Association Between Carcass Maturity Grade and Genes Involved in Bone Growth, in Young Heifers

A Thesis
Presented in Partial Fulfillment of the Requirements for the
Degree of Master of Science
with a
Major in Animal Science
in the
College of Graduate Studies
University of Idaho
by

Antonetta M. Colacchio

Major Professor: Gordon K. Murdoch, Ph.D.
Committee Members: Brenda Murdoch, Ph.D.; Phillip Bass, Ph.D.; Michael Colle, Ph.D.
Department Administrator: Amin Ahmadzadeh, Ph.D.

August 2019

Authorization to Submit Thesis

This thesis of Antonetta Colacchio, submitted for the Master of Science with a Major in Animal Science and titled "Association Between Carcass Maturity Grade and Genes Involved in Bone Growth, in Young Heifers," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: _____ Date: _____
Gordon K. Murdoch, Ph.D.

Committee members: _____ Date: _____
Brenda M. Murdoch, Ph.D.

_____ Date: _____
Phillip Bass, Ph.D.

_____ Date: _____
Michael J. Colle, Ph.D.

Department Administrator: _____ Date: _____
Amin Ahmadzadeh

Abstract

Over the last several decades, there has been an increase in the number of heifers, under 30 months of age chronologically, categorized as B and C-maturity based upon skeletal maturity at harvest. The hypothesis for this project is that heifers with advanced physiological maturity have differences within genes that encode proteins involved in bone ossification. Specifically, the genetic variance within genes that promote advanced rates of ossification relating to bone growth and extracellular matrix mineralization have resulted in young heifers reflecting maturity grades more representative of cattle over 30 months of age. This can result in substantial economic losses for producers since the value of an A-maturity carcass is as much as $350/head greater than when an animal is graded below a B-maturity. Despite the USDA change to allow use of dentition to assess maturity, there are still losses with deduction costs associated with advanced maturity in chronologically young beef cattle. In this study, we assessed eight candidate genes known to play an important role in osteogenesis, for sequence variation linkage to advanced maturity. We used targeted gene capture followed by sequencing for candidate genes, in commercial cattle graded as A, B, and C maturity groups (N=90 heifers) but confirmed to be under 30 months using records and dentition. We identified novel and significant genetic sequence variations in association with maturity score within the alkaline phosphatase (ALPL), estrogen receptor 1 (ESR1), estrogen receptor 2 (ESR2), collagen type X alpha 1 (COL10A1), and parathyroid hormone like protein (PTHLH). This represents the first report of such associations with this phenotype and could lead to opportunities to genetically screen beef cattle for susceptibility for early skeletal maturity.

Acknowledgements

First and foremost, I would like to acknowledge my major professor, Dr. Gordon K. Murdoch for giving me this opportunity to grow and learn as a researcher. I am truly grateful for his invaluable help, support, guidance and overall patience with me throughout my graduate program. Without his knowledge and expertise, I would not have made it to where I am today.

I would like to also thank Dr. Brenda M. Murdoch for all her expertise and guidance in the genetic association studies. Without her support and guidance this program would have been much more difficult.

I would like to thank my committee members, Dr. Phillip Bass and Dr. Michael Colle for their support, guidance, encouragements, and patience during my graduate program. I thank the entire AVS department for their unwavering support. I also want to give a special thanks to Ron Richard for his support and wisdom throughout the time I have known him.

We gratefully acknowledge the financial support of the USDA Hatch NC1184, the Idaho Beef Council and the UI college of Agricultural and Life Sciences. Additionally, we are very grateful to Washington Beef, Dr. Matt Doumit and Kacie Salove for the procurement and sharing of products used during this research. Also, we appreciate the technical support and efforts of the UI Genomics core.

Thank you to my fellow graduate students for their support, assistance and friendship during my time at the University of Idaho. I would also like to thank the members of the Murdoch lab that have provided support and assistance throughout my research program.

Further, I would not have made it through my graduate program without the undying support from my friends; Alex Edwards, Marco Mendoza, Kim Davenport, and many more, as well as my family; especially my parents and little sister.

Dedication

I would like to dedicate this thesis to several people who have provided me with the support and drive to reach my goals.

First, I would like to dedicate this body of work to my family who have always provided the positivity throughout my academic career. They always provided words of encouragement and wisdom that gave me the strength to move forward through difficult times.

I would also like to dedicate this thesis to my best friend Alex Edwards. Without her support and sass, I would not have been able to push the boundaries necessary for me to reach my goals.

It is also with my deepest gratitude that I dedicate this thesis to my professor Dr. Gordon Murdoch who has been the most inspirational mentor that I could have ever hoped for.

Last, I would like to dedicate this body of work to the late Ron Richard. Though our time was brief, the impact you've made in my life will last forever.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AD	Allelic depth
ALPL	Alkaline phosphatase
Alt	Alternate
BLINK	Bayesian information and linkage disequilibrium iteratively nested key
BMP2	Bone morphogenetic protein 2
CNP	2',3'-Cyclic-nucleotide 3'-phosphodiesterase
COL10A1	Collagen type X alpha 1
CYP19A1	cytochrome P450 family 19 subfamily A member 1
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EREs	Estrogen-responsive elements
ESR1	Estrogen receptor 1
ESR2	Estrogen receptor 2
GATK	Genome Analysis Tool Kit
GH1	Growth hormone
GPER	G protein-coupled estrogen receptor 1
GWAS	Genome wide association study
HCL	Hydrogen chloride
HWE	Hardy Weinberg equilibrium
IGF-I	Insulin growth factor 1
IHH	Indian hedgehog

LD	Linkage disequilibrium
MAF	Major allele frequency
mRNA	Messenger ribonucleic acid
MYCBP2	Myc-binding protein 2
MYH2	Sarcomeric myosin heavy chain 2a
NaCl	Sodium chloride
NCBI	National center for biotechnology information
NLN	Neurolysin
NR3C1	Nuclear receptor subfamily 3 group C member 1
PCA	Principle component analysis
PTH	Parathyroid hormone
PTHLH	Parathyroid hormone like protein
PTHrP	Parathyroid hormone-related peptide
Ref	Reference
RNA	ribonucleic acid
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
SVS	SNP & Variation Suite
TBA	Trenbolone acetate
TNAP	Tissue non-specific alkaline phosphatase
TNFSF11	Tumor necrosis factor (ligand) superfamily, member 11

QUAL	Quality
VCF	Variant calling file
VDR	Vitamin D receptor
VEGF	Vascular endothelial growth factor

CHAPTER 1

Review of Literature

The United States Standards for Grades of Carcass Beef

The tentative standards for dressed beef grades were formulated in 1916 (USDA, 2017). These standards provided a uniform report to beef markets and were accepted nationally in 1917. Throughout the following century amendments and changes were added when necessary to the grade specifications. In August 1924 the "Market Classes and Grades of Dressed Beef" was published (USDA, 2017). Finally, in 1927 these standards were used for voluntary beef grading.

The official standards were amended in July 1939 to provide a single standard for the grading and labeling of steer, heifer, and cow beef according to similar inherent quality characteristics (USDA, 2017). In this amendment certain grade terms of "Medium," "Common," and "Low Cutter" were changed to "Commercial," "Utility," and Canner." In April 1989, the official standards were revised to allow the official grade to consist of the quality grade only, yield grade only, or a combination of both. No changes were made in the actual yield grade or quality grade requirements. The change was made to allow the industry greater flexibility in the use of the beef grading system in order to provide consumers with the trimness levels desired (USDA, 2017). Currently, the 8 quality grades available for beef carcasses, in descending order are: Prime, Choice, Select, Standard, Commercial, Utility, Cutter and Canner. To facilitate the application of these principles, the standards recognize five different maturity groups. The five maturity groups are identified in Figure 1.1 as A, B, C, D, and E in order of increasing maturity

(USDA, 2017). The quality grades of carcasses are determined based on the relationship between maturity and marbling (Figure 1.1). Marbling scores are measured by the amount of intramuscular fat located at the 12th and 13th rib interface of the ribeye. There are nine levels of marbling utilized in the beef standards. These levels are abundant, moderately abundant, slightly abundant, moderate, modest, small, slight, traces, and practically devoid. When beef carcasses are presented for quality grading, USDA graders examine visible indicators of physiological maturity (i.e., size and shape of the ribs and ossification of the bones and cartilages along the vertebral column) which are used to classify carcasses into maturity groups before the dentition amendment in 2017 (Tatum, 2011). Carcasses classified as A-maturity typically would qualify for the Prime, Choice, Select, or Standard grades; those classified as B-maturity usually would grade Prime, Choice, or Standard; and those classified as C-, D-, or E-maturity, if graded, would be identified as Commercial, Utility, Cutter, or Canner (USDA, 2017). Animals that are 9-30 months are potentially classified as A maturity. Animals 30-42 months are classified as B maturity. Animals that are 42-72 months are C maturity and animals that are 72-96 months are classified as D maturity. Any animals older than 96 months are classified as E maturity.

In April 2016 a petition was requested for an amendment to the beef standards. With the increasing rates of cattle being under thirty months of age chronologically but being graded as older in the B and C maturity range it has created an economic loss for producers (Agricultural Marketing Service 2017). Lawrence and colleagues (2011), conducted two studies to determine the relationship between the bone ossification maturity system and the number of permanent incisors present at slaughter. Their studies

showed that animals with 2 adult incisors or less were under 30 months of age and that by determining physiological maturity by dentition rather than the current method would be more accurate in sorting carcasses in less variable age groups (Lawrence et al., 2011). A study done in 2014 by Acheson and associates compared sensory properties of beef from A maturity and B maturity carcasses produced by grain-finished steers and heifers classified as less than 30 months old at the time of slaughter using dentition. The results in that study showed no significant differences between A maturity and B-C maturity carcasses and these no longer support the current grading concept of skeletal maturity. This research played a supportive role in the USDA grading dentition amendment. This amendment called for the use of dentition or documentation of chronological age to be used when determining maturity grouping for quality grading in heifers and fed steers. In further support of this legislation, Salove and Doumit (in preparation) determined that there were no sensory attributes of beef associated with the B and C maturity carcasses that were actually under 30 months as confirmed by dentition. In July 2017, the USDA informed the public of the proposed changes that would occur in the revision of the beef standards. This amendment stated that dentition and chronological age documentation as means of determining maturity groupings, and thus eligibility for Prime, Choice, and Select grade designations (Agricultural Marketing Service, 2017) Dental age is more closely related to actual chronological age than is USDA physiological maturity (Raines et al., 2008). On December 18, 2017, the new standards for age documentation and dentition allowed additional methods of classifying maturity of carcasses presented to the USDA for official quality grading (Agricultural Marketing Service, 2017). Cattle with 2 or fewer adult incisors are considered less than 30 months of age, whereas cattle with 3 or

more adult incisors are considered older than 30 months of age (National Cattlemen's Beef Association, 2015).

With these changes, determining the quality of lean in steers, heifers and cow beef, the maturity of the carcasses is determined by one or all of 3 accepted methods: 1) dentition, 2) age documentation, and/or 3) degree of bone ossification and color and texture of the lean flash. One challenge is that within the US system, slaughter cattle often are transferred from producer to packer without documentation of actual animal ages (USDA, 2005). Correspondingly, carcass indicators of physiological maturity are considered when assigning USDA quality grades to beef carcasses in an effort to reflect age-related differences in beef tenderness (Tatum, 2011).

Physiological Maturity and Chronological Age

Physiological maturity refers to the physiological age of the animal rather than its chronological age. Indicators of physiological maturity include bone ossification, cartilage ossification, and the color/texture of muscle (Tatum, 2011). Cartilage and bone ossification receive more emphasis when grading because color and texture can be impacted by pre- and postmortem factors (Hale et al., 2013). When cartilage is being evaluated for physiological maturity the cartilage in the vertebrae of the backbone is assessed. The cartilage between and on the dorsal edges of the individual sacral and lumbar vertebrae as well as the cartilage located on the dorsal surface of the spinous processes of the thoracic vertebrae (USDA, 2017) The vertebral buttons are the most prominent, softest and least ossified in the younger animals. As maturity proceeds from B to E, more ossification becomes evident as the replacement of the soft cartilaginous

matrix is replaced with bone matrix (USDA, 2017). When carcass maturity score was first implemented there was a consistent relationship with cartilage ossification and the age of the animal since the ossification of the buttons began at 30 months of age (USDA, 2017). This relationship became less consistent with modern day cattle as the ossification and overall animal maturity proceeded earlier than 30 months resulting in young animals being classified as older maturity grades (Tatum, 2011).

Through the evaluation of size, shape, and ossification of the bones, carcasses can be determined under or over thirty months of age during grading. The final quality grade would be given with the adjustment factors from advanced skeletal maturity characteristics (Agricultural Marketing Service, 2017). Ossification begins in the sacral region of the vertebrae and as the animal ages, ossification progresses to the lumbar then thoracic region of the vertebrae (Hale et al., 2013). As ossification progresses the sacral region of the vertebrae will fuse. In beef cattle that are very advanced in maturity, there will be no red color, and cartilage on the end of the vertebrae is entirely ossified (USDA, 2017).

