livestockgenomics.csiro.au/ibiss). The initial selection of markers included 1536 SNP markers, of which 148 were generated either by direct sequencing or by selection from NCBI or IBISS databases. A total of 429 markers could not be genotyped across the RH panel, and therefore were not included in the analysis. The remaining 264 markers were determined to form parts of different linkage groups (different chromosomes) through Lod scores in CarthaGene.

3.2.2 Primer design and sequencing of BES

Primer design for SNPs originating from BES was carried out using primer3 (http://frodo.wi.mit.edu/) using the following settings (min opt max): primer size: 22 24 26; primer tm: 58 60 62; primer GC%: 40 50 60. Genomic

DNA from 12 Angus animals was amplified using a PCR program with initial denaturing for 10 min at 94°C, denaturing for 30 sec at 94°C , annealing (55°C, 60°C or 65°C) and elongation (72°C) for 30 sec in 35 cycles.

PCR products were subjected to a clean up stage consisting of 0.5ul of an equal mixture of exonuclease I and shrimp alkaline phosphatase enzymes (Invitrogen) for 15min at 37°C and 15 min at 85°C. Clean PCR products were sequenced using BigDye-terminator chemistry (Applied Biosystems) and a 3730 DNA sequencer (Applied Biosystems). Sequence lengths ranged from 350 bp to 650 bp. The individual SNP sequence data were submitted to GenBank and are publicly available. SNPs from Baylor College of Medicine were submitted to Illumina (Illumina, Inc) and passed an internal quality control that predicted complementarity of primers and secondary structures (dimers, hairpin etc.). Only SNPs with an internal score of >0.6 (out of 1) were selected for genotyping.

3.2.3 Genotyping

The Illumina BeadStation 5.2 genotyping instrument (Illumina, Inc) was used for high throughput genotyping across the 12K radiation hybrid panel according to methods described by McKay *et al.* (2007). The software used for the genotyping analysis was Gencall version 5.2 (Illumina, Inc). Loci were scored based on the absence or presence of amplification. Markers that showed amplification in a particular clone were marked as 1, while markers showing no amplification were marked as zero. Markers whose amplification was uncertain were given a score of 2.

3.2.4 Construction of the 12K RH BTA14 map

RH analysis was carried out using CarthaGene software package (Schiex & Gaspin 1997). Previously mapped markers were used to assign markers to cattle chromosome 14 using a LOD score of 14 and a maximum distance of 100. After the linkage analysis was performed 843 markers were determined to be part of one linkage group. Markers were initially analyzed to identify any double markers (same retention pattern). These markers were merged to be part of the same bin. One marker of every bin was then selected to be mapped using the comparative mapping approach. This approach exploits a comparative 2-point model using RH data and the bovine sequence assembly Btau_3.1 as a reference order. This newly developed algorithm incorporated in CarthaGene (de Givry *et al.* 2005) is described in detail by Faraut *et al*. (2007). The expected number of breakpoints was set to 1 (default setting) and several 2-point reductions (Base TSP+MLE, Extended TSP+MLE, 2-point LOD distance) (Agarwala *et al.* 2000) were solved using the LKH heuristic methods (Helsgaun 2000). The final map was further improved by iteratively testing all the marker permutations in a small sliding window of size 7.

3.2.5 Comparative analysis with of the bovine assembly (Btau_3.1) and human chromosome 8 (HSA8)

Genomic sequence coordinates for SNPs were obtained by performing BLAST (Altschul *et al.* 1990) comparisons (using an E-value cutoff of 1e-50) between SNP flanking sequences and the latest bovine genome assembly (Btau_3.1). SNPs producing BLAST hits to multiple locations in the bovine

genome with the same coverage and sequence identity were removed from CarthaGene's marker input file during RH map construction. Approximate coordinates of the putative orthologous SNP regions in the human genome were obtained by performing BLAST searches (using an E-value cutoff of 1e-3) against the most recent human genome assembly (NCBI build 36). When bovine genome coordinates were available, the 3' end of the 3' flanking sequence of each SNP was extended (using sequence from the bovine genome assembly) prior to performing the comparison with the human genome, to give a total flanking sequence length of 20 kbp. This sequence extension step was performed because the existing flanking sequence did not produce a human genome BLAST hit in most cases. Homologous conserved synteny blocks and inversions between BTA14 and HSA8 were determined according to a set of rules described by Murphy *et al.* (2005).

3.2.6 Graphical representation of BTA14 map and comparative map

Visual representation of map alignments was achieved using AutoGRAPH (Derrien *et al.* 2007).

3.3 Results

3.3.1 Radiation hybrid map

A total of 843 single nucleotide polymorphism (SNP) markers were mapped to bovine chromosome 14 using the 12K rad bovine whole genome radiation panel (Rexroad *et al.* 2000). The majority of the SNP markers are derived from the bovine sequence database

((ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/). Twenty-four had been previously mapped using the 3K panel, 64 are from unmapped bac end sequences (BES) and 3 are from within genes known to be on BTA14. The RH map obtained has a log10-likelihood of -3835.03, with a total length of 4690.3 centirays (cR) and an average marker spacing of 96Kbp. The average retention frequency for all the markers mapped to BTA14 was 18%, with 478 unique retention patterns. A list of all mapped markers and their respective RH positions are given in Appendix One.

3.3.2 Alignment with $RH_{3,000}$ BTA14 map

There are 25 common markers between the high resolution BTA14 map presented here and the BTA14 $RH_{3,000}$ map described in McKay *et al.* (2007). Overall, there is a high consistency in marker order, except for two regions where closely mapped markers are inverted (Figure 3.1). The first inconsistency is comprised of markers SCAFFOLD105570_18245, SCAFFOLD230838_1182, SCAFFOLD135027_2960, SCAFFOLD135027_3247 and BES9_contig292_918. In our map, their positions range from 543.1 to 844 cR. The other region involves three flanking markers showing an inversion in their positions: SCAFFOLD40049_15114 at 3803.6 cR, BES7_Contig136_464 at 3819.9 cR and BES3_Contig324_378 at 3829.5 cR

3.3.3 *Alignment with bovine sequence assembly (Btau_3.1)*

Of the 843 markers mapped, 20 had multiple hits on different chromosomes when compared to the bovine sequence assembly (Btau_3.1) using BLAST. Most

of these hits occurred between BTA14 and an unassigned chromosome with similar BLAST scores (Table 3.1). There are several regions of discrepancies between our BTA14 RH map and the bovine sequence assembly (Btau_3.1) (Figure 3.2 and 3.3). A major region of inconsistency is near the centromere with smaller regions throughout the chromosome showing flips between sets of markers. Overall, the inconsistencies can be summarized by:

A) Single markers or group of closely mapped markers mapping somewhere else in the bovine sequence assembly. ie: BTA-12497 and BTA-20131 to BTA-86950

B) Inversion of flanking markers. ie: BTA-11589 and BTA-34555

C) Inversion of closely mapped markers. ie: BTA-04776 to BTA-06606

All inversions between closely mapped markers were analyzed and suggest incorrectly ordered scaffolds. The first case is represented by markers BTA-20131 (Scaffold: NW_001493188.1) and BTA-86950 (Scaffold: NW_001493187.1). In both maps, these markers map close together, however in our RH map, marker BTA-20131 maps before marker BTA-86950. According to the assembly these markers are approximately 23,000 base pairs away. The log10-likelihood for our order is -3835.05, while the assembly's order log10-likelihood is -3842.27. The second case showed problems in the arrangement within scaffolds. Markers BTA-11589 (NW_001493217.1) and BTA-34555 (NW_001493217.1) show a flip in their positions when compared to the sequence assembly. The log10-likelihood for the assembly, in this case, is -3850.65, while the log10-likelihood for our order

is still -3835.03. Both markers are part of the same scaffold indicating a possible mis-assembly within the scaffold.

The inconsistencies observed between our RH map and the assembly cannot be resolved by comparing previously published maps since there are no other maps of BTA14 with a comparable resolution. A complete list of markers showing inconsistent locations when compared to Btau_3.1 is presented in Table 3.2.

3.3.4 *Alignment with human chromosome 8*

Of the 843 markers ordered on the map, 828 markers (98%) have putative orthologs on the human chromosome 8 (HSA8) (NCBI build 36). Comparative analysis between bovine chromosome 14 and human chromosome 8 identified 4 homologous conserved synteny blocks (HSB): three previously published (Everts-van der Wind *et al.* 2005) and an extra conserved synteny block close to the telomere (Figure 3.4 and 3.5). This additional HSB block is comprised of 29 markers (BES8_Contig464_1373 to BTA-96554) and lies in a region with high consistency between our RH map and the assembly, therefore confirming the identification of a new evolutionary breakpoint. A number of gaps from a previous published map (Everts-van der Wind *et al.* 2005) have been filled and 18 small inversions identified. These inversions were predicted using a set of rules described by Murphy *et al.* (2005).

When two or more markers mapped to the same location, their relative positions were decided using a combination of their bovine assembly and human

chromosome 8 coordinates, with the bovine coordinates taking precedence over the human ones. In this case, the human coordinates were used to determine whether or not markers appeared in ascending or descending order, depending on which HSB they were in. Once a particular trend was observed, their relative positions were established based on the bovine sequence assembly; meaning that markers were arranged sequentially in an either ascending or descending trend, even if there were disagreements with the human coordinates. For example, according to their human coordinates, the order of the four markers mapping to position 1185.9 cR should be BTA-42142, BTA-42148, BTA-42161 and BTA-42153, however according to their assembly coordinates, BTA-42153 precedes BTA-42161 making the order of all four markers sequential (Appendix One).

3.4 Discussion

3.4.1 Comparison with other maps and Btau_3.1

In this study a comprehensive BTA14 RH map was built using the bovine assembly information (Btau_3.1) as a reference order. Traditionally, Lod scores have been used to determine the best fit map; however as the number of markers increases, it becomes more difficult to establish the next best map solely on the basis of these Lod scores. In the comparative method, the best map is a compromise between the RH data and the assembly and it works by comparing the likelihoods and breakpoints for the different maps. Briefly, if two maps have the same likelihood but different breakpoints, the order with fewer breakpoints is preferable. This approach demonstrates extreme robustness when building dense maps, as shown on simulated data and the dog genome (Faraut *et al.* 2007).

Comparison between our map and the previously released 3K RH BTA14 map (McKay *et al.* 2007) demonstrated a high degree of consistency except for regions where markers were in close proximity. In our map, those markers still map close to each other but with some slight shifts in order, particularly when the positions were just a few centirays apart. Perhaps the resolution of the 3K panel was not adequate for determining the order for those closely linked markers, since the number of cell lines for this panel is lower (94) than in the 12K panel (180).

Previously released radiation hybrid maps (Everts-van der Wind *et al.* 2005; Jann *et al.* 2006; McKay *et al.* 2007) have indicated regions that are inconsistent with the bovine sequence assembly. According to Jann *et al.* (2006), BTA14 was not among the chromosomes with a high number of discrepancies with the assembly (Btau_2.0). The inconsistencies observed referred mainly to the assignment of markers to other chromosomes. Such inconsistency still occurred in the latest assembly, but it was most likely due to repeated sequences assigned to multiple chromosomes. McKay *et al.* (2007) also indicated incorrectly assigned markers as well as some small inversions in scaffold ordering between their RH map and Btau_2.0 for some chromosomes; confirming that some discrepancies in scaffold arrangement were already present in previous assembly releases. Table 3.3 summarizes and compares the various BTA14 RH maps.

The vast number of markers made available through the bovine sequencing initiative has made possible the compilation of very closely linked markers. However, it is recognized that even this latest assembly contains a possible 20% error in scaffold assembly (George Weinstock, personal

communication), with no reports on the specific error rates for the scaffolds discussed here. Mammalian genomes are characterized by large duplications and abundant repetitive sequences which can complicate the final assembly (Salzberg & Yorke 2005). Finishing a genome does not necessarily indicate that the mis-assemblies will be resolved. It only means that the gaps are closed but that the sequence itself is not confirmed (Salzberg & Yorke 2005). Software limitations in assembling large, repeated sequences can cause incorrect ordering of large segments of DNA (Pop *et al.* 2004).

Differences in the animal resources used to produce the RH map and the bovine assembly for BTA14 seem unlikely to be the cause of the discrepancies discussed here. For instance, a high similarity in marker order should be expected between the genome of the line-bred Hereford bull represented in the BAC map and the genome of his daughter, which was used for the assembly. The pedigree relationship between this sire and daughter is 0.954 (Mike MacNeil, personal communication). However, comparisons between the BAC map, the 12K BTA14 RH map and the assembly showed that the highest agreement is between the BAC map and the RH map (Warren Snelling, personal communication), with the latter panel being constructed from an Angus bull (JEW38) fibroblast cells (Rexroad *et al.* 2000). Based on this and the fact that the likelihood for our best map is substantially higher (-3835.03) than the likelihood for the assembly order (-4541.33), the notion that the differences we observed are due to rearrangement of individual animal's genomes seems unlikely.

The ultimate map for a species is the correctly assembled genome sequence with the latest assembly having a 7.1 fold-coverage. The bovine sequence assembly used the whole genome shotgun sequencing approach as well as information from a minimum tiling path of BAC clones across the genome. Contigs, which are referred to as the basic units of contiguous bases, are linked together using information from read pairs at the end of clones. Linked contigs will form scaffolds which are, in turn, arranged along the chromosome using mapping information from MARC 2004 (Ihara *et al.* 2004) map. Therefore the observed error rate in scaffold arrangement for the assembly is most likely due to the error rate observed in the MARC 2004 linkage map.

A combination of multiple mapping approaches such as linkage and RH maps have demonstrated their feasibility for improving the assembly (Snelling *et al.* 2004; Weikard *et al.* 2006). A number of mapping approaches have aided the arrangement of scaffolds from the first release of the assembly until now (Ihara *et al.* 2004; Everts-van der Wind *et al.* 2005). Certain high resolution maps such as the one of BTA6 published by Weikard *et al.* (2006) presented a gene based comparative radiation hybrid map providing a platform for the assembly. All of these studies have contributed considerable information to the assembly, but mis-assemblies and inconsistencies are still present.

3.4.2 Comparison with human chromosome 8 (HSA8)

As the density of markers increases, new HSBs and evolutionary breakpoints are likely to be identified through comparative studies. Previously reported HSBs from an independent study (Everts-van der Wind *et al.* 2005) are

in overall agreement with those reported here. The new HSB identified in our map is supported since marker order in this region is highly consistent with the assembly order (http://www.ensembl.org/Bos_taurus). The number of inversions observed in our map (18) was higher than the number identified by a previous BTA14 map (3) (Everts-van der Wind *et al.* 2005). This is not surprising considering the increase in marker density. This increase in marker density coupled with certain limitations of the panel prevented some markers from mapping to unique positions, but by consolidating the human coordinates with the bovine assembly positions for these RH markers with the same position, it was possible to reduce the number of inversions from 25 to 18. A comparative genome assembly approach uses the information from a reference genome to build and arrange the sequenced genome (Pop *et al.* 2004). Therefore, using the high resolution RH map built here in addition to the cattle-human comparative maps already available, it should be possible to resolve rearrangements in the bovine genome assembly.

Table 3.1 List of specific markers mapped to 12K RH BTA14 map with BLAST hits to multiple chromosomes on Btau_3.1.

Accession number	Marker name	BTA14 $RH_{12,000}$ position (cR)	NCBI Accession number (scaffold)/ BTA	BLAST Score/ E-value	
ss61534608	BTA-34290	76.2	NW_001502201.1|BtUn_WGA3528_3	920	0.0
			NW_001493186.1|Bt14_WGA1694_3		
ss69374948	CC513828-C70T	313.1	NW_001497844.1|BtUn_WGA12210_3	279	3e-73
			NW_001493182.1|Bt14_WGA1690_3		
ss69374951	CC517185-A407G	320.9	NW_001508604.1|BtUn_WGA9931_3	361	2e-97
			NW_001493182.1|Bt14_WGA1690_3		
ss69374954	CC517185-A286G	338.7	NW_001508604.1|BtUn_WGA9931_3	366	5e-99
ss61535184	BTA-35317	413.3	NW_001504912.1|BtUn_WGA6239_3	920	0.0
ss69374970	BZ879040-A200G	725.9	NW_001508460.1|BtUn_WGA9787_3	366	5e-99
			NW_001493193.1|Bt14_WGA1701_3		
ss61473730	BTA-114222	938.9	NW_001493195.1|Bt14_WGA1703_3	894	0.0
			NW_001495461.1|Bt8_WGA1109_3		
ss38334682	BTA-12630	1705.8	NW_001493441.1|Bt16_WGA1949_3	920	0.0
			NW_001493205.1|Bt14_WGA1713_3		
ss61480494	BTA-34395	1847.2	NW_001503381.1|BtUn_WGA4708_3	754	0.0
ss61480492	BTA-34393	1847.2	NW_001503381.1|BtUn_WGA4708_3	769	0.0
ss61508240	BTA-34679	2593.8	NW_001493216.1|Bt14_WGA1724_3	920	0.0
			NW_001505586.1|BtUn_WGA6913_3		
ss61508239	BTA-34678	2593.8	NW_001493216.1|Bt14_WGA1724_3	920	0.0
			NW_001505586.1|BtUn_WGA6913_3	915	0.0
ss61508238	BTA-34677	2593.8	NW_001505586.1|BtUn_WGA6913_3	920	0.0
			NW_001493216.1|Bt14_WGA1724_3		
ss61508235	BTA-34674	2606.1	NW_001505586.1|BtUn_WGA6913_3	887	0.0
			NW_001493216.1|Bt14_WGA1724_3		
ss61475994	BTA-16955	2767.3	NW_001497831.1|BtUn_WGA12197_3	915	0.0
			NW_001493220.1|Bt14_WGA1728_3		
ss61485958	BTA-55549	3011.7	NW_001494573.1|Bt2_WGA221_3	920	0.0
			NW_001493224.1|Bt14_WGA1732_3	898	0.0
ss38337066	BTA-15014	3271.8	NW_001502681.1|BtUn_WGA4008_3	915	0.0
ss61564952	BTA-90430	3343.8	NW_001507724.1|BtUn_WGA9051_3	512	6e-143
			NW_001493239.1|Bt14_WGA1747_3		
ss61547746	BTA-58540	3401.2	NW_001494258.1|Bt24_WGA2566_3	869	0.0
			NW_001493238.1|Bt14_WGA1746_3	867	0.0
ss61467380	BTA-35166	3679.4	NW_001493243.1|Bt14_WGA1751_3	920	0.0
			NW_001493244.1|Bt14_WGA1752_3		

Table 3.2 Markers with inconsistent positions when comparing the 12k RH BTA14 and Btau_3.1 maps.

Accession number	Marker name	Discrepancy case (Results Section)	BTA14 $RH_{12,000}$ position (cR)	Btau_3.1 position (bp)	NCBI Accession number (scaffold)
ss61480708	BTA-35408		33.8	3259417	NW_001493183.1
	to	A	to	to	
ss61480590	BTA-34737		41.3	3369381	NW_001493183.1
ss61534608	BTA-34290	A	76.2	5490676	NW_001493186.1
ss61567744	BTA-95738		94	4166573	NW_001493184.1
	to	A,C	to	to	
ss69374927	CC550917-A78G		105.8	3824944	NW_001493184.1
ss61467365	BTA-34867		115.9	5087204	NW_001493185.1
	to	A,C	to	to	
ss69374937	CC516254-A103G		159.7	4217777	NW_001493185.1
ss61476791	BTA-20131		161.4	5582192	NW_001493188.1
	to	A,B	to	to	
ss61563126	BTA-86950		165.7	5558546	NW_001493187.1
ss38328040	BTA-05988		170	5842011	NW_001493188.1
	to	A	to	to	
ss69374939	BZ945547-A231G		181.1	6310161	NW_001493188.1
ss38334549	BTA-12497	A	193	7717936	NW_001493191.1
ss61497128	BTA-98667		202.2	3101373	NW_001493182.1
	to	A	to	to	
ss69374946	CC514645-T214G		284.5	2323067	NW_001493182.1
ss61494244	BTA-87742	A	288.2	7621720	NW_001493190.1
ss61563555	BTA-09947		295.6	2010260	NW_001493182.1
	to	A,C	to	to	
ss69374951	CC517185-A407G		320.9	1392617	NW_001508604.1
ss61501188	BTA-35343		338.7	1289089	NW_001493182.1
	To	A	to	to	
ss69374955	CC530516-A378G		367	1319207	NW_001493182.1
ss69374978	CL605960-C177T		893.2	12779054	NW_001493195.1
	to	A	to	to	
ss69374979	CL605960-C179T		924.1	12779056	NW_001493195.1
ss61519384	BTA-114242		931.5	12655112	NW_001493195.1
	to	A,C	to	to	
ss61522828	BTA-120525		1022.3	10938710	NW_001493195.1
ss38328882	BTA-06830		1085.9	17140783	NW_001493201.1
	to	A	to	to	
btcn20869	NDUFB9-G249T		1173.5	18689115	NW_001493201.1
ss61508639	BTA-42136		1181.9	13680761	NW_001493200.1
	to	A	to	to	
ss61535591	BTA-36054		1358.4	16668155	NW_001493200.1
ss61525638	BTA-17314	A	1367.1	20599967	NW_001493203.1
	to		to	to	
ss61562835	BTA-86411		1440.6	22258733	NW_001493203.1
ss61477078	BTA-21240		1444.8	18869062	NW_001493202.1
	to	A	to	to	
ss61570309	BTA-34296		1519.8	20507528	NW_001493202.1

Accession number	Marker name	Discrepancy case (Results Section)	BTA14 $RH_{12,000}$ position (cR)	Btau_3.1 position (bp)	NCBI Accession number (scaffold)
ss38336702	BTA-14650		1524	23793721	NW_001493204.1
	to	C	to	to	
ss61534593	BTA-34271		1605.6	22342243	NW_001493204.1
ss61570240	BTA-24548		1612.2	27017291	NW_001493207.1
	to	A,C	to	to	
ss61496808	BTA-97344		1647.1	26555321	NW_001493207.1
ss61534613	BTA-34310		1662	24203843	NW_001493205.1
	to	A	to	to	
ss61478665	BTA-27436		1799.1	25574206	NW_001493205.1
ss61480485	BTA-34380		1811.8	27510984	NW_001493209.1
	to	A,C	to	to	
ss61534647	BTA-34396		1840.6	27074160	NW_001493208.1
ss61480499	BTA-34410		1845.4	26302678	NW_001493206.1
	to	A,C	to	to	
ss61476316	BTA-18140		1872.6	25756553	NW_001493206.1
ss38333641	BTA-11589		2369.8	37781488	NW_001493217.1
	to	A,B	to	to	
ss61498127	BTA-34555		2383.4	37335273	NW_001493217.1
ss61534788	BTA-34656	A	2385.6	38446496	NW_001493219.1
ss61470582	BTA-101611		2393.5	37967708	NW_001493218.1
	to	A	to	to	
ss61498127	BTA-107719		2412.8	38297010	NW_001493218.1
ss61515968	BTA-107731		2438.5	33614205	NW_001493216.1
	to	A	to	to	
ss61530605	BTA-26688		2740.2	36743502	NW_001493216.1
ss61534887	BTA-34802		2995.8	45426787	NW_001493225.1
	to	A	to	to	
ss38330676	BTA-08624		3001	45522923	NW_001493225.1
ss61480605	BTA-34832		3009	44447078	NW_001493224.1
	to	A	to	to	
ss61480597	BTA-34816		3042.4	44900551	NW_001493224.1
ss38325277	BTA-03225		3145	48585654	NW_001493231.1
	to	A,B	to	to	
ss38336085	BTA-14033		3166.1	48182373	NW_001493230.1
ss61521093	BTA-117438		3179.6	50195538	NW_001493232.1
	to	C	to	to	NW_001493232.1
ss61534932	BTA-34907		3246.6	49134359	
ss61516731	BTA-109297		3256.8	47947044	NW_001493229.1
	to	A,B	to	to	
ss61473093	BTA-111412		3271.8	47645693	NW_001493229.1
ss69374987	BTA-25649		3275.8	50835960	NW_001493234.1
	to	B	to	to	
ss61527579	BTA-20961		3277.8	50801932	NW_001493234.1
ss61517555	BTA-110811		3301.9	52015677	NW_001493235.1
	to	B	to	to	
ss61487100	BTA-60159		3322.7	51140625	NW_001493235.1
rs29013644	SCAFFOLD106433_368	A	3329	54323295	NW_001493240.1

Accession number	Marker name	Discrepancy case (Results Section)	BTA14 $RH_{12,000}$ position (cR)	Btau_3.1 position (bp)	NCBI Accession number (scaffold)
ss61564952	BTA-90430		3343.8	54282640	NW_001507724.1
	to	B	to	to	
ss61506794	BTA-115692		3409.1	52898920	NW_001493238.1
ss61564183	BTA-88967	A	3428.4	62480674	NW_001493250.1
ss38326828	BTA-04776		3432.4	57846844	NW_001493242.1
	to	A,B	to	to	
ss38328658	BTA-06606		3525	55263181	NW_001493242.1
rs29022898	SCAFFOLD75393_2246		3550.6	62192328	NW_001493247.1
	to	A,B	to	to	
ss61520620	BTA-116472		3563.9	61892052	NW_001493247.1
ss61478975	BTA-28611		3570	61164407	NW_001493245.1
	to	A,C	to	to	
ss61471416	BTA-104921		3582.5	61498397	NW_001493245.1
ss61472941	BTA-110841		3601.2	57998973	NW_001493243.1
	to	A	to	to	
ss61490518	BTA-72921		3761.8	60955585	NW_001493244.1
ss61535147	BTA-35242		3781.1	65514461	NW_001493253.1
	to	A	to	to	
ss38329727	BTA-07675		3819.9	66203557	NW_001493254.1
ss61535181	BTA-35306		3827.1	63014890	NW_001493251.1
	to	A	to	to	
ss38332752	BTA-10700		3943.7	65256182	NW_001493251.1
ss38333031	BTA-10979		4190	74543809	NW_001493260.1
	to	A	to	to	
ss38336767	BTA-14715		4201.6	75391577	NW_001493261.1
ss61535281	BTA-35498		4209.8	71858373	NW_001493258.1
	to	A, C	to	to	
ss61535276	BTA-24001		4236.9	71420092	NW_001493258.1
ss61529051	BTA-23998		4265.2	72032931	NW_001493259.1
	to	A	to	to	
ss61564556	BTA-89757		4382.3	74239095	NW_001493259.1
ss61526929	BTA-19771	A	4612.3	81377972	NW_001493269.1

A-Single markers or group of closely mapped markers mapping somewhere else in the bovine sequence assembly.
B-Inversion of flanking markers.
C-Inversion of closely mapped markers

Table 3.3 Comparison of different BTA14 RH maps including the radiation panel used, techniques used and number of common markers.

	McKay *et al* (2007)	Jann *et al*. (2006)	This Study
Panel type	WG-3K (Williams *et al.* 2002)	WG-3K (Williams *et al.* 2002)	WG-12K (Rexroad *et al.* 2000)
Number of markers mapped to BTA14	215	222	843
Number of markers in common with this study	25	0	-
Number of cell lines	94	94	180
Methodologies	Traditional RH approach and Illumina-based RH typing method	Traditional RH approach and conventional RH typing method	Comparative approach and Illumina-based RH typing method

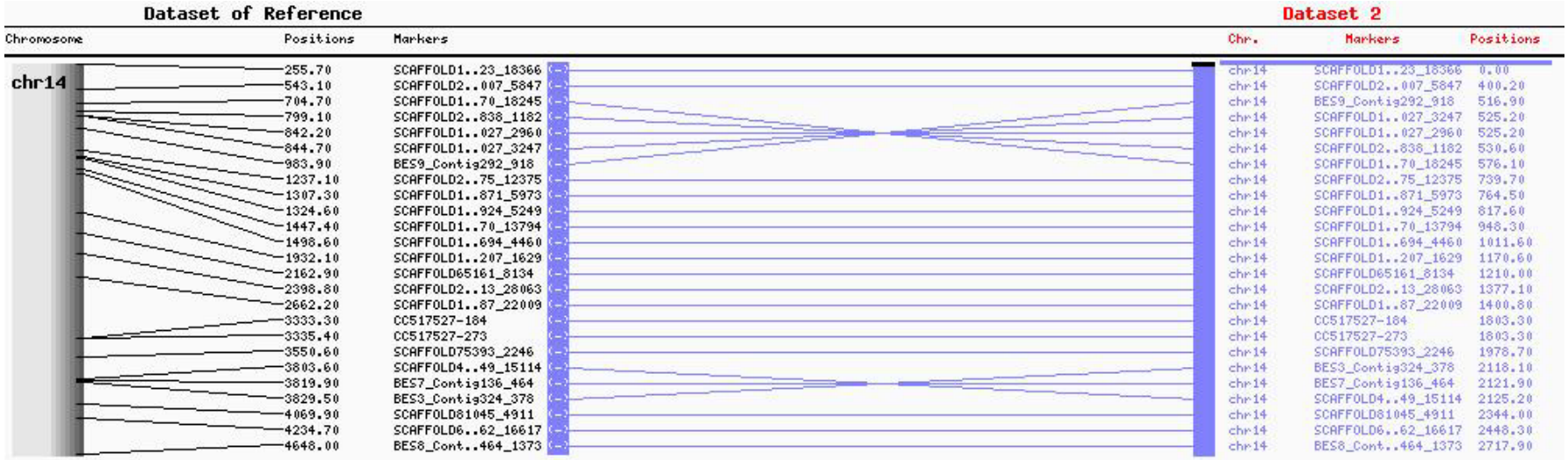

Figure 3.1 12K RH map of BTA14 compared with the UofA $RH_{3,000}$.The right side map refers to 12K BTA14 while the left side refers to 3K BTA14. Common markers are highlighted in bold and connected through blue lines. Distances on both maps are represented in cR.

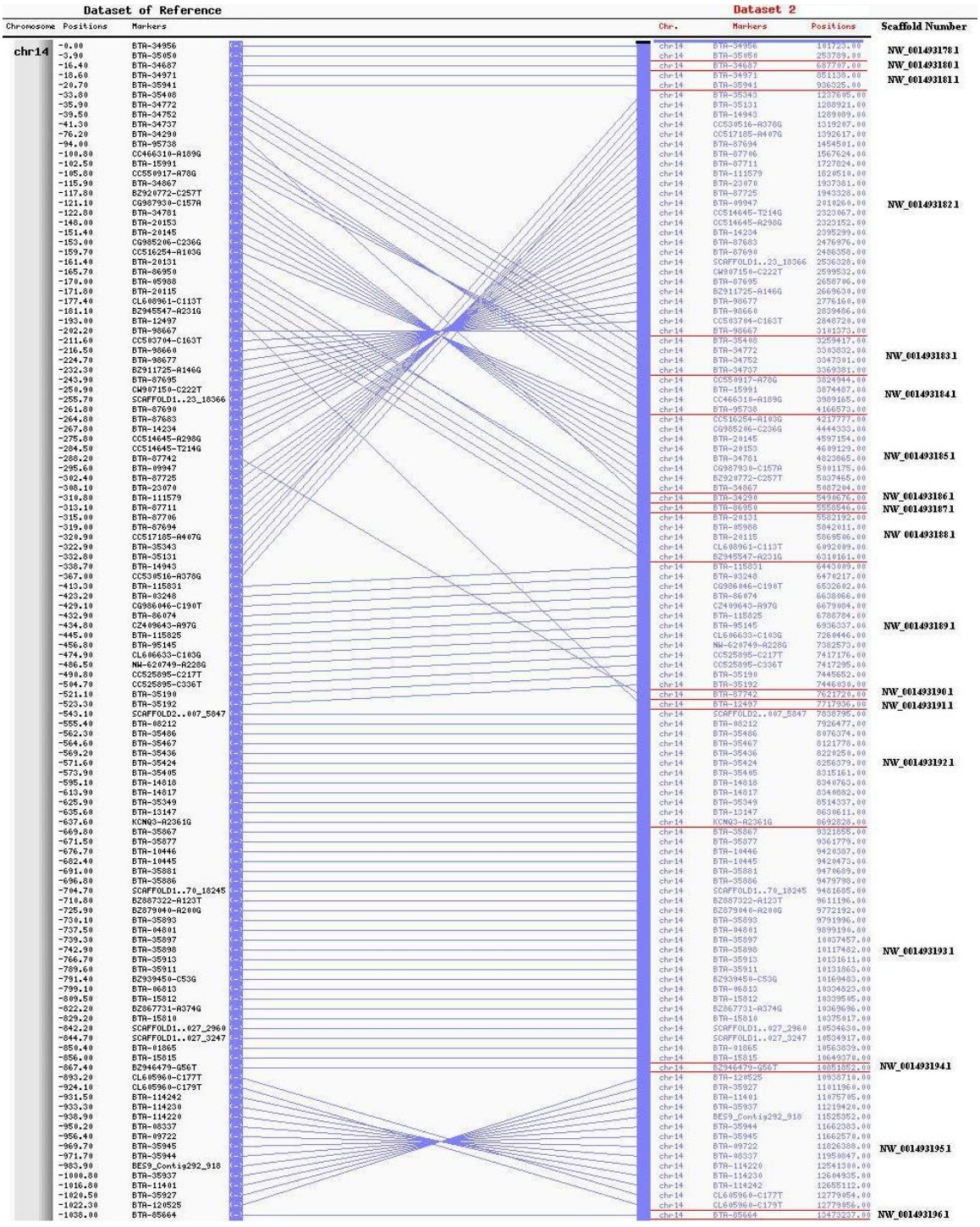

Figure 3.2 12K RH map of BTA14 compared to the corresponding Btau_3.1 map. This figure shows the upper quartile of the map. For the full image see figure 3.4. The right side map refers to 12K BTA14 while the left side map refers to Btau_3.1. Common markers are connected through blue lines. For legibility purposes, only markers with unique retention patterns are displayed. Distances on 12K RH map are represented in cR, on the Btau_3.1 map in base pair positions. Red lines represent breaks in the scaffold numbers.

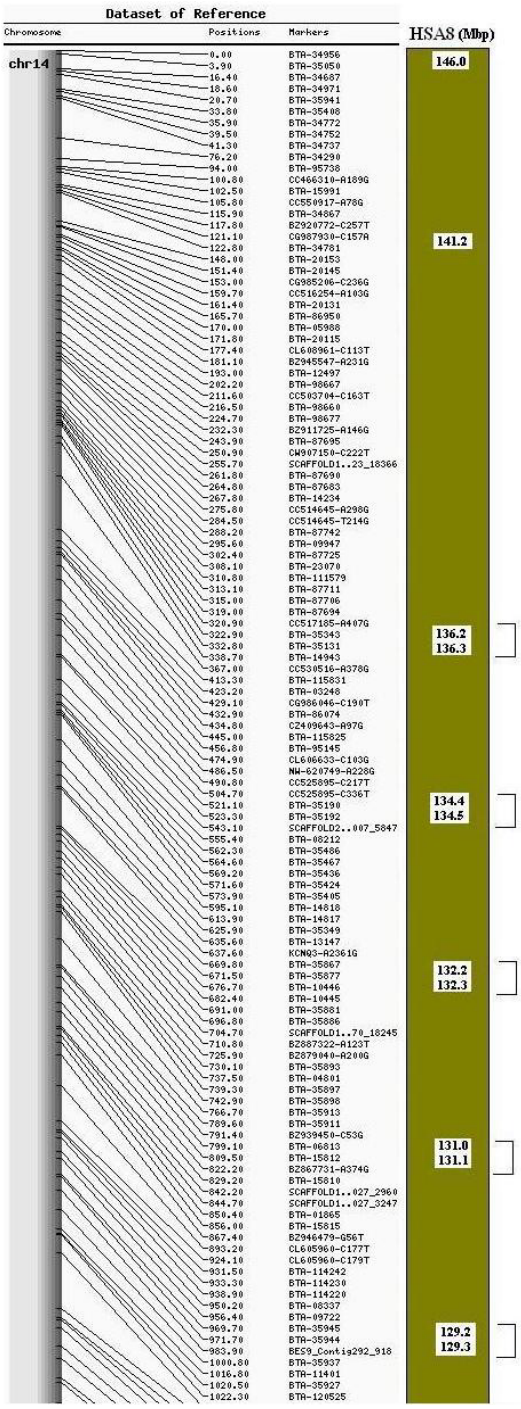

Figure 3.3 12K RH map of BTA14 with Homologous Conserved Synteny Blocks from HSA8.This figure shows the upper quartile of the map. For the full image please see figure 3.5. Comparative map between the 12K radiation hybrid map of bovine chromosome 14 and human chromosome 8. Human positions are represented in Mbp. Radiation hybrid map distances are scaled in cR. The grey bar represents the chromosome with black lines pointing to the markers. Brackets indicate small inversions, which would otherwise make the order in respect to the human chromosome 8 perfect.

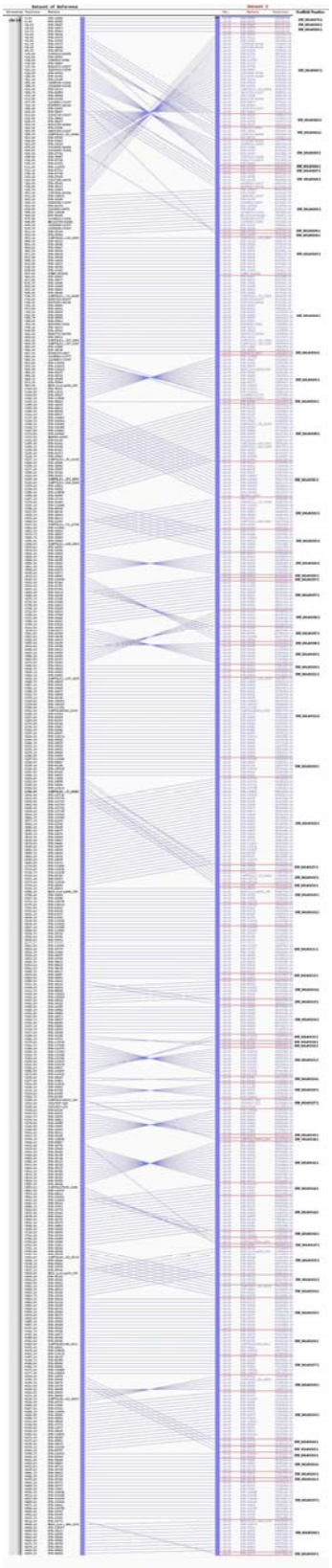

Figure 3.4 Full image of 12K RH map of BTA14 compared to the corresponding Btau_3.1 map

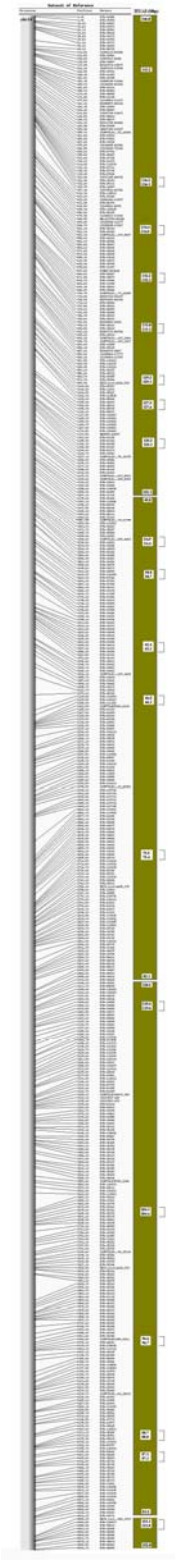

Figure 3.5 Full image of 12K RH map of BTA14 with Homologous Conserved Synteny Blocks from HSA8

3.5 Literature Cited

Bovine Sequence database. [ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/].

CHORI-240 Bovine BAC Library. [http://bacpac.chori.org/bovine240.htm].

Ensembl Genome Browser. [http://www.ensembl.org/Bos_taurus].

Interactive Bovine In Silico SNP. [www.livestockgenomics.csiro.au/ibiss].

Primer3. [http://frodo.wi.mit.edu/].

Agarwala R., Applegate D.L., Maglott D., Schuler G.D. & Schaffer A.A. (2000) A fast and scalable radiation hybrid map construction and integration strategy. *Genome Res* **10**, 350-64.

Altschul S.F., Gish W., Miller W., Myers E.W. & Lipman D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403-10.

Boehnke M., Lange K. & Cox D.R. (1991) Statistical methods for multipoint radiation hybrid mapping. *Am J Hum Genet* **49**, 1174-88.

Casas E., Keele J.W., Shackelford S.D., Koohmaraie M. & Stone R.T. (2004) Identification of quantitative trait loci for growth and carcass composition in cattle. *Anim Genet* **35**, 2-6.

de Givry S., Bouchez M., Chabrier P., Milan D. & Schiex T. (2005) CARTHA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* **21**, 1703-4.

Derrien T., Andre C., Galibert F. & Hitte C. (2007) AutoGRAPH: an interactive web server for automating and visualizing comparative genome maps. *Bioinformatics* **23**, 498-9.

Everts-van der Wind A., Larkin D.M., Green C.A., Elliott J.S., Olmstead C.A., Chiu R., Schein J.E., Marra M.A., Womack J.E. & Lewin H.A. (2005) A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. *Proc Natl Acad Sci U S A* **102**, 18526-31.

Faraut T., de Givry S., Chabrier P., Derrien T., Galibert F., Hitte C. & Schiex T. (2007) A comparative genome approach to marker ordering. *Bioinformatics* **23**, e50-6.

Helsgaun K. (2000) An effective implementation of the Lin-Kernighan traveling salesman heuristic. *European Journal of Operational Research* **126**, 106-30.

Ihara N., Takasuga A., Mizoshita K., Takeda H., Sugimoto M., Mizoguchi Y., Hirano T., Itoh T., Watanabe T., Reed K.M., Snelling W.M., Kappes S.M., Beattie C.W., Bennett G.L. & Sugimoto Y. (2004) A comprehensive genetic map of the cattle genome based on 3802 microsatellites. *Genome Res* **14**, 1987-98.

Itoh T., Watanabe T., Ihara N., Mariani P., Beattie C.W., Sugimoto Y. & Takasuga A. (2005) A comprehensive radiation hybrid map of the bovine genome comprising 5593 loci. *Genomics* **85**, 413-24.

Jann O.C., Aerts J., Jones M., Hastings N., Law A., McKay S., Marques E., Prasad A., Yu J., Moore S.S., Floriot S., Mahe M.F., Eggen A., Silveri L., Negrini R., Milanesi E., Ajmone-Marsan P., Valentini A., Marchitelli C., Savarese M.C., Janitz M., Herwig R., Hennig S., Gorni C., Connor E.E., Sonstegard T.S., Smith T., Drogemuller C. & Williams J.L. (2006) A

second generation radiation hybrid map to aid the assembly of the bovine genome sequence. *BMC Genomics* **7**, 283.

McKay S.D., Schnabel R.D., Murdoch B.M., Aerts J., Gill C.A., Gao C., Li C., Matukumalli L.K., Stothard P., Wang Z., Van Tassell C.P., Williams J.L., Taylor J.F. & Moore S.S. (2007) Construction of bovine whole-genome radiation hybrid and linkage maps using high-throughput genotyping. *Anim Genet* **38**, 120-5.

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C. & Benkel B. (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. *J Anim Sci* **81**, 1919-25.

Murphy W.J., Larkin D.M., Everts-van der Wind A., Bourque G., Tesler G., Auvil L., Beever J.E., Chowdhary B.P., Galibert F., Gatzke L., Hitte C., Meyers S.N., Milan D., Ostrander E.A., Pape G., Parker H.G., Raudsepp T., Rogatcheva M.B., Schook L.B., Skow L.C., Welge M., Womack J.E., O'Brien S J., Pevzner P.A. & Lewin H.A. (2005) Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. *Science* **309**, 613-7.

Pitel F., Abasht B., Morisson M., Crooijmans R.P., Vignoles F., Leroux S., Feve K., Bardes S., Milan D., Lagarrigue S., Groenen M.A., Douaire M. & Vignal A. (2004) A high-resolution radiation hybrid map of chicken

chromosome 5 and comparison with human chromosomes. *BMC Genomics* **5**, 66.

Pop M., Phillippy A., Delcher A.L. & Salzberg S.L. (2004) Comparative genome assembly. *Brief Bioinform* **5**, 237-48.

Rexroad C.E., 3rd, Owens E.K., Johnson J.S. & Womack J.E. (2000) A 12,000 rad whole genome radiation hybrid panel for high resolution mapping in cattle: characterization of the centromeric end of chromosome 1. *Anim Genet* **31**, 262-5.

Salzberg S.L. & Yorke J.A. (2005) Beware of mis-assembled genomes. *Bioinformatics* **21**, 4320-1.

Schiex T. & Gaspin C. (1997) CARTHAGENE: constructing and joining maximum likelihood genetic maps. *Proc Int Conf Intell Syst Mol Biol*, 258-67.

Snelling W.M., Gautier M., Keele J.W., Smith T.P., Stone R.T., Harhay G.P., Bennett G.L., Ihara N., Takasuga A., Takeda H., Sugimoto Y. & Eggen A. (2004) Integrating linkage and radiation hybrid mapping data for bovine chromosome 15. *BMC Genomics* **5**, 77.

Weikard R., Goldammer T., Laurent P., Womack J.E. & Kuehn C. (2006) A gene-based high-resolution comparative radiation hybrid map as a framework for genome sequence assembly of a bovine chromosome 6 region associated with QTL for growth, body composition, and milk performance traits. *BMC Genomics* **7**, 53.

Williams J.L., Eggen A., Ferretti L., Farr C.J., Gautier M., Amati G., Ball G., Caramorr T., Critcher R., Costa S., Hextall P., Hills D., Jeulin A., Kiguwa S.L., Ross O., Smith A.L., Saunier K., Urquhart B. & Waddington D. (2002) A bovine whole-genome radiation hybrid panel and outline map. *Mamm Genome* **13**, 469-74.

Womack J.E., Johnson J.S., Owens E.K., Rexroad C.E., 3rd, Schlapfer J. & Yang Y.P. (1997) A whole-genome radiation hybrid panel for bovine gene mapping. *Mamm Genome* **8**, 854-6.

Yerle M., Pinton P., Robic A., Alfonso A., Palvadeau Y., Delcros C., Hawken R., Alexander L., Beattie C., Schook L., Milan D. & Gellin J. (1998) Construction of a whole-genome radiation hybrid panel for high-resolution gene mapping in pigs. *Cytogenet Cell Genet* **82**, 182-8.

CHAPTER 4

High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle[2].

4.1 Introduction

In previous studies, large variations in linkage disequilibrium (LD) have been reported (Farnir *et al.* 2000; Tenesa *et al.* 2003; Vallejo *et al.* 2003; Gautier *et al.* 2007; McKay *et al.* 2007). Different measures of LD such as r^2 and D' are known to yield different conclusions in terms of the extent of LD. In studies using microsatellites and D' as a primary measure of LD (Farnir *et al.* 2000; Tenesa *et al.* 2003; Vallejo *et al.* 2003) it was reported that LD extended for several megabases. On the other hand, when r^2 was used, LD was shown to be at background levels (r^2 at approximately 0.1) after only 500 kilo base pairs (kbp) (Gautier *et al.* 2007; McKay *et al.* 2007). Differences in marker types used in these studies are also potential causes for LD variation, with microsatellites being more suitable for detecting long range LD than SNPs (Varilo *et al.* 2003).

High resolution LD maps can provide information on specific markers that are part of haplotype blocks used in association analysis (Gautier *et al.* 2007; Khatkar *et al.* 2007). Previous whole genome linkage disequilibrium maps in cattle (Khatkar *et al.* 2007; McKay *et al.* 2007) have been used to analyze different aspects of LD. In the case of McKay *et al.* (2007), approximately 3,000 markers (microsatellites and SNPs) were used to assess the extent of LD in eight

[2] A version of this chapter has been published. Marques et al., 2008. BMC Genetics: 9:45

different cattle breeds, while Khatkar *et al.* (2007) analyzed the haplotype block diversity in Holstein-Friesan cattle using approximately 15,000 SNPs. The latter also used the Btau_3.1 build to arrange markers along the genome, however it is now known that BTAu_3.1 build has inconsistencies with other independently built cattle maps (Marques *et al.* 2007; Snelling *et al.* 2007).

In addition, such LD maps can be considered a crucial tool for researchers looking to confirm or exclude potential polymorphisms as causative mutations. Recent studies using breed specific LD information have shed light on the importance of using LD information to link potential markers to economically relevant traits in cattle. In 2007, Olsen *et al.* (2007) reported that a mutation in *ABCG2*, a gene responsible for secreting important substrates into milk (Jonker *et al.* 2005), is the most likely candidate for affecting the observed milk yield quantitative trait loci (QTL) on BTA6 (Georges *et al.* 1995). The approach used included constructing a dense marker map spanning the QTL region and using linkage and linkage disequilibrium information to assess polymorphisms in *ABCG2* and other genes.

