COMPARATIVE MAPPING OF BOVINE CHROMOSOME 14 FOR IDENTIFICATION OF CANDIDATE GENES UNDERLYING ECONOMICALLY IMPORTANT TRAITS IN BEEF CATTLE

By

TITO ARIEF WIBOWO

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

WASHINGTON STATE UNIVERSITY
Department of Animal Sciences

DECEMBER 2008

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation/thesis of TITO ARIEF WIBOWO find it satisfactory and recommend that it be accepted

_____
(Chair)

_____

_____

_____

ACKNOWLEDGMENT

Without a doubt I would want to first thank God because of His grace and blessings throughout this process of a long and hard work of my study. Only because of Him this achievement was made possible. With this opportunity, I also want to thank Him for putting my mom and dad beside me during these years to always keep me on track and give me the biggest moral support in my life.

Secondly, I want to thank the chair of my committee, Dr. Zhihua Jiang, who spent countless hours and efforts to help me accomplish this project, because it was just simply impossible without his help. I am also grateful to the other committee members who assisted and guided throughout the progression of my project. For Dr. Charles Gaskins, who taught and helped me to untangle numerous statistics problems, Dr. Ruth Newberry, who was always ready to give valuable input for the project, and for Dr. Gary Thorgaard, who gave me perspectives from "outside the box," I am greatly thankful.

Finally, I also want to show my great appreciations for Tyler Daniels, who always kept my sanity during these times, my colleague Jennifer Michal, who has read and assured that these chapters make sense, Dr. Michael MacNeil, who provided the DNA samples and phenotypes for this research, and Merial Ltd., who funded these projects. And for others, whose names I could not possibly mention here because of the limited time and space that contributed significantly for the accomplishment of this project, thank you.

COMPARATIVE MAPPING OF BOVINE CHROMOSOME 14 FOR IDENTIFICATION OF CANDIDATE GENES UNDERLYING ECONOMICALLY IMPORTANT TRAITS IN BEEF CATTLE

Abstract

By Tito Arief Wibowo, Ph.D.
Washington State University
December 2008

Chair: Zhihua Jiang

It has been reported that bovine chromosome 14 (BTA14) is rich in quantitative trait loci (QTL) that are responsible for many economically important traits in cattle. Therefore, the goal of this project was to comparatively annotate the chromosome based on its current genome assembly and explore new potential candidate genes underlying beef carcass, meat quality and eating quality using a Wagyu x Limousin F_2 population. First, we reviewed all reported QTL studies previously conducted on BTA14, standardized the QTL locations and linked the QTL to functional gene regions. We believe that such an anchored QTL map would further improve our ability to understand the genetic complexity of economically important traits located on BTA14 in both dairy and beef cattle. Second, we developed a sequence-based comparative map of BTA14 using human chromosome 8 (HSA8) as an anchor to increase the efficiency and accuracy of identifying and characterizing promising positional candidate genes. Last, we performed a chromosome wide analysis using two different methods, a single gene association with

multiple comparison tests and an interval mapping, to target candidate genes on BTA14 responsible for these economically important traits in cattle. Our study provided a unique set of genetic markers for genomic improvement of carcass, meat quality and healthier products in beef cattle. Such a breeding program based on a marker assisted selection has the potential for a more accurate and rapid animal improvement by predicting genetic merits based on genotypes and using genomic information to design precision breeding systems. More importantly, our work offers new opportunities and tools to develop precision management programs for optimizing meat products based on an animal's genotype, thus leading to consumers' satisfaction for their diet and health.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Dedication

This dissertation/thesis is dedicated to Ham Wibowo and Lieke K. Dirdjosapoetro. My father, who disciplined and taught me to be who I am, and mother, who always cares and loves me of who I always be.

Chapter One:

INTRODUCTION

The beef industry is a major component of the U.S. agricultural economy. It is estimated that the U.S. beef industry is worth $175 billion, and is run by approximately 800,000 ranchers and cattlemen. They conduct business in all 50 states and contribute economically to nearly every county in the nation (http://www.beefusa.org). Statistics show that in January 2004, there were 94.9 million cattle in the United States. Beef is the number one protein source in American diets according to 2003 USDA consumption data and the demand for beef continues to grow. The bottom line is consumers are willing to pay more for the beef they love. Specifically, consumers in general prefer, and will pay more, for steaks that are tastier, juicier, tenderer and more flavorful. However, carcass, meat quality and eating quality in beef are quantitative in nature because they are influenced by complex interactions between numerous environmental factors and alleles of many genes. Furthermore, selection of breeding animals based on any carcass traits requires tremendous effort, expense and time given that measurements must be made on relatives of the animals undergoing selection. Therefore, genomic technology may lead to more accurate and rapid animal improvement for those phenotypic traits that are difficult to measure (Green et al., 2007).

In fact, the bovine genome mapping community began to map quantitative trait loci (QTL) for these economically important traits as early as 1995 when George et al. published one of the first QTL studies in cattle whose impact was very profound in identifying and characterizing QTL and positional candidate genes. Among 29 autosomes and a pair of sex chromosomes in the bovine genome, Bos Taurus 14 (BTA14) took our attention because it was reported in many studies to harbor many QTL

for important traits in cattle. Recently, we found that there are more than 40 investigations dealing with QTL or genes for various traits on BTA14, which reported a total of 126 QTL spanning the chromosome alone in both dairy and beef cattle. Unfortunately, although many strong positional or physiological candidate genes have been proposed, the results from different reports are still inconclusive. Therefore, the goal of this project was to pursue a comparative annotation of BTA14 and apply a chromosome wide analysis in order to further characterize and identify potential geneshat are responsible for carcass, meat quality and eating quality in a Wagyu x Limousine F2 cattle population.

In order to achieve the goal described above, we proposed four objectives as followings. The first objective was to review QTL investigations on BTA14 to standardize QTL locations and link the QTL to the functional gene regions (chapter 2). We believe that such an anchored QTL map would further improve our ability to understand the genetic complexity of economically important traits located on BTA14 in both dairy and beef cattle. The second objective was to develop an *in-silico* comparative map of BTA14 using HSA8 as an anchor to provide a solid foundation to increase the efficiency and accuracy of identifying and characterizing promising positional candidate genes these traits (chapter 3). Moreover, our third objective was to identify and characterize a strong positional candidate gene on BTA14 (chapter 4). Finally the last objective was to perform a chromosome wide analysis using two different methods, single gene association tests and interval mapping, to discover or confirm more QTL on BTA14 for many economically important traits in cattle (chapter 5). We believe that identification of

genetic markers that could pinpoint animals with superior carcass, beef quality and eating quality would ultimately lead to products with higher consumer satisfaction. In fact, as early as in 1998, the U.S. beef industry initiated a Carcass Merit Project to validate markers for economically important carcass and customer satisfaction traits, such as marbling, tenderness and composition and thus provide the tools and mechanisms to genetically identify superior animals in the U.S. beef population which would produce progeny with the greatest potential for meeting the demands of the consumers of today and tomorrow.

Chapter Two:

LITERATURE REVIEW

2.1 Introduction

Livestock species, especially cattle, play an important role in economic development worldwide (Williams et al. 2002). To increase profitability, cattle with superior, economically important traits such as growth, milk production and meat quality, have been selected and used as breeding stock. Traditionally, phenotypic and pedigree data have been used to select and pair the best sires and dams, which would in many cases result in offspring with improved phenotype values over the previous generation. This simple method of data collection and selection tripled the U.S. milk production from 1940-1991 with fewer cows (Ashwell et al. 2002). Therefore, any sophisticated techniques to choose sires and dams that are genetically superior will further enhance the process with more improvement and economical gain.

Finding genes responsible for these economically important traits, however, is challenging because they are quantitative in nature (Guitierrez-Gill et al. 2001; Zhang et al. 1998). In other words, these traits are polygenic and are controlled by the accumulative action of many Mendelian genes. Moreover, the number of genes involved is unknown, and environmental factors can also complicate the process because they can have a confounding effect on phenotypes (Boichard et al. 2003; Zhang et al. 1998; Guitierrez-Gill et al. 2001; Heyen et al. 1999). Nonetheless, recent developments in molecular biology and statistical methodologies allow the possibility to localize regions/genes in the genome or chromosome that are responsible for traits of interest (Kim et al. 2003; MacNeil et al. 2002; Casas et al. 2003).

To date, bovine chromosome 14 (BTA14) has been one of the most widely studied chromosomes for quantitative trait loci (QTL) related to many economically important traits in cattle (Marques et al. 2007). There are more than 40 investigations dealing with QTL or genes for various traits, which reported a total of 126 QTL spanning the chromosome alone. In the present review, we reviewed some aspects related to QTL mapping and then surveyed QTLs or genes on BTA14 discovered in both dairy and beef cattle. Since many of these studies used different sets of markers, mostly microsatellites, we integrated them into the current genome assembly (Btau 4_0) (http://www.ncbi.nlm.nih.gov/projects/genome/guide/cow/) in order to standardize QTL locations. Such a genome assembly anchored QTL map provides the best view on QTL density for each phenotype and the potential links between QTL and functional genes for future study.

2.2 QTL mapping: basics

Generally speaking, developing an accurate genetic map and performing an unbiased statistical analysis on the population data are the heart of QTL analysis study. There are many programs available that can be utilized for analyzing a research data. The main practical problem in the past, nonetheless, had been the availability of genetic markers (Erickson et al. 2004, Doerge et al. 1997). The advancement of genetic markers in the last ten years, therefore, is truly the one that results in the explosion of QTL studies in many species. Besides the availability of genetic markers and statistical software to analysis the data, there are two other main factors that can influence the data analysis: population size and the experimental design that will also be described briefly in this chapter.

2.2.1 Genetic markers

There are, in general, two kinds of genetic markers that have been widely used in the QTL mapping study: codominant and dominant markers. The former markers, such as microsatellite markers and SNPs, can reveal the identity of all three genotypes. The latter markers, on the other hand, such as amplified fragment length polymorphisms (AFLPs) and random amplified polymorphic DNA (RAPDs), cannot distinguish the heterozygous from the dominant homozygous genotypes. Hence, this method only recognizes the dominant type (the presence) and the recessive homozygous genotype (the absence). As the result of masking one of the genotype states, using dominant markers needs to increase the sample size for the analysis. In an F_2 intercross design, for instance, using dominant markers in the repulsion phase, as many as 20 times more individuals must be used in comparison of employing the counterpart codominant markers (Erickson et al. 2004). Not only the types of the genetic markers used in the experiment will affect the QTL study, but the distribution of the markers also influences the accuracy of locating a QTL and the effect of it. More markers added in the experiment analysis increase the accuracy and precision in locating the QTL. Nevertheless, more markers also mean less genetic interval and fewer recombination events between a pair of markers that result in the un-replicated genotypes problem (Erickson et al. 2004).

2.2.2. Reference population

When developing a reference population for QTL mapping, crossbreeding of inbred lines is one of the best methods because it results in, not only maximizing linkage

disequilibrium between allele markers and the QTL, but also ensuring that only two QTL alleles segregate (Mackay et al. 2001). This reference population can be acquired by backcrossing an F_1 generation with the inbred parental lines or intercrossing and crossing the F_2 population. Developing an inbred population, however, can be impractical because of generation times, for instance, and limit the number of alleles at anyone of the QTL locations (Erickson et al. 2004). Therefore, outbred population, such as outcrossed designs and sib-pair method, is an option, especially in searching for multiple QTL in one or more populations. An experimental design can, furthermore, affect the sample size required in the data analysis. An experiment that uses a backcross, for instance, needs to double the sample size to infer QTL with the same accuracy of an F_2 design experiment (Erickson et al. 2004). Increasing the sample size can also enhance the accuracy and precision in observing the recombinant events (Doerge et al. 2001) and the ability to detect QTL with smaller effect (Mackay et al. 2001). Using a small sample size, on the other hand, can result in a profound statistical problem called the Beavis effect (Beavis et al. 1994), where the effect of a QTL is exaggerated.

2.2.3. QTL detection

The simplest method for a QTL mapping approach is by using a single marker association. This method was used by Sax in 1923 and was perfected by Thoday (1961), who determined that a positive correlation of genotype and phenotype was, in fact, an evidence for a QTL-marker interaction (Erickson et al. 2004). There are, moreover, basically two general methods for statistical analysis of a single marker association: Least Square (LS) and Maximum Likelihood (ML) methods. The LS method, which is more

commonly used since it is less computationally intensive, utilizes ANOVA or regression analysis to find the difference between the marker class means (MacKay et al. 2001). This method, nonetheless, holds the assumptions that the error terms have a constant variance, which is not always true. Therefore, maximum likelihood method, offers another alternative to analyze the marker-trait data by taking into accounts all QTL data being mixtures of normal distributions (MacKay et al. 2001).

Doerge et al. in 1997, however, discovered that a single marker analysis had a major problem with a confounding effect between the magnitude and the location of the marker (Doerge et al. 1997). Therefore, an interval mapping approach by Lander and Botstein using a maximum likelihood method is used to locate the likelihood of a QTL present between intervals of two markers (Lander et al. 1989). These two early methods of detecting one QTL at a time, nonetheless, still bias the result because it ignores the possibility of multiple linked QTL. Therefore, more powerful analysis, such as composite interval mapping (CIM) by Basten et al. 2001 and multiple interval mapping (MIM) by Zeng et al, in 1999, were developed. CIM combines maximum likelihood mapping and multiple regression analysis to reduce both the bias of estimating a QTL position and effect by multiple linked QTL (MacKay et al. 2001). However, MIM provides the best stable model for multiple QTL because it, not only provides the position effect, but also the main and interaction effect of multiple QTL. One true weakness for both of CIM and MIM method, however, is that they are highly dependent on the model used. Thus, the effect and position of the QTL can varies based on the difference of the marker cofactors used in the model (MacKay et al. 2001).

Some of the QTL programming program that are available in the public databases include: MAPMARKER/QTL, QTL CARTOGRAPHER, BMAPQTL and QTL EXPRESS. Mapmarker/QTl uses an interval mapping approach, whereas BmapQTL and QTL express use Bayesian interval mapping and outbred QTL methods, respectively. QTL cartographer, furthermore, has the option to use CIM, MIM or Bayesian interval mapping approach.

2.3. QTL detected on BTA14 in dairy cattle

In dairy cattle, the majority of QTL mapping on BTA14 are related to milk production traits, such as milk yield, fat %, fat yield, protein % and protein yield (Ashwell et al. 1997, 2002 and 2004; Bagnato et al. 2007; Bennewitz et al. 2003 and 2004; Boichard et al. 2003; Coppieters et al. 1998; Harder et al. 2006; Heyen et al. 1999; Kuhn et al. 2004; Kaupe et al. 2007; Looft et al. 2001; Mosiq et al. 2001; Rodriguez-zas et al. 2002; Ron et al. 1999; Schnabel et al. 2005; Thaller et al. 2003; Viitala et al. 2003 and Zhang et al. 1998). Other traits, such as reproduction (pregnancy, ovulation and twining) (Ashwell et al. 2002; Cobanoglu et al. 2005; Gonda et al. 2004; Kaupe et al. 2007; Schnabel et al. 2005), health (somatic cell score, SCS, and clinical mastitis) (Ashwell et al. 1998; Kaupe et al. 2007; Klungland et al. 2001; Rodriguez-Zas et al. 2002; Rupp et al. 2003; Viitala et al. 2003 and Zhang et al. 1998) and udder related traits (Ashwell et al. 1997 and 2002; and Schnabel et al. 2005) were also investigated on BTA14. References, significance levels (p values or F values), peak or flanking marker(s), genome-anchored locations (in

Mb), and linkage map locations (in cM) for each of these QTL are listed in Supplemental Table 2.1.

As shown in Figure 2.1, the region around 0 – 10 Mb has many QTLs. There are 56 QTL for five milk production traits, and seven QTL for other traits. Fifty-two of these 56 milk production related QTLs were clustered in a region of ~3.6 Mb (0.26-3.81 Mbp), including 15 QTL for fat % or fat content (Ashwell et al. 2004; Bennewitz et al. 2003 and 2004; Boichard et al. 2003; Heyen et al. 1999; Kaupe et al. 2007; Kuhn et al. 2004; Rodriguez-Zas et al. 2001; Ron et al. 1999; Thaller et al. 2003), 13 QTL for fat yield (Ashwell et al. 2004; Bennewitz. et al. 2003 and 2004; Boichard et al. 2003; Heyen et al. 1999; Kaupe et al. *2007;* Looft et al. 2001; Thaller et al. 2003; Viitala et al. 2003; Winter et al. 2002), 10 QTL for milk yield (Ashwell et al. 1997; Bennewitz. et al. 2004; Boichard et al. 2003; Harder et al. 2006; Kaupe et al. 2007; Looft et al. 2001; Rodriguez-Zas et al. 2002; Thaller et al. 2003), 10 QTL for protein% or protein content (Bennewitz et al. 2004; Boichard et al. 2003; Heyen et al. 1999; Kaupe et al. 2007; Thaller et al. 2003), and four QTL for protein yield (Bennewitz et al. 2003; Kaupe et al. 2007, Looft et al. 2004; Thaller et al. 2003), respectively. Interestingly, this region of ~3.6 Mb on the assembly was expanded to a region of ~30 cM reported by different groups in the linkage map (Supplemental Table 2.1). Therefore, our genome anchored QTL map significantly narrows the physical distance of QTL regions and perhaps provides the precise locations for identification of candidate genes.

The high density of QTL for milk production traits in this small region of BTA14 has led to exploration of candidate genes in the region. In 2002, Grissart and colleagues proposed that bovine *diacylglycerol acyltransferase 1* (*DGAT1*) (at 0.44 Mbp) is a promising candidate gene because a missense mutation (Lysine232Alanine) in the gene could explain the phenotypic variance in milk fat content and other milk characteristics. DGAT1 is a microsomal enzyme that utilizes diacylglycerol and fatty acyl CoA as substrates in order to catalyze the final stage of triacylglycerol synthesis (Cases et al. 1998). Therefore, this gene should affect the fat metabolism, including fat yield and percentage in the milk (Cases et al. 2002). A knock-out study showed that both male and female $Dgat^{-/-}$ mice, even those fed a high fat diet, stabilized their weights and resisted fat storage (Smith et al. 2000), indicating the importance of *DGAT1* in fat metabolism. Winter et al. (2002) further found that the lysine variant was associated with higher milk fat content compared to its counterpart alanine variant in several cattle breeds.

Interestingly, Bennewitz and colleagues (2004) observed with a genome-wide scan a conditional QTL effect on fat percentage at the proximal end of the chromosome and for protein percentage at a more distal chromosomal region in addition to the diallelic *DGAT1* effects on milk, fat, and protein yield and fat and protein percentage. The author argued that this conditional QTL effect might be caused by one or more additional alleles segregating at DGAT1 that were not previously detected, or by a second quantitative trait locus affecting these traits. Kuhn et al. (2004) reported strong evidence for segregation of at least three alleles in the promoter region of the *DGAT1* gene that affects milk fat percentage. In the centromeric region of BTA14, cytochrome P450, family 11, subfamily

B (*CYP11B1*) was also suggested to be the causative gene for the QTL related to fat metabolism (De Roos et al. 2007). The *CYP11B1* gene was negatively associated with milk yield and protein yield, but positively associated with fat content (Kaupe et al. 2007).

An additional 13 QTL for milk production traits were also identified outside this first 10Mb region on BTA14, including two at 33.62 Mb and 11.78 Mb – 33.62 Mb for fat percentage or fat content (Ashwell et al. 1997; Zhang et al. 1998), two at 24.67 Mb – 27.34Mb and 61.48 Mb – 76.75 Mb for fat yield (Ashwell et al. 2004; Harder et al. 2006), three at 33.62 Mb, 34.158 Mb – 51.17 Mb and 65.03 Mb for milk yield (Ashwell et al. 1997; Heyen et al. 1999; Schnabel et al. 2005), five (one at 7.87 Mb – 39.55 Mb, one at 34.16 Mb – 51.17 Mb and three at 73.84 Mb) for protein % (Ashwell et al. 2002; Mosiq et al. 2001; Rodriguez-Zas et al. 2002; Schnabel et al. 2005; Viitala et al. 2003) and one at 65.03 Mb – 75.88 Mb for protein yield (Ashwell et al. 2004) (Figure 2.1 and Supplemental Table 2.1). However, no candidate genes have been identified for these QTL on BTA14.

During the past several decades, the high intensity of selection for milk yield has led to a significant improvement in milk production, but it has also led to a decline in reproductive efficiency in dairy cattle. In the United States, Washburn and colleagues (2002) reported that from 1976 to 1999 the average number of days open increased from 122 to 152 days for Jerseys and from 124 to 168 days for Holsteins. Services per conception also increased, from 1.91 to 2.94 services for both breeds during the same

period. There are at least 6 QTL for reproductive traits detected on BTA14 (Figure 2.1 and Supplemental Table 2.2). Calving ease QTL is located at 1.294 Mb (Kaupe et al. 2007), while both ovulation rate and twining rate might share a common QTL region from 34.16 Mb to 65.38 Mb (Cobanoglu et al. 2005; Gonda et al. 2004). However, it seems that three pregnancy related phenotypes have no common QTL locations: non-return rate is placed at 0.44 Mb (Kaupe et al. 2007), pregnancy rate at 3.81 Mb – 11.78 Mb (Ashwell et al. 2004), and daughter pregnancy rate (DPR) at 34.16 Mb – 51.17 Mb (Schnabel et al. 2005) (Figure 2.1). Defining QTL ontology should be considered by the community in the future to search for common QTLs for a given phenotype.

Mastitis affects every dairy farm and up to 50% of all dairy cattle in the United States (Owen et al. 2000). Economic losses are an estimated $180 per cow per year or $2 billion annually in the United States of America (Nash et al. 2000). A total of 10 QTLs for the disease were reported on BTA14, including three for clinical mastitis (Klungland et al. 2001; Rupp et al. 2003; Viitala et al. 2003) and seven for somatic cell score (SCS) (Ashwell et al. 1997; Ashwell et al. 1998; Kaupe et al. 2007; Rodriguez-Zas et al. 2002; Rupp et al. 2003; Zhang et al. 1998). A region around 1.29 Mb with CYP11B1 harbors a significant QTL for SCS, but two other regions, one at 7.86 Mb – 39.55 Mb and the other at 61.48 Mb - 73.84 Mb might share QTL for both clinical mastitis and SCS. These data indicate that QTL information on SCS would help reduce incidence of clinical mastitis in dairy cattle. In addition, SCS has been widely used to indirectly measure udder traits in dairy cattle (Poso et al. 1996). As shown in Figure 2.1 and Supplemental Table 2.2, QTL

for udder traits are also found in the same region as QTL for SCS, indicating that these traits are either controlled by the same gene or by genes in a linkage.

2.4. QTL detected on BTA14 in beef cattle

In beef cattle, QTL mapping has been mainly focused on growth traits, carcass and meat quality. The targeted phenotypes include hot carcass weight (Kim et al. 2003; Mizoshita et al. 2004), rib eye area (REA) (Stone et al. 1999; MacNeil et al. 2002), average daily gain (ADG) (Kneeland et al. 2004; Mizoshita et al. 2004), intramuscular fat deposition (marbling) (Casas et al. 2003) and subcutaneous fat depth (backfat-EBV or fat thickness) (Casas et al. 2000; Casas et al. 2003; MacNeil et al. 2002, Moore et al. 2003) (Supplemental Table 2.3). *Bos indicus* (Brahman) and *Bos taurus* (Angus or Hereford) breeds have been heavily used in QTL mapping by Casas et al. (2003), Kneeland et al. (2004), Stone (2004) and Kim et al. (2003). Other commonly used breeds include: Wagyu (Mizoshita et al. *2004*), Belgian Blue (Casas et al. *2003*) or composite breeds (½ Red Angus, ¼ Tarentaise and ¼ Charolais) (MacNeil et al. 2002), or mixed breeds, such as ¼Angus, ¼ Hereford, ¼ Red Poll and ¼ Pinzgauer cows with Piedmontese X Angus sires (Casas et al. 2000).

In contrast to the high number of QTL discovered in dairy cattle, the region between 0 Mb – 10 Mb on BTA14 harbor only a few QTLs for beef cattle traits, including two for post-weaning ADG (Kneeland et al. 2004), one for fat thickness (Casas et al. 2003, Moore et al. 2003), one for %USDA choice (Casas et al. 2003) and one for mean body weight (Mizoshita et al. 2004) (Figure 2.1 and Supplemental Table 2.3). *DGAT1*, which is responsible for lipid metabolism traits in dairy cattle, was also tested in beef cattle.

Unfortunately, Moore and colleagues (2003) failed to observe any significant association of the *DGAT1* polymorphism with backfat EBV. On the other hand, Barendse et al. (1999) found that the *TG* gene, which is located at 7.658 Mb on the chromosome, is significantly associated with marbling score in beef cattle. The association was further confirmed by the same group using another population of feedlot cattle (Wood et al. 2006). In contrast, Moore and colleagues (2003) failed to confirm the association of polymorphisms in the *TG* gene with any lipid metabolism traits (such as backfat EBV). Even so, Moore et al (2003) did not exclude the possibility of other polymorphisms in *DGAT1* or *TG* that might have significant effect on backfat reported in other studies.

Most QTL on BTA14 discovered in beef cattle fall into a 30 Mb region, from 15 Mb to 45 Mb, including two QTL for post-weaning ADG, four for pre-weaning ADG (Kneeland et al. 2004), three for birth weight (Kneeland et al. 2004), three for carcass weight (Kim et al. 2003; Mizoshita et al. 2004), one for fat thickness (Casas et al. 2003), one for marbling score (Casas et al. 2003) and two for mean body weight (Mizoshita et al. 2004) (Figure 2.1 and Supplemental Table 2.3). In addition, there is evidence that somewhere along these two regions on BTA14, there are also QTL for some eating quality traits (24.67-46.69 Mbp): pH at 24 hour post-mortem and grilled beef flavor intensity bracketed by RM11-PZ271, and grilled abnormal flavor intensity bracketed with CSSM066-RM11 markers (Guiterrez-Gil et al. 2007).

Since both *DGAT1* and *TG* are located in the region of 0 Mb – 10 Mb on BTA14, it is obvious that they cannot serve as candidate genes for QTLs located in the region of 15

Mb – 45 Mb. Therefore, we targeted two candidate genes in the 15 Mb – 45 Mb region: corticotrophin releasing hormone (*CRH*) and fatty acid binding protein 4 (*FABP4*). The former is located at position 31.49 Mb, while the latter is positioned at 41.95 Mb on the newly assembled bovine genome map. CRH, which is a stress hormone, is released to the anterior pituitary to stimulate the secretion of adrenocorticotrophic hormone (ACTH), which upregulates cortisol. Cortisol has many significant metabolic effects that include stimulating gluconeogenesis in the liver, inhibiting glucose uptake in muscle and adipose tissue, and stimulating fat breakdown in adipose tissue (Wibowo et al. 2007). In addition, transgenic mice with overexpression of *CRH* exhibit muscle wasting, decreased linear growth and obesity (Stenzel-Poore et al. 1992), whereas in porcine, CRH was reported to be significantly associated with backfat thickness, carcass length, average daily mass gain and REA (Murani et al. 2006). In cattle population, Buchanan et al. 2005 showed three SNPs that were associated with REA ($P < 0.034$) and hot carcass mass ($P < 0.0015$) in a Charolais-cross steer population. In 2007, Wibowo et al., discovered a new SNP in the bovine *CRH* gene that is highly associated with backfat thickness ($P <0.001$) in a Wagyu X Limousin F_2 population. Thus, *CRH* is a good positional candidate gene for fat-related traits.

FABP4, on the other hand, is a member of the fatty acid binding protein family that is thought to play a major role in the regulation of lipid and glucose homeostasis through its interaction with perioxisome proliferator-activated receptors (PPARs) (Damcott et al. 2004). In addition, FABP4 is also shown to interact with lipase, a primary enzyme involved in lipid catabolism, which regulates lipid hydrolysis and intracellular fatty acid

trafficking (Michal et al. 2006). Hence, *FABP4* is a strong candidate gene for obesity as it is also located in the region of quantitative trait loci (QTL) for serum leptin levels in mice (Ogino et al. 2003). In 2006, Michal et al. reported a significant association of *FABP4* gene with marbling score and fat thickness in a Wagyu X Limousin F_2 cross with a P-value of 0.0246 and 0.0398, respectively.

2.5. Conclusion

We reviewed more than 40 studies related to identification of QTLs for economically important traits on BTA14 in both dairy and beef cattle and anchored 126 QTL into the current bovine genome assembly. Such a process standardized the QTL locations by avoiding many conflicts reported on different linkage maps and linked the QTL to the functional gene regions. We believe that such an anchored QTL map would further improve our ability to understand the genetic complexity of economically important traits located on BTA14 in both dairy and beef cattle.

2.6. *Reference*

Ashwell MS, Da Y, Vanraden PM, Rexroad, Jr. CE, Miller RH, 1997. Detection of loci affecting milk production and health traits in an elite US Holstein population using microsatellite markers. Anim Genet 28(3): 216-222.

