IDENTIFICATION OF QUANTITATIVE TRAIT LOCI LINKED MARKERS AND

CHARACTERIZATION OF POSITIONAL CANDIDATE GENES FOR BEEF

MARBLING IN WAGYU x LIMOUSIN F_2 CROSSES

By

QIANJUN XIAO

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

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Animal Sciences

MAY 2006

To the Faculty of Washington State University:

   The members of the Committee appointed to examine the dissertation of
QIANJUN XIAO find it satisfactory and recommend that it be accepted.

_____

Chair

_____

_____

_____

ACKNOWLEDGMENT

Education is not the end, but a beginning of a career. At this point, I would like to thank Dr. Zhihua Jiang, my major advisor, for his instructive and patient guidance through my PhD program. Dr. Jiang, I am grateful that you introduced me to this interesting project. You were very patient in instruction and always ready to help. You not only encouraged me to work hard on this project, but you yourself often worked overtime in your office. I did learn a lot from you during my years in your lab. Honestly speaking, I could not have asked for a better major advisor. My appreciation also goes to all the members of my committee, Drs. Jerry Reeves, Gary Thorgaard, and Jan Busboom. It was such a great pleasure to learn from them and have discussions with them. Dr. Reeves, thank you for your patience and support for the parentage test. I also learned a lot of beef cattle production from you and I really enjoyed your inspirational lectures in your beef cattle production class. Dr. Thorgaard, thank you for your kindness and special support for my AFLP technique either theoretically or in the lab. Dr. Busboom, thank you for your nice help of all those beautiful meat pictures and knowledge. I am so proud that there are so many outstanding professors at Washington State University. I would also thank Jennifer Michal for her quick and kind lab assistance. Without her help, I couldn't imagine I would finish my project at this time.

I would like to thank my parents Hongman and Yuyun from the depth of my heart. Thank you for always trusting me and making me believe that I can go further. I am so lucky to have such wonderful parents as you. Special thanks go to my husband Xiao-Lin and my

IDENTIFICATION OF QUANTITATIVE TRAIT LOCI LINKED MARKERS AND

CHARACTERIZATION OF POSITIONAL CANDIDATE GENES FOR BEEF

MARBLING IN WAGYU x LIMOUSIN F$_2$ CROSSES

Abstract

By Qianjun Xiao, Ph.D.
Washington State University
May 2006

Chair: Zhihua Jiang

Selection for marbling has been recognized as an important objective for the production of high quality beef worldwide. The objective of this study is to identify genes or markers that are associated with beef marbling. In experiment 1, we presented a simplified approach for screening and mapping of QTL linked markers for beef marbling using a Wagyu x Limousin F$_2$ reference population. This simplified approach involves integration of the amplified fragment length polymorphism (AFLP) with DNA pooling and selective genotyping and comparative bioinformatics tools. AFLP analysis yielded four AFLP markers. Sequencing and *in silico* characterization assigned two of these AFLP markers to bovine chromosomes 13 (BTA13) and 1 (BTA1), which are orthologous to human chromosomes HSA10p11.23 and HSA21q22.2 with both regions harboring QTL for obesity-related phenotypes. Both AFLP markers showed significantly large additive genetic effects on beef marbling score (BMS). Overall, this approach is less expensive and less time consuming than current QTL mapping approaches. In

particular, this approach is suitable for screening and mapping QTL linked markers when targeting one or a few complex traits. In experiment 2, we identified one candidate gene of poly (A) polymerase associated domain containing 1 (*PAPD1*) associated with beef marbling. This gene is in close proximity to the human gene *KIAA1462* which is orthologous to the AFLP marker located on BTA13 revealed in experiment 1. *PAPD1* gene is a newly identified mitochondrial gene that encodes polymerase required for the polyadenylation and stability of mammalian mitochondrial mRNAs. Ten genetic markers were detected in the promoter and exon 1 region of *PAPD1* gene. Among seven markers assayed on 246 Wagyu x Limousin F_2 animals, two single nucleotide polymorphisms (SNPs) in the promoter region were significantly associated with BMS. However, there was a significant interaction between a third SNP, which causes amino acid changes derived from coding exon 1, and each of these two promoter SNPs on BMS. In particular, the differences between double heterozygous animals and other genotypes animals were from 0.67 to 2.3 standard deviations for the trait in both cases. Our study provides evidence for a new mechanism – the compound heterosis involved in extreme obesity, which warrants further examination.

TABLE OF CONTENTS   Pages

LIST OF TABLES

LIST OF FIGURES

CHAPTER THREE

CHAPTER FOUR

APPENDIX

Dedication

    This dissertation is dedicated to my parents Hongman and Yuyun, my husband Xiao-lin and my daughter Linda (Keyi), who is always my solid and positive support and encouragement while I pursued my educational goals.

CHAPTER ONE

Introduction

CHAPTER 1
INTRODUCTION

There are two major considerations in beef quality grading: maturity and marbling. Maturity is an estimation of the physiological age of the carcass, whereas marbling describes the appearance of white flecks or streaks of fatty tissue between the muscle fibers in meat. Although marbling only slightly affects meat tenderness, it contributes substantially to the palatability traits of juiciness and flavor, since highly marbled beef has an extremely low melting point, giving it a more acceptable mouth feel. Beef marbling is identified as intramuscular fat in a cross section of the longissimus muscle (Cameron *et al.* 1994).

Beef grading systems have been established to correspond to market requirement in several countries, such as USA, Canada, Japan and Australia. Higher marbling scores usually result in grades that receive high prices. Therefore, selection for high marbling and overall quality has been an important breeding objective worldwide. As a quantitative trait, however, beef marbling is difficult to predict. Selection of beef marbling requires tremendous effort, expense and time. If using the selection via progeny test, for example, progeny of a given individual animal must be raised to market weight, its carcass evaluated for these traits, and the data analyzed before any conclusions can be made about the individual's genetic potential to deposit intramuscular fat. This process takes a minimum of three to four years and it increases genetic lag and prediction error. In recent years, molecular genetic techniques have been developed rapidly, which

facilitate identifying and utilizing genes that contribute to the genetic variation (also referred to quantitative trait locus, QTL) of marbling.

The amplified fragment length polymorphism (AFLP) technique is a polymerase chain reaction (PCR)-based DNA fingerprinting method that allows the analysis of multiple DNA fragments of different length in a single reaction. It is estimated that a single enzyme combination will permit the amplification of 100,000 unique AFLP fragments, thus revealing a large number of restriction fragment polymorphisms. The AFLP technique can, therefore, be used to perform whole genome scanning of QTL linked markers, and thus improve our ability to map QTL and estimate their effects and position. On the other hand, selective DNA pooling not only reduces the number of samples required for genotyping but it also reduces the number of markers for which genotyping need be carried out in order to obtain a certain power for detection of a linked QTL. The objective of this study is to use a strategy that combines an AFLP technique with a selective DNA pooling strategy to identify QTL linked genes or markers that are responsible for genetic variation of beef marbling.

This objective is accomplished by the following specific aims: 1) detecting and characterizing QTL closely linked markers for marbling by selective DNA pooling and AFLP techniques; 2) identifying genes or haplotypes those are responsible for marbling within identified regions. Our long term goal is to develop a QTL based marker-assisted selection (Q-MAS) strategy to produce high marbling beef for the USA markets as well as for exports. This should significantly benefit beef cattle farmers. Ultimately, the

present research could contribute to cloning of QTLs underlying genetic variation of beef

marbling.

CHAPTER TWO

Literature Review

CHAPTER 2
LITERATURE REVIEW

2.1 **Beef Marbling**

2.1.1 Measurement of beef marbling

According to USDA, graders evaluate the amount and distribution of marbling in the

ribeye muscle at the cut surface after the carcass has been ribbed between the 12th and

13th ribs ([www.ams.**usda**.gov/lsg/stand/standards/**beef**-car.pdf](www.ams.usda.gov/lsg/stand/standards/beef-car.pdf)) (Figure 1).
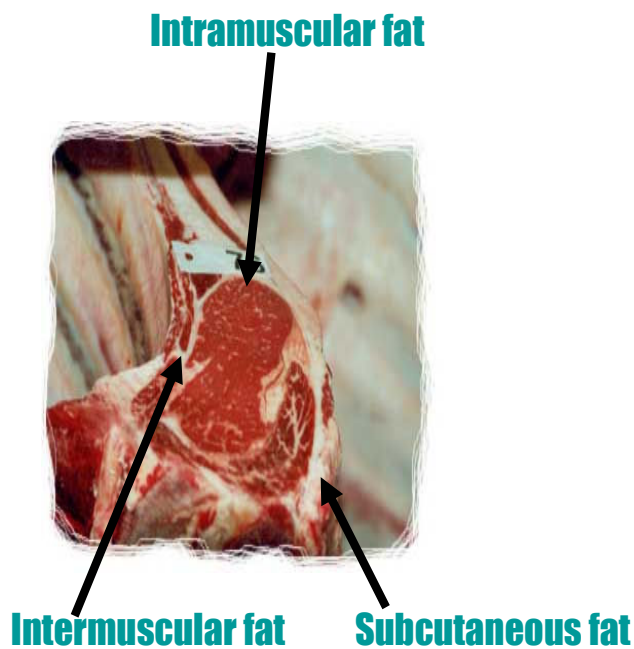
Figure 1 Three different types of fat: intramuscular, intermuscular, and subcutaneous fat

2.1.2 Grades of beef marbling

There are two types of beef grades in the United States: quality grades (expected

palatability or eating satisfaction of the meat) and yield grades (estimates of the

percentage of boneless, closely trimmed retail cuts from the round, loin, rib and chuck)

(Burson, 1997). The USDA Beef Quality Grades are Prime, Choice, Select, Standard,

Commercial, Utility, Cutter, and Canner. Degree of marbling is the primary determinant

of quality grade. Nine of the USDA Degrees of marbling are show in Figure 2 (Devoid

and Practically Devoid are not shown). Table 1 illustrates how degree of marbling and

degrees of maturaty are combined to determine USDA Beef Quality Grades.

Table 1. Relationship between marbling, maturity and carcass quality grade

Relationship between marbling, maturing and carcass quality grade

Very Abundant	Abundant	Moderately Abundant
Slightly Abundant	Moderate	Modest
Small	Slight	Traces

Figure 2. Marbling degrees standard

Beef marbling grading systems in other countries often differ from the US system. In Japan, for example, the beef marbling grading system is more precise, consisting of 12 scores described in Table 2 and illustrated in Figure 3 (JMGA. 1988). In Canada, beef marbling scores has similar system to USA (Kazala *et al*., 1999).

BMS # 1 Quality Grade 1   BMS # 2 Quality Grade 2   BMS # 3 Quality Grade 3

BMS # 4 Quality Grade 3   BMS # 5 Quality Grade 4   BMS # 6 Quality Grade 4

BMS # 7 Quality Grade 4   BMS # 8 Quality Grade 5   BMS # 9 Quality Grade 5

BMS # 10 Quality Grade 4   BMS # 11 Quality Grade 5   BMS # 12 Quality Grade 5
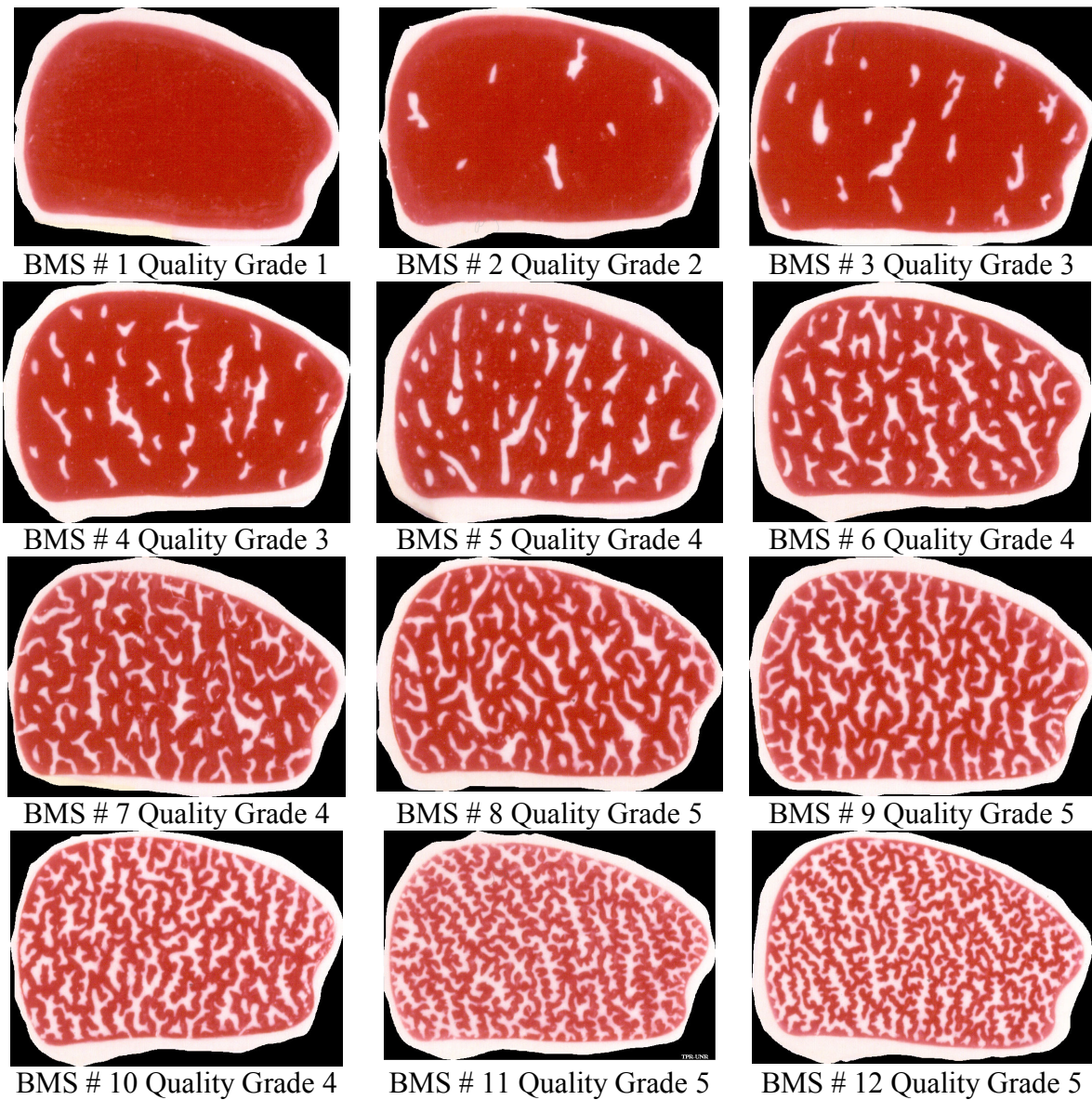
Figure 3 Beef Marbling Standard (Japan)

Table 2 Equivalence of U.S. and Japanese Marbling Scores

U.S.D.A. Quality Grade	U.S.D.A. Marbling Score	BMS Number	Japanese Quality Grade
	Extremely Abundant 50+	11 or 12	5
	Extremely Abundant 0-49	10	5

	Very Abundant 50-99	9	5
	Very Abundant 0-49	8	5
	Abundant	7	4
	Moderately Abundant	6	4
Prime	Slightly Abundant	5	3
	Moderate	4	3
	Modest	3	3
Choice	Small		
Select	Slight		
Standard	Traces		

2.2 Molecular Markers

2.2.1 General review of molecular markers

The availability of various genetic markers makes it possible to detect genetic variation of quantitative traits (Dekkers *et al*., 2002). Genetic markers may represent "signs" for genes affecting the phenotype of interest, when they are closely linked to or located in the proximity of causal genes. Through selecting these markers/genes, one may select the desired phenotype.

Before the development of PCR, there were already a number of techniques to study the genetic variations in population, such as chromosome polymorphisms, morphologically detectable polymorphisms, protein electrophoresis, and restriction fragment length polymorphism (RFLP) (Weeks *et al*., 2000). Among these, the most attention was focused on biochemical genetics based on protein polymorphism analysis, because almost every protein had genetic variations (Cavalli-Sforza and Feldman, 2003). But these techniques often had limitations or drawbacks and could not be used on many organisms.

After the development of PCR techniques in 1986 (Mullis *et al*., 1986), modern molecular genetic markers based on DNA polymorphisms became broadly used. There are three main types of DNA polymorphism: SNPs (single nucleotide polymorphisms), Indel (deletion/insertion), and Repeats (tandemly repeated DNA: microsatellites, minisatellites, and satellites). So far, these molecular markers have been widely used in studies of genetic variations such as population structure, relatedness among individuals, phylogenetic relationships and genomic mapping (Weeks *et al.*, 2000; Collard *et al*., 2005). With the development of automated DNA sequencing in the early 1990's, molecular markers have been used in almost every organism from bacteria (Goulding *et al*., 2000; Kassama, *et al.* 2002), to plants (Becker, *et al.*, 1995), to animals (Vignal, *et al.,* 2002), and to human (Cavalli-Sforza, *et al.* 2003).

Today, there are many techniques to produce molecular genetic DNA markers that can be used for identifying/mapping QTLs. In terms of the band information provided at a single locus, these techniques could be categorized into three types: the bi-allelic dominant which includes RAPDs (random amplification of polymorphic DNA) and AFLPs (amplified fragment length polymorphism); the bi-allelic co-dominant which includes RFLPs (restriction fragment length polymorphism) and SSCPs (single stranded conformation polymorphism), and the multi-allelic co-dominant which includes the microsatellites. On the other hand, these techniques could be divided into two groups based on the method: hybridization-based techniques and PCR-based techniques. Hybridization-based techniques consist of RFLP and DNA fingerprinting. These are relatively old methods because they need a lot of DNA and they are time-consuming.

PCR-based techniques are relatively new methods and include PCR-RFLP, PCR-SSCP, PCR-RAPD, PCR-AFLP, etc. These techniques are relatively inexpensive and need little DNA.

Among those PCR-based techniques, co-dominant genetic markers were thought to be better than those of dominant, since they provide more information in allelic segregation (Sunnucks *et al.,* 2000). However, those co-dominant markers, such as RFLPs, SSCPs, and microsatellites, usually need prior sequence information and are more costly, which inhibited their use to some organisms. Dominant markers, such as RAPD and AFLP, do not need prior sequence information and could be used in almost every organism. But RAPD uses random primers which have less reproducibility. Thus, AFLP is a relatively reliable and lowcost technique.

2.2.2 AFLP

AFLP technique is a molecular fingerprinting method, which was developed by Vos *et al*. (1995). Based on the restriction enzymes and the PCR technique, AFLP technique could yield numerous DNA polymorphisms. The basic procedure of AFLP consists of three steps:1) digestion of genomic DNA with two restriction enzymes (one rare cutter and one frequent cutter) and ligate the adapters to digested DNA fragments; 2) Pre-amplification using primers with one selective nucleotide; 3) Selective amplification using primers with more than one selective nucleotides. Then, it comes to visualize the bands and conduct the data analysis. Potentially, the AFLP technique simultaneously screens the whole

organism genome restriction site and primer selective nucleotides variation and produces almost an unlimited number of genetic markers.

There are many advantages of AFLP technique. The AFLP method is a highly efficient way to generate a large number of genetic markers and it does not require prior sequence information. It also doesn't need a large amount of DNA. AFLP is a powerful molecular tool to generate and identify genetic markers for phylogeny studies, for genome mapping and for quantitative trait loci detection. Moreover, with the usage of PCR, AFLP is a relatively fast technique. In contrast to RAPD, AFLP is reproducible. Once the laboratory procedure is worked out, data could be produced for different species by using exactly the same reagents and conditions. However, there are still several disadvantages of AFLP technique.