Correct marker order is crucial for construction of linkage disequilibrium and haplotype maps, as well as for future candidate gene searches on chromosomes harboring economically important traits. Bovine chromosome 14 (BTA14) is widely known to harbor quantitative trait nucleotides (QTN) with large effect on milk fat percentage (Grisart *et al.* 2004) and marbling (Barendse 1999). In addition, several QTL affecting other economically important traits have been identified on BTA14 (Moore *et al.* 2003; Mizoshita *et al.* 2005).

This study focuses on the comparison of linkage disequilibrium (r^2) between Holstein and Angus cattle using over 500 BTA14 single nucleotide polymorphism (SNP) markers on 331 Holstein and 137 Angus animals. As well, it identifies specific haplotype blocks and tagged SNPs for BTA14 which will be useful for future whole genome association studies.

4.2 Materials and Methods

4.2.1 Animal Resource

Three hundred and thirty-one Holstein bulls provided by Semex Canada and one hundred and thirty seven American Angus bulls were used in this study. The Holstein bulls represent an eight generation extended pedigree. Angus families were selected to consist of one grandparent, one parent and three or more progeny. This pedigree structure has previously produced efficient estimates of phased haplotypes. Pedigree information for Holstein animals was obtained from the Animal Improvement Program Laboratory of the USDA (http://www.holstein.ca/english/AnimalInq/animalinq.asp). Pedigree information for Angus bulls was provided by the American Angus Association (http://www.angus.org/pr/pr_main.html).

4.2.2 Selection and Genotyping of Markers

Single Nucleotide Polymorphisms (SNPs) included in this study were selected from the Bovine genome project (http://www.angus.org/pr/pr_main.html) previously mapped onto BTA14 according to procedures described by Marques *et*

al. (2007). SNPs were analyzed using an Illumina BeadStation 5.2 genotyping instrument (Illumina, Inc) and SNP genotypes were assigned using BeadStudio (Illumina, Inc) software.

4.2.3 LD Analysis

Only markers successfully mapped on BTA14 were used in this study even if they were successfully genotyped on both breeds. Initially all 843 markers from Marques *et al.* (2007) were genotyped. Thirty-one did not successfully amplify on both breeds. These markers were then filtered to exclude loci with a Minor Allele Frequency (MAF) < 0.02 or that had greater than 10% missing genotypes within a breed. This filtering resulted in 518 and 505 candidate loci in Angus and Holstein respectively which were used for further analysis.

Genotype quality and haplotypes were estimated with GENOPROB 2.0 (Thallman *et al.* 2001b, 2001a) using the map coordinates of Marques *et al.* (2007) and the extended pedigree relating all animals within each breed. GENOPROB estimates the probability that a genotype is correct (pGmx) as well as identifies the most likely phase relationship between the alleles. Only high probability (pGmx≥0.95) genotypes were considered for further analysis with no restriction used for order probability. Recent reports on GENOPROB showed that Holstein and Angus breeds produced the most accurately estimated genotypes and phased chromosome due to their complex pedigree structure (McKay *et al.* 2007). Once paternal and maternal haplotypes were estimated they were inserted onto HAPLOVIEW (Barrett *et al.* 2005) to verify their quality. The settings used

included: min genotyped %: 50, Hardy Weinberg (HW) p-value cutoff: 0.0010, Minimum minor allele freq: 0.0010, Maximum # Mendel errors: 1. Overall, 9 markers dropped out from the Angus genotypes and 3 from Holstein.

Markers passing the above filtering criteria were used to estimate LD using only the maternal haplotype and the program HAPLOXT (Abecasis & Cookson 2000). The average marker spacing using this subset of markers was approximately 170 kbp, with the smallest and largest gaps between markers being 0.03 kbp and 2256.72 kbp, respectively. Maternal haplotypes were used in order to avoid biasing the linkage disequilibrium values due to the pedigree structure, which were solely paternal lineages.

Marker positions were inferred using the marker order from the 12K RH map of Marques *et al.* (2007) which is in high agreement with the recently released physical map based on the independent whole genome map of Snelling *et al.* (2007). Preliminary comparison between our marker order and the recently published bovine sequence assembly Btau_4.0 (ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/) shows agreement for markers common (737 markers) to both maps. Relative bp positions were calculated by dividing the highest centiray (cR) position by the corresponding bp position in the bovine sequence assembly Btau_3.1. The resultant average was approximately 17 kbp per cR. In regions where multiple markers had the same cR position, the sequence assembly distance was used. These closely mapped markers were in agreement with the assembly, according to results presented by Marques *et al.* (2007).

Correlation of r-value used 419 markers common to both breeds. The r-value was calculated according to the formula (Hastings 1985):

$$r = \frac{(freq(A1\_B1) * freq(A2\_B2) - freq(A1\_B2) * freq(A2\_B1))}{\sqrt{freq(A1) * freq(B1) * freq(A2) * freq(B2)}}$$

Where A1 is the first allele of the first marker making up the haplotypes A1_B1 or A1_B2, A2 is the second allele of the first marker, B1 is the first allele of the second marker and B2 is the second allele of the second marker. Marker phase analysis was performed as follows: First, marker pairs had their r-values and inter-marker distances calculated. Next, the correlation of r-values (same marker pair) for each breed was calculated. Markers were then binned according to their inter-marker distance (category) and their correlation results averaged for each category. Haplotype and allele frequencies were calculated using SAS version 1.1.3 (SAS, Inc).

Calculation of extended haplotype homozygosity (EHH) was performed using a EHH calculation web tool (Mueller & Andreoli 2004) designed according to procedures described by Sabeti *et al.* (2002). Only maternal haplotypes were loaded.

4.3 Results and Discussion

Markers were binned according to marker distances (kbp) and r^2 was averaged and plotted for each category (Figure 4.1). LD drops from an average of 0.687 for Holstein and 0.648 for Angus to 0.328 and 0.317, respectively, when

going from 1 kbp to 50 kbp marker distance in both breeds. Moderate levels of LD (r^2 at approximately 0.2) are reached at around 100 kbp and background levels (r^2 at approximately 0.1) at around 500 kbp. Both breeds show an inverse relationship between LD and marker distance, confirming recent studies on r^2 measures in cattle (Gautier *et al.* 2007; McKay *et al.* 2007). The average r^2 value for Holstein in McKay *et al.* (2007) was higher (0.91) than in our study (0.687) for the 1kbp inter-marker distance. This difference in value can be attributed to the wide range (0.005 to 1) in LD in our study (Table 4.1 and 4.2). McKay *et al.* (2007) calculated LD using 81 markers for BTA14 compared to 502 in our study. The range in LD in our study is most likely a result of sampling of gametes to form successive generations (Weir & Hill 1980) which is dependent on finite population size and not so much on the sample size. In this case, there could have been ancestral recombination between certain markers in close proximity, but not others. This is plausible, in the case of maternal haplotypes, when one considers the complexity of the pedigrees for both populations, with dams sometimes contributing information to multiple families. Another important aspect to mention in this analysis is the half-sib relationship among some dams in the Holstein population, causing inflated LD values. In addition to these findings, there is a more rapid decline in LD for Angus compared to Holstein overall. Differences in effective population sizes for both breeds are a plausible explanation for this observed difference.

There are a number of algorithms used to define haplotype blocks (Daly *et al.* 2001; Gabriel *et al.* 2002; Zhang *et al.* 2002; Greenspan & Geiger 2006). The

confidence interval algorithm (Gabriel *et al.* 2002) used by Khatkar *et al.* (2007) relies on D' measures between markers to define blocks. The other approach used in LD analysis in dogs (Lindblad-Toh *et al.* 2005) and more recently in cattle (Gautier *et al.* 2007) utilizes the four gamete rule (Wang *et al.* 2002) which defines blocks based on all 4 possible two-marker haplotypes existing with observed frequencies of at least 0.01. Using this method incorporated in HAPLOVIEW (Barrett *et al.* 2005), 122 blocks (33 bp to 1338 kbp) were identified in Holstein and 122 blocks (45 bp to 1767 kbp) were identified in Angus (Figure 4.2 and 4.3). The confidence interval method used by Khatkar *et al.* (2007) found 27 blocks for BTA14. Khatkar *et al.* (2007) included 303 BTA14 markers on Holstein-Friesan cattle compared to 502 BTA14 markers in this study, so it is expected that as the number of markers increases more haplotype blocks are identified. However, Khatkar *et al.* (2007) not only used a different haplotype finding method, but also a different marker order causing differences in the number of blocks found. Another difference to take into consideration is that our haplotype block evaluation did not focus on coding regions, unlike Khatkar *et al.* (2007). Indeed, knowledge of LD within candidate genes is important, however non-coding elements such as miRNAs might also play a role in many inherited traits (Davis *et al.* 2005).

It is important to note that even though the extent of LD between these two breeds is similar, implementation of marker assisted selection based on the information from one breed cannot always be used for the other. In some cases, two markers at the same distance can show similar r^2 values in different breeds,

but can be in different LD phase. For example: BTA-113824 and BTA-113826 have r^2 value of 0.988 in Holstein and 0.923 for Angus. In order to verify if the same phase of LD between markers persisted for both breeds, the correlation of r values was calculated including all the same markers genotyped on both breeds. In order for markers to be in the same LD phase in both breeds, the r statistic has to be the same (value and sign) in both breeds (Goddard *et al.* 2006). Correlation of r statistic between Holstein and Angus indicates that a high correlation persists up to 10 kbp (Figure 4.4), agreeing with results from Goddard *et al.* (2006). This is not surprising since LD phase is less likely to be preserved between different breeds for longer distances. Therefore, careful examination of linkage disequilibrium measurement is necessary before applying genomic selection using the same SNP markers across these breeds.

Identification of haplotype blocks can be very useful in planning for association studies. The idea of selecting the minimum number of SNPs that define a particular haplotype of interest has been widely used in human genetics (Zhang *et al.* 2002; Barrett *et al.* 2005; Consortium. 2005; Hinds *et al.* 2005; Zhang *et al.* 2005; Pe'er *et al.* 2006). Together, haplotype blocks and SNP tagging focus on reducing the number of SNPs required for future association studies; thereby decreasing the cost associated with genotypes without the loss of precision in those studies. Using the tagger option (de Bakker *et al.* 2005) incorporated in HAPLOVIEW (Barrett *et al*. 2005), 410 SNP markers were tagged in Holstein and 420 in Angus (Appendix Two). Briefly, this procedure defines a threshold for r^2 (default: 0.8) and SNPs tagged have LD measure higher

than the threshold set. Of the total number tagged, 304 markers are common to both breeds. Using this approach, Hayes *et al.* (2006) identified sites of preferential recombination when evaluating SNPs in four casein genes in goat milk. They were able to tag 11 SNPs that form part of different haplotypes, thereby reducing the cost of haplotype assisted selection (HAS) while identifying specific haplotypes associated with protein and fat percentage as well as milk volume.

Minor allele frequencies (MAFs) plotted against marker distances were used to observe any trends in decreased MAF. Such regions can indicate areas where alleles are reaching fixation, possibly because of selective pressure. In Holstein, acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*) lysine variant has been increasingly selected for in this breed due to its association with increased milk fat % (Grisart *et al.* 2004). This frequency can vary between populations depending on the breeding goal implemented (high or low milk fat %) (Weller *et al.* 2003). Using human coordinates from Marques *et al.* (2007), the region between SNPs BTA-35050 and BTA-35941 were shown to be flanking the location of *DGAT1*. Calculation of MAFs in this region showed an average MAF equal to 0.43 (Appendix Two). When analyzing nearby regions, a small cluster of low MAF SNPs is observed 7400 kbp away (Figure 4.5). Considering our average estimate of LD reaching background levels (r^2 at approximately 0.1) at 500 kbp inter-marker distance, it is unlikely that these particular SNPs are in high LD with alleles from *DGAT1*; thereby implying that a higher density set of markers is needed in this region in order to make conclusions regarding the allele frequency

trends around *DGAT1*. Screening the Angus breed for obvious signs of low MAF, approximately at the 30 Mbp region showed a cluster of low MAF SNPs (Figure 4.6). These SNPs are located approximately 0.5 Mbp from a region of BTA14 where a carcass weight QTL has been detected (Mizoshita *et al.* 2005).

In order to evaluate and compare the extent of LD for a candidate region between both Holstein and Angus animals, the extended haplotype homozygosity (EHH) approach (Sabeti *et al.* 2002) was used. Analyzing the extent of LD decay at various distances away from a specific candidate region can give insights into the selection histories of populations (Sabeti *et al.* 2006). Basically, the EHH of an unselected allele increased to a specific frequency under neutrality will be different from the EHH of a selected allele raised to the same frequency under selection pressure. The method analyses the relationship between the allele's frequency and the extent of linkage disequilibrium surrounding it. A similar approach has recently been used in studies of signatures of selection in humans population, looking for candidate genes involved in different local adaptations (Voight *et al.* 2006). Haplotypes with long range LD and with high frequency signify a recent positive selection or population bottlenecks (Sabeti *et al.* 2002). The challenge is to determine whether the signature is due to selection or effects of population demography (Sabeti *et al.* 2006). However, regardless of the LD causes, estimating and analyzing LD within a candidate region using appropriate algorithms can indicate selection on genes within this particular region. Segments of the chromosome where selected alleles are located will increase in frequency in a specific population as these selected alleles are pressured to reach fixation.

In our analysis, the *DGAT1* region was selected for EHH analysis and for comparison between the two breeds. In Holsteins, the second highest frequency haplotype, 33.3 % (AA) showed the highest EHH when plotted up to 10 Mbp from the candidate region (Figure 4.7). Another haplotype, AC with a frequency of 15.1%, showed steady EHH values up until approximately 4 Mbp from the candidate region and consistently declined reaching EHH values under the AA haplotype. Within this same region, approximately 1.5 Mbp from *DGAT1*, lies *CYP11B1*, another gene linked to milk production traits in dairy cattle (Kaupe *et al.* 2007)**.** EHH analysis on Angus using SNPs in the same region showed little extended LD away from the candidate region (Figure 4.8). Haplotype AA, with frequency of 61.3 %, showed declining EHH values after approximately 600 kbp away from the candidate region. EHH plots can be used to evaluate not only potential regions showing extended long range LD, but also long range LD between two gene variants, as shown with *DGAT1* in Grisart *et al.* (2004). In this case, EHH values for the fat increasing haplotype (lysine allele) was consistently higher than for the alanine variant.

Table 4.1 Linkage disequilibrium summary statistics for Holstein markers used to plot Figure 4.1

Category	Number of markers	Mean	Standard Deviation	Min	Max
1kb	58	0.687	0.375	0.005	1
5kb	29	0.548	0.403	0.007	1
50kb	163	0.328	0.351	0	1
100kb	170	0.252	0.278	0	1
500kb	1260	0.123	0.171	0	1
1Mb	1383	0.083	0.120	0	0.939
5Mb	11039	0.063	0.092	0	0.871
10Mb	111649	0.013	0.027	0	0.77

Table 4.2 Linkage disequilibrium summary statistics for Angus markers used to plot Figure 4.1					
Category	**Number of markers**	**Mean**	**Standard Deviation**	**Min**	**Max**
1kb	73	0.649	0.367	0.001	1
5kb	39	0.487	0.399	0.001	1
50kb	174	0.318	0.347	0	1
100kb	181	0.219	0.299	0	1
500kb	1401	0.098	0.147	0	1
1Mb	1436	0.076	0.108	0	0.801
5Mb	11057	0.038	0.060	0	0.737
10Mb	114925	0.012	0.018	0	0.434

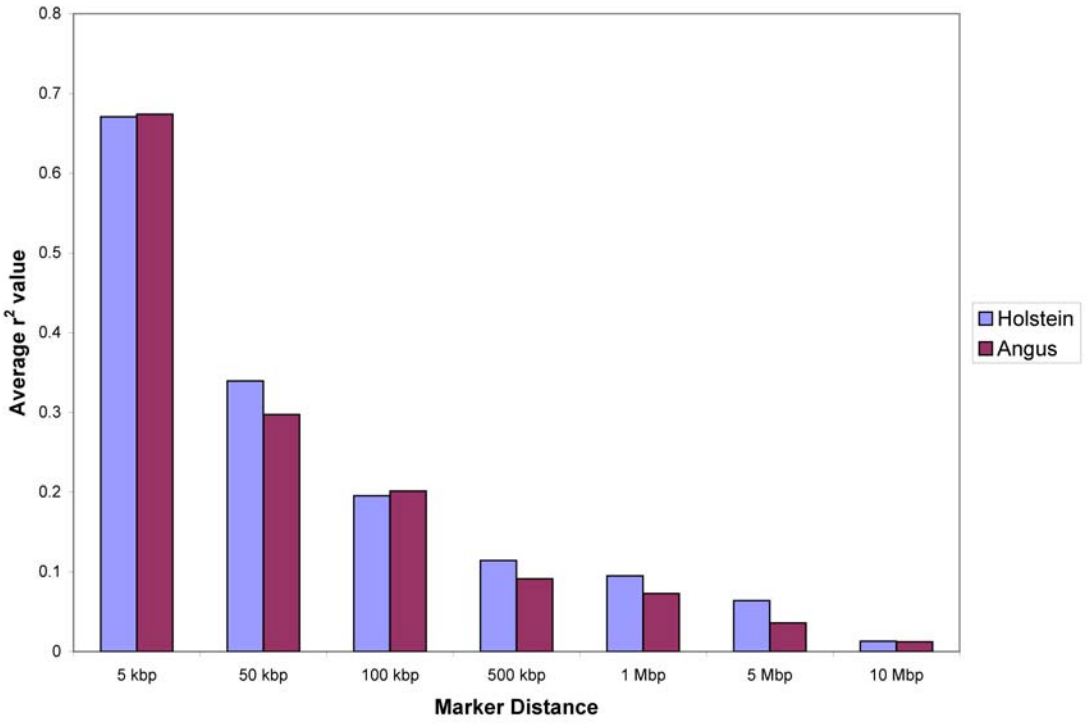

Figure 4.1 Bovine chromosome 14 (BTA14) marker detail. Average r^2 value for different marker distances (kbp) using 509 SNPs on Angus and 502 SNPs on Holstein animals.

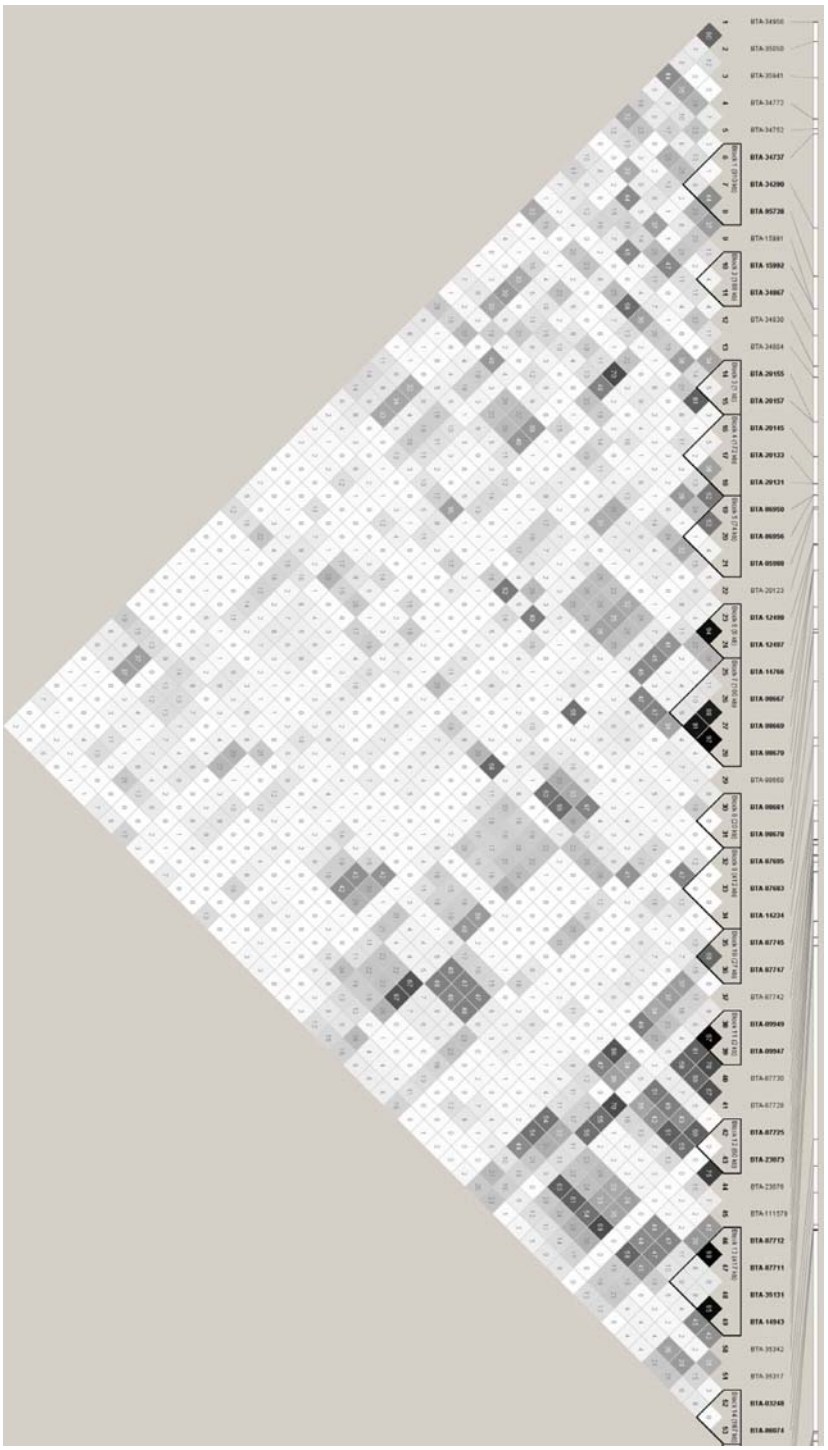

Figure 4.2 Linkage disequilibrium (LD) map for Holstein cattle. LD map of 502 SNP markers on Holstein cattle created using HAPLOVIEW (Barrett *et al.* 2005). For legibility purposes, only the first 53 markers are represented. The dark squares represent high r^2 values and triangles surrounding markers represent haplotype blocks under the four gamete rule (Wang *et al.* 2002). A complete list of haplotype blocks is in Additional file 2.

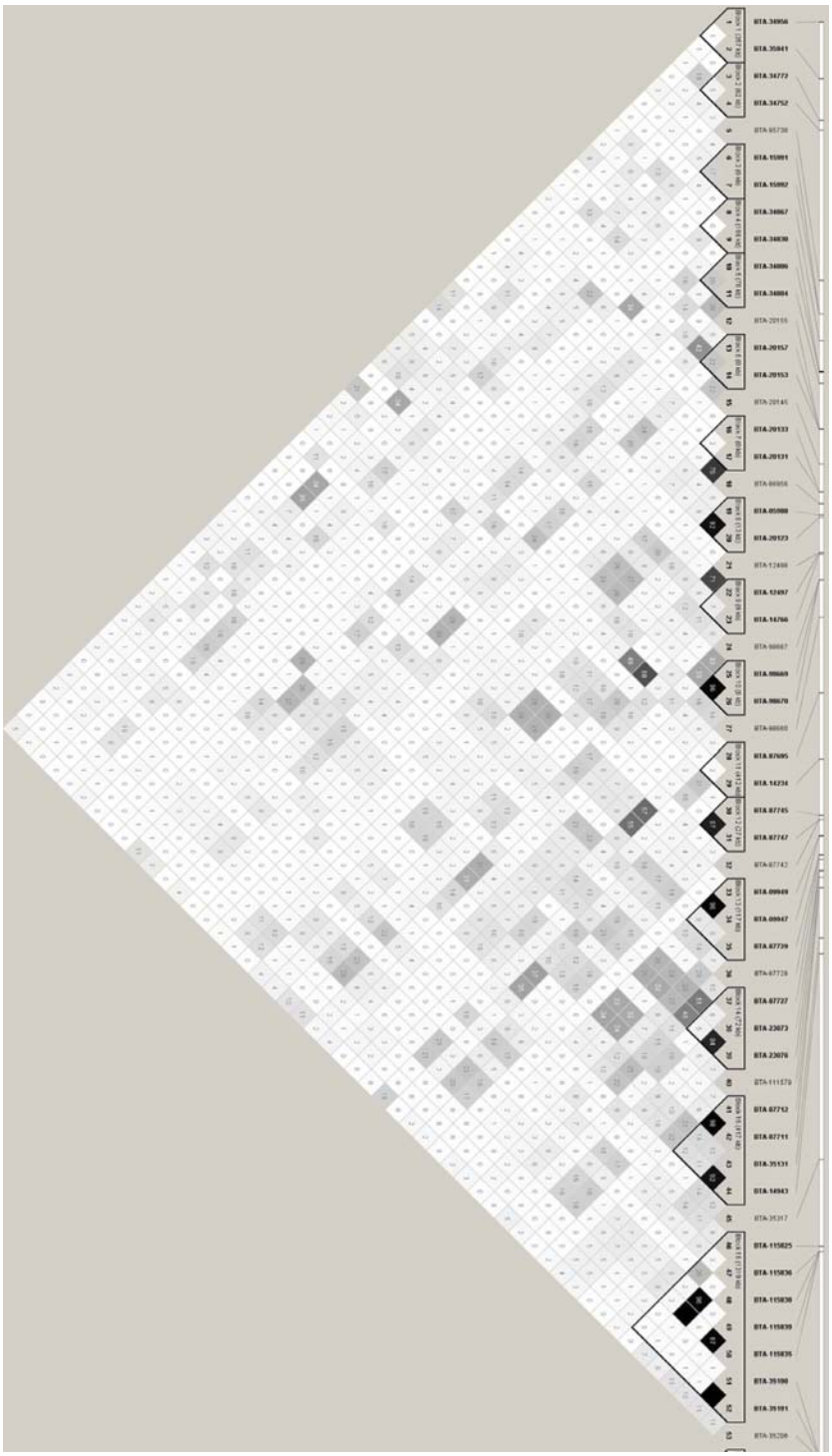

Figure 4.3 Linkage disequilibrium (LD) map for Angus cattle.LD map of 509 SNP markers on Angus cattle created using HAPLOVIEW (Barrett *et al.* 2005). For legibility purposes only the first 53 markers are represented. Dark squares represent high r^2 values and triangles surrounding markers represent haplotype blocks under the four gamete rule (Wang *et al.* 2002). A complete list of haplotype blocks is in Additional file 2.

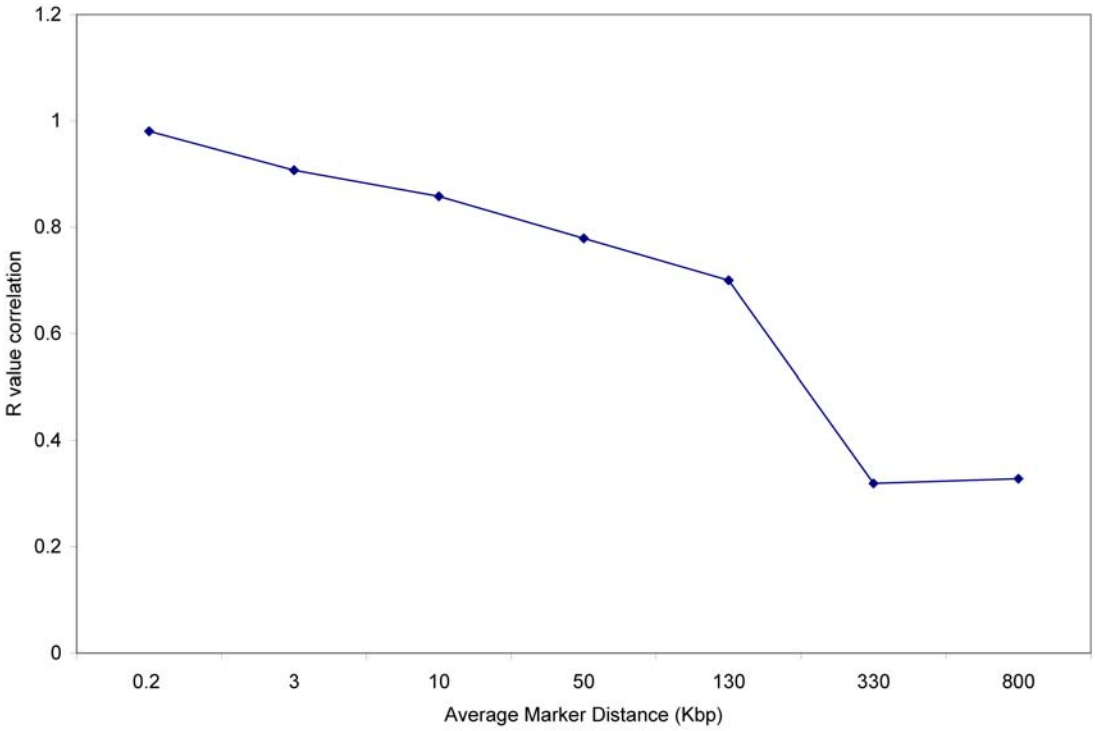

Figure 4.4 Graph depicting the correlation of r-value for Holstein and Angus cattle. Correlation of r-values between Holstein and Angus using 419 markers genotyped on both breeds. Values are plotted against average marker distances (kbp).

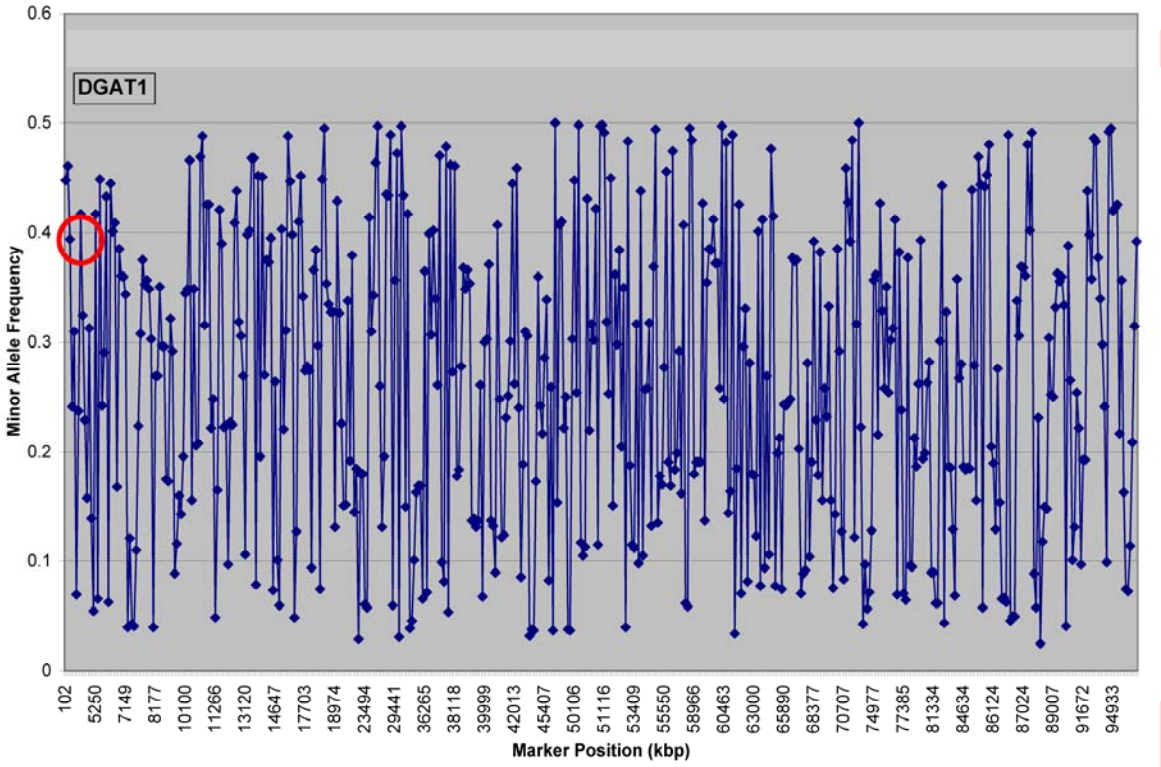

Figure 4.5 Minor Allele Frequency (MAF) for 502 SNPs genotyped on Holstein. MAFs were plotted against marker positions (kbp) on bovine chromosome 14. Red circle depicts the position of acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*).

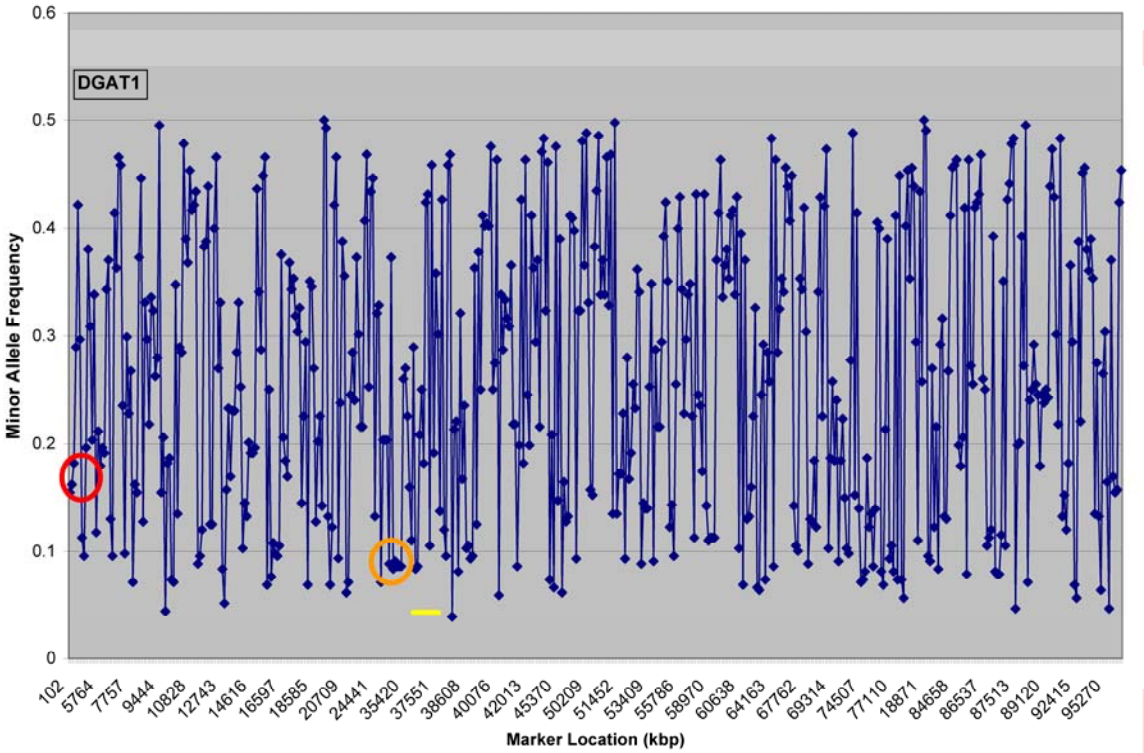

Figure 4.6 Minor Allele Frequency (MAF) for 509 SNPs genotyped on Angus. MAFs were plotted against marker positions (kbp) on bovine chromosome 14. Red circle depicts the position of acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*). Orange circle represents the region of low MAF near a previously identified carcass weight QTL (Mizoshita *et al.* 2005) represented by a yellow line.

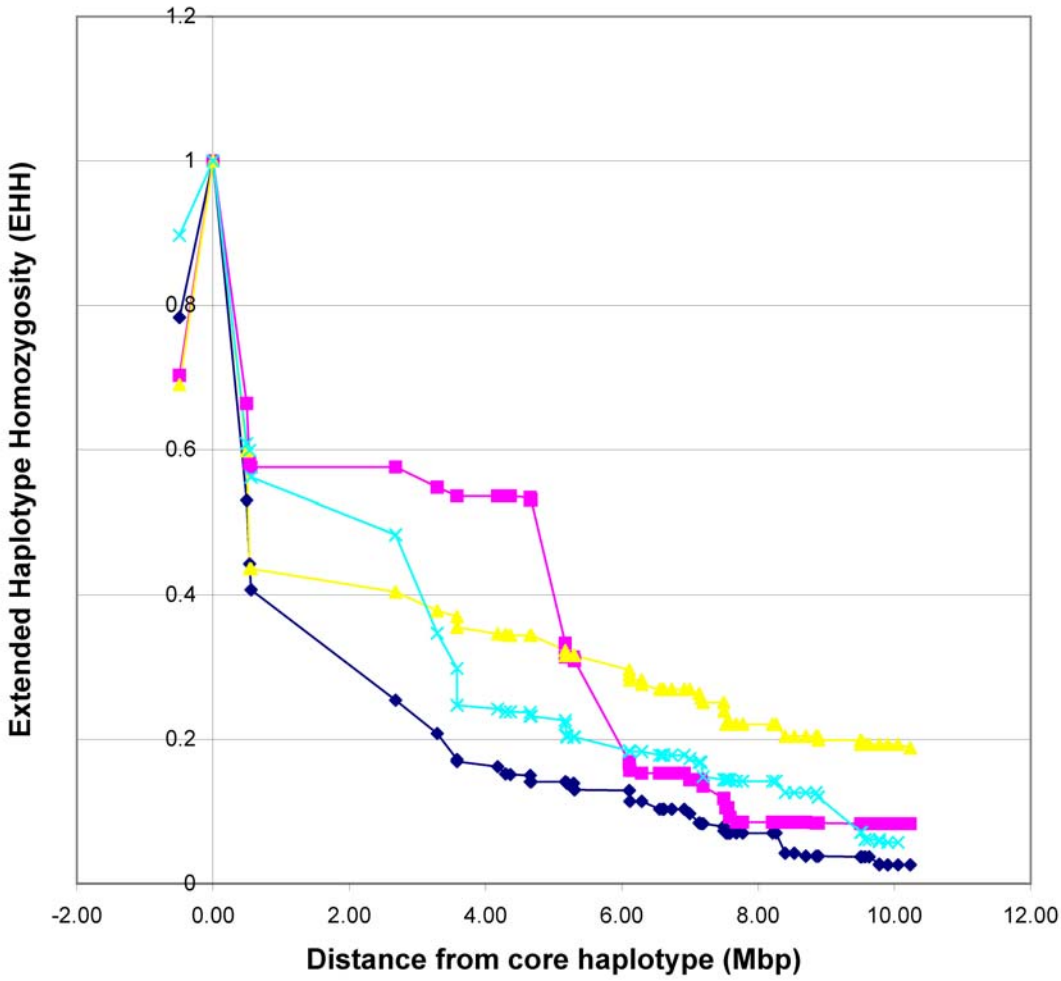

Figure 4.7 Extended haplotype homozygosity (EHH) graph for Holstein cattle. EHH values in Holstein evaluating the decay of LD on either side of the core haplotypes. Values plotted as a function of increasing marker distance. Markers BTA-35050 and BTA-35941 were used to make up the candidate region near acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*). Marker positions are represented in mega base pair (Mbp). The pink plot represents haplotype AC with 15.1% frequency. The blue plot represents haplotype GA with 16.9% frequency. The yellow plot represents haplotype AA with 33.7% frequency. The light blue plot represents haplotype GC with 34.3% frequency.

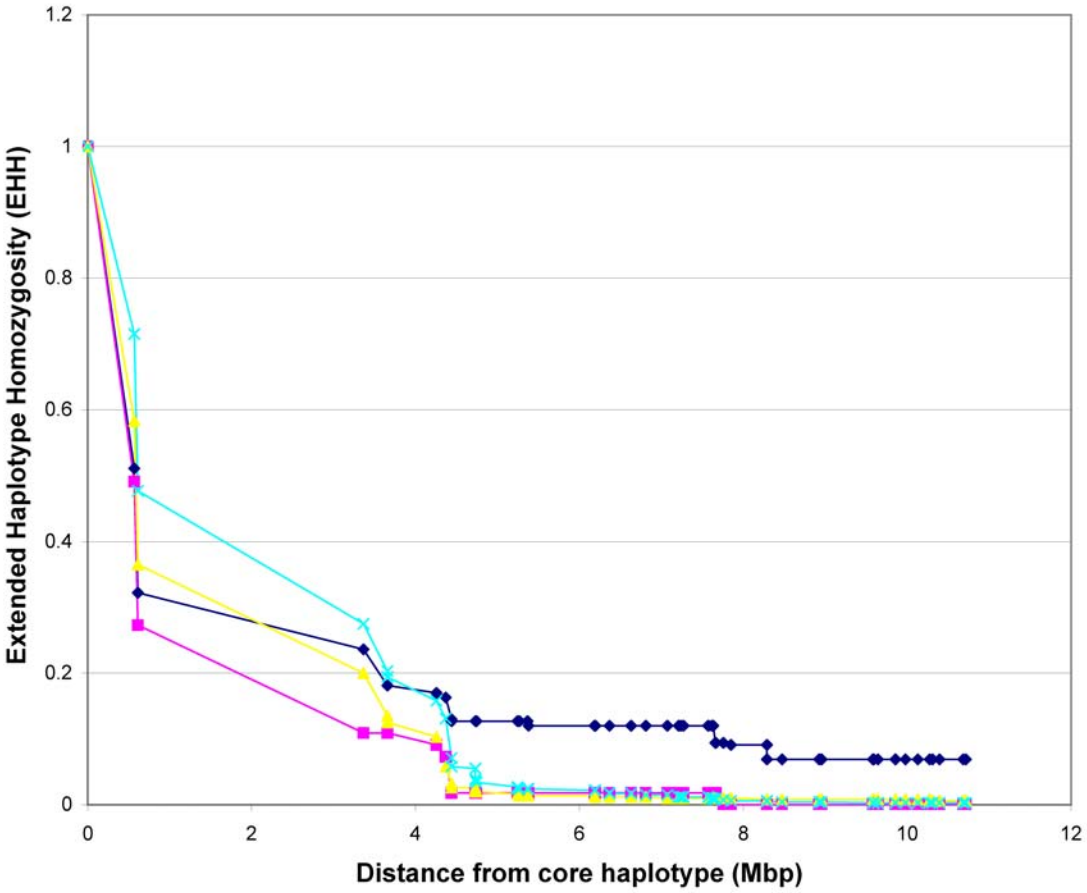

Figure 4.8 Extended haplotype homozygosity (EHH) graph for Angus cattle. EHH values in Angus evaluating the decay of LD on either side of the core haplotypes. Values are plotted as a function of increasing marker distance on Angus cattle. Markers BTA-34956 and BTA-35941 were used to make up the candidate region near acyl-CoA:diacylglycerol acyltransferase 1 (*DGAT1*). Marker positions are represented in mega base pair (Mbp). The pink plot represents haplotype CC with 6.41% frequency. The blue plot represents haplotype CA with 13.9% frequency. The yellow plot represents haplotype AC with 18.5 % frequency. The light blue plot represents haplotype AA with 61.3 % frequency.

4.4 Literature Cited

American Angus Association. [http://www.angus.org/pr/pr_main.html].

Bovine Sequence database. [ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/].

Holstein Canada. [http://www.holstein.ca/english/AnimalInq/animalinq.asp].

Abecasis G.R. & Cookson W.O. (2000) GOLD--graphical overview of linkage disequilibrium. *Bioinformatics* **16**, 182-3.

Barendse W.J. (1999) Assessing lipid metabolism. *Patent, International Publication Number: WO 99/23248. World International Property Organization.*

Barrett J.C., Fry B., Maller J. & Daly M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-5.

Consortium. I.H. (2005) A haplotype map of the human genome. *Nature* **437**, 1299-320.

Daly M.J., Rioux J.D., Schaffner S.F., Hudson T.J. & Lander E.S. (2001) High-resolution haplotype structure in the human genome. *Nat Genet* **29**, 229-32.

Davis E., Caiment F., Tordoir X., Cavaille J., Ferguson-Smith A., Cockett N., Georges M. & Charlier C. (2005) RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr Biol* **15**, 743-9.

de Bakker P.I., Yelensky R., Pe'er I., Gabriel S.B., Daly M.J. & Altshuler D. (2005) Efficiency and power in genetic association studies. *Nat Genet* **37**, 1217-23.

Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D. & Georges M. (2000) Extensive genome-wide linkage disequilibrium in cattle. *Genome Res* **10**, 220-7.

Gabriel S.B., Schaffner S.F., Nguyen H., Moore J.M., Roy J., Blumenstiel B., Higgins J., DeFelice M., Lochner A., Faggart M., Liu-Cordero S.N., Rotimi C., Adeyemo A., Cooper R., Ward R., Lander E.S., Daly M.J. & Altshuler D. (2002) The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9.

Gautier M., Faraut T., Moazami-Goudarzi K., Navratil V., Foglio M., Grohs C., Boland A., Garnier J.G., Boichard D., Lathrop M., Gut I. & Eggen A. (2007) Genetic and haplotypic structure in 14 European and African cattle breeds. *Genetics*.

Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R., Pasquino A.T., Sargeant L.S., Sorensen A., Steele M.R., Zhao X. & et al. (1995) Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* **139**, 907-20.

Goddard M.E., Hayes B., Chamberlain A. & Mcpartlan H. (2006) Can the same markers be used in multiple breeds? *8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Brazil.*

Greenspan G. & Geiger D. (2006) Modeling haplotype block variation using Markov chains. *Genetics* **172**, 2583-99.

Grisart B., Farnir F., Karim L., Cambisano N., Kim J.J., Kvasz A., Mni M., Simon P., Frere J.M., Coppieters W. & Georges M. (2004) Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc Natl Acad Sci U S A* **101**, 2398-403.

Hastings A. (1985) Multilocus population genetics with weak epistasis. I. Equilibrium properties of two-locus two-allele models. *Genetics* **109**, 799-812.

Hayes B., Hagesaether N., Adnoy T., Pellerud G., Berg P.R. & Lien S. (2006) Effects on production traits of haplotypes among casein genes in Norwegian goats and evidence for a site of preferential recombination. *Genetics* **174**, 455-64.

Hinds D.A., Stuve L.L., Nilsen G.B., Halperin E., Eskin E., Ballinger D.G., Frazer K.A. & Cox D.R. (2005) Whole-genome patterns of common DNA variation in three human populations. *Science* **307**, 1072-9.

Jonker J.W., Merino G., Musters S., van Herwaarden A.E., Bolscher E., Wagenaar E., Mesman E., Dale T.C. & Schinkel A.H. (2005) The breast cancer resistance protein BCRP (ABCG2) concentrates drugs and carcinogenic xenotoxins into milk. *Nat Med* **11**, 127-9.

Kaupe B., Brandt H., Prinzenberg E.M. & Erhardt G. (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. *J Anim Sci* **85**, 11-21.

Khatkar M.S., Zenger K.R., Hobbs M., Hawken R.J., Cavanagh J.A., Barris W., McClintock A.E., McClintock S., Thomson P.C., Tier B., Nicholas F.W. & Raadsma H.W. (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**, 763-72.

Lindblad-Toh K., Wade C.M., Mikkelsen T.S., Karlsson E.K., Jaffe D.B., Kamal M., Clamp M., Chang J.L., Kulbokas E.J., 3rd, Zody M.C., Mauceli E., Xie X., Breen M., Wayne R.K., Ostrander E.A., Ponting C.P., Galibert F., Smith D.R., DeJong P.J., Kirkness E., Alvarez P., Biagi T., Brockman W., Butler J., Chin C.W., Cook A., Cuff J., Daly M.J., DeCaprio D., Gnerre S., Grabherr M., Kellis M., Kleber M., Bardeleben C., Goodstadt L., Heger A., Hitte C., Kim L., Koepfli K.P., Parker H.G., Pollinger J.P., Searle S.M., Sutter N.B., Thomas R., Webber C., Baldwin J., Abebe A., Abouelleil A., Aftuck L., Ait-Zahra M., Aldredge T., Allen N., An P., Anderson S., Antoine C., Arachchi H., Aslam A., Ayotte L., Bachantsang P., Barry A., Bayul T., Benamara M., Berlin A., Bessette D., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L., Bonnet C., Boukhgalter B., Brown A., Cahill P., Calixte N., Camarata J., Cheshatsang Y., Chu J., Citroen M., Collymore A., Cooke P., Dawoe T., Daza R., Decktor K., DeGray S., Dhargay N., Dooley K., Dooley K., Dorje P., Dorjee K., Dorris L., Duffey N., Dupes A., Egbiremolen O., Elong R., Falk J., Farina A., Faro S., Ferguson D., Ferreira P., Fisher S., FitzGerald M., Foley K., Foley C., Franke A., Friedrich D., Gage D., Garber M., Gearin G., Giannoukos G.,

Goode T., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N., Hagopian D., Hagos B., Hall J., Healy C., Hegarty R., Honan T., Horn A., Houde N., Hughes L., Hunnicutt L., Husby M., Jester B., Jones C., Kamat A., Kanga B., Kells C., Khazanovich D., Kieu A.C., Kisner P., Kumar M., Lance K., Landers T., Lara M., Lee W., Leger J.P., Lennon N., Leuper L., LeVine S., Liu J., Liu X., Lokyitsang Y., Lokyitsang T., Lui A., Macdonald J., Major J., Marabella R., Maru K., Matthews C., McDonough S., Mehta T., Meldrim J., Melnikov A., Meneus L., Mihalev A., Mihova T., Miller K., Mittelman R., Mlenga V., Mulrain L., Munson G., Navidi A., Naylor J., Nguyen T., Nguyen N., Nguyen C., Nguyen T., Nicol R., Norbu N., Norbu C., Novod N., Nyima T., Olandt P., O'Neill B., O'Neill K., Osman S., Oyono L., Patti C., Perrin D., Phunkhang P., Pierre F., Priest M., Rachupka A., Raghuraman S., Rameau R., Ray V., Raymond C., Rege F., Rise C., Rogers J., Rogov P., Sahalie J., Settipalli S., Sharpe T., Shea T., Sheehan M., Sherpa N., Shi J., Shih D., Sloan J., Smith C., Sparrow T., Stalker J., Stange-Thomann N., Stavropoulos S., Stone C., Stone S., Sykes S., Tchuinga P., Tenzing P., Tesfaye S., Thoulutsang D., Thoulutsang Y., Topham K., Topping I., Tsamla T., Vassiliev H., Venkataraman V., Vo A., Wangchuk T., Wangdi T., Weiand M., Wilkinson J., Wilson A., Yadav S., Yang S., Yang X., Young G., Yu Q., Zainoun J., Zembek L., Zimmer A. & Lander E.S. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803-19.