It is well documented that estrogen plays an important role in skeletal ossification (Tatum, 2011; Grumbach and Auchus, 1999; Weise et al., 2001). It is known that in female mammals, estrogen is increased at puberty. Hence adolescent females of various mammalian species exhibit more advanced skeletal maturity when compared with males of the same age (Grumbach and Auchus, 1999). The estrogenic effects are mediated by estrogen receptor expression directly within the tissues that comprise the growth plates, since the female's endogenous production of estrogens is greater than it is in males there

is a notable difference in the timing of osteogenesis between the sexes. Specifically, skeletal maturity in heifers progresses at a much faster rate than steers, as a result, heifers that are 16- to 27-months old were about seven times more likely to produce carcasses with B-maturity skeletal characteristics and almost eleven times more likely to produce carcasses with C-maturity or older skeletal characteristics than 16- to 27-month-old steers (Tatum, 2011). Three direct physiological effects of estrogen are that it decreases the rate of chondroblast proliferation, rate of chondrocyte maturation and hypertrophy while accelerating the rate of osteoblast maturation (Weise et al., 2001). Given that ossification involves the palisading replacement of a cartilaginous template to bony matrix via the action of osteoblasts, the shift in the balance away from cartilage depositing chondrocytes toward ossifying osteoblasts ultimately results in the scenario where no more cartilage exists to ossify and the bony plate fuses (Weise et al., 2001).

Bone Ossification

Bones in different parts of the skeleton develop through two distinct processes, intramembranous ossification and endochondral ossification. Intramembranous ossification, which occurs in the flat bones of the skull, involves direct differentiation of embryonic mesenchymal cells into osteoblasts to form the bone (Mackie et al., 2011). Endochondral ossification is a type of bone formation where pre-existing hyaline cartilage is ossified to bone (Mackie et al., 2011). Endochondral ossification is a complex process, involving a carefully regulated changes in chondrocyte behavior that is coordinated with the actions of blood vessels, osteoclasts and the other cells that are brought with them (Mackie et al., 2008). This occurs at the epiphyseal plate in long bones and during the process bone elongates. The bone shaft elongates between the epiphysis

and diaphysis (Hunziker et al., 1988). The important determinant of the growth rate in the epiphyseal plate is the total number of proliferating chondrocytes (Hunziker et al., 1988), which depends on the size of the active pool of chondroblast cells and the rate at which these cells are proliferating (Karp et al., 2000).

In the limbs, where extensive growth is required for proximal–distal extension of the long bones, endochondral development begins with the aggregation of undifferentiated mesenchymal cells to form condensations, which by their position, shape, and size already predetermine and provide structural scaffolding of the future skeletal elements (Hinchcliffe and Johnson 1980). The initial stage of endochondral ossification is the formation of the cartilage template during fetal development and postnatal growth (Hinchcliffe and Johnson 1980). When it is time for the cartilage to be replaced by bone, several ordered stages occur. The first stage of cartilage replacement and bone formation is the resting zone. Within the resting zone the cartilage cells (chondrocytes) are inactive and not undergoing any changes. The proliferation of chondrocytes and deposition of the new matrix are both responsible for initial growth of skeletal elements (Hunziker, 1988). The region where this occurs is accordingly referred to as the proliferation zone. The chondroblast cells become active and begin rapidly dividing, creating a zone of growth. Under the control of signaling by Indian hedgehog (IHH), bone morphogenetic proteins (BMPs) and fibroblast growth factor 18, a region of resting chondrocytes feeds into a zone of proliferating chondrocytes that then undergo hypertrophy and subsequently apoptosis (Ortega et al., 2004). During the hypertrophic zone, starting at the pre-hypertrophic chondrocytes within the middle of the zone begin to undergo maturation, enlarging and signaling the secretion of the

extracellular matrix that then will enter the calcification zone, along with interacting with other genes thus creating a signal for the surrounding matrix to calcify (Mackie et al., 2011). Changes in the composition and properties of the cartilage matrix in the hypertrophic zone, allow invasion by capillaries, accompanied by apoptosis of terminally differentiated chondrocytes, degradation of the calcified cartilage matrix, and its replacement by the trabecular bone matrix secreted by invading osteoblasts (St-Jacques et al., 1999). Bone continues to form a palisading replacement of the cartilage until the epiphyseal plate fuses.

The next stage is the calcification zone where the matrix becomes mineralized. The matrix surrounding late hypertrophic chondrocytes is mineralized through the deposition of hydroxyapatite (Mackie et al., 2007). Hydroxyapatite crystals (composed primarily of calcium and phosphate) are deposited in the extracellular matrix (Mackie et al., 2007). The matrix vesicles released by these cells contain a combination of proteins including phosphate transporters, phosphatases and annexins and provide the nucleation site for mineralization (Anderson, 1969; Kirsch et al. 1997; Kirsch, 2006). The phosphatases PHOSPHO1 and tissue non-specific alkaline phosphatase, both contribute to mineralization of cartilage ECM (Mackie et al., 2011).

It has been noted that blood vessels appear to precede osteoclasts as the first cells that enter the lacunae recently vacated by dying hypertrophic chondrocytes (Mackie et al., 2011). Osteoblasts then infiltrate the cartilage and replace it with bone tissue, forming the bone trabeculae (Mackie et al., 2011). Changes in the composition and properties of the cartilage matrix in the hypertrophic zone, allow invasion by capillaries, accompanied by apoptosis of terminally differentiated chondrocytes, degradation of the calcified

cartilage matrix, and its replacement by the trabecular bone matrix secreted by invading osteoblasts (St-Jacques et al., 1999). Continued deposition of cortical bone by the periosteum (outer bone margins) leads to radial growth (St-Jacques et al., 1999). This process will continue until the chondrocytes begin to go through atresia and the plate will fuse. Finally, bone remodeling results from the combined activities of the osteoclasts; bone resorbing cells of hematopoietic origin, and the mineral depositing osteoblasts (Mackie et al., 2011). Coordinating each of these processes is critical in determining the size, shape, rate of growth and mechanical properties of the bones (St-Jacques et al., 1999). Collectively, these observations indicate that chondrocytes of the growth plate release soluble factors that allow them to delicately control the behavior of the invading vascular endothelial cells, osteoclasts and osteoblasts, which all have important roles in transforming growth plate cartilage into dense hard bone tissue (Mackie et al., 2011).

When the rate of ossification surpasses the rate of chondrocytes replacement this ultimately leads to the disappearance of the growth plate cartilage also known as epiphyseal fusion to yield a remnant as the epiphyseal line (Mackie et al., 2011). Once the plate becomes the epiphyseal line no more growth can occur via endochondral ossification and the bone has reached its terminal length (Mackie et al., 2011). Bone growth is a polygenic trait meaning many different genes impact it and we selected specific genes that influence this process as our target genes. These genes and their specific roles are reviewed in latter sections of this chapter.

Growth Implants

Anabolic implants, which are used routinely in conventional U.S. beef production systems to improve growth rate and feed conversion in growing/finishing steers and heifers, have been shown to accelerate skeletal maturation in cattle (Tatum, 2011). The primary active ingredient in commercially available implants is either, estrogen, an androgen or a combination of both. Implants that contain zeranol or estradiol (either alone or in combination with Trenbolone acetate (TBA)) all have been found to increase skeletal maturity (Turner et al., 1981; Vanderwert et al., 1985; Apple et al., 1991; Foutz et al., 1997; Reiling and Johnson, 2003). The rate of skeletal maturation is dependent on the dose and frequency of an implant given before harvest. Tatum (2011) and Reiling and colleagues (2003) showed that among cattle harvested at 21 to 27 months of age, implanting greatly increases the risk of an animal producing a B-maturity or more advanced maturity carcass, especially in heifers.

Maturity in Livestock

Genetic selection for early puberty and growth traits may have also led to the increase in bone ossification leading to animals maturating at a faster rate and increasing the incidence of advanced maturity. This has become a rising occurrence in the last twenty years. More carcasses are being evaluated at an advanced maturity compared to the animals chronological age (Tatum, 2011). This led to economic losses to producers and the petition to add dentition as a form of identification for age in steers and heifers (USDA 2017). However, despite this amendment to the USDA grading system, animals are still undergoing accelerated ossification and creating a loss to producers when age

cannot be verified. With the increase of 15% in processed heifers being graded at a lower maturity based on skeletal ossification (Moore et al., 2012), the identification of the causes of advanced maturity could lead to opportunities in the industry to adapt to this trend.

Candidate Genes

While many of these candidate genes have multiple functions, the purpose of the subsequent brief review is to outline their specific roles in the growth, formation and maturation of skeletal tissue especially as it pertains to mineralized bone. Figure 1.2 gives a brief overview and summary of each of the candidate genes selected and how they are involved in endochondral ossification.

Growth Hormone

Growth hormone (*GH1*) is located on chromosome 19 in the *Bos taurus* genome. This protein-coding gene is at 48,768,617 bp - 48,772,049 bp with the cognate protein also referred to as growth hormone (NCBI, 2019). It is released from the pituitary gland and is controlled by the hypothalamus. This gene has been known for many years to be an important regulator of longitudinal bone growth (Nilsson et al., 2005). Growth hormone enhances the anabolic responsiveness of the growth plate and stimulates the expansion of the chondrocyte progenitor cells (Nilsson et al., 2005). It also stimulates alkaline phosphatase and collagen X expression, both of which play important roles in osteogenesis. Briefly, when *GH* is released into the systemic circulation, it triggers bone deposition and can accelerate its rate of accretion (Olney et al., 2003). Another function of *GH* is that it stimulates the chondrocyte progenitor cells determining the mature size

and it also stimulates the beginning of the cartilage templates (Olney et al., 2003). Ultimately, *GH* triggers and accelerates bone deposition while stimulating cell proliferation at growth plates and will therefore increase the terminal or penultimate size and length of long bones (Nilsson et al., 2005). Long bone growth primarily occurs at the epiphyseal growth plates and is the result of the proliferation and differentiation of chondrocytes. *GH1* has direct effects on these chondrocytes, but primarily regulates this function through a specific autocrine and paracrine action of IGF-I, which stimulates the proliferation of and matrix production by the chondrocytes (Olney et al., 2003). With greater deposition of cartilaginous template, greater extension of the bone tissue can occur. Notably and unsurprisingly, the absence of *GH1* results in a reduced rate of bone remodeling and a gradual loss of bone mineral density also GH1 deficiency severely limits bone growth and hence the accumulation of bone mass (Olney et al., 2003), and is one direct cause of dwarfism. Since *GH* is also an important mediator of skeletal muscle growth it is likely that as cattle have been selected for enhanced lean growth that variants in the GH gene have been at least one component of this production driven selection.

Indian Hedgehog

Indian hedgehog (*IHH*) is located 107,722,665bp-107,728,939bp on chromosome 2 in the *Bos taurus* genome (NCBI, 2019). Indian hedgehog signaling from pre-hypertrophic chondrocytes has been implicated in the control of chondrocyte maturation through the feedback control of a second secreted factor parathyroid hormone-related peptide (*PTHrP*) creating a feedback loop influencing the zone of growth (St-Jacques et al., 1999). During maturation, expression of Indian hedgehog becomes

progressively restricted to post mitotic pre-hypertrophic chondrocytes adjacent to the *PTH/PTHrP-R* (another candidate) expressing proliferative zones (Bitgood et al.,1995). The Indian Hedgehog signaling pathway plays a pivotal role in coordinating several different cellular processes, which are essential for morphogenesis of the vertebrate skeleton (Bitgood et al.,1995). Indian hedgehog plays an important role in osteoblast development (Minina et al., 2002). Indian hedgehog is produced by pre-hypertrophic chondrocytes, promotes proliferation of the adjacent chondrocytes and additionally induces the expression of several bone morphogenic genes (*BMP*s) in the perichondrium and in the proliferating chondrocytes (Minina et al., 2002). Therefore, *IHH* can increase bone growth and reduce epiphyseal plate closure by stimulating rapid and continuous cartilage deposition, ultimately resulting in longer terminal length and later or delayed maturation (Minina et al., 2001). Mineralization of matrix vesicles and cartilage matrix is dependent on alkaline phosphatase activity, as demonstrated by results obtained from tissue non-specific alkaline phosphatase (TNAP)-deficient mice (Anderson et al., 2004, Fedde et al., 1999). When pre-hypertrophic chondrocytes fully differentiate into hypertrophic chondrocytes, they no longer express *IHH*. Consequently, the negative feedback mechanism is attenuated, and new chondrocytes are allowed to initiate the differentiation process (Vortkamp et al., 1996). A mutation in this gene can lead to bone malformation and stunting of bone growth. In a study by St-Jacques and colleagues (1999) analysis of an *Ihh* null mutant in mice suggests a more extensive role for *Ihh* in skeletal development. Mutants' display markedly reduced chondrocyte proliferation, maturation of chondrocytes at inappropriate position, and a failure of osteoblast development in endochondral bones (Vortkamp et al., 1996). Together, the results in this

study suggest a model in which *IHH* coordinates diverse aspects of skeletal morphogenesis through *PTHrP*-dependent and independent processes (St-Jacques et al. 1999). This makes *IHH* an important candidate gene to evaluate in the context of our bone maturation study.

Estrogen Receptor 1

Estrogen receptor 1 (*ESR1*) is found on autosomal chromosome 9 in the *Bos taurus* genome. It is located between 89,969,586bp-90,255,801bp and is a protein coding gene (NCBI, 2019). The cognate protein for this gene is accordingly named estrogen receptor 1, and it is a cytosolic receptor. In heifers, it is well documented (Grumbach, 1999), (Tatum, 2011) and (Weise et al., 2001) that estrogen increases the rate of skeletal ossification. Estrogen decreases the rate of chondroblast duplication, accelerates chondrocytes maturation and hypertrophy while further accelerating the rate of osteoblasts activity until there is no more cartilage to ossify and the plate fuses (Weise et al., 2001). Previous reports suggest that estrogen accelerates growth plate ossification by stimulating vascular and bone cell invasion of the growth plate cartilage, causing ossification to advance beyond the hypertrophic zone into the proliferative and resting zones (Weise et al., 2001). Exogenous estrogenic implants also impact the total circulating estrogen levels in implanted livestock (Tatum, 2011). Estrogen plays an important role in the rate of closure of the epiphyseal plate, where in general higher estrogen levels are associated with earlier closure and reduced mature size (Väänänen and Härkönen., 1996). Since estrogen decreases the duplication and replenishment of the chondrocytes, while accelerating the rate of osteoblast mediated mineralization estrogen

increases the rate of ossification and fusion of the plate (Väänänen et al., 1996). Changes to the structure of the estrogen receptors could change an animal's affinity and responsiveness to both endogenous and exogenous estrogen. This receptor is an important candidate gene target for this study and may help explain how selection for early reproductive maturity may have inadvertently accentuated the early maturation of bone.