The main disadvantage of AFLP was that it generates dominant markers and precise genotypes cannot be assessed directly in diploid or polyploidy organisms, because heterozygotes cannot be distinguished from homozygotes (Weeks *et al.*, 2000). Nevertheless, it is possible to distinguish a heterozygote from a homozygote with the development of automatic sequencing systems. The AFLP dominant band could be converted into simple single locus markers and then combined the co-dominant band information, which could be used for large-scale screening in industry (Brugmans *et al.,* 2003; Sasazaki *et al.* 2004). Additionally, the selection of enzyme and primer combinations in the AFLP technique is somewhat labor-consuming. An improvement to this technique is called post selective fluorescent AFLP (fAFLP), which would be more efficient, inexpensive and less multiplexed (Vatcher *et al.,* 2002). Software packages

(such as *GeneMapper*) are available to analyze large datasets generated by the AFLP technique.

The AFLP technique was broadly utilized on lots of aspects: phylogeny studies (biodiversity, identification of closely linked DNA markers) (Buntjer *et al*., 2002; Nijman *et al*., 2003), genome mapping (construction of genetic maps or physical maps) (Gorni *et al.,* 2004), and linked-marker screening (QTL mapping) (Tsuji *et al*., 2004).

AFLP was used for genome-wide screening and was considered theoretically to be able to identify all genes responsible for variation in quantitative traits especially those economically important traits such as beef marbling score, milk yield production, etc. Tsuji *et al.* (2004) used AFLP genotyped forty-eight cows with extreme genetic merit for beef marbling score (25 highest and 23 lowest) from a population of 4462 Japanese Black cattle. Eighteen polymorphic fragments had significantly different frequencies between the high and low beef marbling score cows. Seven AFLP markers derived from this study and were considered effective markers to discriminate high and low beef marbling score cattle, which could be used for marker-assisted selective breeding.

2.2.3 SNP

SNP stands for single nucleotide polymorphism (Brooks, 1999). It is a nucleotide site of a given sequence for which substitution polymorphism has been observed at a significant frequency (>1% of the population) between different individuals. Several commonly used techniques such as various RFLP and AFLP often reveal SNPs, but not at a detailed molecular level, and not necessary in distinction from small sequence insertions or

14

deletions (indel). True SNPs are specific base-pair variants, and they are abundant in most genomes (Kendal, 2003). The initial sequence information can be specially produced for SNP discovery by sequencing a given fragment in different genotypes. The SNP genotyping techniques detect sequence polymorphisms with a specific PCR fragment (STS), and they make it possible to visualize all the possible single nucleotide changes within the fragment. The basic principles that are used to distinguish between the possible nucleotides at the SNP sites are based on: (1) oligonucleotide hybridization, (2) template-dependent DNA strand elongation by a DNA polymerase, (3) double-strand-dependent ligation, (4) mismatch detection. So far, many laboratory approaches have been developed explicitly to screen large numbers of SNPs in well-characterized model genomes (Winzeler *et al.*, 1998; Wang *et al.*, 1998).

Alternatively, SNPs can also be a by-product of large EST sequencing programs performed with cDNA libraries from different genotypes. In the latter, the discovery of significant SNPs can be carried out completely *in silico* from existing databases. Then, the SNPs are located in transcribed sequences, which is of interest for high-throughput mapping of genes, for example in candidate gene approaches. In principle, SNPs provide a wellspring of molecular markers that are well suited to genomic analysis such as studies of linkage and linkage disequilibrium (Setphen *et al.* 2001). In animals, SNPs have already been used in QTL mapping experiments (Smith *et al.*, 2000; Moore *et al.*, 2003; Li *et al.*, 2004).

2.2.4 Microsatellite

Microsatellites, also called SSR (simple sequence repeats), are tandem repetitions of mono-, di-, tri-, or tetranucleotide units. They are very abundant in eukaryotes, but their distribution in genome is considered nonrandom across coding and noncoding regions (Li et al., 2002). SSR is not only a powerful molecular marker source but also functionally affect gene transcription, translation, chromation organization, cell cycle, etc. (Li et al., 2002).

PCR is used to reveal codominant polymorphic locus-specific microsatellite markers. However, the development of microsatellite markers is difficult. First, a genomic library for the target species needs to be constructed and clones containing microsatellite repeats are screened in this library. Second, positive clones are sequenced and the flanking sequences are obtained to design primers for the unique microsatellite repeats. Finally, a large number of individuals can be screened for microsatellite alleles by using the designed primers (Ciofi *et al*., 1998; Goldstein *et al*., 1999). Recently, Zane and colleagues (2002) presented a relatively fast and effective protocol to isolate microsatellites: fast isolation by AFLP of sequences containing repeats (FIASCO), which is an effective way to screen for microsatellites.

Microsatellites are broadly used to construct linkage maps which are substantially for QTL analysis (Rohrer *et al*., 1996; McCouch et al., 1997). With more and more microsatellite markers available, QTL mapping is more precise than using a single marker, since multiple microsatellite markers have low recombination rates with the QTL (Ashwell *et al*., 1997; Haley 1995). To date, PCR based Microsatellites techniques were

utilized to construct genome maps in most livestock species because the advantages of its abundance in the genome, the specificity of the primers, and its high degree of polymorphisms (Bishop *et al.*, 1995).

2.3 Mapping/identifying QTL for beef marbling

QTL is the location of a gene or genes that affect a trait measured on a quantitative scale, such as growth rate, carcass traits, milk yield, etc. Thus, the task of QTL mapping is to localize these genes, estimate their effect and determine their genetic mode. Mapping/detecting QTL requires three essential stages: 1) Collection of accurate phenotypic data within properly developed/existing pedigree/populations; 2) Collection of accurate genotypic data (DNA/genetic markers) within pedigrees; 3) Statistical analysis correlating phenotypic and genotypic data, reflecting pedigree organization and structure.

2.3.1 QTL analysis in experimental crosses

There are basically two types of experimental crosses that are used in QTL mapping: inbred crosses and outbred crosses. Next, we briefly discuss QTL analysis using the two types of experimental crosses below.

QTL analysis in line-crosses By crossing two inbred lines, linkage disequilibrium is created between loci that differ between the lines, and this in turn creates association between marker loci and linked segregation QTLs. This is the idea behind using marker information to map and characterize QTLs in inbred crosses. There are commonly two types of line-cross populations for generating disequilibrium, F2 vs. backcross

17

populations. The F2 design examines marker-trait associations in the progeny from a cross of F1s, while the backcross design examines marker-trait association in the progeny formed by backcrossing the F1 to one of the parent lines. More complex designs are also considered where individuals are genotyped in one population, while trait values are scored in a future population derived from the genotyped individuals. For example, it is possible to genotype and cross the F2 individuals and estimate the trait value associated with a genotyped individual as the mean value of the resulting F3 family. This is called the F2:3 design. Theoretically, scoring the phenotype as the mean of several individuals, as opposed to measurement of a single individual, can offer increased power over a standard F2 design by reducing the sampling variance. Designs combining information from multiple crosses are starting to be considered since it is expected to be more powerful than those involving single crosses.

In view of the unit of marker analysis used, marker-trait association can be assessed using one-, two-, or multiple-locus marker genotypes. In a single-marker analysis, marker-trait association test is performed for each locus, independent of information of all others. Single marker analysis is generally a good choice when the goal is simply to detect the QTL, but not to estimate precisely its position and effects. Interval mapping is also referred to as flanking-marker analysis, in which a separate analysis is performed for each pair of adjacent marker loci. Interval mapping offers a further increase in power of detection and more precise estimation of QTL effects and position. As an improvement over interval mapping, composite interval mapping (Zeng, 1993, 1994; Jansen, 1993, 1994) considers a marker interval plus a few other well-chosen single markers as cofactors in each analysis. Further, it is possible to consider all of the linked markers on a

chromosome simultaneously (Kearsey and Hyne, 1994; Hyne and Kearsey, 1995; Wu and Li, 1994), which is referred to as multipoint mapping.

QTL analysis in outbred crosses In large animals, however, it is more practical to identify QTL using outbred crosses, rather than the luxury manipulation of inbred crosses. While QTL detected in inbred-line crosses usually represent fixed differences between lines, QTL detected in outbred populations are responsible for within-population variation. This is a fundamental distinction between QTL mapping using inbred crosses and outbred crosses. This explains why a QTL effect in outbred crosses is often detectable by its variance.

There are a variety of designs using oubred crosses, such as the sib family design and grand-daughter design. The Sibship-based method for detecting linked QTL historically started with single sib families. For example, animal breeders typically rely on a few large half-sib families, each resulting from a single sire (often via artificial insemination). In the sib analysis, one can use family data to search for QTL by comparing offspring carrying alternative marker alleles from the same parent. The sib design has now been extended to several sib families. In the granddaughter design (Weller *et al*. 1990), each sire produces a number of half-sib sons that are scored for the marker genotypes. The character value for each son is determined by progeny testing, with the trait value being scored in a large number of daughters (again half-sibs) from each son. This design was developed for milk-production characters in dairy cows, where the offspring were

19

granddaughters of the original sires. More general designs allowing for both full- and half-sib families are examined by van der Beck *et al*. (1995).

2.3.2 Selective Genotyping and DNA Pooling

In many cases, quantitative traits are less expensive to score than marker genotypes. Therefore, DNA pooling and selective genotyping (Lander *et al*., 1989; Darvasi *et al*., 1992) strategies have been developed, in order to reduce the cost in QTL mapping experiments. In selective genotyping, the whole population phenotypes are collected and only those extreme phenotype individuals that are on two tails of the phenotype distribution are genotyped. The main idea of selective genotyping is that the distributions of the two genotypes are quite similar close to the means, but very different in the tails of the distributions. Therefore, the individuals with the extreme phenotypic values represent the most information of QTL. According to Darvasi *et al*. (1992), there is no need to genotype more than the upper and lower 25% of the population. This strategy is most effective if there is only one or two quantitative traits are of interest. Sample pooling is also a way to detect QTL (Plotsky *et al*., 1993; Lipkin *et al*., 1998). In sample pooling, those samples with similar phenotype data are assumed to share the same marker alleles and thus pooled into one pool (Pareek *et al*., 2002). Sample pooling should be applied together with selective genotyping (named selective DNA pooling), which was considered the most useful method if the number of quantitative traits of interest is relatively low (Weller, 2001). The procedure of selective DNA pooling technique is described below. Firstly, phenotype data are collected and ranked. Secondly, those individuals in upper and lower tail of the phenotype distribution were selected and equal

amounts of DNA from each individual in the same tail weres pooled. Finally, DNA pools are genotyped for the marker(s) having frequency differences between pools. Some caution of this technique is the accuracy of the amount of DNA and the stutter bands by PCR. Generally speaking, this is an effective approach for QTL detection and has been widely used in humans, animals, and plants (Mosig *et al.*, 2001; Baro *et al.*, 2001., Pareek *et al.*, 2002).

2.3.3 Candidate Gene Approaches

In some cases, there may be sufficient information to suspect that certain known loci influence character expression. It is therefore possible to directly test for population–level association between trait value and specific alleles at such candidate loci. In principle, candidate gene approaches facilitate discovering and localizing causative genes for quantitative traits (Campbell *et al.* 2003). Advantages of the candidate gene analysis also include its relative robustness to genetic heterogeneity and the ability to detect small QTL effect sizes (Craddock *et al.*, 2001).

Then, how are candidate loci chosen? One obvious approach is to consider loci with known biological actions involved in the development or physiology of the trait of interest, which are referred to functional candidate genes. Alternatively, genes can be chosen in the neighborhood of previously identified QTLs, which are called positional candidate genes. Often, a candidate gene can be both functional and positional. For example, *DGAT1* was a functional and positional candidate gene for milk fat content (Grisart *et al.*, 2002; Wu *et al.* 2005). Once a candidate gene is chosen, its

polymorphisms within selected candidate genes can be tested for association with the variation in quantitative traits.

In recent years, comparative genomics provides a very promising tool for the candidate gene approach, which takes advantages of genome sequencing and conservation in well mapped/studied species such as human, mouse, cattle and other mammalian species to study those species having less genomic information such as pig, chicken or even wild animals, and vice versa. For example, 228 out of 1321 porcine microsatellites were identified to have human orthologs by *in sillico* comparative mapping (Jiang *et al*., 2003). The conservation ratio is low here since noncoding regions are less conserved than coding regions, and microsatellites are in noncoding regions. It was observed that map position of homologous genes is tending to be conserved in a related species (Nadeau *et al*., 1998). Retrieval of the same gene sequences in the target species or orthologous sequences in other species can provide genes structure and expression information. For example, Jiang and colleagues took advantage of human ESTs which have known functions to explore and annotate orthologous cattle ESTs (Jiang *et al*., 2004). Comparative genomics also contributes on QTL mapping. The middle region of bovine chromosome 6 (BTA6) dwell a number of different QTLs affecting milk production, functional and comformation traits, growth and body composition traits. Recently, Weikard and colleagues constructed a high-resolution radiation hybrid (RH) map for this region by linking densely spaced bovine markers and genes to the annotated human genome sequence (Weikard *et al*., 2006). There were 43 new loci identified.

2.3.4 Mapping QTL for beef marbling

Many experiments have been conducted in search for QTL affecting beef marbling using whole or partial genome scanning. MacNeil and Grosz (2002) reported one significant QTL for marbling on BTA2 and two suggestive QTL for marbling on BTA16 and BTA29 in two backcross families by using microsatellite markers genome scanning. One significant QTL for marbling was revealed on BTA23 and probably six suggestive QTL for marbling on bovine chromosome 3 (BTA3), BTA5, BTA9, BTA10, BTA14 and BTA27, respectively (Cacas *et al*., 2003). There was also a significant marbling score QTL on BTA2 plus two suggestive QTL intervals on BTA3 and BTA16 were presented by Casas *et al*. (2004). Those QTLs for marbling on BTA2 and BTA14 were confirmed in similar locations in an Angus population (Taylor and Schnabel, 2004) (<http://animalgenomics.missouri.edu/>). One significant QTL for marbling was showed on BTA6 in Korean Hanwoo cattle (Yeo *et al*., 2004). Significant QTL for marbling were also reported on BTA2, BTA9, BTA14, BTA21 and BTA24 in Japanese black cattle (Kobayashi *et al*., 2004; Imai *et al*., 2004 and Mizoguchi *et al*., 2004).

The candidate gene approach also played an important role in identifying QTLs for beef marbling. For examples, allele "3" of a polymorphism in the 5' promotor region of the Thyroglobulin (*TG*) gene was associated with higher marbling score (Barendse *et al*., 1999); mutations in the leptin gene (*LEP*) caused beef cattle to reach slaughter weight sooner with more marbling (Buchanan *et al*., 2002); a nonconservative K232A substitution in the diacylglycerol O-acyltranferase gene (*DGAT1*) affected intramuscular fat deposition (marbling) in beef (Thaller *et al*. 2003). Additionally, the mitochondrial

transcription factor A (*TFAM*) were found to be associated with beef marbling and backfat (Jiang *et al*., 2005).

As a quantitative trait, the marbling trait could be genetically controlled by a multiple-gene system, and some important genes have not been identified yet. Both genome scan and candidate approach will continue to contribute to identifying markers or genes significantly associated with beef marbling, and toward understanding of the genetic architecture of the beef marbling trait.

2.4 Toward the goals

2.4.1 QTL-based Marker Assisted Selection (Q-MAS)

The goal of animal breeding is to develop improved genotypes to produce a desired superior phenotype. For some traits, the process of selection is hard because of low heritability, or difficulty of measurement, etc. These traits include disease resistance, carcass yield and quality, fertility and reproductive efficiency, milk production and growth performance. In such situations, marker-assisted selection (MAS) would accelerate the selection. Q-MAS is the selection program utilizing the information of genetic markers associated with favorable combinations of QTLs and thus increase the rate of genetic gain.

However, not all identified markers or genes associated with favorable QTLs could be used eventually in MAS. Experimental identified QTLs need to be confirmed to be real and segregating in the breeding population before they can be used in the Q-MAS (Spelman *et al*., 1997). Davis and DeNise discussed the three phases of MAS program

(Davis *et al*., 1998). The first is the detection phase, in which DNA polymorphisms are used to detect the association of specific genes /markers with QTL in a segregating population. The second is the evaluation phase, in which the identified genes/markers are tested in target populations to confirm the genes/markers and QTL linkage. The third is the implementation phase, in which the identified genes/markers linked to QTL are used to obtain phenotypic data which are then combined with pedigree information to predict the genetic merit of individuals within the population. Additionally, two-stage selection strategies were investigated by modeling to improve the selection of cattle (Gomez-Raya *et al*., 1999). The first-stage selection was carried out within families. Then five different second-stage selection strategies were compared in 10 generations (~ 20 yrs). The results showed that the favorable allele was fixed gradually and the strategy of a selection index incorporating marker-QTL information and standard animal model BLUP required the lowest cost. Now along with the release of draft bovine genome sequence in 2005, the information of bovine genomics could be used in breeding programs to improve a range of traits (Sonstegard *et al*., 2004). Among all the QTLs, the most effective markers are the functional candidate genes in selecting animals with superior desired phenotypes. On beef marbling, *DGAT1* gene, which affects the synthesis of triglycerides, has been proved to be an effective marker (Grisart *et al*., 2002; Thaller *et al*., 2003).

2.4.2 QTL Cloning

Ultimately, QTL underlying the genetic variation of quantitative traits can be cloned. Cloning a QTL is to transfer the DNA sequence containing a gene or genes to a plasmid or other manipulate vector, which allows the complete power of modern molecular

biology to be used in the study and breeding application of this QTL. When a gene is cloned, one can sequence it; study its expression in different tissues and at different developmental times, and isolate homologous DNA sequences from other populations/species. As a breeding application, one can place modified copies of this gene back into the organism, or into different species, creating transgenic individuals. A variety of molecular approaches can be used to clone a gene whose product is known (Maniatis *et al.*, 1982). The product of a QTL, however, is typically unknown. So, two different cloning strategies have been suggested for genes with unknown products but discernible phenotypes: transposon tagging and positional cloning.

A transposon refers to a mobile genetic element, which may insert itself into or near a gene. Consequently, it can disrupt expression and creating a visible mutation, providing the basis for cloning by transposon tagging (Bingham *et al.*, 1981). In several species, it is now possible to introduce transposons modified for high insertion rates into the germline either by genetic or micro-injection techniques. Standard molecular techniques using these elements can then be used to isolate any region of DNA within which an element has inserted. For detection of QTL by transposon tagging, the use of inbred lines is essential in order to reduce variation from segregation at other loci. Typically, one starts with a line with little or no genetic variation and then selects for new mutations affecting the character, which are then examined for indications that at least one scored element has moved.

From the theoretical view, if a QTL can be localized to a sufficiently small region, we can possibly examine all of the genes in that region. For example, by using rare chromosome deletions translocations that correlate with the presence of a disease in humans, Collins(1992,1995) have successfully delimited the region in which the disease gene resides. Positional cloning can be done by brute force (sequencing of the entire region) or by more clever methods such as zoo blotting, in which sequences from related species are used to isolate only those parts of the region of interest that are conserved between species. Alternatively, we may use schemes that directly isolate genes, such as exon amplification or exon trapping methods (Duyk *et al*., 1990; Buckler *et al*., 1991). The timing and pattern of gene expression of genes localized within a region can suggest exclusion or inclusion of others as candidate QTLs. Hopefully, we will be able to recognize candidate genes from unique features of the DNA sequence itself, as we become increasingly better at associating particular amino acid sequences with particular functional units (e.g. known DNA binding region or specific catalytic sites).