Marques E., De Givry S., Stothard P., Murdoch B., Wang Z., Womack J. & Moore S. (2007) A high resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. *BMC Genomics* **8**, 254.

McKay S.D., Schnabel R.D., Murdoch B.M., Matukumalli L.K., Aerts J., Coppieters W., Crews D., Dias Neto E., Gill C.A., Gao C., Mannen H., Stothard P., Wang Z., Van Tassell C.P., Williams J.L., Taylor J.F. & Moore S.S. (2007) Whole genome linkage disequilibrium maps in cattle. *BMC Genet* **8**, 74.

Mizoshita K., Takano A., Watanabe T., Takasuga A. & Sugimoto Y. (2005) Identification of a 1.1-Mb region for a carcass weight QTL on bovine Chromosome 14. *Mamm Genome* **16**, 532-7.

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C. & Benkel B. (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. *J Anim Sci* **81**, 1919-25.

Mueller J.C. & Andreoli C. (2004) Plotting haplotype-specific linkage disequilibrium patterns by extended haplotype homozygosity. *Bioinformatics* **20**, 786-7.

Olsen H.G., Nilsen H., Hayes B., Berg P.R., Svendsen M., Lien S. & Meuwissen T. (2007) Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. *BMC Genet* **8**, 32.

Pe'er I., de Bakker P.I., Maller J., Yelensky R., Altshuler D. & Daly M.J. (2006) Evaluating and improving power in whole-genome association studies using fixed marker sets. *Nat Genet* **38**, 663-7.

Sabeti P.C., Reich D.E., Higgins J.M., Levine H.Z., Richter D.J., Schaffner S.F., Gabriel S.B., Platko J.V., Patterson N.J., McDonald G.J., Ackerman H.C., Campbell S.J., Altshuler D., Cooper R., Kwiatkowski D., Ward R. & Lander E.S. (2002) Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832-7.

Sabeti P.C., Schaffner S.F., Fry B., Lohmueller J., Varilly P., Shamovsky O., Palma A., Mikkelsen T.S., Altshuler D. & Lander E.S. (2006) Positive natural selection in the human lineage. *Science* **312**, 1614-20.

Snelling W.M., Chiu R., Schein J.E., Hobbs M., Abbey C.A., Adelson D.L., Aerts J., Bennett G.L., Bosdet I.E., Boussaha M., Brauning R., Caetano A.R., Costa M.M., Crawford A.M., Dalyrmple B.P., Eggen A., Everts-van der Wind A., Floriot S., Gautier M., Gill C.A., Green R.D., Holt R., Jann O., Jones S.J., de Jong P.J., Kappes S.M., Keele J.W., Larkin D.M., Lewin H.A., McEwan J.C., McKay S., Marra M.A., Mathewson C.A., Matukumalli L.K., Moore S.S., Murdoch B., Nicholas F., Osoegawa K., Roy A., Salih H., Schibler L., Schnabel R., Silveri L., Skow L.C., Smith T.P., Sonstegard T.S., Taylor J., Tellam R., Van Tassell C.P., Williams J.L., Womack J.E., Wye N.H., Yang G. & Zhao S. (2007) A physical map of the bovine genome. *Genome Biol* **8**, R165.

Tenesa A., Knott S.A., Ward D., Smith D., Williams J.L. & Visscher P.M. (2003) Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes. *J Anim Sci* **81**, 617-23.

Thallman R.M., Bennett G.L., Keele J.W. & Kappes S.M. (2001a) Efficient computation of genotype probabilities for loci with many alleles: I. Allelic peeling. *J Anim Sci* **79**, 26-33.

Thallman R.M., Bennett G.L., Keele J.W. & Kappes S.M. (2001b) Efficient computation of genotype probabilities for loci with many alleles: II. Iterative method for large, complex pedigrees. *J Anim Sci* **79**, 34-44.

Vallejo R.L., Li Y.L., Rogers G.W. & Ashwell M.S. (2003) Genetic diversity and background linkage disequilibrium in the North American Holstein cattle population. *J Dairy Sci* **86**, 4137-47.

Varilo T., Paunio T., Parker A., Perola M., Meyer J., Terwilliger J.D. & Peltonen L. (2003) The interval of linkage disequilibrium (LD) detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Hum Mol Genet* **12**, 51-9.

Voight B.F., Kudaravalli S., Wen X. & Pritchard J.K. (2006) A map of recent positive selection in the human genome. *PLoS Biol* **4**, e72.

Wang N., Akey J.M., Zhang K., Chakraborty R. & Jin L. (2002) Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am J Hum Genet* **71**, 1227-34.

Weir B.S. & Hill W.G. (1980) Effect of mating structure on variation in linkage disequilibrium. *Genetics* **95**, 477-88.

Weller J.I., Golik M., Seroussi E., Ezra E. & Ron M. (2003) Population-wide analysis of a QTL affecting milk-fat production in the Israeli Holstein population. *J Dairy Sci* **86**, 2219-27.

Zhang K., Deng M., Chen T., Waterman M.S. & Sun F. (2002) A dynamic programming algorithm for haplotype block partitioning. *Proc Natl Acad Sci U S A* **99**, 7335-9.

Zhang K., Qin Z., Chen T., Liu J.S., Waterman M.S. & Sun F. (2005) HapBlock: haplotype block partitioning and tag SNP selection software using a set of dynamic programming algorithms. *Bioinformatics* **21**, 131-4.

CHAPTER 5

Identification of polymorphisms on bovine chromosome 14 affecting meat quality in beef cattle

5.1 Introduction

A number of quantitative trait loci (**QTL**) affecting meat quality traits have been identified to date (Stone *et al.* 1999; MacNeil & Grosz 2002; Casas *et al.* 2003; Moore *et al.* 2003; Casas *et al.* 2004). Identification of additional markers is crucial for narrowing down a number of these QTL. It is not unusual to come across studies that used a few hundred markers for an entire genome scan (Gutierrez-Gil *et al.* 2008).

As the number of polymorphisms identified in the bovine genome increases, so does the issues of selecting the highest quality markers for a QTL scan. There is no doubt that the increase in markers will aid in narrowing down the QTL, but computational time limits the amount of markers that can be used. The advantage of selecting these markers from a pool of genotyped ones comes from the idea that markers can be selected on the basis of sire heterozygosity and linkage disequilibrium information. Selecting them on the basis of heterozygosity ensures that most of the sires are segregating for the particular marker. This way any QTL detected across family will likely not be skewed due to appearance in one family and not the others.

High density SNP panels have the advantage of allowing for further evaluation of marker-marker relationships. Using publicly available software such as HAPLOVIEW Barrett *et al.* (2005), one can analyze the specific amount of

linkage disequilibrium between markers. As a result, any pairs of markers showing high levels of LD can have one representative from that group for inclusion in a QTL scan. If markers are in high LD, then essentially the same effect is being included in the scan. Eliminating redundant polymorphisms and selecting for marker informativeness can increase the power of detecting significant QTL. Chapter 4 examined the use of over 500 SNP markers on bovine chromosome 14 to characterize the pattern of LD along this chromosome for both Holstein and Angus breeds. This procedure can essentially be used in any genome wide analysis where markers are only base pairs apart.

Another issue arising in the era of high density marker sets is the possibility of using those markers in commercial DNA tests without further validation and disregard for conflicting reports. QTL affecting meat quality in beef cattle has been widely reported in several populations (Stone *et al.* 1999; MacNeil & Grosz 2002; Casas *et al.* 2004) and it is population specific. It is not unusual to read conflicting reports on associations between polymorphisms and economically relevant traits. For instance, Barendse (1999) reported a polymorphism in the promoter region of thyroglobulin gene associated with marbling. This polymorphism, however, does not show significant associations with marbling in other studies (Moore *et al.* 2003; Casas *et al.* 2004; Casas *et al.* 2007). Another example comes from *ABCG2* and *OPN* genes. This example is conflicting because at one point, both *OPN* (Schnabel *et al.* 2005) and *ABCG2* (Cohen-Zinder *et al.* 2005) were reported to harbor the causative mutation for a milk QTL on BTA6. It wasn't until later that (Olsen *et al.* 2007) established

genetic support for the polymorphism in *ABCG2*. Regardless of the outcome from independent studies, validation of the identified polymorphisms needs to be evaluated in other populations before commercial utilization. The objective of this study was to identify positional candidate markers affecting meat quality in our experimental beef cattle and to validate these markers in another beef cattle population.

5.2 Materials and Methods

5.2.1 Animals and Management

Four hundred and sixty four steers from twenty eight half sib families from an experimental line of Angus, Charolais or Alberta Hybrid Bulls and the University of Alberta's Hybrid dam line previously described by (Nkrumah *et al.* 2004) were used in this study. The animals test diets were the same for years 2 and 3, but differed in year 1 with the substitution of Barley and oat grain with dry-rolled corn due to shortage of feed barley that particular year, however both diets contained a similar ME content, Briefly, the test diet for year 1 contained 80% dry-rolled corn, 13.5% alfalfa hay pellet, 5% feedlot supplement (32% CP beef supplement) and 1.5% canola oil. Year two and three diets contained 64% barley grain, 20% oat grain, 9% alfalfa hay pellet, 5% beef feedlot supplement and 1.5% canola oil. Details of the animals' diets have been described in (Nkrumah *et al.* 2004). Animals used in the study were cared for according to the guidelines of the Canadian Council on Animal Care (Canadian Council on Animal Care 1993).

5.2.2 Traits Analyzed

Ultrasound and carcass merit data were collected on beef steers over a period of 3 yr (November 2002 to June 2005). Carcass traits were evaluated according to the Canadian beef carcass grading system (Agriculture Canada 1992). Carcass and ultrasound measurements have been previously described by (Nkrumah *et al.* 2004). Briefly, ultrasound measurements of 12th-/13th-rib fat depth (**UBF**), longissimus muscle area (**ULMA**) and marbling score (Lindblad-Toh *et al.*)were obtained with an Aloka 500V real-time ultrasound with a 17-cm, 3.5-MHz linear array transducer at 28-d intervals according to procedures described by (Brethour 1992). After these tests, animals were shipped to a commercial plant and carcass grade fat (**GRFAT**), carcass backfat (**CBF**), longissimus muscle area (**LMA**) and carcass marbling (**CMAR**) measurements were collected at the 12 th/13th rib following a 24-h chill at -4 °C. Ultrasound and carcass marbling score are a measure of intramuscular fat being classified as 1 to < 2 units = trace marbling (Canada A quality grade); 2 to < 3 units = slight marbling (Canada AA quality grade); 3 to < 4 units = small to moderate marbling (Canada AAA quality grade) and >= 4 units = slightly abundant or more marbling (Canada Prime). Lean meat yield (**LMY**) is an estimate of saleable meat calculated according to (Jones 1984). Yield grade (**YGRADE**) Classes are based on the proportion of lean meat and is classified as YGRADE1 = >59 %, YGRADE2 = 54 to 59% and YGRADE3 = <54%.

5.2.3 Compilation of SNP markers and Genotyping on BTA14.

SNPs included in this analysis were compiled and selected from Baylor College of Medicine bovine database publicly available at

ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus. Additional SNPs were derived from bac end sequences (BES) of the CHORI-240 library (http://bacpac.chori.org/bovine240.htm) and IBISS (Interactive Bovine In Silico SNP) database (http://www.livestockgenomics.csiro.au/ibiss). Haplotypes for all animals and their sires were generated using fastPHASE (Scheet & Stephens 2006). Once determined, only sire haplotypes were loaded onto HAPLOVIEW in order to identify markers with observed heterozygosity greater than 0.5. The settings used included: min genotyped %: 50, Hardy Weinberg (HW) p-value cutoff: 0.0010, Minimum minor allele freq: 0.0010, Maximum # Mendel errors: 1. When all animals' haplotypes were loaded, only markers showing heterozygosity greater then 0.5 were selected and run through the tagger option incorporated into HAPLOVIEW. This procedure defines a threshold for r^2 (default: 0.8) and SNPs tagged have LD measure higher than the threshold set. This method yielded 96 SNP markers to be used for QTL analysis. Genotypes for SNPs were generated by high throughput Illumina BeadStation 500G genotyping system and analyzed using Illumina's GenCall Software (version 1.014).

5.2.4 SNP marker location.

Marker positions were inferred using the marker order from the 12K RH map of Marques *et al.* (2007) which is in high agreement with the recently released physical map based on independent whole genome map of Snelling *et al.* (2007). Preliminary comparison between our marker order and the recently published Btau_4.0 shows agreement for markers common to both maps (Paul Stothard, personal communication). Relative genetic distances in centimorgans

(cM) were estimated by using the centiray positions (cR) and the estimated length of chromosome 14 (108 cM).

5.2.5 Statistical analyses

The total number of animals genotyped in our study was 464 belonging to 28 families. However, 374 animals belonging to 15 half-sib families (range 9 to 56 progeny) were used for QTL analysis. The remaining animals belonged to families with less than 2 animals and therefore were not utilized for the purpose of QTL analysis. QTL analysis performed in this study used the multiple marker interval mapping approach described by Knott *et al.* (1996). The conditional probabilities that a calf inherited the first allele of a putative QTL from its sire were obtained from QTL Express (Seaton *et al.* 2002) which uses the information from the closest informative flanking markers at 1 cM intervals. Our analysis is similar to that used by de Koning *et al.* (1999) in which the conditional probabilities of inheriting the sire allele were nested within half-sib families. This is because not only the linkage phase between a marker and a QTL can differ between families, but also because not all sires are heterozygous for the QTL. In addition, sire effects were also included as random. The conditional probabilities from QTL Express were input into SAS (SAS Inst. Inc., Cary, NC) and QTL analysis was performed using the mixed model described by:

$$Y = \mathrm{X}\beta + Gs + Q\alpha + e\,,$$

where Y is a vector of observations on the progeny of each sire, X is the known incidence matrix relating observations to their fixed effect levels, β is the vector of fixed effects (breed, test batch and age), G is the known incidence matrix

relating observations to random sire effects, s is the vector of random additive polygenic effects of sires, Q is a vector of the conditional probabilities, at each interval, that a calf inherited the first allele of a putative QTL from a sire, α is the regression coefficient corresponding to the fixed allele substitution effect for a putative QTL within half-sib families. Significance thresholds at 10% and 5% were determined using 25,000 permutation tests in SAS (SAS Inst. Inc., Cary, NC) by randomly shuffling the phenotypic records of the 374 animals and maintaining the QTL probabilities unchanged, according to the procedure described by Nkrumah *et al.* (2007). This procedure was performed every time specific SNPs were removed from the analysis.

5.2.6 Association Analysis

Associations of the genotypes for each polymorphism and carcass merit were analyzed by regressing phenotypes on genotypes using MIXED procedure (SAS Inst. Inc., Cary, NC). Four hundred and sixty-four animals were available for this analysis. The statistical analyses model included fixed effects of SNP genotype, test batch, breed and age of animal at the beginning of the test, and random effects of sire of animal. Allele substitution effect was calculated regressing phenotypes on the number of copies of one allele for each SNP.

5.2.7 False Discovery Rate

False discovery rate (FDR) minimizes false positives and it takes into consideration the number of tests performed, the ranking of the marker within the

analysis and their significance (*P*-value) rank from lowest to highest (Benjamini & Hochberg 1995; Weller *et al.* 1998). Since FDR assumes independence between traits and since these traits are correlated, FDR was calculated within each trait according to the formula:

$$FDR = \frac{n \times P(k)}{k}$$

Where, k is the individual's ranking , P(k) is the P-value for the association and n is the number of tests performed within a trait.

5.2.8 Validation

At the time of the selection, quantitative trait analysis had not yet been performed and so selection of SNPs was based on allele substitution analysis previously performed using the sire model reaching 5 % significance threshold. Validation process consisted of genotyping 18 selected SNPs on a panel of 1000 beef animals from the University of Guelph. The genotyping assay was developed using the MassARRAY® iPLEX Gold platform technology, run on the Sequenom MassARRAY System and genotypes were provided by Sequenom Inc (San Diego). An animal model was used fitting fixed effects of contemporary group, age at end of test, breed and heterosis. An allele substitution model was used.

5.3 Results and Discussion

5.3.1 Quantitative Trait Loci Analysis

Quantitative trait loci analysis for traits affecting meat quality in our beef population resulted in 1 significant QTL for ultrasound longissimus muscle area

(ULMA, $P < 0.05$) at 2 cM and 2 putative QTL: ultrasound marbling (UMAR, $P < 0.10$) at 2 cM and ultrasound backfat (UBF, $P < 0.10$). The most likely positions for the QTL are listed in Table 5.1. No other QTL peaks were observed for the other meat quality traits studied. The analysis was also performed within-family in order to account to the differences in QTL phase and heterozygosity of sires used in the experiment. A within-family analysis allows for evaluation of sires segregating for a particular QTL which can affect the overall location of an across-family analysis. Work by Nkrumah *et al.* (2007) and Sherman *et al.* (2009) showed that family size and significance of a within family QTL analysis affected the significance and location of an across family analysis for QTL affecting residual feeding intake.

In our study, the across family QTL analysis for UMAR shows that sire 9 is the only family segregating at $P < 0.05$ for a UMAR at the 2 cM position ($P = 0.003$). Seven other families are segregating at other positions ($P < 0.15$) ranging from 1 to 103 cM, but not at 2 cM (Table 5.2). The same trend is observed for ULMA at 2 cM. Sire 13 has the lowest p-value among all families and it is segregating at position 1 cM ($P < 0.004$). The second lowest p-value ($P < 0.03$) are from sires 1 and 14, however sire 1 is segregating at 3 cM and sire 14 at 103 cM. Overall, there are seven families that show segregation in or around 2 cM, contributing to the consensus QTL at 2 cM. For UBF, the most likely QTL position was at 101 cM. Table 5.4 shows that three sires showed segregation in and around 101 cM at $P < 0.05$ and one at $P < 0.10$. Together, these families

contribute to the most likely QTL for UBF being at 101 cM, even though sire 11 shows that the most likely segregation within his family is at 26 cM ($P = 0.002$).

The phase of the QTL can also differ when comparing segregating families from the within-family analysis. The ULMA QTL analysis shows that among the families segregating in our around 2 cM, 5 families have a negative QTL effect, while 2 others have a positive QTL effect. The same trend is observed with the other traits, with at least one segregating family showing the opposite effects as the other ones (Table 5.3).

5.3.2 Positional Candidate Markers

Among the traits with reported QTL, ULMA had the highest number of polymorphisms ($P < 0.05$) under the 2 cM peak (Table 5.1), followed by both UBF (9 markers) and UMAR (9 markers). A list of all markers under the 3 QTL is presented in Table 5.6. Allele frequencies for these markers ranged from 0.01 (ss69374920:A>G) to 0.48 (ss61497130:A>G and ss61569297:A>C).

One polymorphism identified under the UMAR QTL is ss61534850:A>G, a marker that maps nearby to CYP11B1. This gene encodes a member of the cytochrome P450 superfamily of enzymes. This family of enzymes is involved in synthesis of cholesterol, steroids and other lipids (Seybert 1990). In dairy cattle, *CYP11B1* showed a positive association with fat content and negative associations with milk yield and protein yield (Kaupe *et al.* 2007). Another marker (ss61524969:A>G), mapping to approximate 2.4 cM, blasts to the protein tyrosine kinase 2 which encodes a cytoplasmic protein kinase whose activation is involved in cell growth and intracellular signal transduction pathway (Dickens *et al.* 1997).

The collagen, type XXII, alpha 1 (*COL22A1*) gene, under the UMAR QTL, encodes for the protein collagen which is the major part of the extracellular matrix. This protein has also been linked to meat quality in sheep, where its presence in the meat was correlated with a decrease in fat content (Okeudo & Moss 2004). Under the UREA QTL 3 markers map nearby RHPN1, a Rho GTPase binding protein 1. This protein has been described to play a role in cell proliferation, apoptosis and the regulation of gene expression (Seasholtz *et al.* 1999). Analysis under the UBF QTL showed 6 markers. One of the polymorphisms mapped nearby LOC785739, a hypothetical protein with no known functions to date.

5.3.3 Validation

The Validation step for any marker of commercial interest comes from genotyping and analyzing their association in other independent populations. Our validation analysis consisted of selecting 18 markers previously associated with several meat quality traits at the 5 % threshold and genotyping these markers in another beef population from the University of Guelph. Among the 18 markers and their associations, 2 did not show any association with the traits available, which also included some growth traits. Table 5.6 lists the markers selected and their associations with traits in both populations.

Among the 16 markers that show significant associations in the validation population ($P < 0.05$), 7 showed associations with other comparable meat quality traits, while the remaining showed association with growth and efficiency traits,

therefore, these markers were not counted as validated in this population. The markers and traits in bold in Table 5.6 are the ones considered validated.

Discrepancies in marker association between the two animal populations are not likely to be from the different statistical models used for the association analysis. The statistical model used by the University of Guelph was specific to account for their herds and included the use of an animal model fitting several effects including heterosis (see Materials and Methods section), while the University of Alberta population, used a sire model fitting all of the appropriate variables specific for that herd.

Table 5.1 Quantitative Trait Loci (QTL) on bovine chromosome 14 (BTA14) using Angus, Charolais and Crossbred animals across fifteen half-sib families

Trait[1]	**Location (cM)**[2]	***P*-value**	**# SNPs (P < 0.05)**[3]
UMAR	2	0.10	9
ULMA	2	0.05	12
UBF	101	0.10	9

[1] UMAR = ultrasound marbling (score), ULMA = ultrasound longissimus muscle area (cm^2), UBF = ultrasound backfat (mm).

[2] Across family QTL position

[3] Markers under quantitative trait loci (QTL) at the 5% threshold for the allele effect analysis

Table 5.2: Estimated effects for within-family QTL positions on BTA 14 for ultrasound marbling (UMAR)

	Across-Family			Within-Family			
Sire	**QTL Effect[1]**	**SE**	**P-Value**	**QTL Effect**	**SE**	**P-value**	**cM[2]**
1	0.70	0.42	0.098	0.81	0.44	0.068	9
2	-0.15	0.35	0.673	-0.57	0.50	0.255	54
3	-0.10	0.30	0.741	0.43	0.35	0.218	44
4	0.17	0.51	0.743	0.76	0.43	0.079	72
5	0.40	0.48	0.409	0.32	0.28	0.243	91
6	-0.01	0.21	0.976	0.21	0.19	0.259	104
7	0.13	0.22	0.545	0.13	0.22	0.545	2
8	-0.33	0.23	0.151	-0.34	0.24	0.149	1
9	-0.88	0.28	0.003	-0.88	0.28	0.003	2
10	-0.12	0.27	0.644	-0.34	0.24	0.169	107
11	-0.07	0.45	0.875	-0.46	0.40	0.258	98
12	0.21	0.31	0.488	0.41	0.28	0.139	10
13	-0.18	0.37	0.618	-0.45	0.30	0.146	103
14	0.48	0.41	0.241	0.50	0.41	0.220	5
15	-0.06	0.67	0.927	0.52	0.43	0.234	38
16	-0.10	0.27	0.716	0.31	0.26	0.235	73
17	-0.26	0.17	0.138	-0.27	0.17	0.129	1
18	0.39	0.21	0.072	0.39	0.21	0.062	3

[1] Estimate of QTL effect measured in score
[2] Most likely position of QTL within each individual family.

Table 5.3: Estimated effects for within-family QTL positions on BTA 14 for ultrasound longissimus muscle area (ULMA)

	Across-Family			Within-Family			
Sire	**QTL Effect[1]**	**SE**	**P-Value**	**QTL Effect**	**SE**	**P-value**	**cM[2]**
1	-9.81	4.68	0.04	-9.91	4.65	0.03	3
2	1.94	3.80	0.61	8.80	5.66	0.12	38
3	-3.81	2.59	0.14	-4.15	2.52	0.10	4
4	-9.53	5.60	0.09	-9.53	5.60	0.09	2
5	-4.20	4.65	0.37	-3.78	3.35	0.26	92
6	-3.10	1.77	0.08	-3.16	1.75	0.07	1
7	2.54	1.95	0.19	-5.03	2.85	0.08	83
8	-1.32	2.22	0.55	-1.51	2.20	0.50	49
9	3.52	2.77	0.20	-6.48	3.85	0.10	59
10	3.54	2.57	0.17	3.73	2.51	0.14	5
11	-0.71	4.87	0.88	-11.43	5.67	0.04	101
12	-1.21	3.06	0.69	-1.66	2.69	0.54	43
13	10.90	3.81	0.004	11.02	3.83	0.004	1
14	5.24	4.34	0.23	8.47	3.79	0.03	103
15	-4.02	7.53	0.59	9.22	10.26	0.37	30
16	0.65	2.66	0.81	6.11	3.06	0.05	71
17	-3.04	1.82	0.10	-3.37	1.85	0.07	4
18	-0.44	1.91	0.82	-3.39	2.60	0.19	91

[1] Estimate of QTL effect measured in cm^2

[2] Most likely position of QTL within each individual family.

Table 5.4: Estimated effects for within-family QTL positions on BTA 14 for ultrasound backfat (UBF)

	Across-Family			Within-Family			
Sire	**QTL Effect[1]**	**SE**	**P-Value**	**QTL Effect**	**SE**	**P-value**	**cM[2]**
1	0.78	1.67	0.640	-2.57	3.13	0.413	20
2	-0.52	1.15	0.649	-2.05	1.55	0.186	1
3	1.48	1.14	0.196	-2.71	1.60	0.091	63
4	-0.67	2.08	0.746	2.18	2.31	0.346	27
5	2.74	1.01	0.007	2.70	1.00	0.007	100
6	0.24	0.73	0.742	0.54	0.96	0.578	5
7	1.37	1.02	0.179	2.32	0.98	0.019	107
8	-0.30	0.73	0.683	-0.59	0.72	0.411	104
9	0.26	0.99	0.796	-2.11	1.55	0.176	19
10	-1.80	0.85	0.036	-2.04	0.84	0.016	107
11	-2.36	2.06	0.252	-5.32	1.68	0.002	26
12	0.44	1.11	0.691	1.52	1.22	0.215	24
13	-0.57	1.60	0.723	-1.44	1.13	0.203	107
14	0.43	1.21	0.720	2.13	1.84	0.253	72
15	1.39	1.27	0.271	2.71	2.15	0.211	33
16	-1.22	0.99	0.217	-1.84	1.18	0.123	5
17	-0.06	0.71	0.932	-1.16	0.78	0.138	1
18	-0.91	0.78	0.246	-1.29	0.70	0.068	107

[1]Estimate of QTL effect measured in mm
[2] Most likely position of QTL within each individual family.

Table 5.5: Summary of candidate single nucleotide polymorphism markers under 3 meat production trait Quantitative Trait Loci

SNPs	MAF[1]	Trait[2]	cM	Estimate[3]	SE	P-value	FDR[4]	Gene
ss69374920:A>G	0.01 (A)	UMAR	0.1	0.53	0.23	0.0230	0.73	no hits
ss61534850:A>G	0.22 (A)	UMAR	0.9	-0.18	0.06	0.0031	0.38	nearby CYP11B1
ss61524969:A>G	0.03 (A)	UMAR	2.4	0.33	0.15	0.0235	0.72	FADK
ss69374927:A>G	0.13 (G)	UMAR	2.4	-0.16	0.07	0.0291	0.67	no hits
ss69374930:A>G	0.06 (A)	ULMA	2.8	-2.56	1.07	0.017	0.82	no hits
ss69374934:A>G	0.17 (A)	UMAR	2.8	0.15	0.07	0.0252	0.68	no hits
ss69374933:A>T	0.07 (T)	ULMA	2.8	-2.62	1.05	0.013	0.82	no hits
ss69374936:C>G	0.14 (C)	UMAR	3.5	0.26	0.07	0.0002	0.16	no hits
ss69374937:A>G	0.42 (A)	ULMA	3.7	1.40	0.51	0.007	0.82	no hits
ss61476793:A>G	0.25 (A)	UMAR	3.7	0.17	0.06	0.0068	0.55	no hits
ss61476791:A>G	0.43 (A)	ULMA	3.7	1.42	0.51	0.006	0.90	no hits
ss61563132:A>G	0.20 (A)	ULMA	3.8	-1.70	0.57	003	1.11	no hits
ss38328040:A>G	0.22 (A)	UMAR	3.9	0.15	0.06	0.012	0.73	COL22A1
ss61527143:A>T	0.32 (A)	UMAR	3.9	0.17	0.06	0.0024	0.44	COL22A1
ss69374938:A>G	0.05 (A)	ULMA	4.1	-2.74	1.15	0.018	0.81	no hits

SNPs	MAF[1]	Trait[2]	cM	Estimate[3]	SE	P-value	FDR[4]	Gene
ss69374939:A>G	0.44 (G)	ULMA	4.2	1.32	0.57	0.021	0.82	no hits
ss38336818:A>G	0.31 (G)	ULMA	4.4	-1.43	0.65	0.029	0.79	no hits
ss61497130:A>G	0.48 (G)	ULMA	4.7	1.48	0.55	0.008	0.71	nearby RHPN1
ss61569297:A>C	0.48 (A)	ULMA	4.7	-1.46	0.56	0.009	0.61	nearby RHPN1
ss61569302:A>G	0.21 (G)	ULMA	4.7	1.63	0.59	0.006	1.07	nearby RHPN1
ss61497126:A>C	0.29 (A)	ULMA	5.0	-1.70	0.59	0.004	1.03	no hits
ss61472670:A>G	0.39 (G)	UBF	96.1	-0.50	0.21	0.017	0.37	nearby LOC785739
ss61472671:C>G	0.38 (G)	UBF	96.1	-0.48	0.21	0.024	0.40	nearby LOC785739
ss61472672:A>G	0.38 (G)	UBF	96.1	-0.48	0.21	0.023	0.40	nearby LOC785739
ss61472673:A>G	0.28 (G)	UBF	96.1	-0.55	0.21	0.009	0.28	no hits
ss38333031:A>T	0.34 (T)	UBF	96.5	0.75	0.24	0.002	0.22	no hits
ss61480722:A>C	0.26 (C)	UBF	96.6	-0.67	0.23	0.004	0.19	no hits
ss61480729:A>G	0.38 (G)	UBF	96.6	-0.66	0.20	0.001	0.22	no hits
ss61494456:A>G	0.46 (A)	UBF	98.3	-0.48	0.20	0.019	0.39	no hits
ss61563931:A>G	0.46 (A)	UBF	98.3	-0.50	0.20	0.014	0.34	no hits

[1]MAF = minor allele frequency

[2] UMAR = ultrasound marbling (score), ULMA = ultrasound longissimus muscle area (cm^2), UBF = ultrasound backfat (mm).

[3] Estimate of the effect expressed in units of the trait.

[4]False discovery rate calculated as $FDR = mP_{(i)} / I$, where m is the total number of tests and $P_{(i)}$ is the *P*-value at rank *i* when the *P*-values are ranked from lowest to highest (Benjamini and Hochberg, 1995; Weller et al., 1998).

Table 5.6 SNP validation using 1000 animals from the University of Guelph. Allele substitute effects were calculated using an animal model previously described in the materials and methods section. Markers in bold correspond were considered to be validated in the independent population.

	University of Alberta					University of Guelph					
SNPs	**Trait[1]**	**MAF[2]**	**Estimate[3]**	**SE**	**P-value**	**Trait[4]**	**MAF[5]**	**Estimate[6]**	**SE**	**P-value**	**gene**
ss38328882:A>G	CMAR	0.17	0.115	0.052	0.0278	MMWT	0.17	-1.332	0.648	0.041	nearby NSE2
ss38336818:A>G	LMA	0.31	-2.477	0.757	0.0012		0.23	not significant			no hits
	LMY		-0.790	0.357	0.0280						
	ULMA		-1.429	0.648	0.0292						
	YGRADE		0.155	0.067	0.0212						
ss38337066:A>T	CMAR	0.17	0.159	0.057	0.0064	WGm_ABC	0.24	-9.886	2.918	0.002	no hits
	LMA		-2.133	0.914	0.0206	ADG		0.040	0.020	0.044	
ss61467412:A>T	AVBF	0.39	0.890	0.349	0.0119	WGm_ABC	0.22	-6.534	1.908	0.001	nearby ADCY8
	LMA		-1.538	0.684	0.0252	BWm_ABC		-0.310	0.140	0.033	
	GRDFAT		0.918	0.350	0.0093						
	LMY		-0.963	0.318	0.0027						
	YGRADE		0.137	0.060	0.0240						
ss61469012:C>G	CMAR	0.46	0.113	0.039	0.0037	BWm_ABC	0.44	0.382	0.157	0.02	no hits
ss61473630:A>G	AVBF	0.29	0.888	0.365	0.0160	WADG	0.43	0.159	0.039	<.001	no hits
	GRDFAT		**1.038**	**0.362**	**0.0046**	**Fat2**		**0.649**	**0.317**	**0.042**	
	LMY		-0.768	0.331	0.0209	WWT		29.990	9.078	0.002	
	YGRADE		0.143	0.062	0.0224						
ss61473728:A>G	AVER_BF	0.36	-0.959	0.338	0.0048	BW	0.42	0.877	0.325	0.01	no hits
	GRDFAT		-1.079	0.339	0.0016	ADG		0.054	0.018	0.003	
	LMY		0.968	0.307	0.0018	FG		-0.193	0.058	0.001	
	YGRADE		-0.144	0.058	0.0143						

	University of Alberta					University of Guelph					
SNPs	**Trait[1]**	**MAF[2]**	**Estimate[3]**	**SE**	**P-value**	**Trait[4]**	**MAF[5]**	**Estimate[6]**	**SE**	**P-value**	**gene**
ss61480689:A>G	LMA LMY YGRADE	0.13	-2.262 -1.050 0.268	0.972 0.453 0.085	0.0207 0.0213 0.0019	WGm_ABC	0.11	8.665	4.070	0.04	nearby MGC148714
ss61480729:A>G	UBF	0.38	-0.660	0.202	0.0012	DMI HCW	0.34 0.34	-0.146 -6.002	0.073 2.462	0.048 0.016	no hits
ss61480731:A>C	**CMAR** **UMAR**	0.32	-0.088 -0.175	0.044 0.057	0.0470 0.0025	ADG **Lean** LM7D RW	0.35 **0.35** 0.35 0.35	0.058 **0.066** 0.210 0.087	0.023 **0.031** 0.098 0.044	0.013 **0.035** 0.034 0.05	no hits
ss61498340:A>G	**CARCWT**	0.37	**6.757**	**2.252**	**0.0029**	BFF HCW **Lean** WGm_ABC	0.41	-0.140 5.589 **0.052** 4.577	0.061 2.542 **0.023** 1.890	0.024 0.029 **0.026** 0.02	no hits
ss61516307:A>T	**LMA**	0.23	**2.044**	**0.774**	**0.0087**	**LM7D** WGm_ABC	0.26	**0.187** 9.295	**0.086** 2.499	**0.03** <.001	CYP7B1
ss61517008:A>G	UBF YGRADE	0.13	-0.664 -0.225	0.288 0.075	0.0217 0.0030		0.32	not significant			no hits
ss61532780:A>G	AVBF GRDFAT **LMY** UBF UMAR YGRADE	0.13	0.845 0.816 **-0.757** 0.827 0.128 0.129	0.333 0.351 **0.319** 0.230 0.057 0.060	0.0116 0.0209 **0.0182** 0.0004 0.0245 0.0328	**UGL**	0.15	**-0.607**	**0.262**	**0.022**	nearby LOC783782

	University of Alberta					University of Guelph					
SNPs	**Trait[1]**	**MAF[2]**	**Estimate[3]**	**SE**	**P-value**	**Trait[4]**	**MAF[5]**	**Estimate[6]**	**SE**	**P-value**	**gene**
ss61535128:A>G	CMAR	0.43	-0.124	0.038	0.0012	WG	0.47	3.893	1.885	0.046	nearby SIAT4A
	AVBF		**1.559**	**0.558**	**0.0055**	BF		-0.577	0.248	0.021	
	GRDFAT		**1.500**	**0.572**	**0.0092**	**SF**		**0.029**	**0.015**	**0.05**	
	LMY	0.08	-1.231	0.522	0.0190		0.05				
ss61535509:A>G	**UBF**		1.116	0.369	0.0026						DDEF1
	ULMA		2.049	0.969	0.0351						
	YGRADE		0.240	0.099	0.0157						
ss69374936:C>G	UMAR	0.14	0.265	0.071	0.0002	WG	0.16	-5.125	2.286	0.031	no hits
	LMA		**1.930**	**0.585**	**0.0011**	**Fat1**		**-1.363**	**0.424**	**0.002**	
ss69374937:A>G	**LMY**	0.42	**0.619**	**0.277**	**0.0266**	BWm_ABC	0.36	0.558	0.120	<.001	no hits
	ULMA		**1.402**	**0.515**	**0.0067**	Fat3		**-0.416**	**0.189**	**0.03**	
						FG		-0.146	0.062	0.021	

[1] UMAR = ultrasound marbling (score), ULMA = ultrasound longissimus muscle area (cm^2), UBF = ultrasound backfat (mm), CMAR = carcass marbling (score), LMA = carcass longissumus muscle area (cm^2), LMY = lean meat yield (%), YGRADE = yield grade (%), AVBF = carcass backfat (mm), GRDFAT = carcass gradefat (mm), CARCWT = carcass weight (kg).

[2] MAF = minor allele frequency

[3] Estimate of the effect expressed in units of the trait.

[4]False discovery rate calculated as $FDR = mP_{(i)} / I$, where m is the total number of tests and $P_{(i)}$ is the *P*-value at rank *i* when the *P*-values are ranked from lowest to highest (Benjamini & Hochberg 1995; Weller *et al.* 1998).

[5] MMWT = Metabolic mid test weight (kg/day), WGm_ABC = Weaning gain maternal ABC (kg/day), ADG = Average daily gain (kg/day), BWm_ABC = Birth weight maternal ABC (kg), WADG = Average daily gain from birth to weaning (kg/day), Fat2 = Fat depth (min fat in second quadrant, mm), WWT = Weaning weight (kg), FG = Feed / gain, DMI = Average dry matter intake (kg/day), HCW = Hot carcass Weight (kg), Lean = Lean weight (kg), LM7D = Shear force of LD aged 7 days (kg), RW = Rib weight (kg), BFF = Weight of body fat trim, as percentage of rib section (%),UGL = Estimated lean meat yield, WG = ABC for WADG (kg/day), BF = Weight of bone as a percentage of whole rib section (%), SF = Subcutaneous fat trim (kg), Fat1 =Fat depth (min fat in first quadrant, mm) , Fat3 = Fat depth (min fat in third quadrant, mm) , FG = Feed / gain

[6] Estimate of the effect expressed in units of the trait.

5.4 Literature Cited

Agriculture Canada (1992) Livestock carcass grading regulations. *Can. Gazette Part II.*, 3821–8.

Barendse W.J. (1999) Assessing lipid metabolism. *Patent, International Publication Number: WO 99/23248. World International Property Organization.*

Barrett J.C., Fry B., Maller J. & Daly M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-5.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B-Methodological*, 57:289-300.

Brethour J.R. (1992) The repeatability and accuracy of ultrasound in measuring backfat of cattle. *J Anim Sci* **70**, 1039-44.

Canadian Council on Animal Care (1993) Guide to the care and use of experimental animals. *Vol. 1. E. D. Olfert, B. M. Cross, and A. A. McWilliams, ed. Can. Counc. Anim. Care, Ottawa, Ontario, Canada.*

Casas E., Keele J.W., Shackelford S.D., Koohmaraie M. & Stone R.T. (2004) Identification of quantitative trait loci for growth and carcass composition in cattle. *Anim Genet* **35**, 2-6.

Casas E., Shackelford S.D., Keele J.W., Koohmaraie M., Smith T.P. & Stone R.T. (2003) Detection of quantitative trait loci for growth and carcass composition in cattle. *J Anim Sci* **81**, 2976-83.

Casas E., White S.N., Shackelford S.D., Wheeler T.L., Koohmaraie M., Bennett G.L. & Smith T.P. (2007) Assessing the association of single nucleotide polymorphisms at the thyroglobulin gene with carcass traits in beef cattle. *J Anim Sci* **85**, 2807-14.

Cohen-Zinder M., Seroussi E., Larkin D.M., Loor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I. & Ron M. (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 Res* **15**, 936-44.

de Koning D.J., Janss L.L., Rattink A.P., van Oers P.A., de Vries B.J., Groenen M.A., van der Poel J.J., de Groot P.N., Brascamp E.W. & van Arendonk J.A. (1999) Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa). *Genetics* **152**, 1679-90.

Dickens M., Rogers J.S., Cavanagh J., Raitano A., Xia Z., Halpern J.R., Greenberg M.E., Sawyers C.L. & Davis R.J. (1997) A cytoplasmic inhibitor of the JNK signal transduction pathway. *Science* **277**, 693-6.

Gutierrez-Gil B., Wiener P., Nute G.R., Burton D., Gill J.L., Wood J.D. & Williams J.L. (2008) Detection of quantitative trait loci for meat quality traits in cattle. *Anim Genet* **39**, 51-61.

Jones S.D., T. D. Bergess, J. W. Wilton, and C. H. Watson (1984) Feedlot performance, carcass composition and efficiency of muscle gain in bulls

and steers of different mature size slaughtered at different level of fatness. *Can. J. Anim. Sci.* **64**, 621–30.

Kaupe B., Brandt H., Prinzenberg E.M. & Erhardt G. (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. *J Anim Sci* **85**, 11-21.

Knott S.A., Elsen J.M. & Haley C.S. (1996) Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theoretical and Applied Genetics*, 93: 71-80.

Lindblad-Toh K., Wade C.M., Mikkelsen T.S., Karlsson E.K., Jaffe D.B., Kamal M., Clamp M., Chang J.L., Kulbokas E.J., 3rd, Zody M.C., Mauceli E., Xie X., Breen M., Wayne R.K., Ostrander E.A., Ponting C.P., Galibert F., Smith D.R., DeJong P.J., Kirkness E., Alvarez P., Biagi T., Brockman W., Butler J., Chin C.W., Cook A., Cuff J., Daly M.J., DeCaprio D., Gnerre S., Grabherr M., Kellis M., Kleber M., Bardeleben C., Goodstadt L., Heger A., Hitte C., Kim L., Koepfli K.P., Parker H.G., Pollinger J.P., Searle S.M., Sutter N.B., Thomas R., Webber C., Baldwin J., Abebe A., Abouelleil A., Aftuck L., Ait-Zahra M., Aldredge T., Allen N., An P., Anderson S., Antoine C., Arachchi H., Aslam A., Ayotte L., Bachantsang P., Barry A., Bayul T., Benamara M., Berlin A., Bessette D., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L., Bonnet C., Boukhgalter B., Brown A., Cahill P., Calixte N., Camarata J., Cheshatsang Y., Chu J., Citroen M., Collymore A., Cooke P., Dawoe T., Daza R., Decktor K., DeGray S.,

Dhargay N., Dooley K., Dooley K., Dorje P., Dorjee K., Dorris L., Duffey N., Dupes A., Egbiremolen O., Elong R., Falk J., Farina A., Faro S., Ferguson D., Ferreira P., Fisher S., FitzGerald M., Foley K., Foley C., Franke A., Friedrich D., Gage D., Garber M., Gearin G., Giannoukos G., Goode T., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N., Hagopian D., Hagos B., Hall J., Healy C., Hegarty R., Honan T., Horn A., Houde N., Hughes L., Hunnicutt L., Husby M., Jester B., Jones C., Kamat A., Kanga B., Kells C., Khazanovich D., Kieu A.C., Kisner P., Kumar M., Lance K., Landers T., Lara M., Lee W., Leger J.P., Lennon N., Leuper L., LeVine S., Liu J., Liu X., Lokyitsang Y., Lokyitsang T., Lui A., Macdonald J., Major J., Marabella R., Maru K., Matthews C., McDonough S., Mehta T., Meldrim J., Melnikov A., Meneus L., Mihalev A., Mihova T., Miller K., Mittelman R., Mlenga V., Mulrain L., Munson G., Navidi A., Naylor J., Nguyen T., Nguyen N., Nguyen C., Nguyen T., Nicol R., Norbu N., Norbu C., Novod N., Nyima T., Olandt P., O'Neill B., O'Neill K., Osman S., Oyono L., Patti C., Perrin D., Phunkhang P., Pierre F., Priest M., Rachupka A., Raghuraman S., Rameau R., Ray V., Raymond C., Rege F., Rise C., Rogers J., Rogov P., Sahalie J., Settipalli S., Sharpe T., Shea T., Sheehan M., Sherpa N., Shi J., Shih D., Sloan J., Smith C., Sparrow T., Stalker J., Stange-Thomann N., Stavropoulos S., Stone C., Stone S., Sykes S., Tchuinga P., Tenzing P., Tesfaye S., Thoulutsang D., Thoulutsang Y., Topham K., Topping I., Tsamla T., Vassiliev H., Venkataraman V., Vo A., Wangchuk T., Wangdi T., Weiand M.,

Wilkinson J., Wilson A., Yadav S., Yang S., Yang X., Young G., Yu Q., Zainoun J., Zembek L., Zimmer A. & Lander E.S. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803-19.

MacNeil M.D. & Grosz M.D. (2002) Genome-wide scans for QTL affecting carcass traits in Hereford x composite double backcross populations. *J Anim Sci* **80**, 2316-24.

Marques E., De Givry S., Stothard P., Murdoch B., Wang Z., Womack J. & Moore S. (2007) A high resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. *BMC Genomics* **8**, 254.

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C. & Benkel B. (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. *J Anim Sci* **81**, 1919-25.

Nkrumah J.D., Basarab J.A., Price M.A., Okine E.K., Ammoura A., Guercio S., Hansen C., Li C., Benkel B., Murdoch B. & Moore S.S. (2004) Different measures of energetic efficiency and their phenotypic relationships with growth, feed intake, and ultrasound and carcass merit in hybrid cattle. *J Anim Sci* **82**, 2451-9.

Nkrumah J.D., Sherman E.L., Li C., Marques E., Crews D.H., Jr., Bartusiak R., Murdoch B., Wang Z., Basarab J.A. & Moore S.S. (2007) Primary genome

scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. *J Anim Sci* **85**, 3170-81.

Okeudo N.J. & Moss B.W. (2004) Interrelationships amongst carcass and meat quality characteristics of sheep. *Meat Science* **69**, 1-8.

Olsen H.G., Nilsen H., Hayes B., Berg P.R., Svendsen M., Lien S. & Meuwissen T. (2007) Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. *BMC Genet* **8**, 32.

Scheet P. & Stephens M. (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* **78**, 629-44.

Schnabel R.D., Kim J.J., Ashwell M.S., Sonstegard T.S., Van Tassell C.P., Connor E.E. & Taylor J.F. (2005) Fine-mapping milk production quantitative trait loci on BTA6: analysis of the bovine osteopontin gene. *Proc Natl Acad Sci U S A* **102**, 6896-901.

Seasholtz T.M., Majumdar M. & Brown J.H. (1999) Rho as a mediator of G protein-coupled receptor signaling. *Mol Pharmacol* **55**, 949-56.

Seaton G., Haley C.S., Knott S.A., Kearsey M. & Visscher P.M. (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339-40.

Seybert D.W. (1990) Lipid regulation of bovine cytochrome P450(11) beta activity. *Arch Biochem Biophys* **279**, 188-94.

Sherman E.L., Nkrumah J.D., Li C., Bartusiak R., Murdoch B. & Moore S.S. (2009) Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle. *J Anim Sci* **87**, 37-45.