Estrogen Receptor 2

Estrogen receptor 2 (*ESR2*) is a protein-coding gene located between 76,706,017bp-76,771,172bp on chromosome 10 of the *Bos taurus* genome (NCBI, 2019). It is a member of the nuclear superfamily and shares similar structural characteristics of *ESR1*. Upon ligand activation, *ER*s can regulate biological processes by divergent pathways (Zhao et al., 2010). The so-called classical signaling occurs through direct binding of ER dimers to estrogen-responsive elements (EREs) in the regulatory regions of estrogen-responsive genes, followed by recruitment of co-regulators to the transcription start site (Zhao et al., 2010). Since the cells within the epiphyseal growth plate express both estrogen receptor 1 and 2, both genes were selected as candidates for screening in this study. Similar, to the relationship of estrogen receptor 1 with reproductive fitness selection in livestock, estrogen receptor 2 also has the potential for inadvertent selection that impacts osteogenesis rates (Zhao et al., 2010).

Alkaline Phosphatase

Alkaline phosphatase (*ALPL*) is located on autosomal chromosome 2 in the *Bos taurus* genome, the same chromosome as Indian hedgehog gene. *ALPL* gene is located at

131,791,668bp-131,857,687bp (NCBI, 2019). This is a protein-coding gene that encodes the functional cell surface enzyme; tissue non-specific alkaline phosphatase. *ALPL* is an osteoblast differentiation marker and provides the inorganic phosphate essential for mineralization of the extracellular matrix. *ALPL* interacts with collagen type X alpha 1 and establishes the conditions deposition of hydroxyapatite crystals representing the mineralization stage of ossification. *ALPL* is expressed on the cell membrane of hypertrophic chondrocytes, osteoblasts, and is also concentrated on the membranes of the matrix vesicles budding from these cells (Orimo, 2010). Alkaline phosphatase contributes to the inorganic phosphate in the matrix, without which the osteogenic process would not transform the cartilaginous matrix to the trabecular matrix of bone (Millan, 2013). *ALPL* controls the levels of calcium and phosphate deposition of the matrix. Mutations in the alkaline phosphatase gene cause hypophosphatasia, a heritable form of rickets and osteocalcin, caused by an arrest in the hydroxyapatite crystals onto the collagenous extracellular matrix due to accumulation of extracellular inorganic pyrophosphate (Millan, 2013).

Bone Morphogenetic Protein 2

Bone morphogenetic protein 2 (*BMP2*) is located on autosomal chromosome 13 in the *Bos taurus* genome. This is a protein-coding gene and is mapped between 49,550,050bp-49,561,304bp (NCBI, 2019). *BMP2* stimulates the production of bone by inducing the formation of both cartilage and bone (Yoon et al., 2006). It influences the rate of production of both hypertrophic chondrocytes and mature osteoblasts (Minina et al., 2002). It has consistently been observed that *BMP* signaling supports proliferation of

chondrocytes in growth cartilage (Minina et al., 2002; Yoon et al., 2006). The bone morphogenetic protein 2 stimulates bone production by inducing the rate and feedback of the cartilage template formation. It influences the hypertrophic chondrocytes increasing the rapid cartilage deposition (Yu et al., 2010). The hypertrophic chondrocytes no longer produce Indian hedgehog so there is a direct interaction between *BMPs* and *IHH*. *BMP* is produced directly by chondrocytes and influences the expression of Indian hedgehog creating a negative feedback loop (Yu et al., 2010). BMP and IHH signals act in parallel to regulate the distinct zones of growth in the epiphyseal plate and the rate and extent of chondrocyte proliferation (Minina et al., 2002). In addition, *BMP* signaling modulates the expression level of *IHH*, thereby integrating the regulation of chondrocyte proliferation and the onset of hypertrophic differentiation (Minina et al., 2002). Active *BMP* receptors leads to enhanced expression of collagen type X, suggesting that BMP signaling induce hypertrophy (Grimsrud et al., 1999, Grimsrud et al., 2001). By acting at several stages of chondrocyte differentiation *BMP* signaling seems to physiologically regulate the size of a skeletal element without disturbing the differentiation process relative to developmental age (Minina et al., 2002). A slight increase of *BMP* signaling would thus result in slightly larger bones, which would however develop at an unaltered pace. Similarly, a slight decrease in BMP signaling would produce the reverse effect leading to shorter bones, also without a change in the rate of development (Minina et al., 2002).

Parathyroid hormone like protein

Parathyroid hormone like protein (*PTHLH*) is located on autosomal chromosome 5 in the *Bos taurus* genome. This gene is most recently mapped to the region from 82,241,708bp-82,258,858bp on the forward strand and is a protein-coding gene (NCBI, 2019). It plays an important role in the regulation of calcium phosphate. It is secreted by proliferating chondrocytes (Lanske et al. 1996). Parathyroid hormone like protein enhances the activity of osteoblasts in favor of rapid mineralization. The parathyroid like protein balances the activity of osteoblasts while also inducing the ossification relation to bone deposition (Karaplis et al. 1994). Several lines of evidence indicate that the signaling factor parathyroid hormone-related peptide (*PTHrP*) regulates the process of chondrocyte maturation (Karaplis et al. 1994). Targeted inactivation of the genes encoding either *PTHrP* or its receptor (parathyroid hormone/parathyroid hormone-related peptide receptor, *PTH*/*PTHrP-R*) in the mouse, result in a lethal phenotype, a feature of which is skeletal dysplasia characterized by premature maturation of chondrocytes leading to excessive bone formation at birth (Karaplis et al. 1994; Lanske et al. 1996). Conversely, overexpression of *PTHrP* in chondrocytes leads to a delay in chondrocyte maturation and bone formation such that mice are born with a completely cartilaginous endochondral skeleton (Weir et al. 1996). *PTH/PTHrP-R* is expressed at low levels throughout the region of immature chondrocytes and at highest levels in mitotically active chondrocytes in the proliferative zone from where post mitotic hypertrophic precursors emerge (Amizuka et al. 1996; Lee, 1996). This gene works in balance with other candidate genes to determine or establish the relative rates of cartilage template formation and ossification (Lee, 1996).

Collagen Type X, alpha 1

Collagen Type X, alpha 1 (*COL10A1*) is a protein-coding gene on the forward strand of chromosome 9. It is located between 34,961,718bp-34,994,278bp (NCBI, 2019). *COL10A1* is a short chain collagen that is expressed by the hypertrophic chondrocytes during the hypertrophic stage of endochondral ossification as a primary signal for the start of matrix ossification Leitinger and Kwan., 2006). It is a major component in the matrix and provides direct scaffolding for the hydroxyapatite crystals. The gene, collagen type 10, alpha 1 is important in the development and formation of the tropocollagen released by the chondrocytes into the extracellular matrix used in bone formation (Luckman et al., 2003). Collagen type X binds to chondrocytes through integrin $\alpha_2\beta_1$, but also through an interaction with *DDR2* that leads to auto phosphorylation of the receptor, suggesting that this protein may indeed exert effects on intracellular signaling pathways in chondrocytes that ultimately influences their physiological activity (Leitinger et al., 2006; Luckman et al., 2003).

Sarcomeric myosin heavy chain 2a

Sarcomeric myosin heavy chain 2a (*MYH2*) is a protein coding gene and is located on chromosome 19 between 30,137,767bp-30,165,109bp (NCBI, 2019). This gene is important for the development and growth of skeletal muscle. Changes in this gene could lead to changes of the properties in muscle. It belongs to a large family of motor proteins that share the common feature of using ATP hydrolysis, energy transduction and actin binding (Novotny et al., 2015). Muscle and bones have a strong

association and create the musculoskeletal system that provides the framework for structure and movement. Muscle and bone are proportionally matched in their functional capacity and can change significantly when one is changed (Novotny et al., 2015). Osteocytes in bone can sense the change in muscle pressure and thus stimulate bone formation (Novotny et al., 2015). The communication between these two physiological systems allow for the anabolic/catabolic modifications to bone and muscle to maintain the constant relationship between function and structure of both tissues (Novotny et al., 2015).

Vascular Endothelial Growth Factor

Vascular Endothelial Growth Factor (*VEGF*) is a member of the growth factor family and encodes for a heparin-binding protein (NCBI 2019). This factor leads to the migration and proliferation of vascular endothelial cells and essential for angiogenesis (Tang et al. 2004). Any mutation to this gene can lead to abnormal embryonic blood vessel formation. When upregulated this gene is known to progress tumor stage and regulation (Tang et al. 2004). *VEGF* is located on chromosome 23, from position 17,255,515bp-17,270,515bp (NCBI, 2019). In research done by Tang et al. 2004, they showed the importance of *VEGF* being an essential survivor factor for muscle capillaries and when inefficient leads to apoptosis in skeletal muscle.

To address this issue of advanced maturity still occurring led to the project proposal of solving the maturity score and carcass quality problem through targeted gene analyses. The hypothesis for this project is that heifers with advanced physiological

maturity have differences within genes that encode proteins involved in bone ossification. During the process of selection for phenotypes of early maturity heifers has also impacted the selection of corelated genetics that influences the rate of skeletal maturation resulting in an increase in B-, and C-maturity heifers at harvest. Looking at genes that are involved in bone growth in beef cattle could lead to the identification of changes in the genes leading to the advanced ossification rates. In this project several genes were screened through sequence analysis. The candidate genes are Growth hormone (*GH*), Indian Hedgehog (*IHH*), Estrogen Receptor 1(*ESR1*), Estrogen Receptor 2 (*ESR2*), Alkaline Phosphatase (*ALPL*), Bone Morphogenetic Protein 2 (*BMP2*), Parathyroid Hormone like Protein (*PTHLH*), Collagen type X Alpha 1 (*COLXA1*), Sarcomeric myosin heavy chain 2a (*MYH2*), and Vascular Endothelial Growth Factor (*VEGF*).

This research project is geared towards identifying genetic variations that may help producers with this issue by identifying a biological cause that can lead to the potential of mitigating the financial loss associated with advance maturity.

References

Acheson, R. J., Woerner, D. R., & Tatum, J. D. (2014). Effects of USDA carcass maturity on sensory attributes of beef produced by grain-finished steers and heifers classified as less than 30 months old using dentition1. *Journal of Animal Science,92*(4), 1792-1799. doi:10.2527/jas.2013-7553

Agricultural Marketing Service. 2017. Notice: United States standards for grades of carcass beef. Fed. Regist. 82:233.

Amizuka, N., Henderson, J. E., Hoshi, K., Warshawsky, H., Ozawa, H., Goltzman, D., & Karaplis, A. C. (1996). Programmed cell death of chondrocytes and aberrant chondrogenesis in mice homozygous for parathyroid hormone-related peptide gene deletion. *Endocrinology,137*(11), 5055-5067. doi:10.1210/en.137.11.5055

Anderson, H. C. (1969). Vesicles Associated With Calcification In The Matrix Of Epiphyseal Cartilage. *The Journal of Cell Biology,41*(1), 59-72. doi:10.1083/jcb.41.1.59

Anderson, H. C., Sipe, J. B., Hessle, L., Dhamyamraju, R., Atti, E., Camacho, N. P., & Millán, J. L. (2004). Impaired Calcification Around Matrix Vesicles of Growth Plate and Bone in Alkaline Phosphatase-Deficient Mice. *The American Journal of Pathology,164*(3), 841-847. doi:10.1016/s0002-9440(10)63172-0

Apple, J. K., Dikeman, M. E., Simms, D. D., & Kuhl, G. (1991). Effects of synthetic hormone implants, singularly or in combinations, on performance, carcass traits, and longissimus muscle palatability of Holstein steers. *Journal of Animal Science,69*(11), 4437-4448. doi:10.2527/1991.69114437x

Bitgood, M. J., & McMahon, A. P. (1995, November). Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. *Developmental Biology* Nov;172(1):126-38 https://www.ncbi.nlm.nih.gov/pubmed/7589793

Fedde, K. N., Blair, L., Silverstein, J., Coburn, S. P., Ryan, L. M., Weinstein, R. S., . . . Whyte, M. P. (1999, December). Alkaline phosphatase knock-out mice recapitulate the metabolic and skeletal defects of infantile hypophosphatasia. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049802/

Foutz, C. P., Dolezal, H. G., Gardner, T. L., Gill, D. R., Hensley, J. L., & Morgan, J. B. (1997). Anabolic implant effects on steer performance, carcass traits, subprimal yields, and longissimus muscle properties. Retrieved October 20, 2018, from https://www.ncbi.nlm.nih.gov/pubmed/9159272

Grimsrud, C. D., Romano, P. R., Dsouza, M., Puzas, J. E., Reynolds, P. R., Rosier, R. N., & Okeefe, R. J. (1999). BMP-6 Is an Autocrine Stimulator of Chondrocyte Differentiation. *Journal of Bone and Mineral Research,14*(4), 475-482. doi:10.1359/jbmr.1999.14.4.475

Grimsrud, C. D., Romano, P. R., Dsouza, M., Puzas, J. E., Schwarz, E. M., Reynolds, P. R., Okeefe, R. J. (2001). BMP signaling stimulates chondrocyte maturation and the expression of Indian hedgehog. *Journal of Orthopaedic Research,19*(1), 18-25. doi:10.1016/s0736-0266(00)00017-6

Grumbach, M. M. (1999). Estrogen: Consequences and Implications of Human Mutations in Synthesis and Action. *Journal of Clinical Endocrinology & Metabolism,84*(12), 4677-4694. doi:10.1210/jc.84.12.4677

Hale, D. S., Goodman, K., & Savell, J. W. (2013, March 8). USDA Beef Quality and Yield Grades. Retrieved May 18, 2019, from https://meat.tamu.edu/beefgrading/

Hinchliffe, J. R., & Johnson, D. R. (1980). *The development of the vertebrate limb: An approach through experiment, genetics and evolution*. Oxford: Clarendon Press.