2.5 Summary

Marbling contributes to beef flavor, texture and palatability. Subsequently high marbling beef are tastier. The beef grading systems were constructed in several advanced countries to meet this market requirement. Higher beef marbling score results in higher prices of carcass. Therefore, selection for marbling has been recognized as an important objective for the production of high quality beef worldwide. However, selection of beef marbling requires tremendous effort, expense and time. Detection and mapping QTLs (genes or markers linked to the genes) associated with beef marbling may accelerate this process

and introduce the marker-assisted selection program into beef cattle breeding. The following techniques are of great useful in QTL mapping for beef marbling. First, AFLP technique scans the whole bovine genome and could yield numerous DNA polymorphisms compared to other DNA markers (e.g. microsatellites). Second, selective DNA pooling greatly reduces the number of samples needed to genotype, which saves money and time. Third, comparing genomics and *in sillico* cloning utilizes information online and thus also saves money and time. Overall, these strategies will provide a more cost-effective way to produce high marbling beef for the USA markets as well as for exports, and will also increase benefit to beef cattle farmers.

References

Ashwell, M. S., Rexroad Jr., C.E.,Miller, R.H., Van Raden, P. M. and Da, Y. 1997. Detection of loci affecting milk production and health traits in an elite US Holstein population using microsatellite markers. Ani. Genet 28: 216-222.

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

Baro J.A., Carleos C., Corral N., Lopez T., Canon J. 2001. Power analysis of QTL detection in half-sib families using selective DNA pooling.Genet Sel Evol. 33: 231-247.

Becker J., Vos P., Kuiper M., Salamini F., and Heum M. 1995. Combined mapping of AFLP and RFLP markers in barley. Mol. Gene Genet. 249: 65-73.

Bingham P.M., Levis R., Rubin G.M. 1981. Cloning of DNA sequences from the white locus of D. melanogaster by a novel and general method. Cell. 25:693-704.

Bishop, M. D., Hawkins, G. A. and Keeler, C. L. 1995. Use of DNA markers in animal selection. Theriogenology 43: 61-70.

Braunschweig M.H., Paszek A.A., Weller J.I., Da Y., Hawken R.J., Wheeler M.B., Schook L.B., Alexander L.J. 2001. Generation and exploration of a dense genetic map in a region of a QTL affecting corpora lutea in a Meishan x Yorkshire cross. Mamm. Genome. 12: 719-723.

Brooks, A.J. 1999. The essence of SNPs. Gene 234: 177-186.

Brugmans B., Hulst R.G.M., Visser R.G.F., Lindhout P. and Eck H.J. 2003. A new and versatile method for the successful conversion of AFLP markers into simple single locus markers. Nucleic Acids Res. 30: e55.

Buchanan F.C., Fitzsimmons C.J., Van Kessel A.G., Thue T.D., Winkelman-Sim D.C., Schmutz S.M. 2002. Association of a missense mutation in the bovine leptin gene with carcass fat content and leptin mRNA levels. Genet Sel Evol. 34:105-116.

Buckler, A.J., Chang D.D., Graw S.L., Brook J.D., Haber D.A., Sharp P.A., and Housman D.E.. 1991. Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc. Natl. Acad. Sci. USA 88: 4005-4009.

Burson D.E., North Central Regional. Extension Publication #357, Quality and Yield Grades for Beef Carcasses, 1997.

Buntjer J.B., Otsen M., Nijman I. J., Kuiper M. T. R. and Lenstra J. A. 2002. Phylogeny of bovine species based on AFLP fingerprinting. Heredity 88: 46-51.

Campbell E.M., Nonneman D., Rohrer G.A. 2003. Fine mapping a quantitative trait locus affecting ovulation rate in swine on chromosome 8. J Anim Sci. 81:1706-1714.

Casas E., Shackelford S.D., Keele J.W., Koohmaraie M., Smith T.P., Stone R.T. 2003. Detection of quantitative trait loci for growth and carcass composition in cattle. J Anim Sci. 81:2976-2983.

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

Cavalli-Sforza L.L. and Feldman W.M. 2003. The application of molecular genetic approaches to study of human evolution. Nat. Genet supl. 33: 266-275.

Ciofi C. and 6 others. 1998. Genotyping with microsatellite markers. Molecular tools for screening biodiversity, Chapman & Hall, London. p 195-205.

Collard, B. C. Y., Jahufer M. Z. Z., Brouwer J. B. and Pang E. C. K., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142: 169-196.

Collins F.S. 1992. Positional cloning: let's not call it reverse anymore. Nat Genet.1: 3-6.

Collins F.S. 1995. Positional cloning moves from perditional to traditional. Nat Genet. 9: 347-350.

Craddock N., Dave S., Greening J. 2001. Association studies of bipolar disorder. Bipolar Disord. 3: 284-298.

Darvasi A., and Soller M.1994. Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. Genetics. 138:1365-1373.

Davis GP, DeNise SK. 1998. The impact of genetic markers on selection. J Anim Sci. 76:2331-2339.

Dekkers J.M. and Hospital F. 2002. The use of molecular genetics in the improvement of agricultural populations. Nature Reviews Genetics 3, 22-32.

Duyk, G.M., Kim S., Myers R.M. and Cox D.R. 1990. Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. Proc. Natl. Acad. Sci. USA 87: 8995-8999.

Eenennaam A.V., 2005. http://animalscience.ucdavis.edu/animalbiotech/ Biotechnology/MAS/index.htm

Elias Calles J.A., Gaskins C.T., Busboom J.R., Duckett S.K., Cronrath J.D., Reeves J.J., Wright RW Jr. 2000. Differences among Wagyu sires for USDA carcass traits and palatability attributes of cooked ribeye steaks. J Anim Sci. 78: 1710-1715.

Goldstein D. B. and Schlotterer C. (eds). 1999. Microsatellites: Evolution and applications. Oxford University Press, Oxford, England.

Gorni C., Williams J.L., Heuven H.C., Negrini R., Valentini A., van Eijk M.J., Waddington D., Zevenbergen M., Marsan P.A., Peleman J.D. 2004. Application of AFLP technology to radiation hybrid mapping. Chromosome Res.12: 285-297.

Goulding J.N., Stanley J., Saunders N., and Arnold C. 2000. Genome-sequence-based fluorescent amplified-fragment length polymorphism analysis of Mycobacterium tuberculosis. J Clin Microbiol. 38: 1121-1126.

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

bovine DGAT1 gene with major effect on milk yield and composition. Genome Res. 12:222-231.

Haley, C. S. 1995. Livestock QTL – bringing home the bacon?. Trends in Genet. 11: 488-492.

Haley C., http://citeseer.ist.psu.edu/cache/papers/cs/ 13408/ http:zSzzSzwww.agbiotechnet.comzSzproceedingszSz3_haley.pdf/ advances-in-quantitative-trait.pdf

Hyne V. and Kearsey M.J.. 1993. QTL analysis: further uses of 'marker regression'. Theor. Appl. Genet. 91:471-476.

Imai, K., Ihara, N., Watanabe, T., Tsuji, S., Fukuma, T., Osaki, Y., Ogata, Y., Matsushige, T., and Sugimoto, Y. 2004. QTL for beef marbling mapped on BTA9 and BTA14 in a paternal half-sib family from purebred Japanese Black cattle (Hiroshima-gyu) population. Proceedings of the 29th International Conference on Animal Genetics, Tokyo, Japan. p.118.

Jansen R.C. 1993. Interval mapping of multiple quantitative trait loci. Genetics. 135: 205-211.

Jansen R.C. 1994. Controlling the type I and type II errors in mapping quantitative trait loci. Genetics. 138: 871-881.

Jiang Z. and Michal J.J. 2003. Linking porcine microsatellite markers to known genome regions by identifying their human orthologs. Genome 46: 798-808.

Jiang Z., Wu X-L, Garcia M.D., Griffin K.B., Michal J.J., Ott T.L., Gaskins C.T., Wright Jr. R.W. 2004. Comparative Gene-based In Silico Transcriptome Analysis of Different Tissues/Organs in Cattle. Genome 47: 1164-1172.

Jiang Z., Kunej T., Michal J.J., Gaskins C.T., Reeves J.J., Busboom J.R., Dovc P., Wright R.W.Jr. 2005. 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. 334: 516-523.

JMGA. 1988. New Beef Carcass Grading Standards. Japan Meat Grading Association, Tokyo.

Kassama Y., Rooney P.J., and Goodacre R. 2002. Fluorescent amplified fragment length polymorphism probabilistic database for identification of bacterial isolates from urinary tract infections. J Clin. Microbiol. 40: 2795-2800.

Kazala E.C., Lozeman F.J., Mir P.S., Laroche A., Bailey D.R, Weselake R.J. 1999. Relationship of fatty acid composition to intramuscular fat content in beef from crossbred Wagyu cattle. J Anim Sci. 77: 1717-1725.

Kearsey, M.J. and Hyne V. 1994. QTL analysis: a simple 'marker regression' approach. Theor. Appl. Genet. 92: 698-702.

Kendal W.S. 2003. An exponential dispersion model for the distribution of human single nucleotide polymorphisms. Mol Biol Evol. 20: 579-590.

Lander E.S., and Bostein D.1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 121: 185-199. Erratum in: Genetics 1994, 136:705

Li C., Basarab J., Snelling W.M., Benkel B., Murdoch B., Hansen C., Moore S.S. 2004. Assessment of positional candidate genes myf5 and igf1 for growth on bovine chromosome 5 in commercial lines of Bos taurus. J Anim Sci. 82: 1-7.

Li Y-C, Korol A. B., Fahima T., Beiles A., and Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol. Ecol. 11: 2453-2465.

Lipkin, E., Mosig, M. O., Darvasi A., Ezra E., Shalom A., Friedman A., and Soller M. 1998. Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: analysis of milk protein percentage. Genetics 149: 1557-1567.

Kobayashi N., Hirano T., Watanabe T., Katoh S., Sobajima H., Hayashi N., Ohtani T., Sugimoto Y. QTL mapping for carcass traits on BTA2 and BTA24 in a paternal half-sib family comprising purebred Japanese Black cattle (Wagyu) population. Proceedings of the 29th International Conference on Animal Genetics, Tokyo, Japan. 2004; p.118.

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

Maniatis, T., Frisch E.F., and Sambrook J. 1982. Molecular cloning. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Marshall D.M. 1994. Breed differences and genetic parameters for body composition traits in beef cattle. J Anim Sci. 72: 2745-2755.

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol. 35: 89-99.

Mizoguchi Y., Iwamoto E., Tatsuda K., Watanabe T., Sugimoto Y. Identification of a major gene locus for marbling spanning 450 kb at the telomeric region of BTA21

using Japanese Black cattle population. Proceedings of the 29th International

Conference on Animal Genetics, Tokyo, Japan. 2004; p.118.

Montaldo H. H., and Meza-Herrera C. A. 1998. Use of molecular markers and major

genes in the genetic improvement of livestock. EJB Electronic Journal of

Biotechnology Vol.1 ISSN: 0717-3458

Moore S.S., Li C., Basarab J., Snelling W.M., Kneeland J., Murdoch B., Hansen C.,

Benkel B. 2003. Fine mapping of quantitative trait loci and assessment of

positional candidate genes for backfat on bovine chromosome 14 in a commercial

line of Bos taurus. J Anim Sci. 81: 1919-1925.

Mosig M.O., 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.Genetics.157: 1683-1698.

Mullis K., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. 1986. Specific enzymatic

amplification of DNA in vitro: the polymerase chain reaction.Cold Spring Harb

Symp Quant Biol. 51 Pt 1: 263-273

Nijman I.J., Bradley D.G., Hanotte O., Otsen M., and Lenstra J.A. 1999. Satellite DNA

polymorphisms and AFLP correlate with Bos indicus-taurus hybridization. Anim.

Genet. 30: 265-273.

Nijman Ij, Otsen M., Verkaar E.L.C., Ruijter C. de, Hanekamp E., Ochieng J.W.,

Shamshad S., Rege J.E.O., Hanotte O., Barwegen M.W., Sulawati T., and Lenstra

J.A. 2003. Hybridization of banteny (*Bos javanicus*) and zebu (*Bos indicus*)

revealed by mitochondrial DNA, satellite DNA, AFLP and microsatellites. Heredity 90: 10-16.

Pareek C.S., Pareek R.S., Walawski K. 2002. Novel linkage mapping approach using DNA pooling in human and animal genetics. II. Detection of quantitative traits loci in dairy cattle. J Appl Genet.43: 309-318.

Plotsky, Y., Cahaner A., Haberfeld A., Lavi U., Lamont SJ, and Hillel J. 1993. DNA fingerprinting bands applied to linkage analysis with quantitative trait loci in chickens. Anim. Genet. 24: 105-110.

Rocha J.L., Pomp D., Vleck L.D., Nielsen M.K. 2001. Predictors of marker-informativeness for an outbred F2 design. Anim Genet. 32: 365-370.

Rohrer G.A., Alexander L.J., Hu Z., Smith T.P., Keele J.W., Beattie C.W. 1996. A comprehensive map of the porcine genome. Genome Res. 6:371-391.

Sasazaki S., Itoh K., Srimitsu S., Imada T., Takasuga A., Nagaishi H., Takano S., Mannen H. and Tsuji S. 2004. Development of breed identification markers derived from AFLP in beef cattle. Meat Sci. 67: 275-280.

Smith T.P., Casas E., Rexroad C.E. 3rd, Kappes S.M., Keele J.W. 2000. Bovine CAPN1 maps to a region of BTA29 containing a quantitative trait locus for meat tenderness. J Anim Sci. 78: 2589-2594.

Stephens J.C., Schneider J.A., Tanguay D.A., Choi J., Acharya T., Stanley S.E., Jiang R., Messer C.J., Chew A., Han J.H., Duan J., Carr J.L., Lee M.S., Koshy B., Kumar A.M., Zhang G., Newell W.R., Windemuth A., Xu C., Kalbfleisch T.S., Shaner S.L., Arnold K., Schulz V., Drysdale C.M., Nandabalan K., Judson R.S., Ruano

G., Vovis G.F. 2001. Haplotype variation and linkage disequilibrium in 313 human genes. Science. 293: 489-493.

Sunnucks P. 2000. Efficient genetic markers for population biology. Trends Ecol. Evol. 15: 199-203.

Thaller G., Kuhn C., Winter A., Ewald G., Bellmann O., Wegner J., Zuhlke H., Fries R. 2003. DGAT1, a new positional and functional candidate gene for intramuscular fat deposition in cattle. Anim Genet. 34: 354-357.

Tsuji, S., Itoh K., Sasazaki S, Mannen H, Oyama K, Shojo M, Mukai F. 2004. An association study using AFLP markers and application to a beef cattle breeding population. Anim Genet. 35: 40-43.

USDA, 1997. Official United Stated standards of grades of carcass beef. AMS. USDA. Wasington, D. C.

Va der Beck, van Arendonk S. J.A.M., and Groen A.F. 1995. Power of two- and three-generation QTL mapping experiments in outbred populations containing full-sib or half-sib families. Theor. Appl. Genet. 91: 1115-1124.

Vatcher G., Smailus D., Krzywinski M., Guin R., Stott J., Tsai M., Chan S., Pandoh P., Yang G., Asano J., Olson T., Prabhu A.L., Cooper, Marziali A., Schein J., Jones S., and Marra M. 2002. FRFLP and fAFLP: Medium-throughput genotyping by fluorescently post-labeling restriction digestion. BioTechniques 33: 549-553.

Vignal A., Milan D., SanCristobal M., Eggen A.2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol.34: 275-305.

Vos P., Hogers R., bleeker M., Reijans M., Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., and Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Resarch, 23: 4407-4414.

Wang D.G., Fan J.B., Siao C.J., Berno A., Young P., Sapolsky R., Ghandour G., Perkins N., Winchester E., Spencer J., Kruglyak L., Stein L., Hsie L., Topaloglou T., Hubbell E., Robinson E., Mittmann M., Morris M.S., Shen N., Kilburn D., Rioux J., Nusbaum C., Rozen S., Hudson T.J., Lipshutz R., Chee M., Lander E.S. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science. 280: 1077-1082.

Weeks, A.R., Opijnen, T.Van, and Breeuwer, J.A.J. 2000. AFLP fingerprinting for assessing intraspecific variation and genome mapping in mites. Exp. Appl. Acarol. 24: 775-793.

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

Weller, J.I. Kashi Y., and Soller M. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. J. Dairy Sci. 73: 2525-2537.

Weller J. I. 2001. Quantitative Trait Loci Analysis in Animals. CABI Publishing, New York

Wertz A.E., Berger L.L., Walker P.M., Faulkner D.B., McKeith F.K., Rodriguez-Zas S.L. 2002. Early-weaning and postweaning nutritional management affect feedlot

performance, carcass merit, and the relationship of 12th-rib fat, marbling score, and feed efficiency among Angus and Wagyu heifers. J Anim Sci. 80: 28-37

Winzeler E.A., Richards D.R., Conway A.R., Goldstein A.L., Kalman S., McCullough M.J., McCusker J.H., Stevens D.A., Wodicka L., Lockhart D.J., Davis R.W. 1998. Direct allelic variation scanning of the yeast genome. Science. 281:1194-1197.

Wu X-L, MacNeil M. D., De S., Xiao Q-J, Michal J. J., Gaskins C. T., Reeves J. J., Busboom J. R., Wright R. W., Jr. and Jiang Z.. 2005. Evaluation of candidate gene effects for beef backfat via Bayesian model selection. Genetica 125: 103–113.

Wu W.R. and Li W.M. 1994. A new approach for mapping quantitative trait loci using complete genetic marker linkage maps. Theor. Appl. Genet. 89: 535-539

Yeo J.S., Lee J.Y, Kim J.W. 2004. DNA marker mining of ILSTS035 microsatellite locus on chromosome 6 of Hanwoo cattle. J Genet. 83: 245-250.

Zane L., Bargelloni L., and Patarnello T. 2002. Strategies for microsatellite isolation: a review. Mol. Ecol. 11: 1-16.

Zeng Z.B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc Natl Acad Sci U S A. 90:10972-10976.

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

CHAPTER THREE

A simplified QTL mapping approach for screening and mapping of novel AFLP markers associated with beef marbling

A simplified QTL mapping approach for screening and mapping of novel AFLP markers associated with beef marbling

Qianjun Xiao[1], Tito A. Wibowo[1], Xiao-Lin Wu[1], Jennifer J. Michal[1], Jerry J. Reeves[1], Jan R. Busboom[1], Gary H. Thorgaard[2] and Zhihua Jiang[1]

[1]Department of Animal Sciences, Washington State University, Pullman, WA99164 – 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

Abstract

Genome screening of quantitative trait loci (QTL) for a complex trait is usually costly and highly laborious, as it requires a large number of markers spanning the whole genome. Here we present a simplified approach for screening and mapping of QTL linked markers for beef marbling using a Wagyu x Limousin F_2 reference population. This simplified approach involves integration of the amplified fragment length polymorphism (AFLP) with DNA pooling and selective genotyping and comparative bioinformatics tools. AFLP analysis on two high and two low marbling DNA pools yielded ten visually different markers. Among them, four were confirmed based on individual AFLP validation. Sequencing and *in silico* characterization assigned two of these AFLP markers to bovine chromosomes 13 (BTA13) and 1 (BTA1), which are orthologous to human chromosomes HSA10p11.23 and HSA21q22.2 with both regions harboring QTL for obesity related phenotypes. Both AFLP markers showed significantly large additive genetic effects (0.54 ± 0.21 on BTA13 and 0.28 ± 0.11 on BTA1) on beef marbling score (BMS) ($P<0.05$). Advantages and disadvantages of this simplified QTL mapping approach are discussed. Overall, this approach is less time consuming, inexpensive and in particular, suitable for screening and mapping QTL linked markers when targeting one or a few complex traits.