Snelling W.M., Chiu R., Schein J.E., Hobbs M., Abbey C.A., Adelson D.L., Aerts J., Bennett G.L., Bosdet I.E., Boussaha M., Brauning R., Caetano A.R., Costa M.M., Crawford A.M., Dalrymple B.P., Eggen A., Everts-van der Wind A., Floriot S., Gautier M., Gill C.A., Green R.D., Holt R., Jann O., Jones S.J., Kappes S.M., Keele J.W., de Jong P.J., Larkin D.M., Lewin H.A., McEwan J.C., McKay S., Marra M.A., Mathewson C.A., Matukumalli L.K., Moore S.S., Murdoch B., Nicholas F.W., Osoegawa K., Roy A., Salih H., Schibler L., Schnabel R.D., Silveri L., Skow L.C., Smith T.P., Sonstegard T.S., Taylor J.F., Tellam R., Van Tassell C.P., Williams J.L., Womack J.E., Wye N.H., Yang G. & Zhao S. (2007) A physical map of the bovine genome. *Genome Biol* **8**, R165.

Stone R.T., Keele J.W., Shackelford S.D., Kappes S.M. & Koohmaraie M. (1999) A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. *J Anim Sci* **77**, 1379-84.

Weller J.I., Song J.Z., Heyen D.W., Lewin H.A. & Ron M. (1998) A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. *Genetics* **150**, 1699-706.

CHAPTER 6

Identification of Candidate Markers on BTA14 under Milk Production Trait Quantitative Trait Loci in Holstein[3]

6.1 Introduction

Molecular technology advances have enabled the identification of many genetic markers affecting economically relevant traits in cattle. Identification of these genetics markers will be crucial for understanding the variation underlying these complex traits. In dairy cattle, most of the genetic markers identified are the ones affecting milk production traits such as milk, fat and protein yields, and fat and protein percentages. These traits have high economic impacts, and identifying genetic markers associated with these phenotypes can bring important success to the industry. In the era of whole genome association analysis, it can be argued that linkage analysis should be skipped, as it lacks power. However, having both association and linkage evidence makes a stronger case for causality.

The vast knowledge of Quantitative Trait Loci (**QTL**) locations has aided the community in identifying specific genes directly affecting and giving rise to these milk production QTL: *DGAT1* (Grisart *et al.* 2002; Winter *et al.* 2002) for milk fat, *OPN* (Schnabel *et al.* 2005a) or *ABCG2* (Cohen-Zinder *et al.* 2005) for milk yield. Perhaps, the most widely known success story in identification of causal mutation is the case of *DGAT1* (acylCoA:diacylglycerol acyltransferase 1) on bovine chromosome 14 (**BTA14**). Grisart *et al*. (2004) was able to establish

[3] A version of this chapter has been submitted. Marques et al., 2009. J. Dairy Science (JDS-09-2386)

causal effect of *DGAT1* K232A mutation on fat percentage. This work was followed by Riquet *et al.* (1999) who fine mapped a previously identified milk QTL on BTA14 (Coppieters *et al.* 1998; Ron *et al.* 1998).

Since then, a number of other studies have been conducted showing this mutation to have a large effect on milk production traits in other dairy populations Spelman *et al.* (2002); Thaller *et al.* (2003) and Naslund *et al.* (2008). There are also reports indicating that *DGAT1* is not solely responsible for the total genetic variation observed in milk production traits (Bennewitz *et al.* 2004; Kaupe *et al.* 2007), meaning that, by itself, *DGAT1* mutation does not explain the total effect produced in association studies in German Hosltein cattle. Bennewitz *et al.* (2004) explained that after accounting for the effects of *DGAT1* as a genetic cofactor, there were still significant residual effects in these traits at the proximal end of BTA14. Their results suggested that other polymorphisms on BTA14 are also responsible for the large QTL effect observed on the proximal end of BTA14. Following this work, Kaupe *et al.* (2007) jointly analyzed *DGAT1* and *CYP11B1*. *CYP11B1* is a gene which encodes 11β-hydroxylase enzyme involved in the production of hormones influencing fluid volume, electrolyte homeostasis and the metabolism of glucose and lipids (Bulow & Bernhardt 2002). According to (Kaupe *et al.* 2004), when analyzed together, those 2 genes explained more of the variation in milk production traits than *DGAT1* alone. Another study initially proposed a variable number of tandem repeat locus in the 5' noncoding region of *DGAT1* as a putative causative variant (Gautier *et al.* 2007). However, further

analysis indicated that no allele at this locus was significantly associated with fat percentage.

All these studies have in common the correction for the effect of *DGAT1* in their analysis. In fact, the high effect of *DGAT1* on these traits across different populations overshadows the identification of other QTL along BTA14. As the gaps in the bovine map are filled and more information pertaining to the specific marker-marker relationship (i.e: linkage disequilibrium) becomes available, the focus will shift towards identifying smaller effect QTL. The aim of our study was to identify markers under the QTL for milk production traits after accounting for the high effect of *DGAT1* on these traits.

6.2 Materials and Methods

6.2.1 Pedigree, QTL mapping and Statistical Analysis

QTL mapping was performed using 321 animals belonging to 7 sire families. QTL analysis performed in this study used the multiple marker interval mapping approach described by (Knott *et al.* 1996). The conditional probabilities that a calf inherited the first allele of a putative QTL from its sire were obtained QTL Express (Seaton *et al.* 2002) under a half-sib model. These probabilities were then modeled using SAS (SAS Inst. Inc., Cary, NC). The statistical analyses model included fixed effects of SNP genotype, linear covariate of *DGAT1* genotype and reliability as a weight variable.

Chromosome-wise thresholds of 5% and 1% for statistical significance and QTL confidence levels were determined for each trait based on a permutation

test of 25,000 iterations (Churchill & Doerge 1994). The phenotypic data used for linkage analysis were EBV (estimated breeding values) for milk yield (kg), fat yield (kg), protein yield (kg), fat contet (%) and protein content (%). EBVs and pedigree information were obtained from Holstein Canada (http://www.holstein.ca/english/Animallnq/animalinq.asp).

Allele substitution effects were estimated using Procedure Mixed in SAS (SAS Inst. Inc., Cary, NC) under the same statistical mode by regressing phenotypes on the number of copies of one allele for each SNP according to procedures described by (Falconer & Mackay 1996). Probability values do no correct for multiple testing (see section on false discovery rate).

6.2.2 Compilation of SNP markers and Genotyping

SNPs included in this analysis were compiled and selected from Baylor College of Medicine bovine database publicly available at ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus. Additional SNPs were derived from bac end sequences (BES) of the CHORI-240 library (http://bacpac.chori.org/bovine240.htm) and IBISS (Interactive Bovine In Silico SNP) database (www.livestockgenomics.csiro.au/ibiss). Haplotypes for all animals and their sires were generated as described by (Marques ***et al.*** 2008). Initially, only sire haplotypes were loaded onto HAPLOVIEW in order to identify markers with observed heterozygosity greater than 0.5. The settings used included: min genotyped %: 50, Hardy Weinberg (HW) p-value cutoff: 0.0010, Minimum minor allele freq: 0.0010, Maximum # Mendel errors: 1. Once all

animals' haplotypes were loaded, only those markers were selected and run through the tagger option incorporated into HAPLOVIEW. This procedure defines a threshold for r^2 (default: 0.8) and SNPs tagged have LD measure higher than the threshold set. This method started with 502 SNP markers and yielded 139 SNP markers to be used for QTL analysis. Genotypes for SNPs were generated by high throughput Illumina BeadStation 500G genotyping system and analyzed using Illumina's GenCall Software (version 1.014). *DGAT1* primers were designed according to Kaupe *et al.* (2004). Primer design for KCBN2 was obtained by using the genomic sequence and BLASTing it to the mRNA sequence. Exonic regions were identified and selected for further SNP discovery using 3730 ABI sequencer. *DGAT1* genotypes for the 321 animals in the QTL analysis were obtained using 3730 ABI sequencer. *DGAT1* genotypes for the validation animals were obtained using RFLP procedure according to procedures described by (Kaupe *et al.* 2004).

6.2.3 SNP marker location and Sequence BLAST to the genome

Marker positions were inferred using the marker order from the 12K RH map of Marques *et al.* (2007) which is in high agreement with the recently released physical map based on independent whole genome map of Snelling *et al.* (2007). Preliminary comparison between our marker order and the recently published Btau_4.0 shows agreement for markers common to both maps (Paul Stothard, personal communication). Relative genetic distances in centimorgans

(cM) were estimated by using the centiray positions (cR) and the estimated length of chromosome 14 (108 cM).

6.2.4 SNP Validation

In our study, 12 SNP markers associated with milk production traits were selected to be genotyped on a additional 726 Holstein animals that were not available at the start of the experiment. These markers were selected based on their P-value in the initial analysis and that they were significant ($P < 0.05$) in at least one of the milk production traits. The allele substitution effect was calculated with the same model as the initial analysis. The analysis did not include the first 321 animals described for QTL and positional candidate marker analysis. The genotyping assay was developed using the MassARRAY® iPLEX Gold platform technology, run on the Sequenom MassARRAY System and genotypes were provided by Sequenom Inc (San Diego).

6.2.5 False Discovery Rate

False discovery rate procedure was applied to our analysis as a mean to minimize false positives according to procedures described by Benjamini & Hochberg (1995) and Weller *et al.* (1998). Briefly, it takes into consideration the number of tests performed, the ranking of the marker within the analysis and their significance (*P*-value) rank from lowest to highest. Since FDR assumes independence between traits and since these traits are correlated, FDR was calculated within each trait according to the formula:

$$FDR = \frac{n \times P(k)}{k},$$

where k is the ranking of each marker, *P* is the P-value associated with the marker, and n is the number of markers analyzed.

6.3 Results and Discussion

The initial number of markers available for the QTL analysis was 502 SNP markers, a number too large to perform the analysis, considering that the software available to date only use information for markers separated by 1 cM. Markers for QTL analysis were selected on the basis of sire heterozygosity and linkage disequilibrium information. Knowledge of the large effect of *DGAT1* K232A mutation on milk production traits (Grisart *et al.* 2002) allowed the use of this genetic information in the statistical model, accounting for its variation and therefore enabling the identification of other milk production trait QTL on bovine chromosome 14. The initial QTL analysis confirming the presence of the large effect QTL for milk production traits in the proximal region of BTA14 was performed excluding *DGAT1* genetic information as a covariate. Results showed that QTL were being identified in this region: Milk Yield at 1 cM ($P < 0.001$) (Figure 6.1), fat yield at 2 cM ($P < 0.0001$) and 32 cM ($P < 0.01$) (Figure 6.2), protein yield at 42 cM ($P < 0.05$) (Figure 6.3), fat Content at 1 cM ($P < 0.0001$) and 31 cM ($P < 0.05$) (Figure 6.4), protein content at 3 cM ($P < 0.0001$), 19 cM ($P < 0.01$) and 31 cM ($P < 0.05$) (Figure 6.5).

6.3.1 Milk yield

Comparing the results between the initial QTL analysis (excluding *DGAT1* as a covariate) with the second QTL analysis (including *DGAT1* as a covariate) showed that addition of *DGAT1* as a genetic covariate completely removed the peak at 1 cM indicating that *DGAT1* was responsible for this QTL. The results showed 2 putative QTL profiles at 42 cM and 61 cM, both at $P < 0.10$ (Table 6.1 and Figure 6.6). Schnabel *et al.* (2005b) reported that after accounting for the *DGAT1* effect, there was still evidence for the existence of a second milk production QTL (position: 61 cM) on BTA14. This study also used a higher number of animals, which could be a reason for the different results achieved in our analysis.

For the purpose of identifying markers, the two putative QTL were treated as one. Analyzing SNPs under or near these putative QTL, there were 139 SNPs that reached 5% significance. However, due to the high number of SNPs analyzed only the ones which reached $P < 0.01$ significance are reported (Tables 6.1 and 6.2). A 1% significance threshold for the *P* - value was selected as a way to minimize false positives. These SNP sequences were BLASTed to the bovine genome in order to identify important genes where these SNPs are located.

SNPs under these two QTL have a minor allele frequency ranging from 0.07 to 0.49. Several of those SNPs map within or nearby genes whose functions include regulation of secretion of adrenocorticotrophin hormone, Ca^{2+} signal transduction pathway, intracellular vesicular trafficking among others (Table 6.2). In particular, corticotrophic hormone (CRH) has been pinpointed as a

promising candidate gene in cattle affecting marbling and subcutaneous fat depth (Wibowo *et al.* 2007). CRH is the major releasing factor for ACTH secretion. ACTH, in turn, regulates glucocorticoids to mediate stress response (Seasholtz *et al.* 2002).

6.3.2 Fat yield and content

The initial QTL analysis (excluding *DGAT1* as covariate) showed a significant fat yield QTL at 2 cM ($P < 0.0001$) and 32 cM ($P < 0.01$) (Figure 6.2). After accounting for *DGAT1* the QTL peaks were observed at 42 cM and 63 cM, both at $P < 0.05$ (Figure 6.7). The proximal QTL is no longer observed, indicating that *DGAT1* was accounting for that particular QTL. Identifying markers under these new QTL resulted in only 2 markers: ss61514555:A>C ($P = 6.88E-05$) and ss61482545:A>G ($P = 8.93E-03$).The latter maps to an armadillo repeat containing 1 (*ARMC1*) gene, with no known function, while ss61514555:A>C showed no hits to any known genes.

In the case of fat content, there is still a strong QTL peak in the proximal region of BTA14 ($P < 0.01$) when the effect of *DGAT1* was accounted for, however, not as significant as when *DGAT1* was not included in the model. Another nearby QTL is still observed, except that it has now shifted from 31 cM to 29 cM (Figure 6.9). The addition or removal of *DGAT1* from the model does not seem to affect the presence of this QTL, which remains at the 5% significance level. Two SNPs were identified for the first peak: ss61497126:A>C and ss61508244:C>G. There were no direct hits of these SNPs to genes, however,

there are several nearby genes (100 kbps away) for ss61508244:C>G. One of which, LOC512826 (*GPIHBP1*), provides a platform for the binding of both lipoprotein lipase (LPL) and chylomicrons (Ioka *et al.* 2003; Beigneux *et al.* 2007). Beigneux *et al*. (2007) showed that mice with induced *GPIHBP1* deficiency had compromised lipolysis.

Under the second peak, one SNP is of particular interest: ss61473395:A>G (P = 8.26E-03). When this SNP sequence was BLASTed to the genomic sequence, it did not show direct hits to any genes, however, when searching for nearby genes, *MCM4* gene was on the list. This gene selectively interacts with ATP (Bochman & Schwacha 2007), an important coenzyme and enzyme regulator. Among its known functions are: involvement in DNA binding (Kaplan *et al.* 2003) and DNA replication (Bailis *et al.* 2008).

6.3.3 Protein yield and content

The initial QTL analysis (excluding *DGAT1* as a covariate) showed a few significant peaks for protein content: 3 cM at $P < 0.001$, 19 cM and 31 cM at $P < 0.05$ (Figure 6.5). Adding *DGAT1* to the analysis as a genetic covariate still showed a peak at 4 cM ($P < 0.05$) (Figure 6.10). This indicated that *DGAT1* was not solely responsible for this QTL peak. On the other hand, the other two peaks were no longer significant at $P < 0.05$ level. Among the SNPs under this QTL is ss61535522:A>C. There were no known genes where this SNP was BLASTing to. One of the nearby genes includes a gene involved in eukaryotic signal transduction pathway (*MGC138052*), also known as mitogen-activated protein

kinase. This family of serine/threonine kinases play important roles in signal transduction in all eukaryotic cells (Widmann *et al.* 1999).

The same occurred for protein yield analysis. Two peaks were identified when including *DGAT1* as a covariate: 42 cM ($P < 0.01$) and 84 cM ($P < 0.05$) (Figure 6.8). Initially only the peak at 42 cM had been identified and only at the 5% significance level. Several SNPs were identified under these QTL. The range in minor allele frequencies is 0.01 to 0.47. One of these SNPs sequence BLASTs to LOC617133 (*RGS22*), a regulator of G-protein signaling 22. (Hollinger & Hepler 2002) provided a review of the cellular regulation of RGS proteins. Among some functions are cell proliferation and differentiation.

A SNP under the protein yield QTL, ss61514147:A>G, which showed association to milk yield, also shows a significant association with protein yield (P = 1.35E-03). This SNP did not show any direct hits to known genes, however, a nearby gene corticotrophin homone (*CRH*) and tripartite motif-containing 55 (*TRIM55*) were identified. *CRH* was previously mentioned as being a promising candidate for a marbling QTL in beef cattle (Wibowo *et al.* 2007). *TRIM55* encodes a protein which contains a ring zinc finger. This motif is believed to work in protein-protein interactions (Centner *et al.* 2001).

There are 6 SNPs common to two or more traits (Table 6.2). One SNP in particular, ss38330358:A>G, shows associations to milk (P = 1.22E-04) and protein (P = 2.98E-03). This SNP maps to LOC529552, a cytochrome P450, family 7, subfamily B, polypeptide 1 gene (*CYP7B1*). *CYP7B1* is also known as

oxysterol 7-hydroxylase which is involved in a important pathway of bile acid synthesis (Vlahcevic *et al.* 1999).

6.3.4 Validation of SNPs and Candidate Gene

Eleven out of the twelve markers chosen for further validation were confirmed to be significantly associated with at least one of the milk production traits. The most significant association was for milk yield and ss61535150:A>C (P = 1.56E-15), followed by protein yield with this same SNP (P = 3.06E-15) (Table 6.3). This SNP BLASTed to a hypothetical protein LOC783036 with no known function and it showed a minor allele frequency of 0.22. Table 6.3 lists the estimates for the allele substitution effect for SNPs genotyped on 726 Holstein animals. Comparing results from SNPs under the QTL (Table 6.2) and the validation of SNPs (Table 6.3), there are 5 markers in common (ss38329727:C>T, ss61466612:A>G, ss61517610:A>G, ss61535150:A>C, ss61535522:T>G). Overall, increasing the number of animals enabled not only validation of 11 out 12 markers, but also showed new marker association to other traits, which had not been previously observed.

Among all the candidate genes listed on table 2, *KCNB2* a gene encoding a potassium voltage-gated channel, Shab-related subfamily, member 2 was selected for further study. The SNP (ss61516059:A>G) whose sequence BLASTed to *KCNB2* gene had an significant association with milk yield (P = 7.80E-03), but most importantly the function of this gene stood out due its complexity probable association with insulin release, considering that insulin secretion is modulated by

different ionic currents (Dukes & Philipson 1996). Voltage-gated potassium channels are known to be involved in the repolarization of excitable cells. Therefore, any defects in this voltage-channel could cause these pancreatic cells to have higher or lower insulin release, depending on how these voltage-gated channels are affected by a particular mutation.

Genomic sequence of *KCNB2* was BLASTed to mRNA sequence and exonic sequences identified. Search of SNPs in these exonic regions resulted in the identification of 7 unique polymorphisms (Table 6.4). Most of these confer silent mutations, while two, ss107795104:C>G and ss63780010:C>T confer changes from glutamine to glutamic acid and from aspartic acid to asparagine, respectively. One SNP, ss107795103:A>C, showed significant associations ($P < 0.01$) to all 5 milk production traits, however its low minor allele frequency (0.09) could have contributed to its significant associations. It is important to note that not all polymorphisms were identified; therefore many other intronic polymorphisms are still missing from this analysis, which can very well have a more significant impact on the potassium voltage gated channels if they are part of any miRNA sequence target. These are likely to make a significant contribution to phenotypic variation in many livestock species (Georges *et al.* 2007).

Accounting for genetic variation from genes already identified is a major step in searching for novel markers affecting economically relevant traits. The addition of genetic information in statistical models enables the discoveries of those genetic markers with smaller effects. Information from linkage disequilibrium and family heterozygosity are important aspects in selecting the

best possible markers for QTL analysis, especially in the era of high density marker maps.

Selection of markers based solely on their association with the trait is not ideal. Information from linkage analysis, association and function of the gene need to be combined to increase the accuracy of selecting a true predictive marker. The use of this information enabled the discovery of several unidentified smaller effect milk production, as well as single nucleotide polymorphism under or near these QTL.

Table 6.1 Quantitative Trait Loci (QTL) on bovine chromosome 14 (BTA14) using Holstein cattle across seven families

Trait[1]	**Location (cM)**	**Estimate**[2]	**SE**	***P*-Value**	**# SNPs (*P* < 0.01)**
Milk Yield	42	-196.66	97.2	0.1	20
	61	-185.06	95.58	0.1	
Fat Yield	42	-7.89	3.54	0.05	2
	63	-7.9	3.49	0.05	
Protein Yield	42	-7.02	2.73	0.01	8
	84	-6.9	2.76	0.05	9
Fat Content	3	-8.52	3.27	0.01	2
	29	-7.02	2.99	0.05	6
Protein Content	4	-4.15	1.71	0.05	5

[1]Milk Yield= volume of milk (kg), Protein Yield = volume of protein (kg), Fat Yield = volume of fat (kg),

Protein Content = percentage of protein (%), Fat Content = percentage of fat (%)

[2] Estimate of the effect expressed in units of the trait

Table 6.2 Summary of candidate single nucleotide polymorphism markers under 5 milk production traits Quantitative Trait Loci

NCBI	MAF[1]	Trait[2]	Estimate[3]	SE	*P*-value	FDR[4]	Gene
ss38327919:A>C	0.46 (C)	Fat Content	-8.0	0.02	1.33E-04	2.15E-02	no hits
ss38328601:A>G	0.31 (G)	Protein Yield	8.34	1.89	1.48E-05	1.91E-03	LOC515582
ss38329727:C>T	0.38 (C)	Protein Yield	-5.27	1.96	7.64E-03	7.24E-02	nearby LOC783036 and LOC790864
ss38329953:A>G	0.47 (G)	Protein Yield	4.99	1.80	5.91E-03	6.04E-02	no hits
ss38330358:A>G	0.41(A)	Milk Yield Protein Yield	249 5.62	63.99 1.88	1.22E-04 2.98E-03	3.58E-03 4.36E-02	LOC529552
ss61512747:A>G	0.06 (G)	Fat Content	-11.0	4.00	6.96E-03	1.32E-01	no hits
ss61470582:A>G	0.08 (A)	Milk Yield Protein Yield	577 13.7	114.60 3.36	8.71E-07 5.63E-05	1.12E-04 3.62E-03	nearby LOC783431
ss61514147:A>G	0.24 (A)	Milk Yield Protein Yield	300 7.52	79.76 2.32	2.05E-04 1.35E-03	5.08E-03 2.71E-02	nearby CRH and TRIM55
ss61514555:A>C	0.17 (A)	Milk Yield Fat Yield Protein Yield	-322 -12.7 -9.92	84.16 3.14 2.43	1.59E-04 6.88E-05 5.76E-05	4.09E-03 8.86E-03 3.09E-03	no hits
ss61515962:A>G	0.26 (G)	Milk Yield	197	73.81	7.95E-03	5.45E-02	no hits
ss61516059:A>G	0.24 (A)	Milk Yield	-200	74.63	7.80E-03	5.58E-02	KCNB2
ss61516307:A>T	0.25 (A)	Milk Yield	-244	72.75	8.89E-04	1.40E-02	LOC529552

NCBI	MAF[1]	Trait[2]	Estimate[3]	SE	*P*-value	FDR[4]	Gene
ss61517008:A>G	0.36 (A)	Milk Yield	-203	65.36	2.07E-03	2.42E-02	no hits
ss61517610:A>G	0.10 (G)	Milk Yield	-521	101.69	5.68E-07	9.14E-05	nearby LOC782438
ss61473395:A>G	0.15 (G)	Fat Content	-7.0	3.00	8.26E-03	1.40E-01	neaby MCM4
ss61520620:A>G	0.12 (G)	Protein Yield	12.0	2.70	1.24E-05	2.67E-03	no hits
ss38336286:A>G	0.11 (A)	Protein Content	4.37	1.52	4.26E-03	2.75E+00	no hits
ss61466612:A>G	0.05 (G)	Protein Content	7.88	2.07	1.79E-04	1.15E-01	no hits
ss61530605:C>G	0.15 (C)	Milk Yield	376	94.98	9.50E-05	3.06E-03	nearby LOC536186
ss61478975:A>G	0.31 (A)	Protein Yield	-6.89	1.94	4.61E-04	1.35E-02	LOC78474
ss61534605:A>G	0.47 (G)	Fat Content	5.16	1.86	5.77E-03	1.38E-01	LOC526726
ss61480452:A>C	0.25 (A)	Protein Yield	-6.11	2.20	5.80E-03	6.03E-02	MGC157244
ss61480461:A>G	0.07 (A)	Protein Yield	-9.32	3.54	8.93E-03	8.10E-02	no hits
ss61508227:A>C	0.07 (C)	Milk Yield	378	130.46	4.06E-03	3.58E-02	LOC519708
ss61480519:A>T	0.31 (T)	Milk Yield	210	71.89	3.80E-03	3.44E-02	BIG1
ss61480576:A>G	0.12 (A)	Milk Yield	-293	104.64	5.45E-03	4.45E-02	LOC536811, nearby LOC536811
ss61480578:A>G	0.10 (G)	Milk Yield	307	114.63	7.88E-03	5.58E-02	LOC536811
ss61534821:A>G	0.11 (A)	Milk Yield	373	97.30	1.58E-04	4.24E-03	LOC512677

NCBI	MAF[1]	Trait[2]	Estimate[3]	SE	*P*-value	FDR[4]	Gene
ss61534844:A>C	0.49 (A)	Milk Yield	225	62.94	4.02E-04	7.85E-03	PKIA
ss61508244:C>G	0.24 (G)	Fat Content	-5.93	2.14	6.03E-03	1.30E-01	nearby LOC512826
		Protein Content	-3.12	1.12	5.71E-03	3.68E+00	
ss61508266:A>C	0.24 (C)	Protein Yield	-7.98	2.05	1.28E-04	4.58E-03	no hits
ss61535150:A>C	0.21 (A)	Protein Yield	-11.2	2.05	1.01E-07	6.49E-05	hypothetical LOC783036
ss61535152:A>G	0.01 (A)	Protein Yield	8.69	3.12	5.75E-03	6.07E-02	nearby LOC783036 and LOC790864
ss61480685:C>G	0.39 (G)	Protein Yield	5.50	1.96	5.42E-03	6.02E-02	no hits
ss61535181:A>G	0.19 (G)	Protein Yield	10.9	2.30	3.57E-06	1.15E-03	LOC617133
ss61535522:G>T	0.40 (G)	Protein Content	3.42	1.10	2.13E-03	1.37E+00	nearby PYCRL
ss61501250:A>C	0.43 (A)	Fat Content	-5.97	1.94	2.25E-03	9.67E-02	hypothetical LOC539014
ss61482545:A>G	0.45 (G)	Fat Yield	6.35	2.41	8.93E-03	1.69E-01	ARMC1
ss61488018:A>C	0.11 (C)	Milk Yield	-361	107.68	9.27E-04	1.39E-02	nearby LOC512677
ss61566174:A>G	0.44 (A)	Milk Yield	188	70.06	7.64E-03	5.53E-02	no hits
ss61566175:A>G	0.25 (A)	Milk Yield	277	73.60	2.08E-04	4.97E-03	no hits
ss61568592:A>C	0.43 (C)	Protein Yield	5.21	1.80	4.16E-03	5.35E-02	no hits
		Fat Content	-4.96	1.86	8.22E-03	1.47E-01	no hits

NCBI	MAF[1]	Trait[2]	Estimate[3]	SE	*P*-value	FDR[4]	Gene
ss61470049:A>G	0.19 (G)	Milk Yield	224	85.28	9.11E-03	5.81E-02	MGC128538
ss61497126:A>C	0.35 (C)	Fat Content	6.0	2.00	4.61E-03	1.24E-01	no hits
ss61470052:C>G	0.46 (C)	Protein Content	-9.82	2.49	9.85E-05	6.34E-02	no hits

[1]MAF = minor allele frequency

[2]Milk yield = volume of milk (kg), Protein yield= volume of protein (kg), Fat = volume of fat (kg), Protein content = percentage of protein (%), Fat content = percentage of fat (%)

[3] Estimate of the effect expressed in units of the trait.

[4]False discovery rate calculated as $FDR = mP_{(i)} / I$, where m is the total number of tests and $P_{(i)}$ is the *P*-value at rank *i* when the *P*-values are ranked from lowest to highest (Benjamini and Hochberg, 1995; Weller *et al.*, 1998).

Table 6.3 Summary of allele substitution effect for SNP markers genotyped on 726 Holstein animals

NCBI Accession	MAF	Trait	Estimate	SE	*P* - value	FDR	Gene
ss38329727:C>T	0.34 (C)	Milk Yield	141	43.7	1.34E-03	4.03E-03	nearby LOC783036 and LOC790864
		Protein Yield	4.97	1.29	1.22E-04	3.67E-04	
		Fat Yield	-5.99	2.06	3.77E-03	1.51E-02	
ss61516949:A>G	0.15 (G)	Milk Yield	-154	56.7	6.95E-03	1.67E-02	no hits
		Protein Yield	-6.43	1.66	1.22E-04	4.89E-04	
		Fat Content	5.72	2.66	3.17E-02	7.61E-02	
ss61517610:A>G	0.07 (G)	Milk Yield	-311	84.9	2.71E-04	1.08E-03	nearby LOC782438
		Protein Yield	-7.24	2.54	4.50E-03	1.08E-02	
		Fat Yield	9.02	1.95	4.41E-06	5.29E-05	
ss61476770:A>G	0.17 (G)	Milk Yield	251	53.8	3.71E-06	2.23E-05	no hits
		Protein Yield	7.25	1.59	5.85E-06	3.51E-05	
		Fat Content	4.83	1.38	5.26E-04	2.10E-03	
ss61476793:A>G	0.49 (G)	Fat Yield	3.79	1.65	2.19E-02	5.25E-02	no hits
		Protein Content	1.77	0.70	1.14E-02	6.85E-02	
ss61466612:A>G	0.06 (G)	Protein Yield	5.38	2.67	4.43E-02	8.86E-02	no hits
		Fat Content	8.59	1.48	1.00E-08	1.20E-07	
ss61534850:A>G	0.28 (A)	Fat Yield	4.89	1.78	6.17E-03	1.85E-02	nearby LOC512826
		Milk Yield	-106	49.2	3.17E-02	6.34E-02	
ss61508252:C>T	0.40 (C)	Fat Content	-4.78	1.67	4.28E-03	1.28E-02	nearby FOXH1
		Fat Yield	-6.54	1.75	2.06E-04	1.24E-03	
ss61535150:A>C	0.23 (A)	Milk Yield	-245	47.3	3.05E-07	3.66E-06	LOC783036
		Protein Yield	-6.75	1.41	2.06E-06	2.48E-05	
ss61535522:T>G	0.41 (G)	Fat Content	-5.61	1.50	2.06E-04	1.24E-03	nearby PYCRL
		Protein Content	-1.97	0.75	9.11E-03	1.09E-01	

NCBI Accession	MAF	Trait	Estimate	SE	*P* - value	FDR	Gene
ss61569304:A>G	0.14 (G)	Fat Content	4.03	1.91	3.50E-02	7.00E-02	no hits
		Protein Content	-2.16	0.95	2.37E-02	9.48E-02	

[1]MAF = minor allele frequency

[2]Milk yield = volume of milk (kg), Protein yield = volume of protein (kg), Fat yield = volume of fat (kg), Protein content = percentage of protein (%), Fat content = percentage of fat (%)

[3] Estimate of the effect expressed in units of the trait.

[4]False discovery rate calculated as $FDR = mP_{(i)} / I$, where m is the total number of tests and $P_{(i)}$ is the *P*-value at rank *i* when the *P*-values are ranked from lowest to highest (Benjamini and Hochberg, 1995; Weller *et al.*, 1998).

Table 6.4 Estimates of allele substitution effect in milk production traits in Holstein animals for seven single nucleotide polymorphisms within the KCBN2 gene (Gene ID: 535990) located on bovine chromosome 14. The analysis was performed using 321 animals.

NCBI Accession	MAF[1]	GenBank Accession No and Base Position (Btau_3.1)	Trait[2]	Estimate[3]	SE	*P*-value	FDR[4]	Type of mutation
ss107795105:C>G	0.08 (C)	NM_001024563.1 - 2263	Fat Content	5.5	0.042	1.89E-01	9.46E-01	silent (proline)
ss107795104:C>G	0.24 (G)	NM_001024563.1 - 2143	Fat Yield	4.43	2.78	1.12E-01	5.60E-01	glutamine to glutamic acid
			Milk Yield	142	76.3	6.41E-02	3.21E-01	
			Protein Yield	3.42	2.20	1.21E-01	6.06E-01	
ss107795103:A>C	0.09 (C)	NM_001024563.1 - 2017	Fat Yield	-8.48	4.36	5.27E-02	2.63E-01	silent (Threonine)
			Fat Content	11.5	0.034	9.39E-04	4.69E-03	
			Milk Yield	-572	116	1.41E-06	7.04E-06	
			Protein Yield	-13.7	3.41	7.58E-05	3.79E-04	
			Protein Content	4.8	0.018	8.48E-03	4.24E-02	
ss63780010:A>G	0.12 (A)	NM_001024563.1 - 1859	Fat Yield	-5.31	3.59	1.40E-01	7.00E-01	Aspartic acid to asparagine
			Milk Yield	-219	96.9	2.44E-02	1.22E-01	
			Protein Yield	-4.13	2.84	1.47E-01	7.33E-01	
			Protein Content	2.7	0.015	6.84E-02	3.42E-01	
ss107795102:C>G	0.23 (G)	NM_001024563.1 - 1858	Fat Yield	2.79	0.22	2.20E-01	1.10E+00	silent (proline)
			Milk Yield	143	75.5	5.91E-02	2.95E-01	
			Protein Yield	3.53	2.20	1.10E-01	5.50E-01	
ss63780012:G>T	0.41 (T)	NM_001024563.1 - 1756	Fat Content	5.0	0.019	9.13E-03	4.57E-02	silent (Leucine)
			Milk Yield	-197	66.4	3.33E-03	1.66E-02	
			Protein Yield	-3.77	1.95	5.44E-02	2.72E-01	
			Protein Content	2.5	0.010	1.63E-02	8.17E-02	

NCBI Accession	MAF[1]	GenBank Accession No and Base Position (Btau_3.1)	Trait[2]	Estimate[3]	SE	*P*-value	FDR[4]	Type of mutation
ss63780014 C>T	0.46 (T)	NM_001024563.1 - 1633	Protein Content	1.2	0.010	2.34E-01	1.17E+00	silent (aspartic acid)

[1]MAF = minor allele frequency

[2]Milk yield= volume of milk (kg), Protein content = volume of protein (kg), Fat yield = volume of fat (kg), Protein content = percentage of protein (%), Fat content = percentage of fat (%)

[3] Estimate of the effect expressed in units of the trait.

[4]False discovery rate calculated as $FDR = mP_{(i)} / I$, where m is the total number of tests and $P_{(i)}$ is the *P*-value at rank *i* when the *P*-values are ranked from lowest to highest (Benjamini and Hochberg, 1995; Weller *et al.*, 1998).

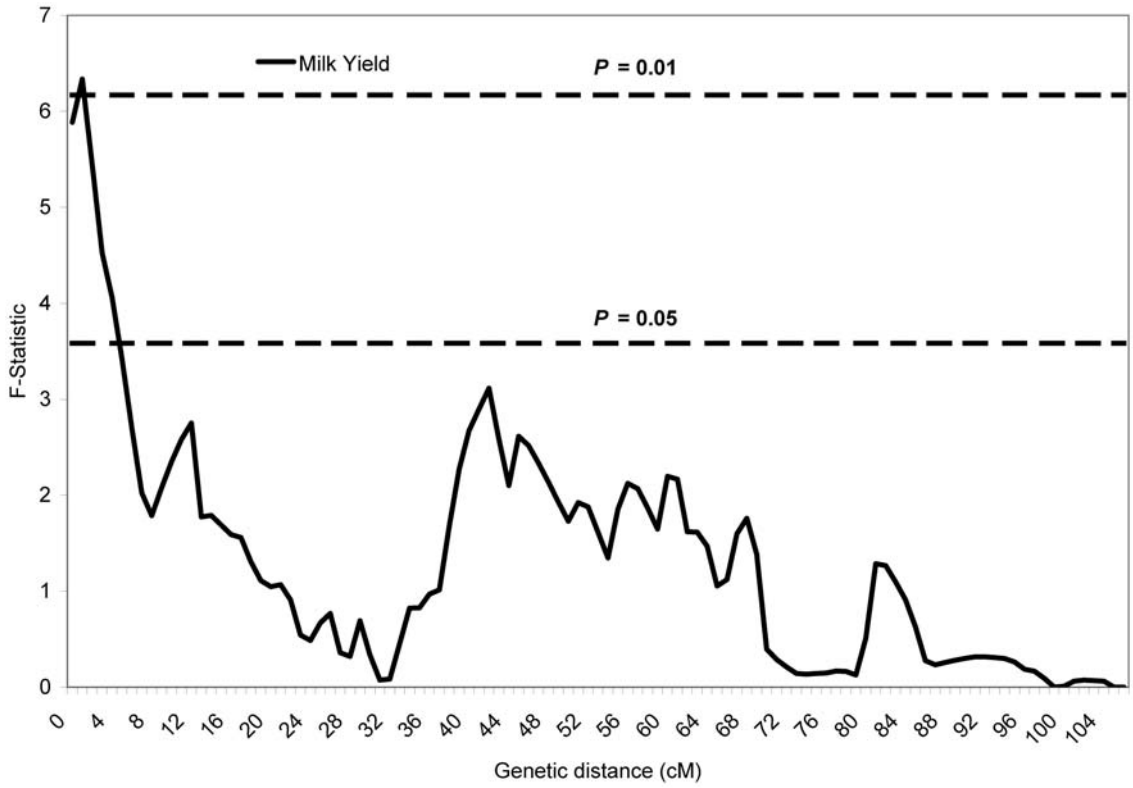

Figure 6.1: Across family F-statistic milk yield quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms excluding D*GAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

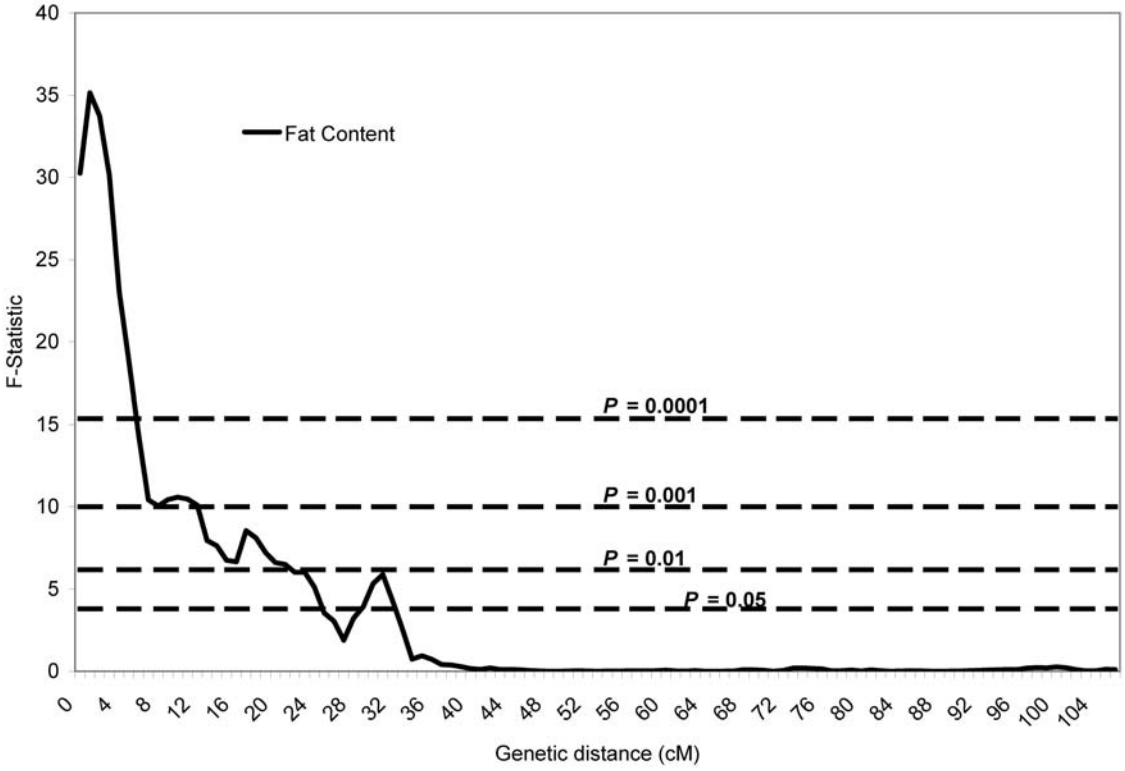

Figure 6.2: Across family F-statistic fat yield quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms excluding D*GAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

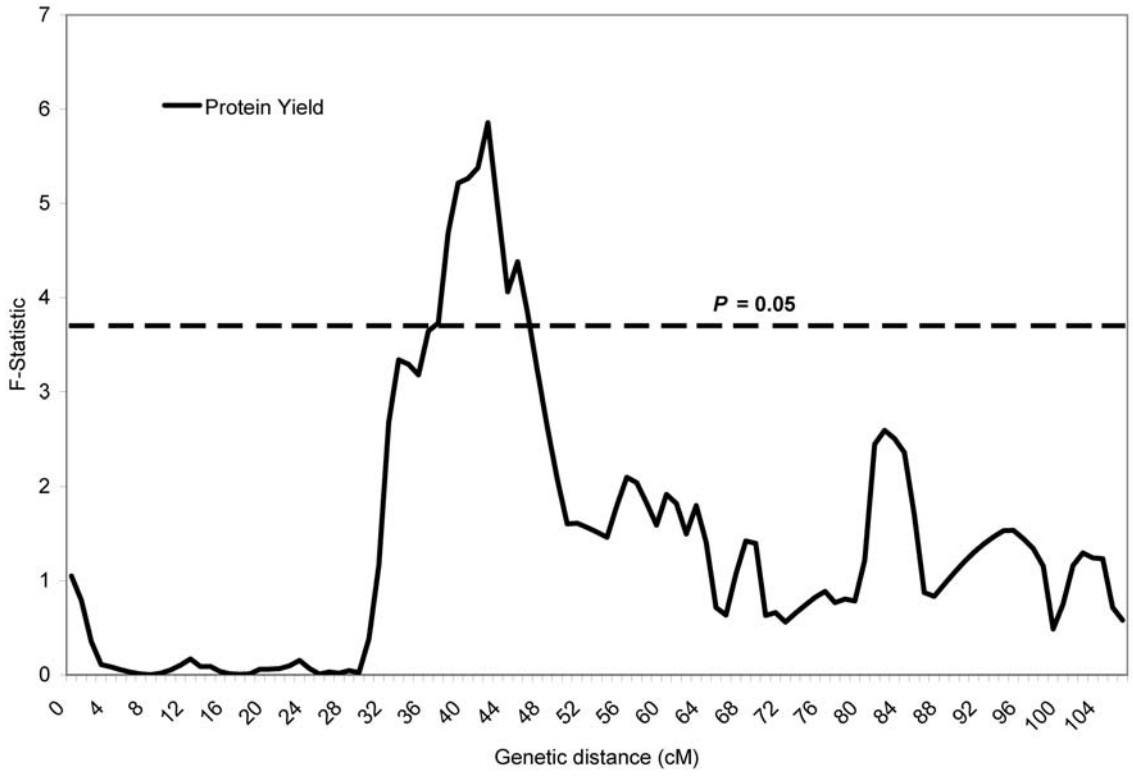

Figure 6.3: Across family F-statistic protein yield quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms excluding D*GAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

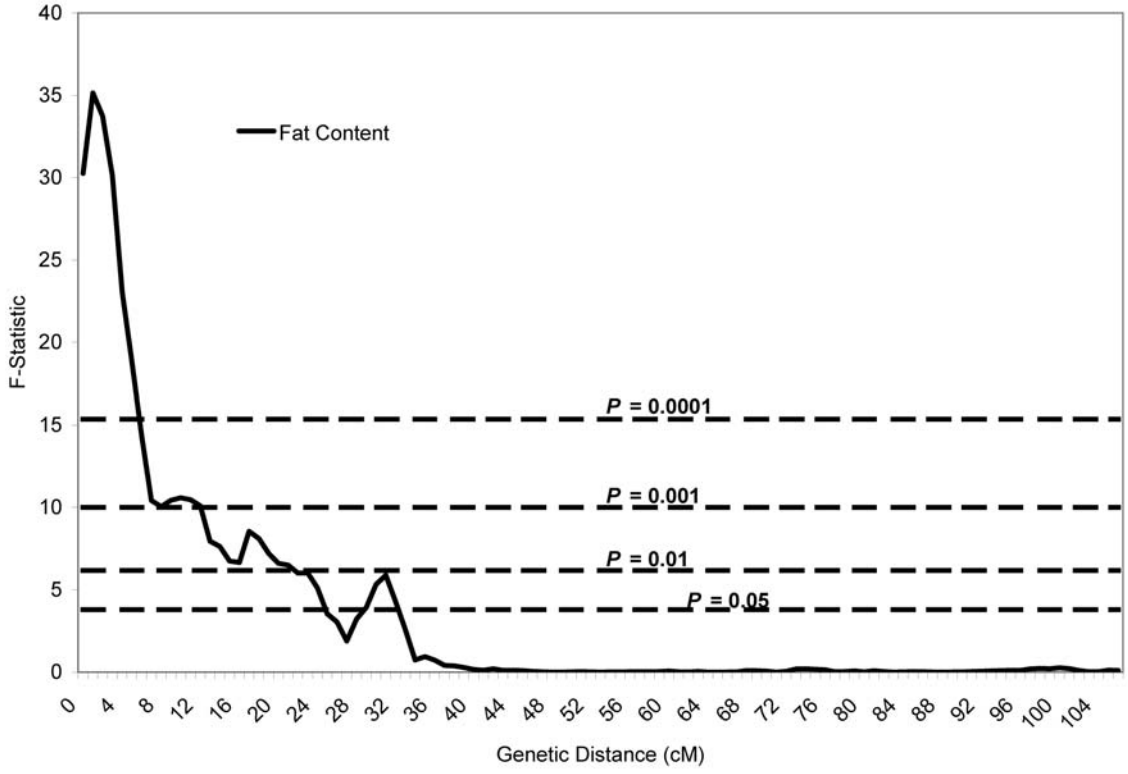

Figure 6.4: Across family F-statistic fat content quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms excluding D*GAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

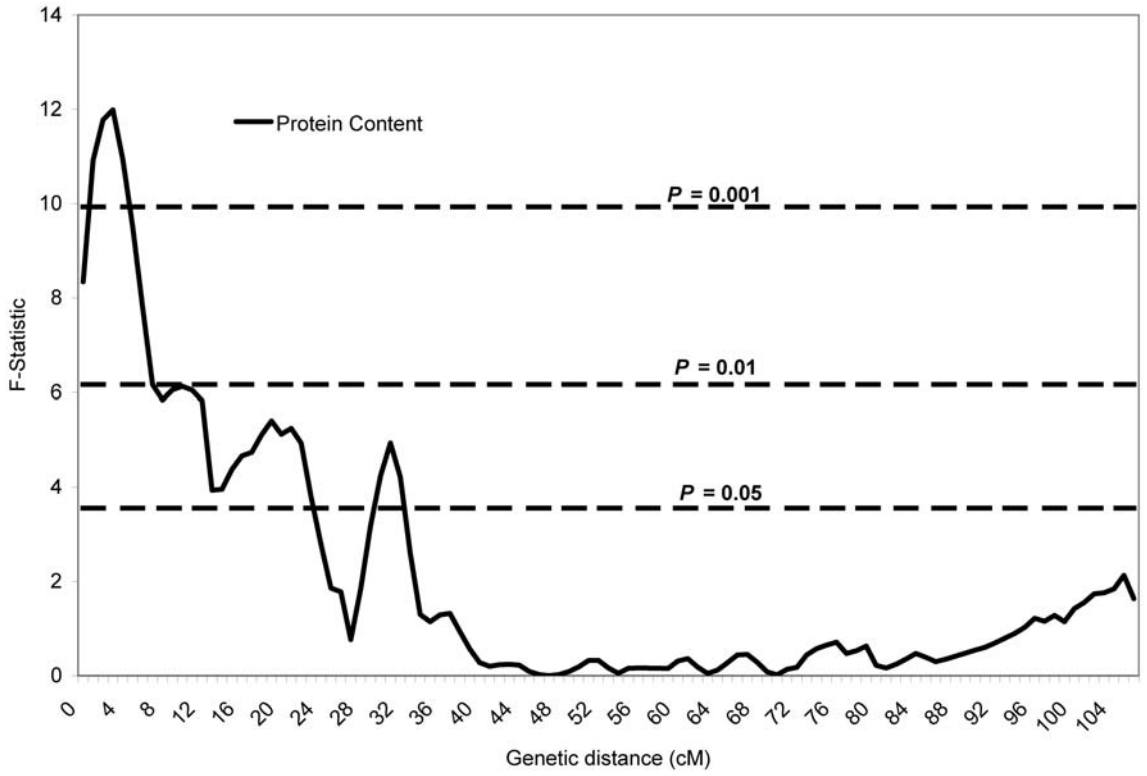

Figure 6.5: Across family F-statistic protein content quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms excluding D*GAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