Hunziker, E. B. (1988). Growth plate structure and function. *Pathol. Immunopathol. Res.* 7, 9-13.

Karaplis, A. C., Luz, A., Glowacki, J., Bronson, R. T., Tybulewicz, V. L., Kronenberg, H. M., & Mulligan, R. C. (1994). Lethal skeletal dysplasia from targeted disruption of the parathyroid hormone-related peptide gene. *Genes & Development,8*(3), 277-289. doi:10.1101/gad.8.3.277

Karaplis, A. C. (1998). Inactivating Mutation in the Human Parathyroid Hormone Receptor Type 1 Gene in Blomstrand Chondrodysplasia. *Endocrinology,139*(12), 5255-5258. doi:10.1210/en.139.12.5255

Karp, S.J., Schipani E., St-Jacques, B., Hunzelman, J., Kronenberg, H. and McMahon, A.P. (2000). Indian hedgehog coordinates endochondral bone growth and morphogenesis via parathyroid hormone related-protein-dependent and -independent pathways. Development 127, 543–548.

Kirsch, T., Nah, H., Shapiro, I. M., & Pacifici, M. (1997). Regulated Production of Mineralization-competent Matrix Vesicles in Hypertrophic Chondrocytes. *The Journal of Cell Biology,137*(5), 1149-1160. doi:10.1083/jcb.137.5.1149

Kirsch, T. (2006). Determinants of pathological mineralization. *Current Opinion in Rheumatology,18*(2), 174-180. doi:10.1097/01.bor.0000209431.59226.46

Lanske, B., Karaplis, A. C., Lee, K., Luz, A., Vortkamp, A., Pirro, A., Karperien, M., Defize, L. H., Ho, C., Mulligan R. C., Abou-Samra, A. B., Juppner, H., Serge G. V., and Kronenberg, H. M. (1996). PTH/PTHrP Receptor in Early Development and Indian Hedgehog--Regulated Bone Growth. *Science,273*(5275), 663-666. doi:10.1126/science.273.5275.663

Lawrence, T. E., Whatley, J. D., Montgomery, T. H., & Perino, L. J. (august 2001). A comparison of the USDA ossification-based maturity system to a system based on dentition. *Journal of Animal Science,79*(7), 1683-1690. doi:10.2527/2001.7971683x

Lee, K. (1996). Parathyroid hormone-related peptide delays terminal differentiation of chondrocytes during endochondral bone development. *Endocrinology,137*(11), 5109-5118. doi:10.1210/en.137.11.5109

Leitinger, B., & Kwan, A. P. (2006). The discoidin domain receptor DDR2 is a receptor for type X collagen. *Matrix Biology,25*(6), 355-364. doi:10.1016/j.matbio.2006.05.006

Luckman, S. P., Rees, E., & Kwan, A. P. (2003). Partial characterization of cell-type X collagen interactions. *Biochemical Journal,372*(2), 485-493. doi:10.1042/bj20021572

Mackie, E., Ahmed, Y., Tatarczuch, L., Chen, K., & Mirams, M. (2008). Endochondral ossification: How cartilage is converted into bone in the developing skeleton. *The*

International Journal of Biochemistry & Cell Biology,40(1), 46-62. doi:10.1016/j.biocel.2007.06.009

Mackie, E. J., Tatarczuch, L., & Mirams, M. (2011). The skeleton: A multi-functional complex organ. The growth plate chondrocyte and endochondral ossification. *Journal of Endocrinology,211*(2), 109-121. doi:10.1530/joe-11-0048

Millán, J. L. (2013). The Role of Phosphatases in the Initiation of Skeletal Mineralization. *Calcified Tissue International,93*(4), 299-306. doi:10.1007/s00223-012-9672-8

Minina, E., Kreschel, C., Naski, M. C., Ornitz, D. M., & Vortkamp, A. (2002). Interaction of FGF, Ihh/Pthlh, and BMP Signaling Integrates Chondrocyte Proliferation and Hypertrophic Differentiation. *Developmental Cell,3*(3), 439-449. doi:10.1016/s1534-5807(02)00261-7

Moore, M. C., Gray, G. D., Hale, D. S., Kerth, C. R., Griffin, D. B., Savell, J. W., . . . Oconnor, M. E. (2012). National Beef Quality Audit–2011: In-plant survey of targeted carcass characteristics related to quality, quantity, value, and marketing of fed steers and heifers1,2. *Journal of Animal Science,90*(13), 5143-5151. doi:10.2527/jas.2012-5550

National Cattlemen's Beef Association. (2015). B CS B - BeefResearch. Retrieved May 20, 2019, from http://www.beefresearch.org/CMDocs/BeefResearch/PE_Fact_Sheets/Quality_Grades_Facts_FINAL.pdf

National Center for Biotechnology Information. (n.d.). Retrieved 2019, from https://www.ncbi.nlm.nih.gov/

Nilsson, O., Marino, R., Luca, F. D., Phillip, M., & Baron, J. (2005, November 21). Endocrine Regulation of the Growth Plate. Retrieved April 2, 2019, from https://www.karger.com/Article/FullText/88791

Novotny, S. A., Warren, G. L., & Hamrick, M. W. (2015). Aging and the muscle-bone relationship. Retrieved October 11, 2018, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4285576/

Olney, R. C. (2003). Regulation of bone mass by growth hormone. *Medical and Pediatric Oncology,41*(3), 228-234. doi:10.1002/mpo.10342

Orimo, H. (2010). The Mechanism of Mineralization and the Role of Alkaline Phosphatase in Health and Disease. *Journal of Nippon Medical School,77*(1), 4-12. doi:10.1272/jnms.77.4

Ortega, N., Behonick, D. J., & Werb, Z. (2004). Matrix remodeling during endochondral ossification. *Trends in Cell Biology,14*(2), 86-93. doi:10.1016/j.tcb.2003.12.003

Raines, C. R., Dikeman, M. E., Unruh, J. A., Hunt, M. C., & Knock, R. C. (2008). Predicting cattle age from eye lens weight and nitrogen content, dentition, and United States Department of Agriculture maturity score1,2. *Journal of Animal Science,86*(12), 3557-3567. doi:10.2527/jas.2007-0445

Reiling, B. A., & Johnson, D. D. (2003). Effects of implant regimens (trenbolone acetate-estradiol administered alone or in combination with zeranol) and vitamin D3 on fresh beef color and quality1. *Journal of Animal Science,81*(1), 135-142. doi:10.2527/2003.811135x

Salove, K. C., Doumit, M., Murdoch, G., & Richard, R. (2018). *Relationship Between Heifer Carcass Maturity and Beef Quality Characteristics*. (master's thesis). University of Idaho, Idaho, United States.

St-Jacques, B., Hammerschmidt, M., & Mcmahon, A. P. (1999). Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes & Development,13*(16), 2072-2086. doi:10.1101/gad.13.16.2072

Tang, K., Breen, E. C., Gerber, H., Ferrara, N. M., & Wagner, P. D. (2004). Capillary regression in vascular endothelial growth factor-deficient skeletal muscle. *Physiological Genomics,18*(1), 63-69. doi:10.1152/physiolgenomics.00023.2004

Tatum, J. D. (2011). Animal Age, Physiological Maturity, and Associated Effects ... Retrieved October 30, 2018, from https://www.beefresearch.org/CMDocs/BeefResearch/PE_White_Papers/Animal_Age.pdf

Turner, H. A., Phillips, R. L., Vavra, M., & Young, D. C. (1981). The Efficacy of an Estradiol-Silicone Rubber Removable Implant in Suckling, Growing and Finishing Steers. *Journal of Animal Science,52*(5), 939-944. doi:10.2527/jas1981.525939x

USDA. (2005). Final Report to the Government of Japan - avc-beef.org. Retrieved October 30, 2018, from http://www.avc-beef.org/AgingCattle-Griffin/USDA-AMS-BeefPhysiologicMaturity.pdf

USDA. (2017). Agricultural Marketing Service Carcass Beef Livestock ... Retrieved October 30, 2018, from https://www.ams.usda.gov/sites/default/files/media/CarcassBeefStandard.pdf

Väänänen, H. K., & Härkönen, P. L. (May 1996). Estrogen and bone metabolism. *Maturitas,23*, S65-S69. doi:https://doi.org/10.1016/0378-5122(96)01015-8

Vanderwert, W., Berger, L. L., Mckeith, F. K., Baker, A. M., Gonyou, H. W., & Bechtel, P. J. (1985). Influence of Zeranol Implants on Growth, Behavior and Carcass Traits in Angus and Limousin Bulls and Steers. *Journal of Animal Science,61*(2), 310-319. doi:10.2527/jas1985.612310x

Vortkamp, A., Lee, K., Lanske, B., Segre, G. V., Kronenberg, H. M., & Tabin, C. J. (1996). Regulation of Rate of Cartilage Differentiation by Indian Hedgehog and PTH-Related Protein. *Science,273*(5275), 613-622. doi:10.1126/science.273.5275.613

Weir, E. C., Philbrick, W. M., Amling, M., Neff, L. A., Baron, R., & Broadus, A. E. (1996). Targeted overexpression of parathyroid hormone-related peptide in chondrocytes causes chondrodysplasia and delayed endochondral bone formation. *Proceedings of the National Academy of Sciences,93*(19), 10240-10245. doi:10.1073/pnas.93.19.10240

Weise M, De-Levi S, Barnes KM, Gafni RI, Abad V, Baron J. 2001. Effects of estrogen on growth plate senescence and epiphyseal fusion. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;98(12):6871-6876. doi:10.1073/pnas.121180498.

Yu, Y. Y., Lieu, S., Lu, C., & Colnot, C. (2010). Bone morphogenetic protein 2 stimulates endochondral ossification by regulating periosteal cell fate during bone repair. *Bone*, *47*(1), 65–73. doi:10.1016/j.bone.2010.03.012

Yoon, B. S., Pogue, R., Ovchinnikov, D. A., Yoshii, I., Mishina, Y., Behringer, R. R., & Lyons, K. M. (2006). BMPs regulate multiple aspects of growth-plate chondrogenesis through opposing actions on FGF pathways. *Development,133*(23), 4667-4678. doi:10.1242/dev.02680

Zhao, C., Dahlman-Wright, K., & Gustafsson, J. Estrogen signaling via estrogen receptor β. J Biol Chem. 2010 Dec 17; 285(51): 39575–39579.

Chapter 2

Association Between Carcass Maturity Grade and Genes Involved in Bone Growth, in Young Heifers

Abstract

Over the last several decades, there has been an increase in the number of heifers, under 30 months of age chronologically, categorized as B and C-maturity based upon skeletal maturity at harvest (USDA statistics). The hypothesis for this project is that heifers with advanced physiological maturity have differences within genes that encode proteins involved in bone ossification. The genetic variance within genes that promote advanced rates of ossification relating to bone growth and extracellular matrix mineralization have resulted in young heifers reflecting maturity grades more representative of cattle over 30 months of age. This can result in substantial economic losses for producers since the value of an A-maturity carcass is as much as $350/head greater than when an animal is graded below a B-maturity. Despite the USDA change to allow use of dentition to assess maturity, there are still losses with deduction costs associated with advanced maturity in chronologically young beef cattle. In this study, eight candidate genes known to play an important role in osteogenesis, for sequence variation linkage to advanced maturity were investigated. Targeted gene capture was used followed by sequencing for candidate genes, in commercial cattle graded as A, B, and C maturity groups (N=90 heifers) but confirmed to be under 30 months using records and dentition. Novel and significant genetic sequence variations were identified in association with maturity score within the alkaline phosphatase (*ALPL*), estrogen receptor 1 (*ESR1*), estrogen receptor 2 (*ESR2*), collagen type X alpha 1 (*COL10A1*), parathyroid hormone like protein (*PTHLH*), and myosin

heavy chain 2 *(MYH 2*). In *ESR1,* one of the most significant SNPs identified was located at 9:901156650 with a p-value of 0.002668 with association to maturity grading This represents the first report of this association and this observation could lead to opportunities to genetically screen beef cattle for susceptibility to early maturity.
Key words: heifer, maturity, targeted gene capture, beef carcass, genetic polymorphism, gene and trait study

Introduction

Conventionally raised steer and heifer carcasses under 30 months of age are always graded within the A maturity category. Physiological maturity is classified into 5 different groups: A, B, C, D, and E being the oldest (USDA 2017). Carcass maturity is determined by evaluating the size, shape, and ossification of the bones and cartilages along the split vertebral column of the carcass along with the color and texture of the lean at the 12^{th} rib interface (O'Connor et al., 2007). Currently for an animal to be considered an A maturity, they must have two or fewer permanent incisors. Historically for an animal to be considered an A maturity, they must have; (1) some evidence of cartilage in all vertebrae, (2) distinct separation of the sacral vertebrae and caps that show considerable evidence of cartilage, (3) caps on the lumbar vertebrae that tend to be partially ossified, (4) no ossification of the thoracic vertebrae, (5) split vertebrae surfaces that tend to be soft, porous, and red, (6) ribs that have some tendency toward flatness, and (7) lean texture that is very fine, and lean that is light red in color (O'Connor et al., 2007). However, in the last few decades there have been reports and studies revealing cattle that are under 30 months of age chronologically have been graded and B and even C maturity

carcasses. In research done by Tatum (2011) found that heifers under 30 months of age were 7 times more likely to produce B maturity carcasses and 11 times more likely to produce C maturity or older carcasses.