Ashwell MS, Da Y, Van Tassell CP, Vanraden PM, Miller RH, Rexroad Jr. CE, 1998. Detection of putative loci affecting milk production and composition, health and type traits in a united states Holstein population. J. Dairy Sci. 81(12): 3309-3314.

Ashwell MS, Van Tassell CP, Sonstegard TS, 2002. A genome scan to identify quantitative trait loci affecting economically important traits in a US Holstein population. J. Dairy Sci 84(11): 2535-2543.

Ashwell MS, Heyen DW, Sonstegard TS, Van Tassell CP, Da Y, VanRaden PM, Ron M, Weller JI, Lewin HA, 2004. Detection of quantitative trait loci affecting milk production, health and reproductive traits in Holstein cattle. J. Dairy Sci 87: 468-475.

Bagnato A, Schiavini, Rossoni A, Maltecca C, Dolezal M, Medugorac I, Solkner J, Russo V, Fontanesi L, Friedmann A, Soller M, Lipkin E, 2008. Quantitative trait loci affecting milk yield and protein percentage in a three-country Brown Swiss population. J. Dairy Sci. 91: 767-783.

Barendse WJ. 1999. Assessing lipid metabolism. International patent application number PCT/AU98/00882. International publication No. WO 99/23248.

Basten CJ, Weir BS, Zeng ZB, 2001. QTL cartographer version 1.15. In: statistics, North Carolina State University: Raleigh NC

Beavis WD, 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the forty-ninth annual corn and sorghum industry research conference. Pp. 250-266. American Seed Trade Association, Chicago, IL.

Bennewitz J, Reinsch N, Grohs C, Leveziel H, Alain M, Thomsen H, Xu N, Looft C, 2003. Combined analysis of data from two granddaughter designs: a simple strategy for QTL confirmation and increasing experimental power in dairy cattle. Genet. Sel. Evol. 35: 319-338.

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: 431-442

Boichard D, Grohs C, Bourgeouis F, Cerqueira F, Faugeras R, Neau A, Rupp R, Amigues Y, Boscher MY, Leveziel H, 2003. Detection of genes influencing economic traits in three French dairy cattle breeds. Genet. Sel. Evol 35: 77-101.

Buchanan FC, Thue TD, Yu P, Winkelman-Sim DC, 2005. Single nucleotide polymorphisms in the corticotrophin-releasing hormone and pro-opiomelancortin genes are associated with growth and carcass yield in beef cattle. Anim Genet 36(2): 127-131.

Casas E, Shackelford SD, Keele JW, Stone RT, Kappes SM, Koohmaraie M, 2000. Quantitative trait loci affecting growth and carcass composition of cattle segregating alternate forms of myostatin. J. Anim Sci 78: 560-569.

Casas E, Shcakelford SD, Keele JW, Koohmaraie M, Smith TPL, Stone RT. 2003. Detection of quantitative trait loci for growth and carcass composition in cattle. J. Anim Sci 81: 2976-2983.

Cases S, Smith SJ, Zheng YW, Myers HM, Lear SR, Sande E, Novak S, Collins S, Welch CB, Lusis AJ, Erickson SK, Farese RV, Jr., 1998. Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc. Natl. Acad. Sci 95: 13018-13023.

Cases S, Stone SJ, Zhou P, Yen E, Tow B, Lardizabal KD, Voelker T, Farese RV, Jr., 2002. Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. J. Biol. Chem 276(42): 38870-38876.

Cobanoglu O, Berger PJ, Kirkpatrick BW, 2005. Genome screen for twinning rate QTL in four North American Holstein families. Anim Genet 36(4): 303-308.

Coppieters W, Riquet J, Arranz JJ, 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: 540-544.

Damcott, C.M., Moffett, S.P., Feingold, E., Barmada, M.M., Marshall, J.A., Hamman, R.F. & Ferrell, R.E, 2004. Genetic variation in fatty acid-binding protein-4 and peroxisome proliferator-activated receptor gamma interactively influence insulin sensitivity and body composition in males. Metabolism 53:303-309.

De Roos APW, Schrooten C, Mullaart E, Calus MPL, Veerkamp RF, 2007. Breeding value estimation for fat percentage using dense marker on Bos Taurus autosome 14. J. Dairy Sci 90: 4821-4829.

Doerge RW, Zheng ZB, Weir BS, 1997. Statistical issues in the search for genes affecting quantitative traits in experimental populations. Stat Sci 12: 195-219.

Erickson DL, fenster CB, Stenoien HK, Prices D, 2004. quantitative trait locus analyses and the study of evolutionary process. Mol Ecol 13: 2505-2522.

Gonda MG, Arias JA, Shook GE, Kirkpatrick BW, 2004. Identification of an ovulation rate QTL in cattle on BTA14 using selective DNA pooling and interval mapping. Anim Genet 35: 298-304.

Grisart B, Coppieter W, Farnir F, Karim L, Ford C, Berzi P, Cambisano N, Mni M, Reid S, Simon P, Spelman R, Georges M, Snell R, 2001. 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: 222-231.

Guiterez-Gil B, Wiener P, Nute GR, Burton D, Wood JD, Williams JL, 2007. Detection of quantitative trait loci for meat quality traits in cattle. Anim Genet 39: 51-61.

Harder B, Bennewitz J, Reinsch N, Thaller G, Thomsen H, Kuhn C, Schwerin M, Erhardt G, Forster M, Reindhardt F, Kalm E, 2006. Mapping of quantitative trait loci for lactation persistency traits in German Holstein dairy cattle. J. Anim Breed and Genet 123: 89-96.

Heyen DW, Weller JI, Ron M, Band M, Beever JE, Feldmesser E, Da Y, Wiggans GR, VanRaden PM, Lewin HA, 1999. A genome san for QTL influencing milk production and health traits in dairy cattle. Physiol Genomics 1: 165-175.

Kaupe B, Brandt H, Prinzenberg EM, 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.

Kim JJ, Farnir F, Savell J, Taylor JF, 2003. Detection of quantitative trait loci for growth and beef carcass fatness traits in a cross between Bos Taurus (Angus) and Bos indiscus (Brahman) cattle. J. Anim Sci 81: 1933-1942.

Klungland H, Sabry A, Heringstad B, Olsen HG, Gomez-Raya L, Vage DI, Ruane J, Aasland M, Ronningen K, Lien S, 2001. Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. Mamm Genome 12: 837-842.

Kneeland J, Li C, Basarab J, Snelling WM, Benkel B, Murdoch B, Hansen C, Moore SS, 2004. Identification and fine mapping of quantitative trait loci for growth traits on bovine chromosome 2, 6, 14, 19, 21 and 23 within one commercial line of Bos Taurus. J. Anim Sci 82: 3405-3414.

Kuhn C, Thaller G, Winter A, Bininda-Emonds ORP, Kaupe B, Erhardt G, Bennewitz J, Scwerin M, Fries R, 2004. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genet 167: 1873-1881.

Lander ES, Botstein D, 1989. Mapping Mendelian factors underlying quantitative trait using RFLP linkage maps. Genet121: 185-199.

Looft C, Reinsch N, Karal-Albrecht C, Paul S, Brink M, Thomsen H, Brockmann G, Kuhn C, Schwerin M, Kalm E, 2001. A mammary gland EST showing linkage disequilibrium to a milk production QTL on bovine chromosome 14. Mamm Genome 12: 646-650.

Mackay TFC, 2001. the genetic architecture of quantitative traits. Annu Rev Genet 35: 303-339.

MacNeil MD, Grosz MD, 2002. Genome-wide scans for QTL affecting carcass traits in Hereford x composite double backcross populations. J. Anim Sci 80: 2316-2324.

Marques E, Givry SD, Stothard P, Murdoch B, Wang Z, Womack J, Moore SS, 2007. A high-resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. BMS Genomics 8: 254.

Michal JJ, Zhang, ZW, Gaskins CT, Jiang Z, 2006. The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu X Limousine F2 crosses. Anim Genet 37(4): 400-402.

Misozhita K, Watanabe T, Hayashi H, Kubota C, Yamakuchi H, Todoroki J, Sugimoto Y, 2004. Quantitative trait loci analysis for growth and carcass traits in a half-sib family of purebred Japanese Black (Wagyu) cattle. J. Anim Sci 82: 3415-3420.

Miyata M, Gasparin G, Coutinho LL, Martinez ML, Machado MA, Barbosa da Silva MMG, Campos AL, Sonstegard TS, Rosario MFD, Regitano LCDA, 2007. Quantitative trait loci (QTL) mapping for growth traits on bovine chromosome 14. Genet and Mol Biol 30: 364-369.

Moore SS, Li C, Basarab J, Snelling WM, 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: 1919-1925.

Mosiq MO, Lipkin E, Khutoreskaya G, Tchourzyna E, Soller M, Friedmann A, 2001. A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a daughter design, using an adjusted false discovery rate criterion. Genet 157: 1683-1698.

Murani E, Muraniova M, Ponsuksili S, Schellander K, Wimmers K, 2006. Molecular characterization and evidencing of the porcine CRH gene as a functional-positional candidate for growth and body composition. Biochem.Biophys.Res.Commun 342: 394-405.

Nash DL, Rogers GW, Cooper JB, Hargrove GL, Keown JF, Hansen LB, 2000. Heritability of clinical mastitis incidence and relationships with sire transmitting

abilities for somatic cell score, udder type traits, productive life, and protein yield. J. Dairy Sci 83:2350-2360.

Ogino T, Moralejo DH, Kose H, Yamada T, Matsumoto K, 2003. Serum leptin concentration is linked to chromosome 2 and 6 in the OLETF rat, an animal model of type 2 diabetes with mild obesity. Mamm Genome 14: 839-844.

Owen JB, Axford RFE, Bishop SC. Mastitis in dairy cattle, 2000. In: Axford R.F.E., Bishop S.C., Nicholas F.W. & Owen J.B. (eds) Breeding for disease resistance in farm animals. CAB International, Wallingord. 243-52.

Poso J, Mantysaari EA, 1996. Relationships between clinical mastitis, somatic cell score, and production for the first three lactations of finish Ayrshire. J. Dairy Sci 79(7):1284-1291.

Rodriguez-Zas SL, Southey BR, Heyen DW, Lewin HA, 2002. Interval and Composite Interval Mapping of Somatic Cell Score, yield and components of milk in dairy cattle. J. Dairy Sci 85: 3081-3091.

Ron M, Heyen DW, Weller JI, Band M, Feldmesser E, Pasternak H, Da Y, Wiggans GR, Vanraden PM, Ezra E, Lewin HA, 1999. Detection and analysis of a locus affecting milk concentration in a US and Israeli dairy cattle population. Proc 6th World Congr.

Genet. Appl. Livest. Prod., 26, 6th WCGALP congress office, 1998, University of New England, Armidale, Australia, pp. 422-425.

Rupp R, Lagriffoul G, Astruc JM, Barillet F, 2003. Genetic parameters for milk somatic cell scores and relationships with production traits in French Lacaune dairy sheep. J. Dairy Sci 86: 1476-1481.

Schnabel RD, Sonstegard TS, Taylor JF, Ashwell MS, 2005. Whole-genome scans to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families. Anim Genet 36: 408-416.

Smith SJ, Cases S, Jensen DR, Chen HC, Sande E, Tow B, Sanan DA, Raber J, Eckel RH, Farese RF, Jr., 2000. Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. Nat Genet 25: 87-90.

Spelman RJ, Ford CA, McElhinney P, Gregory GC, Snell RG, 2002. Characterization of the DGAT1 gene in the New Zealand dairy population. J. Dairy Sci 85: 3514-3517.

Stenzel-Poore Mp, Cameron VA, Vaughan J, Sawchenko PE, Vale W, 1992. Development of Cushing's syndrome in corticotrophin-releasing factor transgenic mice. Endocrinology 130: 3378-3386.

Stone RT, Keele JW, Shackelford SD, Kappes SM, Koohmaraie M, 1999. A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. J. Anim Sci 77: 1379-1384.

Thaller G, Kramer W, Kaupe B, Erhardt G, Fries R, 2003. Effects of DGAT1 variants on milk production traits in German cattle breeds. J Anim Sci 81: 1911-1918.

Viitala SM, Schulman NF, De Koning DJ, Elo K, Kinos R, Virta A, Virta J, Maki-Tanila A, Vilkki JH, 2003. Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle. J. Dairy Sci 86: 1826-1836.

Washburn SP, Silvia WJ, Brown CH, McDaniel BT, McAllister AJ, 2002. Trends in reproductive performance in Southeastern Holstein and Jersey DHI herds. J. Dairy Sci 85: 244-51.

Wibowo TA, Michal JJ, Jiang Z, 2007. Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. Genome 50: 939-945.

Winter A, Kramer W, Werner FAO, Koller S, Kata S, Durstewitz G, Buitkamp J, Womack JE, Thaller G, Fries R, 2002. Association of 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. USA 99: 9300-9305.

Williams JL, Eggen A, Ferretty L, Farr CJ, Gautier M, Amati G, Ball G, Caramorr T, Critcher R, Costa S, Hextall P, Hills D, Jeulin A, Kiguwa S, Ross O, Smith AL, Saunier K, Urquhart B, Waddington S, 2002. A bovine whole-genome radiation hybrid panel and outline map. Mamm Genome 13: 469-474.

Wood IA, Moser G, Burrell DL, Mengersen KL, Hetzel DJ, 2006. A meta-analytic assessment of a thyroglobulin marker for marbling in beef cattle. Genet Sel Evol. 38:479-494.

Zhang Q, Boichard D, Hoeschele I, Ernst C, Eggen A, Murkve B, Pfister-genskow M, Witte LA, Grignola FE, Uimari P, Thaller G, Bishop MD, 1998. Mapping Quantitative Trait Loci for milk production and health of dairy cattle in a large outbreed pedigree. Genet 149: 1959-1973

Supplemental Table 2.1. QTL reported on BTA14 for fat %, fat yield, milk yield, protein % and protein yield in dairy cattle.

Phenotype	Pval	Markers	Mb	cM
FAT% (or FAT CONTENT)				
Ashwell et al. 1997	0.0181	BM302	33.62	52.37
Ashwell et al. 2002	0.001	BMS1678	9.19	14.01
Ashwell et al. 2002	0.073	BMS1678	9.19	14.01
Ashwell et al. 2002	<0.0001	BMS1678	9.19	14.01
Ashwell et al. 2004	23.1*(Fval)	ILSTS039-BMS1678	1.20-9.19	0-14.01
Bennewitz et al. 2003	<0.01	KIEL_E8	0.26	0
Bennewitz et al. 2003	<0.01	KIEL_E8-RM180	0.26-17.16	0-33.31
Bennewitz et al. 2004	<0.001	DGAT1	0.44	18.70
Boichard et al. 2003	<0.001	CSSM066	3.81	5.13
Boichard et al. 2003	0.0004	CSSM066	3.81	5.13
Boichard et al. 2003	0.11	CSSM066	3.81	5.13
Heyen et al. 1999	0.0023	BM1508	8.27	17.85
Heyen et al. 1999	<0.00001	ILSTS039	1.2	0
Kaupe et al. 2007	<0.001	CYP11B1	1.29	29.80
Kaupe et al. 2007	<0.001	DGAT1	0.44	18.70
Kuhn et al. 2004	<0.0001	CSSM066-ILSTS039	1.20-3.81	0-5.13
Rodriguez-Zas et al. 2002		ILSTS039	1.2	0
Ron et al. 1999	0.0003	CSSM066	3.81	5.13
Thaller et al. 2003	<0.001	CSSM066-ILSTS039	1.20-3.81	0-5.13
Viitala et al. 2003	<0.0029	ILSTS039-BMS1747	1.20-7.87	0-10.50
Zhang et al. 1998		ILSTS011-BM302	11.78-33.62	25.71-52.37
FAT YIELD				
Ashwell et al. 2002	<0.0001	BMS1678	9.19	14.01
Ashwell et al. 2004	12.1*(Fval)	ILSTS039-BMS1678	1.20-9.19	0-14.01
Ashwell et al. 2004	10.5*(Fval)	BMS1941-BM8215	24.67-27.34	41.71-48.23
Bennewitz et al. 2003	<0.01	KIEL_E8-CSSM066	0.26-3.81	5.13
Bennewitz et al. 2003	<0.01	KIEL_E8-CSSM066	0.26-3.81	5.13
Bennewitz et al. 2003	<0.01	KIEL_E8	0.26	0
Bennewitz et al. 2004	<0.01	DGAT1	0.44	18.70
Boichard et al. 2003	0.0011	CSSM066	3.81	5.13
Harder et al. 2006	0.01	BM4513-BL1036	61.48-76.75	79.79-100.16
Heyen et al. 1999	0.0005	CSSM066	3.81	5.13
Heyen et al. 1999	0.00002	ILSTS039	1.2	0
Kaupe et al. 2007	<0.001	CYP11B1	1.29	29.80
Looft et al. 2001	<0.01	ILSTS039-CSSM066	1.20-3.81	0-5.13
Thaller et al. 2003	<0.001	ILSTS039-CSSM066	1.20-3.81	0-5.13
Viitala et al. 2003	0.0398	ILSTS039-BMS1747	1.20-7.87	0-10.50
Winter et al. 1998	<0.0001	ILSTS039-BM1508	1.20-8.27	0-17.85
Zhang et al. 1998	2.25*(Fval)	ILSTS011-BM302	11.78-33.62	25.71-52.37
MILK YIELD				

Ashwell et al. 1997	0.0302	BM302	33.62	52.37
Bagnato et al. 2008	0.0501	CSSM066	3.81	5.13
Bagnato et al. 2008	0.0485	BMS1747	7.87	10.5
Bagnato et al. 2008	0.00148	BMS947	51.274	69.8
Bagnato et al. 2008	0.000311	BL1036	76.75	100
Bennewitz et al. 2003	<0.01	KIEL_E8	0.26	0
Bennewitz et al. 2003	<0.01	KIEL_E8-CSSM066	0.26-3.81	5.13
Boichard et al. 2003	0.02	CSSM066	3.81	5.13
Boichard et al. 2003	0.0002	CSSM066	3.81	5.13
Herder et al. 2006	<0.01	KIEL_EB-CSSM066	0.26-3.81	5.13
Heyen et al. 1999	0.0052	BM4305	65.03	83.31
Kaupe et al. 2007	<0.01	CYP11B1	1.29	29.80
Kaupe et al. 2007	<0.001	DGAT1	0.44	18.70
Looft et al. 2001	<0.01	ILSTS039-CSSM066	1.20-3.81	0-5.13
Rodriguez-Zas et al. 2002		CSSM066	3.81	5.13
Schnabel et al. 2005	16.32*(Fval)	BMC1207-BMS1899	34.16-51.17	51.94-69.01
Thaller et al. 2003	<0.001	CSSM066-ILSTS039	1.20-3.81	0-5.13
PROTEIN % (or PROTEIN CONTENT)				
Ashewell et al. 2004	9.5*(Fval)	BMS1678-ILSTS011	9.19-11.78	14.01-25.71
Bagnato et al. 2008	0.014	ILSTS039	1.20	0
Bagnato et al. 2008	0.0045	CSSM066	3.81	5.13
Bagnato et al. 2008	0.015	DIK2201	6.378	8.1
Bagnato et al. 2008	0.012	BMS2055	74.473	93.7
Bennewitz et al. 2004	<0.001	DGAT1	0.44	18.70
Boichard et al. 2003	<0.001	CSSM066	3.81	5.13
Boichard et al. 2003	<0.00001	CSSM066	3.81	5.13
Heyen et al. 1999	0.0048	ILSTS039	1.2	0
Kaupe et al. 2007	<0.001	CYP11B1	1.29	29.80
Mosiq et al. 2001	<0.01	BM6425	73.84	95.14
Rodriguez-Zas et al. 2002		BM6425	73.84	95.14
Schnabel et al. 2005		BMC1207-BMS1899	34.16-51.17	51.94-69.01
Thaller et al. 2003	<0.001	ILSTS039-CSSM066	1.20-3.81	0-5.13
Viitala et al. 2003	<0.0029	BMS1747-BMS740	7.87-39.55	10.50-60.69
Bennewitz et al. 2003	<0.01	RM180-CSSM066	3.81	5.13-35.31
Bennewitz et al. 2003	<0.01	KIEL_E8	0.26	0
Kaupe et al. 2007	<0.001	DGAT1	0.44	18.70
Kaupe et al. 2007	<0.001	CYP11B1	1.29	29.80
PROTEIN YIELD				
Ashwell et al. 2002	0.0005	BM6425	73.84	95.14
Ashwell et al. 2004	14.7*(Fval)	BM4305-INRA100	65.03-75.88	83.31
Bennewitz et al. 2003	<0.01	KIEL_E8	0.26	0
Kaupe et al. 2007	<0.001	DGAT1	0.44	18.70
Looft et al. 2001	<0.01	ILSTS039-CSSM066	1.20-3.81	0-5.13
Thaller et al. 2003	<0.01	ILSTS039-CSSM066	1.20-3.81	0-5.13

Supplemental Table 2.2. QTL reported on BTA14 for health (clinical mastitis, CM and somatic cell score, SCS), reproduction (calving ease, CE; daughter pregnancy rate, DPR; non return rate, NRR; ovulation rate, OVR; pregnancy rate, PR and twinning) and udder traits (fore udder attachment, FUA; front teat placement, FTP; rear udder width, RUW and udder-SCS) in dairy cattle.

Phenotype	Pval	Markers	Mb	cM
CLINICAL MASTITIS				
Klungland et al. 2001		BM4513-BM6425	61.48-73.84	79.79-95.14
Rupp et al. 2003	<0.01	BM6425	73.84	95.14
Viitala et al. 2003	0.01	BMS1747-BMS740	7.86-39.55	10.50-60.69
REPRODUCTIVE TRAITS				
Kaupe et al. 2007 for CE	<0.05	CYP11B1	1.29	29.80
Schnabel et al. 2005 for DPR		BMC1207-BMS1899	34.16-51.17	51.94-69.01
Kaupe et al. 2007 for NRR	<0.05	DGAT1	0.44	18.70
Gonda et al. 2004 for OVR	0.014	BMS947-BM4305	51.27-65.03	83.31
Ashwell et al. 2004 for PR	0.01	ILSTS011-CSSM066	3.81-11.78	5.13-25.71
Cobanoglu et al. 2005 for twinning	0.001	BMC1207-BM2934	34.16-65.38	51.94
SOMATIC CELL SCORE				
Ashwell et al. 1998	0.0096	BM302	33.62	52.37
Kaupe et al. 2007	<0.001	CYP11B1	1.29	29.80
Kaupe et al. 2007	<0.05	CYP11B1	1.29	29.80
Rodriguez-Zas et al. 2002		BM6425	73.84	95.14
Rupp et al. 2003	<0.10	ILSTS011-BM302	11.78-33.62	25.71-52.37
Zhang et al. 1998		ILSTS011-BM302	11.78-33.62	25.71-52.37
UDDER TRAITS				
Ashwell et al. 1997 for udder	0.0052	BM302	33.62	52.37
Ashwell et al. 1997 for Udder-SCS	0.006	BM302	33.62	52.37
Ashwell et al. 2002 for FUA	0.0703	BM302	33.62	52.37
Ashwell et al. 2002 for FTP	0.04	BM302	33.62	52.37
Schnabel et al. 2005 for RUW	0.01	BM4305-BL1036	65.03-76.75	83.31-100.16

Supplemental Table 2.3. QTL reported on BTA14 for ADG (pre and post weaning), birth weight, carcass weight, fat thickness, marbling score, mean body weight, rib eye area and eating quality (%USDA choice, CHOICE; grilled beef flavor intensity, GBFI; grilled abnormal flavor intensity, GAFI and pH (24 hour), pH) in beef cattle.

Phenotype	Pval	Markers	Mb	cM
Average Daily Gain: PRE-WEANNING				
Kneeland et al. 2004	0.046	BMS1941-BMC1207	24.67-34.16	41.71-51.94
Kneeland et al. 2004	0.041	BM1577-BMS108	41.23-46.69	63.16-67.67
Kneeland et al. 2004	0.023	BMC1207-BM1577	34.16-41.23	51.94-63.16
Mizoshita et al. 2004		BM1508-BMS1941	8.27-24.67	17.85-41.71
Average Daily Gain: POST-WEANNING				
Kneeland et al. 2004	0.026	BMS1747-TG	7.66-7.87	10.50-11.95
Kneeland et al. 2004	0.025	CSSM66-BMS1747	3.81-7.87	5.13-10.50
Kneeland et al. 2004	0.039	BMC1207-BM1577	34.16-41.23	51.94-63.16
Mizoshita et al. 2004		BM8125-ILSTS008	27.34-32.08	50.92-66.48
BIRTH WEIGHT				
Kneeland et al. 2004	0.031	BMS1899-RM137	51.17-67.66	69.01-85.18
Kneeland et al. 2004	0.006	BMS1678-BMS1941	9.19-24.67	14.01-41.71
Kneeland et al. 2004	0.049	BMC1207-BM1577	34.16-41.23	51.94-63.16
CARCASS WEIGHT				
Kim et al. 2003	0.611	RM011-BM4513	27.20-61.48	43.63-79.79
Mizoshita et al. 2004	0.016	BMS1941-INRA094	24.67-28.80	41.71-49.83
Mizoshita et al. 2004		BM8125-ILSTS008	27.34-32.08	50.92-66.48
FAT THICKNESS				
Casas et al. 2000	0.47	RM180-RM011	17.16-27.2	35.31-43.63
Casas et al. 2003	0.24	ILSTS039-DIK5082	1.20-9.86	0-21.30
Moore et al. 2004	0.0058	CSSM066	3.81	5.13
MARBLING SCORE				
Casas et al. 2003	0.45	DIK2008-DIK4087	16.43-68.33	31.26-86.63
MEAN BODY WEIGHT				
Mizoshita et al. 2004		BM8125-ILSTS008	27.34-32.08	50.92-66.48
Mizoshita et al. 2004		MNB-14-BMS1941	16.97-24.67	32.12-41.71
EATING QUALITY				
Casas et al. 2003 for CHOICE	0.26	ILSTS039-DIK4681	1.20-12.58	0-25.71
Guiterez-Zas et al. 2007 for pH	0.0013	RM011-PZ271	27.20-54.15	43.63
Guiterez-Zas et al. 2007 for GBFI	0.0084	BM302	33.62	52.37
Guiterez-Zas et al. 2007 for GAFI	0.0166	BM302	33.62	52.37
GROWTH TRAITS				
Miyata et al. 2004	0.05	CSSM066-ILSTS011	3.81-11.78	5.13-25.71
RIB EYE AREA				
Stone et al. 1999		DIK5377-DIK5082	8.54-9.86	17.85-21.23

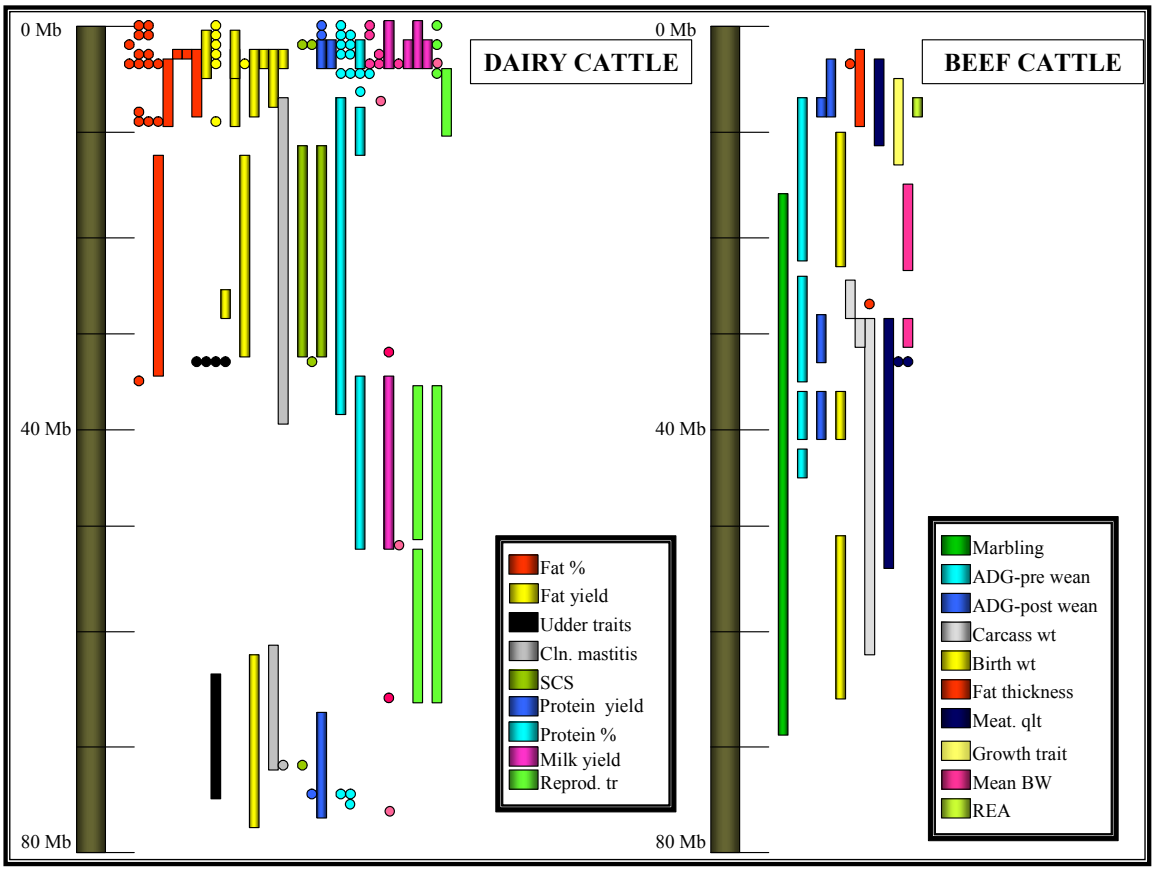

Figure 2.1. Genome assembly anchored QTL map of BTA14 (see Supplemental Tables 2.1-2.3). The gray bar on the left represents the visualization of BTA14 from 0-80 Mega base pairs. Studies on dairy cattle are graphed on the left side with phenotypes represented by different colors, whereas studies in beef cattle are summarized on right side of the graph with phenotypes represented by different colors.