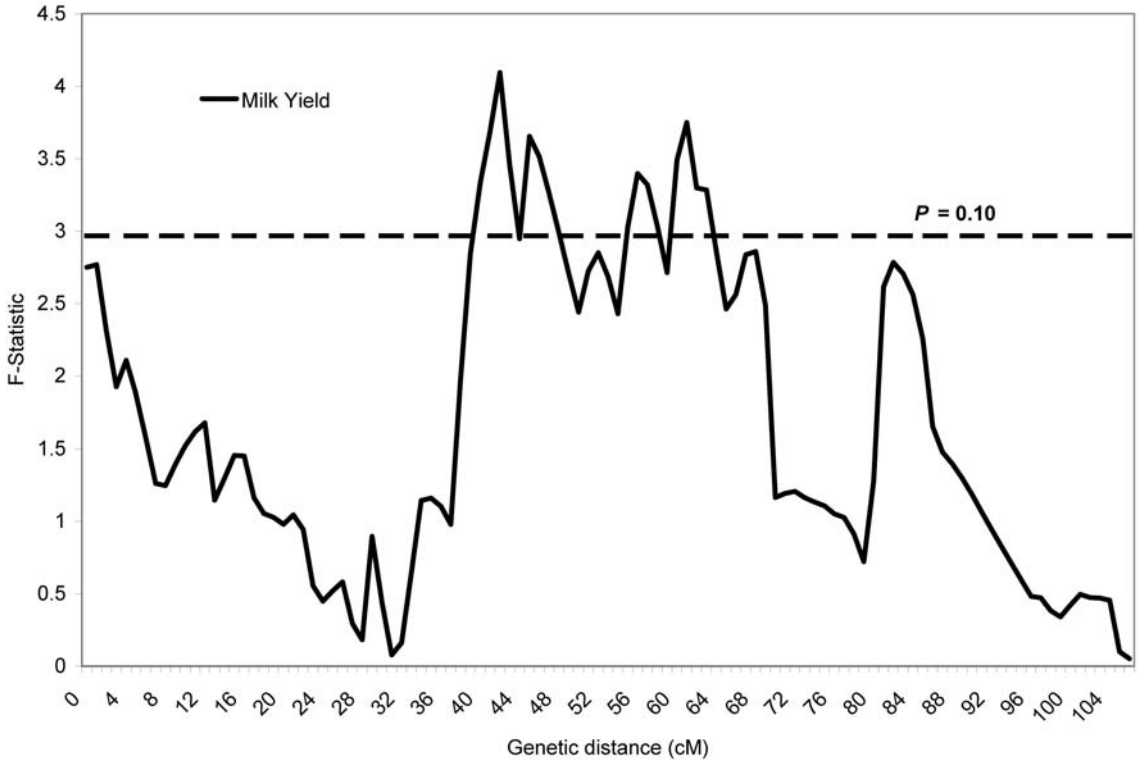

Figure 6.6: Across family F-statistic milk yield quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms and including *DGAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

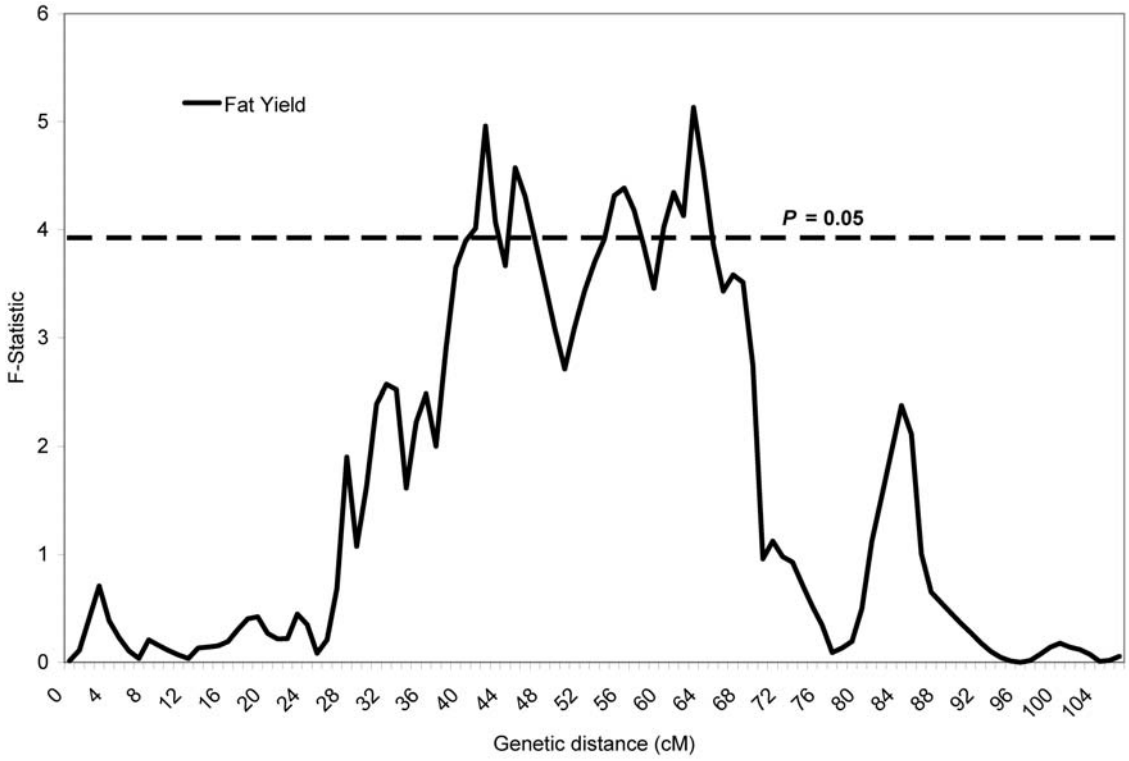

Figure 6.7: Across family F-statistic fat yield quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms and including *DGAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

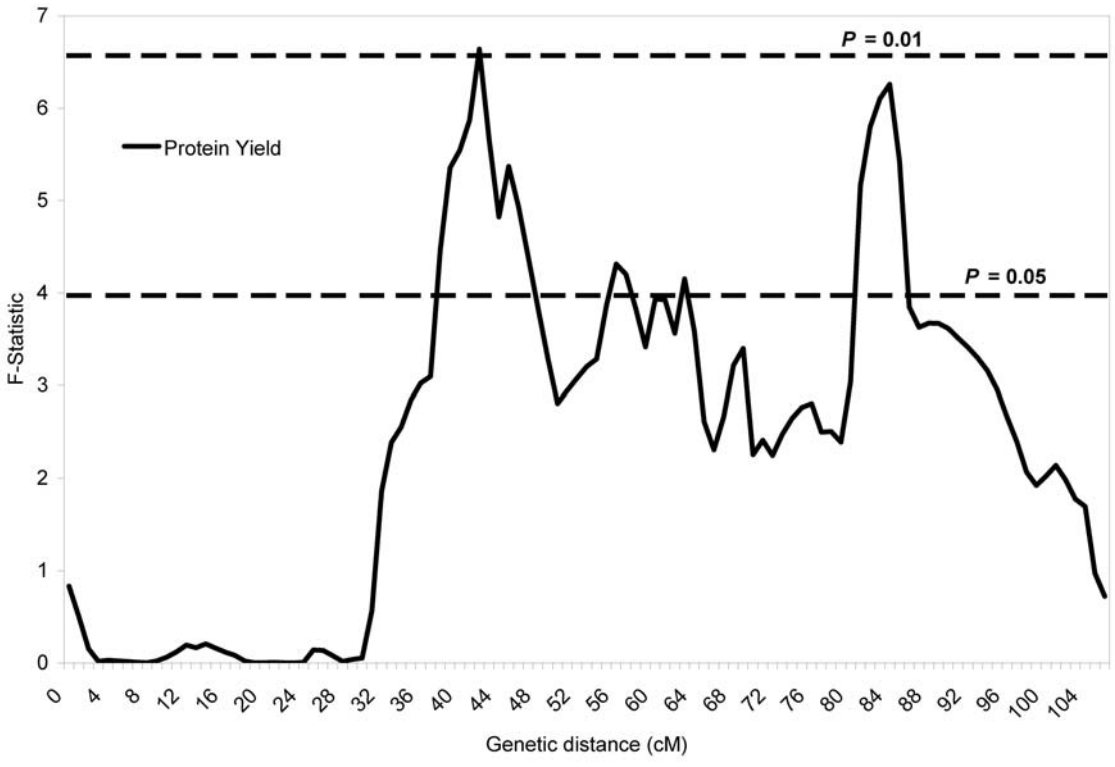

Figure 6.8: Across family F-statistic protein yield quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms and including *DGAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

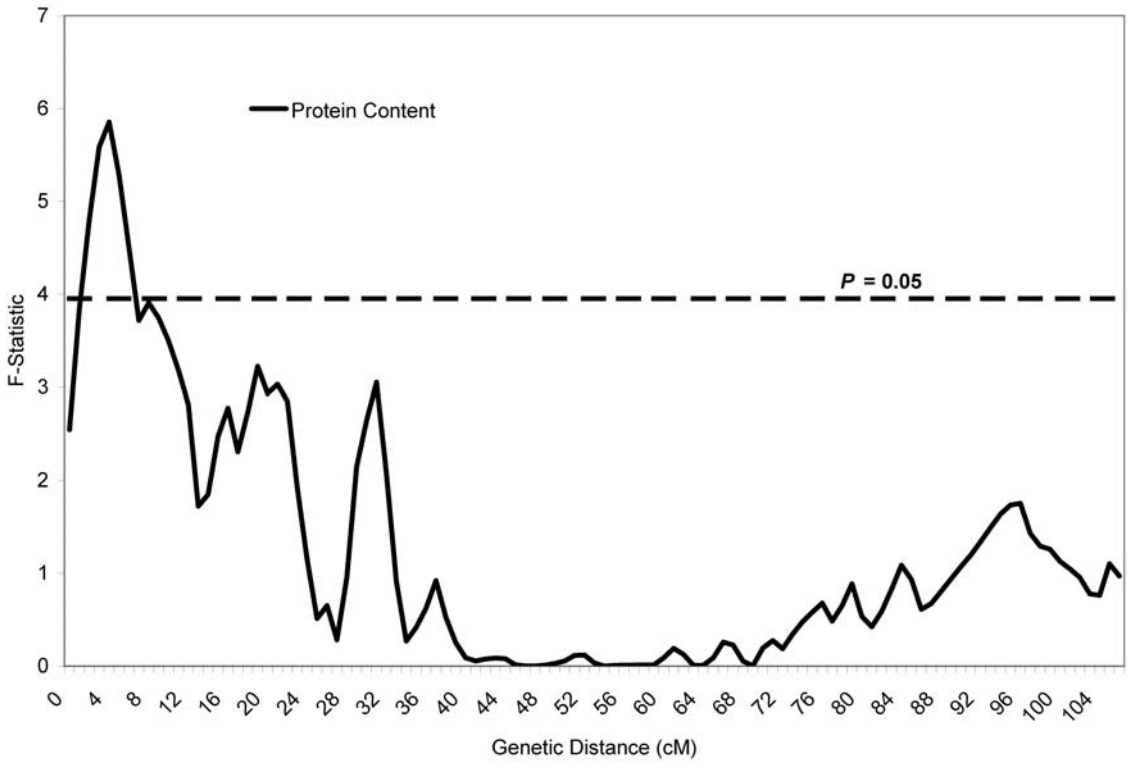

Figure 6.9: Across family F-statistic fat content quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms and including *DGAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

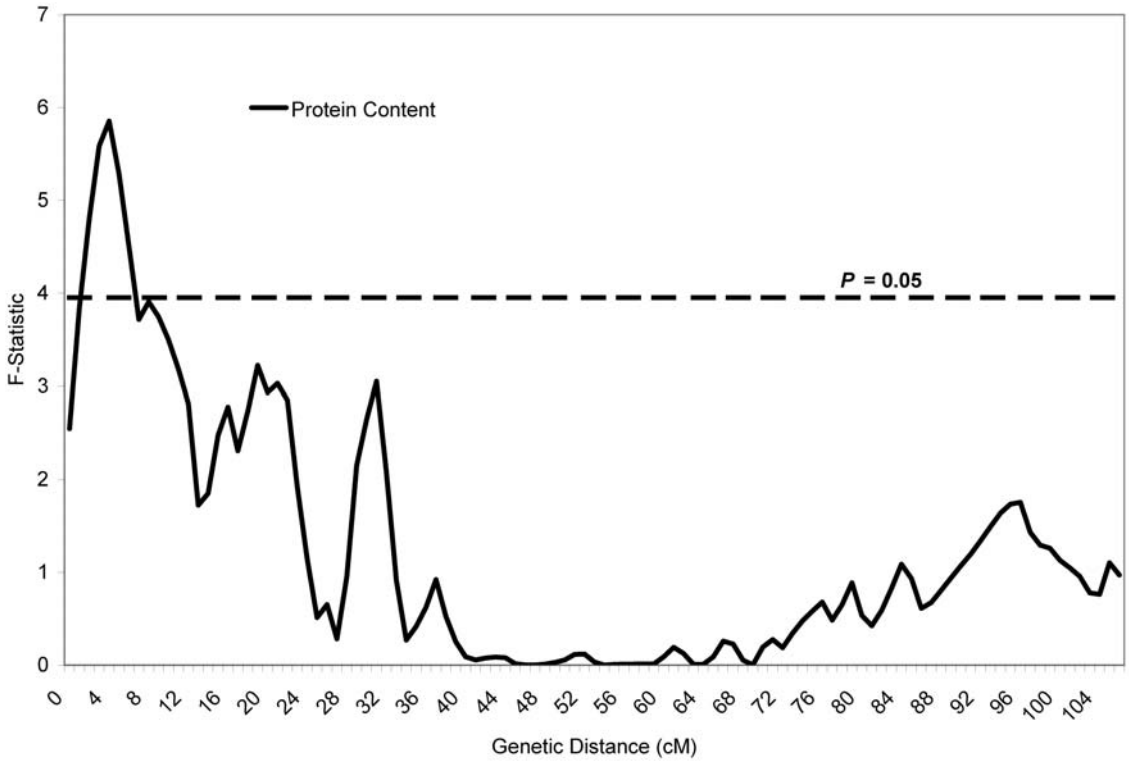

Figure 6.10: Across family F-statistic protein content quantitative trait loci profile on bovine chromosome 14 using 139 single nucleotide polymorphisms and including *DGAT1* genotype as a genetic cofactor. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. *DGAT1* position was estimated to be at approximately 0.3 cM.

6.4 Literature Cited

Bailis J.M., Luche D.D., Hunter T. & Forsburg S.L. (2008) Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote s-phase genome stability. *Mol Cell Biol* **28**, 1724-38.

Beigneux A.P., Davies B.S., Gin P., Weinstein M.M., Farber E., Qiao X., Peale F., Bunting S., Walzem R.L., Wong J.S., Blaner W.S., Ding Z.M., Melford K., Wongsiriroj N., Shu X., de Sauvage F., Ryan R.O., Fong L.G., Bensadoun A. & Young S.G. (2007) Glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 plays a critical role in the lipolytic processing of chylomicrons. *Cell Metab* **5**, 279-91.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B-Methodological*, 57:289-300.

Bennewitz J., Reinsch N., Paul S., Looft C., Kaupe B., Weimann C., Erhardt G., Thaller G., Kuhn C., Schwerin M., Thomsen H., Reinhardt F., Reents R. & Kalm E. (2004) The DGAT1 K232A mutation is not solely responsible for the milk production quantitative trait locus on the bovine chromosome 14. *J Dairy Sci* **87**, 431-42.

Bochman M.L. & Schwacha A. (2007) Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step. *J Biol Chem* **282**, 33795-804.

Bulow H.E. & Bernhardt R. (2002) Analyses of the CYP11B gene family in the guinea pig suggest the existence of a primordial CYP11B gene with aldosterone synthase activity. *Eur J Biochem* **269**, 3838-46.

Centner T., Yano J., Kimura E., McElhinny A.S., Pelin K., Witt C.C., Bang M.L., Trombitas K., Granzier H., Gregorio C.C., Sorimachi H. & Labeit S. (2001) Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain. *J Mol Biol* **306**, 717-26.

Churchill G.A. & Doerge R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963-71.

Cohen-Zinder M., Seroussi E., Larkin D.M., Loor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I. & Ron M. (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 Res* **15**, 936-44.

Coppieters W., Riquet J., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Nezer C., Simon P., Vanmanshoven P., Wagenaar D. & Georges M. (1998) A QTL with major effect on milk yield and composition maps to bovine chromosome 14. *Mamm Genome* **9**, 540-4.

Dukes I.D. & Philipson L.H. (1996) K+ channels: generating excitement in pancreatic beta-cells. *Diabetes* **45**, 845-53.

Falconer D.S. & Mackay T.F. (1996) Introduction to quantitative genetics. *Longman Scientific and Technical, New York, NY* **4th edition**.

Gautier M., Capitan A., Fritz S., Eggen A., Boichard D. & Druet T. (2007) Characterization of the DGAT1 K232A and variable number of tandem repeat polymorphisms in French dairy cattle. *J Dairy Sci* **90**, 2980-8.

Georges M., Coppieters W. & Charlier C. (2007) Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease. *Curr Opin Genet Dev* **17**, 166-76.

Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M. & Snell R. (2002) Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res* **12**, 222-31.

Hollinger S. & Hepler J.R. (2002) Cellular regulation of RGS proteins: modulators and integrators of G protein signaling. *Pharmacol Rev* **54**, 527-59.

Ioka R.X., Kang M.J., Kamiyama S., Kim D.H., Magoori K., Kamataki A., Ito Y., Takei Y.A., Sasaki M., Suzuki T., Sasano H., Takahashi S., Sakai J., Fujino T. & Yamamoto T.T. (2003) Expression cloning and characterization of a novel glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein, GPI-HBP1. *J Biol Chem* **278**, 7344-9.

Kaplan D.L., Davey M.J. & O'Donnell M. (2003) Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex. *J Biol Chem* **278**, 49171-82.

Kaupe B., Brandt H., Prinzenberg E.M. & Erhardt G. (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. *J Anim Sci* **85**, 11-21.

Kaupe B., Winter A., Fries R. & Erhardt G. (2004) DGAT1 polymorphism in Bos indicus and Bos taurus cattle breeds. *J Dairy Res* **71**, 182-7.

Knott S.A., Elsen J.M. & Haley C.S. (1996) Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theoretical and Applied Genetics*, 93: 71-80.

Marques E., De Givry S., Stothard P., Murdoch B., Wang Z., Womack J. & Moore S. (2007) A high resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. *BMC Genomics* **8**, 254.

Marques E., Schnabel R.D., Stothard P., Kolbehdari D., Wang Z., Taylor J.F. & Moore S.S. (2008) High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle. *BMC Genet* **9**, 45.

Naslund J., Fikse W.F., Pielberg G.R. & Lunden A. (2008) Frequency and effect of the bovine acyl-CoA:diacylglycerol acyltransferase 1 (DGAT1) K232A polymorphism in Swedish dairy cattle. *J Dairy Sci* **91**, 2127-34.

Riquet J., Coppieters W., Cambisano N., Arranz J.J., Berzi P., Davis S.K., Grisart B., Farnir F., Karim L., Mni M., Simon P., Taylor J.F., Vanmanshoven P., Wagenaar D., Womack J.E. & Georges M. (1999) Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle. *Proc Natl Acad Sci U S A* **96**, 9252-7.

Ron M., Heyen D.W., Weller J.I., Band M., Feldmesser E., Pasternak H., Da Y., Wiggans G.R., Vanraden P., Ezra E. & Lewin H.A. (1998) *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, 6WCCALP Congress, Armidale, Australia*, 422-5.

Schnabel R.D., Kim J.J., Ashwell M.S., Sonstegard T.S., Van Tassell C.P., Connor E.E. & Taylor J.F. (2005a) Fine-mapping milk production quantitative trait loci on BTA6: analysis of the bovine osteopontin gene. *Proc Natl Acad Sci U S A* **102**, 6896-901.

Schnabel R.D., Sonstegard T.S., Taylor J.F. & Ashwell M.S. (2005b) Whole-genome scan to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families. *Anim Genet* **36**, 408-16.

Seasholtz A.F., Valverde R.A. & Denver R.J. (2002) Corticotropin-releasing hormone-binding protein: biochemistry and function from fishes to mammals. *J Endocrinol* **175**, 89-97.

Seaton G., Haley C.S., Knott S.A., Kearsey M. & Visscher P.M. (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339-40.

Snelling W.M., Chiu R., Schein J.E., Hobbs M., Abbey C.A., Adelson D.L., Aerts J., Bennett G.L., Bosdet I.E., Boussaha M., Brauning R., Caetano A.R., Costa M.M., Crawford A.M., Dalyrmple B.P., Eggen A., Everts-van der Wind A., Floriot S., Gautier M., Gill C.A., Green R.D., Holt R., Jann O., Jones S.J., de Jong P.J., Kappes S.M., Keele J.W., Larkin D.M., Lewin H.A., McEwan J.C., McKay S., Marra M.A., Mathewson C.A., Matukumalli L.K., Moore S.S., Murdoch B., Nicholas F., Osoegawa K., Roy A., Salih H., Schibler L., Schnabel R., Silveri L., Skow L.C., Smith T.P., Sonstegard T.S., Taylor J., Tellam R., Van Tassell C.P., Williams J.L., Womack J.E., Wye N.H., Yang G. & Zhao S. (2007) A physical map of the bovine genome. *Genome Biol* **8**, R165.

Spelman R.J., Ford C.A., McElhinney P., Gregory G.C. & Snell R.G. (2002) Characterization of the DGAT1 gene in the New Zealand dairy population. *J Dairy Sci* **85**, 3514-7.

Thaller G., Kramer W., Winter A., Kaupe B., Erhardt G. & Fries R. (2003) Effects of DGAT1 variants on milk production traits in German cattle breeds. *J Anim Sci* **81**, 1911-8.

Vlahcevic Z.R., Pandak W.M. & Stravitz R.T. (1999) Regulation of bile acid biosynthesis. *Gastroenterol Clin North Am* **28**, 1-25, v.

Weller J.I., Song J.Z., Heyen D.W., Lewin H.A. & Ron M. (1998) A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. *Genetics* **150**, 1699-706.

Wibowo T.A., Michal J.J. & Jiang Z. (2007) Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. *Genome* **50**, 939-45.

Widmann C., Gibson S., Jarpe M.B. & Johnson G.L. (1999) Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. *Physiol Rev* **79**, 143-80.

Winter A., Kramer W., Werner F.A., Kollers S., Kata S., Durstewitz G., Buitkamp J., Womack J.E., Thaller G. & Fries R. (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. *Proc Natl Acad Sci U S A* **99**, 9300-5.

CHAPTER 7

Polymorphisms in positional candidate genes on BTA14 and BTA26 affect carcass quality in beef cattle[4]

7.1 Introduction

Identification of polymorphisms associated with economically relevant traits in cattle is crucial for understanding the mechanisms underlying their genetic variation. Chromosomes known to harbor meat quality trait QTL such as bovine chromosome 14 (**BTA14**) and bovine chromosome 26 (**BTA26**) (Stone *et al.* 1999; Moore *et al.* 2003; Casas *et al.* 2004; Mizoshita *et al.* 2005) are primary sites for the presence of functionally important genes affecting lipid metabolism.

Among the important genes on BTA26 lies the fibroblast growth factor 8 (***FGF8***). It has been linked to a number of quantitative trait loci affecting obesity in mice, indicating its potential for regulating adiposity in other species. Because of its consistent links to obesity QTL in mice, it was suggested that it might act as a master regulator or interacting element controlling multiple genes that contribute to adiposity in mice (Stylianou *et al.* 2006).

Two genes on BTA14 have been linked to effects on lipid metabolism in other species: 2, 4 dienoyl CoA reductase 1 (***DECR1;*** Amills *et al.* (2005) and core binding factor, runt domain, alpha subunit 2; translocated to 1 gene (***CBFA2T1***; Wolford *et al.* (1998). In pigs, *DECR1* mapped under a linoleic acid content QTL located on chromosome 4 (Perez-Enciso *et al.* 2000) and sequencing

[4] A version of this chapter was accepted for publication. J. Anim Sci. Apr 24. [Epub ahead of print]

analysis identified 2 single nucleotide polymorphisms (**SNPs**) showing associations with linoleic content (Amills *et al.* 2005). In 2006, *CBFA2T1* was part of the human obesity map (Rankinen *et al.* 2006) for being associated with fat percentage in studies in Pima Indian males (Wolford *et al.* 1998).

The association between these genes and lipid metabolism in other species make them plausible candidates when searching for associations in cattle and perhaps contributing to meat quality QTL peaks observed on BTA14 and BTA26. The objectives of this study were to identify polymorphisms in candidate genes previously reported to affect lipid metabolism in other species and to evaluate their associations with meat production traits in cattle.

7.2 Materials and Methods

7.2.1 *Animals and Management*

Four hundred and sixty four steers from 28 half sib families from an experimental line of Angus, Charolais, or Alberta Crossbred Bulls and the University of Alberta's Crossbred dam line previously described by Nkrumah *et al.* (2004) were used in this study. The dam line was produced from crosses among 3 composite cattle lines, namely beef synthetic 1, beef synthetic 2, and dairy X beef synthetic. Beef synthetic 1 was composed of 33% each of Angus and Charolais, approximately 20% Galloway, and the remainder from other breeds. Beef synthetic 2 was comprised of approximately 60% Hereford and 40% other beef breeds. The dairy X beef line was composed of approximately 60% dairy breeds (Holstein, Brown Swiss or Simmental) and 40% beef breeds, mainly Angus and Charolais (Goonewardene *et al.* 2003). The animal test diets were the

same for yr 2 and 3, but differed in yr 1 with the substitution of barley and oat grain with dry-rolled corn due to a shortage of feed barley in that particular year; however, both diets had similar ME content. Briefly, the test diet for yr 1 contained 80% dry-rolled corn, 13.5% alfalfa hay pellet, 5% feedlot supplement (32% CP beef supplement), and 1.5% canola oil. Year 2 and 3 diets contained 64% barley grain, 20% oat grain, 9% alfalfa hay pellet, 5% beef feedlot supplement, and 1.5% canola oil. Details of the animals' diets have been described in Nkrumah *et al.* (2004). Animals used in the study were cared for according to the guidelines of the Canadian Council on Animal Care (Canadian Council on Animal Care 1993).

7.2.2 Traits Analyzed

Ultrasound and carcass merit data were collected on beef steers over a period of 3 yr (November 2002 to June 2005). Carcass traits were evaluated according to the Canadian beef carcass grading system (Agriculture Canada 1992). Carcass and ultrasound measurements have been previously described by Nkrumah *et al.* (2004). Briefly, ultrasound measurements of 12th-/13th-rib fat depth (**UBF**), longissimus muscle area (**ULMA**), and marbling score were obtained with an Aloka 500V real-time ultrasound with a 17-cm, 3.5-MHz linear array transducer at 28-d intervals according to procedures described by (Brethour 1992). After these tests, animals were shipped to a commercial plant and carcass grade fat (**GRFAT**), carcass backfat (**CBF**), and longissimus muscle area (**LMA**) measurements were collected at the 12 th/13th rib following a 24-h chill at -4°C. Ultrasound and carcass marbling score are a measure of intramuscular fat, being

classified as 1 to < 2 units = trace marbling (Canada A quality grade); 2 to < 3 units = slight marbling (Canada AA quality grade); 3 to < 4 units = small to moderate marbling (Canada AAA quality grade); and >= 4 units = slightly abundant or more marbling (Canada Prime). Lean meat yield (**LMY**) is an estimate of saleable meat calculated according to Jones (1984). Yield grade (**YGRADE**) classes are based on the proportion of lean meat and are classified as YGRADE1 = >59%, YGRADE2 = 54 to 59%, and YGRADE3 = <54%.

7.2.3 DNA Isolation and Genotyping

A 10-mL blood sample from 464 steers and their sires was collected by jugular venipuncture from which genomic DNA was extracted using a standard saturated salt phenol/chloroform procedure. Sequences from *DECR1* (Gene ID: LOC509952), *CBFA2T1* (Gene ID: LOC538628), and *FGF8* (Gene ID: LOC326284) were BLASTed to the bovine genome assembly using the NCBI BLAST tool to design primers for both intronic and exonic regions of those genes. Primer design for *CBFA2T1*, *DECR1*, and *FGF8* sequences was carried out using primer3 (http://frodo.wi.mit.edu/) with the following settings (min opt max): primer size: 22 24 26; primer tm: 58 60 62; primer GC%: 40 50 60. Genomic DNA was amplified using standard PCR conditions. PCR products were subjected to a clean up stage consisting of an equal mixture of Exonuclease I and Shrimp Alkaline Phosphatase (2/1 concentration) enzymes (Invitrogen, Carlsbad, CA) for 15 min at 37°C and 15 min at 85°C. Clean PCR products were sequenced using BigDye-terminator chemistry (Applied Biosystems, Norwalk, CT) and a 3730

DNA sequencer (Applied Biosystems, Norwalk, CT). Genotyping of microsatellites was performed by automated fragment analysis using an ABI PRISM 3730 DNA sequencer (Applied Biosystems, Norwalk, CT). SNP genotyping was carried out using the Illumina GoldenGate assay on the BeadStation 500G Genotyping System (Illumina Inc., San Diego, CA).

7.2.4 Sequencing Analysis

After sequencing analysis was performed on test animals, sequence products were BLASTed to the bovine sequence assembly Btau_3.1 using the NCBI BLAST tool to verify that the correct gene sequences were being analyzed. Multiple sequence alignment for humans, mice, and bovine sequences was performed using the online tool ClustalW2 available at http://www.ebi.ac.uk/Tools/clustalw2/index.html.

7.2.5 Quantitative Trait Loci Analysis

A total of 75 SNP markers and 3 microsatellites (RM26, BM804, BM7237 were used for QTL analysis on BTA26, in addition to ss95214675:A>G *FGF8* SNP. Another QTL analysis was performed on BTA14 using 112 SNP markers and 11 microsatellites (DIK2359 BMS1747 NRKM-003 NRKM-052 RM011 DIK4730 DIK2570 BL1029 BMS947 BL1036 BMS1941), in addition to ss95214671:C>T *DECR1* SNP and ss95215669:G>T *CBFA2T1* SNP. These markers on both chromosomes were selected because they showed the highest number of heterozygous sires and minor allele frequency of 0.14 (*CBFA2T1*) and 0.37 (*DECR1* and *FGF8*) (Table 7.1). Marker locations across both chromosomes

were obtained by Snelling *et al.* (2007). Gene SNP locations were estimated by analyzing the location of their flanking markers. Both *DECR1* and *CBFA2T1* BLASTed between the same markers, and, therefore, were given similar cM estimates of approximately 92.7 cM on BTA14. *FGF8* was estimated to be around 27.0 cM on BTA26.

The total number of animals genotyped in our study was 464 belonging to 28 families. However, 396 animals belonging to 20 half-sib families (range 10 to 56 progeny) were used for QTL analysis. The remaining animals belonged to families with less than 2 animals and therefore were not utilized for the purpose of QTL analysis. QTL analysis performed in this study used the multiple marker interval mapping approach described by Knott *et al.* (1996). The conditional probabilities that a calf inherited the first allele of a putative QTL from its sire were obtained from QTL Express (Seaton *et al.* 2002), which uses the information from the closest informative flanking markers at 1 cM intervals. This analysis is similar to that used by de Koning *et al.* (1999) in which the conditional probabilities of inheriting the sire allele were nested within half-sib families. This is because not only the linkage phase between a marker and a QTL can differ between families, but also because not all sires are heterozygous for the QTL. In addition, sire effects were also included as random effects (Nagamine & Haley 2001; Nkrumah *et al.* 2007; Van Eenennaam *et al.* 2007). The conditional probabilities from QTL express were input into SAS (SAS Inst. Inc., Cary, NC) and QTL analysis was performed using the mixed model described by:

$$Y = \mathrm{X}\beta + Gs + Q\alpha + e\,,$$

where Y is a vector of observations on the progeny of each sire, X is the known incidence matrix relating observations to their fixed effect levels, β is the vector of fixed effects (breed, test batch, and age), G is the known incidence matrix relating observations to random sire effects, s is the vector of random additive polygenic effects of sires, Q is a vector of the conditional probabilities, at each interval, that a calf inherited the first allele of a putative QTL from a sire, α is the regression coefficient corresponding to the fixed allele substitution effect for a putative QTL within half-sib families. Significance thresholds at 5% and 1% were determined using 25,000 permutation tests in SAS (SAS Inst. Inc., Cary, NC) by randomly shuffling the phenotypic records of the 396 animals and maintaining the QTL probabilities unchanged, according to the procedure described by Nkrumah *et al.* (2007). The permutation procedure was carried out for when the SNPs were included or excluded from the QTL analysis. Exclusion of SNPs from the analysis was carried out by removing the SNP genotypes of all animals and obtaining new conditional probabilities for each calf according to procedures described above. The reported permutation threshold was an average between the thresholds for each trait when analyzed separately. The difference between them was at most 0.02, which was not enough to modify the significance of the QTL after the average was calculated. The same was performed for the inclusion and exclusion of the candidate gene SNPs.

7.2.6 Association Analysis

Associations of the genotypes for each polymorphism and carcass merit were analyzed by regressing phenotypes on genotypes using the MIXED procedure (SAS Inst. Inc., Cary, NC). Four hundred and sixty four animals were available for this analysis. The statistical analyses model included fixed effects of SNP genotype, test batch, breed, and age of animal at the beginning of the test, and random effects of sire of animal. Allele substitution effect was calculated by regressing phenotypes on the number of copies of one allele for each SNP using the mixed model procedure in SAS (SAS Inst. Inc., Cary, NC).

7.2.7 False Discovery Rate

A false discovery rate procedure was applied to our analysis to minimize false positives according to procedures previously described by Benjamini & Hochberg (1995) and Weller *et al.* (1998). Briefly, the procedure takes into consideration the number of tests performed, the ranking of the markers within the analysis, and their significance (*P*-value) rank from lowest to highest. Because FDR assumes independence between traits and because these traits are correlated, FDR was calculated within each trait according to the formula:

$$FDR = \frac{n \times P(k)}{k},$$

where k is the ranking of each marker, *P* is the P-value associated with the marker, and n is the number of markers analyzed.

7.3 Results and Discussion

7.3.1 Polymorphisms Detected

In total, 4 polymorphisms were detected in intronic regions of *CBFA2T1*. Single nucleotide polymorphisms detected in the *CBFA2T1* gene were initially BLASTed against the bovine sequence assembly mRNA reference sequence, with none of the SNPs BLASTing to this mRNA sequence. Initial BLASTing analysis of the first 2 SNPs showed sequence complimentarity to an unknown contig (NW_001502787.1). However, when the same sequence segment that did not contain the SNPs was compared to the mRNA reference sequence, this sequence BLASTed to exonic regions, implying that our sequence was indeed part of the *CBFA2T1* gene. It is known that there are still segments in BTA14 that are showing sequence complementarity to multiple regions or to unknown regions in the latest sequence assembly (Marques *et al.* 2007). In order to confirm that, in fact, our sequences were mis-assigned, *CBFA2T1* mRNA (NW_001099385) sequence was compared to the genomic sequence. Parts of the mRNA sequence were BLASTed to the same unassigned contig where the first 2 SNPs BLASTed, suggesting that the SNPs resided in the *CBFA2T1* gene.

Sequencing analysis of *DECR1* identified 9 SNPs. Sequences were BLASTed to the bovine sequence assembly to determine if they were coding or non-coding polymorphisms. Four of the SNPs were located on exonic regions of *DECR1*. Two of these SNPs were shown to change the amino acid constitution of *DECR1*: isoleucine to valine and valine to methionine substitutions, whereas the other 2 were silent polymorphisms (Table 7.1). Multiple sequence alignment

between human, mouse, and bovine *DECR1* showed that all 4 amino acid changing polymorphisms are found in conserved sequence regions across those species (Figure 7.1).

Sequencing results of *FGF8* detected 4 SNPs. One of the SNPs was intronic, and the other 3 exonic. Two of the exonic SNPs produced amino acid substitutions (glycine to arginine and glutamic acid to lysine). Only 1 exonic SNP produced a silent mutation (glycine to glycine). Table 7.1 summarizes the list of SNPs, including the allele frequency and location in the genome. Multiple sequence alignment between human, mouse, and bovine *FGF8* showed that none of the SNPs reported are part of the conserved sequence across those species.

7.3.2 Association Analysis

Single locus association analysis for *CBFA2T1* showed that 2 out of the 4 SNPs in that gene were associated with at least 1 meat quality trait. ss95214667:C>G was found to be significant ($P = 0.012$) with ULMA and ss95215669:G>T with UBF($P = 0.019$) and ULMA ($P = 0.006$). Table 7.2 lists the trait estimates, overall P-value, and the allele substitution effects for significant SNPs.

The same regression analysis was performed for *DECR1*. Results showed associations for 6 of the 9 SNPs with UBF ($P = 0.010$ to $P = 0.026$) (Table 7.2). Association tests were also performed on other meat traits, with no significant levels being achieved. Other significant associations were observed between 2 of the SNPs in *FGF8* and other meat traits. For example, ss95214675:A>G showed

significant associations with LMY (P = 0.005), CBF (P=0.004), GRFAT (P = 0.011), and LMA (P = 0.005), whereas ss95214676:C>G presented significant associations with UBF (P = 0.048), GRFAT (P = 0.033), LMY (P = 0.042), and LMA (P = 0.005).

This study evaluated 17 SNPs across 9 traits and detected 16 significant associations for the fixed effects model and 9 significant associations for the allele substitution model, both at $P < 0.05$. Because the meat quality traits used in this study were correlated, the most appropriate procedure was to perform FDR tests within each trait, as presented in Table 7.2. It is important to note that the candidate genes presented here were selected based on their known functions and evidence of links to lipid metabolism pathways.

Minor allele frequencies of the SNPs identified ranged from 0.28 to 0.49 for *DECR1*, 0.07 to 0.15 for *CBFA2T1*, and 0.01 to 0.37 for *FGF8*. Because of the low minor allele frequency for some of these SNPs, it is expected that some of the associations reported are biased. For instance, ss95215669:G>T has a minor allele frequency of 0.14, meaning that among 464 animals analyzed, 9 are homozygous GG; which could explain the non-significant association for this SNP for the allele substitution effect (Table 7.2).

Single point polymorphisms conferring amino acid substitutions were further evaluated to indicate what types of interactions with other proteins or receptors are being affected. In the case of both *DECR1* and *FGF8*, the encountered amino acid substitutions are of further interest, because the amino acids involved have different properties. Substitutions involving amino acids with

similar properties such as isoleucine and valine will most likely not cause major changes to the protein. Both valine and isoleucine are hydrophobic and possess an additional non-hydrogen substituent attached to their Cβ carbon. The other substitution - valine to methionine – may have a bigger impact. Even though valine and methione are apolar, methionine is less lipophilic because of its thiogroup, which might reduce stability overall. Functional studies can elucidate how these changes in amino acid sequence will affect the functions of *DECR1* and *FGF8* in cattle, and, in turn, why and how they affect variation in the traits studied.

7.3.3 Quantitative Trait Loci Analysis

Quantitative trait loci analysis was performed to further validate the effects of the SNPs on the traits under study. The objective of this analysis was to determine if the polymorphisms discovered were near a QTL peak, as well as to evaluate any changes in the QTL peak.

A QTL scan was performed using a set of 123 markers previously genotyped in our beef population, in addition to 1 *DECR1* SNP (see Materials and Methods). This QTL analysis on BTA14 yielded significant results for UBF at 91 cM, LMY at 86 cM, GRFAT at 15 cM, and YGRADE at 87 cM, all at the $P < 0.05$ level (Figure 7.2). These results are consistent with previously reported QTL analysis on this chromosome (MacNeil & Grosz 2002; Moore *et al.* 2003; Casas *et al.* 2004). In addition, yield grade and ultrasound backfat showed the highest F-value only 2 to 4 cM from *DECR1* (Figure 7.2 and 7.3). For yield grade, the most significant location was 87 cM, whereas for UBF it was at 91 cM (*DECR1*

location: 92.7 cM*)*. This is consistent with association results between *DECR1* and UBF in single locus analysis (Table 7.2).

When *DECR1* was subsequently removed from the QTL analysis, no UBF QTL was present at this location (Figure 7.4), providing additional support for the effect of *DECR1* SNPs on this trait. The same was done for YGRADE and GRFAT. In these cases, the QTL peaks shifted to 80 cM for YGRADE (Figure 7.3) compared to the previous peaks, with no change for GRFAT. Evaluation of the SNP effect on the UBF QTL included using SNP ss95214671:C>T as a covariate in the analysis. This step resulted in a decrease in threshold for this QTL (F statistic = 2.03 vs. 1.83), indicating that this term explained some of the QTL variation. When *CBFA2T1* SNPs were included in the analysis, similar QTL profiles were observed. This could also be due to the effects of *DECR1* SNPs, because both genes are relatively close to each other.

A QTL scan was also performed on BTA26 using 78 markers in addition to 1 of the *FGF8* SNPs (see Materials and Methods). The results showed the presence of peaks for LMY at 2 cM and for YGRADE at 25 cM, both at $P < 0.01$, and for CBF at 25 cM ($P < 0.05$) (Figure 7.5). When *FGF8* was subsequently removed from the analysis, LMY was the only one that did not change. A YGRADE QTL was still present, although at a lower F-value (2.04 compared to 2.51) (Figure 7.6). No QTL was observed for CBF at this position when the *FGF8* SNP was removed (Figure 7.7). Using SNP ss95214675:A>G genotype as a fixed effect, along with the most significant QTL position in CBF, resulted in a non-significant QTL ($P = 0.0524$), indicating that the SNP had a significant

contribution in this QTL. When SNPs were used as covariates for YGRADE and LMY, the QTL significance did not change.

The mitochondrial enzyme encoded by *DECR1* participates in the β-oxidation pathway catalyzing the reduction of trans-2-cis-4-dienoyl-CoA to 3-enoyl-CoA (Kunau & Dommes 1978) and it is therefore an interesting candidate influencing the genetic variation observed in meat quality. The *FGF8* androgen induced property was first discovered in earlier experiments by Tanaka *et al.* (1992). That study reported that a mouse mammary carcinoma cell line was stimulated to secrete a number of FGFs when induced by androgens. These FGFs, in turn, demonstrated growth like properties on this carcinoma cell line. Isolation and characterization of the activity determined that *FGF8* contributed to some of the growth effects. In humans, it is present in increased levels in breast cancer cells (Zammit *et al.* 2002) and its 22-protein family has a wide range of effects from wound healing repair (Clarke *et al.* 1993; Cuevas 1998) to tumorigenesis (Davies *et al.* 1996). In cattle, there are no reports linking *FGF8* to variations in carcass or meat quality. However, other studies (Stone *et al.* 1999; Casas *et al.* 2004) have also reported the presence of QTL affecting yield grade and grade fat on BTA26. In the literature, there is increasing evidence on the influence of FGF receptors on the regulation of glucose and lipid homeostasis (Hart *et al.* 2000; Huang *et al.* 2007). With the bovine gene annotation project being carried out to verify and confirm certain gene structures, it is likely that more information will come out for *FGF8*, including a receptor binding site. Combining the linkage and functional analysis of these candidate genes, is still necessary to conclude that at

least one of the polymorphisms detected and analyzed here are major contributors to the variation observed in carcass quality.

Nonetheless, there are several pieces of information that point to one of the polymorphisms in *DECR1* as a major contributor to the variation observed in backfat thickness. 1) This mutation affects the amino acid composition (valine to methionine) at a highly conserved region among humans, mice, and cattle. 2) The QTL becomes significant when *DECR1* is added to the analysis. 3) The QTL is partially removed when correcting the analysis for the effect of *DECR1* SNP ss95214671:C>T. 4) *DECR1* participates in the β-oxidation pathway reducing trans-2-cis-4-dienoyl to 3-enoyl-CoA (Kunau & Dommes 1978). The evidence for *FGF8* being a causative mutation is not as strong as for *DECR1*, because the function of *FGF8* is not clearly linked to a lipid metabolism pathway in cattle and its polymorphisms did not affect a conserved region among the three species.

Table 7.1. Summary of *DECR1*, *CBFA2T1*, and *FGF8* genes including dbSNP, nucleotide position, and SNP alleles

dbSNP	Gene Name	BTA	GenBank Accession No and Base Position (Btau_3.1)	Type of Change	Minor Allele Frequency
ss95214665:C>T	*CBFA2T1*	14	NW_001502787.1 - 83885	intron	0.07 (C)
ss95214666:G>T	*CBFA2T1*	14	NW_001502787.1 - 84976	intron	0.15 (C)
ss95215669:G>T	*CBFA2T1*	14	NW_001493258.14 - 665697	intron	0.14 (G)
ss95214667:C>G	*CBFA2T1*	14	NW_001493258.14 - 665710	intron	0.13 (G)
ss95214668:C>T	*DECR1*	14	NW_001493259.1 – 733467	intron	0.28 (C)
ss95214669:C>T	*DECR1*	14	NW_001493259.1 – 733476	intron	0.45 (C)
ss95214757:C>T	*DECR1*	14	NW_001493259.1 – 733677	isoleucine to valine	0.49 (T)
ss95214758:G>T	*DECR1*	14	NW_001493259.1 – 733738	alanine to alanine	0.43 (T)
ss95214759:A>G	*DECR1*	14	NW_001493259.1 – 733869	intron	0.43 (A)
ss95214670:C>G	*DECR1*	14	NW_001493259.1 – 736133	alanine to alanine	0.49 (G)
ss95214671:C>T	*DECR1*	14	NW_001493259.1 – 736141	valine to methione	0.37 (T)
ss95214672:C>T	*DECR1*	14	NW_001493259.1 – 736359	intron	0.49 (C)
ss95214673:A>T	*DECR1*	14	NW_001493259.1 – 736398	intron	0.49 (A)
ss95214674:G>T	*FGF8*	26	NW_001494359.1 - 166776	glycine to glycine	0.14 (T)
ss95214675:A>G	*FGF8*	26	NW_001494359.1 - 166804	glutamic acid to lysine	0.37 (G)
ss95214676:C>G	*FGF8*	26	NW_001494359.1 - 167062	intron	0.23 (C)
ss95214677:C>G	*FGF8*	26	NW_001494359.1 - 170040	glycine to argininine	0.01 (C)

Table 7.2. Estimates of allele substitution effect for meat quality in beef cattle for SNP on Core-binding factor, runt domain, alpha subunit 2; translocated to 1 (*CBFA2T1*) , mitochondrial 2,4-dienoyl CoA reductase1 (*DECR1*), and fibroblast growth factor 8 (*FGF8*).

				Fixed effect		Allele substitution Effect[4]			
SNP	Gene	BTA	Trait[1]	Overall *P*-value[2]	*FDR*[3]	Estimate[5]	SE	*P*-value	*FDR*
ss95215669:G>T	*CBFA2T1*	14	UMAR	0.019	0.318	0.027	0.045	0.544	NA[6]
			ULMA	0.006	0.099	-0.755	0.798	0.345	0.029
ss95214667:C>G	*CBFA2T1*	14	ULMA	0.012	0.099	1.036	0.821	0.208	0.029
ss95214758:G>T	*DECR1*	14	UBF	0.016	0.075	0.220	0.101	0.029	0.117
ss95214759:A>G	*DECR1*	14	UBF	0.012	0.075	-0.239	0.099	0.017	0.117
ss95214670:G>C	*DECR1*	14	UBF	0.018	0.075	-0.203	0.099	0.040	0.117
ss95214671:C>T	*DECR1*	14	UBF	0.026	0.075	-0.340	0.138	0.014	0.117
ss95214672:C>T	*DECR1*	14	UBF	0.013	0.075	0.206	0.101	0.041	0.117
ss95214673:A>T	*DECR1*	14	UBF	0.010	0.075	0.211	0.101	0.038	0.117
ss95214675:A>G	*FGF8*	26	LMY	0.005	0.353	-0.780	0.281	0.006	0.099
			CBF	0.004	0.070	0.957	0.307	0.002	0.033
			GRFAT	0.011	0.281	0.830	0.306	0.007	0.350
ss95214676:C>G	*FGF8*	26	UBF	0.048	0.075	0.142	0.110	0.201	0.117
			GRFAT	0.033	0.281	0.738	0.353	0.041	0.350
			LMY	0.042	0.353	-0.544	0.328	0.099	0.099
			LMA	0.005	0.084	-1.182	0.713	0.099	0.099

[1]UMAR = ultrasound marbling (score), ULMA = ultrasound longissimus muscle area (cm^2), UBF = ultrasound backfat (mm), LMY = carcass lean meat yield (%), CBF = carcass backfat (mm), GRFAT = carcass gradefat (mm), LMA = carcass longissimus muscle area (cm^2)

[2] *P*-value from overall F-test.