Until modification to the USDA-AMS grading standards to allow dentition in 2017, physiological maturity was used to determine quality grade via carcass maturity. Despite the modifications for allowing dentition to be utilized to determine quality grade there is still an economic loss associated with cattle graded over 30 months of age and falling below the A maturity category (AMS, 2017). In 2017 before the dentition amendment, carcasses with a quality grade of Standard had a $36.00 per hundredweight discount, while carcasses C maturity or older receives a $39.00 per hundredweight discount compared to choice carcasses (USDA, 2017). This created an economic loss for producers with heifers having a lower maturity due to skeletal maturity. The current research was conducted to determine if genes impacting osteogenesis accelerated the ossification rates in heifers from carcasses verified to be less than 30 months of age by dentition.

Materials and Methods

Carcass Selection

Ninety beef heifer carcasses were selected on four days over a three-month period from a commercial beef processing facility (Toppenish, WA) in late 2015 and early 2016. The left side of the carcass was commercially evaluated for quality and yield grades. The cattle selected were finished in feedlots in Canada and the western United States. Thirty carcasses were selected within each physiological maturity category of A^{00} – A^{99} (A), B

^{00} - B ^{99} (B), and C ^{00} – C^{99} (C). All carcasses were determined to be less than 30 months of age by dentition, in accordance with USDA guidelines. Carcasses were selected to ensure marbling scores between Small^{00} (SM) to Modest ^{99} (MT). The initial selection of the carcasses was based on overall maturity scores form USDA graders and marbling scores obtained from a USDA approved grading system instrument.

Sample Collection

The strip loin [Institutional Meat Purchase Specifications (IMPS) 180; *Longissimus thoracis* (USDA, 2014) of the selected carcasses were used for collecting ninety 5-g samples of skeletal muscle. Thirty samples were collected from A maturity heifers, thirty from B maturity grade heifers and thirty from C maturity grade heifers. All heifers were verified to be under 30 months of age using dentition. The muscle tissue was stored in a15-milliliter (ml) conical tube, at -20 degrees Celsius (ºC). DNA was extracted from the stored samples, as per the subsequent outlined protocol.

DNA Extraction

One gram-wet weight from each frozen sample was taken and ground into a fine powder using chilled mortar and pestles under liquid nitrogen to avoid thawing. Grinding the tissue into a fine powder under liquid nitrogen allowed for more effective DNA extraction while maintaining DNA integrity. The powdered muscle tissue was stored in 15- ml conical tubes and kept in the -20 ºC freezers until needed. Genomic DNA was extracted using a protocol by Gross Bellard and colleagues 1972 (Appendix A). Powdered tissue was weighed out to 0.2 g wet weight and placed in a 15-millilier (ml) conical tube. Digestion buffer (25mM of 100mM EDTA pH 8.0,100mM of 1M NaCl,

0.5% of 10% SDS pH7.2, and 10mM of 1M Tris HCl pH 8.0.) and Qiagen proteinase K (20mg/ml) were added to the tissue. Samples were briefly vortexed in a -Genie Mixer, Scientific Industries, Inc., catalog #22456 to mix and incubated in a 50ºC water bath (Isotemp Water Bath, Fisher Scientific, catalog #16-460-5) for 12-18 hours. Following incubation, 2.5 mL of 25:24:1 phenol/chloroform/isoamyl alcohol from Fisher's was added to each sample and centrifuged at 1700xg for 10 minutes in a Sorvall ST16R centrifuge. The aqueous layer, with an approximate volume of 1.5 ml, was transferred to a new 15 ml tube. An equal volume of 24:1 chloroform/isoamyl alcohol (Sigma) was added to each sample, mixed, and then centrifuged again at 1700xg for 10 minutes. Once again, the aqueous layer was transferred to a new conical tube. Half of the volume of the aqueous layer of 7.5 M ammonium acetate was added to each sample. Chilled (-20 degrees Celsius) 100% ethanol was added at twice the volume of the remaining solution. The tubes were inverted 20 times to mix and then placed in the freezer for an hour to allow DNA precipitation. Samples were centrifuged in a Sorvall ST16R centrifuge at 1700xg for 2 minutes and repeated until a pellet was visible at the bottom of the tube. The supernatant was carefully removed using a Gilson pipettor and discarded. Chilled (-20 degrees Celsius) 70% ethanol (1mL) was added to the pellets to help precipitate any DNA remaining on the wall of the tubes. Finally, the entire samples were centrifuged for 4 minutes at 1700xg. The ethanol was removed from the tubes, and pelleted DNA was air-dried in the fume hood. After drying DNA samples for an hour, 500 µl of nuclease free water (Ambion), was added to each sample and pipetted until DNA was in solution. The solution was transferred to a 1.5-ml microcentrifuge tube and DNA was evaluated for quality and quantity then stored at-20 ºC freezer.

DNA Integrity assessment

Double stranded DNA sample quality and quantity were assessed using a Nano Drop 1000 spectrophotometer (Thermo Fisher). Absorbance values were required to be above 1.8 for A260/A280 and A260/A230. The A260/A280-ratio value indicates purity of the DNA and usually the value is around 1.8. The A260/A230-ratio value is for "pure" nucleic acid and often higher than the actual A260/A280 values. The range for the A260/A230 values needs to be between 2.0-2.2 otherwise any lower indicated contaminants. If the values for the extracted DNA were lower than the standards a repeat DNA extraction for the failed sample was completed. Once the DNA quality was confirmed for all 90 samples, they were transferred at a fixed concentration (100 ng of DNA (10ul of 10ng/ul) into individual wells of a 96 well plate and delivered to the University of Idaho Genomics core for cDNA library preparation, RNA bait set sequence capture and targeted DNA sequencing.

RNA Bait Set Design

Gene locations and sequences for each of the candidate genes and controls were obtained from the NCBI- database, including upstream and downstream sequence information. RNA bait sets were designed utilizing the candidate gene sequences that were complimentary to the sense strand. The sequences of both upstream and downstream regions of the genes were included in the bait set design to extend the pulldown and the evaluation of the candidate genes outside of the coding region. The bait sets themselves are complimentary RNA specific sequences for the targeted genes. This

allowed for effective RNA bait-set capture pull down of the entire region of DNA associated with the target genes. The bait sets allowed for specific capture of the genes selected so that the sequencing was focused on specific targets, thus more effectively evaluate the target genes versus the entire genome. The advantage of using RNA bait-sets was their high binding affinity (target efficiency) as well as the ability to selectively digest the bait sets following DNA capture and prior to sequencing so that all of the reads were from the original DNA sources. This experiment represents the first time the University of Idaho has used RNA bait set capture for targeted sequencing. Bait-sets were generated that would span some previously reported bovine single nucleotide polymorphisms to serve as landmarks for successful pulldown, this is discussed further in the results section.

The SNPs were selected from the following sources:

1) The ~2900 SNPs from the Illumina Bovine3K Genotyping BeadChip downloaded from http://bioinformatics.tecnoparco.org/SNPchimp/index.php/download/download-cow-data;

2) The 7,931 SNPs from the Illumina BovineLD v2.0 set: https://www.illumina.com/products/by-type/microarray-kits/bovineld.html; downloaded from http://bioinformatics.tecnoparco.org/SNPchimp/index.php/download/download-cow-data;

3) The ISAG Cattle Molecular Markers and Parentage Testing set "ISAG cattle core + additional SNP panel 2013"; downloaded from

https://www.isag.us/committees.asp?autotry=true&ULnotkn=true; https://www.isag.us/Docs/Cattle-SNP-ISAG-core-additional-panel-2013.xlsx.

All three lists were uploaded into the statistical program "R", and duplicates were removed. Seventeen targeted genes were selected based on their functions and were predicted to related to beef carcass maturity (Table 2.1). The list of SNP loci and targeted gene loci were provided to MYcroarray. MYcroarray generated several sets of targets based on stringency of multiple mapping through MYbaits design pipeline. The least stringently filtered set was chosen to ensure the maximum capture while still retaining specificity. SNPs were tiled at 1x (1 probe per SNP). Gene regions were tiled at 2x (meaning that all regions were covered by 2 overlapping capture baits). The set of targeted probes were synthesized by MYcroarray and used by the IBEST Genomics Resources Core at the University of Idaho to capture shotgun libraries prior to sequencing on the MiSeq.

Targeted Sequence Capture

Single nucleotide polymorphisms (SNPs) were extracted from the bosTau8 genome at 120 bp lengths from 2875 unique locations. In addition, 17 larger non-SNP target loci were also extracted from the bosTau8 genome with a total target length of 1,293,110bp. Stretches of 10 or fewer unknown nucleotides (N) were replaced with T's. 120nt baits with 2x~ tiling density were designed; 120bp regions would have only 1 bait = 18,220 unfiltered baits. Each bait candidate was "BLASTed" against the bosTau8 genome and a hybridization melting temperature (T_m) was estimated for each hit assuming standard MYbaits® buffers and conditions. Only the top 1000 hits are

considered; hits >1000 are ignored because they are non-specific regions. This allowed for the capture of only the targeted DNA regions. The biotinylated RNA library baits were mixed with the genomic samples and a hybridization (HYB) buffer then incubated to allow for the hybridization. Streptavidin coated magnetic beads were added to the mix. These beads bind the RNA baits that have captured the targeted gene regions. Using a magnetic pull, the bound RNA sets were separated from any unbound DNA fragments, this was the targeted bead capture step. Following capture, the magnetic beads were washed away, and the RNA bait sets were digested leaving only the desired specific DNA segments. The DNA segments were then amplified for high-throughput sequencing. Using Nextera Libraries with Mybaits is how the libraries were generated and used a V3 2x300 MiSeq run to generate the sequences.

Gene Alignment

From the ninety sequenced samples, data was received as fastq.gz files with sequence data. The files were downloaded and imported into Qiagen's CLC genomics workbench version 11.0 lab then realigned to Bos_taurus_UMD_3.1.1 primary assembly. Of the 90 DNA samples sent, eight had very poor capture yield and sequencing quality and were therefore excluded from further analysis. With the loss of eight animals the final experimental animal data set was comprised of twenty-seven A maturity, twenty-eight B maturity, and twenty-seven C maturity. A workflow was created that aligned the samples to the most recent Bovine reference genome (Bos_taurus_UMD_3.1.1), read the tracks mapped, called insertion/deletions (indels) and structural variants, performed local alignment, fixed ploidy variant detection, filtered against known variants, and generated

statistical analysis for target regions (Figure 2.1). After alignment for all samples to the reference genome, any variations or changes within the sample sequences from the reference were highlighted. The aligned sequences were then saved as "VCF files" and exported.

SNP Variation Calling

Using SNP & Variation Suite v8 (SVS) from Golden Helix, the VCF files were imported in as individual samples, the fields G_T (genotypes), AD (allelic depth), DP (read depth), ref/alt (reference/alternate), and QUAL (quality) were selected to be included in the spreadsheets created, once imported. This yielded a spreadsheet including all 82 samples and the thousands of SNPs, indels, deletions, and insertions called between all the samples (Figure 2.2). The spreadsheet was broken down by bovine autosomal chromosomes 2, 5, 9, 10, 13, 19; where the genes of interest resided. Variants called with a minor allele frequency (MAF) less than 0.05 were filtered out of the spreadsheet. Next all insertions and indels were removed that did not meet the read stringency and therefore had a low confidence level. Subsequently, the isolated each gene regions within the chromosomes on the spreadsheet. The next criterion was a minimal read depth of 10 reads if homozygous for the reference and 15minimum read depth for heterozygous to be considered true. Read depth was evaluated by manual scrutiny of the specific sequences, of every SNP, at each location, for every sample to determine the accurate genotype (If I may editorialize; "time consuming human labor"). If less than the required number of reads for a SNP, of any one of the samples was detected, it was marked as "missing data" and removed from further analyses. If more than 17 of the 82 samples (20%) have

missing data, in a particular SNP, then it was removed from the spreadsheet since the confidence in the SNP called was less than 80%. Having completed the initial evaluation of all the SNPs for each gene, another spreadsheet with the maturity (phenotype) for each sample was merged to yield a finalized spreadsheet. More specifically, this determined in what manner the called significant polymorphism was associated with the maturity trait. For instance, it was tested for either basic or recessive association, as well as with the A versus B versus C maturity phenotype or the A versus B and C combined maturity phenotype. The phenotype column was then highlighted as the categorical column, which allowed for genotype association tests to be run, identifying significant trait associated SNPs. There are 5 genotype association tests that can be run: the basic allelic test, genotypic test, additive model, dominant model and recessive model. These tests will be described in further detail in a forthcoming section of this thesis.

GATK Variant Calling

The raw reads were cleaned using HTStream (https://github.com/ibest/HTStream) to remove the duplicated reads, trim adapter sequences and low-quality ends. Clean reads were again mapped to the reference bovine genome (Bos_taurus_UMD3.1.1/bosTau8) with BWA version 0.7.17 (Li and Durbin, 2009) using default parameters and further processed with SAMtools (version 1.5). The variant calling was performed by GATK version 4.0.2.0 (McKenna A et al., 2010) HaplotypeCaller at the designed genomic intervals. GATK generate 32,252 SNPs in total. The SNPs from eight underperforming samples (S5, S19, S20, S53, S59, S68, S85, S87) were removed. Then the SNPs were filtered through two steps: Firstly, all the SNPs with the missing data is greater than 20%

and MAF less than 5% were removed. After filtration, 11,259 SNPs remained. Secondly, at each genotype, only the SNPs with DP (sequence depth) that is equal or greater than 10 were retained. Otherwise, the genotype was treated as missing data. Then SNPs with the missing data (>20%) and MAF (<5%) were removed. After filtration, 575 SNPs remained. This analysis was only used for the GWAS portion of this study. The rest of the analysis on the candidate genes was done using the SNP & Variation Suite v8 (SVS) from Golden Helix.

BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway)

A total of 11,259 SNPs with a MAF greater than 0.05 and marker call rate greater than 80% were used to carry out the Genome Wide Association Study (GWAS). Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al, 2017) performed the GWAS. The SNP set was used to calculate the genetic covariance matrixes from genotypes by the R/Bioconductor package SNPRelate (Zheng et al., 2012). The phenotype file, genotype file and covariance were provided to BLINKC (https://github.com/Menggg/BLINK) to conduct GWAS analysis.

Statistical and Bioinformatics analysis

Five different genotype association tests were used. The basic allelic test, genotypic test, additive model, dominant model and recessive model test were run for each gene to identify significant SNPs. The two phenotype sheets that were used in this study were the 1v2v3 (A versus B versus C maturity) phenotype or the 1v2 (A versus B and C combined maturity) phenotype. The 1v2v3 (A versus B versus C maturity)

phenotype sheet was comparing the three maturities against each other to distinguish SNPs that are leaning to one polar end of the maturities or the other. The 1v2 (A versus B and C maturity) phenotype tested the desired maturity A against both of the less desirable B and C maturity. Both phenotype sheets were paired with the final SNP list for each gene then ran through either the basic allele association test or the recessive allele association test. Having these two phenotype sheets allowed for a comparison of the maturities (phenotypes) and allowed for better assessment of the SNPs and what maturity they were associated with. For each test alleles can be classified based on reference or alternate allele status as specified by a marker map field. This allows the association testing of DNA-Sequence data where the reference alleles are known for variants and all tests need to be in terms of the alternate allele(s) or alternate sequence. The reference field should only contain information about the reference allele(s) or sequence (SVS Golden Helix, 2018). The genotype association test requires datasets that contain genotype data with either a case/control or quantitative traits. For the data ran, the phenotype of A, B, C or A, BC was selected as the quantitative trait. Once selecting the column going to the genotype tab and selecting genotype association test from the menu, this “opens up” a window that allows for the model wanted and different parameters to be selected. There are then 3 tabs within the test window, the association test parameters, PCA parameters and the overall marker statistics tab. The association test parameters tab contains all the parameters necessary for the association test plus options to select principal components analysis for stratification corrections. The second tab, the PCA parameters contain all the remaining parameters for PCA. The final tab, overall marker

statistics contains the parameters for obtaining statistics that are independent of any association test.

The case/control dependent variable that was selected for each test is the "classify alleles by allele frequency". Then the genetic model test and the different test statistics can be selected. Also missing values and the multiple testing corrections can be selected along with any additional outputs as seen in figure 2.3. On the overall marker statistics tab, the options selected were call rates, number of alleles, allele frequencies, Hardy-Weinberg Equilibrium (HWE) p-value, and genotype counts for the association study that was ran as seen on figure 2.4. After running the genotype association test the results are placed on an output spreadsheet. Each statistics method calculated, is presented in its own column.

The first association test ran was the basic allele test. For the basic allelic test, the genotypes dd, Dd, and DD are resolved into pairs of alleles dd, Dd, or DD. Both elements of each genotype are considered to correspond to the same value of the dependent variable. This test simply tests for the frequency of D v d. The associations with these individual alleles are then tested.

The next test is the genotypic test. This test refers to comparing the genotypes dd, Dd, and DD without regard to any "order" or allelic count or allelic pairing they might have. This test can reveal associations without regard to any specific genotype model. No associations are "hidden" because no model is assumed (SVS Golden Helix, 2018).

Under the additive model, testing is designed specifically to reveal associations, which depends on the allele classification. If the alleles are classified according to allele frequency then the associations depends on the minor allele–that is, where having two

minor alleles (DD) rather than having no minor alleles (dd) is twice as likely to affect the outcome in a certain direction as is having just one minor allele (Dd) rather than no minor alleles (dd) (SVS Golden Helix, 2018). This model says the heterozygote is approximately halfway between the ref and alt homozygous.

In the Dominant test, if the alleles are classified according to allele frequency then this model specifically tests the association of having at least one minor allele D (either Dd or DD) versus not having it at all (dd). If the alleles are classified according to reference/alternate alleles then this model specifically tests the association of having at least one alternate allele A (either Ar or AA) versus not having it at all (rr) (SVS Golden Helix, 2018).

The final test that can be used is the recessive test. If the alleles are classified according to allele frequency, then this model specifically tests for the association of having the minor allele D as both alleles (DD) versus having at least one major allele d (Dd or dd). If the alleles are classified according to reference/alternate alleles then this model specifically tests the association of having the alternate allele A as both alleles (AA) versus having at least one reference allele r (Ar or rr) (SVS Golden Helix, 2018).

Under each genotype association test there is also test statistics. One test statistic available is the correlation/trend test. This test works best in the additive, dominant and recessive models, but is not available for the genotype model. The correlation/ trend test is the only test that is available when Principal Components Analysis (PCA) is used for stratification correlation. This test will show the p-value for the dependent variable value having any correlation with or trend that depends on the count value of the genotype (SVS Golden Helix, 2018). In the additive model, the count of the minor allele D, which

is zero within genotype dd, one within genotype Dd, and two within genotype DD, where d is the major allele. The dominant model, the count is one for genotypes DD and Dd and zero for genotype dd. Finally, in the recessive model: The count is one for genotype DD and zero for genotypes Dd and dd. The "Corr/Trend R" output from this test indicates the effect direction. A positive direction means that a greater count of the minor or alternate allele was present, versus the major or reference allele correlating with an increased effect (SVS Golden Helix).

Another test that was used during the analysis is the F-Test. This test is available for a quantitative dependent variable. It can be used in all genotype models except for the additive model. This tests whether the distributions of the dependent variable within each category are significantly different between the various categories of the predictor variable (SVS Golden Helix, 2018). Any SNPs with a p-value ≤ 0.05 were declared significant.

Linkage Disequilibrium

After identifying and creating the final spreadsheets for each candidate gene with the polymorphisms that the researchers were confident in was ran through the linkage disequilibrium (LD) test. Linkage disequilibrium looks at the non-random association of alleles at multiple loci within a population and do not occur randomly with respect to each other. Loci are said to be in linkage disequilibrium when the frequency of the association of their different alleles is higher or lower than what is expected if the loci were independent and randomly associated or distributed (Slatkin, 2008). The statistics from this test can be ran as either R squared of D'. Within the LD plot areas that are

shades of red are in different levels of LD whereas the areas in blue are not. This test identifies what SNPs are in linkage disequilibrium and if they were in haplotype blocks or not. Haplotype blocks show high levels of linkage disequilibrium and is a region in the organism's genome that there are low genetic recombination events. To compute haplotype blocks, on the LD screen select marker block tab and select visible blocks in the options then click run. The algorithm with produce the blocks in a black outlined pentagon on the top of the LD plot. Once the block is created selecting on the block gives details about it. Using the LD plot, researchers were able to determine what significant SNPs were in LD with each other.

Results

The subsequent results section summarizes the phenotypically associated SNPs identified in the candidate and control genes, within this research's population of A-maturity (n=27), B-maturity (n=28) and C-maturity (n=27) beef cattle. Though originally collected tissue from 90 animals, 8 animals were excluded from analyses due to library or related sequencing failure. Table 2.2 is a brief overview of the locations of candidate genes selected in this project.

Significant variants within Alkaline Phosphatase

Alkaline phosphatase (*ALPL*) is located on autosomal chromosome 2 specifically 131,791,668bp-131,857,687bp in the *Bos taurus* genome (NCBI, 2019). This gene is transcribed on the reverse strand and consists of 12 exons and 11 introns within the

annotated gene coding sequence. *ALPL* is an osteoblast differentiation marker and provides the inorganic phosphate essential for mineralization of the extracellular matrix.

Within the *ALPL* gene, thirty-three SNPs were reported that have sufficient read depth and number of reads and from those SNPs three significant SNPs were identified at positions 2:131,827,143, 2:131,837,201 and 2:131,837,038 in intron 1 on the A versus BC phenotype sheet which can be observed in table 2.3. The SNP 2:131,837,201 is significant (P= 0.015), the SNP at 2:131,827,143 is significant (P= 0.014), and finally the SNP at 2:131,837,038 has a p-value equal to 0.021 which can be observed in figure 2.5.

The SNP identified at 2:131,827,143 had an allele change from guanine (G) to adenine (A) nucleotide. The major allele frequency relative to the SNP is shown in table 2.4. At this particular SNP the major allele is the same as the bovine reference allele G. This means that the A maturity animals have more of the reference allele G within their population whereas the B&C maturity animals have the alternate A nucleotide allele within their population. This SNP has been reported on Ensembl previously however this is the first reported association to the less desirable maturity phenotype.

There is a nucleotide change from the reference genomes version Bos_taurus_UMD_3.1.1 that reports a thymine (T) to the alternate allele cytosine (C) at position 2:131,837,201. In figure 2.6 researchers were able to capture an accurate representation of the allele ratio between maturities at this particular SNP. In this figure it shows calling of the alternate allele C alleles in the A maturity animals, a mix of C and T in the B maturity animals and finally mostly the reference allele T in the C maturity animals. Figure 2.6 shows an example Angus animal not from this study and its genotype in comparison to those in this study and giving insight that at this particular location the

SNP identified and linked to advanced maturity is out in the population. Within the A maturity animals (n=27) there were thirteen animals with the T allele with 5 homozygous T/T, 8 being heterozygous T/C, 8 being homozygous C/C and 6 having data missing. In the combine B&C maturity animals (n=55) 20 homozygous T/T, 20 being heterozygous T/C, 6 being homozygous C/C and 9 having data missing so unable to call. The major allele frequency relative to the SNP is shown in table 2.4. When looking at the different maturities, it was clear that the B and C maturity had more of the T allele called similar to the bovine reference genome animal, whereas the A maturity animals have more of the alternate C allele called at this locus, in the study population. This is a novel discovery as it's the first ever report of a specific nucleotide variation in this gene linked to the quality trait of carcass maturity.

The final SNP identified within the *ALPL* gene is located at 2:131,837,038. At this position there is an allele change from the reference genome adenine (A) to the alternate guanine (G) nucleotide. At this SNP the major allele is the reference allele A. The less desired maturity phenotype B&C tended to have more of the reference allele A within its population. The A maturity animals had more of the alternate allele G within its population. This demonstrates that the more desired A maturity actually had the alternate allele as compared to the less desired maturities, at this specific locus. This is another SNP that had been previously identified but was not previously linked to the advanced maturity phenotype.

There were no significant SNPs identified within the recessive test with the A versus BC phenotype sheet. Also, no significant SNPs were identified with the A versus B versus C phenotype sheet on either the basic allele or recessive association test. No

SNPs were identified using the Genomic Analysis ToolKit (GATK) variant calling. GATK is an industry standard used for identifying SNPs and other polymorphisms in DNA and RNAseq data. This analysis was used to identify possible SNPs that were missed when using the SVS software and possibly excluded due to type 1 error. The software also allowed the researchers to run an GWAS study across the animal population.

While the polymorphisms that were reported had all been previously reported on Ensembl, this study is the first to link them to the economically important trait of advanced maturity phenotype. Given the role of alkaline phosphatase in ossification, the discovery of an association of these SNPs with an increased rate of bone deposition (skeletal maturity), this discovery has physiological merit.

Significant variants within Parathyroid Hormone like Protein

Parathyroid hormone like protein (*PTHLH*) is located on autosomal chromosome 5 in the *Bos taurus* genome. This gene is most recently mapped to the region from 82,241,708bp-82,258,858bp on the forward strand and is a protein-coding gene (NCBI, 2019). This gene has 2 transcripts, so depending on which transcript is being read it either has 3 or 4 introns and 4 or 5 exons (NCBI, 2019). The parathyroid like protein balances the activity of osteoblasts while also inducing the ossification relation to bone deposition (Karaplis et al. 1994). Within the *PTHLH* gene there were 23 SNPs in the final dataset after excluding SNPs that failed to meet the minimum stringency requirements. After phenotypic association analyses only one polymorphism was identified as significant within this gene. The variant detected is located at 5:82,249,219 in int ron 3 and is a

deletion of the reference (Bos_taurus_UMD_3.1.1) alleles CC. The p-value of the deletion is 0.045 and was identified with the A versus BC maturity phenotype comparison on the basic allele association test. The major allele is CC with the minor allele being a deletion of this cytosine repeat. The A maturity animals had more of the reference allele compared to the B&C maturity animals. The less desirable maturities of B and C had the alternate allele containing the deletion.

There were no significant SNPs identified within the A versus B versus C maturity phenotype in either the basic allele or recessive association test or in the A versus BC phenotype within the recessive association test in the sampled heifers. Using the GATK variant calling, a single SNP was identified within *PTHLH* in intron 3. This SNP was located at 5:82,248,608 and has a p-value of 0.0379. The reference (Bos_taurus_UMD_3.1.1) allele is cytosine (C) and the alternate being a thymine (T). The major allele is the reference allele. The A maturity animals had more of the reference allele C compared to the B and C maturities. The location of the SNPS within the *PTHLH* gene can be observed in figure 2.7 which shows the location and p-value of each SNP. All significant SNPs identified have been previously reported on Ensembl but were not yet linked to the advanced maturity phenotype. Therefore, it was observed that the following SNP located at 5:82,249,219 is associated with maturity grade in young cattle with A-maturity having the reference allele CC and B and C maturity more commonly having the alternate deletion. The following SNP 5:82,248,608 is associated with A maturity animals having the reference allele C and the less desirable maturities B and C having the alternate allele T.