Chapter Three:

COMPARATIVE MAPPING OF BOVINE CHROMOSOME 14 (BTA14) TO HUMAN CHROMOSOME 8 (HSA8) AS MEANS TO UNDERSTANDING THE MAMMALIAN GENOME EVOLUTION AND DISCOVERING PROMISING POSITIONAL CANDIDATE GENES FOR QUANTITATIVE TRAITS

Comparative mapping of bovine chromosome 14 (BTA14) to Human chromosome 8 (HSA8) as means to understanding the mammalian genome evolution and discovering promising positional candidate genes for quantitative traits

Tito A. Wibowo[1], Charles T. Gaskins[1], Ruth C. Newberry[1], Gary H. Thorgaard[2], Jennifer J. Michal[1] and Zhihua Jiang[1]

[1]Department of Animal Sciences, Washington State University, Pullman, WA 99164 – 6351, USA and [2]School of Biological Sciences, Washington State University, Pullman, WA 99164- 4236, USA.

Correspondence to: **Dr. Zhihua Jiang**

Phone: 509 – 335 8761

Fax: 509 – 335 4246

Email: jiangz@wsu.edu

3.1. Abstract

Comparative mapping has been used as an important tool for investigating chromosome conservation or rearrangement events among species, thus providing unique information on candidate genes that underlie quantitative traits in the species of interest. In the present study, we developed a comparative map of bovine chromosome 14 (BTA14) based on its orthologous chromosome, human chromosome 8 (HSA8) using an *in-silico* mapping technique. Among a total of 671 genes annotated in a region of 47.6-146.2 Mb on HSA8, 566 genes/loci were found to have orthologs in the bovine genome. However, only 411 of them were conserved on BTA14, thus leading to development of a comparative map between HSA8 and BTA14 with nine homologous synteny blocks (HSBs) revealed by the AutoGRAPH program. Therefore, our comparative map would provide a solid foundation to increase the efficiency and accuracy of identifying and characterizing promising positional candidate genes for many economically important traits on BTA14.

3.2. Introduction

Following the first six vertebrates, which include human, mouse, rat, zebra fish and two puffer fish (Thomas et al. 2002), the decision to sequence the genome of the next organism was a dilemma. Nevertheless, cattle, which are in a Cetartiodactyla group, are separated from primates 94 million years ago, make its genome more valuable because it can be used for investigating the molecular basis for adaptive evolution (Everts-van der Wind et al. 2004) as well as for annotating the human genome for genes and other non-regulatory elements (Thomas et al. 2002). Moreover, cattle can also serve as a model for many human genetic diseases that cannot be represented by the rodent strains (O'Brien et al. 2001). Most importantly livestock species, such as cattle, have high economic value worldwide (Williams et al 2001; O'Brien et al. 2001). Therefore, in 2003, the U. S. National Institutes for Health (NIH) announced Bos Taurus genome with a priority to be sequenced (Larkin et al 2003).

With the advancement in molecular techniques for development of genetic markers, the cattle genome mapping has progressed immensely. Bos taurus autosome 14 (BTA14), in particular, seemed to be one of the chromosomes in the cattle genome that are highly explored because it harbors many QTL in both dairy and beef cattle (Marques et al. 2007 and Wibowo et al., 2008). However, candidate genes responsible for these QTL are still unknown or inconclusive. The diacylglycerol O-acyltransferase 1 (DGAT1) gene, for instance, was reported as the promising candidate gene for many economically important traits in cattle by Grisart et al. 2002 and Winter et al. 2002, but some associations could not be verified by other studies (for example, Moore et al. 2003). Therefore, comparative

mapping technique can be used to assist the assembly of genes in BTA14 by comparing orthologous genes between cattle genome and a gene-rich genome, such as human genome. According to previous reports, it was known that BTA14 are homologous to the q-arm of HSA8 (Solinas-Toldo et al. 1995, Lyons et al. 1997, Everts-van der Wind et al. 2005, Itoh et al. 2005 and Marques et al. 2007). Therefore, HSA8 can be used as an anchor chromosome to further annotate and assemble BTA14, with the final goal to identify promising positional candidate genes to improve economically important traits.

The importance of comparative mapping technique has long been realized because it facilitates the identification of many positional candidate genes (Everts-van der Wind et al. 2004; Solinas-Toldo et al. 1995) as well as the determination of many important features related to mammalian genome evolution (or karyotype evolution). Candidate genes for milk productions (Grisart et al. 2002) and double muscling (Grobet et al. 1997), for example, were discovered because of the availability of comparative mapping between cattle and human (Everts-van der Wind et al. 2004). Therefore, integrating well-annotated genes from human into the cattle gene map is important for further isolation and characterization of critical candidate genes (Connor et al. 2006; Coppieters et al. 1998; Womack et al. 1995). Furthermore, as far as understanding the mammalian genome evolution, comparative mapping among different species can be used to analyze chromosome rearrangement and conservation by comparing orthologous genes based on their chromosomal locations (Ehrlich et al 1997; Nadeau et al 1998).

3.3. Materials and Methods

3.3.1. Genome resources

The data used in the present study were downloaded from the Genome Resources databases at the National Center for Biotechnology Information (NCBI). The Human Genome Resources are available as the result of the Human Genome Project (HGP), which is funded by the international research community. Originally, the strategy used by the HGP is to minimize the size of the genome to several hundred thousand base pair long and insert them into artificial chromosomes, called the bacterial artificial chromosome (BAC). These sequences were then 'shotgunned' into smaller pieces to ease the sequencing process and aligned using computer algorithm based overlapping fragments to construct the complete sequence. A "working draft" of the human genome was released in 2000. Since then, the NCBI staff has provided various levels of computation, analysis, and curation as needed for the human genome. Basically, the NCBI annotation pipeline annotates the genomic reference sequence data with features such as genes, RNAs, proteins, variation (SNPs), STS markers, and FISH mapped clones. All sequences (genomic, RNAs, proteins) are available for customized BLAST searches, which are displayed on NCBI's Map Viewer.

Aside from completing the human genome sequence, researchers at the Baylor College of Medicine Human Genome Sequencing Center have also been working on the draft for the Bovine Genome Project, together with the Genome British Columbia Sequencing and Mapping Platform at the British Colombia Cancer Agency (BCCA). This project includes BAC libraries (using Hereford bull L1 Domino 99375, with a registration number 41170496, USDA-ARS) and Whole Genome Sequence (WGS) libraries (using a

daughter of La Domino 99375, Hereford L1 Diminette 01449, with a registration number 42190680).

In the present study, we used the current assembly of human genome (build 36.3) and the latest bovine genome (built Btau_4.0) for comparative analysis and characterization of HSA8 and BTA14.

3.3.2. Gene Classifications

Based on the information deposited in the GenBank database (http://www.ncbi.nlm.nih.gov/), all annotated genes on HSA8 can be categorized into five groups: 1) functional genes, 2) human chromosome open reading frame (ORF) genes, 3) hypothetical protein/hypothetical genes, 4) similar to other genes and 5) pseudogenes. Functional genes are those with well-annotated sequences and they have known biological functions. ORF genes are those derived from an annotated genomic sequence using gene prediction method: GNOMON with supports of mRNA and EST evidence. Hypothetical proteins or hypothetical genes are predicted based on computational method with no experimental, *in vivo*, evidence. Similar to other genes are those derived from a computational analysis (GNOMON) based on similarity with other protein. Pseudogene is defined as an inactive gene that is derived from a previously active gene whose sequence does not have exon-intron boundaries.

3.3.3. Genome conservation and rearrangement

Previous studies have shown that a region from 46.1 Mb to 146.2 Mb on HSA8 is orthologous to the entire BTA14. Therefore, HSA8 was used as an anchor chromosome for our comparative mapping of BTA14. All genes annotated on HSA8 (http://www.ncbi.nlm.nih.gov/) were used to determine their orthologs located on BTA14 based on one-one homologous matches. The process was relatively easy to identify the bovine functional genes that have the same gene symbols and high sequence similarity with those on HSA8. For other annotated loci on HSA8, we used the human sequences as references for BLAST searches to retrieve the bovine orthologs on BTA14. The BLAST searches against the bovine genome resources produced different outcomes, including 1) un-annotated ortholog when a BLAST search showed sequence similarity between the human reference and a locus on the bovine genome without the same gene symbol; 2) no ortholog when a BLAST search showed no significant similarity found between the human sequence and the bovine genome; 3) unassigned ortholog when a BLAST search showed sequence similarity between the reference sequence and the bovine genomic DNA, but the locations were still unknown on BTA14 due to the poor development of the current built of the bovine genome and 4) a run-away ortholog when a BLAST search indicated sequence similarity between the human sequence and the bovine genome, but it was found on other chromosome rather than on BTA14.

All orthologous matches were then aligned using AutoGRAPH in order to achieve a visual representation of the comparative map between HSA8 and BTA14 (Derrien et al. 2007). Based on this program, a homologous synteny block (HSB) and internal rearrangement were defined using the most acceptable and common definition described

by Murphy (Murphy et al. 2005), which was also used by Everts-van der Wind et al. 2005 and Marques et al. 2007. HSB is defined as uninterrupted region (by other HSB from the same or different chromosome, for that matter) that consists of at least two or more markers on the same chromosome in the human and cattle genomes. An internal rearrangement, furthermore, is defined by a minimum of three consecutive markers on the cattle HR map in the same order as in the human genome with adjacent markers separated by a span of >1 human-Mb.

3.4. Results

3.4.1. Current genome annotation of HSA8 and BTA14

Based on the current build of human genome (build 36.3), the region of ~100 Mb from 46.1 Mb to 146.2 Mb on HSA8 harbors a total of 671 annotated genes, including 399 functional genes, 26 open reading frame genes, 64 hypothetical protein genes, 33 similar to other genes and 149 pseudogenes. The Bovine Genome Resources (Btau_4.0) indicate that BTA14 has a total of 501 genes annotated at the present, including 317 functional genes, 10 open reading frame genes and 178 hypothetical protein genes. Among these 317 functional genes on BTA14, we found that four genes: *CRH, CA1, CSMD3* and *TRNAM-CAU* were annotated twice. Overall, more work needs to be done on BTA14 for a complete annotation. The distributions of different groups of annotated genes presented on the orthologous region between HSA8 and BTA14 is demonstrated in Figure 3.1.

3.4.2. Identification of human orthologs on BTA14

The numbers of the HSA8 orthologs identified in the bovine genome are summarized according to the five categories of genes in Table 3.1. Among 399 human functional genes investigated in the present study, 377 (94.48%) of them were found to have orthologs in the bovine genome, but with 338 genes (84.71%) placed on BTA14, 10 (2.51%) on other chromosomes and 29 (7.27%) whose locations remain unknown in cattle (Table 3.1). However, the percentage of the human orthologs found to be located on BTA14 decreased to 69.23% (18/26) for ORF genes, 42.42% (14/33) for similar-to-other genes, 25.00% (16/64) for hypothetical protein genes and 16.78% for pseudogenes, respectively (Table 3.1). On the other hand, the number of no-orthologs found increased to 15.38% (4/26) for ORF genes, 20.81% (31/149) for pseudogenes, 33.33 % (11/33) for similar-to-other genes and 56.25% (36/64) for the hypothetical protein genes. The details on one-one orthologous search are listed in Supplemental Table 3.1.

3.4.3. Not annotated genes on BTA14

As indicated in Tables 3.1, we found 43 functional genes, 8 ORF genes, 15 hypothetical protein genes, 14 similar-to-other genes and 25 pseudogenes that were not well annotated on BTA14. For the first two categories of genes, we realized that some of them did not have ESTs to support a complete annotation. In particular, seven functional genes: *CNBD1, CPA6, MIRN124-2, MIRN599, PSCA, PSKH2* and *SLCO5A1*, and one ORF *C8orf39* have no ESTs developed so far in cattle. However, for the last three categories of genes that were not well annotated in the bovine genome, we found that only nine hypothetical genes: *FLJ46284, KIA1833, LOC100128550, LOC100128687, LOC100129242, LOC100129885, LOC100130155, LOC644727* and, *MGC39715*; seven

similar-to-other genes: *LOC1001218414, LOC1001218419, LOC100129173, LOC100129963, LOC130742, LOC130862, LOC286187* and *LOC728967*; and eleven pseudogenes: *LOC100127983, LOC100128271, LOC100128418, LOC100128955, LOC100129377, LOC100130095, LOC100130861, LOC100131371, LOC100132280, LOC392225, LOC643228,* and *LOC644199* were conserved with their neighboring genes on BTA14 as they were seen on HSA8.

3.4.4. Run-away orthologs

For nine functional genes: (*COX6C, LOC389672, LOC728638, RPL7, SEC11B, SLC39A4, WDR21C, ZNF251*, and *ZNF517*) and 2 ORF genes (*C8orf62* and *C8orf70*), the current assembly of bovine genome showed that they were no longer conserved on BTA14. Instead, they were located on other chromosomes. Based on the sequence similarity and evidence with ESTs support, we are confident that these genes are orthologs of human genes. However, the other 6 hypothetical protein genes, 6 similar-to-other genes and 82 pseudogenes that were placed on other chromosomes in the cattle genome, were rather vague to determine at the moment if they were really orthologous to the human genes annotated on HSA8. Basically such an ortholog cannot be only judged by the sequence similarity.

3.4.5. Comparative Mapping

All orthologs, including those categorized as annotated orthologs and those as unannotated orthologs that were conserved between HSA8 and BTA14, were used to define the conservation segments between the two these chromosomes. Based on the

result from the AutoGRAPH program, there were nine conserved segments or HSBs (Figure 3.2; table 3.2) with different gene density per segment. HSBs with the highest gene density were observed in the region of: 87.595-120.497 and 48.812-82.553 Mb on HSA8with 127 and 115 genes and loci per conserved segment, respectively. The lowest gene density, on the other hand, was observed in the region of: 145.123-145.136 and 120.638-121.619 Mb with only 4 and 8 genes and loci per HSB, respectively. Finally, the other five regions on HSA8: 82.732-87.180, 122.694-142.435, 143.736-144.522, 144.590-144.987 and 145.1778-146.075 Mb had variable gene densities: 18, 62, 17, 16 and 44 genes and loci per HSB, respectively.

3.5. Discussion

To further confirm the comparative map built previously by Lyons et al. 1997, Everts-van der Wind et al. 2005, Itoh et al. 2005, and Marques et al. 2007, our result of *in-silico* comparative map also showed a high degree of homology between BTA14 and HSA8. The highest degree of homologous was observed on genes and loci on HSA8 that were categorized as functional genes and ORFs with 84.71% and 69.23% of conservation, respectively. Nevertheless, although 295 genes and 10 ORFs were conserved between the two chromosomes, 43 functional genes and eight ORFs were not well annotated on BTA14. Therefore, to further annotate these genes and loci on the bovine genome, we decided to find an early evidence of the existence of these genes and loci on the bovine genome. The reference human sequence of these genes and loci were BLAST-ed against the ESTs sequence on the bovine genome. We discovered that 36 functional genes and seven ORFs had ESTs sequence in the cattle genome. Hence, HSA8 can be used here as the anchor chromosome to place these genes and loci on the cattle genome. Nevertheless,

our comparative map was built based on the availability of the data on the NCBI database. Therefore, we did not exclude the possibility that the other seven functional genes and one other ORF, whose ESTs sequences could not be found in the cattle genome, had human orthologs genes but did not have the data just yet by the time this project was accomplished.

Interestingly, aside from those genes and loci that were mapped on BTA14, nine functional genes and two ORFs were placed outside the conserved chromosome with ESTs evidence to support the same human orthologs genes. Eleven of these genes and ORFs were not well annotated yet in the bovine genome, but three genes: *COX6C*, *RPL7* and *SLC39A4* were well annotated but still placed outside BTA14, at BTA 9, X, and 17, respectively. One possible reason of this run-away orthologous genes were found in different chromosome was that these genes had evolved to be pseudogenes in the cattle genome. Hence, it was not surprising that *COX6C* and *RPL7* had many copies of pseudogenes on the cattle genome. Even though Btau_4.0 placed the structural gene for *COX6C* gene on BTA9, this gene also had one pseudogene that was placed on BTA14 with conserved neighboring genes as seen in HSA8. This gene was bracketed by *RGS22* and *VPS13B*, which were about ~ 82.9 Kb downstream and ~864.8 Kb upstream on HSA8, respectively, and about ~95.8 Kb downstream and ~ 959.0 Kb upstream on BTA14, respectively. Furthermore, *RPL7* was placed on BTA X on Btau_4.0 with only one exon, whereas the ortholog human sequence had six exons on the human genome, which indicated that this gene might act only as pseudogene in the bovine genome. The result of BLAST search engine, furthermore, showed more than 30 homologous regions

of this gene on the bovine genome with a relatively high conservation between the ortholog human sequences. Fifteen of these homologous regions had more than 90% sequence similarities. In addition, homologous sequence on BTA 2, 7 and 8 showed that *RPL7* pseudogene evolved to have two exons that were separated by an intron in the cattle genome. Nonetheless, having many processed pseudogenes in the genome is a typical character of the encoding ribosomal protein, such as *RPL7*. However, this example can be used to study the mechanism of how pseudogene evolves since the functional gene for *RPL7* has six exons, instead of only two. Finally, *SLC39A4* was placed on BTA17, without any pseudogene. This suggested that, with this early evidence, this gene was functional gene that was not conserved on BTA14. Yet, since the bovine genome was still not well annotated further research need to confirm the actual location of *SLC39A4* in the cattle genome.

Twenty three functional genes and four ORFs from HSA8 did not have neither genomic nor ESTs orthologs in the bovine genome. Cattle and man were separated 94 million years ago (Everts-van der Wind et al. 2004) from the common ancestor. Therefore, it was tempting to conclude that the absence of these functional genes or ORFs in the cattle genome was caused by the difference adaptive evolution between the two mammals. Nevertheless, our comparative map was limited to the availability of the data. Hence, again, these absences of ESTs sequence on the cattle genome could be the result of lacking of resources.

The other categories of genes and loci on HSA8: hypothetical protein genes, similar-to-other and pseudogene were less conserved on BTA14 with only 25.00%, 42.42% and 16.78% of conservations, respectively. Only one locus *C8ORFK36* on hypothetical protein gene category was well annotated on BTA14 with EST evidence found on the cattle genome. The other 15 hypothetical protein gene loci, 14 similar-to-other loci, and pseudogene loci that were mapped to BTA14 were not well annotated. Since these genes or loci categories might not have ESTs orthologs in the cattle genome, we used different approach to annotate these loci in the genome. Comparison of conserved neighboring genes on HSA8 was used to determine the most accurate location to place these loci on BTA14. Although not all loci could be placed, at least nine hypothetical protein gene loci, seven similar-to-other loci and eleven pseudogene loci had conserved neighboring genes between their locations on HSA8 and the current placements on BTA14. This could suggest that these loci, which were not well annotated yet, could possibly be confirmed on BTA14 by using HSA8 as the anchor chromosome.

From all of these five categories of the genes annotated on HSA8, nevertheless, only those that were orthologous to BTA14, including both annotated orthologous genes and un-annotated orthologous genes, were used to determine the HSBs between the two chromosomes using the AutoGRAPH program. Discrepancies, however, occurred when comparing our comparative map with other previously built map established with different techniques. Everts-van der Wind et al. (1995), for instance, reported three synteni blocks: 48.9-83.4, 120.3-82.6 and 146.2-124.1 Mbp on HSA8 using 5000 rad-RH panel, and then confirmed by Itoh et al. 2005 using 7000 rad- RH panel. On the other

hand, Lyons et al. 1997 and Marques et al. 2007, using CATS analysis and 12k rad- RH panel, respectively, discovered a new evolutionary break point region, at position: 121.0-122.0 Mbp on HSA8 (Marques et al. 2007) that had an opposite order from one another. Here, based on the current bovine genome (Btau_4.0) and human genome (build 36.3), we reported nine HSBs. Although it appears that we report more HSBs than the previous reports, our result is actually comparable with the previously reported comparative map between BTA14 and HSA8. HSB in the region of 48.9-83.4 and 121.0-122.0 Mb from HSA8, for instance, were in great consensus previous reports. Nevertheless, in other regions, we found more internal rearrangements within the HSB that caused them to be recognized as a separate HSB by the AutoGRAPH program. A region at 120.3-82.6 Mb, for example, was similar between our report and other study. Our result, however, presented this region in two separate HSBs, 82.732-87.180 and 87.595-120.497 Mb, with the same order of the loci as other reports. Unfortunately, we also observed some regions with high discrepancies, such as in the region of 146.2-124.1Mb. Instead of identifying this region with only one HSB, our result showed five different HSBs covering the region. One region in particular, 143.7-144.5 Mb, had a reversed order than other maps. Thus, although it was possible that other maps built with linkage or radiation hybrid (RH) technique flawed, it was more likely that the lack of well annotation of the current bovine genome assembly that needed to be improved.

3.6. *Reference*

Band MR, Larson JH, Rebeiz M, Green CA, Heyen DW, Donovan J, Windish R, Steining C, Mahyuddin P, Womack JE, Lewin HA, 2000. Genome Res 10: 1359-1368.

Connor EE, Ashwell MS, Schnabel R, Williams JL, 2006. Comparative mapping of bovine chromosome 27 with human chromosome 8 near a dairy form QTL in cattle. Cytogenet Genome Res 112: 98-102.

Coppieters W, Riquet J, Arranz JJ, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Nezer C, Simon P, Vanmanshoven P, Wagenaar D, Georges M, 1997. A QTL with major effect on milk yield and composition maps to bovine chromosome 14. Mamm Genome 9: 540-544.

Ehrlich J, Sankoff D, Nadeau H, 1997. Synteny conservation and chromosome rearrangements during mammalian evolution. Genet 147: 289-296.

Everts-van der Wind A, Kata SR, Band MR, Rebeiz M, Larkin DM, Everts RE, Green CA, Liu L, Natarajan S, Goldammer T, Lee JH, McKay S, Womack JE, Lewin HA, 2004. Genome Res 14: 1424-1437.

Everts-van der Wind A, Larkin DM, Green CA, Elliott JS, Olmstead CA, Chiu R, Schein JE, Marra MA, Womack JE, Lewin HA, 2005. A high-resolution whole-genome

cattle-human comparative map reveals details of mammalian chromosome evolution. Proc. Natl. Acad. Sci. 102 (51): 18526-18531.

Grisart B, Coppieters W, Farnir F, Karim L, Ford C, Berzi P, Cambisano N, Mui 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: 222-231.

Grobet L, Martin LJ, Poncelet D, Pirottin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Menissier F, Massabanda J, Fries R, Hanset R, Geroges M, 1997. A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. Nat Genet 17: 71-74.

Itoh T, Watanabe T, Ihara N, Mariani P, Beattie, CW, Sugimoto Y, Takasuga A, 2005. A comprehensive radiation hybrid map of the bovine genome comprising 5593 loci. Genomics 85: 413-424.

Larkin DM, Everts-van der Wind A, Rebeiz M, Schweitzer PA, Bachman S, Green C, Wright CL, Campos EJ, Benson LD, Edwards J, Liu L, Osoegawa K, Womack JE, de Jong PJ, Lewin HA, 2003. Genome Res 13: 1966-1972.

Lyons LA, Laughlin TF, Copeland NG, Jenkins NA, Womack JE, O'Brien SJ, 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nat Genet 15: 47-56.

Marques E, Givry SD, Stothard P, Murdoch B, Wang Z, Womack J, Moore SS, 2007. A high-resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. BMS Genomics 8: 254.

Marra MA, Hillier L, Waterson RH, 1998. Expressed sequence tags-ESTablishing bridges between genomes. Trends Genets 14 (1): 4-7.

Moore SS, Li C, Basarab J, Snelling WM, 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: 1919-1925.

Nadeau JH, sankoff D, 1998. Counting on comparative maps. Trends Genets 14 (12): 495-501.

O'Brien SJ, Eizirik E, Murphy WJ, 2001. On choosing mammalian genomes for sequencing. Science 292: 2264-2266.

Solinas-Toldo S, Lengauer C, Fries R., 1995. Comparative genome map of human and cattle. Genomics 27: 489-496.

Thomas JW, Touchman JW, 2002. Vertebrate genome sequencing: building a backbone for comparative genomics. Trends Genets 18(2): 104-108.

Williams JL, Eggen A, Ferretty L, Farr CJ, Gautier M, Amati G, Ball G, Caramorr T, Critcher R, Costa S, Hextall P, Hills D, Jeulin A, Kiguwa S, Ross O, Smith AL, Saunier K, Urquhart B, Waddington S, 2002. A bovine whole-genome radiation hybrid panel and outline map. Mamm Genome 13: 469-474.

Winter A., Kramer W, Werner FAO, Koller S, Kata S, Durstewitz G, Buitkamp J, Womack JE, Thaller G, Fries R, 2002. Association of 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. USA 99: 9300-9305.

Womack JE, Kata SR. 1995. Bovine genome mapping: evolutionary inference and the power of comparative genomics. Current Opinion in Genetics and Development 5(5): 725-733

Table 3.1. Summary of *in-silico* comparative chromosome based on annotated genes on HSA8

	HUMAN	CATTLE				
	HSA8	BTA 14			Other	Unassigned
		Orthologous		No		
Description		Annotated	Not Annotated	Orthologous		
Functional gene	399	295	43	23	9	29
ORF	26	10	8	4	2	2
Hypothetical protein	64	1	15	36	6	6
Similar to other	33	0	14	11	6	7
Pseudogene	149	0	25	31	82	11
Total	671	306	105	105	105	50

Table 3.2. The comparison of HSBs between HSA8 (build 36.3) and BTA14 (Btau_4.) based on AutoGRAPH program and previous reports.