[3]False discovery rate calculated as $FDR = mP_{(i)} / I$, where m is the total number of tests and $P_{(i)}$ is the *P*-value at rank *i* when the *P*-values are ranked from lowest to highest (Weller et al., 1998).

[4] Allele substitution effect was calculated by regressing phenotypes on the number of copies of one allele for each SNP.

[5] Estimate of the effect expressed in units of the trait.

[6] NA = *P*-values were not significant at the $P < 0.05$ level; therefore, *FDR* was not calculated.

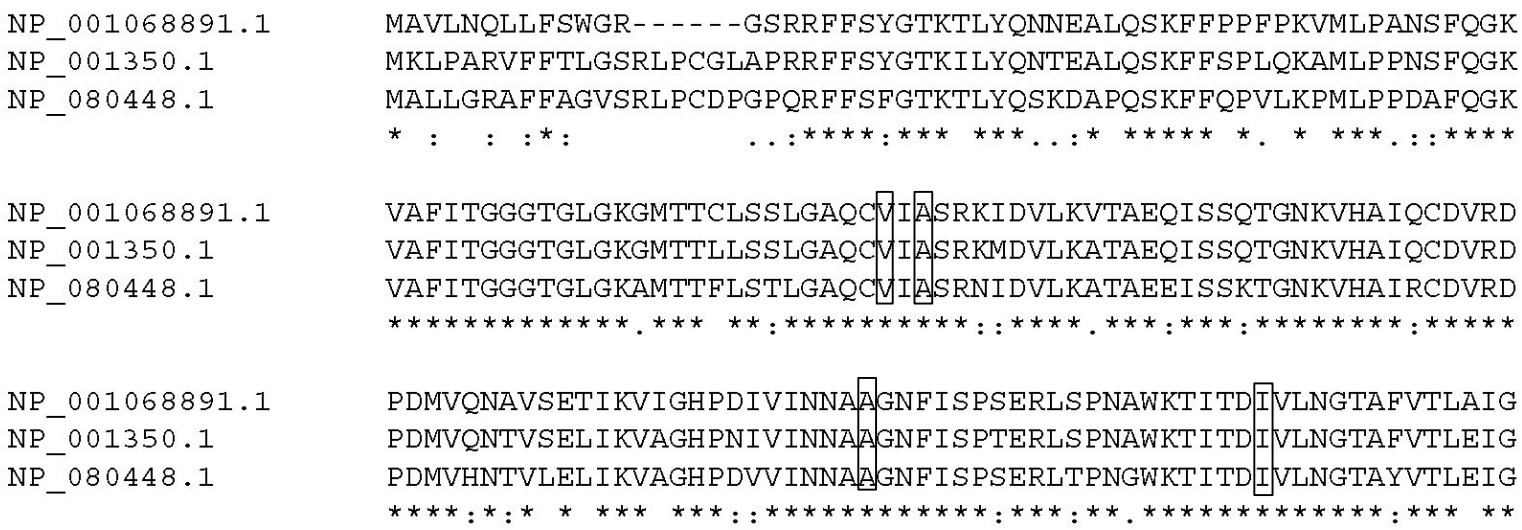

Figure 7.1. Multiple sequence alignment for *DECR1* among cattle (NP_001068891.1), humans (NP_001350.1), and mice (NP_080448.1) using ClustalW2 online tool. The dark boxes indicate polymorphism locations as described in Table 7.1. Symbols: '*' indicates that the residues in the column are identical across all sequences, ':' indicates that conserved substitutions were identified, and '.' indicates that semi-conserved substitutions were identified according to the ClustalW2 online tool.

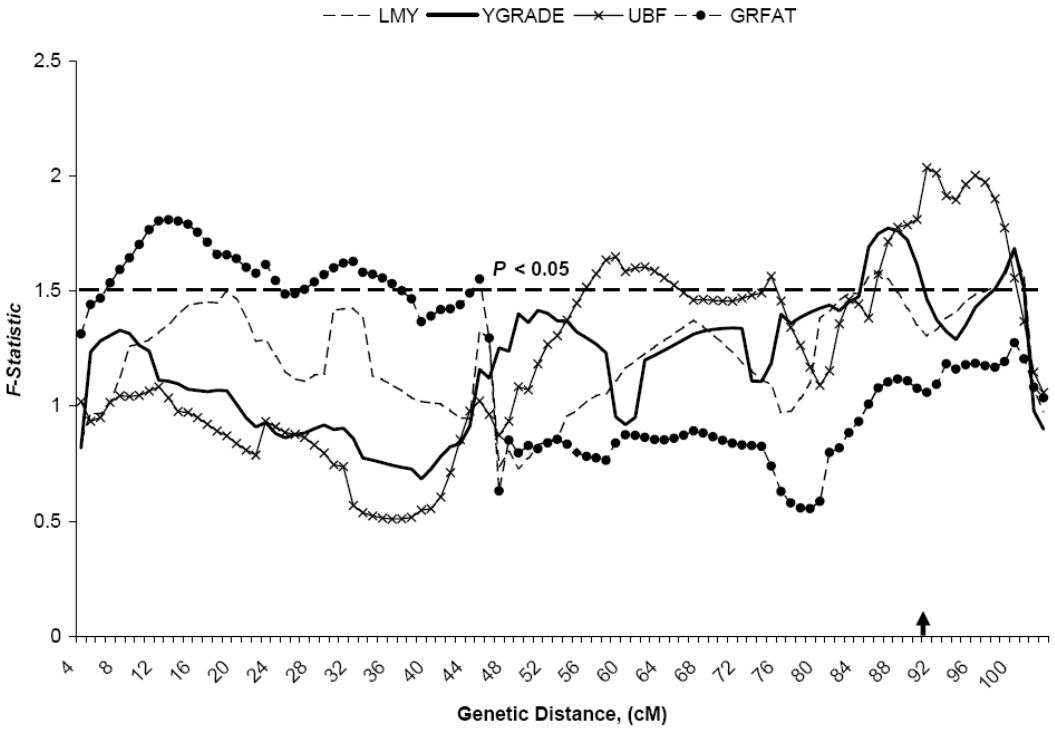

Figure 7.2: Across family F-statistic profiles for lean meat yield (LMY), yield grade (YGRADE), ultrasound backfat (UBF), and grade fat (GRFAT) using 123 markers in addition to ss95214671:C>T *DECR1* SNP on bovine chromosome 14. The horizontal line represents the chromosome wise threshold from 25,000 permutations. Relative *DECR1* SNP position is indicated by an arrow on the horizontal axis.

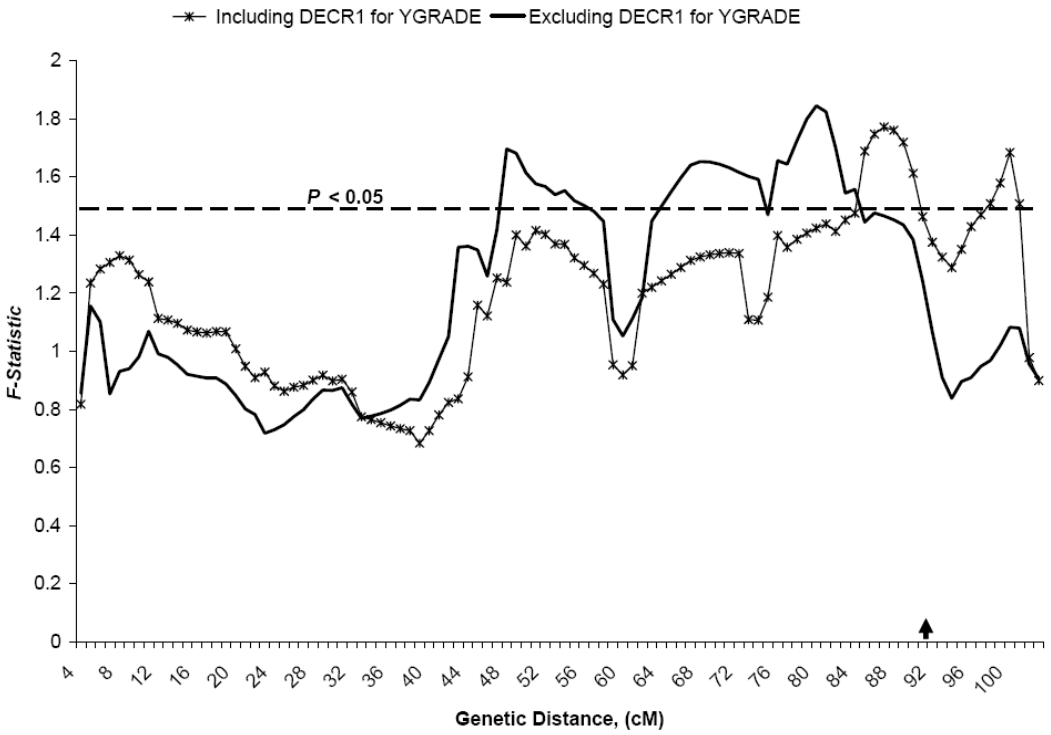

Figure 7.3: Across family F-statistic profiles using 123 markers in addition to ss95214671:C>T *DECR1* SNP on bovine chromosome 14 when including and excluding *DECR1* from the analysis for yield grade (YGRADE). The horizontal line represents the chromosome wise threshold from 25,000 permutations. Relative *DECR1* SNP position is indicated by an arrow on the horizontal axis.

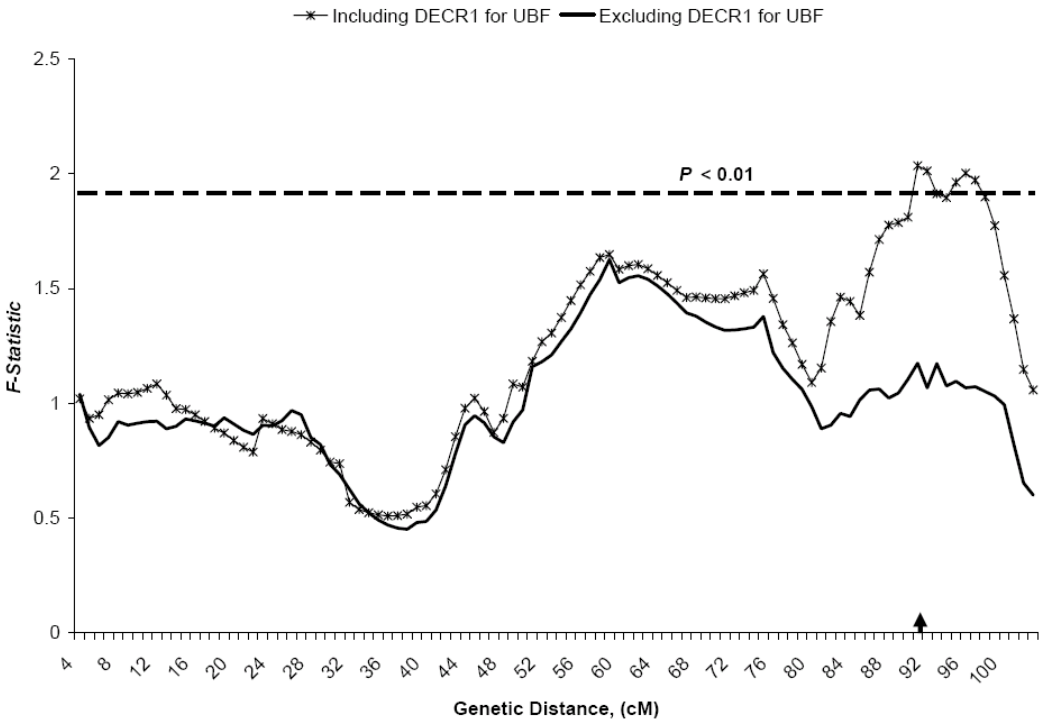

Figure 7.4. Across family F-statistic profiles using 123 markers in addition to ss95214671:C>T *DECR1* SNP on bovine chromosome 14 when including and excluding *DECR1* from the analysis for ultrasound backfat (UBF). The horizontal line represents the chromosome wise threshold from 25,000 permutations. Relative *DECR1* SNP position is indicated by an arrow on the horizontal axis.

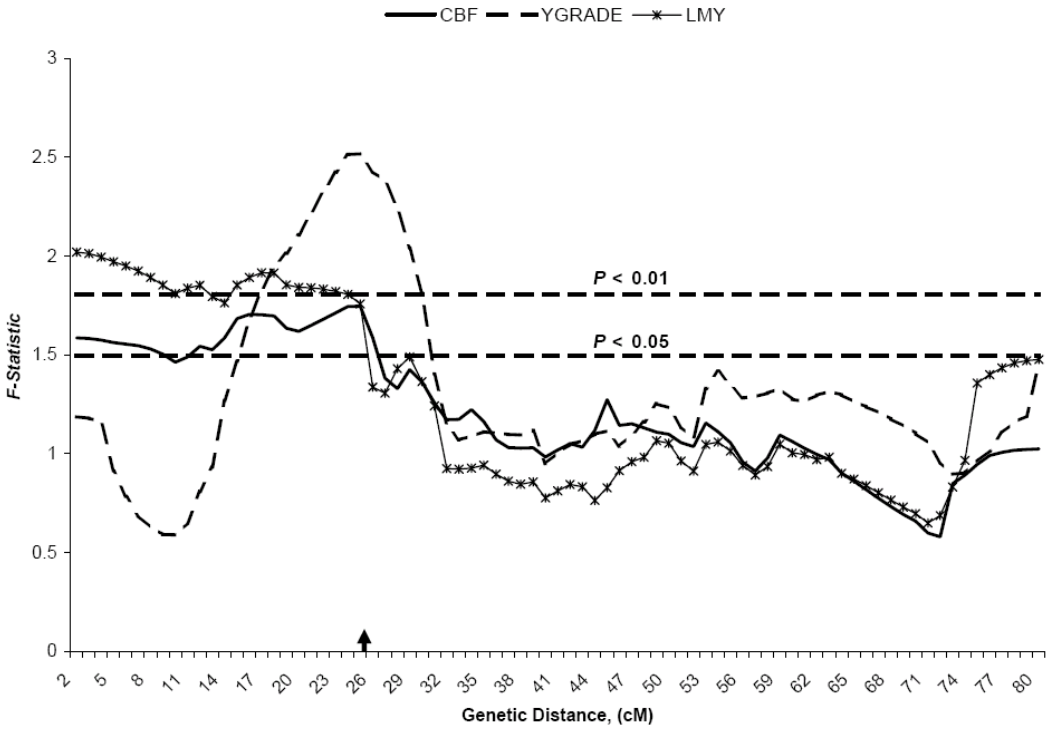

Figure 7.5: Across family F-statistic profiles for carcass backfat (CBF), yield grade (YGRADE), and lean meat yield (LMY) using 80 markers in addition to ss95214675:A>G *FGF8* SNP on bovine chromosome 26. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. Relative *FGF8* SNP position is indicated by an arrow on the horizontal axis.

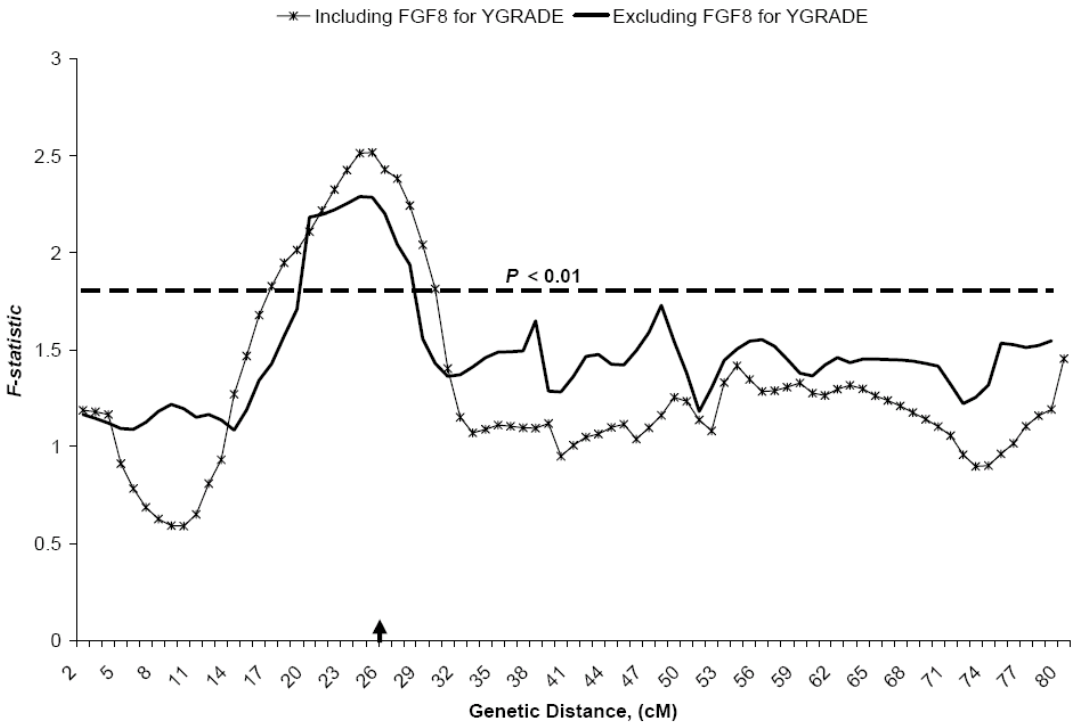

Figure 7.6: Across family F-statistic profiles using 80 markers in addition to ss95214675:A>G *FGF8* SNP on bovine chromosome 26 when including and excluding *FGF8* from the analysis for YGRADE. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. Relative *FGF8* SNP position is indicated by an arrow on the horizontal axis.

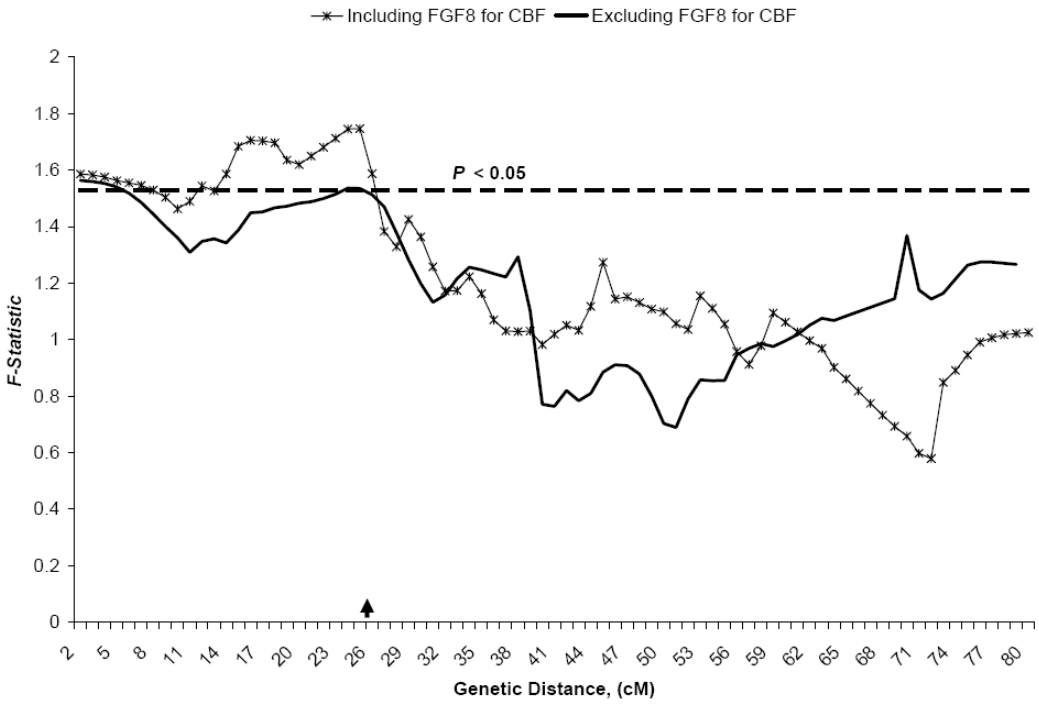

Figure 7.7: Across family F-statistic profiles using 80 markers in addition to ss95214675:A>G *FGF8* SNP on bovine chromosome 26 when including and excluding *FGF8* from the analysis for CBF. The horizontal line(s) represents the chromosome wise threshold from 25,000 permutations. Relative *FGF8* SNP position is indicated by an arrow on the horizontal axis.

7.4 Literature Cited

Agriculture Canada (1992) Livestock carcass grading regulations. *Can. Gazette Part II.*, 3821–8.

Amills M., Vidal O., Varona L., Tomas A., Gil M., Sanchez A. & Noguera J.L. (2005) Polymorphism of the pig 2,4-dienoyl CoA reductase 1 gene (DECR1) and its association with carcass and meat quality traits. *J Anim Sci* **83**, 493-8.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B-Methodological*, 57:289-300.

Brethour J.R. (1992) The repeatability and accuracy of ultrasound in measuring backfat of cattle. *J Anim Sci* **70**, 1039-44.

Canadian Council on Animal Care (1993) Guide to the care and use of experimental animals. *Vol. 1. E. D. Olfert, B. M. Cross, and A. A. McWilliams, ed. Can. Counc. Anim. Care, Ottawa, Ontario, Canada.*

Casas E., Keele J.W., Shackelford S.D., Koohmaraie M. & Stone R.T. (2004) Identification of quantitative trait loci for growth and carcass composition in cattle. *Anim Genet* **35**, 2-6.

Clarke M.S., Khakee R. & McNeil P.L. (1993) Loss of cytoplasmic basic fibroblast growth factor from physiologically wounded myofibers of normal and dystrophic muscle. *J Cell Sci* **106 (Pt 1)**, 121-33.

Cuevas P. (1998) Hormone-like effects of fibroblast growth factors (FGFs) in the nervous system. *Surg Neurol* **50**, 571-3.

Davies B.R., Fernig D.G., Barraclough R. & Rudland P.S. (1996) Effect on tumorigenicity and metastasis of transfection of a diploid benign rat mammary epithelial cell line with DNA corresponding to the mRNA for basic fibroblast growth factor. *Int J Cancer* **65**, 104-11.

de Koning D.J., Janss L.L., Rattink A.P., van Oers P.A., de Vries B.J., Groenen M.A., van der Poel J.J., de Groot P.N., Brascamp E.W. & van Arendonk J.A. (1999) Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa). *Genetics* **152**, 1679-90.

Goonewardene L.A., Wang Z., Price M.A., Yang R.-C., Berg R.T. & Makarechian M. (2003) Effects of udder type and calving assistance on weaning traits of beef and dairy x beef calves. *Livestock Production Science*, 81: 47 - 56.

Hart A.W., Baeza N., Apelqvist A. & Edlund H. (2000) Attenuation of FGF signalling in mouse beta-cells leads to diabetes. *Nature* **408**, 864-8.

Huang X., Yang C., Luo Y., Jin C., Wang F. & McKeehan W.L. (2007) FGFR4 prevents hyperlipidemia and insulin resistance but underlies high-fat diet induced fatty liver. *Diabetes* **56**, 2501-10.

Jones S.D., T. D. Bergess, J. W. Wilton, and C. H. Watson (1984) Feedlot performance, carcass composition and efficiency of muscle gain in bulls and steers of different mature size slaughtered at different level of fatness. *Can. J. Anim. Sci.* **64**, 621–30.

Knott S.A., Elsen J.M. & Haley C.S. (1996) Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theoretical and Applied Genetics*, 93: 71-80.

Kunau W.H. & Dommes P. (1978) Degradation of unsaturated fatty acids. Identification of intermediates in the degradation of cis-4-decenoly-CoA by extracts of beef-liver mitochondria. *Eur J Biochem* **91**, 533-44.

MacNeil M.D. & Grosz M.D. (2002) Genome-wide scans for QTL affecting carcass traits in Hereford x composite double backcross populations. *J Anim Sci* **80**, 2316-24.

Marques E., De Givry S., Stothard P., Murdoch B., Wang Z., Womack J. & Moore S. (2007) A high resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. *BMC Genomics* **8**, 254.

Mizoshita K., Takano A., Watanabe T., Takasuga A. & Sugimoto Y. (2005) Identification of a 1.1-Mb region for a carcass weight QTL on bovine Chromosome 14. *Mamm Genome* **16**, 532-7.

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C. & Benkel B. (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. *J Anim Sci* **81**, 1919-25.

Nagamine Y. & Haley C.S. (2001) Using the mixed model for interval mapping of quantitative trait loci in outbred line crosses. *Genet Res* **77**, 199-207.

Nkrumah J.D., Basarab J.A., Price M.A., Okine E.K., Ammoura A., Guercio S., Hansen C., Li C., Benkel B., Murdoch B. & Moore S.S. (2004) Different measures of energetic efficiency and their phenotypic relationships with growth, feed intake, and ultrasound and carcass merit in hybrid cattle. *J Anim Sci* **82**, 2451-9.

Nkrumah J.D., Sherman E.L., Li C., Marques E., Crews D.H., Jr., Bartusiak R., Murdoch B., Wang Z., Basarab J.A. & Moore S.S. (2007) Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. *J Anim Sci* **85**, 3170-81.

Perez-Enciso M., Clop A., Noguera J.L., Ovilo C., Coll A., Folch J.M., Babot D., Estany J., Oliver M.A., Diaz I. & Sanchez A. (2000) A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross. *J Anim Sci* **78**, 2525-31.

Rankinen T., Zuberi A., Chagnon Y.C., Weisnagel S.J., Argyropoulos G., Walts B., Perusse L. & Bouchard C. (2006) The human obesity gene map: the 2005 update. *Obesity (Silver Spring)* **14**, 529-644.

Seaton G., Haley C.S., Knott S.A., Kearsey M. & Visscher P.M. (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339-40.

Snelling W.M., Chiu R., Schein J.E., Hobbs M., Abbey C.A., Adelson D.L., Aerts J., Bennett G.L., Bosdet I.E., Boussaha M., Brauning R., Caetano A.R., Costa M.M., Crawford A.M., Dalrymple B.P., Eggen A., Everts-van der Wind A., Floriot S., Gautier M., Gill C.A., Green R.D., Holt R., Jann O.,

Jones S.J., Kappes S.M., Keele J.W., de Jong P.J., Larkin D.M., Lewin H.A., McEwan J.C., McKay S., Marra M.A., Mathewson C.A., Matukumalli L.K., Moore S.S., Murdoch B., Nicholas F.W., Osoegawa K., Roy A., Salih H., Schibler L., Schnabel R.D., Silveri L., Skow L.C., Smith T.P., Sonstegard T.S., Taylor J.F., Tellam R., Van Tassell C.P., Williams J.L., Womack J.E., Wye N.H., Yang G. & Zhao S. (2007) A physical map of the bovine genome. *Genome Biol* **8**, R165.

Stone R.T., Keele J.W., Shackelford S.D., Kappes S.M. & Koohmaraie M. (1999) A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. *J Anim Sci* **77**, 1379-84.

Stylianou I.M., Korstanje R., Li R., Sheehan S., Paigen B. & Churchill G.A. (2006) Quantitative trait locus analysis for obesity reveals multiple networks of interacting loci. *Mamm Genome* **17**, 22-36.

Tanaka A., Miyamoto K., Minamino N., Takeda M., Sato B., Matsuo H. & Matsumoto K. (1992) Cloning and characterization of an androgen-induced growth factor essential for the androgen-dependent growth of mouse mammary carcinoma cells. *Proc Natl Acad Sci U S A* **89**, 8928-32.

Van Eenennaam A.L., Li J., Thallman R.M., Quaas R.L., Dikeman M.E., Gill C.A., Franke D.E. & Thomas M.G. (2007) Validation of commercial DNA tests for quantitative beef quality traits. *J Anim Sci* **85**, 891-900.

Weller J.I., Song J.Z., Heyen D.W., Lewin H.A. & Ron M. (1998) A new approach to the problem of multiple comparisons in the genetic dissection of complex traits. *Genetics* **150**, 1699-706.

Wolford J.K., Bogardus C. & Prochazka M. (1998) Polymorphism in the 3' untranslated region of MTG8 is associated with obesity in Pima Indian males. *Biochem Biophys Res Commun* **246**, 624-6.

Zammit C., Coope R., Gomm J.J., Shousha S., Johnston C.L. & Coombes R.C. (2002) Fibroblast growth factor 8 is expressed at higher levels in lactating human breast and in breast cancer. *Br J Cancer* **86**, 1097-103.

CHAPTER 8

General Discussions and Future Directions

8.1 General Discussions

Genetic improvement in beef and dairy cattle can bring significant advances in satisfying the global food demand, which is expected to double by 2050. Milk production and carcass quality traits are economically relevant traits with a large economic impact. There have been increasing efforts world wide targeting the genetic improvement of both meat and milk production traits in livestock species. The objective of the research described in this thesis was to identify regions on bovine chromosome 14 (BTA14) contributing to the genetic variation underlying milk and meat production traits in cattle in efforts to characterize and identify genes influencing these traits.

The first study (Chapter 3) aimed at correctly ordering genetic markers along BTA14 and comparing the order to the latest bovine sequence assembly in order to aid collaborative efforts in improving the future versions of the assembly. Accurately defining the marker order on chromosomes is crucial for correct identification of quantitative trait loci (**QTL**), haplotype map construction and refinement of candidate gene searches.

The study resulted in the construction of 12K Radiation hybrid map of bovine chromosome 14 using 843 single nucleotide polymorphism markers. The resulting map was aligned with the latest version of the bovine assembly (Btau_3.1) as well as other previously published RH maps. The resulting map identified distinct regions on bovine chromosome 14 where discrepancies

between this RH map and the bovine assembly occur. A major region of discrepancy was found near the centromere involving the arrangement and order of the scaffolds from the assembly.

The map further confirms previously published conserved synteny blocks with human chromosome 8. As well, it identifies an extra breakpoint and conserved synteny block previously undetected due to lower marker density. This conserved synteny block is in a region where markers between the RH map presented here and the latest sequence assembly (Btau_3.1) are in very good agreement. The high resolution map produced by this study and the conclusions arising from it were also in agreement with other independently released maps (Snelling *et al.* 2007).

In Chapter 4, we assessed the extent of linkage disequilibrium in both Angus and Holstein cattle. Using the information from the high resolution radiation hybrid map (Chapter 3), LD maps were constructed for both breeds. LD maps can provide a wealth of information on specific marker-phenotype relationships, especially in areas of the genome where positional candidate genes with similar functions are located. Over 500 Single Nucleotide Polymorphism (SNP) markers from both Angus and Holstein animals had their phased haplotypes estimated and their pairwise r^2 values compared. For both breeds, results showed that average LD extends at moderate levels up to 100 kilo base pairs (kbp) and falls to background levels after 500 kbp. Correlation analysis for marker pairs common to these two breeds confirmed that there are no substantial correlations between r-values at distances over 10 kbp, congruent with the

relationship between Angus and Holstein. More similar breeds will show high phase correlation at further distances. Comparison of extended haplotype homozygosity (Sabeti *et al.* 2002), which calculates the LD decay away from a core haplotype, shows that in Holstein there is long range LD decay away from the *DGAT1* region consistent with the selection for milk fat percentage in this population.

Overall, the results obtained from this study can be applied in future single or haplotype association analysis for both populations, aiding in confirming or excluding potential polymorphisms as causative mutations, especially around QTL regions. In addition, knowledge of specific LD information among markers will aid the research community in selecting appropriate markers for whole genome association studies. Evidence of the power of linkage disequilibrium in excluding or confirming causality is seen in the study of (Olsen *et al.* 2007) who by means of LD eliminated one of the *OPN* polymorphisms as the causal mutation underlying a milk QTL on BTA6. Several studies characterizing the amount of linkage disequilibrium genome wide continue to emerge (Kim & Kirkpatrick 2009; Villa-Angulo *et al.* 2009).

Work performed in earlier chapters of this thesis led to studies carried out in Chapters 5 and 6. In these chapters QTL scans affecting both milk and meat production traits were characterized leading to the identification of candidate polymorphisms under these QTL. Chapter 5 described using linkage disequilibrium information described in Chapter 4 to narrow down the number of markers needed for a QTL scan for traits affecting meat quality in beef cattle, as

well as identifying candidate markers under those QTL. It also described the analysis of an independent validation using 1000 beef animals from an independent population. The results showed that several of the selected markers showed association in this beef population, confirming the association of those markers with meat production traits.

In Chapter 6, the genetic effect of the *DGAT1* gene was used as a covariate in the analysis to identify smaller effect markers also associated with milk production traits. Using the genetic effect of *DGAT1* in the model eliminated the significant QTL peaks arising from the effect of *DGAT1* on milk production traits (Grisart *et al.* 2002; Winter *et al.* 2002). Briefly, work presented in Chapter 6 used linkage disequilibrium information from 502 single nucleotide polymorphisms to select markers for a Quantitative Trait Loci scan on bovine chromosome 14 for milk production traits in 321 Holstein animals using *DGAT1* genotype information as a covariate. The use of a genetic effect as a covariate was also used by (Olsen *et al.* 2007) to explain the effect of the *ABCG2* polymorphism on a milk QTL on BTA6. Results showed QTL peaks that otherwise would not have been identified. Overlaying information between QTL and allele effect analysis enabled the identification of 45 SNPs under those milk production trait QTL. Further testing of the SNPs using 726 additional Holstein animals enabled the identification of other marker-trait associations not previously identified. Searches for positional candidate genes under these QTL yielded promising results with one specific candidate gene encoding a voltage-gated channel which is known to be involved in the repolarization of excitable cells (Dukes &

Philipson 1996). Further analysis between SNPs discovered within this candidate gene showed that 2 polymorphisms confer a change from aspartic acid to asparagine and from glutamine to glutamic acid. Overall, combining information from marker-marker relationships, familial informativeness, marker quality and genetic knowledge of traits enabled the characterization of additional markers with significant associations with milk production traits.

The final chapter (Chapter 7) examined the associations between 2 positional candidate genes with meat production traits. Several studies have reported the presence of meat quality QTL on bovine chromosome 14 (BTA14) with no specific genes being conclusively linked as their cause. Two genes located on bovine chromosome 14 (BTA14): 2, 4 dienoyl CoA reductase 1 (*DECR1*) and core binding factor alpha domain 2 (*CBFA2T1*) have been previously evaluated in other species and found to contain polymorphisms influencing lipid metabolism. Using phenotypic information from four-hundred and sixty four Angus, Charolais, and crossbred animals associations were identified with ultrasound marbling score (*CBFA2T1*, $P = 0.019$) and ultrasound backfat (*DECR1*, $P = 0.012$), for example. Additional scrutiny in independent samples needs to be performed to validate these results. Future recommendations for follow up work include validating these markers in other cattle populations, similar to the work done at the end of Chapter 5 where an unrelated population was used to validate some of the candidate markers.

The original hypothesis was that considerable genetic variation existed in milk production and carcass quality traits in Holstein and Angus cattle,

respectively, which could be identified by quantitative trait loci and by genetic marker associations. The genomics-based tools applied in this research led to the identification of these quantitative trait loci and genetic marker associations. The studies performed and discussed in detail in Chapters 5, 6 and 7 supported the original hypothesis demonstrating opportunities for genetic improvement in Holstein and Angus cattle.

8.2 Future recommendations and directions

One direction to follow emerging from this work is the functional analysis of the genes reported in this thesis. More specifically, the assessment of how the function of the genes will be affected when the specific mutation is introduced. The use of transgenic mice for the task has employed for *DGAT1* (Grisart *et al.* 2004) and *ABCG2* (Cohen-Zinder *et al.* 2005) demonstrating its uselfulness in determining causality for a specific trait. For example, in the case of *KCNB2*, the potassium voltage-gated channel gene, one of the polymorphisms can be introduced in transgenic mice and the function of the gene can be assessed. If potassium ions cannot cross the membrane to aid in the change of action potential, insulin secretion can be compromised. Monitoring glucose levels outside cells that present this mutated gene can shed light if this gene, in fact, affects insulin secretion or not.

Another recommendation is the analysis of whole genome high density SNP markers. This process would encompass using markers along all of the chromosomes and it calculates the association of all markers together in the

analysis, as it has been performed in humans (Hirschhorn & Daly 2005). Previous whole genome LD work (McKay *et al.* 2007) estimated that based on the extent of LD genome wide, it will be necessary to have 30,000 SNPs to accurately carry out a whole genome association analysis. Whole genome association studies have already been effective in mapping monogenic traits in domestic animals, such as the identification of the gene causing congenital muscular dystonia (CMD) in Belgian Blue Cattle (Charlier *et al.* 2008). Another approach, genomic selection, is beginning to make its mark in livestock, more specifically in dairy cattle (Meuwissen *et al.* 2001; VanRaden 2008). This approach uses closely linked markers which will capture all the quantitative trait loci effects (small and large) across the genome. Simulation analysis showed promise, however, work still needs to be carried out when considering that the accuracy between different studies can vary from 20 to 67% (Hayes *et al.* 2009).

Comparison of the quantitative trait loci scan methods performed in Chapters 5 and 6 with Bayesian analysis incorporated in LOKI (Heath 1997) and a combination of linkage disequilibrium and linkage analysis (Meuwissen *et al.* 2002) can further aid in narrowing the markers association with the traits of interest. More specifically, if the QTL peaks across multiple methods overlap, this can be used as confirmation of their existence and provide more confidence in their true location.

Another exciting avenue for this work lies in combining genetics with gene expression analysis, also called genetical genomics. In this case, gene expressions are measured at specific tissues relevant to the trait of interest. The

levels of transcripts for these genes are then referred to as expression QTL (eQTL). Overlapping the analysis of eQTL with the QTL affecting the phenotype of interest can point out towards specific candidate genes. This approach is still in its infancy in livestock (Haley & de Koning 2006), but it can be very helpful specially if there is a strong correlation between the transcript levels and the phenotypes both across and within the genotypes (Schadt *et al.* 2005). The polymorphisms presented here cannot directly affect the transcript levels, since their location lies outside the region of the promoter. However, they can affect the existence of exonic splicing enhancers (Chew *et al.* 1999). These enhancers are DNA sequence motifs within an exon involved in accurate splicing of pre-RNA in mRNA.

8.3 Literature Cited

Charlier C., Coppieters W., Rollin F., Desmecht D., Agerholm J.S., Cambisano N., Carta E., Dardano S., Dive M., Fasquelle C., Frennet J.C., Hanset R., Hubin X., Jorgensen C., Karim L., Kent M., Harvey K., Pearce B.R., Simon P., Tama N., Nie H., Vandeputte S., Lien S., Longeri M., Fredholm M., Harvey R.J. & Georges M. (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat Genet* **40**, 449-54.

Chew S.L., Liu H.X., Mayeda A. & Krainer A.R. (1999) Evidence for the function of an exonic splicing enhancer after the first catalytic step of pre-mRNA splicing. *Proc Natl Acad Sci U S A* **96**, 10655-60.

Cohen-Zinder M., Seroussi E., Larkin D.M., Loor J.J., Everts-van der Wind A., Lee J.H., Drackley J.K., Band M.R., Hernandez A.G., Shani M., Lewin H.A., Weller J.I. & Ron M. (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 Res* **15**, 936-44.

Dukes I.D. & Philipson L.H. (1996) K+ channels: generating excitement in pancreatic beta-cells. *Diabetes* **45**, 845-53.

Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M. & Snell R. (2002) Positional candidate cloning of a QTL in dairy cattle: identification of a

missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res* **12**, 222-31.

Grisart B., Farnir F., Karim L., Cambisano N., Kim J.J., Kvasz A., Mni M., Simon P., Frere J.M., Coppieters W. & Georges M. (2004) Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc Natl Acad Sci U S A* **101**, 2398-403.

Haley C. & de Koning D.J. (2006) Genetical genomics in livestock: potentials and pitfalls. *Anim Genet* **37 Suppl 1**, 10-2.

Hayes B.J., Bowman P.J., Chamberlain A.J. & Goddard M.E. (2009) Invited review: Genomic selection in dairy cattle: progress and challenges. *J Dairy Sci* **92**, 433-43.

Heath S.C. (1997) Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am J Hum Genet* **61**, 748-60.

Hirschhorn J.N. & Daly M.J. (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* **6**, 95-108.

Kim E.S. & Kirkpatrick B.W. (2009) Linkage disequilibrium in the North American Holstein population. *Anim Genet*.

McKay S.D., Schnabel R.D., Murdoch B.M., Matukumalli L.K., Aerts J., Coppieters W., Crews D., Dias Neto E., Gill C.A., Gao C., Mannen H., Stothard P., Wang Z., Van Tassell C.P., Williams J.L., Taylor J.F. & Moore S.S. (2007) Whole genome linkage disequilibrium maps in cattle. *BMC Genet* **8**, 74.

Meuwissen T.H., Hayes B.J. & Goddard M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819-29.

Meuwissen T.H., Karlsen A., Lien S., Olsaker I. & Goddard M.E. (2002) Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics* **161**, 373-9.

Olsen H.G., Nilsen H., Hayes B., Berg P.R., Svendsen M., Lien S. & Meuwissen T. (2007) Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. *BMC Genet* **8**, 32.

Sabeti P.C., Reich D.E., Higgins J.M., Levine H.Z., Richter D.J., Schaffner S.F., Gabriel S.B., Platko J.V., Patterson N.J., McDonald G.J., Ackerman H.C., Campbell S.J., Altshuler D., Cooper R., Kwiatkowski D., Ward R. & Lander E.S. (2002) Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832-7.

Schadt E.E., Lamb J., Yang X., Zhu J., Edwards S., Guhathakurta D., Sieberts S.K., Monks S., Reitman M., Zhang C., Lum P.Y., Leonardson A., Thieringer R., Metzger J.M., Yang L., Castle J., Zhu H., Kash S.F., Drake T.A., Sachs A. & Lusis A.J. (2005) An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* **37**, 710-7.

Snelling W.M., Chiu R., Schein J.E., Hobbs M., Abbey C.A., Adelson D.L., Aerts J., Bennett G.L., Bosdet I.E., Boussaha M., Brauning R., Caetano A.R., Costa M.M., Crawford A.M., Dalrymple B.P., Eggen A., Everts-van der Wind A., Floriot S., Gautier M., Gill C.A., Green R.D., Holt R., Jann O.,

Jones S.J., Kappes S.M., Keele J.W., de Jong P.J., Larkin D.M., Lewin H.A., McEwan J.C., McKay S., Marra M.A., Mathewson C.A., Matukumalli L.K., Moore S.S., Murdoch B., Nicholas F.W., Osoegawa K., Roy A., Salih H., Schibler L., Schnabel R.D., Silveri L., Skow L.C., Smith T.P., Sonstegard T.S., Taylor J.F., Tellam R., Van Tassell C.P., Williams J.L., Womack J.E., Wye N.H., Yang G. & Zhao S. (2007) A physical map of the bovine genome. *Genome Biol* **8**, R165.

VanRaden P.M. (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* **91**, 4414-23.

Villa-Angulo R., Matukumalli L.K., Gill C.A., Choi J., Van Tassell C.P. & Grefenstette J.J. (2009) High-resolution haplotype block structure in the cattle genome. *BMC Genet* **10**, 19.

Winter A., Kramer W., Werner F.A., Kollers S., Kata S., Durstewitz G., Buitkamp J., Womack J.E., Thaller G. & Fries R. (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. *Proc Natl Acad Sci U S A* **99**, 9300-5.

APPENDICES

APPENDIX ONE

List of markers mapped using the 12K radiation hybrid map as described in Chapter 3.