Keywords: QTL mapping, AFLP, DNA pooling, bioinformatics tools, beef marbling.

Introduction

Marbling is a term commonly used to describe the appearance of white flecks or streaks of fatty tissue between the muscle fibers in meat. As an indicator of intramuscular fat, this trait has attracted a great deal of publicity and interest, because deposition of fat in the muscle of a beef carcass contributes positively to the taste, texture and flavor of the meat (Elias Calles *et al*., 2000). Obviously, beef marbling is of high economic importance, but progress is currently limited because selection for beef marbling requires tremendous effort, expense and time. A trait such as beef marbling is therefore ideally suited to capitalize on molecular genetic technologies (Parnell, 2004). Identifying, mapping, and understanding the function and control of genes for beef marbling will permit the development of new genetic technologies and open the way to realize the full genetic potential for improvement of beef production for maximum profits.

As usual, multiple genes and environmental factors determine complex genetic traits such as beef marbling. The individual loci that make up the genetic component of a quantitative trait are called "quantitative traits loci (QTL)." QTL mapping is defined as a process to localize chromosome regions harboring genetic variants that affect a continuously distributed, polygenic phenotype (DiPetrillo et al., 2005). Genome-wide linkage studies of complex traits conducted by utilizing highly informative microsatellite markers have proven to be a feasible means of detecting QTLs in different species. However, due to costliness and high labor requirements, these genome scans were usually performed with relatively few markers spanning the whole genome, and thus provided a

43

low resolution of mapped QTL locations, perhaps 20 cM or more. These distances make it difficult to move from the mapped QTL to identification of actual genes. For example, although QTL analysis started in the early 1990s, investigators have identified only ~30 causal genes underlying QTL in mice so far (Flint et al., 2005). Thus, the major hurdle to identifying QTL genes is not identification and localization of a QTL in the genome, but rather to the expensive and time-consuming process of narrowing QTL to a few candidate genes for a detailed characterization and functional analysis.

Here we present a simplified, inexpensive QTL mapping approach for genome-wide scans of QTL-linked markers and for narrowed locations of QTL regions. This simplified approach involves integration of the amplified fragment length polymorphism (AFLP) with DNA pooling and selective genotyping and comparative bioinformatics tools. AFLP simultaneously screens high numbers of loci for polymorphisms and detects many more polymorphic DNA markers than any other PCR based detection systems (Vos et al., 1995). Thus, the technique provides the capacity to reveal "responsible mutations" for a QTL. Meanwhile, DNA pooling and selective genotyping are applied to reduce the numbers of samples (Darvasi, 1992). The main mechanism involved in DNA pooling and selective genotyping is that distributions of genotypes are quite similar close to the means, but very different in the tails of the distributions. Therefore, individuals with extreme phenotypic values provide the most information for the QTL (Plotsky et al., 1993; Lipkin et al., 1998). Comparative bioinformatics tools take advantage of genome sequencing and conservation in human, cattle and other mammalian species. In particular, retrieval of the same gene sequences in the target species or orthologous sequences in

other species can immediately place QTL-linked markers to narrowed chromosome regions. Using such a simplified approach, we identified QTL-linked AFLP markers for marbling in Wagyu x Limousin F_2 crosses, which have relevant evidence for obesity observed in humans.

Materials and Methods

Animals, marbling scores and genomic DNA

The animals used in the present study were F_2 progeny derived from Wagyu x Limousin F_1 sires and F_1 dams at the Fort Keogh Livestock and Range Research Laboratory, ARS, USDA. The Wagyu breed of cattle has been selected for high marbling for a long time, whereas the Limousin breed has been selected for heavy muscle, which has led to a low marbling score. The difference in marbling scores between these two breeds makes them very unique for mapping QTL for this economically important trait in beef cattle. Development of the reference population has been described previously by Wu and colleagues (2005). Beef marbling score (BMS) was a subjective measure of the amount of intramuscular fat in the *longissimus* muscle based on USDA standards (http://www.ams.usda.gov/). Subcutaneous fat depth (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. Phenotypic data have been adjusted for effects of year, gender, and age at harvest (linear) before they were used in the association analysis. Both DNA samples and performance data on these F2 animals were

kindly provided by Dr. MacNeil. Based on the availability of both phenotypic data and DNA samples, 246 F_2 animals were used in the study. Thirty samples with the highest BMS were used to form two high pools whereas 30 samples with the lowest BMS were used to form two low pools by adding equal amounts (20 ng) of DNA from each of 15 individuals to a pool.

AFLP analysis

The AFLP analysis on these four DNA pools was performed using a procedure including adapter and primer sequences described previously by Ajmone-Marsan and colleagues (1997) with minor modifications (Table 1; Table 2). In brief, 200 ng of genomic DNA were digested with two restriction enzymes: *EcoRI* and *TaqI* (New England Biolabs, Beverly, MA, USA) based on the manufacturer's instruction. The digested products were then ligated to 5 pMol *EcoRI* – adapters and 50 pMol *TaqI* – adapters in 50 μl of solution containing 25 U T4 ligase and 1X T4 ligase buffer under incubation overnight at room temperature. The ligated DNA templates were diluted 1:10 with 10mM Tris-HCl, 0.1mM EDTA (pH 8.0) for pre-amplification.

The pre-amplification PCR condition was: 10 ng of DNA template, 1x Platinum Taq Buffer (20 mM Tris-HCl, PH 8.4, 50 mM KCl; Invitrogen), 3.0 mM $MgCl_2$, 0.3 mM each of the four dNTPs, 1 U Platinum Taq polymerase (Invitrogen), 0.2 pMol of pre-amplification *EcoRI* primer (E01: 5'-GAC TGC GTA CCA ATT CA-3') and 2 pMol of pre-amplification *TaqI* primers (T01: 5'-GAT GAG TCC TGA CCG AA-3' or T02: 5'-

GAT GAG TCC TGA CCG AC-3') in a total volume of 50 μl. The PCR program is as follows: 2 min at 94°C, 2 min at 72°C, 25 cycles of 10 seconds at 94°C, 30 seconds at 56°C and 2 min at 68°C, followed by 30 min at 60°C, and ended at 4°C. PCR products were analyzed using 1.6% agarose gels, stained with ethidium bromide and photographed. The pre-amplification products were diluted 1:20 with 10mM Tris-HCl, 0.1mM EDTA (pH 8.0) and then used as selective amplification DNA templates.

For selective amplification, 8 *EcoRI* primers and 8 *TaqI* primers were used, resulting in 64 primer combinations. *EcoRI* primers were 5' end fluorescently labeled. The following PCR reaction mix was used: 3.0 μl of diluted pre-amplification products, 1x Platinum *Taq* Buffer (20 mM Tris-HCl, PH8.4, 50 mM KCl; Invitrogen), 3.0 mM $MgCl_2$, 0.3 mM each of the four dNTPs, 1 U Platinum Taq polymerase (Invitrogen) and 5 ng of selective-amplification *EcoRI* primer and 25 ng selective-amplification *TaqI* primer in a total volume of 20 μl. A touchdown thermal protocol was used in the selective PCR amplification (Ajmone-Marsan *et al*., 1997). The selective-amplification products were prepared as a mix of the following: 1.0 μl of each fluorescently- labeled PCR product, 12.0 μl of formamide, and 0.5 μl of Gene Scan[TM] 500 LIZ[TM] size standard (Applied Biosystems). The mixed products were then separated on an ABI 3730 sequencer in the Laboratory for Biotechnology and Bioanalysis (Washington State University) using a standard protocol. Data were collected automatically and analyzed using software *GeneMapper3.7* (Applied Biosystems).

AFLP markers sequencing, in silico flanking walking and PCR-RFLP genotyping

Sixty-four primer combinations were used to generate AFLP patterns on four DNA pools, including two high-marbling pools and two low-marbling pools. Comparison of peak heights yielded ten potential AFLP markers that had the most visually-striking differences between high pools and low pools (see examples, Figure 1), which were then selected for individual AFLP validation using a same protocol as described above. The validation was performed on 24 high and 24 low BMS samples and the presence or absence of AFLP bands of interest was scored individually. The Fisher's exact test was used to test the difference in fragment frequencies between the extreme individuals. The significant primer combinations, fragment frequencies and significance levels are listed in Table 1.

To isolate the AFLP fragments, selective amplification products that contained the marker fragments of interest were separated on a 5% polyacrylamide gel. The bands representing AFLP fragments of interest were excised using a scalpel. After excision, gel fragments were placed in 15 µl of 1X TE and frozen at -80°C for ~30 min, thawed at room temperature and then refrozen at -20°C. After thawing again, samples were centrifuged for 15 min at 15 000 g and 4.0 µl was taken for PCR re-amplification using pre-amplification AFLP primers. Fragments were sequenced directly using the same pre-amplification primers on ABI 3730 automatic capillary sequencing system following standard Big Dye protocols.

Among these four fragment isolates, only a fragment obtained from primer combination

E+AAC/T+ACA did not yield a readable sequence. Three readable sequences of AFLP

markers were used as queries to perform BLAST searches against the 6X bovine genome

sequence database (http://www.hgsc.bcm.tmc.edu/projects/bovine/) in order to obtain the

flanking sequence of the same locus in cattle. Unfortunately, the sequence derived from

the primer combination E+AAG/T+CAT hit a SINE (short interspersed nuclear elements)

and hence it was discarded for further use. Sequences obtained from primer

combinations E+AGT/T+CAT and E+AGT/T+ACT each hit a bovine genomic contig

with highly matched sequences. Primers were designed to further characterize these two

AFLP fragments. The primer sequences designed for the primer combination

E+AGT/T+CAT were: forward, 5'-TTT GGA GCA GTG ACA GGA TCA GAC-3'; and

reverse: 5'-AGA GAG CCT GCG TCC TTA TCT CAC-3' (GenBank accession no.

AAFC02027857) (Figure 2A). The primers designed for the primer combination

E+AGT/T+ACT were: forward, 5'-AAA CTG TCC TTC AAG GTA GTC AAC A-3'

and reverse, 5'-GGG GCA CTA GAG TGG GTT GCC ATT T-3' (GenBank accession

no. **AAFC02113318**) (Figure 3A). PCR amplification was performed on six F_1 Wagyu x

Limousin bulls in order to reveal molecular causes for the AFLP polymorphisms and

determine the strategies for genotyping the markers on all F_2 progeny.

Statistical analysis

Initially, the difference in allelic frequency between the high and low BMS pools or

individual groups served as the first step to detect the association between the marker and

the traits. Allelic frequencies were estimated based on a conjugate Bayesian gamma-Poisson hierarchical model with Markov chain Monte Carlo implementation (Kunej et al., 2005). The comprehensive marker-trait association analysis was based on the following mixed model as

$$y = X\beta + Zu + e$$

where y is a vector of observations, β is a vector for all fixed effects including the overall mean and the candidate gene effect (exactly, marker associated effects), $u \sim N(0, A\sigma_u^2)$ is a vector containing residual genetic effects, other than the current gene (marker) effect, with A being the additive genetic relationship matrix of all individuals, X and Z are incidence matrices relating the effects in β and u, respectively, to observations in y, and $e \sim N(0, I\sigma_e^2)$ is a vector of residual errors. The association analysis was conducted using the PROC GLM procedure in the SAS system (SAS Institute, Cary, NC, USA). The additive effect was estimated as one half of two homozygous markers means, and the dominance effects as the deviation of the heterozygous mean from the average of homozygous means, under the assumption of complete linkage between the marker and the candidate gene.

Results

Genome-wide screening of QTL linked AFLP markers for beef marbling

In total, 64 AFLP primer combinations with eight *EcoR*I primers and eight *Taq*I primers were employed in a genome-wide screening of QTL linked markers for beef marbling on

two-high and two-low BMS DNA pools derived from a population of Wagyu x Limousin

F_2 crosses. Each primer combination generated about 30 - 120 clearly scorable fragments

with a size range of 75 – 500 bp. The primer combination of E + AAC/T + ACT and E +

AAG/T + ACT yielded more fragments (~ 100) than other primer combinations. The

type of fluorescence may affect fragments numbers, because we found that PET labeled

primers yielded the least fragments among the four types of fluorescently labeled

primers. Analysis using *GeneMapper3.7* (Applied Biosystems) demonstrated ten

potential AFLP markers with the most striking visual differences in terms of peak height

between high and low performance pools. These markers can be classified into two

categories. One category included markers present in both high pools but absent in both

low pools or vice versa, such as primer combinations E+ACA /T+CAC (Figure 1A as an

example), E+AGA/T+ACT, E+ACA/T+AAC, E+AAC/T+ACA and E+AAG/T+AAC.

The other category of markers showed the differences in peak heights: the peaks in both

high pools are remarkably higher than those in both low pools, or vice versa, such as

primer combinations E+AAG/T+CAT (Figure 1B as an example), E+AGT/T+ACT,

E+ATC/T+ACT, E+AGT/T+CAT, and E+ATC/T+CAC.

In order to exclude any false positive markers, we performed individual AFLP analyses

on the 24 top and 24 bottom marbling samples. Among the ten markers identified above,

only four consistently showed differences in AFLP fragment frequencies between high

and low groups of animals (Table 3). Fragments derived with primer combinations

E+AAC/T+ACA, E+AGT/T+CAT and E+AGT/T+ACT were significantly different

(P<0.01) between high and low marbling groups. . In comparison, Fisher's exact test

revealed that the difference between high and low animals only approached significance (P<0.1) when the primer combination E+AAG/T+CAT was used. All four AFLP fragments were excised from a 5% polyacrylamide gel, re-amplified and sequenced. All primer combinations, except E+AAC/T+ACA produced products that generated readable sequences. Among the three readable sequences, BLAST search indicated that most of the sequence from the E+AAG/T+CAT marker was SINE related and could not be used as a marker and was subsequently discarded. Products amplified with primer combinations E+AGT/T+CAT and E+AGT/T+ACT each hit a bovine genomic contig, respectively.

Characterization of QTL-linked AFLP makers for beef marbling

Figure 2 illustrates both the AFLP (E+AGT/T+ACT) marker sequence and its flanking sequence with primers designed to further characterize the AFLP fragment. The flanking sequence was simply extracted from a bovine genomic contig - Con118216 (GenBank accession number **AAFC02113318**). Both *EcoR*I and *Taq*I restriction enzyme recognition sites were clearly identified in the fragment (Figure 2A), which span a sequence of 238 bp in length. By adding 22 bp of adaptor sequence to the enzyme cut fragment, the total length exactly matched the AFLP size of 260 bp (22 + 238 bp) identified on gels (Table 3). BLAST search found the bovine genomic contig - Con118216 is orthologous to a human genomic sequence with GenBank accession number **AL158036**, which contains a novel gene KIAA1462 on HSA10p11.23. *In-silico* mapping could place this AFLP marker or the bovine genomic contig to a region of 45.589 and 46.63 cM on bovine chromosome 13 (BTA13). Sequencing analysis of the

amplified products spanning the AFLP (E+AGT/T+ACT) marker on six F1 Wagyu x

Limousin bulls revealed a single nucleotide polymorphism (SNP) at the *Taq*I cut site, but

nothing at the *EcoR*I cut site (Figure 2B). Interestingly, an additional SNP was also

detected within the selective primer-extended region beside the *Taq*I cut site. Therefore,

two G/A SNPs are responsible for the AFLP at this locus (Figure 2B).

BLAST search using the sequence derived from the primer combination E+AGT/T+CAT

retrieved a contig Con28119 (GenBank accession number: **AAFC02027857**) from the 6X

bovine genome sequence database (http://www.hgsc.bcm.tmc.edu/projects/bovine/).

Both *EcoR*I and *Taq*I restriction enzymes make cuts for a fragment of 217 bp in length

(Figure 3A), which corresponded exactly to this AFLP marker of 239 bp (217 bp + 22

bp) identified in the gel when extra adaptor sequence of 22 bp was included in the

product (Table 1). The bovine genomic contig Con28119 was then found to be

homologous to human genomic contig AP001725, containing chromosome 21 open

reading frame 5 (*C21orf5*) on HSA21q22.2. The current draft map of the bovine genome

(NCBI builder 2.1) indicated that the bovine ortholog of human C21orf5 is located on

BTA1 at position 96.18 Mb (http://www.ncbi.nlm.nih.gov/projects/genome/guide/cow/).

A pair of primers was designed to amplify this AFLP marker, but sequencing of PCR

products on 6 F1 Wagyu x Linmousin bulls failed to show any mutations at either *EcoR*I

or *Taq*I cut sites. Instead, a C/T transition occurred at the 2nd extended base for the

selective *Taq*I primer, which certainly caused the AFLP polymorphism at the locus

(Figure 3B). The association of this AFLP marker with beef marbling will be reported

separately, as its human ortholog *C21orf5* is located in the Down syndrome critical region (http://www.ncbi.nlm.nih.gov/projects/genome/guide/human/).

Association of AFLP (E+AGT/T+ACT) marker on BTA13 with beef marbling and SFD

As indicated above, two SNPs were responsible for the AFLP marker derived from primer combination E+AGT/T+ACT. Obviously, the G/A substitution within the *Taq*I restriction site could be genotyped with the *Taq*I enzyme. Fortunately, the G/A substitution within the primer extension region could be revealed by digestion with restriction enzyme *Msp*I. Initially, we genotyped both SNPs on 30 high and 30 low BMS individuals, however, we found that the SNP with *Taq*I cut site was not very informative between the high and low BMS individuals. Therefore, restriction enzyme *Msp*I was used to genotype 246 F2 individuals by PCR-RFLP (Figure 4).

Using selective DNA genotyping data, Bayesian analysis showed dramatic differences in the posterior allelic frequencies between the high and low BMS groups. The 95% quantile range of either allele, G or A, did not overlap between the high and low BMS groups (Table 4). To visualize the difference, we plotted the posterior distributions of allele A for the high and low BMS groups, respectively, using the 9900 posterior samples, which were saved at every one-tenth of the 90000 updates after the 1000 burn-in Markov chain updates (Figure 5). Very clearly, the overlapping samples between the two posterior distributions of allele A consisted of less than 5% posterior samples. In the Bayesian paradigm, the posterior probability can be understood as a measure of uncertainty or degree of belief. Therefore, we would interpret this result as significant

difference in allelic frequencies between the high and low BMS groups, and the latter was indication of the linkage between the AFLP marker and a gene affecting BMS (Figure 5).

We further conducted the association analysis using the data from all individuals, based on an animal model with the marker as the fixed effect variable and residual individual genetic effect as the random effect variable. We included the latter in the model in order to account for effects of genes other than the one under investigation. Based on the F statistics constructed for the fixed effects, the AFLP marker was significantly associated with BMS (F=4.68, P=0.0102). Estimated marker means of BMS were -0.109±0.072, 0.276±0.155, 0.978±0.414 for genotypes GG, GA and AA, respectively. Obviously, the AA genotype was associated with significantly higher marbling score than the GG genotypes (P<0.05). The difference in SFD, however, was very slight among the three genotypes. Candidate gene effects were estimated under the assumption of complete linkage. We estimated the additive effect on BMS was 0.54±0.21, which was close to the highly significant threshold level (P = 0.0114). However, estimated dominance effect on BMS was not significant (P>0.05) (Table 5). Additionally, for this candidate gene, both additive and dominance effects on SFD were not significant (P>0.05). These results would strongly suggest that this candidate gene affected BMS in an additive genetic mode.