HSA8	BTA14		Other reports			
(Mbp)	(Mbp)	ORDER	Everts van der Wind et al. 1995; Itoh et al. 2005		Marques et al. 2007; Lyons et al. 1997	
48.812-82.553	18.785-41.955	↑	48.812-82.553	↑	48.812-82.553	↑
82.732-87.180	70.902-75.435	↓	82.6-120.3	↓	82.6-120.3	↓
87.595-120.497	75.081-42.130	↓		↓		↓
120.638-121.619	79.985-80.962	↑		↓	121.0-122.0	↑
122.694-142.435	17.680-1.888	↓	124.1-146.2	↓	124.1-146.2	↓
143.735-144.522	1.090-1.543	↑		↓		↓
144.590-144.987	0.965-0.775	↓		↓		↓
145.123-145.136	1.065-1.059	↓		↓		↓
145.178-146.075	0.599-0.013	↓		↓		↓

Supplemental Table 3.1. Orthologous analysis of functional genes between the synteny region of HSA8 and the bovine genome by targeting BTA14

	HOMO SAPIENS		BOS TAURUS		
Symbol	**Start**	**Stop**	**Start**	**Stop**	**Description**
ANNOTATED ORTHOLOGOUS GENES ON BTA14					
ABRA	107842321	107851649	55610645	55621461	actin-binding Rho activating protein
ADCK5	145568539	145589265	404677	418268	aarF domain containing kinase 5
ADCY8	131861729	132123854	9201726	9426613	adenylate cyclase 8 (brain)
ADHFE1	67507272	67543598	30582423	30615404	alcohol dehydrogenase, iron cont 1
ANGPT1	108330886	108579430	54783106	55084883	angiopoietin 1
ANKRD46	101602176	101641188	61841203	61882331	ankyrin repeat domain 46
ANXA13	124762215	124818828	15634803	15650685	annexin A13
ARFGEF1	68272451	68418466	31219174	31406114	ADP
ARMC1	66677618	66708986	29948961	29973726	armadillo repeat containing 1
ASPH	62578374	62789560	26877708	27065682	aspartate beta-hydroxylase
ATAD2	124401922	124477868	16007999	16033231	ATPase family, AAA domain cont 2
ATP6V0D2	87180255	87235570	75435337	75490139	ATPase, H+ transporting, lysosomal
ATP6V1C1	104102424	104154461	59589168	59629227	ATPase, H+ transporting, lysosomal
ATP6V1H	54790668	54918403	21706742	21754129	ATPase, H+ transporting, lysosomal
AZIN1	103907712	103945573	59765276	59795942	antizyme inhibitor 1
BAALC	104222097	104311709	59443816	59530484	brain and acute leukemia, cytoplasmic
BOP1	145456864	145485928	467652	486604	block of proliferation 1
CA1	86427709	86477594	76046197	76056557	carbonic anhydrase I
CA13	86345259	86383554	76145079	76169785	carbonic anhydrase XIII
CA2	86563383	86580973	75898446	75914801	carbonic anhydrase II
CA3	86537710	86548526	75932466	75942462	carbonic anhydrase III, muscle specific
CA8	61263977	61356508	25769686	25851726	carbonic anhydrase VIII
CALB1	91140014	91164283	72314241	72336158	calbindin 1, 28kDa
CCNE2	95961628	95976660	67782182	67794832	cyclin E2
CDH17	95208566	95289986	68428155	68599288	cadherin 17, LI cadherin (liver-intestine)
CEBPD	48812029	48813279	18785576	18787169	CCAAT/enhancer binding protein
CHCHD7	57286869	57293732	23265198	23271074	coiled-coil-helix-coiled-coil-helix
CHD7	61753893	61942019	26137555	26327347	chromodomain helicase DNA binding
CHMP4C	82807243	82834305	79826972	79856186	chromatin modifying protein 4C
CHRAC1	141590586	141596434	2424520	2427765	chromatin accessibility complex 1
CNGB3	87655277	87825017	74895793	75066368	cyclic nucleotide gated channel beta 3
COL22A1	139669660	139995418	3475101	3714674	collagen, type XXII, alpha 1
COLEC10	120148627	120188388	42390816	42436251	collectin sub-family member 10
COMMD5	146046355	146049736	118906	120941	COMM domain containing 5
COPS5	68117868	68137116	31084644	31107073	COP9 constitutive photomorphogenic
CPNE3	87595788	87642842	75081894	75138489	copine III
CPSF1	145589254	145605541	396499	404637	cleavage and polyadenylation specific fct
CRH	67251173	67253252	30476483	30478099	corticotropin releasing hormone
CSMD3	113304333	114518418	48322004	48323228	CUB and Sushi multiple domains 3

CSPP1	68139157	68271052	31119055	31218158	centrosome and spindle pole
CTHRC1	104452962	104464393	59271909	59288176	collagen triple helix repeat containing 1
CYC1	145221948	145224416	572202	574603	cytochrome c-1
CYHR1	145660016	145661657	249944	263515	cysteine/histidine-rich 1
CYP11B1	143950775	143958238	1294946	1302624	cytochromeP450, fam11, subfamB, pol 1
DDEF1	131133535	131483399	9726354	10065407	differentiation enhancing factor 1
DECR1	91082756	91133403	72348154	72391666	2,4-dienoyl CoA reductase 1, mitoch
DENND3	142207902	142275083	1997561	2047080	DENN/MADD domain containing 3
DEPDC2	69027157	69306451	32988865	33279685	DEP domain containing 2
DEPDC6	120955146	121131939	80376637	80492328	DEP domain containing 6
DERL1	124094769	124123722	16269050	16293939	Der1-like domain family, member 1
DGAT1	145510762	145521375	444097	446810	diacylglycerol O-acyltransferase
DNAJC5B	67096345	67175309	30300948	30393961	DnaJ (Hsp40) homolog, subfamily C
DPY19L4	95801327	95873245	67860002	67921974	dpy-19-like 4 (C. elegans)
DPYS	105460829	105548453	58209503	58299412	Dihydropyrimidinase
E2F5	86276871	86314005	76202989	76231677	E2F transcription factor 5, p130-binding
EBAG9	110621105	110646568	52646250	52653220	estrogen receptor binding site associated
EEF1D	144733041	144750726	925709	938385	eukaryotic translation elongation factor 1
EFCAB1	49798520	49810344	19473254	19481501	EF-hand calcium binding domain 1
EFR3A	132985541	133094956	8559689	8601402	EFR3 homolog A (S. cerevisiae)
EIF2C2	141610446	141714827	2373025	2409522	eukaryotic translation initiation factor
EIF3E	109283148	109330135	54054335	54102285	eukaryotic translation initiation factor
EIF3H	117726236	117837243	44953973	45051260	eukaryotic translation initiation factor
ENPP2	120638500	120720287	79985613	80114337	ectonucleotide
ENY2	110415812	110425087	52877993	52887448	enhancer of yellow 2 homolog
EXOSC4	145205510	145207539	589223	591098	exosome component 4
EXT1	118880783	119193239	43445832	43759712	exostoses (multiple) 1
EYA1	72272222	72437021	35074955	35233610	eyes absent homolog 1 (Drosophila)
FABP4	82553481	82558004	41955210	41959600	fatty acid binding protein 4, adipocyte
FABP5	82355340	82359563	41734584	41741360	fatty acid binding protein 5
FABP9	82533173	82536313	41939300	41942980	fatty acid binding protein 9, testis
FAM110B	59069667	59224831	24168669	24235498	family with sequence similarity 110
FAM49B	130922898	131021182	10132033	10233191	family with sequence similarity 49
FAM84B	127633869	127639648	13381539	13386603	family with sequence similarity 84
FAM92A1	94781949	94809850	68960049	68985320	family with sequence similarity 92
FBXL6	145549899	145552940	428720	431721	F-box and leucine-rich repeat protein 6
FBXO32	124584539	124622627	15820177	15852302	F-box protein 32
FBXO43	101214835	101227252	62314040	62336121	F-box protein 43
FER1L6	124933408	125201483	15353599	15527783	fer-1-like 6 (C. elegans)
FOXH1	145670306	145672526	239928	242419	forkhead box H1
FZD6	104380276	104414270	59325628	59391516	frizzled homolog 6 (Drosophila)
GDF6	97223734	97242196	66262264	66279562	growth differentiation factor 6
GEM	95330657	95343733	68377777	68390687	GTP binding protein expressed in musc
GGH	64090192	64113940	27930715	27953360	gamma-glutamyl hydrolase
GML	143913219	143925264	1258011	1286241	GPI anchored molecule like protein
GOLSYN	110655581	110773196	52633928	52640360	Golgi-localized protein
GPAA1	145209512	145213107	584696	587934	glycosylphosphatidylinositol anchor attc.
GPR172A	145553033	145555753	425960	428588	G protein-coupled receptor 172A

GPR20	142435767	142446547	1888054	1889139	G protein-coupled receptor 20
GPT	145700231	145703357	209123	212128	glutamic-pyruvate transaminase
GRHL2	102574162	102750995	60933845	61055008	grainyhead-like 2 (Drosophila)
GRINA	145136214	145139571	1059703	1060406	glutamate receptor, ionotropic,
GSDMDC1	144711635	144716375	952755	959272	gasdermin domain containing 1
HAS2	122694719	122722811	17680016	17696287	hyaluronan synthase 2
HEY1	80838800	80842653	40040138	40043034	enhancer-of-split related with YRPW
HNF4G	76614758	76641624	35606077	35635567	hepatocyte nuclear factor 4, gamma
HRSP12	99183743	99198594	64295345	64303544	heat-responsive protein 12
HSF1	145486078	145509193	448053	453004	heat shock transcription factor 1
IL7	79807560	79880313	38974841	39051983	interleukin 7
IMPA1	82732751	82761115	79902881	79925291	inositol(myo)-1(or 4)-monophosphatase
INTS8	95904710	95961897	67794760	67839649	integrator complex subunit 8
JRK	143735876	143748403	1090119	1091705	jerky homolog (mouse)
KCNK9	140693986	140784481	2992665	2993414	potassium channel, subfamily K
KCNQ3	133210438	133562186	8183590	8475094	potassium voltage-gated channel
KCNS2	99508426	99512201	64001954	64004712	potassium voltage-gated channel
KCNV1	111048409	111056135	52285517	52291616	potassium channel, subfamily V
KHDRBS3	136538898	136729031	5764809	5908681	KH domain containing, RNA binding
KIAA1429	95569776	95634864	68091769	68155921	KIAA1429
KLF10	103730188	103737128	59950195	59955729	Kruppel-like factor 10
LACTB2	71712045	71743946	34604796	34630808	lactamase, beta 2
LAPTM4B	98856985	98934006	64553152	64615031	lysosomal associated protein
FAM91A1	124850063	124896873	15565881	15603899	family with sequence similarity 91, A1
KIAA0146	48336095	48811028	18787723	19031421	KIAA0146
FAM83A	124263933	124291499	16144075	16162914	family with sequence similarity 83, A
BHLHB5	65655368	65658740	29204116	29206913	basic helix-loop-helix domain
KIAA1688	145725371	145809699	149247	185759	KIAA1688 protein
SULF1	70541427	70735701	33566339	33660746	sulfatase 1
KIAA1875	145234623	145245206	558672	559703	KIAA1875
LRP12	105570640	105670344	58095917	58186678	low density lipoprotein-related protein
LRRC14	145714199	145721365	195643	199106	leucine rich repeat containing 14
LRRC24	145718580	145723216	189486	194858	leucine rich repeat containing 24
LRRC6	133653629	133756995	8049136	8120139	leucine rich repeat containing 6
LRRCC1	86206629	86245567	76239546	76283565	leucine rich repeat and coiled-coil dom
LY6D	143863300	143865010	1155359	1156829	lymphocyte antigen 6 complex, locus D
LY6E	144171300	144175199	1385601	1389305	lymphocyte antigen 6 complex, locus E
LY6H	144310706	144313116	1448989	1451445	lymphocyte antigen 6 complex, locus H
LYN	56954926	57085685	23085065	23133147	v-yes-1 Yamaguchi sarcoma viral
LYNX1	143842758	143856642	1141694	1143793	Ly6/neurotoxin 1
LYPD2	143828630	143830954	1127249	1129299	LY6/PLAUR domain containing 2
LYPLA1	55121491	55177130	21846406	21864400	lysophospholipase I
MAF1	145231293	145234503	563802	566837	MAF1 homolog (S. cerevisiae)
MAL2	120289791	120327094	42269086	42302331	mal, T-cell differentiation protein 2
MAPK15	144870495	144876621	852901	858569	mitogen-activated protein kinase 15
MATN2	98950487	99118124	64360737	64536349	matrilin 2
MCM4	49036047	49052621	19210501	19222447	minichromosome maintenance complex
MED30	118602211	118621682	44067815	44086895	mediator complex subunit 30

MFSD3	145705360	145707397	205787	207864	major facilitator superfamily domain
MIRN151	141811844	141811933	2275084	2275152	microRNA 151
MIRN30B	135881944	135882031	6477865	6477952	microRNA 30b
MIRN30D	135886300	135886369	6473521	6473590	microRNA 30d
MLZE	130829624	130868316	10253498	10268900	melanoma-derived leucine zipper
MMP16	89118576	89408833	73040445	73436731	matrix metallopeptidase 16
MOS	57188055	57189095	23188216	23189525	v-mos Moloney murine sarcoma viral
MRPL13	121477264	121526828	80835927	80876848	mitochondrial ribosomal protein L13
MRPL15	55210334	55223014	21870983	21876539	mitochondrial ribosomal protein L15
MRPS28	80993650	81105061	40190694	40292939	mitochondrial ribosomal protein S28
MTBP	121526847	121605056	80877191	80953529	Mdm2, transformed 3T3 cell
MTDH	98725583	98807714	64683940	64738109	metadherin
MTERFD1	97320821	97342972	66167788	66185533	MTERF domain containing 1
MTFR1	66719528	66785340	29988540	30047774	mitochondrial fission regulator 1
MTSS1	125632206	125809911	14853262	15012900	metastasis suppressor 1
MYBL1	67636968	67687729	30719638	30753289	v-myb myeloblastosis viral oncogene
MYC	128817498	128822856	12151194	12156367	v-myc myelocytomatosis viral oncogene
NAPRT1	144728098	144731636	939528	942677	nicotinate phosphoribosyltransferase
NBN	91014740	91066075	72408561	72464234	nibrin
NCALD	102767946	103206311	60872727	60873116	neurocalcin delta
NCOA2	71186821	71478574	34114745	34291940	nuclear receptor coactivator 2
NDRG1	134318596	134378680	7510051	7566221	N-myc downstream regulated gene 1
NDUFB9	125620524	125631408	15012769	15020175	NADH dehydrogenase 1 beta subc
NFKBIL2	145624984	145640589	266557	277562	nuc. factor of kappa light polypept. gene
NIBP	140811770	141537860	2471901	2859130	NIK and IKK{beta} binding protein
NKAIN3	63324055	64066182	27559142	27876365	Na+/K+ transporting ATPase
NOV	120497882	120505776	42130892	42139270	nephroblastoma overexpressed gene
NPAL2	99273563	99375797	64134342	64225549	NIPA-like domain containing 2
NRBP2	144987904	144996188	775439	780963	nuclear receptor binding protein 2
NSMAF	59658617	59734940	24580454	24633790	neutral sphingomyelinase activation
NSMCE2	126173277	126448544	14310926	14548932	non-SMC element 2, MMS21 homolog
NUDCD1	110322324	110415491	52887684	52980985	NudC domain containing 1
OC90	133105910	133131298	8496373	8556541	otoconin 90
ODF1	103633036	103642422	60023783	60033244	outer dense fiber of sperm tails 1
OPLAH	145178155	145186959	599578	608671	5-oxoprolinase (ATP-hydrolysing)
OPRK1	54300829	54326747	21567240	21588999	opioid receptor, kappa 1
OSGIN2	90983269	91009271	72473994	72538707	oxidative stress induced growth inhibitor
OSR2	100025862	100033504	63509619	63517269	odd-skipped related 2 (Drosophila)
OTUD6B	92151600	92168499	71902397	71997879	OTU domain containing 6B
OXR1	107739270	107834097	55660321	55683533	oxidation resistance 1
PAG1	82042600	82186858	41424122	41441978	phosphoprotein associated with
PARP10	145123308	145132623	1065748	1072558	poly (ADP-ribose) polymerase family
PCMTD1	52892700	52936350	20767384	20815604	protein-L-isoaspartate O-methtransferase
PENK	57516070	57521143	23430381	23434397	Proenkephalin
PGCP	97726675	98224900	65173815	65734542	plasma glutamate carboxypeptidase
PHF20L1	133856786	133930234	7943651	7990482	PHD finger protein 20-like 1
PKHD1L1	110443882	110612676	52759141	52852787	polycystic kidney and hepatic disease 1
PKIA	79590891	79678040	38758243	38856050	protein kinase inhibitor alpha
PLAG1	57236037	57286392	23219718	23221723	pleiomorphic adenoma gene 1

PLEKHF2	96215208	96238089	67508106	67528592	pleckstrin homology domain containing
PMP2	82515116	82522274	41923319	41928456	peripheral myelin protein 2
POLR2K	101232015	101235406	62313104	62317319	polymerase (RNA) II (DNA directed)
POP1	99199244	99239816	64260129	64288971	processing of precursor 1
PPM2C	94998338	95007472	68769096	68777486	protein phosphatase 2C
PPP1R16A	145692917	145698312	214035	219209	protein phosphatase 1
PRKDC	48848222	49035296	19081946	19209665	protein kinase, DNA-activated
PTDSS1	97343343	97415950	66097468	66164491	phosphatidylserine synthase 1
PTK2	141737683	142080514	2116865	2309780	PTK2 protein tyrosine kinase 2
PUF60	144970535	144983525	786305	798546	poly-U binding splicing factor 60KDa
PXMP3	78055049	78075079	37070697	37085976	peroxisomal membrane protein 3, 35kDa
PYCRL	144757532	144762887	913501	920774	pyrroline-5-carboxylate reductase-like
RAB2A	61592113	61696183	26047338	26092362	RAB2A, member RAS oncogene family
RAD21	117927355	117956182	44689347	44718908	RAD21 homolog (S. pombe)
RAD54B	95453364	95556486	68168222	68286312	RAD54 homolog B (S. cerevisiae)
RALYL	85258008	85996634	76464451	76648147	RALY RNA binding protein-like
RB1CC1	53697571	53789579	21285639	21314094	RB1-inducible coiled-coil 1
RBM35A	95722576	95788870	67937077	67994462	RNA binding motif protein 35A
RDHE2	57375122	57395795	23365427	23398159	epidermal retinal dehydrogenase 2
RECQL4	145707479	145714008	199408	205730	RecQ protein-like 4
RGS20	54926921	55034416	21779728	21808549	regulator of G-protein signaling 20
RGS22	101042452	101187520	62374124	62495106	regulator of G-protein signaling 22
RHPN1	144522400	144537533	1543285	1552286	rhophilin, Rho GTPase binding protein 1
RIMS2	104582152	105334627	58418881	58449007	regulating synaptic membrane
RIPK2	90839110	90872433	72556691	72595294	receptor-interacting serine-threonine
RLBP1L1	62363104	62576758	26716330	26874238	retinaldehyde binding protein 1-like 1
RNF139	125556189	125570040	15052229	15061031	ring finger protein 139
RNF19A	101338464	101391503	62125690	62198945	ring finger protein 19A
RP1	55691180	55705947	22193803	22202968	retinitis pigmentosa 1
RPL30	99123118	99126949	64350193	64353512	ribosomal protein L30
RPL8	145985958	145988609	92447	95043	ribosomal protein L8
RPS20	57148167	57149623	23167503	23168751	ribosomal protein S20
RRM2B	103285907	103320522	60314878	60352055	ribonucleotide reductase M2 B
RRS1	67503817	67505522	30578864	30580630	RRS1 ribosome biogenesis regulator
RSPO2	108980721	109164742	54198066	54371110	R-spondin 2 homolog (Xenopus laevis)
RUNX1T1	93040328	93176619	70855365	70968428	runt-related transcription factor 1
SAMD12	119270875	119703365	42906357	43347772	sterile alpha motif domain containing 12
SCRIB	144945078	144969537	840294	848808	scribbled homolog (Drosophila)
SCXB	145393505	145395033	480388	482097	scleraxis homolog B (mouse)
SDC2	97575058	97693213	65772340	65789606	syndecan 2
SDCBP	59628282	59657973	24561204	24579915	syndecan binding protein (syntenin)
SGK3	67787445	67936811	30836526	30946287	serum/glucocorticoid regulated kinase
SHARPIN	145225524	145231128	567045	571373	SHANK-associated RH domain
SLA	134118155	134184479	7732247	7750788	Src-like-adaptor
SLC10A5	82768446	82769762	79893698	79895261	solute carrier family 10, member 5
SLC26A7	92330692	92479554	71722604	71871605	solute carrier family 26, member 7
SLC45A4	142290052	142307855	1959573	1988305	solute carrier family 45, member 4
SLC7A13	87295404	87311720	76285955	76312690	solute carrier family 7

SLURP1	143819364	143820831	1121843	1123776	secreted LY6/PLAUR domain
SNAI2	49992789	49996541	19605726	19609330	snail homolog 2 (Drosophila)
SNTB1	121619226	121893910	80962904	81331929	syntrophin, beta 1
SNTG1	50987150	51867980	20268374	20407189	syntrophin, gamma 1
SNX16	82874373	82916990	79767735	79801461	sorting nexin 16
SOX17	55533048	55536009	22087847	22089782	SRY (sex determining region Y)-box 17
SQLE	126079901	126103707	14605537	14627772	squalene epoxidase
ST3GAL1	134540312	134653344	7372446	7387594	ST3 beta-galactoside alpha
STK3	99536037	99907085	63654907	63959646	serine/threonine kinase 3
STMN2	80685935	80739792	39879856	39938981	stathmin-like 2
TAF2	120812195	120914255	80208724	80296937	TAF2 RNA polymerase II,150kDa
TATDN1	125569930	125620492	15020221	15051718	TatD DNase domain containing 1
TCEA1	55041669	55097561	21814089	21833010	transcription elongation factor A (SII), 1
TG	133948387	134216325	7658632	7894999	Thyroglobulin
TGS1	56848345	56900559	22927623	22953424	trimethylguanosine synthase homolog
TM7SF4	105421230	105438092	58325569	58332561	transmembrane 7 superfamily member 4
TMEM55A	92075675	92122227	72033012	72105209	transmembrane protein 55A
TMEM64	91704778	91727112	72191506	72229634	transmembrane protein 64
TMEM65	125392340	125454121	15167067	15211459	transmembrane protein 65
TMEM67	94836321	94898323	68865365	68910525	transmembrane protein 67
TMEM68	56813874	56848439	22891755	22927546	transmembrane protein 68
TMEM74	109864522	109868946	53422494	53423411	transmembrane protein 74
TNFRSF11B	120004977	120033564	42556262	42584443	tumor necrosis factor receptor
TOX	59880531	60194321	24763903	25074447	thymocyte selection-assc high mobility
TP53INP1	96007377	96030767	67736463	67746215	tumor protein p53 inducible nuclear
TPD52	81109658	81246391	40301192	40351554	tumor protein D52
TRAM1	71648227	71683158	34551562	34583855	translocation associated membrane
TRHR	110168900	110200989	53120296	53165263	thyrotropin-releasing hormone receptor
TRIB1	126511745	126519827	14240296	14248453	tribbles homolog 1 (Drosophila)
TRIM55	67201832	67250274	30414873	30461943	tripartite motif-containing 55
TRMT12	125532229	125534448	15108499	15110329	tRNA methyltransferase 12 homolog
TRNAM-CAU	124238651	124238723	16175577	16175649	transfer RNA methionine
TRNAS-AGA	96351061	96351142	67353675	67353756	transfer RNA serine (anticodon AGA)
TRNAY-GUA	66772086	66772173	30403968	30404061	transfer RNA tyrosine (anticodon GUA)
TRNAY-GUA	67188156	67188248	30404552	30404640	transfer RNA tyrosine (anticodon GUA)
TRPS1	116489900	116750402	46109178	46331384	trichorhinophalangeal syndrome I
TSPYL5	98354890	98359352	65036781	65038897	TSPY-like 5
TSTA3	144765933	144770038	906366	911205	tissue specific transplantation antigen
TTC35	109525029	109568312	53766048	53816056	tetratricopeptide repeat domain 35
TTPA	64135985	64161166	27959235	27966715	tocopherol (alpha) transfer protein
UBR5	103334745	103493671	60168917	60306581	ubiquitin protein ligase E3
UQCRB	97311910	97316987	66190479	66195192	ubiquinol-cytochrome c reductase b.p.
VCPIP1	67705042	67742006	30764507	30790163	valosin containing protein (p97)/p47 cpl
VPS28	145619808	145624735	278704	282724	vacuolar protein sorting 28 homolog
WDR67	124154101	124233571	16185133	16251964	WD repeat domain 67
WISP1	134272494	134310753	7576123	7590191	WNT1 inducible signaling pathway
XKR4	56177571	56601268	22620262	22806799	Kell blood group complex subunit
YTHDF3	64243675	64287900	28026852	28060347	YTH domain family, member 3

YWHAZ	102000090	102034745	61615879	61648185	tyr 3-monooxygenase/tryp5
ZBTB10	81561003	81597165	40745608	40779921	zinc finger and BTB domain
ZC3H3	144590968	144694763	965869	1027369	zinc finger CCCH-type containing 3
ZFAND1	82776513	82796085	79866472	79892713	zinc finger, AN1-type domain 1
ZFAT1	135559213	135794463	6550449	6698034	ZFAT zinc finger 1
ZFHX4	77756070	77942076	36743509	36948573	zinc finger homeobox 4
ZHX1	124329877	124355728	16088074	16110874	zinc fingers and homeoboxes 1
ZHX2	123863082	124055936	16345059	16347575	zinc fingers and homeoboxes 2
ZNF250	146075775	146097616	13417	23591	zinc finger protein 250
ZNF34	145969303	145983529	79180	85106	zinc finger protein 34
ZNF572	126054733	126060809	14639106	14665722	zinc finger protein 572
ZNF623	144789516	144807043	881706	885473	zinc finger protein 623
ZNF696	144444971	144451539	1501423	1503299	zinc finger protein 696
ZNF7	146023707	146039410	103440	110560	zinc finger protein 7
ZNF704	81713324	81949571	41020369	41209765	zinc finger protein 704
UN-ANNOTATED GENES ON BTA14					
ASNSL1	47610000	47645724	31890000	33620000	asparagine synthetase-like 1
CNBD1	87947792	88464071	74042000	74142000	cyclic nucleotide binding domain
COL14A1	121206533	121453454	80591000	80763000	collagen, type XIV, alpha 1 (undulin)
CPA6	68496962	68821134	31778000	31878000	carboxypeptidase A6
CYP11B2	143988977	143996261	1197000	1314000	cytochrome P450, fam11,sub B
CYP7A1	59565292	59575275	24437000	24537000	cytochrome P450, family 7, subfamily A
CYP7B1	65671246	65873902	29198000	29408000	cytochrome P450, family 7, subfamily B
DCC1	120915397	120937330	80307000	80407000	defective in sister chromatid cohesion
EPPK1	145012206	145019422	688000	788000	epiplakin 1
FAM135B	139211448	139578247	3882000	4072000	family with sequence similarity 135
FAM83H	144878091	144887902	798000	898000	family with sequence similarity 83
GLI4	144420982	144430476	1438000	1538000	GLI-Kruppel family member GLI4
HASNT	122720767	122726114	17618000	17718000	hyaluronan synthase 2 antisense
KIAA0196	126105684	126173243	14528000	14628000	KIAA0196
KIFC2	145662546	145670307	195000	295000	kinesin family member C2
LOC377711	145419563	145456704	444000	544000	KIAA1833-like
MAFA	144582658	144583719	1537000	1637000	v-maf musculoaponeurotic fibrosarcoma
MIRN124-2	65454259	65454367	29004000	29104000	microRNA 124-2
MIRN548D1	124429454	124429550	31344000	31444000	microRNA 548d-1
MIRN599	100618039	100618133	62643000	62943000	microRNA 599
NECAB1	91872954	92040806	72107000	72207000	N-terminal EF-hand calcium binding
PDE7A	66793867	66916297	30048000	30173000	phosphodiesterase 7A
PLEC1	145061309	145121531	679000	779000	plectin 1, intermediate fil. B.p 500kDa
PRDM14	71126575	71146116	34021000	34121000	PR domain containing 14
PSCA	143758877	143761145	1061000	1161000	prostate stem cell antigen
PSKH2	87129807	87150967	75475000	75575000	protein serine kinase H2
PVT1	128875961	129182681	10840000	12680000	Pvt1 oncogene homolog, MYC activator
RBM12B	94812904	94822400	68903000	69003000	RNA binding motif protein 12B
SAS-ZFAT	135679496	135682114	6560000	6660000	zinc finger protein 406 antisense trans
SLCO5A1	70747129	70909762	33646000	33918000	solute carrier organic anion transporter
SNORA72	99123490	99123621	30040000	30140000	small nucleolar RNA, H/ACA box 72
SNORD87	67997263	67997338	30946000	31046000	small nucleolar RNA, C/D box 87

SPAG1	101239439	101323306	62210000	62314000	sperm associated antigen 1
SPATC1	145158595	145174003	566000	666000	spermatogenesis and centriole assoc 1
ST18	53185952	53484856	20915000	21015000	suppression of tumorigenicity 18
TIGD5	144751364	144753627	874000	974000	tigger transposable element derived 5
TMEM71	133791376	133842010	7957000	8057000	transmembrane protein 71
VPS13B	100094670	100958984	62440000	63550000	vacuolar protein sorting 13 homolog B
WDSOF1	104496118	104524376	59066000	59166000	WD repeats and SOF1 domain
ZFP41	144400484	144416250	0	100000	zinc finger protein 41 homolog (mouse)
ZFPM2	106400323	106885943	56710000	57270000	zinc finger protein, multitype 2
ZNF16	146126548	146147078	59066000	59166000	zinc finger protein 16
ZNF706	102278442	102287166	61327000	61427000	zinc finger protein 706
NO ORTHOLOGS FOUND IN THE BOVINE GENOME					
BREA2	144851480	144852166			breast cancer estrogen-induced apopt 2
DDEF1IT1	131376783	131377961			DDEF1 intronic transcript 1
FAM150A	53609251	53640574			family with sequence similarity 150, A
FLJ43860	142513111	142586512			FLJ43860 protein
FLJ45872	138890869	139164970			FLJ45872 protein
HHLA1	133139449	133197055			HERV-H LTR-associating 1
LOC338328	144366443	144370418			high density lipoprotein-binding protein
LOC646486	82599836	82606105			fatty acid binding protein
LOC728724	130297893	130322678			hCG1814486
LY6K	143778533	143782613			lymphocyte antigen 6 complex, locus K
MIRN548A3	105565772	105565868			microRNA 548a-3
MIRN661	145091346	145091434			microRNA 661
MIRN875	100618189	100618264			microRNA 875
MIRN937	144967114	144967199			microRNA 937
MIRN939	145590171	145590252			microRNA 939
PXDNL	52394690	52884558			peroxidasin homolog (Drosophila)-like
SLC30A8	118216518	118258134			solute carrier family 30 (zinc transport)
SNHG6	67996719	68000331			small nucleolar RNA host gene 6
SNORD54	57148952	57149014			small nucleolar RNA, C/D box 54
TMEM75	129029046	129030151			transmembrane protein 75
TOP1MT	144462901	144488425			topoisomerase (DNA) I, mitochondrial
XKR9	71744154	71810731			Kell blood group complex subunit
ZNF707	144838610	144849543			zinc finger protein 707
UNASSIGNED ORTHOLOGOUS GENES IN THE BOVINE GENOME					
ARC	143689412	143692835			activity-regulated cytoskeleton-asc. Prot
BAI1	143542379	143623370			brain-specific angiogenesis inhibitor 1
CRISPLD1	76059531	76109348			cysteine-rich secretory protein LCCL
FAM82B	87553694	87590125			family with sequence similarity 82, B
GDAP1	75425173	75441900			ganglioside-induced differentiation
IMPAD1	58033042	58068981			inositol monophosphatase domain containing 1
JPH1	75309493	75396117			junctophilin 1
KCNB2	73612180	74013138			potassium voltage-gated channel
LOC441376	118019664	118024403			AARD protein
LY96	75066144	75103859			lymphocyte antigen 96
MSC	72916331	72919285			musculin (activated B-cell factor-1)