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61534971	1	BTA-34956	0	31.7	14	101723	8	145991701
ss61508252	2	BTA-35050	3.9	26.7	14	253789	8	145658078
ss69374920	2	NW-405528-A260G	3.9	------	14	195925	8	145716378
ss69374921	2	NW-405528-A556G	3.9	------	14	196221	8	145716082
ss61534805	3	BTA-34687	16.4	36.3	14	687707	8	145155228
ss61508249	4	BTA-34971	18.6	34.9	14	851138	8	144954444
ss61535522	5	BTA-35941	20.7	28.6	14	936325	8	144450095
ss61480708	6	BTA-35408	33.8	36.4	14	3259417	8	144385978
ss61508244	7	BTA-34772	35.9	35.1	14	3303832	8	144299699
ss61534850	8	BTA-34752	39.5	37.5	14	3347301	8	144179804
ss61480590	9	BTA-34737	41.3	17.7	14	3369381		
ss61534608	10	BTA-34290	76.2	26.9	14 & Unknown	5490676	8	143023198
ss61567744	11	BTA-95738	94	36.3	14	4166573	8	142412953
ss69374922	12	CC466310-A189G	100.8	39.3	14	3989165	8	142155041
CC466310-T108G	12	CC466310-T108G	100.8	------	14	3989086	8	142154960
ss69374923	12	CC547030-C180T	100.8	------	14	3926578	8	142016216
ss69374924	12	CC547030-C391G	100.8	------	14	3926367	8	142016005
ss69374926	13	CC765708-T176G	102.5	------	14	3923126	8	142012764
ss69374925	13	CC765708-C266T	102.5	------	14	3923039	8	142012677
ss61524968	13	BTA-15991	102.5	36.8	14	3874487	8	141941418
ss61524969	13	BTA-15992	102.5	------	14	3874407	8	141941338

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss69374927	14	CC550917-A78G	105.8	30.7	14	3824944	8	141848305
ss61467365	15	BTA-34867	115.9	37.7	14	5087204	8	141344246
ss69374928	16	BZ920772-C257T	117.8	39.3	14	5037465	8	141234980
ss69374931	17	CZ413501-A295G	121.1	------	14	5013006	8	141203287
ss69374930	17	CZ413501-A235G	121.1	------	14	5012946	8	141203227
ss69374929	17	CG987930-C157A	121.1	41.4	14	5001175	8	141191456
ss69374932	18	CC502260-A167G	122.8	------	14	4980890	8	141124792
ss69374933	18	CC578276-A191T	122.8	------	14	4974072	8	141129448
ss61534902	18	BTA-34830	122.8	------	14	4973169	8	141109141
ss69374934	18	CL609530-A134G	122.8	------	14	4954420	8	141069215
ss61467359	18	BTA-34806	122.8	------	14	4902528		
ss61534889	18	BTA-34804	122.8	------	14	4900939		
ss61534868	18	BTA-34781	122.8	22.9	14	4823865		
ss61466612	19	BTA-20155	148	------	14	4610848	8	140638520
ss61466614	19	BTA-20157	148	------	14	4610553	8	140638225
ss61466610	19	BTA-20153	148	38.6	14	4609129	8	140637590
ss61527158	20	BTA-20145	151.4	40.9	14	4597154	8	140581644
ss69374935	20	BZ853464-C187T	151.4	------	14	4435159	8	140335321
ss69374936	21	CG985206-C236G	153	34.4	14	4444333	8	140344495
ss69374937	22	CC516254-A103G	159.7	39.4	14	4217777	8	140102121
ss61476793	23	BTA-20133	161.4	------	14	5582076	8	140080417
ss61476791	23	BTA-20131	161.4	37	14	5582192	8	140080301
ss61563126	24	BTA-86950	165.7	37.2	14	5558546	8	139842651
ss61563132	24	BTA-86956	165.7	------	14	5558344	8	139842449
ss38328040	25	BTA-05988	170	38.9	14	5842011	8	139745889
ss61527143	25	BTA-20123	170	------	14	5855685	8	139732715
ss61476780	26	BTA-20115	171.8	34.5	14	5869506	8	139718894
ss69374938	27	CL608961-C113T	177.4	36.5	14	6092009	8	139442071

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss69374939	28	BZ945547-A231G	181.1	29.2	14	6310161	8	139080876
ss69374940	29	SCAFFOLD4277_4521	193	------	14	7717451	8	138929055
ss38334550	29	BTA-12498	193	------	14	7717793	8	138927599
ss38334549	29	BTA-12497	193	29.9	14	7717936	8	138927456
ss38336818	29	BTA-14766	193	------	14	7727081	8	138898699
ss69374941	30	CW896678-A194G	202.2	------	14	3091100	8	138736991
ss69374943	30	CW896678-G316T	202.2	------	14	3091222	8	138736869
ss69374942	30	CW896678-C366T	202.2	------	14	3091280	8	138736811
ss61497128	30	BTA-98667	202.2	29	14	3101373	8	138726718
ss61497130	30	BTA-98669	202.2	------	14	3101722	8	138726369
ss61569297	30	BTA-98670	202.2	------	14	3102379	8	138725712
ss61569302	30	BTA-98676	202.2	------	14	3105309	8	138721574
ss69374944	31	CC503704-C163T	211.6	31.6	14	2848720	8	138333647
ss61497126	32	BTA-98660	216.5	31.3	14	2839486	8	138313605
ss61470052	33	BTA-98681	224.7	------	14	2796992	8	138266634
ss61569304	33	BTA-98678	224.7	------	14	2776446	8	138247596
ss61569303	33	BTA-98677	224.7	32	14	2776160	8	138247310
ss63389435	34	BZ911725-A146G	232.3	28	14	2669630	8	138115218
ss61563531	35	BTA-87695	243.9	29	14	2658706	8	138083459
ss69374945	36	CW907150-C222T	250.9	28.6	14	2599532	8	138022483
rs29011059	37	SCAFFOLD160323_18366	255.7	28.3	14	2536328	8	137918538
ss61494227	38	BTA-87690	261.8	30.4	14	2486358	8	137826593
ss61494220	39	BTA-87683	264.8	30.1	14	2476976	8	137821538
ss61494221	39	BTA-87684	264.8	------	14	2476909	8	137821471
ss61494223	39	BTA-87686	264.8	------	14	2476778	8	137821340
ss61494224	39	BTA-87687	264.8	------	14	2476669	8	137821231
ss38336286	40	BTA-14234	267.8	27.3	14	2395299	8	137713652
ss63388325	41	CC514645-A298G	275.8	28.4	14	2323152	8	137605467

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss69374946	42	CC514645-T214G	284.5	32.2	14	2323067	8	137605382
ss61494247	43	BTA-87745	288.2	------	14	2256776		
ss61494249	43	BTA-87747	288.2	------	14	2283901	8	137544402
ss69374947	43	BZ887867-C132T	288.2	------	14	7620224	8	137494733
ss61494244	43	BTA-87742	288.2	28.5	14	7621720	8	134266279
ss38332001	44	BTA-09949	295.6	------	14	2013182	8	137211909
ss38331999	44	BTA-09947	295.6	27.9	14	2010260	8	137208502
ss61494233	45	BTA-87730	302.4	------	14	1980878	8	137175244
ss61494232	45	BTA-87729	302.4	------	14	1974851	8	137164493
ss61563558	45	BTA-87728	302.4	------	14	1948350	8	137149950
ss61563557	45	BTA-87727	302.4	------	14	1943739	8	137145683
ss61563555	45	BTA-87725	302.4	29.3	14	1943328	8	137145272
ss61528550	46	BTA-23070	308.1	33.2	14	1937381	8	137130373
ss61528551	46	BTA-23071	308.1	------	14	1937352	8	137130344
ss61528553	46	BTA-23073	308.1	------	14	1917871	8	137111363
ss64899266	46	BTA-23076	308.1	------	14	1912249	8	137105741
ss61517993	47	BTA-111579	310.8	32.5	14	1820510	8	137001220
ss69374950	48	CZ414665-A199G	313.1	------	14	1869137	8	137054275
ss69374949	48	CZ414665-A141G	313.1	------	14	1869080	8	137054218
ss61563543	48	BTA-87712	313.1	------	14	1727884	8	136876467
ss61563542	48	BTA-87711	313.1	30.4	14	1727824	8	136876407
ss69374948	48	CC513828-C70T	313.1	------	14	1703180	8	136847313
ss61469546	49	BTA-87706	315	26.5	14	1567624	8	136596396
ss61511263	50	BTA-87694	319	28.1	14	1454501	8	136476873
ss69374951	51	CC517185-A407G	320.9	29.6	14 & Unknown	1392617	8	136384084
ss69374952	51	CC517185-C362T	320.9	------	14 & Unknown	1392572	8	136384039
ss61501188	52	BTA-35343	322.9	25.7	14	1237605	8	136165211
ss61480657	53	BTA-35131	332.8	31.7	14	1288921	8	136272530

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss69374953	54	BZ945595-C234T	338.7	------	14	1396811	8	136388278
ss69374954	54	CC517185-A286G	338.7	------	14 & Unknown	1392495	8	136383962
ss38336995	54	BTA-14943	338.7	19.1	14	1289089	8	136272698
ss61501187	54	BTA-35342	338.7	------	14	1237671	8	136165277
ss69374955	55	CC530516-A378G	367	14	14	1319207	8	136301675
ss61466079	56	BTA-115831	413.3	23.4	14	6443009	8	135924512
ss61535184	56	BTA-35317	413.3	------	Unknown			
ss38325300	57	BTA-03248	423.2	24.9	14	6470217	8	135897804
ss69374956	58	CG986046-C190T	429.1	28.8	14	6532602	8	135821299
ss61504937	59	BTA-86074	432.9	32.9	14	6638066	8	135663086
ss69374957	60	CZ409643-A97G	434.8	22	14	6679084	8	135599575
ss61520277	61	BTA-115825	445	21.6	14	6788784	8	135411221
ss61570111	61	BTA-115836	445	------	14	6818771	8	135382326
ss61570113	61	BTA-115838	445	------	14	6819210	8	135381887
ss61570114	61	BTA-115839	445	------	14	6819323	8	135381774
ss61506800	61	BTA-115835	445	------	14	6819545	8	135381552
ss61570109	61	BTA-115833	445	------	14	6823953	8	135377644
ss61567464	62	BTA-95145	456.8	18.5	14	6936337	8	135248414
ss69374958	63	CL606633-C103G	474.9	20.3	14	7260446	8	134719164
ss69374959	64	NW-620749-A228G	486.5	25.9	14	7382573	8	134554853
ss69374960	64	NW-620749-A246G	486.5	------	14	7382591	8	134554835
ss69374961	65	CC525895-C217T	490.8	22.8	14	7417176	8	134511367
ss69374962	65	CC525895-T255G	490.8	------	14	7417214	8	134511329
ss69374963	66	CC525895-C336T	504.7	21.9	14	7417295	8	134511248
ss69374964	67	CZ019069-A153G	521.1	------	14	7441330	8	134487712
ss61535114	67	BTA-35190	521.1	28.7	14	7445652	8	134473613
ss61535115	67	BTA-35191	521.1	------	14	7445857	8	134473408
ss61535128	67	BTA-35206	521.1	------	14	7457456	8	134461009

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61535134	67	BTA-35217	521.1	------	14	7506991	8	134402078
ss61535116	68	BTA-35192	523.3	18.8	14	7446030	8	134473235
rs29026768	69	SCAFFOLD240007_5847	543.1	28.2	14	7838795		
ss61535159	70	BTA-35264	555.4	------	14	7597988	8	134310589
ss61535160	70	BTA-35265	555.4	------	14	7597575	8	134297881
ss38330264	70	BTA-08212	555.4	26.8	14	7926477	8	134174403
ss61480731	71	BTA-35486	562.3	30.3	14	8076374	8	133994596
ss61535272	72	BTA-35474	564.6	------	14	8103960	8	133956798
ss61535265	72	BTA-35467	564.6	27.2	14	8121778	8	133942231
ss61535249	73	BTA-35436	569.2	28.8	14	8220250	8	133845444
ss61535250	73	BTA-35437	569.2	------	14	8220366	8	133845328
ss61535243	74	BTA-35424	571.6	31.2	14	8256379	8	133788980
ss61501204	75	BTA-35405	573.9	19.4	14	8315161	8	133689855
ss38336870	76	BTA-14818	595.1	20.7	14	8340763	8	133654742
ss38336872	77	BTA-14820	613.9	------	14	8340641	8	133654864
ss38336869	77	BTA-14817	613.9	25.7	14	8340882	8	133654623
ss61467384	78	BTA-35349	625.9	28	14	8514337	8	133448508
ss38335199	79	BTA-13147	635.6	30.9	14	8630611	8	133295327
ss61505440	79	BTA-93854	635.6	------	14	8632196	8	133293742
ss61480781	79	BTA-35689	635.6	------	14	8663005	8	133240774
ss61535391	79	BTA-35707	635.6	------	14	8694800	8	133204669
btcn12326	80	KCNQ3-A2361G	637.6	16.2	14	8692828	8	133206641
ss61535485	81	BTA-35867	669.8	37.8	14	9321855	8	132288864
ss61535493	82	BTA-35877	671.5	38.6	14	9361779	8	132242327
ss61535481	83	BTA-35863	676.7	------	14	9289523	8	132338991
ss61535491	83	BTA-35873	676.7	------	14	9325964	8	132303052
ss61535489	83	BTA-35871	676.7	------	14	9322247	8	132288472
ss61535490	83	BTA-35872	676.7	------	14	9322332	8	132288387

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss38332498	83	BTA-10446	676.7	31.5	14	9420387	8	132132423
ss38332499	83	BTA-10447	676.7	------	14	9420286	8	132132524
ss38332497	84	BTA-10445	682.4	30.4	14	9420473	8	132132337
ss61467412	85	BTA-35881	691	35	14	9470689	8	132080400
ss61508294	85	BTA-35883	691	------	14	9477500	8	132052425
ss61508297	86	BTA-35886	696.8	33.5	14	9479798	8	132065748
rs29013586	87	SCAFFOLD105570_18245	704.7	35.4	14	9481685	8	132054536
ss69374969	88	NW-206732-C386G	710.8	------	14	9600943	8	131932292
ss69374968	88	NW-206732-A255G	710.8	------	14	9601074	8	131932161
ss69374965	88	BZ887322-A123T	710.8	27.2	14	9611196	8	131918000
ss69374966	88	BZ887322-A206T	710.8	------	14	9611253	8	131917943
ss69374967	88	BZ887322-C256T	710.8	------	14	9611303	8	131917893
ss69374970	89	BZ879040-A200G	725.9	33.2	14 & Unknown	9772192	8	131708277
ss61480835	90	BTA-35893	730.1	31.3	14	9791996	8	131681322
ss61480836	90	BTA-35894	730.1	------	14	9796366	8	131680669
ss61477434	91	BTA-22322	737.5	------	14	9893779	8	131588195
ss38326853	91	BTA-04801	737.5	34.4	14	9899190	8	131583985
ss61517788	91	BTA-111174	737.5	------	14	9979791	8	131489751
ss61480838	92	BTA-35897	739.3	32.9	14	10037457	8	131433012
ss61535498	93	BTA-35898	742.9	23	14	10117482	8	131356240
ss61535499	93	BTA-35899	742.9	------	14	10117618	8	131356104
ss61535502	93	BTA-35902	742.9	------	14	10120837	8	131352076
ss61535503	93	BTA-35903	742.9	------	14	10128127	8	131345407
ss61535505	93	BTA-35905	742.9	------	14	10128281	8	131345253
ss61535506	93	BTA-35906	742.9	------	14	10128343	8	131345191
ss61535507	93	BTA-35907	742.9	------	14	10128663	8	131344871
ss61535508	93	BTA-35908	742.9	------	14	10128851	8	131344683
ss61535509	93	BTA-35909	742.9	------	14	10128854	8	131344680

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61535513	94	BTA-35913	766.7	23.9	14	10131611	8	131343076
ss61535511	95	BTA-35911	789.6	35.9	14	10131863	8	131342824
ss69374971	96	BZ939450-C53G	791.4	32.2	14	10169483	8	131312175
rs29021171	97	SCAFFOLD230838_1182	799.1	------	14	10334402	8	131083681
ss38328865	97	BTA-06813	799.1	32.6	14	10334823	8	131079277
ss69374972	97	BZ867731-C318T	799.1	------	14	10369641	8	131037035
ss61524851	97	BTA-15809	799.1	------	14	10378604	8	131033749
ss69374973	97	CC466235-T77G	799.1	------	14	10430708	8	130957627
ss61475686	98	BTA-15812	809.5	30.3	14	10339505	8	131071971
ss69374974	99	BZ867731-A374G	822.2	33.6	14	10369696	8	131036980
ss61524852	100	BTA-15810	829.2	31.1	14	10375017	8	131031659
rs29012817	101	SCAFFOLD135027_2960	842.2	37.6	14	10534630	8	130779978
rs29012823	102	SCAFFOLD135027_3247	844.7	34.2	14	10534917	8	130779691
ss38323917	103	BTA-01865	850.4	34.6	14	10563839	8	130754152
ss61499653	103	BTA-15807	850.4	------	14	10576268	8	130741723
ss69374975	104	CC767106-A322G	856	------	14	10632616	8	130681753
ss69374976	104	CC767106-C246T	856	------	14	10632692	8	130681677
ss61524855	104	BTA-15815	856	30.7	14	10649370	8	130655302
ss69374977	105	BZ946479-G56T	867.4	23.1	14	10851852	8	130448882
ss69374978	106	CL605960-C177T	893.2	19.9	14	12779054	8	130391058
ss69374979	107	CL605960-C179T	924.1	30.8	14	12779056	8	130391056
ss61519385	108	BTA-114243	931.5	------	14	12655306	8	130261065
ss61519384	108	BTA-114242	931.5	34.5	14	12655112	8	130260871
ss69374983	109	CC525898-A134G	933.3	------	14	12643203	8	130250049
ss69374984	109	CC525898-C322T	933.3	------	14	12643015	8	130249861
ss69374982	109	BTA-114234	933.3	------	14	12612900	8	130220923
ss69374981	109	BTA-114232	933.3	------	14	12612871	8	130220894
ss69374980	109	BTA-114230	933.3	34.3	14	12604935	8	130217173

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61473728	110	BTA-114220	938.9	28.9	14	12541300	8	130171973
ss61473730	110	BTA-114222	938.9	------	14 & 8	12541710		
ss38330392	111	BTA-08340	950.2	------	14	11956265	8	129533212
ss61480864	111	BTA-35970	950.2	------	14	11956060	8	129533007
ss38330389	111	BTA-08337	950.2	32.8	14	11950847	8	129527794
ss38330390	111	BTA-08338	950.2	------	14	11950745	8	129527692
ss61480861	111	BTA-35966	950.2	------	14	11947037	8	129524484
ss61480858	111	BTA-35963	950.2	------	14	11940258	8	129514529
ss38333168	111	BTA-11116	950.2	------	14	11936750	8	129514197
ss61535537	112	BTA-35960	956.4	------	14	11841230	8	129416365
ss61535536	112	BTA-35959	956.4	------	14	11837307	8	129412942
ss61535535	112	BTA-35958	956.4	------	14	11837253	8	129412888
ss61535533	112	BTA-35956	956.4	------	14	11836851	8	129412486
ss38331774	112	BTA-09722	956.4	26.8	14	11826388	8	129402993
ss61535532	112	BTA-35955	956.4	------	14	11809664	8	129384076
ss61535531	112	BTA-35954	956.4	------	14	11806276	8	129383382
ss61535525	113	BTA-35945	969.7	33.5	14	11662570	8	129216976
ss61480853	114	BTA-35950	971.7	------	14	11693484	8	129251704
ss61535528	114	BTA-35948	971.7	------	14	11666770	8	129221536
ss61535524	114	BTA-35944	971.7	24.1	14	11662383	8	129216789
AAFC02053718	115	BES9_Contig292_918	983.9	18.8	14	11525352	8	129070366
ss61480851	116	BTA-35937	1000.8	22.1	14	11219420	8	128688915
ss38333453	117	BTA-11401	1016.8	34	14	11075705	8	128522575
ss61508302	117	BTA-35938	1016.8	------	14	11100514	8	128546888
ss61480848	118	BTA-35927	1020.5	35.2	14	11011960	8	128440672
ss61522828	119	BTA-120525	1022.3	26.6	14	10938710	8	128347660
ss61562418	120	BTA-85664	1038	12.1	14	13473237	8	128293464
ss61562421	120	BTA-85667	1038	------	14	13476259	8	128282160

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61562420	120	BTA-85666	1038	------	14	13476617	8	128292249
ss38328882	121	BTA-06830	1085.9	25.4	14	17140783	8	127522584
ss38332745	122	BTA-10693	1093.9	------	14	17042763	8	127649316
ss38328883	122	BTA-06831	1093.9	30.4	14	17140915	8	127522452
ss61493685	123	BTA-85436	1095.9	------	14	17281272	8	127347083
ss61493684	123	BTA-85435	1095.9	26.2	14	17281426	8	127346929
ss61493686	124	BTA-85437	1102	19.9	14	17367043	8	127246029
ss61516354	125	BTA-108467	1117.9	20.7	14	17794725		
ss61472371	126	BTA-108362	1129.3	18.2	14	17933719	8	126495650
ss61472377	127	BTA-108368	1145.3	24.1	14	18232101	8	126159418
ss61472375	128	BTA-108366	1149.7	------	14	18232544	8	126158975
ss61472374	128	BTA-108365	1149.7	19	14	18232631	8	126158888
btcn47613	128	SQLE-C949G	1149.7	------	14	18278134	8	126107568
ss61506359	129	BTA-108461	1167.5	31.2	14	18555661	8	125781458
ss61516355	130	BTA-108468	1170.5	30.7	14	18613270	8	125709584
btcn20869	131	NDUFB9-G249T	1173.5	26	14	18689115	8	125627645
ss61508639	132	BTA-42136	1181.9	29.6	14	13680761	8	125540099
ss61538773	133	BTA-42142	1185.9	27.6	14	13811675	8	125361286
ss61482478	133	BTA-42148	1185.9	------	14	13882045	8	125271739
ss61538779	133	BTA-42153	1185.9	------	14	13936674	8	125184927
ss61538787	133	BTA-42161	1185.9	------	14	13939055	8	125201961
ss61508642	134	BTA-42145	1192.1	27.6	14	13844451	8	125320825
ss61538794	135	BTA-42168	1198.9	------	14	13963653	8	125160772
ss61538793	135	BTA-42167	1198.9	26.9	14	13963921	8	125160504
ss61482479	136	BTA-42171	1204.4	22.7	14	14093965	8	125015932
ss61538798	137	BTA-42173	1218.7	------	14	14231234	8	124802057
ss61538803	137	BTA-42178	1218.7	------	14	14235337	8	124798452
ss61508645	137	BTA-42181	1218.7	------	14	14236828	8	124796961

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss38329715	137	BTA-07663	1218.7	20.5	14	14249525	8	124782514
ss38329718	137	BTA-07666	1218.7	------	14	14249657	8	124782382
ss38329720	137	BTA-07668	1218.7	------	14	14250061	8	124781978
rs29027332	138	SCAFFOLD250075_12375	1237.1	16.7	14	14250142	8	124782149
ss61535549	139	BTA-35987	1265.4	------	14	14338668	8	124695158
ss61535548	139	BTA-35986	1265.4	26.7	14	14339086	8	124694740
ss61535551	139	BTA-35989	1265.4	------	14	14352915	8	124674414
ss61501250	140	BTA-35993	1275.3	------	14	14700498	8	124319675
ss61501249	140	BTA-35992	1275.3	33.4	14	14700698	8	124319475
ss61535553	140	BTA-35994	1275.3	------	14	14715136	8	124296242
ss61535555	140	BTA-35996	1275.3	------	14	14769294		
ss61535556	141	BTA-35997	1277.3	22.4	14	14837064	8	124137971
ss61535570	142	BTA-36018	1295.8	28.6	14	15450747	8	123466027
ss61535572	143	BTA-36020	1302.4	29.8	14	15455701	8	123471749
rs29013315	144	SCAFFOLD100871_5973	1307.3	22.5	14	15506284	8	123413692
rs29012803	145	SCAFFOLD134924_5249	1324.6	22.7	14	15867530	8	123095485
ss61527019	146	BTA-19924	1339.2	------	14	16030001	8	122972247
ss61527018	146	BTA-19923	1339.2	------	14	16030126	8	122955923
ss61527016	146	BTA-19921	1339.2	21.1	14	16030242	8	122955823
ss61527020	147	BTA-19925	1349.4	22.4	14	16074327	8	122906446
ss61475000	148	BTA-119505	1355.6	29	14	16429657	8	122577951
ss61535591	149	BTA-36054	1358.4	26.1	14	16668155	8	122352021
ss61525640	150	BTA-17316	1367.1	------	14	20599868	8	122130057
ss61525638	150	BTA-17314	1367.1	20.7	14	20599967	8	122130031
ss61480878	151	BTA-36064	1379.6	22.8	14	21391995	8	48970973
ss61473395	152	BTA-112886	1387.1	27.1	14	21523304	8	49259471
ss61562834	153	BTA-86408	1395.2	22.2	14	21646020	8	49655715
ss61562833	154	BTA-86406	1407.5	19.6	14	21907297		

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss38328592	155	BTA-06540	1425	21.1	14	22133376	8	50913559
ss61562835	156	BTA-86411	1440.6	30	14	22258733	8	51088247
ss61477079	157	BTA-21242	1444.8	------	14	18869030	8	51365252
ss61477078	157	BTA-21240	1444.8	29.8	14	18869062	8	51365284
ss61531696	157	BTA-28726	1444.8	------	14	18897571	8	51459127
ss61534577	157	BTA-34246	1444.8	------	14	18927616	8	51488672
rs29010516	158	SCAFFOLD155270_13794	1447.4	26.9	14	19092774	8	51953548
ss61506535	159	BTA-111851	1452.8	28	14	19658489	8	53264865
rs29010386	159	SCAFFOLD15358_17120	1452.8	------	14	19661447	8	53260740
ss61534587	159	BTA-34265	1452.8	------	14	19676363	8	53283737
ss61534586	160	BTA-34263	1461.3	27.1	14	19774380	8	53381664
ss38327919	161	BTA-05867	1472.7	27.1	14	19844421	8	53499273
ss61534584	162	BTA-34259	1486.7	------	14	19826607	8	53457609
ss61534582	162	BTA-34257	1486.7	------	14	19826780	8	53457782
ss38327921	162	BTA-05869	1486.7	26.7	14	19844748	8	53499273
ss61534580	163	BTA-34252	1494.4	26.4	14	20194230	8	54241443
ss61534632	163	BTA-34365	1494.4	------	14	20402564	8	54927279
rs29009981	164	SCAFFOLD145694_4460	1498.6	22.9	14	20207371	8	54223331
ss61480444	165	BTA-34250	1508.6	23.2	14	20252345	8	54427081
ss61570309	166	BTA-34296	1519.8	29.7	14	20507528	8	55106963
ss61565351	167	BTA-91255	1524	------	14	23809216	8	55753105
ss61565348	167	BTA-91252	1524	------	14	23809006	8	55753315
ss38336702	167	BTA-14650	1524	31.7	14	23793721	8	55768600
ss61512747	168	BTA-101386	1528	------	14	23500283	8	56245067
ss38328378	168	BTA-06326	1528	23.3	14	23498353	8	56246997
ss61534605	169	BTA-34285	1544.4	20.7	14	22741955	8	57380404
ss62007800	170	BTA-34283	1558.9	28.9	14	22565488	8	57670193
ss61534604	171	BTA-34282	1561	18.6	14	22542534	8	57695548

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61534596	172	BTA-34274	1585.6	20	14	22342561	8	57981272
ss61534598	173	BTA-34276	1605.6	------	14	22345845	8	57978578
ss61534599	173	BTA-34277	1605.6	------	14	22345581	8	57978842
ss61534597	173	BTA-34275	1605.6	------	14	22342579	8	57981254
ss61534595	173	BTA-34273	1605.6	------	14	22342502	8	57981331
ss61534593	173	BTA-34271	1605.6	27	14	22342243	8	57981590
ss61570240	174	BTA-24548	1612.2	30.6	14	27017291	8	58756033
ss61568592	174	BTA-97370	1612.2	------	14	26788233	8	58897820
ss61520603	175	BTA-116449	1614.3	------	14	26823318	8	58710740
ss61520602	175	BTA-116448	1614.3	24.3	14	26837754	8	58684668
ss61568588	176	BTA-97366	1623	20.7	14	26758400	8	58958874
ss61568589	176	BTA-97367	1623	------	14	26758349	8	58958925
ss61568580	177	BTA-97353	1640.2	------	14	26628236	8	59118771
ss61568577	177	BTA-97350	1640.2	26.1	14	26627628	8	59119379
ss61496808	178	BTA-97344	1647.1	21.9	14	26555321	8	59203649
ss61496809	178	BTA-97345	1647.1	------	14	26555193	8	59203777
ss61568593	178	BTA-97371	1647.1	------	14	26421462	8	59406109
ss61534613	179	BTA-34310	1662	29.4	14	24203843	8	59784856
ss61480461	180	BTA-34309	1666	22.4	14	24288016	8	59876811
ss61534611	181	BTA-34304	1676.7	------	14	24461708	8	60037728
ss61534610	181	BTA-34303	1676.7	------	14	24466572	8	60043263
ss61476007	181	BTA-17025	1676.7	16.9	14	24484683	8	60064464
ss61478618	182	BTA-27288	1703.4	30.2	14	24999237	8	60794627
ss38334682	183	BTA-12630	1705.8	30.3	14 & 16	25046839	8	60874901
ss61467336	184	BTA-34289	1708.1	25.6	14	25232504	8	61155390
ss61480452	185	BTA-34291	1722.5	34.5	14	25367652	8	61341218
ss38329951	186	BTA-07899	1725.6	29.3	14	25407176	8	61403961
ss38329953	186	BTA-07901	1725.6	------	14	25407507	8	61404292

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61500656	186	BTA-28254	1725.6	------	14	25502133	8	61486967
ss61500655	186	BTA-28253	1725.6	------	14	25502307	8	61487141
ss61534642	187	BTA-34386	1733.9	18.2	14	25531009	8	61573188
ss61480486	188	BTA-34381	1765.1	17.7	14	25560113	8	61607484
ss61478665	189	BTA-27436	1799.1	25.8	14	25574206	8	61636587
ss61480485	190	BTA-34380	1811.8	32.5	14	27510984	8	61788694
ss61534638	191	BTA-34376	1814.1	------	14	27470182	8	61825371
ss61534637	191	BTA-34375	1814.1	------	14	27470152	8	61825401
ss61534635	191	BTA-34373	1814.1	28.8	14	27469936	8	61825617
ss61480482	192	BTA-34369	1821.5	33.2	14	27341440	8	61946468
ss61501125	193	BTA-34405	1823.8	32.4	14	27251226	8	62052297
ss61501127	193	BTA-34407	1823.8	------	14	27250727	8	62052796
ss61534652	194	BTA-34403	1826.8	29.4	14	27179549	8	62110652
ss61534653	194	BTA-34404	1826.8	------	14	27179374	8	62110827
ss61480496	195	BTA-34400	1835.8	------	14	27116475	8	62156554
ss61480495	195	BTA-34399	1835.8	32.7	14	27116422	8	62156607
ss61534647	196	BTA-34396	1840.6	32.4	14	27074160	8	62193851
ss61480499	197	BTA-34410	1845.4	35.7	14	26302678	8	62426659
ss61480494	198	BTA-34395	1847.2	------	Unknown		8	62349600
ss61480492	198	BTA-34393	1847.2	------	Unknown		8	62349722
ss61534657	198	BTA-34416	1847.2	------	14	26325839	8	62388677
ss61534655	198	BTA-34414	1847.2	------	14	26325803	8	62388713
ss61534654	198	BTA-34413	1847.2	------	14	26325597	8	62388919
ss61534645	198	BTA-34390	1847.2	28.1	14	26232485	8	62509890
ss61480490	199	BTA-34388	1856.2	28.2	14	26005637	8	62779268
ss61526043	200	BTA-18146	1863.5	------	14	25819551	8	63170947
ss61526042	200	BTA-18145	1863.5	------	14	25819529	8	63170969
ss61499843	200	BTA-18141	1863.5	23.5	14	25776745	8	63250785

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61476316	201	BTA-18140	1872.6	12.8	14	25756553	8	63274481
ss61534661	202	BTA-34421	1906.6	30.1	14	27651295	8	63109459
ss61534663	203	BTA-34423	1908.7	------	14	27651729	8	63109025
ss61534662	203	BTA-34422	1908.7	20.4	14	27651424	8	63109330
ss61534664	203	BTA-34424	1908.7	------	14	27660193	8	63550520
ss61480508	204	BTA-34463	1924.1	30.4	14	27808284	8	64281074
ss61508227	205	BTA-34462	1926.1	28.6	14	27810889	8	64283679
rs29011751	206	SCAFFOLD174207_1629	1932.1	23.5	14	28435571	8	64594206
ss61501128	207	BTA-34429	1945.7	23.7	14	28566213	8	64725331
ss61534670	207	BTA-34435	1945.7	------	14	28584137	8	64742757
ss61534671	207	BTA-34437	1945.7	------	14	28670421	8	64820294
ss61480506	208	BTA-34436	1957	7.5	14	28594468	8	64756208
ss61534697	209	BTA-34468	1984.1	13.7	14	28920570	8	65075057
ss61534700	209	BTA-34471	1984.1	------	14	28976145	8	65102884
ss61534704	209	BTA-34475	1984.1	------	14	28976451	8	65103190
ss61534701	210	BTA-34472	1995.7	16.8	14	28976258	8	65102997
ss61534723	211	BTA-34528	2023.7	11.1	14	29188413	8	65405087
ss38330358	212	BTA-08306	2079.2	13.5	14	29438434	8	65716960
ss61516307	213	BTA-108350	2124.6	21.8	14	29618743	8	65930184
ss61539390	213	BTA-43271	2124.6	------	14	29769177	8	66179687
ss69374985	213	BTA-111518	2124.6	------	14	29817331	8	66244253
ss61472367	214	BTA-108347	2139.5	20.1	14	29662679	8	66069918
ss61517772	215	BTA-111152	2158.8	27.9	14	29885426	8	66315669
rs29018581	216	SCAFFOLD65161_8134	2162.9	27.8	14	29957904	8	66386905
ss61501854	217	BTA-43143	2165.2	------	14	29962710	8	66397703
ss61501852	217	BTA-43140	2165.2	29.4	14	29963305	8	66398298
ss61482545	218	BTA-42489	2167.6	30.2	14	30149295	8	66686350
ss61538859	219	BTA-42290	2169.9	29.3	14	30193103	8	66752605

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61539954	220	BTA-44326	2174.5	30.7	14	30293993		66862685
ss61529618	221	BTA-24981	2176.4	26.5	14	30517798	8	67124501
ss61529619	222	BTA-24986	2182.4	19.5	14	30545925	8	67146427
ss61478079	223	BTA-24987	2207.2	19.3	14	30575511	8	67160808
ss61514147	224	BTA-104210	2236.2	27.9	14	30591455	8	67191314
ss61467346	225	BTA-34536	2244	31.4	14	30767537	8	67505640
ss61467347	225	BTA-34537	2244	------	14	30767645	8	67505748
ss61467345	226	BTA-34535	2246.1	19.1	14	30771388	8	67508551
ss61480519	227	BTA-34504	2272.1	------	14	31427042	8	68302970
ss61534713	227	BTA-34502	2272.1	30.8	14	31527477	8	68358248
ss61467342	228	BTA-34499	2276.7	------	14	31552144	8	68380967
ss61534710	228	BTA-34491	2276.7	31.6	14	31557987	8	68386810
ss61534711	229	BTA-34492	2279.1	27	14	31558411	8	68387234
ss61480517	230	BTA-34488	2286.5	------	14	31584573	8	68409197
ss61480513	230	BTA-34484	2286.5	23.7	14	31588543		68411973
ss61506756	231	BTA-114994	2297.5	15	14	31826528	8	68778054
ss61469626	232	BTA-89587	2325.6	21.2	14	32458039	8	69425824
ss61508493	233	BTA-40148	2335	21.2	14	32678774	8	69768761
ss61514555	234	BTA-105019	2344.4	21.6	14	32890428	8	70000421
ss61481947	235	BTA-40164	2351.5	22.8	14	33106788	8	70243434
ss61534731	236	BTA-34541	2356.1	------	14	33192273	8	70347011
ss61534730	236	BTA-34540	2356.1	------	14	33205330	8	70359916
ss61480530	236	BTA-34538	2356.1	18.6	14	33231453	8	70394080
ss61534734	237	BTA-34544	2369.8	------	14	37827104	8	70635746
ss38333641	237	BTA-11589	2369.8	19.2	14	37781488	8	70672907
ss61534739	238	BTA-34555	2383.4	22.9	14	37335273	8	71117121
ss61534790	239	BTA-34658	2385.6	------	14	38446417	8	71524454
ss61534788	239	BTA-34656	2385.6	22.5	14	38446496	8	71524533

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61470582	240	BTA-101611	2393.5	25.2	14	37967708	8	72014569
rs29022424	241	SCAFFOLD270113_28063	2398.8	25.3	14	37988553	8	72031470
ss61515962	242	BTA-107716	2404	22.3	14	38167343	8	72218118
ss61498127	243	BTA-107719	2412.8	16.4	14	38297010	8	72331030
ss61515968	244	BTA-107731	2438.5	27.5	14	33614205	8	72694237
ss61498133	245	BTA-107733	2442.9	29.7	14	33625834	8	72705366
ss61515954	246	BTA-107702	2445	26.6	14	33665984	8	72744986
ss61515956	247	BTA-107705	2449.2	6.7	14	33951451	8	73053774
ss61472242	248	BTA-107891	2522	12.2	14	34529244	8	73837515
ss61472243	249	BTA-107892	2568.7	19.2	14	34529300	8	73837571
ss61516059	249	BTA-107904	2568.7	------	14	34721554	8	74005120
ss61528182	250	BTA-22294	2577.7	22.5	14	34980294	8	74307517
ss61466720	251	BTA-22306	2582.1	25.9	14	35120789	8	74335856
ss61534803	252	BTA-34685	2586.4	23	14	35215056	8	74543081
ss61534804	252	BTA-34686	2586.4	------	14	35215089	8	74543114
ss61508238	253	BTA-34677	2593.8	21.7	14 & Unknown	35465868	8	74829551
ss61508239	253	BTA-34678	2593.8	------	14 & Unknown	35466304	8	74829987
ss61508240	253	BTA-34679	2593.8	------	14 & Unknown	35466395	8	74830078
ss61508235	254	BTA-34674	2606.1	18.3	14 & Unknown	35469395	8	74829170
ss61534799	255	BTA-34668	2630	25	14	35626209	8	75055030
ss61534798	256	BTA-34667	2635.4	27.6	14	35659562	8	75084550
ss61480572	257	BTA-34666	2639.5	27	14	35738041	8	75158136
ss61480576	258	BTA-34690	2645.8	27.8	14	35908704	8	75317084
ss61480578	259	BTA-34692	2652	26.7	14	35908974	8	75317515
ss61480579	260	BTA-34693	2659.7	29.6	14	35941867	8	75348717
ss38323652	261	BTA-01600	2662.2	27.8	14	35942345	8	75349195
rs29012557	261	SCAFFOLD130087_22009	2662.2	------	14	35942595	8	75349658
ss61480581	261	BTA-34695	2662.2	------	14	35949271	8	75356121

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61480584	261	BTA-34703	2662.2	------	14	36030630	8	75431290
ss61534815	261	BTA-34707	2662.2	------	14	36139358	8	75539444
ss61534809	262	BTA-34698	2668.4	22.2	14	35957127	8	75363477
ss61534810	263	BTA-34700	2685.5	16.8	14	35963488	8	75369838
ss61519163	264	BTA-113824	2710.5	28.1	14	36249385	8	75652763
ss61473631	265	BTA-113816	2713.6	29.8	14	36281160	8	75683267
ss61519165	266	BTA-113826	2716.7	------	14	36249698	8	75653076
ss61473630	266	BTA-113815	2716.7	27.8	14	36284677	8	75683352
ss61566175	267	BTA-92747	2720.9	------	14	36409410	8	75796123
ss61566174	267	BTA-92746	2720.9	25.9	14	36414255	8	75800928
ss61495681	268	BTA-92751	2723	22.2	14	36515654	8	75877536
ss61506685	269	BTA-113819	2730.4	22.1	14	36629130	8	75988844
ss61506686	269	BTA-113820	2730.4	------	14	36629303	8	75989017
ss61530607	270	BTA-26690	2740.2	------	14	36743446	8	76118489
ss61530605	270	BTA-26688	2740.2	27.3	14	36743502	8	76118545
ss61496460	271	BTA-95927	2742.3	------	14	38683822	8	76391415
ss61531756	271	BTA-28831	2742.3	18.5	14	38744448	8	76474444
AAFC02023116	272	BES9_Contig495_375	2756	22.3	14	38967777	8	76743537
ss38334702	273	BTA-12650	2758.6	20.7	14	39055968	8	76821850
ss61494093	273	BTA-87125	2758.6	------	14	39137678	8	76883065
ss61475994	274	BTA-16955	2767.3	23.2	14 & Unknown	39226689	8	77011988
ss61475993	275	BTA-16953	2770.1	------	14	39229854	8	77013602
ss61517008	275	BTA-109776	2770.1	23.3	14	39317422	8	77103558
ss61495336	275	BTA-91674	2770.1	------	14	39401470	8	77261771
ss61516109	276	BTA-108010	2774.6	21.7	14	39588241		
ss61550620	277	BTA-63823	2781.5	25.1	14	39694154	8	77474965
ss61550616	278	BTA-63819	2783.6	17	14	39788481	8	77602702
ss61488018	279	BTA-63807	2802.1	21.5	14	40255202	8	77978691

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61570527	280	BTA-63802	2806.3	------	14	40282106	8	78001107
ss61534821	280	BTA-34717	2806.3	------	14	40334913	8	78055293
ss61534820	280	BTA-34716	2806.3	------	14	40335013	8	78055393
ss61527681	280	BTA-21190	2806.3	18.5	14	40759860	8	78483038
ss61517623	281	BTA-110915	2814.9	26	14	41060494	8	78746647
ss61517610	282	BTA-110902	2819	25.1	14	41121103	8	78803973
ss61519245	283	BTA-113985	2827	34.9	14	41336333	8	79062635
ss61519243	284	BTA-113983	2828.9	30.4	14	41340597	8	79064106
ss61534832	285	BTA-34728	2834.7	31.1	14	41713465	8	79451135
ss61534835	285	BTA-34731	2834.7	------	14	41714013	8	79455338
ss61480591	285	BTA-34753	2834.7	------	14	41757351	8	79485916
ss61534845	286	BTA-34746	2836.6	26.2	14	41885965	8	79629758
ss61534840	287	BTA-34741	2842.6	31.5	14	41896522		
ss61534844	288	BTA-34745	2844.6	------	14	41889718	8	79633012
ss61508242	288	BTA-34738	2844.6	21.8	14	41912581	8	79652382
ss61563137	289	BTA-86970	2861.4	------	14	42136404	8	79873177
ss61473951	289	BTA-115192	2861.4	------	14	42243265	8	79965780
ss61519942	289	BTA-115191	2861.4	18.2	14	42250451	8	79972966
ss61467357	290	BTA-34774	2890.2	18.6	14	42524247	8	132801230
ss38335076	291	BTA-13024	2912.3	24.6	14	42528906	8	132796571
ss38326927	292	BTA-04875	2918.2	32.8	14	42759532	8	80479811
ss61534871	293	BTA-34784	2920.2	31.8	14	42918931	8	80634541
ss61470039	294	BTA-98621	2924.3	31	14	42984766	8	80698574
ss61470040	294	BTA-98622	2924.3	------	14	42984802	8	80698610
ss61470041	294	BTA-98623	2924.3	------	14	42984854	8	80698662
ss61470049	294	BTA-98632	2924.3	------	14	43007339	8	80720647
ss61569280	295	BTA-98634	2930.6	27.3	14	43013073	8	80724766
ss61505809	296	BTA-98615	2941.4	28.8	14	43085799	8	80778933

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61569276	297	BTA-98617	2945.7	15.5	14	43090084	8	80783687
ss38327009	298	BTA-04957	2978.5	29	14	43276492	8	81787123
ss61478221	299	BTA-25540	2980.5	19.2	14	43731398	8	81437396
ss61534880	299	BTA-34794	2980.5	------	14	43598783	8	81581929
ss61534887	300	BTA-34802	2995.8	23.6	14	45426787	8	82164752
ss61534888	300	BTA-34803	2995.8	------	14	45426885	8	82164850
ss38330676	301	BTA-08624	3001	22.4	14	45522923	8	82254890
ss61480605	302	BTA-34832	3009	25.9	14	44447078	8	120498230
ss61485958	303	BTA-55549	3011.7	19.4	14 & 2	44562269	8	120342984
ss61519991	304	BTA-115282	3029.4	29.2	14	44762965	8	120136513
ss61498730	305	BTA-115280	3032.2	26.8	14	44784891	8	120114824
ss61480602	306	BTA-34822	3037.8	------	14	44896577	8	119999093
ss61480600	306	BTA-34819	3037.8	------	14	44900298	8	119995872
ss61480599	306	BTA-34818	3037.8	27.4	14	44900349	8	119995821
ss61480597	307	BTA-34816	3042.4	21.3	14	44900551	8	119995619
ss61480611	308	BTA-34855	3059.5	25.8	14	45700527	8	119800964
ss61534912	309	BTA-34858	3068.3	29.8	14	45938338	8	119568155
ss61501142	309	BTA-34860	3068.3	------	14	46015111	8	119498065
ss61534915	309	BTA-34862	3068.3	------	14	46037094	8	119476200
ss61534918	309	BTA-34865	3068.3	------	14	46044612	8	119468459
ss61534913	310	BTA-34859	3072.5	24.8	14	45941614	8	119564207
ss61501145	311	BTA-34873	3083.5	------	14	46160993	8	119346241
ss61501143	311	BTA-34871	3083.5	24.7	14	46161205	8	119346029
ss61542262	312	BTA-48659	3089.7	------	14	46318025	8	119209763
ss69374986	312	BTA-48660	3089.7	------	14	46374418	8	119152116
ss38327709	312	BTA-05657	3089.7	27.4	14	46392861	8	119136926
ss61509061	313	BTA-48665	3091.8	19.2	14	46447832	8	119084700
ss61534921	314	BTA-34884	3107.1	22.3	14	46739670	8	118795779

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61480615	314	BTA-34887	3107.1	------	14	46871329	8	118692623
ss61534924	314	BTA-34890	3107.1	------	14	46880664	8	118683288
ss61534925	315	BTA-34891	3119.7	21.2	14	47045981	8	118518038
ss61480617	316	BTA-34899	3137.8	26.2	14	47159901	8	118386968
ss38325277	317	BTA-03225	3145	18.6	14	48585654	8	118078492
ss38325279	317	BTA-03227	3145	------	14	48582700		
ss38336085	318	BTA-14033	3166.1	20.3	14	48182373	8	117720853
ss61480626	318	BTA-34928	3166.1	------	14	48239815	8	117660034
ss61480625	318	BTA-34927	3166.1	------	14	48240104	8	117659745
ss61534946	318	BTA-34925	3166.1	------	14	48346730	8	117585573
ss61474497	319	BTA-117443	3179.6	------	14	50208027	8	117190133
ss61521097	319	BTA-117442	3179.6	------	14	50204313	8	117188414
ss61521093	319	BTA-117438	3179.6	26.7	14	50195538	8	117179173
ss61521092	320	BTA-117436	3181.7	17.9	14	50146044	8	117131605
ss61498653	321	BTA-113373	3198.2	17	14	49777446	8	116760144
ss61498658	321	BTA-113380	3198.2	------	14	49773744	8	116754608
ss61498659	322	BTA-113383	3208	14.3	14	49636441	8	116627569
ss61498662	323	BTA-113386	3220.7	19.8	14	49632670	8	116623811
ss61465948	324	BTA-113396	3229.4	23.5	14	49567934	8	116563852
ss61473518	325	BTA-113400	3234.8	23.2	14	49549376	8	116545519
ss61521599	326	BTA-118379	3241.5	25.2	14	49231478	8	116306879
ss61534934	327	BTA-34909	3246.6	------	14	49134448	8	116197061
ss61534932	327	BTA-34907	3246.6	23.4	14	49134359	8	116196972
ss61476770	328	BTA-20070	3256.8	------	14	48760567	8	115805067
ss61516731	328	BTA-109297	3256.8	20.7	14	47947044	8	115632080
ss61473093	329	BTA-111412	3271.8	28.3	14	47645693		
ss38337066	329	BTA-15014	3271.8	------	Unknown		8	115332440
ss69374987	330	BTA-25649	3275.8	32.3	14	50835960	8	80698199

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61527579	331	BTA-20961	3277.8	18.3	14	50801932	8	115018994
ss61517555	332	BTA-110811	3301.9	28.2	14	52015677	8	114513883
ss61480627	333	BTA-34933	3306.2	30.6	14	51658215	8	114241869
ss61534951	333	BTA-34934	3306.2	------	14	51657832	8	114241486
ss61563286	334	BTA-87278	3310.1	22.5	14	51475269	8	114088920
ss61487096	335	BTA-60155	3320.6	28.3	14	51174136	8	113871292
ss61487100	336	BTA-60159	3322.7	26.5	14	51140625	8	113839671
ss61570502	336	BTA-60163	3322.7	------	14	51060765		
ss61548674	336	BTA-60172	3322.7	------	14	51020400	8	113735864
ss61548677	336	BTA-60175	3322.7	------	14	51016269	8	113730193
rs29013644	337	SCAFFOLD106433_368	3329	26.2	14	54323295	8	113661857
CC517527	338	CC517527-184	3333.3	28.5	14	52750593	8	113365131
CC517527	339	CC517527-273	3335.4	26.1	14	52750681	8	113365043
ss61487114	340	BTA-60218	3339.6	27.5	14	52819972	8	113306909
ss61487121	340	BTA-60225	3339.6	------	14	52822701	8	113301403
ss61487120	340	BTA-60224	3339.6	------	14	52823142	8	113300962
ss61487119	340	BTA-60223	3339.6	------	14	52823187	8	113300917
ss61564952	341	BTA-90430	3343.8	18.2	14 & Unknown	54282640	8	113232085
ss61534991	342	BTA-34986	3363.7	------	14	54116008		
ss61534984	342	BTA-34979	3363.7	------	14	54075678	8	129678681
ss61534981	342	BTA-34975	3363.7	24.2	14	54071850	8	112393956
ss61480632	343	BTA-34961	3374.3	28.6	14	53789620	8	112145604
ss61480628	344	BTA-34955	3379.8	31.8	14	53650548	8	111939245
ss61534964	345	BTA-34948	3382.5	30.9	14	53585379	8	111832373
ss61534960	346	BTA-34943	3388.5	25.5	14	53505981	8	111736342
ss61547746	347	BTA-58540	3401.2	------	14	53124052	8	111349666
ss61535008	347	BTA-35011	3401.2	31.2	14 & 24	53083611	8	111293641
ss61535012	348	BTA-35015	3403.1	27.5	14	53083323	8	111293353

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61535020	348	BTA-35023	3403.1	------	14	53077475	8	111288001
ss61535032	349	BTA-35035	3409.1	------	14	53022816	8	111211845
ss61535038	349	BTA-35041	3409.1	------	14	52985374	8	111174901
ss38334432	349	BTA-12380	3409.1	------	14	52917265	8	111062333
ss61506794	349	BTA-115692	3409.1	20.8	14	52898920	8	111056876
ss61520212	349	BTA-115696	3409.1	------	14	52892658	8	111039516
ss61520211	349	BTA-115695	3409.1	------	14	52892202	8	111039060
ss61564183	350	BTA-88967	3428.4	29.4	14	62480674	8	110651477
ss38326828	351	BTA-04776	3432.4	21.9	14	57846844	8	110488715
ss61535045	351	BTA-35076	3432.4	------	14	57800877	8	110446186
ss61535048	351	BTA-35080	3432.4	------	14	57744069	8	110389406
ss61535049	351	BTA-35081	3432.4	------	14	57743897	8	110389234
ss61535052	352	BTA-35084	3440.8	24	14	57647098	8	110317803
ss61535060	353	BTA-35092	3443	21.7	14	57593454	8	110295713
ss61480652	353	BTA-35100	3443	------	14	57493824	8	110170068
ss61535070	354	BTA-35105	3449.6	25.9	14	57444600	8	110109698
ss61535071	355	BTA-35106	3452.8	26.8	14	57406622	8	110063521
ss61467377	356	BTA-35113	3456	19.3	14	57289969	8	109940293
ss61570317	356	BTA-35117	3456	------	14	57288347	8	109932505
ss61535076	356	BTA-35119	3456	------	14	57246696	8	109877133
ss61535078	357	BTA-35122	3473.9	27	14	57140964	8	109772210
ss61535083	358	BTA-35127	3480.6	17.7	14	57003839	8	109657928
ss61494923	359	BTA-90006	3501.1	26	14	56012740	8	108728381
ss61564967	360	BTA-90473	3503.3	------	14	55991213	8	108707354
ss61535097	360	BTA-35154	3503.3	20.6	14	55725538	8	108466006
ss61535096	361	BTA-35153	3510.3	16.6	14	55697084	8	108440244
ss61535095	362	BTA-35150	3520.1	19.6	14	55269422	8	108053902
ss38328658	363	BTA-06606	3525	11.5	14	55263181	8	108057407