Discussion

We presented a simplified QTL mapping approach in the study by integration of AFLP, selective DNA pooling and bioinformatics tools. The first step was to apply the AFLP technique in screening of QTL linked markers for a complex trait on DNA pools of animals with extreme phenotypes. In the second step, the potential QTL-linked markers were validated individually on high and low marbling score animals and truly significant QTL linked markers were further characterized by DNA sequencing. The *in-silico* tools were then employed in the third step to identify same gene sequences of AFLP markers in the targeted species or orthologous sequences in other species and place the AFLP markers in the targeted genome. Finally, the flanking sequence of an AFLP marker was used to design primers for revealing molecular causes responsible for the amplified fragment length polymorphisms and thus determining the genotype assay for marker-trait association analysis. Clearly, this simplified QTL mapping approach has several advantages.

The simplified QTL mapping approach is neither expensive nor time-consuming

 In the present study, we performed a genome-wide scan using 64 primer combinations on four DNA pools. Theoretically, this process just requires a total of 256 PCR reactions. We estimated that these 64 primer combinations would generate a total of 3840 (64 x 60) fragments. If we assume that 10% of these fragments (Ajmone-Marsan et al., 1997; Felip et al., 2005) are polymorphic, the screening was done with a total of 384 markers. If we pursued a conventional genome-wide screening approach with 384 markers on 250 F_2 progeny used in the present study, we would have conducted at least 96,000 PCR

reactions. An additional 480 reactions were used for individual validation of ten potential

QTL linked markers. The *in silico* retrieval of flanking sequences of AFLP markers and

the *in silico* mapping of them to the target genome was free. Clearly this simplified QTL

mapping approach saves a lot of time and laboratory expenses. The approach is

particularly powerful when we deal with one or two traits at a time.

AFLP assay detects a variety of genetic polymorphisms in the genome

For the AFLP derived from primer combination E+AGT/T+ACT, two G/A mutations

were confirmed by sequencing six F1 Wagyu x Limousin bulls. One G/A occurred

within the *Taq*I enzyme recognition site, and another G/A just five bases apart from the

first mutation affected selective primer binding. For the AFLP derived from primer

combination E+AGT/T+CAT, a C/T substitution was detected within the extended

region. Clearly, the C/T mutation does not affect enzyme digestion, but it does affect the

selective primer binding, which caused the amplified fragment length polymorphisms.

Therefore, these results provided clear evidence that the AFLP assay can detect

polymorphisms within the restriction enzyme recognition site as well as the surrounding

areas where the selective primer can reach. Theoretically, any deletion/insertion or short

tandem repeats with the AFLP fragment can also be detected by the technique (Savelkoul

et al., 1999). Interestingly, of the three mutations identified in these two AFLP markers,

none were located in the *EcoR*I cut site or flanking regions. Rather, they were all located

in the *Taq*I cut site and the surrounding regions. The reason could be due to a

dinucleotide CpG that presents at the center of *Taq*I recognition site. When it is present,

the cytosine within the dinucleotide is usually methylated. The methylated cytosine has a propensity to undergo spontaneous deamination to form thymidine (Caiafa and Zampieri, 2005). That is why C/T (G/A) transition is so frequent in mammalian genomes.

AFLP driven markers for beef marbling make sense

The human orthologous regions for both AFLP markers associated with beef marbling were determined in this study: one AFLP marker amplified with primer combination E+AGT/T+ACT was orthologous to a novel gene *KIAA1462* on HSA10p11.23, while another marker derived from the primer combination E+AGT/T+CAT, was orthologous to a novel gene *C21orf5* on HSA21q22.2. Both regions in the human genome harbor QTLs for obesity-related phenotypes. On HSA10p12-11, a major quantitative trait locus was revealed that had significant linkage to obesity and was confirmed in different populations (Hager et al., 1998; Hinney et al., 2000; Price et al., 2001). Similarly, on HSA21q21-23, quantitative trait loci were discovered to have effects on diabetes, obesity, or total cholesterol and triglycerides (Lindgren et al., 2002; Li et al., 2004; North et al., 2005). Also, in mice, a significant linkage was found to adiposity chromosome 16 (Reed et al., 2003) as this mice chromosome is orthologous to HSA21.

In silico mapping of AFLP markers points to candidate genes for beef marbling

Using the AFLP technique to screen QTL-linked markers for longevity in *Drosophila melanogaster*, Luckinbill and Golenberg (2002) found that the furthest distance between

58

an AFLP marker and a known QTL was less than 3.0 cM. Most were within 1.0 or 1.5 cM and two were within the peak of mapped QTL limits. With such a short distance, candidate genes can be easily identified and nucleotide variants can be easily sequenced. In the vicinity of both obesity QTLs on HSA10p11.23 and on HSA21q22.2, two candidate genes, *PAPD1* and *DSCR1*, have drawn our attention. The former gene is a nuclear-encoded mitochondrial poly(A) polymerase (hmtPAP) and has a role in mitochondrial RNA processing (Tomecki et al., 2004), while the latter has a crucial role in the maintenance of mitochondrial function and integrity (Chang and Min 2005). In mice, a profound decrease of approximately 50% in the levels of transcripts for nuclear-encoded mitochondrial genes was found to accompany the onset of obesity (Wilson-Fritch et al., 2004). Recently, Jiang and colleagues (2005) found that the bovine mitochondrial transcription factor A (*TFAM*) was associated with BMS in the same population of cattle as described above. Therefore, we will further characterize these candidate genes to see if they contribute to obesity-related phenotypes.

No doubt, this simplified QTL mapping approach; in particular the AFLP assay itself has several drawbacks. First, of ten potential QTL linked AFLP markers identified for beef marbling, only four were confirmed to show significant differences between high and low groups of animals by individual AFLP genotyping. This means that AFLP screening on DNA pools generated a relatively high percentage of false positive markers. Second, as over 600 genes, markers and chromosomal regions have been identified as associated or linked with human obesity phenotypes (Perusse et al., 2005), the two enzyme (*EcoR*I and *Taq*I) combination will not detect all markers linked to beef marbling in the population.

Third, among three readable sequences of AFLP markers generated in the study, one (E+AAG/T+CAT) completely hit the SINE element and another (E+AGT/T+ACT) has almost of half of its sequence relevant to SINE element. This implies that *EcoR*I-*Taq*I combination favors amplification of repetitive regions. Therefore, the enzyme combination best suitable to mammalian genome needs to be further considered.

All together, two novel positional candidate gene regions were identified to have significant effects on beef marbling score in Wagyu x Limousin F_2 crosses through a simplified QTL mapping approach. Further studies are needed to confirm and characterize these genes on how they are functionally involved in the genetic control of beef marbling variation. In particular, the two genes that are orthologous to human genes *KIAA1462* and *C21orf5* should be our next targets, because both HSA10p and HSA21q harbor QTLs for obesity-related phenotypes. This study may also provide important information to unravel genetic complexity of obesity on these two positions in humans.

References

Ajmone-Marsan P., Valentini A., Cassandro M., Vecchiotti-Antaldi G., Bertoni G., Kuiper M., 1997. AFLP markers for DNA fingerprinting in cattle. Anim. Genet. 28, 418-426.

Caiafa P and Zampieri M., 2005. DNA methylation and chromatin structure: the puzzling CpG islands. J. Cell. Biochem. 94, 257-264.

Chang KT, Min KT., 2005. Drosophila melanogaster homolog of Down syndrome critical region 1 is critical for mitochondrial function. Nat Neurosci. 8, 1577-85.

Darvasi A, Weller JI., 1992. On the use of the moments method of estimation to obtain approximate maximum likelihood estimates of linkage between a genetic marker and a quantitative locus. Heredity 68 (Pt 1), 43-46.

DiPetrillo K, Wang X, Stylianou IM, Paigen B., 2005. Bioinformatics toolbox for narrowing rodent quantitative trait loci. Trends Genet. 21, 683-692.

Elias Calles JA, Gaskins CT, Busboom JR, Duckett SK, Cronrath JD, Reeves JJ, Wright RW Jr., 2000. Differences among Wagyu sires for USDA carcass traits and palatability attributes of cooked ribeye steaks. J. Anim. Sci. 78, 1710-1715.

Felip A, Young WP, Wheeler PA, Thorgaard GH., 2005. An AFLP-based approach for the identification of sex-linked markers in rainbow trout (Oncorhynchus mykiss). Aquaculture 247, 35-43.

Flint J, Valdar W, Shifman S, Mott R., 2005. Strategies for mapping and cloning quantitative trait genes in rodents. Nat. Rev. Genet. 6, 271-286.

Hager J, Dina C, Francke S, Dubois S, Houari M, Vatin V, Vaillant E, Lorentz N, Basdevant A, Clement K, Guy-Grand B, Froguel P., 1998. A genome-wide scan

for human obesity genes reveals a major susceptibility locus on chromosome 10. Nat Genet. 20, 304-308.

Hinney A, Ziegler A, Oeffner F, Wedewardt C, Vogel M, Wulftange H, Geller F, Stubing K, Siegfried W, Goldschmidt HP, Remschmidt H, Hebebrand J., 2000. Independent confirmation of a major locus for obesity on chromosome 10. J Clin Endocrinol Metab. 85, 2962-2965.

Jiang Z., Kunej T., Michal J.J., Gaskins C.T., Reeves J.J., Busboom J.R., Dovc P., Wright R.W. Jr., 2005. 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. 334, 516-523.

Kunej T., Wu X.-L., Berlic T.M., Michal J.J., Jiang Z., Dovc P., 2005. Frequency distribution of a Cys430Ser polymorphism in peroxisome proliferator-activated receptor-gamma coactivator-1 (PPARGC1) gene sequence in Chinese and Western pig breeds. J. Anim. Breed. Genet. 122, 7–11.

Li W.D., Dong C., Li D., Zhao H., Price R.A., 2004. An obesity-related locus in chromosome region 12q23-24. Diabetes 53, 812-820.

Lindgren C.M., Widen E., Tuomi T., Li H., Almgren P., Kanninen T., Melander O., Weng J., Lehto M., Groop L.C., 2002. Contribution of known and unknown susceptibility genes to early-onset diabetes in scandinavia: evidence for heterogeneity. Diabetes 51, 1609-1617.

Lipkin, E., Mosig, M. O., Darvasi A., Ezra E., Shalom A., Friedman A., and Soller M., 1998. Quantitative trait locus mapping in dairy cattle by means of selective milk

DNA pooling using dinucleotide microsatellite markers: analysis of milk protein percentage. Genetics 149, 1557-1567.

Luckinbill L.S., Golenberg E.M., 2002. Genes affecting aging: mapping quantitative trait loci in Drosophila melanogaster using amplified fragment length polymorphisms (AFLPs). Genetica. 114, 147-156.

North K.E., Miller M.B., Coon H., Martin L.J., Peacock J.M., Arnett D., Zhang B., Province M., Oberman A., Blangero J., Almasy L., Ellison R.C., Heiss G., 2005. Evidence for a gene influencing fasting LDL cholesterol and triglyceride levels on chromosome 21q. Atherosclerosis 179, 119-125.

Parnell P.F., 2004. Industry application of marbling genetics: a brief review. Aust. J. Exp. Agr. 44, 697-703.

Plotsky, Y., Cahaner A., Haberfeld A., Lavi U., Lamont SJ, and Hillel J., 1993. DNA fingerprinting bands applied to linkage analysis with quantitative trait loci in chickens. Anim. Genet. 24, 105-110.

Perusse L., Rankinen T., Zuberi A., Chagnon Y.C., Weisnagel S.J., Argyropoulos G., Walts B., Snyder E.E., Bouchard C., 2005. The human obesity gene map: the 2004 update. Obes Res. 13. 381-490.

Price R.A., Li W.D., Bernstein A., Crystal A., Golding E.M., Weisberg S.J., Zuckerman W.A., 2001. A locus affecting obesity in human chromosome region 10p12. Diabetologia 44, 363-366.

Reed D.R., Li X., McDaniel A.H., Lu K., Li S., Tordoff M.G., Price R.A., Bachmanov A.A., 2003. Loci on chromosomes 2, 4, 9, and 16 for body weight, body length,

and adiposity identified in a genome scan of an F2 intercross between the 129P3/J and C57BL/6ByJ mouse strains. Mamm Genome. 14, 302-313.

Savelkoul P.H., Aarts H.J., de Haas J., Dijkshoorn L., Duim B., Otsen M., Rademaker J.L., Schouls L., Lenstra J.A., 1999. Amplified-fragment length polymorphism analysis: the state of an art. J Clin Microbiol. 37, 3083-3091.

Tomecki R., Dmochowska A., Gewartowski K., Dziembowski A., Stepien P.P., 2004. Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. Nucleic Acids Res. 32, 6001-6014.

Vos P., Hogers R., Bleeker M., Reijans M., Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., and Zabeau M., 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23, 4407-4414.

Wilson-Fritch L., Nicoloro S., Chouinard M., Lazar M.A., Chui P.C., Leszyk J., Straubhaar J., Czech M.P., Corvera S., 2004. Mitochondrial remodeling in adipose tissue associated with obesity and treatment with rosiglitazone. J Clin Invest. 114, 1281-1289.

Wu X.-L., MacNeil M.D., De S., Xiao Q.J., Michal J.J., Gaskins C.T., Reeves J.J., Busboom J.R., Wright Jr, R.W., Jiang Z., 2005. Baysian evaluation of candidate genes for beef backfat quantitative trait loci using the SAS system. Genetica 125, 103-113.

Table 1   Pooled DNA samples with adjusted marbling score

Group	Pools	Range	Mean
Low Beef Marbling Score			
	LM1	-2.30 ~ -1.44	-1.7638
	LM2	-1.40 ~ -1.10	-1.2650
High Beef Marbling Score			
	HM2	1.70 ~ 1.15	1.3635
	HM1	3.50 ~ 1.70	2.2438

Table 2   Adapters and primers used in the AFLP analysis

	Name	Sequences
Adapters *EcoRI*	Eco top strand	5'-AGCTGTAGACTGCGTACC
	Eco bottom strand	5'-AATTGGTACGCAGTCTAC
Adapters *TaqI*	Taq top strand	5'- CGGTCAGGACTCATCA
	Taq bottom strand	5'-GATCTGATGAGTCCTGAC

Pre-amplification

EcoRI primers	E01	5'-GACTGCGTACCAATTC A
TaqI primers	T01	5'-GATGAGTCCTGACCGA A
	T02	5'-GATGAGTCCTGACCGA C

Selective-amplification

EcoRI primers	EN	5'-GACTGCGTACCAATTC A NN
TaqI primers	T1N	5'-GATGAGTCCTGACCGA A NN
	T2N	5'-GATGAGTCCTGACCGA C NN

N: nucleotide A, T, C or G

Table 3 Selection of AFLP markers based on fragment frequency.

Primer Combination	Fragments Size (base pairs)	H1	L1	Peak differences [1]
E+AAC/T+ACA	251	0.08	0.23	p<0.01
E+AAG/T+CAT	256	0.67	0.54	p<0.1
E+AGT/T+CAT	239	0.5	0.19	p<0.01
E+AGT/T+ACT	260	0.21	0.02	p<0.01

H1: fragment present ("1") frequencies in high BMS group. L1: fragment present ("1") frequencies in low BMS group. 1: Peak difference between the high and low marbling score groups. The P values were calculated based on Fisher's Exact Test.

Table 4 Bayesian posterior estimation of allelic frequencies for *MspI* AFLP markers.

	Allele G				Allele A				MCE[3]
	pfG[2]	Q2.5	Q50	Q97.5	pfA	Q2.5	Q50	Q97.5	
HB[1]	0.7814	0.6493	0.7863	0.8815	0.2186	0.1185	0.2137	0.3507	6.56E-4
LB	0.9600	0.8962	0.9654	0.9938	0.0400	0.0061	0.0346	0.1038	2.91E-4
F2	0.8847	0.8523	0.8857	0.9129	0.1153	0.0871	0.1143	0.1477	1.60E-4

[1] HB = high BMS pool; LB = low BMS pool; F2 = F2 crosses. [2] pfG and pfA are posterior frequencies of alleles G and A, respectively. Q2.5, Q50, Q97.5 represent 2.5%, 50% (median) and 97.5% quantile of posterior samples, respectively. [3] MCE = Markov chain error.

Table 5 Additive and dominance effects of the candidate gene on beef marbling score (BMS) and subcutaneous fat depth (SFD)

Traits	Genetic effect	Estimate	Standard Error	t Value	Pr > |t|
SFD	additive	-0.0330	0.0349	-0.95	0.3452
	dominance	0.0184	0.0400	0.46	0.6463
BMS	additive	0.5437	0.2130	2.55	0.0114
	dominance	-0.1582	0.2445	-0.65	0.5183
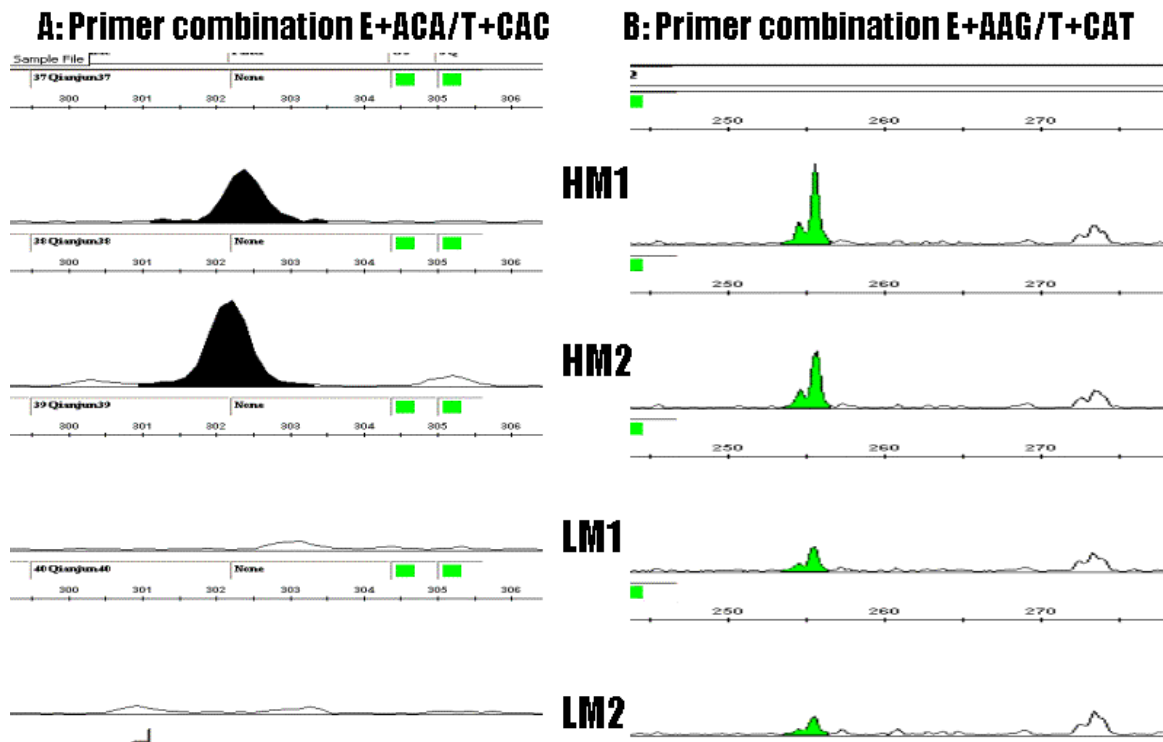
Figure 1. Identification and selection of visually significant AFLP markers for beef marbling. A, an example shows presence/absence patterns of a particular AFLP band between high (two DNA pools, HM1 and HM2) and low (two DNA pools, LM1 and LM2) animals. B, an example shows high/low frequency patterns of a particular AFLP band between high (two DNA pools, HM1 and HM2) and low (two DNA pools, LM1 and LM2) animals.