NPBWR1	54015021	54016007			neuropeptides B/W receptor 1
PABPC1	101784320	101803491			poly(A) binding protein, cytoplasmic 1
PI15	75899327	75929819			peptidase inhibitor 15
PTP4A3	142501189	142510802			protein tyrosine phosphatase type IVA, 3
RDH10	74369891	74399347			retinol dehydrogenase 10 (all-trans)
REXO1L1	86755947	86762978			REX1, RNA exonuclease 1 homolg S
RPESP	74141808	74168061			RPE-spondin
SCRT1	145525262	145530751			scratch homolog 1, zinc finger protein
SLC25A32	104480042	104496447			solute carrier family 25, member 32
STAU2	74624394	74821611			staufen, RNA binding protein, homolg 2
TCEB1	75021188	75046900			transcription elongation factor B (SIII), 1
TERF1	74083661	74122281			telomeric repeat binding factor 1
TMEM70	75050984	75057570			transmembrane protein 70
TRPA1	73096040	73150373			transient receptor potential cation channel, subfamily A
TSNARE1	143291348	143482444			t-SNARE domain containing 1
UBE2V2	49083548	49137007			ubiquitin-conjugating enzyme E2 var 2
UBE2W	74865394	74953664			ubiquitin-conjugating enzyme E2W
WWP1	87424110	87549297			WW domain containing E3 ubiquitin
RUN-AWAY ORTHOLOGOUS GENES LOCATED ON OTHER CHROMOSOME RATHER THAN BTA14					
COX6C	102297476	102297893			cytochrome c oxidase subunit VIc
hCG1984468	81633557	81634622			hCG1984468
hCG1988300	62653256	62655044			hCG1988300
RPL7	74365428	74368423			ribosomal protein L7
SEC11B	55597892	55626571			SEC11 homolog B (S. cerevisiae)
SLC39A4	145608606	145613081			solute carrier family 39 member 4
WDR21C	88952095	88955412			WD repeat domain 21C
ZNF251	145917103	145951775			zinc finger protein 251
ZNF517	145995065	146005333			zinc finger protein 517

Supplemental Table 3.2. Orthologous analysis of ORF genes between the synteny region of HSA8 and the bovine genome by targeting BTA14

	HOMO SAPIENS		BOS TAURUS		
Symbol	**Start**	**Stop**	**Start**	**Stop**	**Description**
ANNOTATED ORTHOLOGOUS GENES ON BTA14					
C8orf22	50147456	50151195	19646209	19650700	chromosome 8 open reading frame 22
C8orf32	124498146	124523441	15978057	15995914	chromosome 8 open reading frame 32
C8orf33	146248630	146252220	74573	76616	chromosome 8 open reading frame 33
C8orf37	96327411	96350613	67354206	67379248	chromosome 8 open reading frame 37
C8orf45	67948884	67976570	30954869	30981236	chromosome 8 open reading frame 45
C8orf46	67568321	67593294	30640296	30666569	chromosome 8 open reading frame 46
C8orf47	99145926	99175014	64311489	64336868	chromosome 8 open reading frame 47
C8orf53	117847923	117853380	44887352	44894591	chromosome 8 open reading frame 53
LOC728071	145408688	145411765	498000	598000	similar to brain protein 16
MGC70857	145722411	145725266	139000	239000	similar to RIKEN cDNA C030006K11
UN-ANNOTATED GENES ON BTA14					
C8orf30A	145264660	145267607	498000	598000	chromosome 8 open reading frame 30A
C8orf34	69512702	69893811	32478000	32849000	chromosome 8 open reading frame 34
C8orf38	96106397	96140114	67562000	67662000	chromosome 8 open reading frame 38
C8orf39	94820534	94822735	68898000	68998000	chromosome 8 open reading frame 39
C8orf55	143805623	143815352	1066000	1166000	chromosome 8 open reading frame 55
C8orf59	86313540	86319895	76144000	76244000	chromosome 8 open reading frame 59
C8orf76	124301412	124322798	16079000	16179000	chromosome 8 open reading frame 76
C8orf81	117955813	117958288	45000000	45500000	chromosome 8 open reading frame 81
NO ORTHOLOGS FOUND IN THE BOVINE GENOME					
C8orf51	144521033	144522174			chromosome 8 open reading frame 51
C8orf71	58354704	58355653			chromosome 8 open reading frame 71
C8orf73	144719505	144726071			chromosome 8 open reading frame 73
C8orf77	146199177	146231895			chromosome 8 open reading frame 77
UNASSIGNED ORTHOLOGOUS GENES IN THE BOVINE GENOME					
C8orf31	144192054	144207095			chromosome 8 open reading frame 31
C8orf44	67751008	67755789			chromosome 8 open reading frame 44
RUN-AWAY ORTHOLOGOUS GENES LOCATED ON OTHER CHROMOSOME RATHER THAN BTA14					
C8orf62	50815329	50816434	56579000	56679000	chromosome 8 open reading frame 62
C8orf70	79740885	79792490	56579000	56679000	chromosome 8 open reading frame 70

Supplemental Table 3.3. Orthologous analysis of hypothetical protein gene loci between the synteny region of HSA8 and the bovine genome by targeting BTA14

	HOMO SAPIENS		BOS TAURUS		
Symbol	**Start**	**Stop**	**Start**	**Stop**	**Description**
ANNOTATED ORTHOLOGOUS GENE ON BTA14					
C8ORFK36	124727096	124734371	15721449	15731757	hypothetical protein LOC340359
UN-ANNOTATED GENES ON BTA14					
FLJ46284	93794362	93867464	69969000	70069000	hypothetical gene AK128161
KIAA1833	145274907	145388831	449000	549000	hypothetical protein KIAA1833
LOC100128550	52892693	52896201	20719000	20819000	hypothetical protein LOC100128550
LOC100128687	91727308	91730288	72143000	72243000	hypothetical protein LOC100128687
LOC100129242	94508323	94577132	69135000	69235000	hypothetical protein LOC100129242
LOC100129885	145461347	145463278	431000	531000	hypothetical protein LOC100129885
LOC100130155	65448260	65453481	29000000	29100000	hypothetical protein LOC100130155
LOC100131637	86276685	86282422	76181000	76281000	hypothetical protein LOC100131637
LOC100132129	145645773	145659812	207000	307000	hypothetical protein LOC100132129
LOC100132559	143758394	143760390	1061000	1161000	hypothetical protein LOC100132559
LOC100132813	110725520	110729830	52495000	52595000	hypothetical protein LOC100132813
LOC100133047	81257811	81306027	40430000	40530000	hypothetical protein LOC100133047
LOC286144	93964934	94047548	69741000	69841000	hypothetical protein LOC286144
LOC644727	53788684	53790926	21275000	21375000	hypothetical LOC644727
MGC39715	101654299	101731069	61770000	61870000	hypothetical protein MGC39715
NO ORTHOLOGS FOUND IN THE BOVINE GENOME					
FLJ10489	104202442	104214370			hypothetical protein FLJ10489
FLJ39080	75674656	75833142			hypothetical gene AK096399
FLJ46365	49665514	49666906			hypothetical LOC401459
LOC100127998	50980886	50987126			hypothetical protein LOC100127998
LOC100128539	70788179	70794511			hypothetical protein LOC100128539
LOC100128545	61488123	61488817			hypothetical protein LOC100128545
LOC100128549	109390427	109516738			hypothetical protein LOC100128549
LOC100128692	66791139	66797284			hypothetical protein LOC100128692
LOC100128829	73104515	73128111			hypothetical protein LOC100128829
LOC100128962	87261763	87293581			hypothetical protein LOC100128962
LOC100129523	82600506	82608143			hypothetical protein LOC100129523
LOC100129525	130497383	130527280			hypothetical protein LOC100129525
LOC100129527	73504410	73504715			hypothetical protein LOC100129527
LOC100129596	146016501	146016953			hypothetical protein LOC100129596
LOC100129809	70500280	70523027			hypothetical protein LOC100129809
LOC100129954	48333807	48335938			hypothetical protein LOC100129954
LOC100130027	146006492	146007368			hypothetical protein LOC100130027
LOC100130588	58669279	58880142			hypothetical protein LOC100130588
LOC100130939	141647523	141647904			hypothetical protein LOC100130939
LOC100131102	104247333	104327946			hypothetical protein LOC100131102
LOC100131144	144726809	144730700			hypothetical protein LOC100131144

LOC100131552	123781689	123799755			hypothetical protein LOC100131552
LOC100131595	68249722	68254995			hypothetical protein LOC100131595
LOC100131721	82354477	82356473			hypothetical protein LOC100131721
LOC100131910	141114782	141115174			hypothetical protein LOC100131910
LOC100132183	94796649	94797517			hypothetical protein LOC100132183
LOC100132374	86321916	86328497			hypothetical protein LOC100132374
LOC100132709	86541332	86564153			hypothetical protein LOC100132709
LOC100133015	145672420	145695490			hypothetical protein LOC100133015
LOC137886	59486377	59526614			hypothetical protein LOC137886
LOC286094	136315668	136377980			hypothetical protein LOC286094
LOC389676	94215699	94248255			hypothetical LOC389676
LOC392242	86742172	86746963			hypothetical gene NM_172239
LOC643972	104101548	104102826			hypothetical LOC643972
LOC728610	104327737	104380185			hypothetical protein LOC728610
LOC729696	74819711	74822229			hypothetical protein LOC729696
UNASSIGNED ORTHOLOGOUS GENES IN THE BOVINE GENOME					
LOC100127988	74432114	74437577			hypothetical protein LOC100127988
LOC100128126	74494863	74516305			hypothetical protein LOC100128126
LOC100129092	56210317	56217472			hypothetical protein LOC100129092
LOC100130101	97417524	97418748			hypothetical protein LOC100130101
LOC100132891	72918191	73040259			hypothetical protein LOC100132891
LOC286177	58335727	58341727			hypothetical protein LOC286177
RUN-AWAY ORTHOLOGOUS GENES LOCATED ON OTHER CHROMOSOME RATHER THAN BTA14					
LOC100128259	107347604	107354303			hypothetical protein LOC100128259
LOC100129098	55540707	55545539			hypothetical protein LOC100129098
LOC100130665	97453369	97468842			hypothetical protein LOC100130665
LOC100132012	85903363	85904269			hypothetical LOC100132012
LOC100133147	122194774	122195341			hypothetical LOC100133147
LOC727887	90982092	90983767			hypothetical protein LOC727887

Supplemental Table 3.4. Orthologous analysis of similar-to-other loci between the synteny region of HSA8 and the bovine genome by targeting BTA14

	HOMO SAPIENS		BOS TAURUS		
Symbol	**Start**	**Stop**	**Start**	**Stop**	**Description**
UN-ANNOTATED GENES ON BTA14					
LOC100128414	95513350	95518531	68168000	68268000	similar to fibrinogen silencer binding prot
LOC100128419	55774637	55896143	22234000	22334000	similar to hCG2036697
LOC100129173	145238276	145243184	509000	609000	similar to KIAA1875 protein
LOC100129963	65629341	65655678	29154000	29254000	similar to hCG1983308
LOC100130274	144860852	144862267	822000	922000	similar to hCG1646697
LOC100130742	92184066	92206187	71893000	71993000	similar to hCG2008599
LOC100130862	71648007	71650116	34517000	34617000	similar to PRO1292
LOC286187	68062920	68103340	30992000	31092000	similar to RIKEN cDNA 1700011J18
LOC442388	57448181	57470478	23344000	23444000	similar to epidermal retinal dehydrogenase 2
LOC642574	144925288	144944766	769000	869000	similar to CG3104-PA, isoform A
LOC650747	82680356	82705924	79865000	79965000	similar to Inositol monophosphatase (IMP)
LOC727967	145388936	145393461	435000	535000	similar to block of proliferation 1
LOC728774	69600306	69600695	21892000	21992000	similar to ribosomal protein S15a
LOC728967	79740915	79742910	38859000	38959000	similar to chromosome 8 orf70
NO ORTHOLOGS FOUND IN THE BOVINE GENOME					
LOC100129104	134650165	134655281			similar to glycoprotein VSP-3
LOC100130098	96154018	96155025			similar to hCG2008575
LOC100130231	127027163	127032609			similar to hCG1814455
LOC100130232	105463956	105464985			similar to LP2209
LOC100130301	74316118	74360696			similar to rCG30376
LOC100131064	145508625	145510391			similar to HSF1 protein
LOC100131819	70907491	70925206			similar to hCG1778814
LOC100132807	100422134	100422777			similar to hCG2040244
LOC392275	145176011	145178536			similar to Sphingomyelin phosphodiesterase
LOC441377	119843187	119843640			similar to 40S ribosomal protein S26
LOC644334	49945392	49989383			similar to Band 4.1-like protein 5
UNASSIGNED ORTHOLOGOUS GENES IN THE BOVINE GENOME					
LOC100129654	68022627	68040962			similar to hCG2037011
LOC392217	48233219	48233685			similar to Ig lambda chain V region 4A prec
RUN-AWAY ORTHOLOGOUS GENES LOCATED ON OTHER CHROMOSOME RATHER THAN BTA14					
LOC100128689	48666940	48667731			similar to ubiquitin-conjugating enzyme E2
LOC100130092	136706149	136707406			similar to MAPRE1 protein
LOC100131726	124283897	124284130			similar to HCC-related HCC-C11_v3
LOC100131813	104093982	104098751			similar to hCG2041218
LOC286157	75677839	75679320			similar to poly(rC) binding protein 2
LOC646463	82328919	82330360			similar to Ubiquitin-conjugating enzyme E2

Supplemental Table 3.5. Orthologous analysis of pseudogenes between the synteny region of HSA8 and the bovine genome by targeting BTA14

	HOMO SAPIENS		BOS TAURUS		
Symbol	**Start**	**Stop**	**Start**	**Stop**	**Description**
UN-ANNOTATED GENES ON BTA14					
BTF3L2	71347952	71349049	64829000	64929000	basic transcription factor 3, like 2
BTF3P1	52796779	52797298	64829000	64929000	basic transcription factor 3, pseudogene 1
LOC100127983	92039504	92066661	72094000	72194000	hypothetical protein LOC100127983
LOC100127992	82709139	82720079	21986000	22086000	hypothetical LOC100127992
LOC100128271	89114835	89116348	73393000	73493000	hypothetical LOC100128271
LOC100128418	105813020	105862998	57740000	58910000	hypothetical LOC100128418
LOC100128548	95971841	95973548	62971000	63071000	hypothetical LOC100128548
LOC100128955	71498914	71553870	34391000	34491000	hypothetical LOC100128955
LOC100128961	101631932	101638258	27194000	27294000	hypothetical LOC100128961
LOC100129096	70090765	70205819	27250000	27980000	hypothetical LOC100129096
LOC100129237	57176257	57177380	9118000	9218000	similar to nucleophosmin
LOC100129377	106062541	106066289	57647000	57747000	hypothetical LOC100129377
LOC100129960	71017383	71019437	24523000	24623000	hypothetical LOC100129960
LOC100130095	109576545	109734724	53600000	53700000	hypothetical LOC100130095
LOC100130861	47746707	47816648	17000000	17100000	hypothetical LOC100130861
LOC100131371	63930417	63935921	27732000	27832000	hypothetical LOC100131371
LOC100131467	104849685	104850433	13990000	1409000	hypothetical LOC100131467
LOC100131925	54253778	54270620	60062000	60162000	hypothetical LOC100131925
LOC100132280	111619124	111637925	51761000	51861000	hypothetical LOC100132280
LOC392225	61356623	61463108	24540000	26100000	similar to phosducin-like 3
LOC643228	97150133	97195221	66250000	66930000	similar to GAPDH
LOC644199	108254600	108333251	53330000	55350000	similar to High mobility group protein B1
LOC644335	110799139	110863693	6577000	6945000	similar to Est1p-like protein B
NPM1P6	62277474	62278314	9118000	9218000	nucleophosmin 1 pseudogene 6
RPL10AP2	48187776	48191592	23845000	23945000	ribosomal protein L10a pseudogene 2
NO ORTHOLOGS FOUND IN THE BOVINE GENOME					
ATP6V1GP2	48224100	48225305			ATPase, H+ transporting, psdge 2
CYCSP23	120699577	120699895			cytochrome c, somatic pseudogene 23
DUXAP2	102447360	102448568			double homeobox A pseudogene 2
IFITM8P	64484177	64484575			interferon induced trans protein 8 psdg
LOC100128541	48224788	48227012			hypothetical LOC100128541
LOC100129370	111890701	111891299			hypothetical LOC100129370
LOC100129371	100427469	100429509			hypothetical LOC100129371
LOC100130225	71581044	71581607			hypothetical LOC100130225
LOC100130299	49369137	49369910			hypothetical LOC100130299
LOC100130300	54962121	54962592			hypothetical LOC100130300
LOC100131958	130945906	130946597			similar to rCG64241
LOC100131992	145411783	145414583			hypothetical LOC100131992
LOC100132812	68644205	68644615			hypothetical LOC100132812

LOC100133058	94300444	94311849			similar to hCG1644482
LOC100133156	54050143	54051229			hypothetical LOC100133156
LOC392221	55266698	55267149			similar to Co-co-hel-co-co-hel dm-prot 2
LOC392262	111186481	111187369			similar to 40S ribosomal protein SA
LOC442396	125233346	125368864			similar to ADP-ribosylation factor 1
LOC644233	109212456	109223136			similar to Serine/threonine-protein kinase
LOC648630	145267755	145270555			similar to testis-specific serine kinase 5
MRP63P7	99987659	99987931			mitochondrial ribosomal protein 63 psd 7
MRPS16P1	92998911	92999326			mitochondrial ribosomal prot S16 psd 1
MRPS36P3	123172890	123173022			mitochondrial ribosomal prot S36 psd 3
POU5F1P2	103701961	103702651			POU class 5 homeobox 1 pseudogene 2
REXO1L2P	86884284	86885361			REX1, RNA exonuclease 1 homolog-2
REXO1L3P	86754080	86756141			REX1, RNA exonuclease 1 homolog-3
REXO1L4P	86859903	86861964			REX1, RNA exonuclease 1 homolog -4
REXO1L5P	86872094	86874155			REX1, RNA exonuclease 1 homolog-5
REXO1L6P	86896198	86898259			REX1, RNA exonuclease 1 homolog-6
REXO1L7P	86908388	86911709			REX1, RNA exonuclease 1 homolog-7
VENTXP6	74725690	74727046			VENT homeobox (Xenopus laevis) psd6
UNASSIGNED ORTHOLOGOUS GENES IN THE BOVINE GENOME					
LOC100128120	74904993	74906764			hypothetical LOC100128120
LOC100128338	144896449	144900494			hypothetical LOC100128338
LOC100129659	66807672	66808547			hypothetical LOC100129659
LOC100130298	62040879	62042894			similar to hCG1816373
LOC100130834	144740647	144742319			similar to LPPA601
LOC100131770	67555439	67556278			hypothetical LOC100131770
LOC100132809	72605900	72631810			similar to hCG1646161
LOC100133234	56524623	56530364			similar to SET binding factor 1
LOC392232	73229701	73326354			similar to Transient receptor potential cat
MAPK6PS4	48002018	48006274			mitogen-activated protein kinase 6 psd4
MAPK6PS5	110551443	110553756			mitogen-activated protein kinase 6 psd5
RUN-AWAY ORTHOLOGOUS GENES LOCATED ON OTHER CHROMOSOME RATHER THAN BTA14					
CKS1A	81719154	81720070			CDC28 protein kin reg subunit 1A psd
CPP	92238687	92241126			ceruloplasmin (ferroxidase) pseudogene
CYCSP22	51837599	51837911			cytochrome c, somatic pseudogene 22
H2AFZP2	71178107	71179157			H2A histone family, member Z, psd 2
HNRPA1P4	83365593	83369814			heterog nuclear ribonucleoprotA1 psd 4
LOC100101127	70868957	70875031			thioredoxin domain containing 1 psd
LOC100127925	144912311	144914671			hypothetical LOC100127925
LOC100127982	98934421	98935214			similar to transmembrane protein 69
LOC100128119	81669523	81670491			hypothetical LOC100128119
LOC100128132	106276268	106278373			hypothetical LOC100128132
LOC100128412	88656977	88657871			hypothetical LOC100128412
LOC100128540	64211520	64213197			hypothetical LOC100128540
LOC100128627	144148702	144149404			similar to Chain A
LOC100128685	107291147	107298802			hypothetical LOC100128685
LOC100128686	51934628	51935175			hypothetical LOC100128686
LOC100128836	82882489	82884185			similar to heterogeneous nuclear

					ribonucleoprotein A1
LOC100128963	80645941	80655256			hypothetical LOC100128963
LOC100129093	48585624	48587835			similar to p47
LOC100129100	89409840	89567439			hypothetical LOC100129100
LOC100129367	138778149	138779892			hypothetical LOC100129367
LOC100129512	103702904	103707267			hypothetical LOC100129512
LOC100129667	54787810	54788306			hypothetical LOC100129667
LOC100129815	108728684	108729440			hypothetical LOC100129815
LOC100129816	100314453	100319258			hypothetical LOC100129816
LOC100130096	59499701	59504922			hypothetical LOC100130096
LOC100130153	84272143	84277221			hypothetical LOC100130153
LOC100130158	127154013	127155664			hypothetical LOC100130158
LOC100130448	126001919	126003505			hypothetical LOC100130448
LOC100130524	95283317	95284456			hypothetical LOC100130524
LOC100131013	55712181	55772240			hypothetical LOC100131013
LOC100131146	143047359	143048109			hypothetical LOC100131146
LOC100131252	74352148	74352835			hypothetical LOC100131252
LOC100131254	55047062	55048895			hypothetical LOC100131254
LOC100131333	110980573	110984558			hypothetical LOC100131333
LOC100131339	87124447	87124906			hypothetical LOC100131339
LOC100131593	102610064	102610972			hypothetical LOC100131593
LOC100131775	95020759	95024281			hypothetical LOC100131775
LOC100131849	99494286	99495188			hypothetical LOC100131849
LOC100132553	92239087	92241390			similar to Ceruloplasmin precursor
LOC100132965	74980143	74989799			hypothetical LOC100132965
LOC157667	94930965	94931484			similar to myosin regulatory light chain
LOC286150	98705607	98706483			hypothetical LOC286150
LOC340443	81375659	81376271			similar to ribosomal protein S5
LOC389655	49459705	49460174			similar to 60S ribosomal protein L29
LOC389662	57551877	57615579			similar to septin 10 isoform 1
LOC392226	63218264	63219901			similar to Serine/threonine-protein kinase
LOC392264	120561463	120612050			hypothetical gene AK122835
LOC392265	123850694	123851623			similar to Cell division protein kinase 5
LOC392266	124367646	124372344			similar to Peroxisomal multifunctional
LOC392267	124481903	124484336			inosine monophosphate dehydro 1 psd
LOC392268	125223304	125224934			similar to neuropilin- and tolloid-like
LOC392271	134407422	134407829			similar to 60S ribosomal protein L32
LOC402342	81340295	81343564			similar to ribosomal protein L13a
LOC442389	61981152	61981713			interferon induced transmembrane psd
LOC642137	87098164	87098471			similar to large ribosomal protein L36a
LOC642367	87562919	87564347			similar to N-terminal Asn amidase
LOC642461	88679804	88683158			similar to mitochondrial isoleucine tRNA
LOC643022	95389391	95435018			hypothetical LOC643022
LOC643494	99986001	99987074			similar to RNA binding motif protein 4B
LOC643831	102952899	102954063			phosducin-like 3 pseudogene
LOC644103	106092457	106092972			similar to 60S ribosomal protein L29
LOC644363	79834004	79837327			hypothetical LOC644363
LOC645551	62984880	62986533			similar to core 1 synthase
LOC646197	74290688	74291416			similar to heat shock protein 1, beta

LOC650095	127033280	127034275			similar to Superoxide dismutase
LOC728381	109064378	109066170			nuclear receptor binding factor 2 psd
LOC728587	47867504	47934176			similar to zinc finger protein 273
MAPK6PS1	54611373	54616779			mitogen-activated protein kinase 6 psd 1
MRPL9P1	77677517	77678742			mitochondrial ribosomal protein L9 psd 1
NACAP1	102450297	102450999			nascent-polypeptide-assc complex psd1
NASPP1	62011956	62015186			nuclear autoantigenic sperm protein psd 1
POU5F1P1	128497294	128498373			POU class 5 homeobox 1 pseudogene 1
PSMC6P	57124694	57126429			proteasome (prosome) 26S subunit, psd
PTTG3	67842186	67842794			pituitary tumor-transforming 3
RPL37P6	57663404	57663993			ribosomal protein L37 pseudogene 6
RPS23P1	98946328	98947032			ribosomal protein S23 pseudogene 1
RPS26P10	93225294	93225965			ribosomal protein S26 pseudogene 10
RPS26P6	101977075	101977622			ribosomal protein S26 pseudogene 6
SEDLP2	72523394	72524129			spondyloepiphyseal dysplasia, late, psd 2
SOX5P	88869379	88871228			SRY (sex determining region Y)- 5 psd
TARBP2P	104739083	104741639			TAR (HIV-1) RNA binding protein 2 psd
TMED10P	146191076	146192348			transmembrane emp24-like trafficking ps

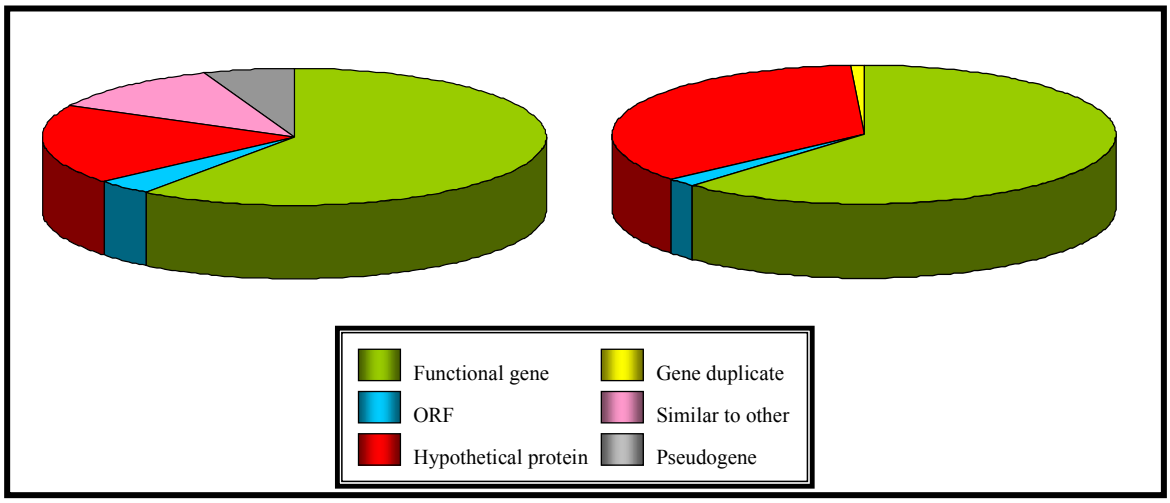

Figure 3.1. Graphic representation and comparison of annotated genes on BTA14 Btau_4.0) and HSA8 (build 36.3).