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
rs29022898	364	SCAFFOLD75393_2246	3550.6	19	14	62192328	8	107301864
ss61520620	365	BTA-116472	3563.9	27.1	14	61892052	8	107027537
ss61478975	366	BTA-28611	3570	25.1	14	61164407	8	106724154
ss61531642	366	BTA-28619	3570	------	14	61177174	8	106723953
ss61531637	366	BTA-28614	3570	------	14	61174568	8	106696318
ss61535108	367	BTA-35181	3582.5	------	14	61441020	8	106464821
ss61471416	367	BTA-104921	3582.5	19.9	14	61498397	8	106413606
ss61472941	368	BTA-110841	3601.2	8.8	14	57998973	8	105864392
ss61480677	369	BTA-35202	3653.2	23.3	14	58815291	8	104984306
ss61532781	370	BTA-30781	3658.1	22.5	14	58924988	8	104864481
ss61508000	371	BTA-30782	3662.8	------	14	58917305	8	104901268
ss61479531	371	BTA-30777	3662.8	------	14	58933270	8	104857396
rs29012948	371	SCAFFOLD135991_1371	3662.8	------	14	59028774	8	104717057
ss61532780	371	BTA-30774	3662.8	17.9	14	59123466	8	104583678
ss61508259	372	BTA-35162	3672.4	20.8	14	59486320	8	104384835
ss61508258	373	BTA-35161	3679.4	22.3	14	59489245	8	104380836
ss61535131	373	BTA-35214	3679.4	------	14	59411569	8	104337534
ss61535132	373	BTA-35215	3679.4	------	14	59411132	8	104337213
ss61467380	373	BTA-35166	3679.4	------	14	59327372	8	104425432
ss61535130	374	BTA-35210	3688.4	------	14	59870595	8	103943749
ss61535103	374	BTA-35176	3688.4	29.2	14	59924844	8	103871572
ss61535105	375	BTA-35178	3691.2	27.4	14	59925054	8	103871362
ss38331915	376	BTA-09863	3696.9	29.4	14	60024423	8	103764070
ss61480672	377	BTA-35171	3699	------	14	60134943	8	103633326
ss61534626	377	BTA-34335	3699	17.9	14	60249512	8	103512712
ss38328601	378	BTA-06549	3718.2	15.1	14	60397674	8	103343187
ss61501116	379	BTA-34339	3741.2	21.3	14	60559645	8	103182927
ss61501120	379	BTA-34343	3741.2	------	14	60563028	8	103171586

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61480468	380	BTA-34349	3746.8	21.6	14	60597520	8	103165296
ss61480475	381	BTA-34358	3752.6	19.8	14	60752849	8	103020599
ss61490518	382	BTA-72921	3761.8	17.6	14	60955585	8	102825122
ss61535147	383	BTA-35242	3781.1	24.2	14	65514461	8	102311606
ss61501169	384	BTA-35239	3787.5	------	14	65531314	8	102293065
ss61480682	384	BTA-35238	3787.5	25	14	65682763	8	102144282
ss61508266	385	BTA-35235	3791.4	18.5	14	65712888	8	102117036
ss61469665	385	BTA-90525	3791.4	------	14	65791163	8	102027571
rs29024269	386	SCAFFOLD40049_15114	3803.6	20.6	14	65911036	8	101692254
ss61508269	387	BTA-35266	3808.1	18.8	14	66006252	8	101601854
ss61535155	388	BTA-35260	3815.1	18	14	66187331	8	101454591
ss38329727	389	BTA-07675	3819.9	15.9	14	66203557	8	101438111
ss61535152	389	BTA-35257	3819.9	------	14	66220842	8	101424951
ss61535151	389	BTA-35256	3819.9	------	14	66224622	8	101417548
ss61535150	389	BTA-35254	3819.9	------	14	66264451	8	101383854
AAFC02097052	389	BES7_Contig136_464	3819.9	------	14	62800944	8	101356597
ss61535181	390	BTA-35306	3827.1	21	14	63014890	8	101151712
AAFC02008385	391	BES3_Contig324_378	3829.5	19.7	14	63042833	8	101120582
ss61535180	392	BTA-35305	3834.2	------	14	63167733	8	100958089
ss61535177	392	BTA-35302	3834.2	21.4	14	63180973	8	100954653
ss61501183	393	BTA-35300	3836.6	------	14	63195069	8	100933960
ss61501176	393	BTA-35293	3836.6	------	14	63199568	8	100926754
ss61480689	393	BTA-35292	3836.6	16.9	14	63216728	8	100912799
ss61535169	394	BTA-35281	3848.5	23.2	14	63423641	8	100711406
ss61480685	395	BTA-35275	3851.2	3.1	14	63473753	8	100646013
ss61535186	396	BTA-35321	3929.7	20.6	14	65105333	8	99004171
ss61480700	397	BTA-35326	3938.2	22.1	14	65224039	8	98868635
ss38332752	398	BTA-10700	3943.7	21.5	14	65256182	8	98853280

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61535191	398	BTA-35328	3943.7	------	14	65364911	8	98714788
ss61535193	399	BTA-35330	3950.3	21.2	14	66356241	8	98598177
ss61480703	399	BTA-35352	3950.3	------	14	66452630	8	98474464
ss61495198	400	BTA-91119	3962.4	------	14	66691559	8	98227270
ss61495197	400	BTA-91118	3962.4	27.6	14	66691853	8	98226976
ss61480705	401	BTA-35365	3964.6	26.9	14	66918422	8	98009674
ss38328775	402	BTA-06723	3969.2	23.2	14	66957007	8	97986039
ss61535217	403	BTA-35368	3976.3	25.9	14	67151726	8	97841467
ss61535219	403	BTA-35370	3976.3	------	14	67151870	8	97841323
ss61535222	404	BTA-35373	3978.6	25.9	14	67224396	8	97751113
ss61535225	404	BTA-35376	3978.6	------	14	67224638	8	97750871
ss61501193	405	BTA-35382	3982.9	24.9	14	67319415	8	97661243
ss61501195	406	BTA-35384	3985.3	25	14	67323356	8	97657304
ss61501199	407	BTA-35388	3987.7	17.7	14	67323569	8	97657091
ss61535230	408	BTA-35392	4007.8	16.2	14	67370377	8	97610410
ss61467385	409	BTA-35398	4028.7	15.2	14	67708921	8	97256581
ss38336925	410	BTA-14873	4043.4	12.1	14	67927901	8	97034586
ss61535232	410	BTA-35394	4043.4	------	14	67944785	8	97033879
ss61535788	411	BTA-36398	4055.9	16.8	14	68138581	8	96839453
ss61535781	412	BTA-36391	4061.6	------	14	68139960	8	96838074
ss61535777	412	BTA-36387	4061.6	------	14	68142633	8	96836129
ss61535776	412	BTA-36386	4061.6	19.8	14	68143192	8	96835570
rs29023181	413	SCAFFOLD81045_4911	4069.9	20.9	14	68416911	8	96605011
ss61570324	414	BTA-36383	4076	------	14	68345881	8	96693733
ss61478478	414	BTA-26641	4076	20.3	14	68512272	8	96536648
ss61498340	415	BTA-109508	4079	21.3	14	68577300	8	96476875
ss61517168	416	BTA-110113	4081.3	------	14	68753373	8	96377619
ss61517167	416	BTA-110112	4081.3	15.4	14	68756445	8	96374389

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61480711	417	BTA-35415	4097.1	17.9	14	68932843	8	96270962
ss61563877	418	BTA-88359	4109.7	8.7	14	69140076	8	96079852
ss61483348	419	BTA-45484	4146.6	17.6	14	69901688	8	95288253
ss38328018	420	BTA-05966	4158.7	17.6	14	70319249	8	94904482
ss61472673	421	BTA-109661	4173.4	------	14	70746226	8	94480264
ss61472672	421	BTA-109660	4173.4	------	14	70771219	8	94455772
ss61472671	421	BTA-109659	4173.4	------	14	70771265	8	94455726
ss61472670	421	BTA-109658	4173.4	24.1	14	70771455	8	94455536
ss61516937	422	BTA-109650	4177.9	16.1	14	70971671	8	94268707
ss61516952	422	BTA-109675	4177.9	------	14	71003893	8	94244437
ss61516949	422	BTA-109672	4177.9	------	14	71004083	8	94244247
ss38333031	423	BTA-10979	4190	20.7	14	74543809	8	93882210
ss61535256	424	BTA-35457	4194.7	------	14	74705082	8	93766517
ss61480728	424	BTA-35454	4194.7	------	14	74732313	8	93751665
ss61480729	424	BTA-35455	4194.7	------	14	74732748	8	93751230
ss61480722	424	BTA-35448	4194.7	------	14	74735773	8	93746232
ss61535254	424	BTA-35445	4194.7	24	14	74755379	8	93729099
ss61467392	425	BTA-35479	4199.3	23.2	14	75220288		
ss61496838	426	BTA-97523	4201.6	------	14	75439474	8	93306515
ss61535266	426	BTA-35468	4201.6	------	14	75408767	8	93252252
ss61535271	426	BTA-35473	4201.6	------	14	75403902	8	93250052
ss38336767	426	BTA-14715	4201.6	17.7	14	75391577	8	93269098
ss61535281	427	BTA-35498	4209.8	17.6	14	71858373	8	93115751
ss61467399	428	BTA-35493	4223.5	23.1	14	71612630	8	92867348
ss61480732	429	BTA-35488	4230.3	23.4	14	71495651	8	92745528
rs29018423	430	SCAFFOLD60962_16617	4234.7	25.8	14	71492408	8	92740671
ss61535276	431	BTA-35487	4236.9	------	14	71467924	8	92718303
ss61480730	431	BTA-35484	4236.9	------	14	71456155	8	92704113

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61477841	431	BTA-24001	4236.9	14.9	14	71420092	8	92701885
ss61529051	432	BTA-23998	4265.2	29.6	14	72032931	8	92443975
ss61529055	433	BTA-24003	4267.2	30.5	14	72037088	8	92444938
ss61529063	433	BTA-24011	4267.2	------	14	72036445	8	92444938
ss61517338	434	BTA-110397	4269.1	------	14	72083974	8	92401561
ss61517337	434	BTA-110396	4269.1	------	14	72084077	8	92401458
ss61517336	434	BTA-110395	4269.1	------	14	72094269	8	92391306
ss61517335	434	BTA-110394	4269.1	------	14	72100333	8	92356212
ss61517334	434	BTA-110393	4269.1	21.5	14	72103746	8	92379320
ss61494456	434	BTA-88465	4269.1	------	14	72146802	8	92319509
ss61563931	434	BTA-88468	4269.1	------	14	72154797	8	92312013
btcn19347	435	CALB-G433C	4284.5	------	14	72593565	8	91137467
ss38331038	435	BTA-08986	4284.5	20.9	14	72614271	8	91117260
ss61535294	435	BTA-35522	4284.5	------	14	72627739	8	91101414
ss61501211	435	BTA-35526	4284.5	------	14	72685173	8	91055569
ss61535304	436	BTA-35542	4295.7	21.2	14	72929979	8	90671493
ss61480745	437	BTA-35535	4303.9	24.7	14	72963566	8	90573302
ss38329822	438	BTA-07770	4306.5	14.3	14	73080081	8	90044964
ss61477211	439	BTA-21577	4335.8	24.9	14	73756936	8	89072006
ss38326677	440	BTA-04625	4338.3	23.4	14	73821351	8	89022655
ss69374988	441	BTA-116521	4343.6	15.5	14	73839723	8	88989761
ss38327136	442	BTA-05084	4371.3	25.1	14	74057382	8	88740462
ss38327137	442	BTA-05085	4371.3	------	14	74057678	8	88740549
ss38327139	442	BTA-05087	4371.3	------	14	74061350	8	88773023
ss61535306	442	BTA-35550	4371.3	------	14	74112299	8	88685855
ss61535307	443	BTA-35551	4373.6	25.8	14	74116691	8	88684111
ss61535348	444	BTA-35639	4375.9	26.6	14	74211249	8	88590753
ss61564574	444	BTA-89775	4375.9	------	14	74231868	8	88574349

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61564568	444	BTA-89769	4375.9	------	14	74235604	8	88571037
ss61564561	444	BTA-89762	4375.9	------	14	74236059	8	88560800
ss61564562	444	BTA-89763	4375.9	------	14	74236120	8	88560625
ss61512063	445	BTA-100003	4378.1	25.2	14	74231664	8	88574553
ss61564556	446	BTA-89757	4382.3	20.6	14	74239095	8	88566660
ss61474355	447	BTA-116903	4399.7	------	14	75509911	8	88333944
ss61474353	447	BTA-116901	4399.7	27.8	14	75509225	8	88333258
ss61485755	448	BTA-54848	4405	23.9	14	76051896	8	88002569
ss61535354	449	BTA-35651	4421.8	------	14	76557417	8	87054499
ss61535352	449	BTA-35646	4421.8	------	14	76443829	8	86966037
ss61535351	449	BTA-35645	4421.8	31.1	14	76443828	8	86965876
ss61535353	450	BTA-35650	4429.6	35.8	14	76557245	8	87054499
ss61535395	451	BTA-35714	4432.6	34.9	14	76634774	8	87149758
ss61535397	451	BTA-35716	4432.6	------	14	76607858	8	87091388
ss61480788	452	BTA-35709	4435.6	18.4	14	76783267	8	87311948
ss61467410	453	BTA-35683	4466.1	32	14	77253392	8	86011697
rs29010281	453	SCAFFOLD15134_29247	4466.1	------	14	77328439	8	85952595
ss61480791	454	BTA-35719	4468	26.2	14	77447818	8	85857797
ss61480805	455	BTA-35746	4475.8	32.4	14	77612774	8	85624735
ss61480807	455	BTA-35766	4475.8	------	14	78029541	8	85435494
ss61535430	456	BTA-35772	4483.3	34.4	14	78102434	8	85360205
ss61535431	457	BTA-35773	4485.7	27.5	14	78242501	8	85231861
ss61469012	458	BTA-73851	4494.9	26.2	14	78395492	8	85038775
ss61514401	459	BTA-104696	4506.3	32.7	14	78750133	8	84784907
ss61512943	460	BTA-101805	4510.1	20.2	14	78872326	8	84656448
ss61512947	461	BTA-101809	4537.9	19.8	14	78872556	8	84656218
ss61470646	462	BTA-101804	4565.6	30.6	14	78878989	8	84655934
ss61512945	462	BTA-101807	4565.6	------	14	78872413	8	84656361

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss38336693	463	BTA-14641	4571.1	23.8	14	79088834	8	84361814
ss61514441	464	BTA-104755	4586.4	31	14	79478060	8	83988480
ss61476703	465	BTA-19794	4588.5	------	14	79595407	8	83857447
ss61526939	465	BTA-19793	4588.5	29.2	14	79676882	8	83782335
ss61526935	466	BTA-19789	4592.6	28.4	14	79677957	8	83784618
ss61507345	467	BTA-19783	4600.3	23.9	14	79752454	8	83703406
ss61563795	467	BTA-88198	4600.3	------	14	79833886	8	83616399
ss61526930	468	BTA-19772	4612.3	------	14	81377927	8	83223592
ss61526929	468	BTA-19771	4612.3	14.6	14	81377972	8	83223547
AAFC02120841	469	BES8_Contig464_1373	4648	33.6	14	80239209	8	121050738
ss61480811	470	BTA-35777	4649.9	------	14	80214786	8	121017714
ss61480812	470	BTA-35778	4649.9	------	14	80214813	8	121037833
ss38335659	470	BTA-13607	4649.9	------	14	80215106	8	121037942
ss61535435	470	BTA-35782	4649.9	------	14	80225195	8	121037942
ss61535436	470	BTA-35783	4649.9	------	14	80228860	8	121037833
ss61480817	470	BTA-35797	4649.9	------	14	80317731	8	121138905
ss61535446	470	BTA-35807	4649.9	------	14	80327434	8	121149039
ss61535444	470	BTA-35805	4649.9	------	14	80327475	8	121149080
ss61474106	470	BTA-115977	4649.9	33	14	80338899	8	121165739
ss61535455	471	BTA-35816	4655.5	------	14	80436526	8	121288112
ss61535454	471	BTA-35815	4655.5	------	14	80436709	8	121288295
ss61535452	471	BTA-35813	4655.5	32.7	14	80436922	8	121288508
ss61535459	471	BTA-35820	4655.5	------	14	80453450	8	121301908
ss61535463	471	BTA-35824	4655.5	------	14	80479977	8	121334335
ss61535466	472	BTA-35827	4661	------	14	80566038	8	121420018
ss61525323	472	BTA-16698	4661	37.3	14	80568271	8	121422251
ss61535468	472	BTA-35829	4661	------	14	80572946	8	121426427
ss61501234	472	BTA-35830	4661	------	14	80581280	8	121433156

Accession number	Position	SNP	BTA14 RH position (cR)	2pt LOD score	BTA[1]	BTAu_3.1 (BTA14) Position (bp)	HSA[2]	HSA8 Position (bp)
ss61501237	472	BTA-35833	4661	------	14	80581439	8	121433315
ss61501242	472	BTA-35838	4661	------	14	80581558	8	121433434
ss61501246	472	BTA-35842	4661	------	14	80584176	8	121437657
ss61501243	472	BTA-35839	4661	------	14	80584626	8	121438107
ss61508291	473	BTA-35846	4662.8	37	14	80681344	8	121552911
ss61480827	474	BTA-35856	4664.5	29.7	14	80800259	8	121683393
ss38331451	475	BTA-09399	4673.7	33.5	14	80877766	8	121770719
ss38331452	476	BTA-09400	4679.1	30.1	14	80877816	8	121770769
ss61568133	477	BTA-96562	4688.5	35.5	14	80964659	8	121857358
ss61568126	478	BTA-96554	4690.3		14	81028848	8	122013437

[1] Bovine Chromosome
[2] Human Chromosome

APPENDIX TWO

List of markers used for linkage disequilibrium analysis as described in Chapter 4

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61534971	0	YES	YES	0.448	0.135
ss61535522	20.7	YES	YES	0.394	0.115
ss61508244	35.9	YES	YES	0.241	0.159
ss61534850	39.5	YES	YES	0.31	0.243
ss61567744	94	YES	YES	0.417	0.396
ss61524968	102.5	YES	YES	0.325	0.295
ss61524969	102.5	YES	YES	0.229	0.094
ss61467365	115.9	YES	YES	0.158	0.082
ss61534902	122.8	YES	YES	0.313	0.157
ss61467359	122.8	YES			0.363
ss61534889	122.8	YES	YES	0.139	0.292
ss61466612	148	YES	YES	0.055	0.191
ss61466614	148	YES	YES	0.417	0.332
ss61466610	148	YES			0.098
ss61527158	151.4	YES	YES	0.066	0.199
ss61476793	161.4	YES	YES	0.449	0.137
ss61476791	161.4	YES	YES	0.242	0.187
ss61563132	165.7	YES	YES	0.433	0.165
ss38328040	170		YES	0.063	0.313
ss61527143	170	YES	YES	0.445	0.333
ss38334550	193	YES		0.401	0.129
ss38334549	193	YES	YES	0.409	0.092
ss38336818	193	YES	YES	0.168	0.422
ss61497128	202.2	YES	YES	0.385	0.357
ss61497130	202.2			0.361	0.406

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61569297	202.2	YES		0.36	0.411
ss61497126	216.5	YES	YES	0.344	0.226
ss61563531	243.9	YES	YES	0.043	0.095
ss38336286	267.8	YES	YES	0.11	0.229
ss61494247	288.2		YES	0.224	0.224
ss61494249	288.2	YES	YES	0.308	0.26
ss61494244	288.2	YES	YES	0.375	0.036
ss38332001	295.6			0.353	0.151
ss38331999	295.6	YES	YES	0.357	0.147
ss61494232	302.4	YES			0.361
ss61563558	302.4	YES	YES	0.303	0.5
ss61563557	302.4	YES			0.094
ss61528553	308.1		YES	0.269	0.359
ss64899266	308.1	YES	YES	0.269	0.325
ss61517993	310.8	YES	YES	0.351	0.16
ss61563543	313.1	YES	YES	0.297	0.347
ss61563542	313.1			0.296	0.332
ss61480657	332.8	YES		0.175	0.206
ss38336995	338.7		YES	0.173	0.2
ss61535184	413.3	YES	YES	0.292	0.479
ss61520277	445	YES	YES	0.16	0.127
ss61570111	445	YES			0.168
ss61570113	445	YES		0.143	0.035
ss61570114	445				0.168
ss61506800	445				0.172
ss61535114	521.1	YES	YES	0.345	0.059
ss61535115	521.1			0.349	0.055
ss61535128	521.1	YES	YES	0.466	0.338
ss61535134	521.1	YES	YES	0.156	0.116

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN(	MAF (ANGUS)
ss61535159	555.4		YES	0.206	0.277
ss61535160	555.4	YES		0.208	0.262
ss38330264	555.4	YES	YES	0.469	0.471
ss61480731	562.3	YES	YES	0.488	0.394
ss61535249	569.2	YES			0.347
ss61535250	569.2	YES			0.421
ss61535243	571.6	YES	YES	0.316	0.381
ss38336870	595.1			0.426	0.398
ss38336869	613.9	YES	YES	0.426	0.405
ss38335199	635.6			0.222	0.071
ss61480781	635.6	YES	YES	0.248	0.085
ss61535481	676.7	YES			0.09
ss61535491	676.7	YES	YES	0.421	0.397
ss38332498	676.7	YES	YES	0.39	0.408
ss61467412	691	YES	YES	0.097	0.383
ss61508294	691	YES		0.227	0.124
ss61508297	696.8		YES	0.225	0.119
ss61480836	730.1	YES			0.434
ss38326853	737.5	YES			0.244
ss61517788	737.5	YES	YES	0.409	0.332
ss61480838	739.3	YES	YES	0.438	0.059
ss61535502	742.9	YES	YES	0.319	0.032
ss38328865	799.1	YES	YES	0.306	0.152
ss61524851	799.1	YES	YES	0.269	0.22
ss61475686	809.5	YES	YES	0.106	0.157
ss61524852	829.2				0.207
ss61519385	931.5	YES		0.468	0.218
ss69374981	933.3	YES		0.452	0.284
ss61473728	938.9	YES	YES	0.196	0.353

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61473730	938.9	YES	YES	0.451	0.24
ss38330392	950.2	YES	YES	0.27	0.091
ss38330389	950.2			0.375	0.133
ss38330390	950.2	YES	YES	0.373	0.137
ss61535537	956.4				0.16
ss61535536	956.4				0.159
ss61535535	956.4				0.154
ss61535533	956.4	YES			0.163
ss38331774	956.4	YES	YES	0.074	0.442
ss61535531	956.4	YES	YES	0.264	0.276
ss61480853	971.7		YES	0.101	0.239
ss61535524	971.7	YES	YES	0.06	0.439
ss61480851	1000.8	YES			0.44
ss38333453	1016.8	YES			0.048
ss61508302	1016.8	YES		0.403	0.21
ss61480848	1020.5		YES	0.221	0.052
ss61562418	1038	YES	YES	0.488	0.099
ss61562421	1038		YES	0.447	0.093
ss61562420	1038				0.093
ss38328882	1085.9	YES	YES	0.398	0.087
ss38332745	1093.9	YES	YES	0.049	0.374
ss38328883	1093.9	YES	YES	0.127	0.202
ss61493685	1095.9				0.142
ss61493684	1095.9	YES			0.142
ss61493686	1102	YES	YES	0.41	0.367
ss61516354	1117.9	YES	YES	0.452	0.314
ss61472371	1129.3	YES	YES	0.342	0.338
ss61472377	1145.3	YES		0.274	0.281
ss61472375	1149.7	YES		0.278	0.275

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61472374	1149.7		YES	0.274	0.3
ss61506359	1167.5	YES	YES	0.094	0.142
ss61516355	1170.5	YES	YES	0.366	0.234
ss61538773	1185.9	YES	YES	0.384	0.276
ss61482478	1185.9	YES	YES	0.297	0.044
ss61538779	1185.9	YES			0.309
ss61538787	1185.9				0.312
ss61538794	1198.9	YES	YES	0.075	0.278
ss61538793	1198.9	YES	YES	0.449	0.14
ss61482479	1204.4	YES	YES	0.495	0.21
ss61538798	1218.7	YES	YES	0.354	0.21
ss61508645	1218.7			0.335	0.117
ss38329715	1218.7	YES	YES	0.328	0.491
ss38329720	1218.7			0.328	0.487
ss61535551	1265.4	YES	YES	0.131	0.104
ss61501250	1275.3	YES	YES	0.429	0.051
ss61535553	1275.3	YES	YES	0.327	0.111
ss61535555	1275.3	YES			0.411
ss61535556	1277.3	YES	YES	0.226	0.453
ss61535572	1302.4	YES			0.197
ss61527018	1339.2	YES	YES	0.151	0.378
ss61527016	1339.2	YES	YES	0.152	0.314
ss61475000	1355.6	YES	YES	0.338	0.055
ss61525640	1367.1	YES	YES	0.192	0.043
ss61473395	1387.1	YES	YES	0.145	0.226
ss61562834	1395.2	YES			0.278
ss61562833	1407.5	YES			0.205
ss38328592	1425	YES	YES	0.185	0.346
ss61562835	1440.6	YES	YES	0.029	0.295

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61477079	1444.8			0.18	0.191
ss61477078	1444.8	YES	YES	0.18	0.191
ss61531696	1444.8	YES	YES	0.061	0.393
ss61506535	1452.8	YES	YES	0.414	0.46
ss61534587	1452.8	YES	YES	0.31	0.224
ss61534586	1461.3	YES	YES	0.343	0.431
ss38327919	1472.7	YES	YES	0.464	0.455
ss38327921	1486.7	YES	YES	0.497	0.104
ss61534580	1494.4	YES	YES	0.26	0.29
ss61534632	1494.4	YES	YES	0.131	0.307
ss61570309	1519.8	YES			0.045
ss61565351	1524			0.435	0.193
ss61565348	1524		YES	0.434	0.19
ss38336702	1524	YES	YES	0.489	0.191
ss61512747	1528	YES	YES	0.06	0.046
ss61534605	1544.4	YES	YES	0.472	0.329
ss61534596	1585.6				0.064
ss61534598	1605.6		YES	0.031	0.079
ss61534597	1605.6	YES			0.067
ss61534595	1605.6				0.074
ss61534593	1605.6				0.071
ss61570240	1612.2	YES	YES	0.497	0.22
ss61568592	1612.2	YES	YES	0.434	0.22
ss61520603	1614.3	YES	YES	0.15	0.192
ss61520602	1614.3	YES	YES	0.417	0.109
ss61568580	1640.2	YES	YES	0.101	0.104
ss61568577	1640.2	YES	YES	0.163	0.264
ss61496808	1647.1	YES	YES	0.169	0.059
ss61496809	1647.1			0.169	0.066

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61568593	1647.1	YES		0.066	0.167
ss61534613	1662	YES	YES	0.365	0.224
ss61480461	1666	YES	YES	0.072	0.153
ss61534611	1676.7			0.399	0.422
ss61476007	1676.7	YES	YES	0.402	0.431
ss38334682	1705.8	YES			0.092
ss61467336	1708.1	YES	YES	0.34	0.429
ss61480452	1722.5	YES	YES	0.261	0.156
ss38329953	1725.6	YES	YES	0.47	0.328
ss61480486	1765.1	YES	YES	0.099	0.31
ss61478665	1799.1	YES	YES	0.082	0.124
ss61480485	1811.8	YES	YES	0.478	0.397
ss61534638	1814.1	YES	YES	0.054	0.065
ss61534635	1814.1	YES			0.051
ss61480482	1821.5	YES	YES	0.462	0.48
ss61501127	1823.8	YES	YES	0.461	0.437
ss61534653	1826.8	YES			0.02
ss61480496	1835.8		YES	0.178	0.179
ss61480495	1835.8	YES		0.184	0.197
ss61534647	1840.6	YES			0.049
ss61480499	1845.4	YES	YES	0.278	0.321
ss61480494	1847.2	YES		0.368	0.122
ss61480492	1847.2		YES	0.349	0.205
ss61534657	1847.2				0.051
ss61534655	1847.2	YES			0.056
ss61534654	1847.2				0.054
ss61534645	1847.2	YES			0.05
ss61480490	1856.2	YES			0.3
ss61526043	1863.5	YES			0.107

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61499843	1863.5	YES		0.366	0.364
ss61476316	1872.6	YES	YES	0.354	0.209
ss61534661	1906.6	YES		0.137	0.393
ss61534663	1908.7		YES	0.139	0.398
ss61534662	1908.7			0.131	0.394
ss61534664	1908.7			0.137	0.399
ss61480508	1924.1	YES	YES	0.261	0.45
ss61508227	1926.1	YES	YES	0.068	0.256
ss61534670	1945.7	YES	YES	0.3	0.28
ss61534671	1945.7	YES	YES	0.303	0.484
ss61480506	1957	YES	YES	0.371	0.036
ss61534697	1984.1	YES			0.299
ss61534700	1984.1	YES		0.137	0.26
ss61534701	1995.7	YES			0.325
ss61534723	2023.7	YES	YES	0.09	0.301
ss38330358	2079.2	YES	YES	0.407	0.287
ss61516307	2124.6	YES	YES	0.248	0.352
ss61539390	2124.6		YES	0.122	0.201
ss69374985	2124.6	YES		0.124	0.194
ss61472367	2139.5	YES	YES	0.231	0.067
ss61517772	2158.8	YES	YES	0.251	0.192
ss61501854	2165.2	YES	YES	0.301	0.446
ss61501852	2165.2	YES			0.132
ss61482545	2167.6	YES	YES	0.445	0.43
ss61538859	2169.9	YES	YES	0.262	0.246
ss61539954	2174.5	YES	YES	0.459	0.161
ss61529618	2176.4	YES			0.373
ss61529619	2182.4	YES			0.322
ss61514147	2236.2	YES	YES	0.24	0.282

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61467346	2244	YES	YES	0.086	0.351
ss61467347	2244	YES			0.195
ss61480519	2272.1	YES	YES	0.31	0.475
ss61534713	2272.1			0.306	0.475
ss61506756	2297.5	YES		0.038	0.321
ss61469626	2325.6	YES			0.443
ss61508493	2335	YES	YES	0.037	0.036
ss61514555	2344.4	YES	YES	0.173	0.169
ss61481947	2351.5	YES	YES	0.36	0.031
ss61534731	2356.1	YES	YES	0.242	0.472
ss61534734	2369.8	YES	YES	0.217	0.11
ss38333641	2369.8	YES	YES	0.286	0.39
ss61534739	2383.4	YES			0.035
ss61515962	2404	YES	YES	0.259	0.124
ss61515954	2445	YES	YES	0.5	0.101
ss61515956	2449.2	YES	YES	0.154	0.134
ss61516059	2568.7	YES	YES	0.222	0.388
ss61528182	2577.7	YES	YES	0.25	0.071
ss61534803	2586.4		YES	0.038	0.307
ss61534804	2586.4	YES		0.037	0.301
ss61508235	2606.1	YES	YES	0.303	0.464
ss61534799	2630	YES	YES	0.448	0.37
ss61534798	2635.4	YES	YES	0.254	0.476
ss61480572	2639.5	YES	YES	0.498	0.35
ss61480576	2645.8	YES	YES	0.117	0.132
ss61480578	2652		YES	0.105	0.143
ss61480579	2659.7			0.113	0.383
ss61480584	2662.2	YES	YES	0.317	0.491
ss61534815	2662.2	YES	YES	0.302	0.472

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61534809	2668.4	YES		0.422	0.329
ss61534810	2685.5	YES		0.115	0.388
ss61519163	2710.5			0.497	0.3
ss61473631	2713.6			0.498	0.467
ss61519165	2716.7	YES	YES	0.491	0.293
ss61473630	2716.7	YES	YES	0.319	0.479
ss61566175	2720.9	YES	YES	0.253	0.1
ss61566174	2720.9	YES	YES	0.45	0.492
ss61495681	2723	YES			0.108
ss61506685	2730.4	YES			0.153
ss61506686	2730.4				0.157
ss61530605	2740.2	YES	YES	0.151	0.187
ss61496460	2742.3	YES			0.051
ss61531756	2742.3	YES	YES	0.362	0.252
ss61494093	2758.6	YES	YES	0.298	0.155
ss61475994	2767.3	YES	YES	0.384	0.152
ss61475993	2770.1	YES	YES	0.205	0.208
ss61517008	2770.1	YES	YES	0.35	0.22
ss61495336	2770.1	YES			0.324
ss61550620	2781.5	YES	YES	0.483	0.309
ss61550616	2783.6	YES	YES	0.188	0.062
ss61488018	2802.1		YES	0.115	0.116
ss61570527	2806.3	YES			0.12
ss61534821	2806.3			0.112	0.114
ss61534820	2806.3	YES	YES	0.317	0.224
ss61527681	2806.3	YES	YES	0.098	0.344
ss61517623	2814.9	YES	YES	0.438	0.056
ss61517610	2819	YES	YES	0.105	0.252
ss61519245	2827	YES		0.257	0.195

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61519243	2828.9	YES	YES	0.258	0.189
ss61534832	2834.7	YES	YES	0.318	0.281
ss61534845	2836.6	YES	YES	0.132	0.384
ss61534840	2842.6	YES	YES	0.369	0.433
ss61534844	2844.6	YES	YES	0.494	0.329
ss61508242	2844.6	YES			0.079
ss61563137	2861.4	YES	YES	0.135	0.12
ss61473951	2861.4	YES	YES	0.178	0.028
ss61519942	2861.4	YES		0.17	0.22
ss61467357	2890.2	YES	YES	0.277	0.362
ss38326927	2918.2	YES			0.403
ss61534871	2920.2	YES	YES	0.456	0.321
ss61470041	2924.3	YES	YES	0.474	0.192
ss61470049	2924.3	YES	YES	0.184	0.294
ss61569280	2930.6	YES	YES	0.199	0.312
ss38327009	2978.5	YES			0.21
ss61478221	2980.5	YES	YES	0.162	0.083
ss61534880	2980.5	YES	YES	0.407	0.387
ss61534887	2995.8	YES	YES	0.062	0.188
ss61534888	2995.8			0.059	0.183
ss38330676	3001	YES			0.134
ss61485958	3011.7	YES	YES	0.484	0.43
ss61519991	3029.4	YES			0.111
ss61480602	3037.8			0.18	0.056
ss61480600	3037.8	YES		0.191	0.062
ss61480599	3037.8		YES	0.191	0.062
ss61480597	3042.4			0.191	0.059
ss61480611	3059.5	YES	YES	0.427	0.341
ss61534912	3068.3	YES	YES	0.137	0.393

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61534915	3068.3	YES	YES	0.355	0.464
ss61534918	3068.3	YES		0.385	0.293
ss61501143	3083.5	YES	YES	0.384	0.362
ss61542262	3089.7	YES	YES	0.412	0.309
ss69374986	3089.7			0.372	0.388
ss38327709	3089.7	YES	YES	0.372	0.407
ss61509061	3091.8	YES	YES	0.258	0.321
ss61534921	3107.1	YES	YES	0.497	0.4
ss61480615	3107.1	YES	YES	0.248	0.071
ss61534925	3119.7	YES	YES	0.482	0.413
ss38325277	3145	YES		0.164	0.032
ss38325279	3145	YES	YES	0.489	0.38
ss38336085	3166.1	YES			0.107
ss61480626	3166.1		YES	0.034	0.107
ss61480625	3166.1	YES	YES	0.185	0.106
ss61521097	3179.6	YES	YES	0.296	0.213
ss61521093	3179.6	YES	YES	0.331	0.321
ss61498653	3198.2				0.032
ss61498658	3198.2	YES			0.032
ss61465948	3229.4	YES	YES	0.082	0.228
ss61473518	3234.8	YES	YES	0.281	0.271
ss61534932	3246.6	YES		0.179	0.035
ss61516731	3256.8	YES	YES	0.401	0.264
ss38337066	3271.8	YES	YES	0.412	0.22
ss61527579	3277.8	YES	YES	0.269	0.483
ss61517555	3301.9	YES	YES	0.106	0.043
ss61534951	3306.2	YES	YES	0.476	0.472
ss61563286	3310.1	YES	YES	0.415	0.226
ss61487096	3320.6	YES		0.078	0.306

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61570502	3322.7	YES	YES	0.213	0.329
ss61548677	3322.7		YES	0.075	0.289
ss61487114	3339.6	YES	YES	0.243	0.408
ss61487121	3339.6			0.242	0.407
ss61487120	3339.6	YES		0.245	0.38
ss61487119	3339.6	YES		0.248	0.417
ss61534991	3363.7	YES		0.377	0.112
ss61534984	3363.7			0.374	0.083
ss61534981	3363.7	YES	YES	0.375	0.075
ss61534964	3382.5	YES		0.089	0.326
ss61547746	3401.2	YES	YES	0.281	0.385
ss61535020	3403.1	YES			0.294
ss61535038	3409.1	YES	YES	0.104	0.035
ss38334432	3409.1	YES	YES	0.191	0.112
ss61506794	3409.1	YES	YES	0.392	0.102
ss61520212	3409.1	YES	YES	0.229	0.15
ss61520211	3409.1	YES	YES	0.179	0.087
ss61564183	3428.4	YES	YES	0.382	0.303
ss61535045	3432.4	YES	YES	0.156	0.387
ss61535048	3432.4	YES	YES	0.258	0.184
ss61535049	3432.4		YES	0.232	0.386
ss61535060	3443	YES	YES	0.333	0.488
ss61480652	3443	YES	YES	0.156	0.077
ss61535070	3449.6	YES			0.147
ss61535071	3452.8	YES	YES	0.076	0.252
ss61467377	3456	YES	YES	0.143	0.152
ss61535078	3473.9	YES	YES	0.292	0.244
ss61535083	3480.6	YES			0.055
ss61564967	3503.3	YES	YES	0.084	0.161

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61535097	3503.3	YES	YES	0.459	0.2
ss61535095	3520.1	YES	YES	0.392	0.106
ss38328658	3525	YES	YES	0.484	0.046
ss61520620	3563.9	YES	YES	0.122	0.048
ss61478975	3570	YES	YES	0.317	0.264
ss61471416	3582.5	YES	YES	0.5	0.496
ss61472941	3601.2	YES	YES	0.223	0.148
ss61480677	3653.2	YES	YES	0.043	0.388
ss61532781	3658.1	YES	YES	0.097	0.118
ss61508000	3662.8	YES	YES	0.057	0.039
ss61479531	3662.8		YES	0.072	0.039
ss61532780	3662.8	YES	YES	0.128	0.055
ss61508258	3679.4	YES			0.19
ss61535131	3679.4			0.357	0.075
ss61535132	3679.4	YES	YES	0.362	0.076
ss61467380	3679.4	YES	YES	0.216	0.059
ss61535130	3688.4	YES	YES	0.427	0.116
ss61535103	3688.4		YES	0.329	0.394
ss61535105	3691.2	YES			0.411
ss61480672	3699	YES	YES	0.351	0.048
ss61534626	3699	YES	YES	0.254	0.059
ss38328601	3718.2	YES	YES	0.302	0.177
ss61480468	3746.8	YES	YES	0.313	0.38
ss61480475	3752.6	YES	YES	0.412	0.071
ss61490518	3761.8	YES	YES	0.07	0.056
ss61535147	3781.1	YES			0.075
ss61501169	3787.5	YES	YES	0.382	0.396
ss61480682	3787.5	YES			0.055
ss61508266	3791.4	YES	YES	0.238	0.453

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61469665	3791.4	YES	YES	0.071	0.048
ss61508269	3808.1	YES			0.028
ss38329727	3819.9	YES	YES	0.377	0.369
ss61535150	3819.9	YES	YES	0.213	0.437
ss61535181	3827.1	YES		0.187	0.349
ss61535180	3834.2	YES	YES	0.262	0.487
ss61480685	3851.2	YES	YES	0.393	0.479
ss61535186	3929.7	YES	YES	0.194	0.273
ss61480703	3950.3	YES			0.104
ss61495197	3962.4	YES	YES	0.263	0.439
ss38328775	3969.2	YES	YES	0.282	0.242
ss61535217	3976.3		YES	0.09	0.46
ss61535219	3976.3	YES		0.09	0.471
ss61535222	3978.6	YES			0.078
ss61535225	3978.6				0.07
ss61535230	4007.8	YES	YES	0.301	0.256
ss61467385	4028.7	YES	YES	0.443	0.095
ss38336925	4043.4	YES	YES	0.044	0.167
ss61570324	4076	YES	YES	0.328	0.055
ss61498340	4079	YES	YES	0.187	0.268
ss61517168	4081.3	YES	YES	0.186	0.311
ss61517167	4081.3	YES	YES	0.129	0.103
ss61563877	4109.7	YES	YES	0.358	0.102
ss61483348	4146.6	YES	YES	0.267	0.238
ss61472673	4173.4	YES		0.187	0.391
ss61472672	4173.4	YES		0.184	0.415
ss61472671	4173.4			0.186	0.413
ss61472670	4173.4		YES	0.185	0.412
ss61516952	4177.9		YES	0.279	0.167

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61516949	4177.9	YES	YES	0.156	0.173
ss38333031	4190	YES	YES	0.469	0.151
ss61480728	4194.7	YES	YES	0.058	0.059
ss61480729	4194.7	YES	YES	0.442	0.472
ss61480722	4194.7	YES		0.453	0.266
ss61467392	4199.3	YES	YES	0.48	0.232
ss61496838	4201.6	YES	YES	0.205	0.417
ss61535266	4201.6		YES	0.19	0.434
ss61535271	4201.6		YES	0.129	0.438
ss38336767	4201.6	YES	YES	0.276	0.476
ss61535281	4209.8	YES	YES	0.154	0.256
ss61467399	4223.5	YES			0.192
ss61480732	4230.3	YES	YES	0.066	0.067
ss61535276	4236.9	YES		0.067	0.074
ss61480730	4236.9	YES	YES	0.063	0.11
ss61477841	4236.9	YES	YES	0.489	0.388
ss61529051	4265.2			0.046	0.065
ss61529055	4267.2	YES		0.05	0.059
ss61529063	4267.2	YES	YES	0.05	0.06
ss61517338	4269.1	YES			0.087
ss61517337	4269.1	YES	YES	0.338	0.326
ss61517336	4269.1	YES	YES	0.306	0.069
ss61517335	4269.1	YES		0.369	0.416
ss61517334	4269.1			0.368	0.393
ss61494456	4269.1		YES	0.361	0.475
ss61563931	4269.1	YES			0.475
btcn19347	4284.5	YES			0.024
ss38331038	4284.5	YES	YES	0.48	0.169
ss61535294	4284.5		YES	0.402	0.162

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61501211	4284.5	YES	YES	0.491	0.399
ss61480745	4303.9	YES	YES	0.089	0.279
ss61477211	4335.8	YES	YES	0.231	0.496
ss38326677	4338.3	YES	YES	0.025	0.04
ss38327136	4371.3		YES	0.15	0.232
ss38327137	4371.3	YES		0.148	0.248
ss38327139	4371.3		YES	0.304	0.282
ss61535306	4371.3	YES		0.252	0.225
ss61535307	4373.6		YES	0.25	0.226
ss61535348	4375.9	YES		0.332	0.19
ss61564568	4375.9		YES	0.363	0.272
ss61564561	4375.9	YES	YES	0.356	0.246
ss61564562	4375.9			0.36	0.27
ss61512063	4378.1	YES	YES	0.334	0.256
ss61564556	4382.3	YES	YES	0.041	0.436
ss61474355	4399.7	YES	YES	0.388	0.479
ss61474353	4399.7	YES			0.399
ss61535397	4432.6	YES	YES	0.101	0.217
ss61480788	4435.6	YES	YES	0.131	0.487
ss61467410	4466.1	YES	YES	0.254	0.112
ss61480791	4468	YES	YES	0.222	0.121
ss61480805	4475.8	YES	YES	0.097	0.109
ss61535431	4485.7	YES	YES	0.438	0.16
ss61469012	4494.9	YES	YES	0.398	0.361
ss61514401	4506.3	YES	YES	0.358	0.269
ss61470646	4565.6				0.042
ss61512945	4565.6	YES	YES	0.483	0.036
ss61514441	4586.4	YES	YES	0.377	0.382
ss61526939	4588.5	YES	YES	0.34	0.198

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61526930	4612.3	YES		0.492	0.447
ss61526929	4612.3		YES	0.495	0.455
ss61480811	4649.9	YES	YES	0.42	0.404
ss61480812	4649.9	YES		0.423	0.401
ss38335659	4649.9			0.426	0.416
ss61535435	4649.9	YES			0.383
ss61535436	4649.9	YES	YES	0.217	0.117
ss61535446	4649.9	YES	YES	0.357	0.257
ss61535444	4649.9	YES			0.077
ss61474106	4649.9	YES			0.039
ss61535455	4655.5	YES			0.262
ss61535463	4655.5	YES	YES	0.073	0.305
ss61535466	4661				0.152
ss61525323	4661	YES	YES	0.114	0.032
ss61535468	4661	YES	YES	0.209	0.318
ss61508291	4662.8	YES			0.162
ss38331451	4673.7				0.133
ss38331452	4679.1	YES			0.14
ss61568133	4688.5	YES	YES	0.315	0.395
ss61568126	4690.3	YES	YES	0.392	0.463
ss61508252	3.9		YES	0.461	
ss61480590	41.3		YES	0.07	
ss61534608	76.2		YES	0.237	
ss61563126	165.7		YES	0.291	
ss61470052	224.7		YES	0.04	
ss61569304	224.7		YES	0.121	
ss61494220	264.8		YES	0.041	
ss61494233	302.4		YES	0.349	
ss61563555	302.4		YES	0.04	

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61501187	338.7		YES	0.322	
ss38325300	423.2		YES	0.089	
ss61504937	432.9		YES	0.116	
ss61570109	445		YES	0.196	
ss61535116	523.3			0.349	
ss61535391	635.6		YES	0.049	
ss61535493	671.5		YES	0.165	
ss38332499	676.7		YES	0.223	
ss38332497	682.4			0.224	
ss38323917	850.4		YES	0.398	
ss61499653	850.4			0.402	
ss61524855	856		YES	0.468	
ss69374982	933.3		YES	0.079	
ss61480858	950.2		YES	0.395	
ss61522828	1022.3		YES	0.311	
ss61480878	1379.6		YES	0.379	
ss61534577	1444.8		YES	0.058	
ss61480444	1508.6		YES	0.196	
ss38328378	1528		YES	0.357	
ss61568588	1623		YES	0.039	
ss61568589	1623			0.046	
ss61534610	1676.7		YES	0.307	
ss61501125	1823.8		YES	0.273	
ss61534704	1984.1		YES	0.132	
ss61467345	2246.1		YES	0.189	
ss61467342	2276.7		YES	0.032	
ss61534788	2385.6		YES	0.339	
ss61470582	2393.5		YES	0.083	
ss61515968	2438.5		YES	0.037	

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61472242	2522			0.407	
ss61472243	2568.7		YES	0.41	
ss38323652	2662.2		YES	0.431	
ss61480581	2662.2		YES	0.22	
ss61516109	2774.6		YES	0.04	
ss61470039	2924.3			0.191	
ss61470040	2924.3		YES	0.169	
ss61569276	2945.7		YES	0.292	
ss61480605	3009			0.495	
ss61480617	3137.8		YES	0.144	
ss61534946	3166.1		YES	0.426	
ss61474497	3179.6		YES	0.071	
ss61534934	3246.6		YES	0.18	
ss61476770	3256.8		YES	0.123	
ss61473093	3271.8		YES	0.078	
ss69374987	3275.8			0.094	
ss61487100	3322.7		YES	0.199	
ss61480632	3374.3			0.203	
ss61480628	3379.8		YES	0.071	
ss61534960	3388.5		YES	0.092	
ss61570317	3456		YES	0.385	
ss61494923	3501.1		YES	0.127	
ss61535096	3510.3		YES	0.428	
ss38331915	3696.9		YES	0.258	
ss61535155	3815.1		YES	0.065	
ss61535152	3819.9		YES	0.096	
ss61535151	3819.9			0.095	
ss61495198	3962.4		YES	0.199	
ss61501193	3982.9			0.062	

NCBI Accession number	BTA14_RH_position (cR)	Tagged SNPs (Angus)	Tagged SNPs (Holstein)	MAF[1] (HOLSTEIN)	MAF (ANGUS)
ss61501195	3985.3		YES	0.062	
ss61480711	4097.1		YES	0.069	
ss38328018	4158.7		YES	0.28	
ss61516937	4177.9		YES	0.439	
ss61535256	4194.7		YES	0.444	
ss38329822	4306.5		YES	0.058	
ss69374988	4343.6		YES	0.118	
ss61485755	4405		YES	0.265	
ss61480807	4475.8			0.193	
ss61535430	4483.3		YES	0.193	
ss61512947	4537.9			0.486	
ss61526935	4592.6		YES	0.298	
ss61507345	4600.3		YES	0.241	
ss61563795	4600.3		YES	0.099	
ss61535452	4655.5		YES	0.163	
ss61535459	4655.5			0.075	

[1] Minor Allele Frequencies