A

```
AAACTGTCCTTCAAGGTAGTCAACA
AATATTTAGGAATTCACCCTCATGT
ATACTTTCTTCAATTTTGCATAGAA
GTTTATAGGAATGACTTGGTATCAG
AGAACCTTGAATTCAGTTCCTAATT
CTGGCTCTTCTGAAGTTTGAGATCC
TAAGGAAGAAATGTAACCACTCTGC
ACATTAGTTTTCACTTCTGTAACAT
GAAATAGCAGAGTTTGCTGTTAGAA
TAAAATGACATATTATTGGTAAAGT
GCCTCAAAGTAATTGGCACAGGGTT
AGAGTTCAGTCACTTCCCTGGTGGC
TCAGACGGTAAAGCGTCTGCCTGCA
GTGCAGGAAACCCGGGTTCGATCCC
TGGGTTGGGAAGATCCCCTGGAGAA
GGAAATGGCAACCCACTCTAGT
```

B

Figure 2. Characterization of an AFLP marker derived from primer combination E+AGT/T+ACT. A: AFLP marker and its flanking sequences. Primer sequences used for the PCR amplification are marked in blue color. Both *EcoR*I and *Taq*I cut sites are marked in purple color. Mutant sites are underlined. B: Two G/A mutations were detected in the AFLP fragment, one occurred within the *Taq*I cut site and one occurred in the selective primer extension region.
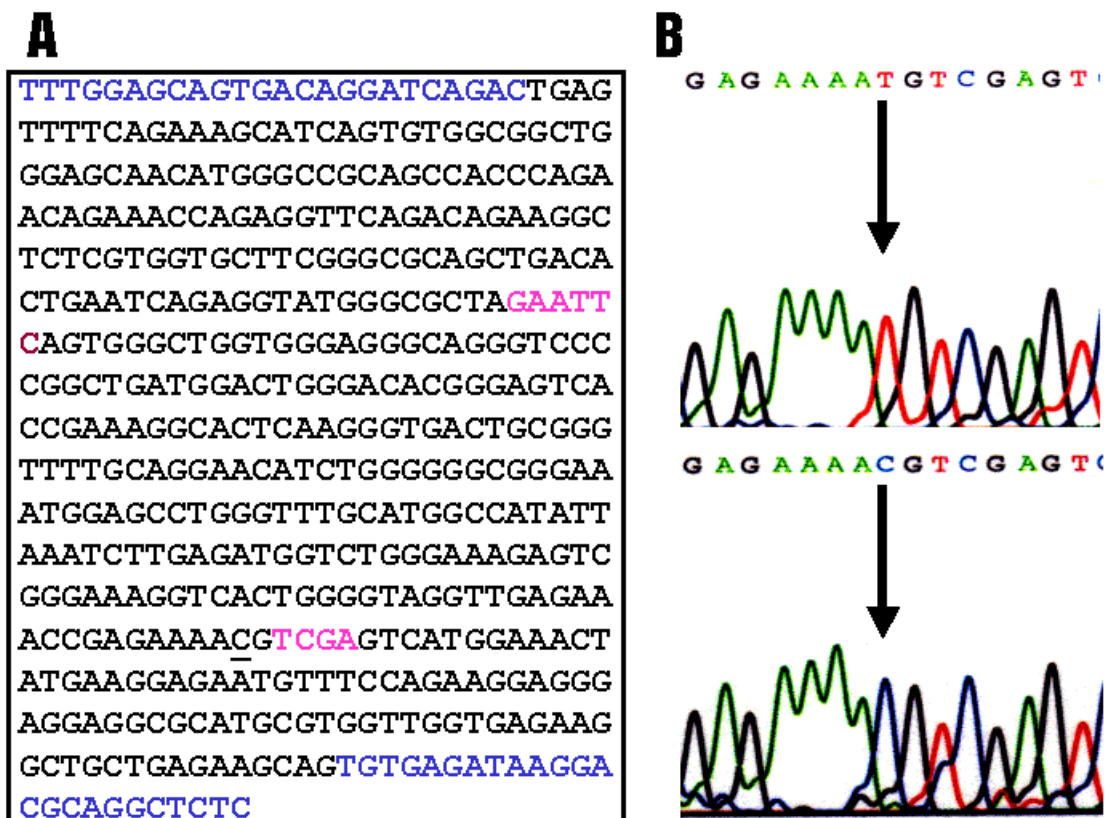
Figure 3. Characterization of an AFLP marker derived from primer combination E+AGT/T+CAT. A: AFLP marker and its flanking sequences. Primer sequences used for the PCR amplification are marked in blue color. Both *EcoR*I and *Taq*I cut sites are marked in purple color. A mutant site is underlined. B: A C/T mutation is detected in the AFLP fragment, but located in the selective primer extension region.
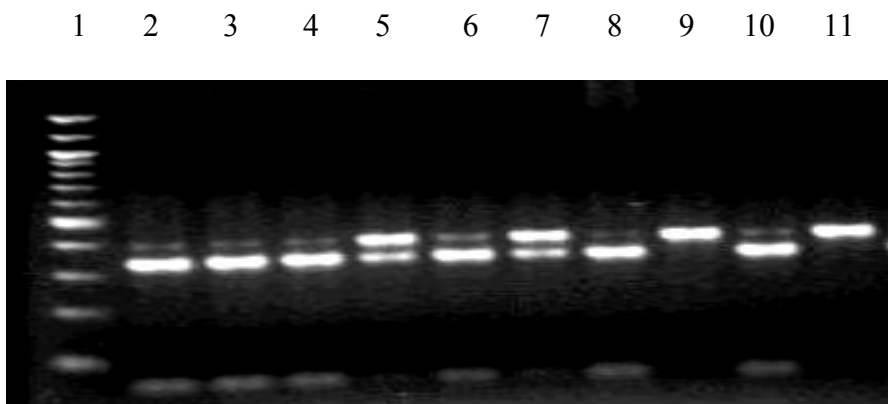
Figure 4. PCR-RFLP genotyping of SNPs in the bovine AFLP marker E+AGT/T+ACT.

Lanes 1, 100 bp ladder. Lane 2 -11, a 402 bp fragment was digested with restriction

enzyme *MspI*. Lanes 2, 3, 4, 6, 8, and lane 10, GG animals (344 + 58 bp); lanes 5 and 7,

GA animals (402 + 344 + 58 bp); lanes 9 and 11, AA animals (402 bp).
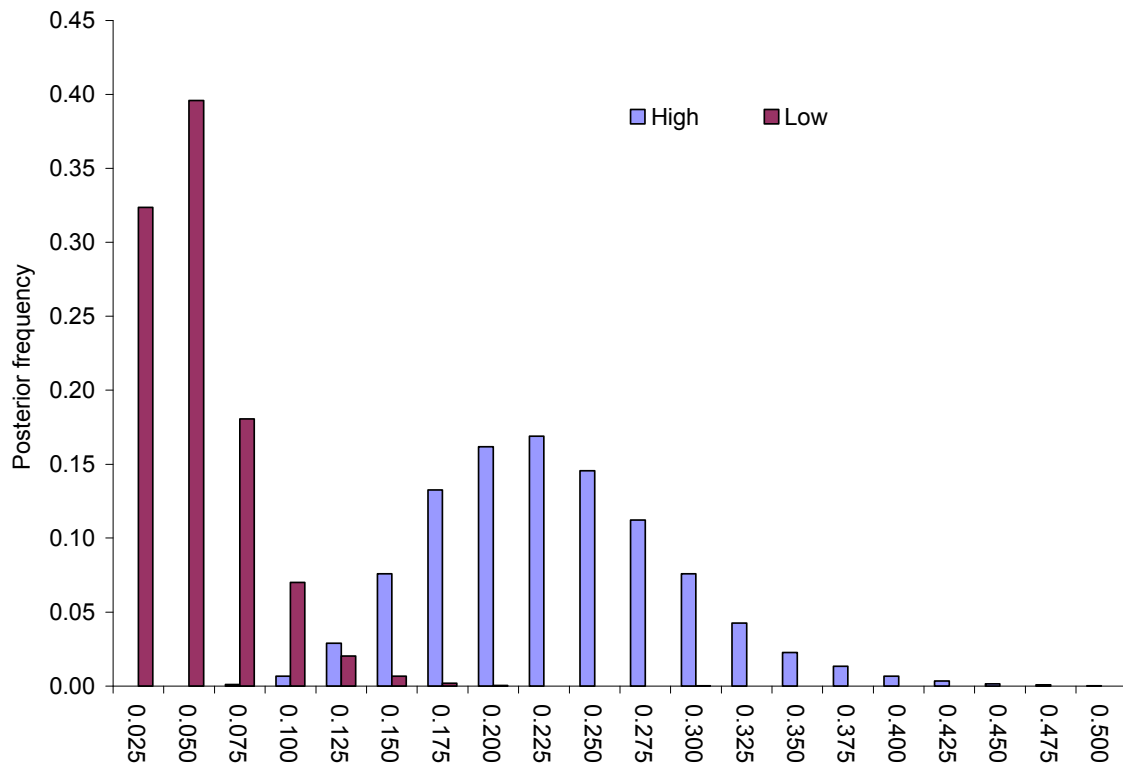
Figure 5. Allele "A" allelic frequency in the high and low BMS groups, respectively

CHAPTER FOUR

A novel nuclear-encoded mitochondrial poly(A) polymerase gene *PAPD1* is a potential candidate for the extreme obesity related phenotypes in Wagyu x Limousin F$_2$ crosses

A novel nuclear-encoded mitochondrial poly(A) polymerase gene *PAPD1* is a potential candidate for the extreme obesity related phenotypes in Wagyu x Limousin F₂ crosses

Qianjun Xiao[1], Xiao-Lin Wu[1], Jennifer J. Michal[1], Jerry J. Reeves[1], Jan R. Busboom[1], Gary H. Thorgaard[2] and Zhihua Jiang[1]

[1]Department of Animal Science, 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

76

Abstract

People with obesity, especially extreme obesity are at risk for many health problems. However, the responsible genes remain unknown in >95% of severe obesity cases. Our previous genome-wide scan using Wagyu x Limousin F_2 crosses revealed a molecular marker significantly associated with intramuscular fat deposition. Characterization of this marker showed that it is orthologous to the human gene *KIAA1462* located on HSA10p11.23, where a major quantitative trait locus for morbid obesity has been reported. In close vicinity of the marker, a newly identified mitochondrial poly(A) polymerase associated domain containing 1 (*PAPD1*) gene has drawn our attention, because the polymerase is required for the polyadenylation and stability of mammalian mitochondrial mRNAs. In the present study, both cDNA and genomic DNA sequences were well annotated for the bovine PAPD1 gene and ten genetic markers were detected in the promoter and exon 1 region. Among seven markers assayed on ~ 250 Wagyu x Limousin F_2 animals, two single nucleotide polymorphisms (SNPs) in the promoter region were significantly associated with intramuscular fat (P<0.05). However, a third SNP, which causes amino acid changes derived from coding exon 1 was significantly interacted with each of these two promoter SNPs on intramuscular fat deposition. In particular, the differences between double heterozygous animals and the slim genotype animals exceeded 2.3 standard deviations for the trait in both cases. Our study provides evidence for a new mechanism – the compound heterosis involved in the extreme obesity, which warrants further examinations.

Keywords: PAP associated domain containing 1; Genetic polymorphism; Compound heterosis; Obesity.

Introduction

Obesity has increased at a fast rate over the past two decades and is now a leading worldwide public health problem. In 1991, only four of 45 participating states in the United States had obesity prevalence rates of 15 to 19% and none had prevalence greater than 20% [1]. By the year 2004, however, seven of 49 participating states had obesity prevalence rates of 15–19 percent, 33 states had rates of 20–24 percent and 9 states had rates more than 25 percent (Centers for Disease Control and Prevention, http://www.cdc.gov/). Overall, more than 65% of US adults are overweight or obese, with nearly 31% of adults (more than 61 million people) meeting criteria for obesity. Furthermore, the greatest increases have been observed in the prevalence of extreme obesity [2]. Unfortunately, although over 600 genes, markers and chromosomal regions have been identified as associated with or linked to human obesity phenotypes [3], the responsible genes are still unknown in >95% of severe obesity cases [4].

Recently, we used an amplified fragment length polymorphism (AFLP) approach to screen genome regions for extreme obesity-related phenotypes on divergent animals derived from Wagyu x Limousin crosses [5]. The Wagyu breed of cattle has been selected for high intramuscular fat for a long time, whereas the Limousin breed has been selected for heavy muscle, which leads to low intramuscular fat. The difference in the trait between these two breeds makes them very unique for mapping genes for obesity-related phenotypes. Characterization of one of the AFLP fragments that were significantly different in frequencies between extreme high or low intramuscular fat

78

animals indicated that it is orthologous to the human gene *KIAA1462* located on HSA10p11.23. Several studies on different populations have shown that this human chromosomal region harbors a major quantitative trait locus linked to or associated with morbid obesity [6-8]. The gene *PAPD1*, poly(A) polymerase associated domain containing 1, which is just adjacent to the human gene *KIAA1462* on the chromosome has drawn our attention. The gene is a newly discovered nuclear-encoded mitochondrial poly(A) polymerase and has a role in mitochondrial RNA processing [9]. The polymerase is required for the polyadenylation and stability of mammalian mitochondrial mRNAs [10].

Mitochondria perform a large number of reactions in mammalian cells. In particular, the major site of fatty acid ß-oxidation occurs in mitochondria [11], which may provide key intermediates for the synthesis of triglycerides via the action of pyruvate carboxylase [12]. Therefore, genes that are involved in regulating and stimulating mitochondrial biogenesis would affect the key aspects of adipogenesis and consequently contribute to development of the obesity-related phenotypes in mammals. Wilson-Fritch and colleagues [13] found a profound decrease of approximately 50% in the levels of transcripts for nuclear-encoded mitochondrial genes accompanying the onset of obesity. Recently, Jiang and colleagues [14] found that the bovine mitochondrial transcription factor A (*TFAM*) was significantly associated with intramuscular fat (marbling scores) and subcutaneous fat depth in the same population of cattle as described above. Here we report for the first time that the *PAPD*1, a novel nuclear-encoded mitochondrial poly(A)

79

polymerase gene contributes significantly to extreme obesity phenotypes due to compound heterosis.

Materials and methods

Compilation of cDNA and genomic DNA sequences for the bovine *PAPD1* gene.
A three-step bioinformatics procedure was used for retrieving both cDNA and genomic DNA sequences of the bovine *PAPD1* gene (Figure 1). In step 1, the cDNA sequence of the human *PAPD1* gene was used as references to retrieve the ortholgous ESTs (expressed sequence tags) against the GenBank database "est_others" with a species option limited to *Bos taurus*. In step 2, several ESTs were chosen and assembled to form a primary cDNA sequence for the bovine gene, which was then used to perform a species-specific ESTs search against the same database in order to expand the primary sequence to a full-length cDNA sequence. Finally, the full-length cDNA sequence was used to search for genomic DNA sequences against the 6X bovine genome sequence database (http://www.hgsc.bcm.tmc.edu/projects/bovine/). Alignment of the full-length cDNA sequence with the genomic DNA sequence determined its genomic organization. The ORF (open reading frame) finder developed by the National Center for Biotechnology Information (NCBI) was used to find all open reading frames and deduce amino acid sequence based on the full-length cDNA sequence of the bovine *PAPD1* gene.

*Primer design for amplification of the bovine **PAPD1** gene.*

Three pairs of primers were designed to detect genetic polymorphisms in the bovine

PAPD1 gene, based on the genomic sequence generated from the bovine genome

sequencing project. The first primer pair (forward sequence, 5'-GAG TGT GGT GGT

TAG GGG TGG TAG-3' and reverse sequence, 5'-TTC ACA GTA GGG TTT CCC

TTC CTC-3') targets a promoter region of 517 bp from bases 9,048 to 9,564 (GenBank

accession no. AAFC02034082). The second primer pair (forward sequence, 5'-CCC

CGT CTC TGG TTC TAT TTT CAA T 3' and reverse sequence, 5' GCG ACT CCA

GAC TCT TCC TCC TG 3') amplifies an entire exon 1 plus partial promoter sequence,

5'UTR and partial intron 1 sequence with a product of 481 bp in length (GenBank

accession no. AAFC02034082). The last primer pair (forward sequence, 5'-TTC ATT

GAG TTA GAC AAG GCT GTG-3' and reverse sequence, 5'-TAG CTC CAA TAC

GTG TCA ATT TTT-3') is based on GenBank accession no. AAFC02139310 by

amplifying a product of 543 bp, including partial intron 6 sequence, entire exon 7,

complete intron 7, entire exon 8 and partial intron 8 sequence.

Detection of the genetic polymorphisms in the bovine *PAPD1* gene.

The animals used in the present research were derived from a Wagyu x Limousin cross,

including 6 F_1 bulls, 113 F_1 dams and 246 F_2 progeny. Two obesity-related phenotypes

included beef marbling score (BMS) and subcutaneous fat depth (SFD). The former was

a subjective measure of the amount of intramuscular fat in the *longissimus* muscle based

on USDA standards, while the latter 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. Genetic polymorphisms were screened by comparison

of sequence patterns on six F1 bulls. The PCR reactions were performed in a total volume of 10 µl, including 25ng of bovine genomic DNA, 1 fold of Platin um *Taq* Buffer (20 mM Tris-HCl, PH8.4, 50 mM KCl; Invitrogen), 3.0 mM $MgCl_2$, 0.3 mM each of the four dNTPs, 1 U Platinum Taq polymerase (Invitrogen) and 25ng of each primer. The PCR reactions started with denaturation at 94°C for 2 min, followed by 8 touch-down cycles: 30 seconds at 94°C, 30 seconds at 71°C to 63°C (i.e. decreasing 1°C per cycle), and 30 seconds at 72°C. Then, the PCR took another 37 cycles of reactions: 30 seconds at 94°C, 30 seconds at 63°C and 30 seconds at 72°C, and finally ended after 5 min at 72°C. The PCR products were then sequenced on an ABI 3730 sequencer in the Laboratory for Biotechnology and Bioanalysis (Washington State University) following a standard protocol.

Assay development for marker genotyping in the bovine **PAPD1** *gene.*

Six single nucleotide polymorphisms (SNPs) were detected in the promoter region, which were then genotyped individually using a direct PCR direct sequencing approach. Four SNPs were found around the exon 1 region, including two coding SNPs and two genomic SNPs, respectively. Of these two coding SNPs, one is a silent mutation and one is a missense mutation. Fortunately, the latter mutation can be revealed using the PCR-RFLP (restriction fragment length polymorphism) approach. The PCR amplicons were digested at 37°C for three hours with 2U of *Stu*I for the G/A SNP, followed by analysis on 1.6% agarose gels and visualized by staining with ethidium bromide. However, no polymorphisms were discovered in the product spanning exon 7 and exon 8. Therefore,

82

among ten SNPs detected in the bovine *PAPD1* gene, seven were successfully assayed on the reference population.