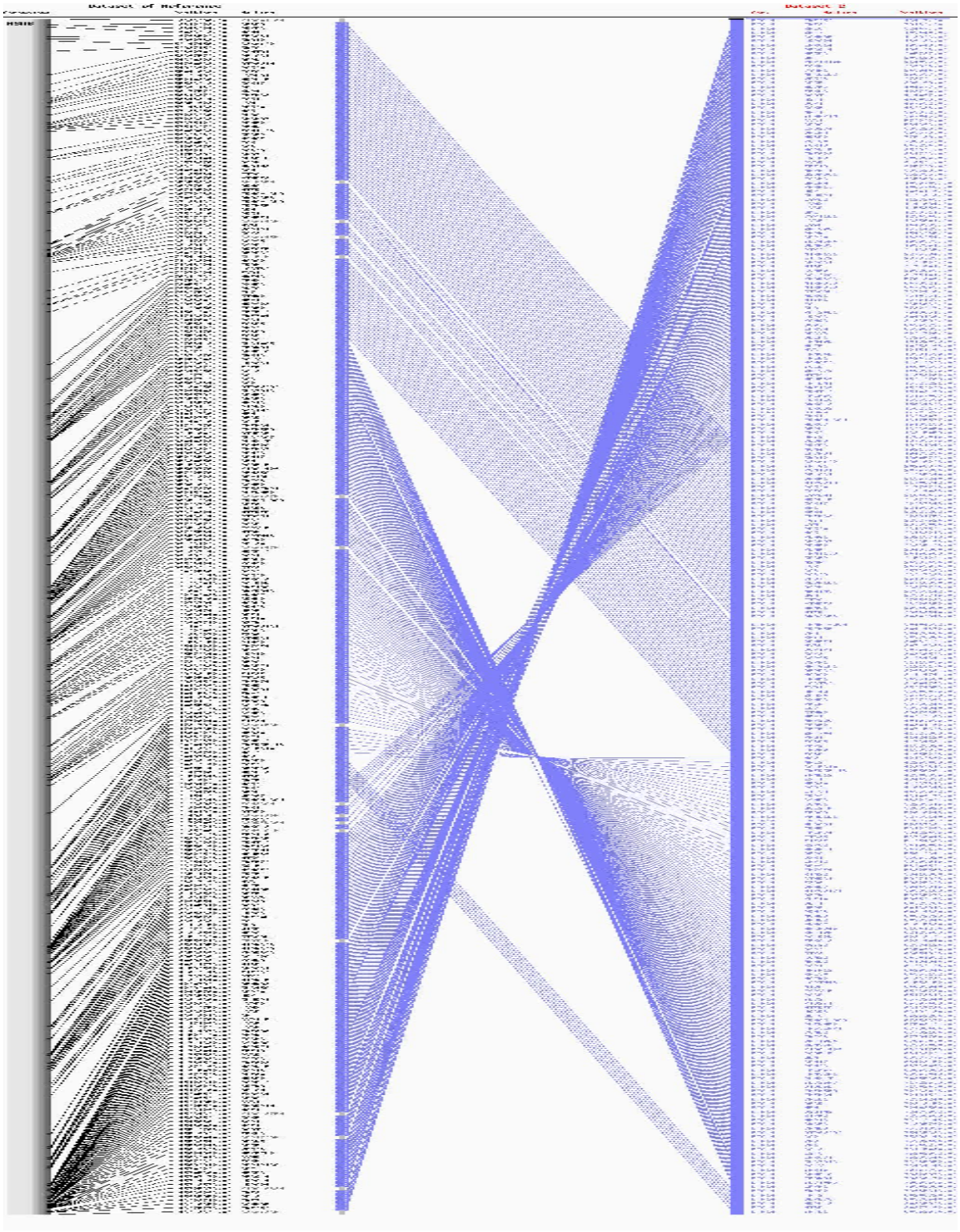

Figure 3.2. Comparative map between HSA8 and BTA14 based on 411 orthologous genes/loci.

Chapter Four:

THE CORTICOTROPIN RELEASING HORMONE AS A STRONG CANDIDATE GENE FOR MARBLING AND SUBCUTANEOUS FAT DEPTH IN BEEF CATTLE

The corticotropin releasing hormone as a strong candidate gene for marbling and subcutaneous fat depth in beef cattle

Tito A. Wibowo, Jennifer J. Michal, Zhihua Jiang

Department of Animal Sciences, Washington State University, Pullman WA 99164-6351

Correspondence to: **Dr. Zhihua Jiang**

Phone: 509 – 335 8761

Fax: 509 – 335 4246

Email: jiangz@wsu.edu

4.1. Abstract

Corticotropin releasing hormone (*CRH*) gene is mapped on bovine chromosome 14 (BTA14), where more than 30 fat-related quantitative trait loci have been reported in dairy and beef cattle. The gene regulates secretion of adrenocortocotrophin hormone, the hypothalamic-pituitary-adrenal axis and multiple hypothalamic functions, therefore, we hypothesize that *CRH* is a strong candidate gene for marbling and subcutaneous fat depth (SFD) in a Wagyu x Limousin F_2 population. Two pairs of primers were designed, one targeting the proximal promoter region and non-coding exon 1 and one amplifying the coding exon 2 of this gene. A total of five single nucleotide polymorphisms were identified, including AAFC03076794.1:*g.9657C>T*, *c.10718G>C*, *c.10841G>A*, *c.10893A>C* and *c.10936G>C*. Among these four cSNPs, *c.10718G>C*, *c.10841G>A*, and *c.10936G>C* are missense mutations, leading to amino acid changes from arginine to proline, from serine to asparagine and from aspartic acid to histidine, respectively. All of these five SNPs were genotyped on ~250 F_2 progeny, but four were selected as tagging SNPs for association analysis due to no historical recombination observed between *c.10718G>C* and *c.10893A>C*. General linear model showed *g.9657C>T*, *c.10718G>C*, and *c.10936G>C* had significant effects on SFD (P=0.002 - 0.001), but only *c.10936G>C* was significantly associated with marbling (P=0.022). Our results provide further evidence that *CRH* is a strong candidate gene for a concordant QTL related to lipid metabolism in mammals.

4.2. Introduction

Genome wide screenings using microsatellite markers have shown that bovine chromosome 14 (BTA14) harbors quantitative trait loci (QTL) for both beef marbling score (BMS) and subcutaneous fat depth (SFD) in beef cattle. In two half-sib families developed by mating a Belgian Blue x MARC III sire and a Piedmontese x Angus sire to MRAC III dams, Casas and collegues (2000) observed a fat depth QTL at 38 cM interacting with the myostatin genotypes on BTA2. In a half-sib family produced by mating a Brahman x Hereford sire to Hereford, Augus, Hereford x Augus and MARCIII dams, the same group identified a fat-depth QTL at 16 cM, while a marbling QTL at 47 cM on BTA14 (Casas et al., 2003). A QTL for marbling on BTA14 was also detected in a half-sib family of 348 purebred Japanese Wagyu steers, but it is located at 53 cM (Mizoshita et al., 2004).

Both diacylglyceril O-acyltransferase 1 (*DGAT1*) and thyroglobulin (*TG*) have been proposed as potential candidate genes for marbling and fat depth QTLs on BTA14, because they have an effect on lipid metabolism. The DGAT1 enzyme utilizes diacylglycerol and fatty acyl CoA as substrates in order to catalyze the final stage of triacylglycerol synthesis, while thyroglobulin is the glycoprotein precursor to the thyroid hormones whose metabolism is important for energy expenditure and dissipation of heat in special tissues. However, associations of these two genes with both marbling and fat depth have been reported inconsistently among different populations. In a Canadian beef population, neither *DGAT1* nor *TG* showed a significant ($P>0.10$) association with the backfat EBV (estimated breeding value) (Moore et al., 2003). In a study using 22

Brahman sire families mated to the Brahman dams, Casas and colleagues (2005) reported a significant association of *TG* with fat thickness ($P<0.05$), but not with marbling score. No significant associations of the *DGAT1* polymorphism were observed for either marbling or fat thickness. Just recently, a meta-analysis contacted on an Australian beef population provided substantial evidence for an additive association between a *TG* marker and marbling using a Bayesian hierarchical model (Wood et al. 2006). All these data indicate that genes underlying QTLs for both marbling and fat depth on BTA14 remain unclear.

Corticotrophin releasing hormone (CRH) has been known for its involvements in many biological and physiological actions and functions. Basically, CRH plays an important role as the major hypothalamic releasing factor for pituitary adrenocorticotropin (ACTH) secretion (Seasholtz et al., 2002), which regulates glucocorticoid and cathecolamines to mediate stress response. Behavioral effects of CRH include increased locomotor activity and inhibition of food intake, while its actions on metabolism are mediated mainly by activation of the sympathetic nervous system (Rothwell 1990). Interestingly, there is increasing evidence supporting involvement of this CRH peptide in the regulation of energy balance and body weight, influencing both food intake and sympathetically-mediated thermogenesis. For example, the increased activity of the hypothalamic-pituitary-adrenal (HPA) axis stimulated by *CRH* was found to be highly associated with abdominal fat (Perusse et al., 2001). Furthermore, ACTH secretion under a stress environment stimulates glucocorticoids, which help return the stress system to homeostasis and mediate many metabolic changes, such as increases of leptin production

(Seasholtz et al., 2002; Buchanan et al., 2005). As the CRH gene is located on the bovine chromosome 14 (BTA14), we decided to validate its candidacy for marbling and SFD in a Wagyu x Limousin F2 cross.

4.3. Materials and Methods

4.3.1. Animals

A Wagyu x Limousin reference population was generated jointly by Washington State University and the Fort Keogh Livestock and Range Research Laboratory, ARS, USDA, as described previously (Jiang et al., 2005). However, DNA extraction on 6 F_1 bulls, 113 F_1 dams and ~250 F_2 progeny plus performance data collection on these F_2 animals was conducted in the USDA laboratory. Marbling was a subjective measure of the amount of intramuscular fat in the *longissimus* muscle based on USDA standards (http://www.ams.usda.gov/), ranging from 4 = Slight^{0} to 9.5 = Moderately Abundant^{50} (SD = 1.00) in this F_2 population. SFD was measured at the 12-13th rib interface perpendicular to the outside surface at a point three-fourths the length of the longissimus muscle from its chine bone end with a range of 0.1 to 1.3 inches (SD = 0.18) in this F_2 population. Based on the availability of both data and DNA samples, 246 F_2 animals were used in the present study.

4.3.2. DNA sequences and primer design

Both cDNA sequence (CO895988) and genomic DNA contig (AAFC03076794) of the bovine *CRH* gene were retrieved from the GenBank databases and used to determine its genomic organization by sequence alignment. Primer design was completed using the

online oligonucleotide design tool Primer3. Two pairs of primers were designed: one targets the proximal promoter and non-exon 1 (forward –5'CCC CTC CCA TTC ACT CTC TTT TCT 3' and reverse – 5'AGT TCT GTC TAG GCG CTC CCT ACC3'), and the other amplifies exon 2 region (forward – 5' GGG TCT GTG GGT GTC GTC CT 3' and reverse – 5' AAA AAT AAA CAT GGT ATC AGA GCA ATG 3') of the bovine *CRH* gene, respectively.

4.3.3. Mutation detection and genotyping

Approximately 50 ng of genomic DNA each from six Wagyu x Limousin F_1 bulls was amplified in a final volume of 10 μL that contained 12.5 ng of each primer, 150 μM dNTPs, 1.5 mM $MgCl_2$, 50 mM KCl, 20 mM Tris-HCl and 0.25U of Platinum Taq polymerase (Invitrogen, Carlsbad, CA). The PCR conditions were carried out as follows: 94°C for 2 min, 35 cycles of 94°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec, followed by a further 5 min extension at 72°C. PCR products were then sequenced for polymorphism detection performed on ABI 3730 sequencer in the Laboratory for Biotechnology and Bioanalysis (Washington State University) using a standard protocol. Only one single nucleotide polymorphism (SNP) was detected in the promoter/exon 1 product and then revealed with digestion of restriction enzyme *Hha*I. However, four SNPs were identified in the exon 2 region so that the same PCR product direct sequencing approach was used to genotype them on ~ 250 F_2 progeny.

4.3.4. Data analysis

The degrees of Hardy-Weinberg equilibrium within each marker and linkage disequilibrium among different markers in the bovine *CRH* gene were estimated using HAPLOVIEW program (Barrett et al. 2005). The phenotypic data for both IMCL and SFD measurements have been previously adjusted for year of birth, sex, age (days), live weight (kilograms), or fat depth (inches), as appropriate. The adjusted phenotypes were then used in a subsequent association analysis using the GLM (general linear model) procedure of SAS v9.1 (SAS institute Inc., Gary, NC). Pair-wise comparisons of least squares means were performed using a protected t-test. Additionally, quantitative transmission disequilibrium test (QTDT) (Abecasis et al. 2000) was performed to further examine the association between markers and adjusted obesity-related phenotype data. *P* value <0.05 was considered statistically significant. For significantly associated mutations, the MatInspector web server (Quandt et al. 1995) was used to screen potential transcriptional regulatory binding site changes caused by promoter polymorphisms, while the Mfold web server (Zuker 2003) was used to predict mRNA secondary structure changes caused by coding polymorphisms.

4.4. Results

4.4.1. Annotation of the bovine CRH gene

Search against the GenBank database surprisingly revealed that the bovine *CRH* gene has not been well annotated yet. Although a sequence (AF340152) with a complete coding sequence for the bovine gene was submitted to the GenBank in 2001, it represents the exon 2 region only. In the present study, therefore, we did a BLAST search against the bovine EST (expressed sequence tags) database using the sequence AF340152 as a

reference and retrieved a full-length cDNA sequence (CO895988) for the bovine *CRH* gene. A genomic DNA contig (AAFC03076794) for the same gene was then retrieved from the bovine genome sequence database and thus alignment between the cDNA sequence and the genomic sequence determined the genomic organization of bovine *CRH* gene (Figure 4.1A). Like its human ortholog, the bovine *CRH* gene has two exons and one intron. The exon 1 is a non-coding exon, but the exon 2 contains 11 bp of non-coding sequence and 573 bp of complete coding sequence. The size of intron 1 is 771 bp in length (AAFC03076794). In addition, five ESTs (EE338630, DV826091, DV825584, DV822182, and EE339662) show that the bovine CRH gene might encode a new splicing form with a prohormone of 130 amino acids, 60 amino acids shorter than the regular one. So far, no new splicing forms of *CRH* gene have been reported in other mammals.

4.4.2. SNPs and haplotypes

A total of five single nucleotide polymorphisms were identified, including one SNP (AAFC03076794.1: *g.9657C>T*) in the proximal promoter region and four SNPs (AAFC03076794.1: *c.10718G>C*, *c.10841G>A*, *c.10893A>C* and *c.10936G>C*) in the exon 2 region. Among these four cSNPs, *c.10718G>C*, *c.10841G>A*, and *c.10936G>C* are missense mutations, leading to amino acid changes from *Arginine* to *Proline*, from *Serine* to *Asparagine* and from A*spartic acid* to *Histidine*, respectively. The minor alleles among these five SNPs are C, G, A, A, C respectively, with a frequency ranging from 0.08 to 0.42. Genotyping on ~250 F2 progeny indicated that all five SNPs fall into the Hardy-Weinberg equilibrium (P>0.05). HAPLOVIEW analysis indicated that among these five SNPs, two SNPs *c.10718G>C* and *c.10893A>C* have no-historical

recombination by forming two haplotypes of *CC* and *GA* (Figure 4.1), and thus eight haplotypes were identified in the population, including *TCGCC*, *CGGAG*, *TCGCG*, *CGAAG*, *CCGCC*, *TGGAG*, *TGAAG* and *CCACG* with a frequency of 0.368, 0.326, 0.204, 0.070, 0.016, 0.012, 0.003 and 0.002, respectively.

4.4.3. Promoter SNP and potential regulatory binding sites

Screening the proximal promoter region using *MatInspector* web server program revealed that allele *g.9657C*, but not a *g.9657T* gains four possible transcription factor binding sites, including neuron restrictive silencer factor (NRSF), E2F transcription factor, CDE-CHR binding factor-1 (CDF-1) and transcription factor CP2 (Figure 4.2). In fact, the promoter region flanking the polymorphic site in cattle is highly conserved in other mammals, such as human (NT_008183.18), chimpanzee (XM_519792.2), rhesus monkey (XM_001094433.1), mouse (NW_001030719.1), rat (M54987.1), sheep (M22853.1), dog (AB162117) and pig (DQ358705.1) (Figure 4.3). However, among the four regulatory binding sites gained by the allele *g.9657C* in cattle, only the NRSF is retained in all of other eight species (Figure 4.3).

4.4.4. Coding SNPs and the mRNA secondary structure

HAPLOVIEW analysis indicated that four cSNP: *c.10718G>C*, *c.10841G>A*, *c.10893A>C* and *c.10936G>C* form five haplotypes: *GGAG*, *CGCG*, *CGCC*, *CACG* and *GAAG*, respectively. The Mfold program (Zuker 2003) was used to predict how these haplotypes affect mRNA secondary structure involving a complete coding sequence of 573 bp for the preprohormone of the bovine *CRH* gene. Figure 4.4a shows that the first

three haplotypes (*GGAG*, *CGCG* and *CGCC*) gave the same secondary structures. The secondary structures of the last two haplotypes were illustrated in Figure 4.4b and 4.4c, respectively, but they just slightly differ from each other (see arrows inside the boxes). However, there was a remarkable difference in the secondary structure between the first three haplotypes and the last two haplotypes. Obviously, polymorphic site *c.10841G>A* play a critical role in determining the secondary structure of the *CRH* mRNA in cattle.

4.4.5. Association analysis of SNPs with SFD and marbling

As both SNPs - *c.10718G>C* and *c.10893A>C* have no-historical recombination events in the population, four tagging SNPs - *g.9657C>T*, *c.10718G>C*, *c.10841G>A* and *c.10936G>C* were used in the association analysis. Except the SNP *c.10841G>A*, all other three SNPs were significantly associated with SFD (Table 4.1). The difference in SFD between two homozygous reached 0.12 inches at *g.9657C>T* ($P<0.01$), 0.10 inches at *c.10718G>C* ($P<0.001$) and 0.11 inches at *c.10936G>C* ($P<0.005$), respectively, which account for 0.56 – 0.67 standard deviation of the trait in the population. However, only one SNP *c.10936G>C* had a significant effect on BMS (Table 4.1). Animals with *CC* genotypes had 0.549 ($P<0.05$) and 0.399 ($P<0.05$) less marbling scores than animals with *GG* and *CG* genotypes, which account for 0.549 and 0.399 standard deviations for the trait, respectively.

4.4.6. Association analysis of haplotypes with SFD

As indicated above, three SNPs *g.9657C>T*, *c.10718G>C*, and *c.10936G>C* were associated with SFD in the reference population. Therefore, we decided to determine

how haplotypes affect the trait, but any haplotype with five or less animals was excluded in the analysis. Figure 4.5a shows an association plot of haplotypes between *c.10718G>C* and *c.10936G>C* with SFD measurements in inches. The haplotyp *GGGG* was 0.146 inches (P<0.05) higher than its *CCCC* counterpart. For the haplotypes between g.9657C>*T* with *c.10718G>C* (Table 4.5b), animals with CC*GG* had SFD value 0.102 inches (P<0.05) higher than animals with *TTCC* haplotype. Figure 4.5c shows an association plot of haplotypes between *g.9657C>T* and *c.10936G>C*. The same trend was observed with a difference of 0.177 inches between *CCGG* and *TTCC* haplotypes.

4.5. Discussion

Corticotrophin releasing hormone (CRH) released from the hypothalamus to the anterior pituitary under stress condition is known as a stressor for stimulating secretion of adrenocorticotrophic hormone (*ACTH*), which up-regulates the cortisol level. Cortisol, as the most important kind of glucorticoids in human, has some profound metabolic effects, such as stimulating gluconeogenesis (in the liver), inhibiting glucose uptake (in muscle and adipose tissue) and stimulating fat breakdown (in adipose tissue). Hence, a lot of research on *CRH* has broadened not only to stress-related studies but also to any metabolic diseases. The transgenic mice showed that over expression of *CRH* causes hair loss, muscle wasting, decreased linear growth and obesity (Senzel-Poore et al., 1992). These conditions were also observed in man and other animals with Cushing syndrome disease, which also caused by an increase of endogenous production of cortisol, with metabolic aberration, muscle wasting and obesity as some of the clinical symptoms. Furthermore, in pigs, significant associations of the *CRH* gene were found with back fat

thickness, carcass length, average daily gain and longissimus muscle area (Murani et al., 2006). In a Charolais-cross steers population, Buchanan and colleagues (2005) reported that three SNPs of the bovine *CRH* gene were highly associated with end-of-test rib eye area ($P < 0.034$) and hot carcass weight ($P < 0.0015$). In the present study, we demonstrated that the bovine *CRH* gene was significantly associated with marbling and SFD in a Wagyu x Limousin F_2 population. All these data indicate that *CRH* is a strong candidate gene for concordant QTLs related to body composition and energy metabolism.

In the present study, one SNP was detected in the promoter region of the bocine *CRH* with a transition of *cytosine* to *thymine* substitution (*g.9657C>T*). This mutation is located 138 bases from the putative transcriptional start site and forms a CpG site when allele C occurs (Figure 4.2). Many studies about CpG island have been done in the past with the majority of the result supporting that methylation on gene regulatory sites causes a severe suppression of the transcription activity. Some of the suggested mechanisms of modulating gene activity by methylation include inhibition of sequence-specific transcription factor binding region because of the alteration by methylated cytosine in the recognition sites, blockage of by some CpG binding protein (such as MeCP-1 and MeCP2) and alteration of chromatin structures (Kudo and Fukuda 1995). On the other hand, the *MatInspector* program revealed that the allele *9657T* eliminates four potential regulatory binding sites: neuron restrictive silencer factor (NRSF), E2F transcription factor, CDE-CHR binding factor-1 (CDF-1) and transcription factor CP2 (Figure 4.2). Cross species alignment (Figure 4.3) indicated that the proximal promoter region flanking the polymorphic site is highly conserved in nine mammalian species, suggesting

evolutionary importance of the region in the transcription regulatory sites of *CRH*. Moreover, the *MatInspector* analysis revealed that among four binding sites described above, only the neurons restrictive silencing factor (NRSF) is conserved among species. Seth et al. (2001) showed that NRSF was found in the first intronic region of *CRH* and it represses the gene expression through HDAC-dependent mechanism. However, NRSF also acts as an enhancer of the transcription activity. When a RE-1/NRSE region is either disrupted or deleted from the intronic region of *CRH*, a significant 1.2-2.5 fold up regulation was observed from the reporter activity (Seth et al., 2001).

However, we cannot exclude the involvement of other three regulatory elements in regulating of *CRH* expression in cattle. E2F is a heterodimeric protein that plays an important role in cell growth and apoptosis. Study in the methylated promoter region of genes that are regulated by E2F, such as in *dhfr, E2F1, cdc2, c-myb* and *c-myc*, showed that methylation can block the E2F elements (Campanero et al., 2000). Furthermore, previous study also showed that methylation of the first *cytosine* residue in *GCGC* motif of the E2F element, which is the case of the bovine *CRH* polymorphic site (Figure 4.2), blocks the binding of E2F protein from the cell extract to interfere with transcription activity of the gene. CDF-1 is a stereospecific transcription regulatory factor that recognizes two binding sites (CDE and CHR region). Inverting the position of CHR and CDE region abolishes the cell-cycle regulated repression, without affecting transcription, suggests that CDF-1 interacts with the activating domain to mediate repression (Zwicker et al., 1997). The functional role of CP-2 transcription factor was still unclear due to its ability to not only act as a ubiquitous transcription factor to most tissues but also it can be

involved in tissue or stage specific transcription of some genes (Kang et al., 2005). Nevertheless, the promoter polymorphism (*g.9657C>T*) leaves us a lot of physiological and functional questions to answer in future.

Four SNPs (*c.10718G>C*, *c.10841G>A*, *c.10893A>C* and *c.10936G>C*) were also detected in the exon 2 coding region of the bovine *CRH* gene. One mutation (*c.10893A>C*) results in silence mutation, where the remaing, *c.10718G>C*, *c.10841G>A*, and *c.10936G>C,* are missense mutations that leads to amino acid alterations from arginine to proline, from serine to asparagine and from aspartic acid to histidine, respectively. According to Majewski and Ott (2003), arginine, aspartic acid and histidine are the least mutable amino acids with the mutability of 0.365, 0.424 and 0.482, respectively. Among these three missense mutations, only *c.10841G>A* (serine to asparagines) was associated with neither marbling nor SFD in the reference population, but it impacts the secondary structure of the bovine *CRH* mRNA significantly (Figure 4.4). The *c.10718G>C* (arginine to praline) was significantly associated with SFD, while the *c.10936G>C* (aspartic acid to histidine) affected both SFD and marbling (Table 4.1). Therefore, it seems no evidence that there is any association between mRNA secondary structures and phenotypes. In addition, three SNPs (*g.9657C>T*, *c.10718G>C* and *c.10936G>C*) had also significant haplotype effects on SFD (Figure 4.5). Overall, our study confirmed that *CRH* is a strong candidate gene that regulates lipid metabolism in mammals. However, the center locations of QTLs detected in different experiments vary for both marbling and SFD on BTA14 (Casas et al., 2000 and 2003; Mizoshita et al.,

2004). This implies that other possible candidate genes on the chromosome might be involved in lipogenesis, which needs to be further explored.

4.6. References

Abecasis GR, Cardon LR, Cookson WO (2000). A general test of association for quantitative traits in nuclear families. Am J Hum Genet (66):279-92.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

Buchanan FC, Thue TD, Yu P, Winkelman-Sim DC. Single nucleotide polymorphisms in the corticotrophin-releasing hormone and pro-opiomelancortin genes are associated with growth and carcass yield in beef cattle. Anim Genet. 2005 Apr;36(2):127-31.

Campanero MR, Armstrong MI, Flemington EK. CpG methylation as a mechanism for the regulation of E2F activity. Proc Natl Acad Sci U S A. 2000 Jun 6;97(12):6481-6.

Casas E, White SN, Riley DG, Smith TP, Brenneman RA, Olson TA, Johnson DD, Coleman SW, Bennett GL, Chase CC Jr. Assessment of single nucleotide polymorphisms in genes residing on chromosomes 14 and 29 for association with carcass composition traits in Bos indicus cattle. J Anim Sci. 2005 Jan;83(1):13-9.

Jiang Z, Kunej T, Michal JJ, Gaskins CT, Reeves JJ, Busboom JR, Dovc P, Wright RW Jr. Significant associations of the mitochondrial transcription factor A promoter polymorphisms with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. Biochem Biophys Res Commun. 2005 Aug 26;334(2):516-23.

Kang HC, Chung BM, Chae JH, Yang SI, Kim CG, Kim CG. Identification and characterization of four novel peptide motifs that recognize distinct regions of the transcription factor CP2. FEBS J. 2005 Mar;272(5):1265-77.

Kudo S, Fukuda M. Tissue-specific transcriptional regulation of human leukosialin (CD43) gene is achieved by DNA methylation. J Biol Chem. 1995 Jun 2;270(22):13298-302.

Majewski J, Ott J. Amino acid substitutions in the human genome: evolutionary implications of single nucleotide polymorphisms. Gene. 2003 Feb 27;305(2):167-73.

Mizoshita K, Watanabe T, Hayashi H, Kubota C, Yamakuchi H, Todoroki J, Sugimoto Y. Quantitative trait loci analysis for growth and carcass traits in a half-sib family of purebred Japanese Black (Wagyu) cattle. J Anim Sci. 2004 Dec;82(12):3415-20.

Moore SS, Li C, Basarab J, Snelling WM, Kneeland J, Murdoch B, Hansen C, Benkel B. 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. 2003 Aug;81(8):1919-25.

Murani E, Muraniova M, Ponsuksili S, Schellander K, Wimmers K. Molecular characterization and evidencing of the porcine CRH gene as a functional-positional candidate for growth and body composition. Biochem Biophys Res Commun. 2006 Apr 7;342(2):394-405.

Perusse L, Rice T, Chagnon YC, Despres JP, Lemieux S, Roy S, Lacaille M, Ho-Kim MA, Chagnon M, Province MA, Rao DC, Bouchard C. A genome-wide scan for abdominal fat assessed by computed tomography in the Quebec Family Study. Diabetes. 2001 Mar; 50(3): 614-21.

Quandt K, Frech K, Karas H, Wingender E, Werner T (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23: 4878-48.

Rothwell NJ. Central effects of CRF on metabolism and energy balance. Neurosci Biobehav Rev. 1990 Fall;14 (3):263-71.

Seasholtz AF, Valverde RA, Denver RJ. Corticotropin-releasing hormone-binding protein: biochemistry and function from fishes to mammals. J Endocrinol. 2002 Oct;175(1):89-97.

Senzel-Poore MP, VA Cameron, J Vaughan, PE Sawchenko, W Vale (1992). Development of Cushing's Syndrome in Corticotropin-Releasing Factor Transgenic Mice. Endocrinology 130 (6): 3378.

Seth KA, Majzoub JA. Repressor element silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) can act as an enhancer as well as a repressor of corticotropin-releasing hormone gene transcription. J Biol Chem. 2001 Apr 27;276(17):13917-23.