Significant variants within Estrogen Receptor 1

Estrogen receptor 1 (*ESR1*) is located on autosomal chromosome 9 in the *Bos taurus* genome. It is located on the forward strand between 89,969,586bp-90,255,801bp and is a protein coding gene (NCBI, 2019). In heifers and steers, it is known that estrogen increases the rate of skeletal ossification. Estrogen decreases the rate of chondroblast duplication, accelerates chondrocyte maturation and hypertrophy while further accelerating the rate of osteoblast activity until there is no more cartilage to ossify and the plate fuses. This gene contains the most called SNPs of all of the candidate genes, this is at least partially explained by the large gene size but may also be indicative of variation in this gene and its relationship with the animal's phenotype. Within the ESR1 gene (AKA ER1, estrogen receptor α), 310 SNPs were identified before being run through genetic association analysis tests with one of the two phenotype sheets either 1v2 (A versus B and C maturity) phenotype or 1v2v3 (A versus B versus C maturity) phenotype. Out of the 310 SNPs, forty-seven SNPs were identified as significant after analyzing the 310 SNP for the various association tests with the two phenotype data sets. Twenty-four of the significant SNPs were identified with the A versus B versus C phenotype comparison on the basic allele association test. Ten SNPs were identified with the A versus B versus C phenotype on the recessive association test. Twelve significant SNPs were identified on the basic allele association test with the 1v2 (A versus B&C) phenotype. With the recessive association test one SNPs were identified as significant

with the 1v2 (A versus B&C) phenotype. The SNPs location, p-value, test found in are as shown in table 2.5.

One of the most significant SNPs identified was at position 9:90,115,650 within intron nine has an allele change from guanine (G) to adenine (A) with a p-value equal to 0.002. This SNP was identified on the basic allele association test with the A versus B versus C phenotype sheet. The major allele at this position was the reference and after calculation the major allele frequency was highest in the A maturity animals compared to the B maturity and C maturity animals. The C maturity animals have more of the alternate allele within its population (see table 2.6). This SNP 9:90,115,650 is also in linkage disequilibrium with three other significant SNPs 9:90,115,988, 9:90,116,049, and 9:90,116,111. These three SNPs were identified on the recessive association test with the A maturity versus B maturity versus C maturity phenotype sheet (see table 2.7). The SNP 9:90,115,988 has an allele change of the reference (Bos_taurus_UMD_3.1.1) T to the alternate C with a p-value of 0.006 and the trend being that more of the reference T is identified in the A maturity animals whereas more of the alternate C is identified within the C maturity. At the position 9:90,116,049 there is an allele change from the reference A to the alternate C. The p-value for this SNP is 0.005 with more of the reference allele being in the C maturity animals whereas more of the alternate allele is in the A maturity population. Finally, the SNP 9:90,116,111 has the allele change of the reference T to the alternate A with the p-value being 0.014 and the trend being more of the reference allele within the C maturity animals whereas the A maturity animals have more of the alternate within them. In addition, 9:90,116,049 and 9:90,116,111 are in the haplotype block 12 within the estrogen receptor 1 gene meaning that within the ESR1 gene that those areas

within the different haplotype blocks are in high levels of linkage disequilibrium (LD). In figure 2.8 the entire linkage disequilibrium map for ESR1 is shown. This figure also reveals all the different haplotype blocks found within this gene and the levels of LD each SNP has with others based on color. Specifically, with this figure the researchers focused on identifying which of the significant SNPs found were in LD with each other so that in future research genetic tests could be created.

Another highly significant SNP identified during this association study is located at 9:90,180,362. This SNP was discovered with the basic allele association test with the phenotype sheet A versus B versus C maturity. It is located in intron 9 with an allele change from the reference (Bos_taurus_UMD_3.1.1) C to the alternate T and a significant p-value of 0.003. After assessing the major allele frequency for this SNP, it was identified that the A maturity animals have more of the reference allele within their population while the C maturity animals has more of the alternate within their population. Furthermore, this SNP is in linkage disequilibrium according to the D' value of 0.89 with another significant SNP 9:90,178,812. The SNP 9:90,178,812 was also identified with the basic allele association test with the A versus B versus C maturity phenotype sheet. It has a significant p-value of 0.012 with an allele change from the reference (Bos_taurus_UMD_3.1.1) A to the alternate G. After calculating the major allele frequency, the trend found was that more animals within the C maturity group had the reference/major allele whereas animals within the A maturity have more of the alternate allele within the group.

The third most significant SNP identified within the basic allele association test with the A maturity versus B maturity versus C maturity phenotype sheet is located at

position 9:90,097,671 in intron 9 with an allele change of a G to an A and the p-value of 0.003. At this SNP it is observed that the C maturity animals and more of the reference (Bos_taurus_UMD_3.1.1) G allele within its group while the A maturity animals have more of the alternate allele A in its population. This particular SNP is in linkage disequilibrium with three significant other SNPs; 9:90,097,772, 9:90,098,969, and 9:90,103,219. All these SNPs are located within intron 9. The SNPs 9:90,097,671 and 9:90,097,772 are both in haplotype block 6 within the ESR1 gene. Both SNPs 9:90,097,772 and 9:90,103,219 were also identified within the basic allele association test with the A maturity versus B maturity versus C maturity phenotype sheet. The SNP 9:90,098,969 was identified within the basic association test with the A maturity versus the B&C maturity combined phenotype sheet. The rest of the significant SNPs identified with the basic allele association test can be viewed on table 2.6. This table identifies the major allele for each SNP, the major allele frequency and the association trend for maturity grade with each SNP.

On the recessive association test with the A versus B versus C maturity phenotype out of the ten SNPs identified three were highly significant. The most significant identified was found at the location 9:90015095 in intron 7 with an allele change from the reference allele C to the alternate allele T. Its p-value is 0.003 and this particular SNP is associated with seeing more of the reference allele C in the A maturity animals whereas the C maturity animals have more of the alternate allele T in their group. The next most significant polymorphism identified is located at 9:90116049 in intron 9 with an allele change from the reference A to the alternate C with a significant p-value of 0.005. After calculating the MAF for this particular SNP the report is that the C maturity animals have

more of the reference allele A within the group whereas the animals in the A maturity group have more of the reference allele C within their group. What is also important about this SNP is that it is in haplotype block 12 in ESR 1 gene with another significant SNP identified, 9:90116111 and it is in high linkage disequilibrium with that SNP at 0.85. This likely means that a genetic test designed for one of these SNPs may validate across a population given that these two SNPs segregate together. The SNP9:90116049 is also in linkage disequilibrium at 0.76 with another polymorphism identified 9:90115988 which is also the third most significant SNP identified on the recessive test with the A versus B versus C maturity phenotype. The SNP 9:90115988 is located within intron 9 with a p-value of 0.006 and the allele change at this SNP is from the reference T to the alternate C. At this particular polymorphism the trend identified within the maturities is that the A and B maturity animals have more of the reference allele within their populations whereas the C maturity animals have more of the alternate allele C in their group. The rest of the SNPs identified with this test, their major allele frequency and the association trend for maturity grade with each SNP can be identified on table 2.7.

Twelve additional SNPs were identified on the basic allele test with the A versus B and C maturity phenotype and two were considered highly significant. The fist polymorphism identified was at the location 9:8990110 in intron 5 with an allele change from the reference C to the alternate T and a p-value of 0.008. At this particular SNP the alternate allele T was also the major allele and after calculating the major allele frequency more of the alternate allele is seen in the combined B and C maturity animals which is the less desired maturities whereas the reference allele C is seen more in the A maturity animals. The next most significant SNP identified is at the location 9:90103736 within

intron 9. Its p-value is 0.009 and there is an allele change from the reference C to the alternate T. After calculating the major allele frequency for this SNP, the trend that was noticed is that the Band C combined maturity had more of the reference allele within their population whereas the animals in the A maturity group has more the alternate allele. This SNP is also in haplotype block 8 with ten other significant SNPs identified from the tests. This SNP is in linkage disequilibrium with multiple other significant SNPs including 9:90098969, 9:90103219, 9:90103583, 9:90103589, 9:90103593, and 9:90104212. The rest of the significant polymorphisms found can be located on table 2.8. This table identifies the major allele for each SNP, the major allele frequency and the association trend for maturity grade with each SNP.

On the final test, recessive association test with the A versus B&C maturity phenotype sheet one SNP was identified. This SNP is located at 9:90097591 with an allele change from the reference T to the alternate C on intron 9. The p-value is 0.039 and after calculating the major allele frequency it was noted that there are more of the reference allele T in the less desired B&C maturity combine animals compared to the A maturity animals which have more of the alternate allele C in their group as seen on table 2.9. This SNP is also a part of the haplotype block 6 within the gene along with four other significant SNPs.

What is unique about the amount of SNP identified within *ESR1* gene is that one is a novel discovery not yet reported by Ensembl. This first reported SNP is found within intron 9 at 9:90,104,171. It was discovered on the basic allele association test with the A versus B versus C maturity phenotype sheet. This SNP has an allele change from the reference allele C to the alternate allele T. Its p-value is 0.036 and after calculating the

major allele frequency it was determined that the trend showed more of the reference allele in the C maturity population whereas the A maturity animals have more of the alternate allele within their group. Given the size and importance of the estrogen receptor 1 gene finding variants not yet identified is important for the next steps in studying these SNPs and their linkage to advance maturity. The rest of the SNPs identified in this study have all been reported before but none of the SNPs found have yet been reported linked to the advanced carcass maturity phenotype.

Using the GATK variant calling, three SNPs remained significant. Two were called with this software whereas not identified or considered significant with the SVS suite. The other SNP identified was found in both software's however had a more significant p-value within the SVS suite software. The two SNPs identified with GATK are 9:90,078,489 and 9:90,097,792. The general location of each SNP, its p-value and what test it was identified on is demonstrated in Figure 2.9 sections A and B. These figures give a general view of how close each SNP is to each other and the multitude of polymorphisms identified within this gene in respect to the intron and exons. These SNPs are all found on the forward strand and cluster mostly within intron 9.

Significant variants within Estrogen Receptor 2

Estrogen receptor 2 (*ESR2*) is located on autosomal chromosome 10 in the *Bos taurus* genome. This is a protein-coding gene and is located between 76,706,017bp-76,771,172bp on the reverse strand (NCBI, 2019). ESR2 has 7 introns and 8 exons. Similar, to the relationship of estrogen receptor 1 with reproductive fitness selection in livestock, estrogen receptor 2 has the similar potential for inadvertent selection that

impacts osteogenesis rates. Within the ESR2 gene (ER2, estrogen receptor β), 47 SNPs were identified before being run through the basic allele and recessive test with one of the two phenotype sheets. Out of the 47 SNPs in the final spreadsheet two were identified as significant. The two significant SNPs were identified with the A versus BC maturity phenotype in the basic allele association test. The two SNPs are located at 10:76,769,026 and 10:76,767,492. Both of the SNPs were identified as intergenic variants, the p-values and the changes in alleles are seen in table 5. The two SNPs, 10:76,769,026 and 10:76,767,492 have been reported on Ensembl but not yet linked to the advanced maturity phenotype. No SNPs were identified using the GATK variant calling. In figure 2.10 a Manhattan plot identifies significant SNPs found using the GATK software. Both of the SNPs identified demonstrated that the B and C maturity group has the reference allele more commonly whereas the A maturity animals have more of the alternate allele among them. These results indicate that the estrogen receptor genes are exceptionally good and novel candidates for future validation in a larger more diverse population of cattle. Furthermore, these results beg the question as to whether beef cattle with a genetic marker and associated propensity for advanced maturity based on estrogen receptor variants may be impacted by common estrogenic based implants?

Significant variants within Myosin Heavy Chain 2

Sarcomeric myosin heavy chain 2a (*MYH2*) is a protein coding gene and is located on chromosome 19 between 30,137,767bp-30,165,109bp (NCBI, 2019). This gene is important for the development and growth of skeletal muscle. It is probable that rapidly growing and/or maturing cattle would be required to express this gene more in

association with the more significant lean growth. This gene has 40 exons and 39 introns. Within the *MYH2* gene on the finalized sheet in SVS before genetic association analysis 13 SNPs were left with the finalized data set prior to the association tests with the different phenotypes. There were no significant results found with the A versus B versus C maturity phenotypes with the basic allele and recessive association test. Also, no significant cells were identified within the A versus BC maturity phenotype with the basic allele and recessive association test. Using the GATK variant calling; a significant SNP was discovered in the MYH2 gene on intron 5. This SNP has a p-value of 0.013. The significant SNP identified has been reported on Ensembl but not yet linked to the advanced maturity phenotype.

Results within Collagen Type X alpha 1

Collagen type X, alpha 1 (*COL10A1*) is a protein-coding gene on the forward strand of chromosome 9. It is located between 34,961,718bp-34,994,278bp (NCBI, 2019). It has 2 introns and 3 exons. *COL10A1* is a short chain collagen that is expressed by the hypertrophic chondrocytes during the hypertrophic stage of endochondral ossification as a primary signal for the start of matrix ossification. Within *COL10A1* gene there were 116 SNPs identified in the final product sheet, this sheet included all the SNPs that passed the early on requirements before running the more advanced analysis of the basic allele or recessive test with the two phenotypes. Subsequent to the advanced maturity analyses no significant SNPs were identified. No SNPs were identified using the GATK variant calling.

Results in Indian Hedgehog

Indian Hedgehog (*IHH*) is found on autosomal chromosome 2 in the *Bos taurus* genome. This gene runs from 107,722,665bp-107,728,939bp on the reverse strand and is a protein coding gene (NCBI, 2019). *IHH* has 2 introns and 3 exons. The Indian Hedgehog signaling pathways plays a pivotal role in coordinating several different cellular processes, which are essential for morphogenesis of the vertebrate skeleton. Within the IHH gene there were no SNP that passed the stringency required to confirm a confident SNP. In accordance with these observations, no SNPs were identified using the GATK variant calling within this gene either.