Association of the bovine **PAPD1** *gene with the obesity-related phenotypes.*

The phenotypic data have been previously adjusted for the difference in year of birth, sex, age (days), live weight (kilograms), or fat depth (inches), as appropriate. The adjusted phenotypes were then used in subsequent association analysis based on the following mixed model:

$$y = X\beta + Zu + e \tag{1}$$

where y was a vector containing all observations, β was a vector for fixed effects including marker all genotypes or genotype combinations, $u \sim N(0, A\sigma_u^2)$ was a vector for residual genetic effects that were not accounted for by the genes (markers) under investigation, A was a numerator relationship matrix, X and Z were incidence matrices which link observations in y to effects in β and u, respectively, and $e \sim N(0, I\sigma_e^2)$ was a vector for residual errors. Model parameters were estimated using REML (residual maximum likelihood) and Bayesian analysis, respectively. In the Bayesian estimation, we assumed normal priors for fixed effects and chi-square distributions for the variance components, σ_u^2 and σ_e^2, respectively. Markov chain Monte Carlo (MCMC) was used to simulated model parameters based on their fully conditional distributions below.

$$\beta, u \mid \sigma_u^2, \sigma_e^2, y \sim N\left(\begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix}, \begin{bmatrix} C_{\beta\beta} & C_{\beta u} \\ C_{u\beta} & C_{\beta\beta} \end{bmatrix} \right) \tag{2}$$

where:

$$\begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} C_{\beta\beta} & C_{\beta u} \\ C_{u\beta} & C_{\beta\beta} \end{bmatrix} \begin{bmatrix} X'y + 1\beta_0 \dfrac{\sigma_e^2}{\sigma_\beta^2} \\ Z'y \end{bmatrix}; \qquad \begin{bmatrix} C_{\beta\beta} & C_{\beta u} \\ C_{u\beta} & C_{\beta\beta} \end{bmatrix} = \begin{bmatrix} X'X + I\dfrac{\sigma_e^2}{\sigma_\beta^2} & X'Z \\ Z'X & Z'Z + A^{-1}\dfrac{\sigma_e^2}{\sigma_u^2} \end{bmatrix}^{-1}$$

$$\sigma_u^2 \mid \beta, u, \sigma_e^2, y \sim \left(\sum_{i=1}^{q} \sum_{i'=1}^{q} a^{ii'} u_i u_{i'} + S_u^2 \right) \chi_{q+\upsilon_u}^{-1} \tag{3}$$

$$\sigma_e^2 \mid \beta, u, \sigma_u^2, y \sim \left(r'r + S_e^2 \right) \chi_{n+\upsilon_e}^{-1} \tag{4}$$

where $a^{ii'}$ was the element in row i and column i' of the $q \times q$ inverse matrix A^{-1},

$r = y - X\beta - Zu$ was a vector containing residual errors, υ_k and S_k^2 are the degrees of

freedom parameter and the scaling factor in the prior distribution of $\sigma_k^2 \sim S_k^2 \chi_{\upsilon_k}^{-1}$

($k = u, e$).

Results

Annotation of the bovine *PAPD1* gene

The BLAST search using the cDNA sequence of the human *PAPD1* gene (NM_018109)

retrieved more than 20 orthologous ESTs in cattle against the "est_others" in the

GenBank database. Three ESTs - DV819803, DV879788 and DN525471 were used to

form a preliminary cDNA sequence of the bovine *PAPD1* gene (Figure 1), which was

then used as a query for a second BLAST run to retrieve both 5' and 3' cDNA end

sequences missed by the first BLAST run using the cDNA sequence of the human gene

as a reference. Adding two ESTs (DV822230 and CV974359) into three ESTs described

as above for a final assembly contributed to a formation of a full-length cDNA sequence

of 2,554 bp for the bovine gene. This newly assembled bovine gene hit more than 30

ESTs in the database with sequence identity of 98 – 100%. ORF finder detected a longest coding sequence of 1,701 bp with total 566 deduced amino acids for the bovine PADA1 protein. The full-length cDNA sequence of the bovine *PAPD1* gene retrieved six genomic DNA contigs (AAFC02034082, AAFC02215345, AAFC02219529, AAFC02142915, AAFC02142935 and AAFC02139310) from the 6X bovine genome sequence database with a total of 50,536 bp (Figure 1). Alignment of the cDNA with genomic DNA sequences revealed that this bovine gene contains nine exons and eight introns. Among these eight introns, however, only introns 2, 7 and 8 have completed sequences.

Genetic polymorphisms and genotyping

Three regions of the bovine *PAPD1* gene were amplified and sequenced on six F1 bulls individually to screen polymorphisms. Sequencing data showed that the promoter region was highly polymorphic: a total of six genetic polymorphisms were detected in a product of 517 bp, with one mutation per every 86 bp of sequence. These six polymorphisms include five SNPs (AAFC02034082.1:g.*9224G>A, 9330A>C, 9367G>C, 9419A>C* and *9435T>G*) and one insertion/deletion (AAFC02034082.1:g.*9409[(G)4]+[(G)5]*, respectively. We decided to genotype all these genetic markers in the promoter region using a direct sequencing PCR product approach. In the exon 1 and its flanking regions, four SNPs were detected, including two genomic SNPs (AAFC02034082.1:g.*10084T>C* and *10228G>T*) and two coding SNPs (AAFC02034082.1:c.*10326C>T* and *10364A>G*). The latter coding SNP (AAFC02034082.1:c.*10364A>G*) changes amino acids from aspartic acid (GAC) to glycine (GGC) and cause a gain/loss of enzyme cut site for *Stu*I.

Therefore, a PCR-RFLP assay was developed to genotype this marker on all individuals. No polymorphisms were discovered around exon 7 and exon 8 region of the bovine *PAPD1* gene.

Associations of bovine *PAPD1* gene with beef marbling and SFD
REML-based analysis of variance suggested significant associations of two SNPs in the promoter region (AAFC02034082.1:*g.9367 G>C* and *9419A>C*) with BMS, and was suggestive of an association of AAFC02034082.1:*c.10364A>G* with SFD (Table 1). Bayesian analysis was in agreement with the REML-based analysis of variance, yet it gave a more intuitive inference of the association by giving their posterior distribution. For AAFC02034082.1:*g.9367G>C*, only CG and GG genotypes exist in the F2 progeny. However, for AAFC02034082.1:*g.9419A>C*, only AC and AA genotypes were considered in the analysis, because only two CC individuals were detected in the F2 progeny. In the Bayesian analysis, over 95% posterior samples showed positive differences between *PAPD1:g.9367CG* and *GG* or negative difference between *PAPD1:g.9419AC* and *AA* (Figure 2). Therefore, these deviates of means between one heterozygous genotype and one homozygous genotype could be interpreted as allele substitution effects estimated under the assumption of complete linkage between the marker and the causal gene.

We further analyzed pair-wise interactions between these two markers (AAFC02034082.1:*g.9367G>C* and *9419A>C*), which were significantly associated with BMS, and other markers in the bovine *PAPD1* gene. REML analyses indicated only AAFC02034082.*1:c.10364A>G* was significantly interacted with these two markers for

beef marbling. Due to the singularity of matrix X'X, interaction effects are not uniquely estimable in the REML analysis. We therefore presented interaction estimates obtained using Bayesian estimation. Interestingly, interactions between two heterozygous genotypes led to the highest marbling scores (Figure 3A and 3B). Animals with both heterozygous genotypes of *PAPD1: c.10364AG* and *g.9367CG* had an additional 2.33 and 0.67 - 0.69 marbling scores compared to the genotype combination of *PAPD1: c.10364GG* with *g.9367CG* and all other combinations, respectively (Figure 3A). The same interaction trend was also observed for markers AAFC02034082.1:*c.10364A>G* and *g.9419A>C*. The marbling score was 2.40 higher in animals with double heterozygote genotypes (*PAPD1: c.10364AG* and *g.9419AC*) than the genotypes with *PAPD1: c.10364GG* and *g.9419AC*. The remaining genotypes at both markers were lower than the double heterozygous genotypes by 0.64 – 0.70 marbling scores (Figure 3B).

Discussion

Oxidative phosphorylation is the most important event performed in mitochondria, which transforms the energy of nutrient-derived substrates into the energy stored in ATP [11]. Mitochondria supply the majority of this ATP to the rest of cell. When energy intake exceeds energy expenditure, the resulting imbalance may expand the size and increase the number of fat cells. This process basically leads to fat deposition in humans and in mammals. Therefore, the genes that regulate and stimulate the mitochondrial biogenesis would play an important role in adipogenesis [13]. In the present study, we found that

the *PAPD1* gene, a novel nuclear-encoded mitochondrial poly(A) polymerase contributes to extreme fat deposition capability in cattle. Evidence has shown that the *PAPD1* gene is involved in regulating the mitochondrial biogenesis by controlling poly(A) synthesis in human mitochondria [10]. Recently, Jiang and colleagues [14-15] studied effects of the mitochondrial transcription factor A (*TFAM*), B1 (*TFB1M*) and B2 (*TFB2M*) on obesity related phenotypes using the same reference population as described above. Overall, we found that involvement of these four genes, *TFAM*, *TFB1M*, *TFB2M* and *PAPD1* in stimulating and regulating of the mitochondrial biogenesis may be tissue-specific or relevant. *TFAM* contributed significantly to both intramuscular fat deposition and SFD [14], while *TFB1M* had no effect on either trait. *PAPD1* promoted more on intramuscular fat, but not significantly on SFD (this study). However, *TFB2M* contributed more to SFD, but almost nothing to intramuscular fat deposition [15].

In 1998, Hager and colleagues [6] first reported significant evidence for linkage of obesity to a chromosome 10 locus with a maximal logarithm of odds score (MLS) near the D10S197 marker based on a genome-wide scan on 158 multiplex French obese Caucasian families with 514 individuals. This obesity-linked peak at D10S197 on HSA10p11-12 was further confirmed in four other ethnic groups, including 170 European-American and 43 African-American families [8], 89 German families [16] and 188 nuclear families around Paris, France[17]. As D10S197 is located in intron 7 of the *GAD2* gene at position 26.54 Mb on HSA10, Boutin and colleagues [17] proposed that the gene *GAD2* encoding the glutamic acid decarboxylase enzyme (GAD65) is a positional candidate gene for obesity on the chromosome. Both a case-control study (575

morbidly obese and 646 control subjects) and a familial-based analysis analyzing *GAD2*

variants confirmed the association with the obesity of SNP +61450 C>A and +83897

T>A haplotype (χ^2 = 7.637, p = 0.02). However, a fine mapping performed on the same

population by the authors revealed a second peak at D10S600 (MLS = 3.4 vs. MLS = 3.2

at D10S197). The marker D10S600 is located at position 28.70 on HSA10, just 2.16 Mb

away from the marker D10S197 (26.54 Mb on HSA10). Recently, Hinney and co-

workers [7] further moved the Chromosome 10 obesity peak down to the marker *TCF8*

with a maximum LOD of 2.32. The authors also found that the two point LOD score was

1.38 only at the marker D10S197. The *TCF8* marker is positioned at 31.65 Mb on

HSA10. Interestingly, the *PAPD1* gene we studied here is located at position 30.64 Mb

on HSA10 between both markers D10S600 and TCF8. These data warrant further

investigations on the candidacy of *PAPD1* gene as a potential candidate gene for the

second obesity peak on HSA10p11-12.

Evidence has shown that heterosis is one of the contributors responsible for the obesity-

related phenotypes in mammals. Vaisse and colleagues [18] reported a high frequency

(4%) of rare heterozygous *MC4-R* mutations in a large population of morbidly obese

patients, but no such mutations were found in controls. Interestingly, *HTR2C* promoter

variation was found to be a risk factor for obesity and, perhaps through heterosis,

influences weight loss in obese women [19]. Among ninety-five obese women who

participated in a randomized trial of psychological treatments for weight loss,

heterozygotes lost less weight during the trial than did homozygotes (6.8 vs. 9.7 kg; P =

0.047) and weighed more 6 months (90.1 vs. 83.6 kg; P = 0.006) and 12 months (91.8 vs.

84.6 kg; P = 0.009) later. Heterozygotes also had higher triglyceride levels than homozygotes, while C/C subjects in the obesity trial did not differ from T/T subjects in terms of weight loss or triglycerides. In an F2 intercross obtained by mating the OLETF and Fischer-344 (F344) rats, Ogino and co-workers [20] found that the Chr 3 QTL exhibits heterosis, heterozygotes showing significantly higher glucose levels than OLETF or F344 homozygotes. In our study, we detected two compound heteroses in the bovine *PAPD1* gene that contribute to the extreme obesity-related phenotypes. The differences between *PAPD1: c.10364AG - g.9367CG* and *PAPD1: c.10364GG - g.9367CG*, and between *PAPD1: c.10364AG - g.9419AC* and *PAPD1: c.10364GG - g.9419AC* were 2.33 and 2.40 marbling scores, respectively (Figure 3a and 3b). In the Wagyu x Limousin F2 population, marbling scores ranged from 4 to 9.5 with a standard deviation of 1 marbling score. Therefore, these differences correspond to 2.33 and 2.40 standard deviations, indicating that double heterozygotes leads to extreme high beef marbling score animals. However, how the double heteroses produces a high amount of intramuscular fat deposition needs to be further examined.

References

[1] National Task Force on the Prevention and Treatment of Obesity, Overweight, obesity, and health risk, Arch Intern Med. 160 (2000) 898-904.

[2] A.M. Spiegel, B.M. Alving, Executive summary of the Strategic Plan for National Institutes of Health Obesity Research, Am J Clin Nutr. 82 (2005) 211S-214S.

[3] L. Perusse, T. Rankinen, A. Zuberi, Y.C. Chagnon, S.J. Weisnagel, G. Argyropoulos, B. Walts, E.E. Snyder, C. Bouchard, The human obesity gene map: the 2004 update, Obes Res. 13 (2005) 381-490.

[4] J.S. Flier, Obesity wars: molecular progress confronts an expanding epidemic, Cell 116 (2004) 337-50.

[5] Q. J. Xiao, T. A. Wibowo, X-L. Wu, J. J. Michal, J. J. Reeves, J.R. Busboom, G.H. Thorgaard, Z. Jiang, A simplified QTL mapping approach for screening and mapping of novel AFLP markers associated with beef marbling, Journal of Biotechnology (submitted).

[6] J. Hager, C. Dina, S. Francke, S. Dubois, M. Houari, V. Vatin, E. Vaillant, N. Lorentz, A. Basdevant, K. Clement, B. Guy-Grand, P. Froguel, A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10, Nat Genet. 20 (1998) 304-8.

[7] A. Hinney, A. Ziegler, F. Oeffner, C. Wedewardt, M. Vogel, H. Wulftange, F. Geller, K. Stubing, W. Siegfried, H.P. Goldschmidt, H. Remschmidt, J. Hebebrand, Independent confirmation of a major locus for obesity on chromosome 10, J Clin Endocrinol Metab. 85 (2000) 2962-2965.

[8] R.A. Price, W.D. Li, A. Bernstein, A. Crystal, E.M. Golding, S.J. Weisberg, W.A. Zuckerman, A locus affecting obesity in human chromosome region 10p12, Diabetologia 44 (2001) 363-6.

[9] R. Tomecki, A. Dmochowska, K. Gewartowski, A. Dziembowski, P.P. Stepien, Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase, Nucleic Acids Res. 32 (2004) 6001-6014.

[10] T. Nagaike, T. Suzuki, T. Katoh, T. Ueda, Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase, J Biol Chem. 280 (2005) 19721-19727.

[11] A.S. Reichert, W. Neupert, Mitochondriomics or what makes us breathe, Trends Genet. 20 (2004) 555-562.

[12] O.E. Owen, S.C. Kalhan, R.W. Hanson, The key role of anaplerosis and cataplerosis for citric acid cycle function, J Biol Chem. 277 (2002) 30409-30412.

[13] L. Wilson-Fritch, A. Burkart, G. Bell, K. Mendelson, J. Leszyk, S. Nicoloro, M. Czech, S. Corvera, Mitochondrial biogenesis and remodeling during adipogenesis and in response to the insulin sensitizer rosiglitazone, Mol Cell Biol. 23 (2003) 1085-1094.

[14] Z. Jiang, T. Kunej, J.J. Michal, C.T. Gaskins, J.J. Reeves, J.R. Busboom, P. Dovc, R.W. Wright 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. 334 (2005) 516-523.

[15] Z. Jiang, T. Kunej, T.A. Wibowo, J. J. Michal, Z.W. Zhang, C.T. Gaskins, J.J. Reeves, J.R. Busboom, P. Dovc, R.W. Wright Jr., The basal nucleus-encoded mitochondrial transcription genes and meat quality in beef cattle, Proceedings of 8th World Congress on Genetics Applied to Livestock Production 2006 Belo Horizonte, MG, Brazil.

[16] K. Saar, F. Geller, F. Ruschendorf, A. Reis, S. Friedel, N. Schauble, P. Nurnberg, W. Siegfried, H.P. Goldschmidt, H. Schafer, A. Ziegler, H. Remschmidt, A. Hinney, J. Hebebrand, Genome scan for childhood and adolescent obesity in German families, Pediatrics 111 (2003) 321-327.

[17] P. Boutin, C. Dina, F. Vasseur, S. Dubois, L. Corset, K. Seron, L. Bekris, J. Cabellon , B. Neve, V. Vasseur-Delannoy, M. Chikri, M.A. Charles, K. Clement, A.

Lernmark, P. Froguel, GAD2 on chromosome 10p12 is a candidate gene for human obesity, PLoS Biol. 1 (2003) E68.

[18] C. Vaisse, K. Clement, E. Durand, S. Hercberg, B. Guy-Grand, P. Froguel, Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. J Clin Invest. 106 (2000) 253-262.

[19] E.C. Pooley, C.G. Fairburn, Z. Cooper, M.S. Sodhi, P.J. Cowen, P.J. Harrison, A 5-HT2C receptor promoter polymorphism (HTR2C - 759C/T) is associated with obesity in women, and with resistance to weight loss in heterozygotes, Am J Med Genet B Neuropsychiatr Genet. 126 (2004)124-127.

[20] T. Ogino, D.H. Moralejo, M. Zhu, K. Toide, S. Wei, K. Wei, T. Yamada, A. Mizuno, K. Matsumoto, K. Shima, Identification of possible quantitative trait loci responsible for hyperglycaemia after 70% pancreatectomy using a spontaneously diabetogenic rat, Genet Res. 73 (1999) 29-36.