Wood IA, Moser G, Burrell DL, Mengersen KL, Hetzel DJ. A meta-analytic assessment of a thyroglobulin marker for marbling in beef cattle. Genet Sel Evol. 2006 Sep-Oct;38(5):479-94.

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003 Jul 1;31(13):3406-15.

Zwicker J, Lucibello FC, Jerome V, Brusselbach S, Muller R. CDF-1-mediated repression of cell cycle genes targets a specific subset of transactivators. Nucleic Acids Res. 1997 Dec 15; 25(24):4926-32.

Table 4.1. Association analysis of the bovine *CRH* gene with SFD and BMS in Wagyu X Limousine F2 crosses.

Marker	Genotype	#Animals	SFD (in inches)			BMS (in inches)		
			LSM ± S.E.	P_{GLM}	P_{QTDT}	LSM ± S.E.	P_{GLM}	P_{QTDT}
g.9657C>T	CC	43	0.485 ± 0.024^{a}	0.001	0.0002	6.010 ± 0.147^{a}	0.669	0.026
	CT	107	0.396 ± 0.015^{b}			5.882 ± 0.093^{a}		
	TT	82	0.365 ± 0.017^{b}			5.809 ± 0.106^{a}		
c.10718G>C	CC	84	0.357 ± 0.017^{a}	0.002	0.0004	5.791±0.106	0.477	0.254
	CG	111	0.403 ± 0.015^{b}			5.939±0.092		
	GG	43	0.458 ± 0.024^{c}			5.973±0.148		
c.10841G>A	AG	36	0.441 ± 0.026	0.067	0.0653	5.921±0.160	0.847	0.846
	GG	204	0.388 ± 0.011			5.887±0.068		
c.10936G>C	CC	33	0.333 ± 0.028^{a}	0.002	0.0005	5.493 ± 0.168^{a}	0.022	0.009
	CG	118	0.383 ± 0.014^{a}			5.892 ± 0.088^{b}		
	GG	88	0.438 ± 0.017^{b}			6.042 ± 0.103^{b}		

* Means within a column with different superscripts are significantly different (P<0.01)

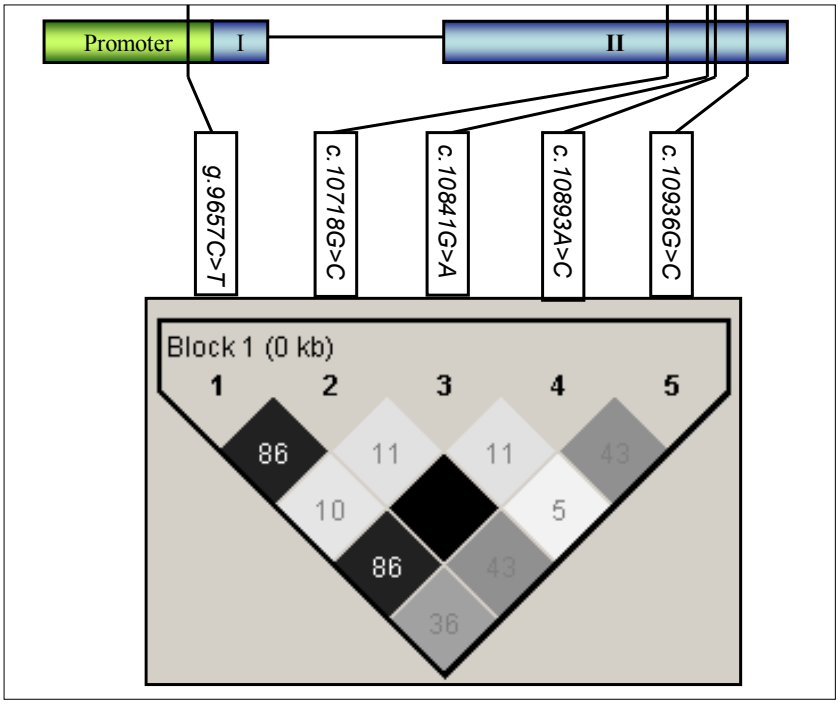

Figure 4.1. Genomic organization and haplotype analysis of the bovine CRH gene. Green bar: promoter region; blue bar: exons (I and II); straight line: intron. Pairwise linkage disequilibrium relationship for 5 mutations (g.9657 C>T, c.10718G>C, c10841 G>A, c. 10893 A>C and c.10936G>C) is illustrated based on r^2 measurements.

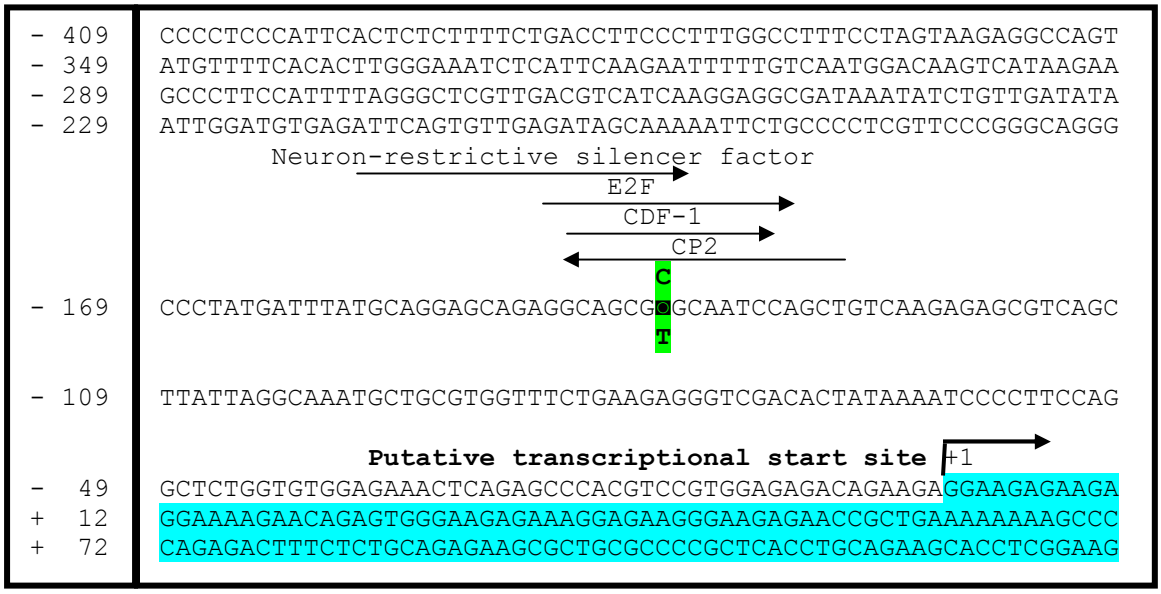

Figure 4.2. Nucleotide sequence of the proximal promoter region of the bovine CRH gene. The putative transcription start site is numbered as +1. The polymorphic site is bold and shadowed by green color and exon 1 is highlighted in turquoise color. Potential transcription regulatory biding sites for NRSF, E2F, CDF-1 and CP2 are associated with allele C only.

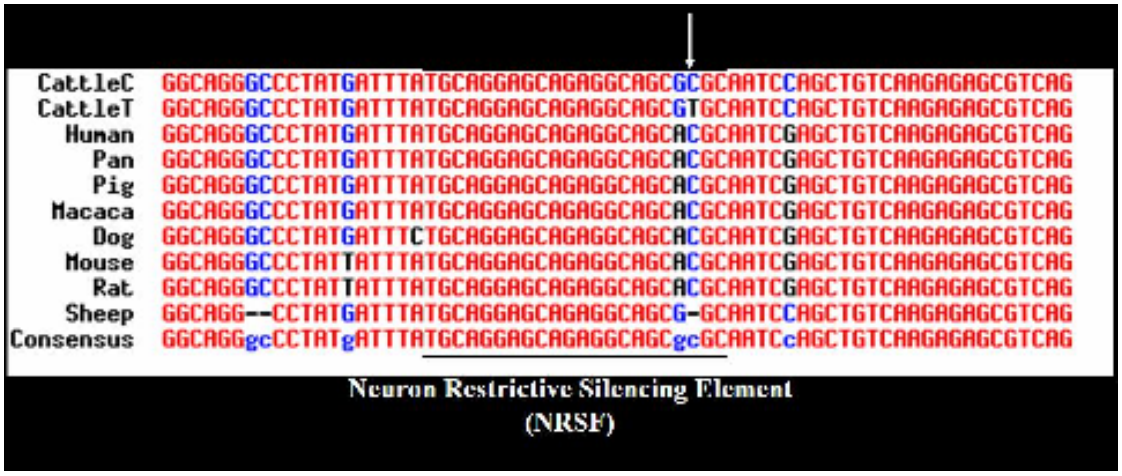

Figure 4.3. Sequence alignment of the CRH proximal promoter region among nine species with a conserved binding site for NRSF (boxed). The polymorphic site in the promoter of bovine CRH was detected (see arrow) with the allele T eliminating the conserved binding site.

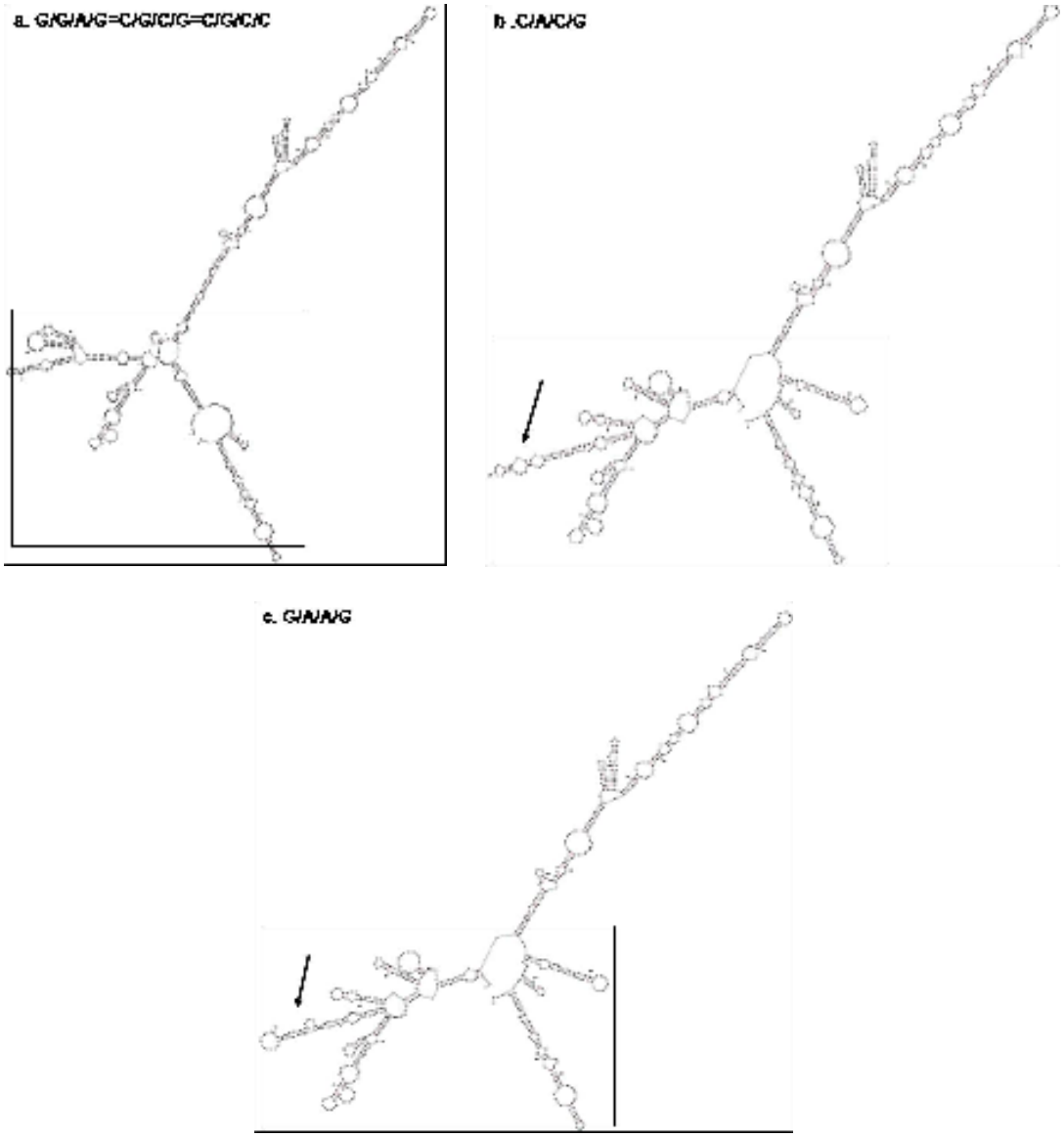

Figure 4.4. Predicted secondary mRNA structure of bovine CRH gene based on five haplotypes. a) represents GGAG, CGCG and CGCC, b) represents CACG and c) represents GAAG haplotype. The arrow shows the slightly different structure discovered between CACG and GAAG haplotypes.

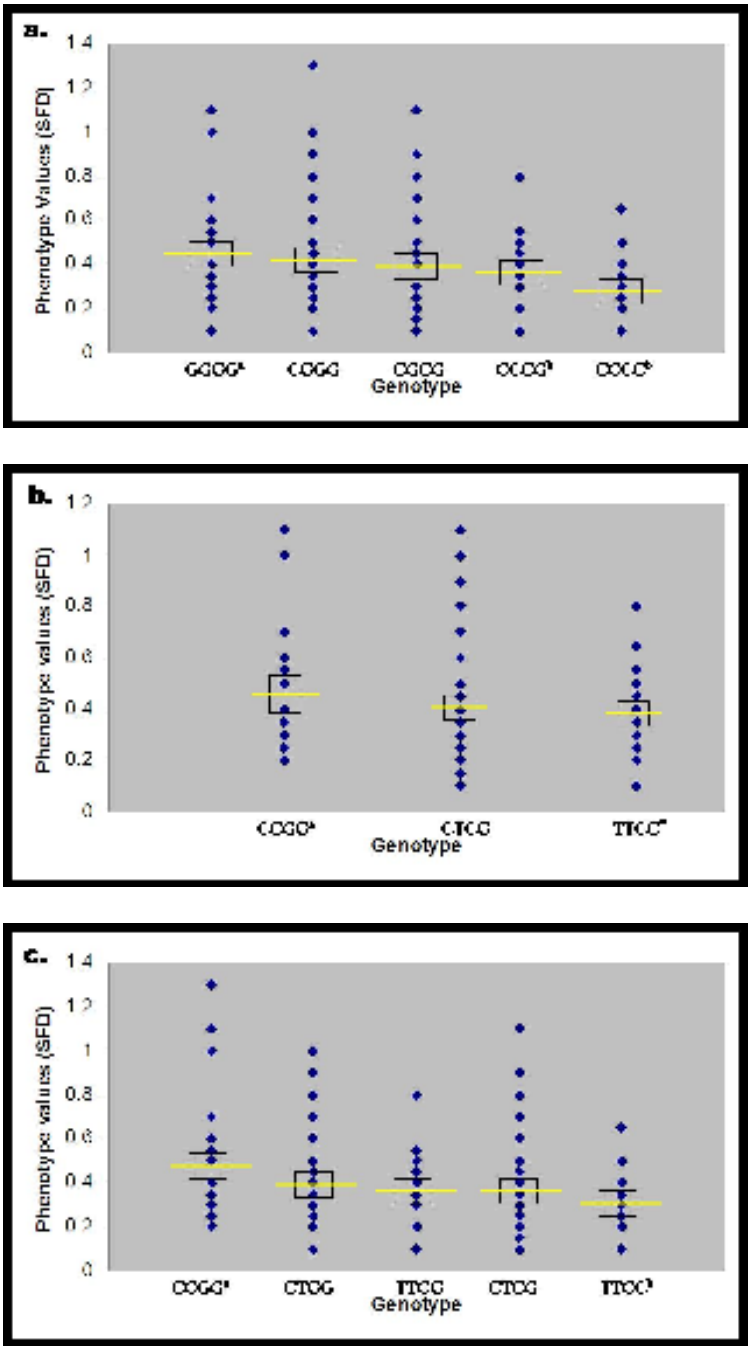

Figure 4.5. Association plot of haplotypes with SFD values (in inches). a) haplotypes between c.10718G>C and c.10936G>C, b) haplotypes between g.9657C>T and c.10718G>C and c) haplotypes between g.9657C>T and c.10936G>C. Different superscript shows significant differences (P values<0.05) between the two compared haplotypes.

Chapter Five:

CHROMOSOME WIDE ASSOCIATION ANALYSIS OF POTENTIAL CANDIDATE GENES ON BTA14 WITH ECONOMICALLY IMPORTANT TRAITS IN BEEF CATTLE

Chromosome Wide Association Analysis of Potential Candidate Genes on BTA14 with Economically Important Traits in Beef Cattle

Tito A. Wibowo[1], Charles T. Gaskins[1], Ruth C. Newberry[1], Gary H. Thorgaard[2], Jennifer J. Michal[1] and Zhihua Jiang[1]

[1]Department of Animal Sciences, Washington State University, Pullman, WA 99164 – 6351, USA and [2]School of Biological Sciences, Washington State University, Pullman, WA 99164- 4236, USA.

Correspondence to: **Dr. Zhihua Jiang**

Phone: 509 – 335 8761

Fax: 509 – 335 4246

Email: jiangz@wsu.edu

5.1. Abstract

Many genomic studies have been performed to identify genes underlying economically important traits in cattle. Here, both single gene association analysis and chromosome-wide analysis were performed on an F_2 Wagyu and Limousin cross population to explore potential candidate genes on bovine chromosome 14 that are responsible for carcass, meat and eating quality in beef cattle. Single gene association in combination with false discovery rate (FDR) and Bonferonni corrections tests revealed that five associations reached significance at 5% threshold. FDR technique, however, was less conservative because it added two more associations between CRH_23 and MUFA at 5% significance and between CRH_P and MUFA at 10% threshold. Furthermore, three QTL were revealed with chromosome wide permutation analysis for two traits: fatty acids related traits (PUFA) and carcass characteristics/quality (SFD). One QTL for PUFA was located at region 16-52 Mb with Fmax=2.42 at 52 Mb. Two QTL for SFD, furthermore, were located at region 14-24 Mb and 32-52 Mb with the highest probability revealed at 36Mb with F value=2.33.

5.2. Introduction

In recent years, much attention has been paid to identify genes that are associated with traits of direct interest in livestock industry. Using genome wide scan and positional or physiological candidate gene approaches many studies have been performed with a purpose to identify genes underlying economically important traits in cattle. As early as 1995, for instance, Georges and colleagues explored one of the first important QTL studies in cattle (Boichard et al 2003; Thaller et al 2003). The authors used a granddaughter design with only 159 microsatellites spanned about 2500 cM (Georges et al 1995). Certainly, this wide average region of 16 cM between markers has made it difficult to localize candidate genes in a given location (Farnir et al. 2004). With the progress made in genetic markers, nevertheless, the resolutions of chromosome mapping have been improved significantly based on both linkage and radiation hybrid (RH) mapping.

Since then, bovine chromosome 14 (BTA14) has been one of the most explored chromosomes for quantitative trait loci (QTL) or candidate genes in the bovine genome. There are more than 40 studies reported in both dairy and beef cattle. In dairy cattle alone, more than 50 QTL were discovered for milk production traits, such as fat yield/percentage, protein yield/percentage and milk yield. Recently, Wibowo and colleagues (2008) observed that most of these QTL are clustered in a region of 0-10 Mb on BTA14 anchored with the current bovine genome assembly (Btau_4.0). Two genes, *DGAT1* (0.444 Mb) and *CYP11B1* (1.294 Mb), were proposed as candidate genes for these QTL in dairy cattle. In particular, two independent studies conducted by Grissart et

al. (2002) and Winter et al. (2002) confirmed the association of *DGAT1* gene with fat-related trait in dairy cattle. The second candidate gene, *CYP11B1* was suggested later by De Roos et al. (2007) to be the causative gene for the same traits. In beef cattle, unfortunately, QTL studies were not done as extensively as in dairy population. Most of the QTL in beef cattle are located in a region 15-45 Mb on BTA14. Two candidate genes, *CRH* (30.476 Mb) and *FABP4* (41.955 Mb), were suggested for the QTL in this region by Wibowo et al. (2007) and Michal et al. (2006), respectively.

Although many candidate genes have been proposed as the underlying genes for QTL in both dairy and beef cattle, reports from different studies were still inconsistent. Moore and colleagues (2003), for instance, failed to confirm the association of polymorphisms in the *DGAT1* gene with any lipid metabolism traits (such as backfat EBV). Therefore, our goal here was to perform chromosome wide association analysis of potential candidate genes on BTA14 with carcass, meat quality and eating quality using an F_2 population of Wagyu and Limousin cross.

5.3. Materials and Methods

5.3.1. Reference Population and Animal Management

A Wagyu x Limousin reference population was generated jointly by Washington State University and the Fort Keogh Livestock and Range Research Laboratory, ARS, USDA. Eight Wagyu bulls were used to sire 121 F_1 cows with 108 Limousine cows over a-3 year period. Furthermore, for this project, six F_1 bulls were *inter se* mated with 113 F_1 dams without inbreeding to produce ~250 F_2 progeny. After weaning the calves at an average

age of 175 days, they were managed on diets in a two-phase system: a growing phase with 50-54% DM, 14.4-15.6% CP and 1.06-1.18 Mcal/kg net energy gain (NEg) and a finishing pahse with 68-70% DM, 11.6-13.4% CP and 1.26-1.31 Mcal/kg NEg. At the age of 450-641 days (avg= 561 days), the calves were transported to the abattoir, held overnight with water but without food and harvested the following morning using standard industry procedures.

5.3.2. Evaluation of phenotypic values

Hot carcass weight (CW) was determined immediately after harvest before rinsing/washing and chilling; marbling score (Marbling) was accessed subjectively, ranging from 4 = Slight^{0} to 9.5 = Moderately Abundant^{50}, from the interface of the 12-13th rib; fat thickness (Fat depth) was measured at the 12th rib interface perpendicular to the outside surface at a point three-fourths the length of the longissimus muscle from its chine bone; and the amount of kidney, pelvic and heart (KPH) fat was estimated as a percentage with respect of carcass weight. Sensory and shear force analysis were evaluated by a team at Colorado State University (CSU). About six to eight cores were used for the measurement and averaged to obtain a single shear force value (Warner-Bratzler shear force; WBSF) from each steak using an Instron testing machine; sensory traits were measured subjectively using trained panelists with an 8-point scale scoring for initial juiciness, myo-fibral tenderness (MFtender), connective tissue amount, overall tenderness (OVtender) and beef flavor intensity. Finally, fatty-acid analysis, which includes conjugated linoleic acid (CLA; per 100 g dry meat), saturated fatty acid (SFA), mono-unsaturated fatty acid (MUFA), poly-unsaturated fatty acid (PUFA), three indexes

of Δ^9 desaturase activity (R1= C14:1 to C14:0, R2= C16:1 to C16:0 and R3= C18:1 to C18:0) and cholesterol traits (CHOL; per 100g dry meat), were done at the University of Wyoming with procedures described by Rule et al. 2002.

5.3.3. Gene selection, annotation and primer design

Since BTA14 is homologous to a region of ~100 Mb from 46.1 Mb to 146.2 Mb on human chromosome 8 (HSA8), we integrated positional candidate genes based on the latest human obesity gene map (the 2005 update) to the bovine chromosome. Our comparative annotation procedures for retrieval of both cDNA and genomic DNA sequences for each bovine gene include three steps. In step 1, we used the cDNA sequences of the human orthologs as references for BLAST searches to retrieve the orthologous cDNA sequences against the GenBank database "nr" or the orthologous ESTs sequences against the GenBank database "est_others" with a species option limited to *Bos taurus*. The cDNA sequences in the "nr" database represents three categories: cDNA sequences derived from a full-length cDNA library, known gene cDNA sequences or annotated cDNA sequences compiled by the GenBank staff. In step 2, we chose the longest cDNA sequence retrieved from the "nr" database or one assembled from several ESTs retrieved from the "est_others" database to form a primary cDNA sequence for each cattle gene. This sequence was then used to perform a species-specific BLAST search against the "est_others" database in order to expand the primary sequence to a full-length cDNA sequence. For step 3, we used the full-length cDNA sequence to search for genomic DNA contigs in the 7.15X bovine genome sequence database (see the *Bovine Genome Resources* at NCBI). Such a process retrieved both cDNA and genomic

DNA sequences of each of potential candidate genes selected in the present study. The list of these selected candidate genes and their primers for SNP development is summarized in Supplemental Table 5.1.

5.3.4. SNPs detection and genotyping

SNPs detection was performed on six Wagyu x Limousin F_1 bulls using approximately 50 ng of genomic DNA amplified in a final volume of 10 μl that contained 12.5 ng of each primer, 150 μM dNTPs, 1.5 mM $MgCl_2$, 50 mM KCl, 20 mM Tris-HCl and 0.25U of Platinum Taq polymerase (Invitrogen, Carlsbad, CA). The following PCR conditions were used: 94^{o}C for 2 min, 35 cycles of 94^{o}C for 30 sec, 60^{o}C for 30 sec and 72^{o}C for 30 sec, followed by a final 5 min extension at 72^{o}C. PCR products were then sequenced on ABI 3730 sequencer in the Laboratory for Biotechnology and Bioanalysis (Washington State University) using a standard protocol. Polymorphisms were detected using Sequence Scanner (v1.0) program or standard sequence comparisons among these six F_1 bulls. A total of 12 SNPs were genotyped on all F2 animal using one of the three techniques: PCR-RFLP, PCR product directing sequencing and the Sequenom iPLEX assay using services provided by the Children's Hospital Oakland Research Institute, Oakland, California.

5.3.5.. SNPs association analysis and haplotype analysis

Association analysis were performed on a total of 19 traits described above: Shear, Juiciness, MFtender, OVtend, Flavor, R1, R2, R3, CLA, CW, REA, CHOL, KPH, SFA, MUFA, PUFA, Marbling, Fat Depth and SFD. A single gene analysis was performed

using a linear model (proc GLM; General Linear Model) on SAS v9.1 (SAS institute Inc., Gary, NC) that include sire, sex, year of culled and genotype as the fixed variables, and age (in days) as the regression variable. Two Pvalues, Pvalue<0.05 and Pvalue<0.1, were used as thresholds to determine the significance. Moreover, multiple comparisons were also explored with significance thresholds that were adjusted using two of the accepted methods: false discovery rate (FDR) (Benjamini et al 1995) and Bonferonni-modified test (Cheverud et al 2001). Furthermore, the degrees of Hardy-Weinberg equilibrium within each marker and linkage disequilibrium among different markers were estimated using HAPLOVIEW program (Barrett et al. 2005).

5.3.6. Chromosome analysis using QTLexpress

The discovered SNPs from the reference population spanned from 0 Mb to 52Mb on BTA14, the QTL express program (Seaton et al. 2002) was used for QTL detection with a half-sib method. One QTL at a time was detected for each of the 19 phenotypes with a permutation on chromosome-wide analysis mode. In order to perform the chromosome wide analysis, we calculated the distance between each marker using the actual physical distance mapped in the current bovine assembly. To simplify, whenever two or more SNPs were found within the same gene, the distance between the SNPs were assumed to be 0.0001 (or by 100 bases). Moreover, all the values for these phenotypes were adjusted using a linear model (proc GLM; General Linear Model) to perform chromosome wide analysis. This model assumed sire, sex and year of culled as the fixed variables, and age (in days) as the regression variable using analysis procedure of SAS v9.1 (SAS institute Inc., Gary, NC).

5.4. Results

5.4.1. SNPs and Haploview

HSA8 region 47.6-146.2 Mb was orthologous to BTA14 (Lyons et al. 1997, Everts-van der Wind et al. 2005, Itoh et al. 2005, and Marques et al. 2007). Based on the current finding on human gene obesity map, there were nine possible candidate genes/markers: *D8S1110, GPR7, CYP7A1, D8S1113, CRH, D8S2324, FABP4, GATA8B01* and *CBFAST1*. *CRH* and *FABP4*, in particular, were already annotated on bovine genome (Btau_3.1 and Btau_4.0) and were selected to be candidate genes as well on BTA14. Other markers, however, were not found or annotated on the bovine genome. Therefore, based on orthologous regions of these markers on BTA14, we integrated ten other positional candidate genes on BTA14: *CCNE2, CSPP1, CYP11B1, DGAT1, ENY2, LY6H, LYPLA1, MRPL15*, *MTFR1*, *MYC, RAB2* and *TG*. Results from genotyping the first six F_1 bulls revealed 12 SNPs: five SNPs on *CRH*, two SNPs on *FABP4*, and one SNP each for: *DGAT1, ENY2, MTFR1*, *RAB2*, and *TG*. Moreover, using HAPLOVIEW analysis, one SNP, from the two SNPs found on FABP4 gene, was not in Hardy-Weinberg equilibrium. Hence, it was eliminated from further analysis.

5.4.2. Association analysis

Our initial statistical analysis was based on a single gene association analysis, which revealed a total of 34 significant associations. Among them, five associations had a P value of <0.001: *RAB2* with CLA and CW, *MTFR1* with flavor, *CRH_109* with MUFA, *CRH_284* with MUFA; 12 associations reached a P value of <0.1: *TG* with BMS, KPH,

and Marbling, *RAB2* with REA, *MTFR1* with Juiciness and SFD, *CRH_P* with CHOL, *CRH_109* with CHOL and SFA, *CRH_232* with Marbling, *CRH_284* with CHOL and SFA; and the remaining17 associations had a P value of <0.05: *TG* with CW and R1, *MTFR1* with CW and Fat depth, *CRH_P* with CW, MUFA, PUFA and R1, *CRH_109* with R1, *CRH_232* with MUFA, R3 and SFA, *CRH_284* with R1, *FABP4* with KPH, MFtender and OVtender, and *ENY2* with REA (Figure 5.1), respectively. Three SNPs: *DGAT1, CRH_327* and one SNP at *FABP4* genes, were not significantly associated with any of the traits in our cattle population.

Furthermore, to reduce type I error, we adjusted the significance thresholds for multiple comparison tests with FDR and adjusted-Bonferonni. Both of these tests confirmed only five associations that reached the significance with a P value <0.05: MUFA with *CRH_109* and *CRH_284*, Flavor with *MTFR1,* and CW and CLA with *RAB2* genes (table 5.1). In addition, one association between MUFA and *CRH_232* was significant at 5% threshold by FDR test but only at 10% significance by Bonferonni correction test. FDR analysis also supported an association between *CRH_P* and MUFA at 10% threshold.

5.4.3. Chromosome wide analysis

As mentioned earlier, 11 polymorphisms were discovered in seven candidate genes. These genes were further mapped on BTA14 based on the current bovine assembly (Btau_4.0): *DGAT1* (at 0.444-0.446 Mb), *TG* (at 7.658-7.895 Mb), *RAB2* (at 26.047-26.0923 Mb), *MTFR1* (at 29.988-30.047 Mb), *CRH* (at 30.476-30.478 Mb), *FABP4* (at 41.955-41.959 Mb), and *ENY2* (at 52.878-52887 Mb). Three QTL were revealed for two

traits: fatty acids related traits (PUFA) and carcass characteristics/quality (SFD). One QTL for PUFA was located at region 16-52 Mb (Figure 5.2) with Fmax=2.42 at 52 cM (Table 5.2). Two QTL for SFD, furthermore, were located at region 14-24 and 32-52 Mb with the highest probability revealed at 36Mb, Fvalue=2.33 (Figure 5.2; Table 5.2).

5.5. Discussion

Many genomic maps, which include linkage and RH map, were developed for bovine genome. Linkage map that was pioneered by George et al. (1995) was improved by Barendse et al. 1997 and Ihara et al. 2005 using 746 DNA polymorphisms and 3802 microsatellites markers (3160 cM), respectively. In order to build a linkage map, however, a polymporphic or informative marker is required because linkage map is established based on the frequency of recombination event that occur between two markers. RH map, on the other hand, uses break induced radiation to estimate the physical distances between markers. Therefore, RH mapping is more efficient and effective to apply because it does not require polymorphic markers, and its resolution depends simply on the radiation dose used (Marques et al. 2007; Williams et al. 2002). Some studies of the bovine genome that utilized RH map include: 5000-rad by Womack et al. 1997, 12 K-rad by Rexroad et al. 2000, 3000-rad by Williams et al. 2002, 5000-rad by Everts-van der Wind et al. 2005, 7000-rad by Itoh et al. 2005, 3000-rad by McKay et al. 2006, 3000-rad by Jann et al. 2006, and 12K-rad by Marques et al. 2007. Nevertheless, since candidate gene approach was used to discover the SNPs for this project, neither of the methods discussed above was used to build our chromosome map. Each SNP found within a candidate gene was integrated on BTA14 based on the current built of bovine genome map (Btau_4.0). Therefore, we used an actual physical distance between

markers, instead of using arbitrary distance measurement. Moreover, SNPs that were discovered in our analysis did not have equal distribution throughout the chromosome. Five SNPs alone, for example, were discovered within the *CRH* gene. Therefore, relying on linkage analysis to build our chromosome map might not be appropriate.

According to our chromosome wide analysis, there were three suggested QTL for PUFA and SFD in our reference population. One QTL with significant effect on PUFA was located in the region of 16-52 Mb with maximum Fstatistic at 52 Mb. Since consumers have become more health conscious, they are more concern about the health-related issued of the meat that they eat. Therefore, a high proportion of PUFA and low fat content are more desirable since these traits showed incidence to reduce obesity and cardiovascular diseases (Hocquette et al. 2006). QTL, that we discovered for PUFA in the region of 16-52 MB, harbors many for economically important traits in beef cattle (Casas et al. 2003; Kim et al. 2003; Kneeland et al. 2004; Miyata et al. 2004; Mizoshita et al. 2004; and Moore et al 2004). Nonetheless, no candidate gene has been suggested on BTA14 for this trait. Two possible candidate genes, however, were located within this region, *CRH* at 30.6 Mb and *FABP4* at 41.9 Mb. Based on multiple comparison analysis, neither *CRH* nor *FABP4* was associated with PUFA. Less conservative statistical method of single gene association technique, however, revealed one SNP on promoter region of *CRH* gene that was significantly associated ($P<0.05$) with PUFA. Two QTL, furthermore, showed effects on SFD at region 14-24 and 32-52 Mb. QTL at region 14-24 Mb seemed to agree with the QTL previously discovered by Casas et al. (2000) for fat thickness, without proposed candidate gene. According to our single gene association analysis,

however, one candidate gene *MTFR1*, which was located within the region, was significantly associated with SFD ($P<0.05$).The second QTL for SFD contained the location at which the Fstatistic reach maximum at 36 Mb. This QTL finding confirmed the previously reported *CRH* and *FABP4* gene as candidate genes again for SFD based on reports from Wibowo et al. (2007) and Michal et al. (2006), respectively.

5.6. Reference

Barendse W, Vaiman D, Kemp SJ, Sugimoto Y, Armitage SM, Williams JL, Eggen A, Agaba M, Aleyasin SA, Band M, Bishop MD, Buitkamp J, Byrne K, Collins F, Cooper L, Coppettiers W, Denys B, Drinkwater RD, Easterday K, Elduque C, Ennis S, Erhardt G, Ferretti L, Flavin N, Gao Q, Georges M, Gurung R, Harlizius B, Hawkins G, Hetzel J, Hirano T, Hulme D, Jorengensen C, Kessler M, Kirkpatrick BW, Konfortov B, Kostia S, Kuhn C, Lenstra JA, Leveziel H, Lewin HA, Leyhe B, Lil L, Burriel IM, McGraw RA, Miller JR, Moody DE, Moore SS, Nakane S, Nijman IF, Olsaker I, Pomp D, Rando A, Ron M, Shalom A, Teale AJ, Thieven U, Urquhart BGD, Vage DI, Van de Weghe A, Varvio S, Velmala R, Vilkki J, Weikard R, Woodside C, Womack JE, Zanoti M, Zaragoza P. (1997). A medium-density genetic linkage map of the bovine genome. Mammalian Genome 98: 21-28.

Barendse WJ, 1999. Assessing lipid metabolism. International patent application number PCT/AU98/00882. International publication No. WO 99/23248.

Barrett JC, Fry B, Maller J, Daly MJ, 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-5.

Benjamini Y and Hochberg Y, 1995. Controlling the false discovery rate in multiple testing under independence. Ann. Stat. 29: 1165-1188.

Boichard D, Grohs C, Bourgeouis F, Cerqueira F, Faugeras R, Neau A, Rupp R, Amigues Y, Boscher MY, Leveziel H (2003). Detection of genes influencing economic traits in three French dairy cattle breeds. Genet. Sel. Evol (35): 77-101.

Cheverud JM, 2001. A simple correction for multiple comparisons in interval mapping genome scans. Heredity 87: 52-58.

Everts-van der Wind A, Larkin DM, Green CA, Elliott JS, Olmstead CA, Chiu R, Schein JE, Marra MA, Womack JE, Lewin HA, 2005. A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. Proc. Natl. Acad. Sci. 102 (51): 18526-18531.

Farnir F, Grissart B, Coppieters W, Riquet J, Berzi P, Cambisano N, Karim L, Mni M, Moisio S, Simon P, Wagenaar D, Vilkki J, Georges M, 2002. Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbreed half-sib pedigrees: revising the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. Genet 161: 275-287.

Georges M, Nielsen D, Mackinnon M, Mishra A, Okimoto R, Pasquino AT, Sargeant LS, Sorensen A, Steele MR, Zhao X, Womack JE, Hoeschele I, 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genet 139: 907-920.

Grisart B, 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. Genom. Res. 12: 222-231.

Hocquette JF. Renand G, Leveziel H, Picard B, Cassar-Malek I, 2006. The potential benefits of genetics and genomics to improve beef quality-a review. Anim. Sci. Papers and Reports (24) 3: 173-189.

Itoh T, Watanabe T, Ihara N, Mariani P, Beattie, CW, Sugimoto Y, Takasuga A, 2005. A comprehensive radiation hybrid map of the bovine genome comprising 5593 loci. Genom 85: 413-424.

Jann OC et al, 2006. A second generation radiation hybrid map to aid the assembly of the bovine genome sequence. BMC Genomics 7: 283.

Kaupe B, 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.

McKay SD, 2006. Construction of bovine whole-genome radiation hybrid and linkage maps using high-throughput genotyping. Anim. Genet 38: 120-125.

Marques E, Givry SD, Stothard P, Murdoch B, Wang Z, Womack J, Moore SS, 2007. A high-resolution radiation hybrid map of bovine chromosome 14 identifies scaffold rearrangement in the latest bovine assembly. BMS Genom 8: 254.

Michal JJ, Zhang, ZW, Gaskins CT, Jiang Z, 2006. The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu X Limousine F2 crosses. Anim Genet 37(4): 400-402.

Moore SS, Li C, Basarab J, Snelling WM, 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: 1919-1925.

Thaller G, Kramer W, Kaupe B, Erhardt G, Fries R, 2003. Effects of DGAT1 variants on milk production traits in German cattle breeds. J. Anim. Sci 81: 1911-1918.

Williams JL, Eggen A, Ferretty L, Farr CJ, Gautier M, Amati G, Ball G, Caramorr T, Critcher R, Costa S, Hextall P, Hills D, Jeulin A, Kiguwa S, Ross O, Smith AL, Saunier K, Urquhart B, Waddington S, 2002. A bovine whole-genome radiation hybrid panel and outline map. Mamm. Genom. 13: 469-474.

Winter A, 2002. Association of 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. USA (99): 9300-9305.

Womack JR, Johnson JS, Owens EK, Rexroad III CE, Schlapfer J, Yang YP, 1997. A whole-genome radiation hybrid panel for bovine gene mapping. Mamm. Genom. 8: 854-856.

Wibowo TA, Michal JJ, Jiang Z, 2007. Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. Genome 50: 939-945.

Table 5.1. Association analysis for 4 genes: *RAB2, MTFR1,* and *CRH* with CLA, CW, Flavor, and MUFA that was significant at 5% chromosome-wide threshold using FDR analysis and Bonferonni correction tests. * indicates association analysis that significant only at 10% threshold.

No	Phenotype	Pvalue	Gene/Marker	Genotype	LS means ± SE
1	**CLA**	<0.0001	*RAB2*	AA	62.5728±4.747[a]
				AT	46.0787±2.075[b]
				TT	45.9397±1.8089[b]
2	**CW**	0.0398	*RAB2*	AA	698.5765±22.7352
				AT	668.2642±10.1207[a]
				TT	713.6076±8.6896[b]
3	**Flavor**	0.0141	*MTFR1*	CC	5.5571±0.09294[a]
				CG	5.7988±0.04506[b]
				GG	5.5746±0.03639[a]
4	**MUFA**	0.0225	*CRH_109*	CC	50.594±0.224[a]
				CG	49.7543±0.1978[b]
				GG	50.5159±0.3385[a]
		0.0225	*CRH_284*	AA	50.5159±0.3385[a]
				AC	49.7453±0.1978[b]
				CC	50.594±0.224[a]
		0.0056	*CRH_232**	AG	49.4051±0.3538
				GG	50.3215±0.1522
		0.0584	*CRH_P**	CC	50.2655±0.3339
				CT	49.7897±0.2025[a]
				TT	50.6269±0.2288[b]

Table 5.2. Statistical profiles and QTL locations for sensory traits, fatty acids-related traits, and carcass traits in a Wagyu x Limousin cattle population.

Traits	Position	F	LR	R RSS	Mean
R1	0Mb	0.99	5.89	10391.229	-2.3801
Marbling	0Mb	1.60	9.45	389.3523	0.1343
Mftender	8Mb	1.19	7.08	106.9244	0.0138
Ctissue	8Mb	1.05	6.26	25.0083	-0.116
Ovtender	8Mb	0.49	2.96	16024.328	2.8327
Flavor	8Mb	1.76	10.4	21.5944	-0.1419
REA	8Mb	0.46	2.78	522.1887	0.0591
KPH	8Mb	0.21	1.25	6032.109	-0.7051
MUFA	8Mb	0.15	0.91	8805.423	1.1067
CW	24Mb	1.27	7.54	1421005.2	13.2663
R3	28Mb	0.77	4.62	322101.7	9.6753
SFA	28Mb	0.45	2.73	877.7277	-0.2999
CHOL	32Mb	0.68	4.08	422187.62	-2.1069
SFD	36Mb	2.33	13.6	4.8592	-0.108
Shear	36Mb	1.39	8.26	85.7266	0.1752
Juiciness	48Mb	1.08	6.45	43.1857	-0.023
R2	52Mb	1.00	6	582.4025	0.3923
CLA	52Mb	1.26	7.51	48292.562	5.0352
PUFA	52Mb	2.42	14.12	239.9283	-0.2981

Supplemental Table 5.1. Primer lists based on the preliminary selected candidate genes on BTA14.

Gene	Location (Mb)	Ampl. region	Primers	Temp.	Size
CCNE2	67.782	3'UTR	F: GGCGAAAAACTTAAGTACCCTTTCA	62.84	528
			R: TTAGTGGCGTTTTCTTTTCAGACAG		
		5'UTR	F: CAGAAAAGATCCCATCTGCTTCC	63.41	548
			R: CTCTCCTCTCTTCCCCTCCTTAACT		
		Promoter	F: CACTACCAAATTCACTCACGGAGAC	63.25	516
			R: AGATAATTTGTTTCCCTCCCAGGTT		
CRH	30.476	Exon2	F: AGTTCTGTCTAGGCGCTCCCTACC	63	819
			R: GGGTCTGTGGGTGTCGTCCT		
		Promoter	F: AAAAATAAACATGGTATCAGAGCAATG	63	563
			R: CCCCTCCCATTCACTCTCTTTTCT		
CSPP1	31.119	5'UTR	F: GCTCTAGGGTGTATGGCCTTGTTAG	63.42	533
			R: TGTGAGGTACAAAACACCACAGAAG		
		Promoter	F: ATTGTTCAGTTGCTAAGTCGTGTCC	62.86	594
			R: CTTTTGCCTAACCTGAGATCACAAA		
		Promoter	F: CATAGTCATTGCTGCTGCTAAGTCA	62.9	443
			R: CTCGCTAATGCTTTGTACGTTCAGT		
CYP11B1	1.294	3'UTR	F: TAGTGTTTTCAAGGTCTCACCTCCA	63.33	500
			R: CAGCAACTCTACACAATCCAGAACA		
		3'UTR	F: CCTGACTTTTGAGTCACACCACTGT	63.91	587
			R: GAGTATCATTGACCCTGATGTCCAG		
		5'UTR	F: GAAGAGAGACGAACATACCCAAGGT	63.11	562
			R: CATGTCCAAGTGCATGTTCTCAG		
DGAT1	0.444		F: TGGGCTCCGTGCTGGCCCTGATGGTCTA	80	
			R: TTGAGCTCGTAGCACAGGGTGGGGGCGA		
ENY2	52.877	Promoter	F: TCAAAAATGCAATCTCTGGGCTCT	63	548
			R: GACTGTGACTTTCCTCACGGTCAA		
FABP4	41.955	Exon3	F: ATATAGTCCATAGGGTGGCAAAGA	55	598
			R: AACCTCTCTTTGAATTCTCCATTCT		
LY6H	1.451	3'UTR	F: AAGCGGCACTTTTTCTCAGACTATC	63.33	501
			R: GTCTCCAGGCTCAGAAGTGATTTTA		
		5'UTR	F: GCTTCTTTTCGCAGTAAACGTTCTT	63.25	667
			R: GTAGGTGTGTCCGGAGAGGAAAG		
		Promoter	F: CCAGCAGGATGAGTAATCTGAGAGA	63.25	636
			R: TCTACACCTTTGGGTCAGTTGGATA		
LYPLA1	21.846	3'UTR	F: AATGATGGACATCAAGCAGTTCATT	63.33	501
			R: ACGGATACTGCAATGGTGACTATGT		
		Promoter	F: TATATCCTTTAGGGCTTCCCAGATG	63.23	536

			R: GGTCAGCATCTGTATGATTTTGGTC		
MRPL15	21.87	3'UTR	F: ATTCTGAAACCTACGGATGAGAAGC	63.12	565
			R: ATCTTGCCAGAAAACACTGCAATAC		
		5'UTR	F: ACACACAGTGACACCAAGTCCTACA	63.08	593
			R: TTACTCCTCTAAACCCGAACCTGAG		
		Promoter	F: CATTAACAGCAAGTGAAGTCAGCAA	62.78	571
			R: AAGAGTCACTGGAAAAGACCCTGAT		
MTFR1	29.988	Promoter	ATGAAGCCAAGTGTGAATGAAAGGA	63	501
			CAGAGAGGACGAATGCTTGAAAGAG		
MYC	12.151	3'UTR	F: GGTTCCCTCTGTCAACTCCTTACAT	63	540
			R: GAAAAGCCCTCACCCTTTCTTAAAT		
		5'UTR	F: CGTTTGGCAGATCCTATTCTCCA	64.58	506
			R: CTCCTCCTCGTCGCAGTAGAAATAA		
		Promoter	F: CCACTTCCCCAATAAATCTAACGTC	63	501
			R: AGGAATCACTGGAGAGAAGAGTGGT		
RAB2	26.047	Exon1	F: GGCCTGAAAATGAGTAGGAAGTGG	63	548
			R: GGAAGTGTTCCTGCACTCTTGCTT		
TG	7.658	Promoter	F: GTGAAAATCTTGTGGAGGCTGTA	63	211
			R: GGGGATGACTACGAGTATGACTG		

	DGAT1	TG	RAB2	MTFR1	CRH_P	CRH_109	CRH_232	CRH_284	CRH_327	FABP4	FABP_1	ENY2
BMS		*										
CHOL					*	*		*				
CLA			***									
CW		**	***	**	**							
Flavor				***								
Fat depth				**								
Juiciness				*								
KPH		*								**		
Marbling		*					*					
Mftender										**		
MUFA					**	***	**	***				
Ovtender										**		
PUFA					**							
R1		**			**	**		**				
R3							**					
REA			*									**
SFA						*	**	*				
SFD				*								

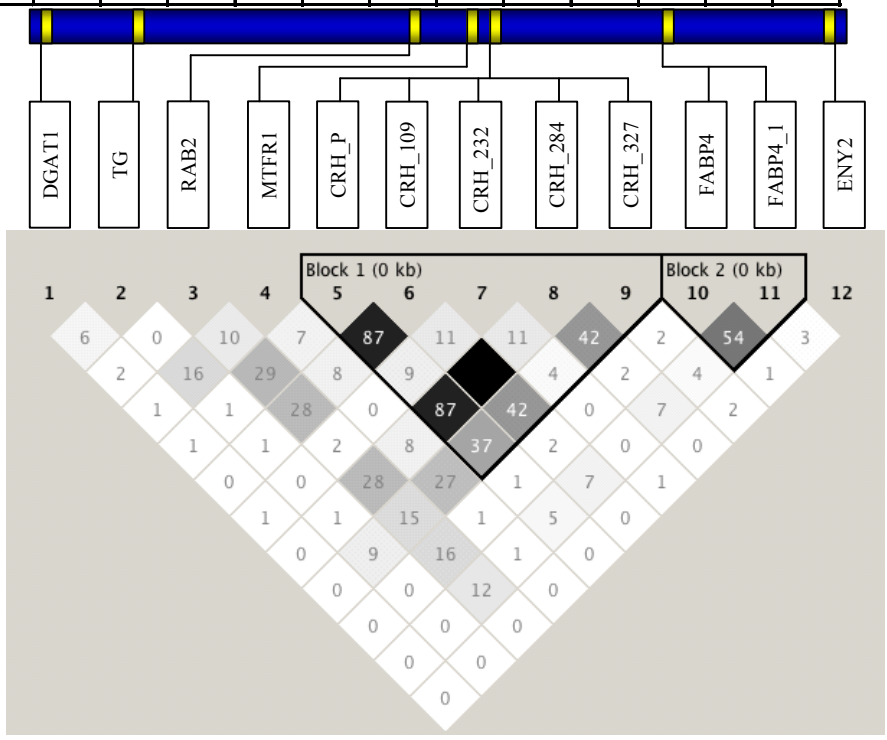

Figure 5.1. Single gene association analysis and HAPLOVIEW profile from 12 SNPs and 18 phenotypes.

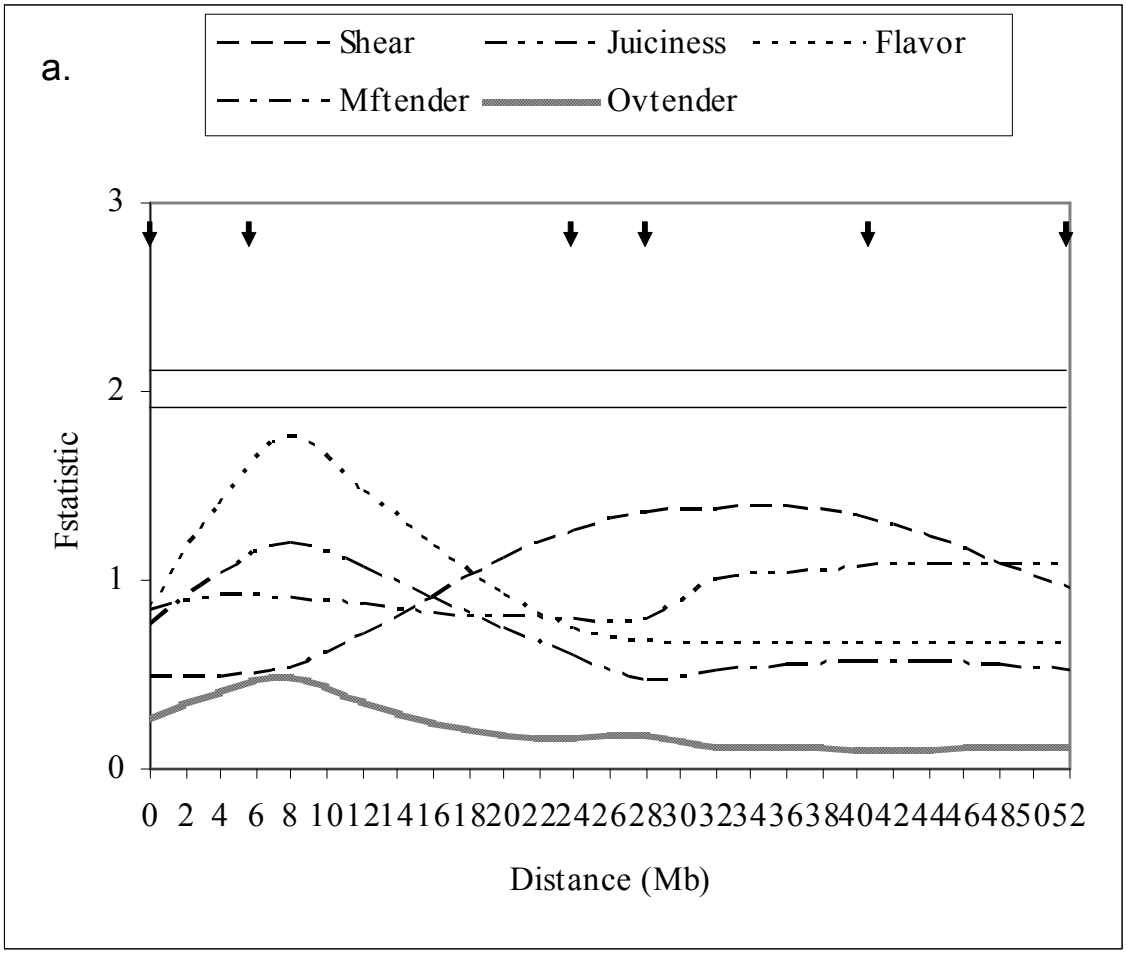

Figure 5.2. BTA14 Fstatistic profile for 3 categorized traits: a. sensory traits (shear, juiciness, flavor, MFtender, and OVtender), b. fatty acids-related traits (R1, R2, R3, CHOL, CLA, SFA, MUFA and PUFA), and c. carcass traits (SFD, Ctissue, CW, REA, KPH and marbling) using six different markers (shown as arrows). There were two chromosome wide significance thresholds of 2.18 and 1.86 as horizontal lines.

Figure 5.2. *Continue*

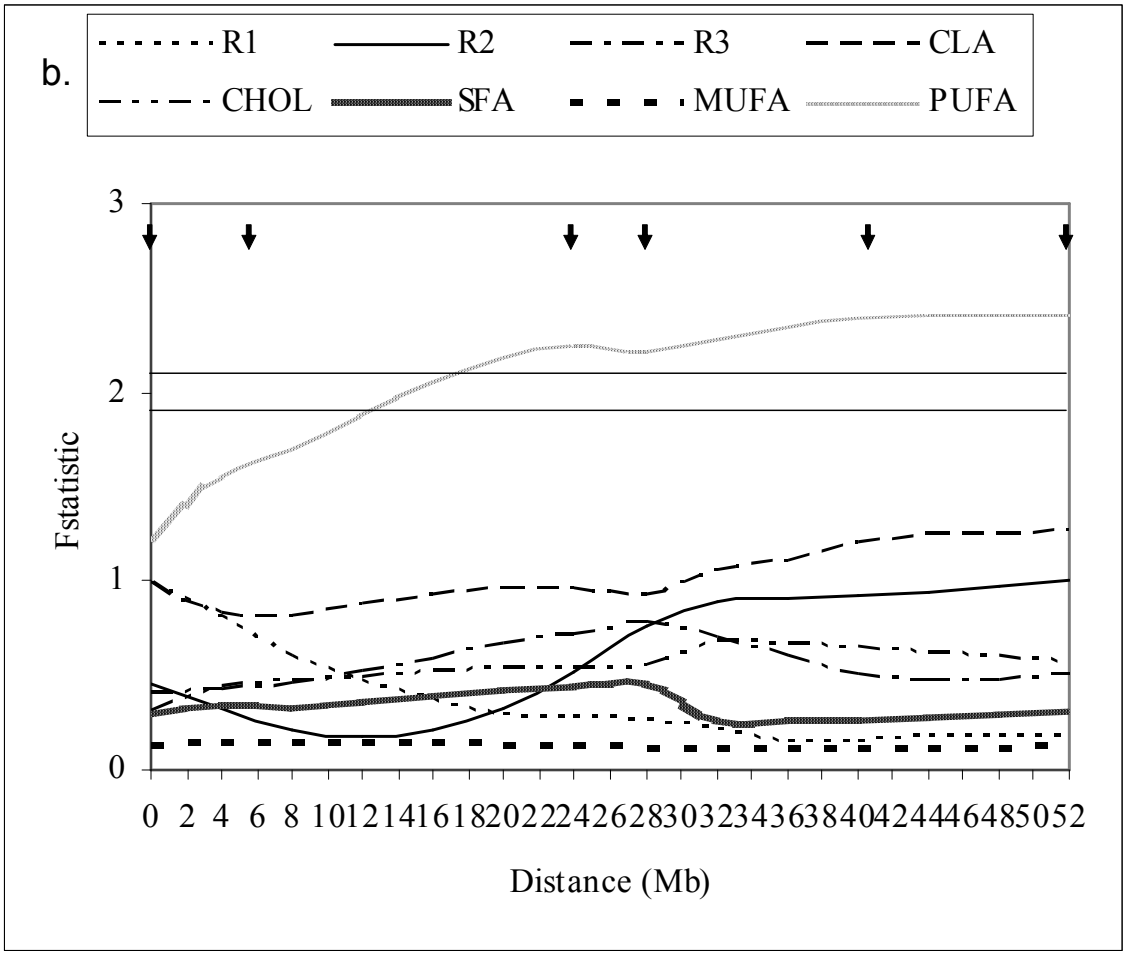

Figure 5.2. *Continue*