Table 1 Analysis of variance for the association tests between polymorphic markers (AAFC02034082) with beef marbling score (BMS) and subcutaneous fat depth (SFD)

| Marker	DF	SS	MS	F value	Pr>F
------ BMS (mean square of errors: 0.94) ------					
g.9224G>A	2	2.06	1.03	1.10	0.3352
g.9330A>C	2	0.35	0.18	0.19	0.8260
g.9367G>C	1	4.54	4.54	4.85	0.0288
g.9409[(G)4]+[(G)5	2	1.15	0.58	0.61	0.5419
g.9419A>C	1	3.78	3.78	4.03	0.0460
g.9435T>G	2	1.47	0.74	0.78	0.4577
c.10364A>G	2	1.08	0.54	0.58	0.5621
------ SFD (mean square of errors: 0.02) ------					
g.9224G>A	2	0.12	0.06	2.44	0.0894
g.9330A>C	2	0.04	0.02	0.80	0.4520
g.9367G>C	1	0.04	0.04	1.54	0.2160
g.9409[(G)4]+[(G)5	2	0.07	0.04	1.50	0.2265
g.9419A>C	1	0.01	0.01	0.49	0.6147
g.9435T>G	2	0.02	0.01	0.46	0.6293
c.10364A>G	2	0.13	0.06	2.64	0.0739
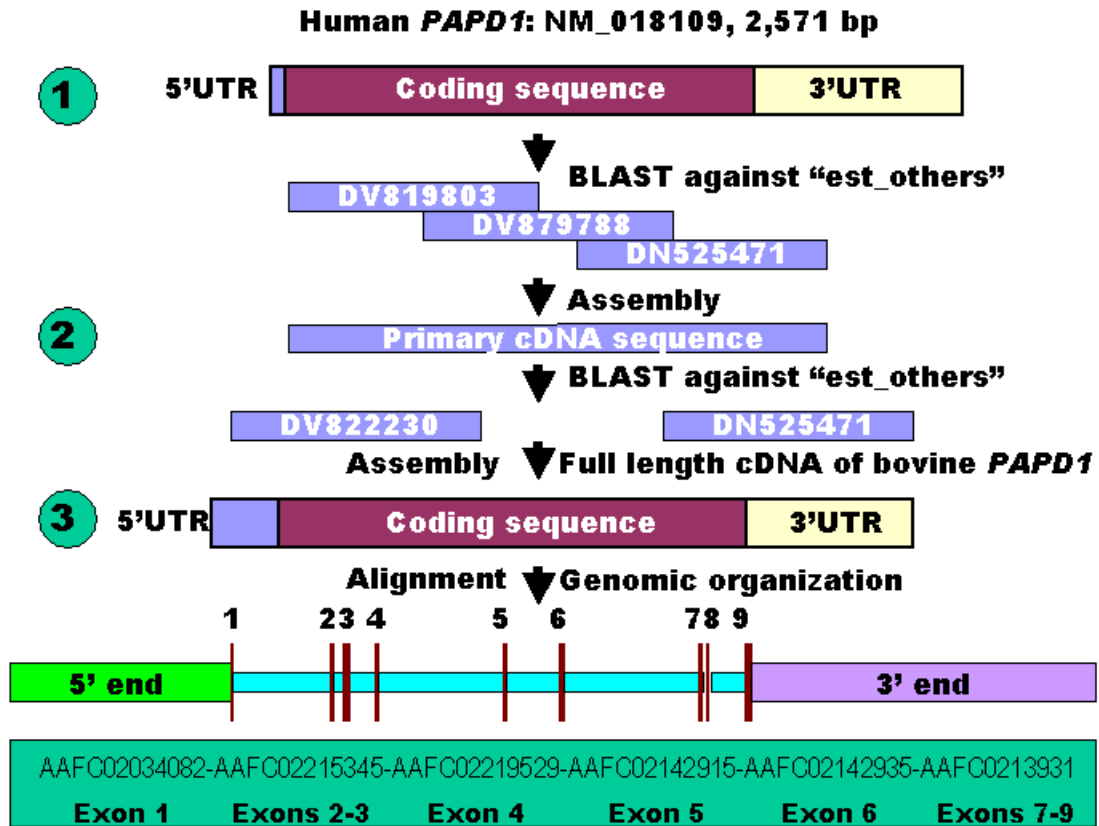
Figure 1. Comparative annotation of both cDNA and genomic DNA sequences of the bovine *PAPD1* gene.
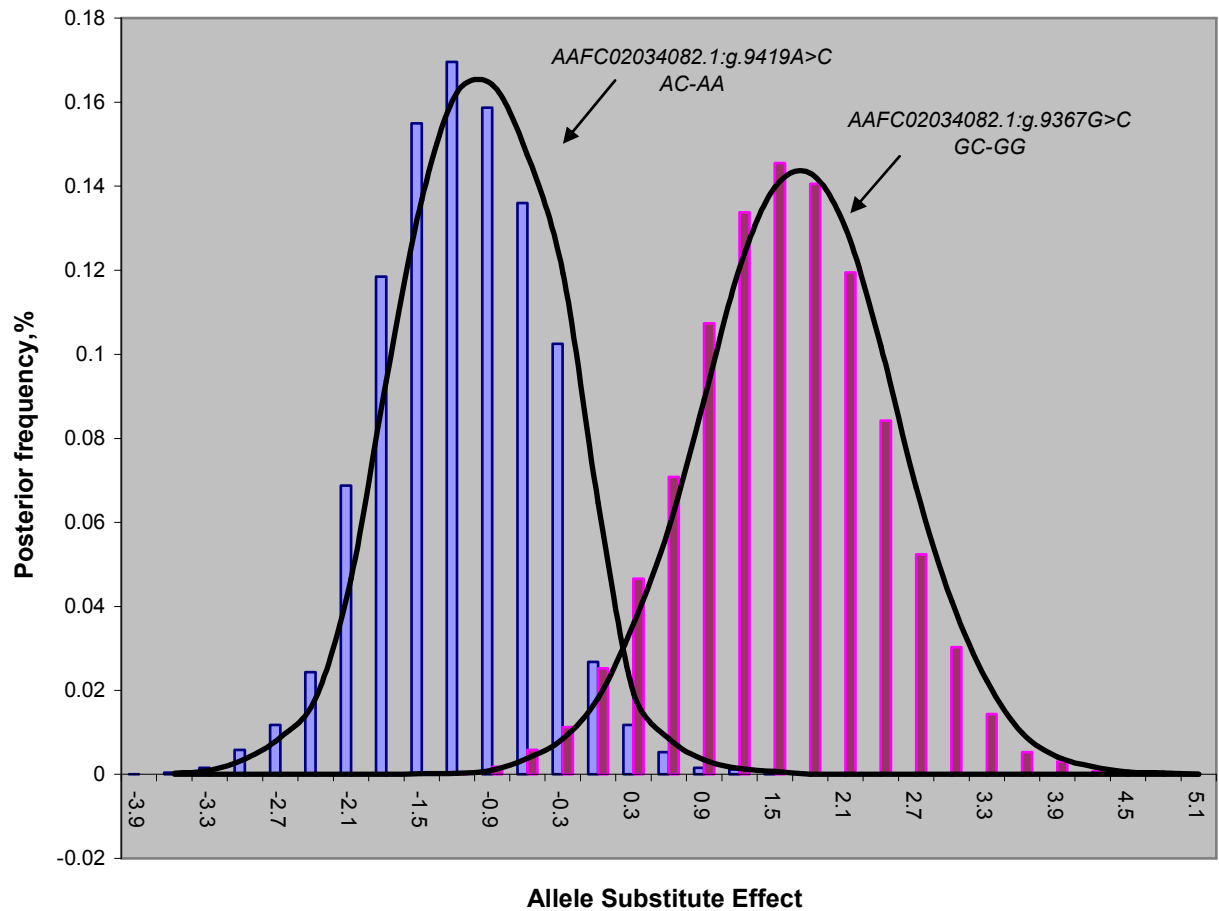
Figure 2 Posterior distribution of estimated allele substitution effect of two SNPs in the promoter region (AAFC02034082.1:*g.9367G>C* and *g.9419A>C*) on beef marbling score under the assumption of complete linkage between the markers and the causal gene.
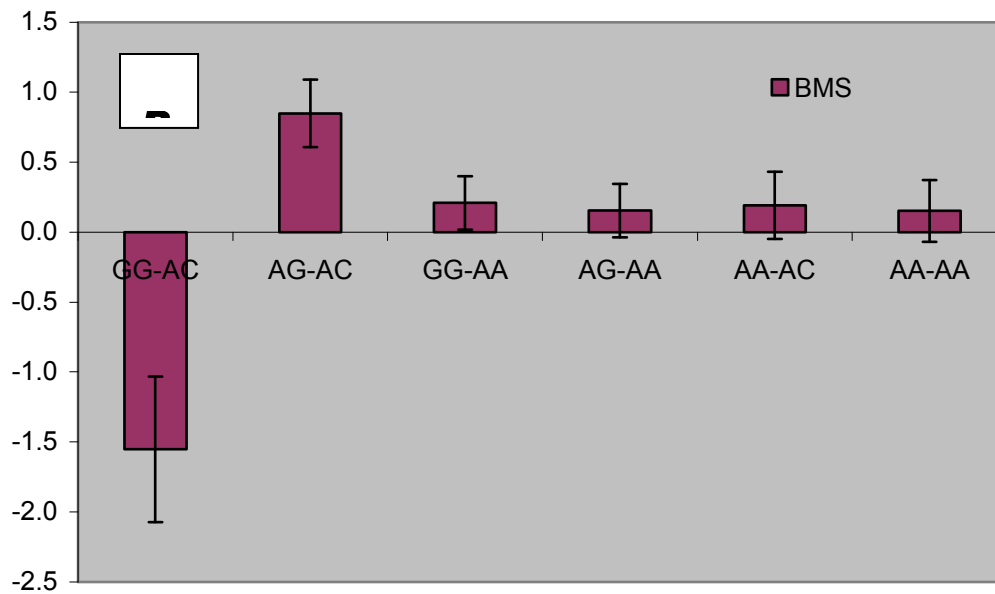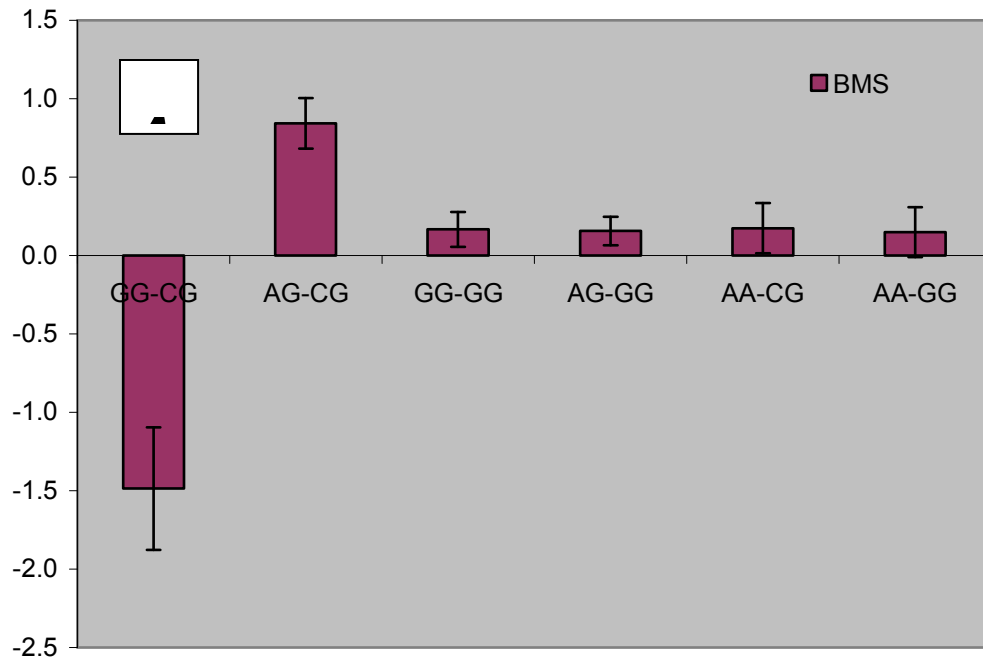
Figure 3   Bayesian estimation of interaction effects between markers: (a) *c.10364A>G - g.9367G>C*, and (b) *c.10364A>G - g.9419A>C*, on beef marbling score.

APPENDIX

APPENDIX

AFLP ANALYSIS PROTOCAL

1) *Genomic DNA Extraction*

Genomic DNA is extracted by using standard phenol:chloroform technique (Ausubel *et al.,* 1989). The DNA concerntration and quality was measured based on absorbance of UV light at 260 nm (A_{260}) and 280 nm (A_{280}) by 384-well microplate spectrophotometer SPECTRAmax (Molecular devices, Sunnyvale, California) through DNA plate blank method. Data were obtained by SoftPro Max (version 4.6) software (Molecular devices, Sunnyvale, California). The quantity was obtained by A and quality was determined by A_{260} : A_{280} ratio.

2) *Genomic DNA Digestion*

Individual or pooled DNA is digested with two restriction enzymes. After restriction enzyme digestion, three types of DNA fragments are generated: *EcoRI – EcoRI* (~1%), *EcoRI – TaqI* (~10%), *and TaqI – TaqI (~ 89%)*. The DNA fragments which have different restriction cutting sites (*EcoRI – TaqI*) are what will be amplified later.

Digestion conditions dependent on what restriction enzymes are used. For *EcoRI/TaqI* pairs, two steps are taken. First, genomic DNA (200ng) is incubated for 3 h at 65°C with 6 U *TaqI* in 1X BSA, 1X *TaqI* buffer and add distilled water to a total volume of 20ul. The digestion reaction is inactivated at 80°C for 20 min. The digestion products are then checked by taking 5.0ul to run in a 1.6% agarose gel which containing 0.0032% Ethidium Bromide. The samples are run 50 min at 130 Volts and then visualized under UV light. Second, the *TaqI* digested DNA is incubated for 3 h at 37°C with 6 U *EcoRI* in 1X *EcoRI* buffer and add distilled water to a total volume of 30ul. The reaction is inactivated at 65°C for 10 min. The digestion products are then checked in a 1.6% agarose gel.

3) *Ligation*

Adapters are ligated to the resulting restriction fragment ends. Oligonucleotide adapters are designed from the restriction enzyme recognition site sequence. They are *EcoRI*

adapters and *TaqI* adapters (Table 1). Adapters are prepared by annealing two strands together to construct double strand adapters. In detail, two strands (same concentration and same volume) are mixed and heated at 96°C for 10 minutes. Cool down gradually.

The restriction digestion products (200ng) are incubated overnight at room temperature with 25 weiss U T4 ligase in a solution containing 5 pMol *EcoRI* – adapters, 50 pMol *TaqI* – adapters, 1X T4 ligase buffer, adding distilled water up to a total volume of 50.0 μl. Template DNA is then diluted 1:10 with 10mM Tris-HCl, 0.1mM EDTA (pH 8.0) before further use.

Table 1 Adapters used in AFLP analysis

	Name	Sequence
Adapters *EcoRI*	Eco top strand	5'-AGCTGTAGACTGCGTACC
	Eco bottom strand	5'-AATTGGTACGCAGTCTAC
Adapters *TaqI*	Taq top strand	5'- CGGTCAGGACTCATCA
	Taq bottom strand	5'-GATCTGATGAGTCCTGAC

4) *Pre-Amplification*

EcoRI primers and *TaqI* primers are designed from the sequence of adapters with one base overhang for each primer (Table 2). Amplify the DNA restriction fragments with ends of *EcoRI* – *TaqI*. The PCR conditions differed depending on the AFLP primers used. The following PCR reaction mix was used: 10 ng of DNA template, 1 fold of Platinum Taq Buffer (20 mM Tris-HCl, PH8.4, 50 mM KCl; Invitrogen), 3.0 mM $MgCl_2$, 0.3 mM each of the four dNTPs, 1 U Platinum Taq polymerase (Invitrogen) and 0.2 pMol of pre-amplification *EcoRI* (E01) and 2 pMol of pre-amplification *TaqI* (T01 or T02) primers in a total volume of 50 μl. The PCR program is as following: 2 min at 94°C, 2 min at 72°C, 25 cycles of 10 seconds at 94°C, 30 seconds at 56°C and 2 min at 68°C, followed by 30 min at 60°C, 4°C store. The pre-amplification products are diluted 1:20 with 10mM Tris-HCl, 0.1mM EDTA (pH 8.0) before further use.

Table 2   Primers used in AFLP analysis

Name	Sequence

101

Pre-amplification

EcoRI primers	E01	5'-GACTGCGTACCAATTC A
TaqI primers	T01	5'-GATGAGTCCTGACCGA A
	T02	5'-GATGAGTCCTGACCGA C

Selective-amplification

EcoRI primers	EN	5'-GACTGCGTACCAATTC A NN
	E32	GACTGCGTACCAATTC A AC
	E33	GACTGCGTACCAATTC A AG
	E35	GACTGCGTACCAATTC A CA
	E38	GACTGCGTACCAATTC A CT
	E39	GACTGCGTACCAATTC A GA
	E42	GACTGCGTACCAATTC A GT
	E44	GACTGCGTACCAATTC A TC
	E45	GACTGCGTACCAATTC A TG
TaqI primers	T1N	5'-GATGAGTCCTGACCGA A NN
	T32	GATGAGTCCTGACCGA A AC
	T33	GATGAGTCCTGACCGA A AG
	T35	GATGAGTCCTGACCGA A CA
	T38	GATGAGTCCTGACCGA A CT
	T2N	5'-GATGAGTCCTGACCGA C NN
	T48	GATGAGTCCTGACCGA C AC
	T49	GATGAGTCCTGACCGA C AG
	T50	GATGAGTCCTGACCGA C AT
	T51	GATGAGTCCTGACCGA C CA

N: A or T or C or G

5) *Selective-amplification*

Amplify the pre-amplification products with three bases overhang *EcoRI* primers and *TaqI* primers (Table 2). *EcoRI* primers are 5' end fluorescent labeled (Table 5). We have 8 *EcoRI* primers and 8 *TaqI* primers, which makes a totally 64 primer combinations. The following PCR reaction mix is used: 3.0 μl of diluted pre-amplify products, 1 fold of Platinum *Taq* Buffer (20 mM Tris-HCl, PH8.4, 50 mM KCl; Invitrogen), 3.0 mM MgCl$_2$, 0.3 mM each of the four dNTPs, 1 U Platinum Taq polymerase (Invitrogen) and 5ng of selective-amplification *EcoRI* primer and 25ng selective-amplification *TaqI* primer in a

total volume of 20 μl. The PCR program is a touchdown program as the following: 4 min at 94°C, 30 seconds at 65°C, 2 min at 68°C, 8 touch-down cycles of 10 seconds at 94°C, 30 seconds at 64°C, 2 min at 68°C. From the first cycle, the annealing temperature is reduced by 1°C down from 64°C to 56°C. Then followed by 26 cycles of 10 seconds at 94°C, 30s at 56°C and 1 min at 68°C, followed by 30 min at 60°C.

6) *Detection and scoring of AFLP markers*

The selective-amplification products produced by 4 different colors fluorescent-labeled *EcoRI* primers (Table 3) in selective-amplification are prepared as following mix: 1.0 μl of each fluorescent color labeled PCR products (totally 4.0 μl), 12.0 μl of formamide, 0.5 μl of Gene Scan^{TM} 500 LIZ^{TM} size standard (Applied Biosystems). Prepared sample are then run in ABI 3730 capillary system auto-sequencer. Data are collected by software *GeneMapper3.7*.

In here, AFLP markers are considered co-dominant. When the band shows up indicates two types: homozygous peak present, heterozygous peak present. In this case, the homozygous peak height should be as twice intensity as those of heterozygous peak.

Table 3 Fluorescents Labeled to AFLP Selective *EcoRI* Primers (Applied Biosystems)

Primer	Fluorescents	Fluorescent color
E32	6-FAM	Blue
E33	VIC	Green
E35	NED	Yellow
E38	PET	Red
E39	6-FAM	Blue
E42	VIC	Green
E44	NED	Yellow
E45	PET	Red

7) *AFLP Fragments Isolation and Sequencing*

Conduct a chi-square test for allele frequencies of the high group and low group. Those AFLP markers which have significant differences frequencies between highest marbling score samples and lowest marbling score samples will be selected for further analysis.

These markers are run in a 5% polyacrylamide gel. The interested bands containing AFLP fragments are excised using a scalpel. After excision, gel fragments are placed in 15 µl of 1X TE and frozen at -80°C for ~30 min, followed by one thawing-refreezing step at -20°C. After thawing, samples are centrifuged for 15 min at 15 000 g and 4.0 µl is taken for PCR re-amplification using pre-amplification AFLP primers. Fragments will be sequenced directly using the same pre-amplification primers on ABI 3730 automatic capillary sequencer.

Ausubel, FM, Brent, R, Kingston, RE, Moore, DD, Seidman, JG, Smith, JA, and Struhl K. 1989. Current Protocols in Molecular Biology. Vol. 1. Greene Publishing Associates and Wiley Interscience, New York.