Results in Bone Morphogenetic Protein 2

Bone morphogenetic protein 2 (*BMP2*) is located on autosomal chromosome 13 in the *Bos taurus* genome. This is a protein-coding gene and is mapped to the forward strand between 49,550,050bp-49,561,304bp (NCBI, 2019). *BMP2* stimulates the production of bone by inducing the formation of both cartilage and bone. It influences the rate of production of both hypertrophic chondrocytes and mature osteoblasts. Very few SNPs were called, and all failed the rigorous stringency tests to be considered true. In fact, only 3 SNPs were on the final *BMP2* data set that passed the stringencies set prior to the advanced maturity analysis with the phenotypes and association tests were ran yielding no reportable significant genetic polymorphisms. No SNPs were identified using the GATK variant calling within this gene either.

Results in Growth Hormone

Growth hormone (*GH1*) is found on autosomal chromosome 19 in the *Bos taurus* genome. This gene runs from 48,768,617bp-48,772,049bp on the reverse strand and is a protein-coding gene, with the cognate protein also referred to as growth hormone (NCBI, 2019).. It has 5 exons and 4 introns. Growth hormone enhances the anabolic responsiveness of the growth plate and stimulates the expansion of the chondrocyte progenitor cells. In the *GH* gene there weren't any significant SNPs found. Similarly, the less rigorous GATK variant calling within this gene failed to identify and significant associations.

Negative control gene Vascular Endothelial Growth Factor

Vascular endothelial growth factor (*VEGF*) is the negative control gene. *VEGF* is located on chromosome 23, from position 17,255,515bp-17,270,515bp (NCBI, 2019). Vascular Endothelial Growth Factor (*VEGF*) is a member of the growth factor family and encodes for a heparin-binding protein (NCBI, 2019). This factor leads to the migration and proliferation of vascular endothelial cells and essential for angiogenesis. This gene was selected based on the physiology premise that variation within this gene is not expected to have an impact on ossification and therefore should not be linked to these maturity traits. When looking at *VEGF* in both the SVS suite software and the GATK variant calling there were no changes in the sequences keeping it as the control.

SNPs identified using GATK variant calling

When creating the bait sets for the library preparations and targeted sequence capture there was extra "room" on the chips allowing for other genes to be selected for capture which is shown in Table 2.10. Of the genes below and besides the candidate genes for this project there were significant SNPs found in *NR3C1, IGF1, TNFSF11, MYCBP2, NLN*, and between gene regions. *NR3C1* is known as nuclear receptor subfamily 3 group C member 1. This gene's function is to encode for glucocorticoid receptor and that can function as both a transcription factor that binds to glucocorticoid elements and acts as a regulator of other transcription factors. Insulin like growth factor one (*IGF1*) is known a protein coding gene that is a member of the family of proteins for mediating growth and development. This gene is similar to insulin and its functions. *TNFSF11* is known as *TNF* superfamily member 11. *TNFSF11* is a gene in the tumor necrosis factor cytokine family and is a dendritic cell survivor factor and is involved in T cell dependent immune response regulation. *MYCBP2* is known as MYC binding protein 2, E3 ubiquitin protein ligase. This gene encodes for the E3 ubiquitin-protein ligase and plays a role in axon guidance and synapse formation in the nervous system as it develops. NLN is known as neurolysin and encodes for a member of the metallopeptidase *M3* protein family. This gene helps with cleaving neurotensin at the Pro10-Tyr11 bond which leads to the formation of neurotensin (1-10) and neurotensin (11-13).

This variant calling software found 18 Significant SNPs. 5 of the SNPs were not within genes but rather in intergenic regions or upstream from a gene. The rest of the

SNPs were found within genes including a few from the original 8 candidate genes. The location, gene and p-value can be observed in table 10. No insertions or deletions were discovered in this part of the study. A basic GWAS was completed for the eighty-two heifers sampled and the Manhattan plot can be observed in figure 2.10.

Discussion

The forthcoming section will provide a brief review of the genetic markers identified through the research project and contextualize the significance of these given that all of these represent the first polymorphisms identified with significant association with the maturity phenotype in beef cattle. All of these SNPs have been validated, observed and reported on the Ensembl Genome Browser 96 from other studies. However, this study is the first to link any of these variations with the carcass maturity phenotype.

It is reported that within the population of heifers used in this research and within the candidate gene; alkaline phosphatase (*ALPL*) gene there are three identified significant SNPs associated with carcass maturity grade. These SNPs are located at positions 2:131,827,143, 2:131,837,201 and 2:131,837,038 in intron 1 and passed the association test using the 1v2 (A versus BC) phenotype sheet. The SNP at 2:131,827,143 has the most significant p-value ($p = 0.014$) followed by the SNP located at position 2:131,837,201 ($p = 0.015$), and finally the SNP at 2:131,837,038 has a reported p-value equal to 0.021. The location of these significant SNPs in the *ALPL* can be observed in figure 2.1.

The SNP at position 2:131,837,201 represents a nucleotide change from the reference thymine (T) reported in the bovine genome (Bos_taurus_UMD_3.1.1) to an alternate allele containing a cytosine (C) substitution at this position. More specifically, it was observed that animals that graded as C-maturity generally have the T-allele similar to the reference genome animal and in contrast A-maturity more frequently have the alternate C-allele. This is a novel discovery as it's the first ever report of a specific nucleotide variation in this gene linked to carcass maturity. This could be a marker for carcass maturity in future research to validate this association across a larger population of beef heifers and perhaps even steers. Most importantly, novel discoveries such as this empower producers with the ability to screen for or possibly select for the more desirable allele in hopes of yielding the more valued A-maturity grades more often.

The SNP identified at 2:131,827,143 is a single nucleotide change from guanine (G) to adenine (A). At this particular SNP location, the major allele is the reference allele G. The A maturity animals more commonly have the reference allele G within their population whereas the B&C maturity animals have the alternate allele A within their population. Once again, report of this SNP and its specific association with the less desired maturity phenotype is a novel discovery. Similarly, the utility of this discovery is that it could become another genetic marker for carcass maturity trait and in future be incorporated into genetic screening tools to help mitigate the rising incidence of advanced maturity in chronologically (<30 month) beef cattle.

The final SNP identified within the *ALPL* gene located at 2:131,837, 038.is characterized by a nucleotide change from the reference adenine (A) to an alternate guanine (G). With this SNP, the less desired maturity phenotypes B&C tended to have

more of the reference allele A within its population whereas the A maturity animals have more of the alternate allele G within its population. Hence, the more desired carcass maturity grade heifers had the alternate allele as compared to the less desired maturities. This has the potential to lead to a genetic screening assay that tests animals for the reference allele and may even lead to production changes for those animals shown to possess the allele type linked to a higher risk for advanced carcass maturity, so that they may still grade in the more valuable A maturity category. This represents yet another SNP that was previously known but had not yet been linked to any economically important trait such as advanced maturity phenotype, until now.

For the same reasons that the *ALPL* gene was justifiably selected for screening the advanced maturity heifers, it is not completely surprising that the researchers did in fact identify novel genetic markers linked to this phenotype. It is possible that given the fact that *ALPL* plays a role in osteogenesis and especially the rate of mineralization (deposition of osteoid) that these specific sequence changes, though located within introns of the gene, may be linked to differences in mRNA splicing functional expression levels of the cognate protein, though to be clear the study did not examine causation only association. However, these observations may be important to producers because an assay could be created to identify the SNPs in young cattle and allow for the producer to modify production steps for those particular animals to reduce risk of economic loss associated with advanced maturity. In research done by Moore et al. (2012), 7.2% of carcasses produced by fed steers and heifers are classified as B maturity or older based on USDA carcass maturity indicators. With the variants identified with this project can lead to the genetic assays that allow the producers for instance, be able to modify which

implants are selected and used or possibly finish animals with the advanced maturity marker at a younger age towards maintaining their A maturity classification at harvest. However, this will require additional research in a larger pool of beef cattle to validate the associations that have been observed, and also additional research to test for interactions of genotype, maturity grade and implant protocols, and/or other production practices. Finally, once validated these genetic markers can also be used to reduce incidence of advanced maturity if selection protocols are utilized to increase desired alleles within the population, meaning the alleles that are associated with better chances of A maturity when harvested under 30 months of age.

Within the Parathyroid hormone like protein (*PTHLH*) gene the initial analyses revealed 23 SNPs after excluding SNPs that failed to meet the minimum read depth stringency requirements. Within this gene one polymorphism was identified as significant following the statistical maturity grade association tests. The variant detected is located at position 5:82,249,219 in intron 3 and rather than being a simple nucleotide substitution, rather is a deletion of the reference (Bos_taurus_UMD_3.1.1) CC. This variant was detected using the basic allele association test where the A versus BC maturity phenotype were statistically evaluated. At this particular location animals yielding carcasses that were categorized in the A maturity grade have more of the reference alleles CC whereas the B and C maturity animals have more of the alternate allelic genotype where the deletion is present. These results, along with the previously reported SNPs will become more useful for creating carcass maturity screening assays to test larger populations of cattle once they are validated across a larger more representative pool of beef heifers. Another SNP within the *PTHLH* genes was identified using the GATK variant calling

software. This SNP is also located in intron 3 at position 5:82,248,608 with the reference allele being a C and the alternate allele being a T. The A maturity animals has more of the reference allele C compared to the less desired B and C maturities. With *PTHLH* playing a vital role in the activity of osteoblasts in ossification and the sensitivity of these cells to regulation by parathyroid hormone, these findings are important towards directing future carcass maturity research especially that research aiming towards the creation and validation of accurate genetic screening assays for carcass maturity traits.

Through the study design and analyses of targeted sequencing in the population of 90 heifers the researchers report the discovery of forty-seven unique polymorphisms within the estrogen receptor 1 (*ESR1*) gene which can be reviewed on table 5. Twenty-four of the significant SNPs were identified as significantly associated with carcass maturity using the A versus B versus C phenotype sheet on the basic allele association test. Ten SNPs were identified with the A versus B versus C phenotype on the recessive association test. Twelve significant SNPs were identified on the basic allele association test with the A versus B&C phenotype. Finally, on the recessive association test, one SNPs was identified as significant with the A versus B&C phenotype. In figure 5A and 5B all 47 SNPs can be observed at the particular locations within the *ESR1* gene. During the linkage disequilibrium (LD) test for Estrogen Receptor 1, it was observed that 25 of the 47 SNPS identified were in haplotype blocks. These results are key to helping limit the number of SNPs tests, allowing for a reduced number of SNPs to be screened since some are inherited in a linked manner rather than independently and at random.

Out of the twenty-four SNPS identified with the A versus B versus C phenotype sheet on the basic allele association test, thirteen of the SNPs are associated with the

advanced maturity phenotype C based on the major allele frequency. Twelve of those thirteen SNPs have the reference genome (Bos_taurus_UMD_3.1.1) alleles associated with the C maturity animals whereas the last SNP in that group of thirteen has the alternate genome (Bos_taurus_UMD_3.1.1) allele associated with the C maturity animals. The last eleven SNPs have the same nucleotide sequence as the reference genome (Bos_taurus_UMD_3.1.1) alleles associated with the A maturity animals. With this information assays could be created around these significant polymorphisms to identify at risk animals for either the advanced maturity trait and change their production protocols or identify animals with the desired A maturity and continue their production methods. After analysis of linkage disequilibrium, the researchers report that 10 of the SNPs identified using the A versus B versus C phenotype comparison on the basic allele association test were spread between 6 haplotype blocks. Haplotypes are sets of alleles linked at a certain locus on a chromosome. The importance of these results and findings in the haplotype blocks is that SNPs grouped in each block are in linkage disequilibrium (LD) meaning that when one SNP is showing a particular allele then those other SNPS in LD with it are linked and also inherited a particular set of alleles. Thus, SNPs discovered to be in high LD are not randomly inherited but rather are usually linked therefore allowing for only one of the SNPs with the strongest LD to be used to create a genetic marker. This still requires that the assay is tested and validates in a larger population with the knowledge when the one allele is inherited at that position then alleles in LD with the SNP will have their particular alleles inherited in a predictable manner. This ultimately will create a more efficient and less expensive experimental design when starting to test larger populations for these advanced maturity markers.

The results of the recessive association test with the A versus B versus C maturity phenotype sheet confirmed 10 significant polymorphisms. Six of the significant SNPs identified have the same genotype as the reference genome (Bos_taurus_UMD_3.1.1) alleles associated with the A maturity animals whereas the other four significant SNPs identified have the reference genome (Bos_taurus_UMD_3.1.1) alleles associated with the less desired C maturity animals. Five of these significant SNPs were also identified within haplotype blocks in the estrogen receptor 1 gene. This is another important discovery because these blocks show which sets of alleles are in LD and can be used to reduce the number of polymorphisms included as genetic markers of advanced maturity on future genetic screening assays. Importantly, knowledge of which SNPs are linked to the "at risk or less desired maturity" can be used to create genetic markers in an assay that tests a larger population and if validated can then be used to create a genetic test to identify the at-risk animals for potential production changes or selection against within the herd. This will help producers in the future locate and identify animals that are at risk of being graded as advanced maturity even though they are chronologically young animals.

Furthermore, within the estrogen receptor gene, twelve significant SNPs were identified through implementation of the basic allele association test with the A versus BC maturity phenotype sheet. Furthermore, eleven of these SNPs identified are such that the reference genome allele (Bos_taurus_UMD_3.1.1) are the ones associated with the less desired B&C combined maturity animals. The other SNP identified has the alternate allele associated with the less desired B&C combined maturity animals or put another way the allele type of the reference genome animal is associated with the A maturity

carcass grade. Also, and not surprisingly, with all these SNPs residing within intron 9 of the ESR1 gene, nine of them are within haplotype blocks and therefore in linkage disequilibrium (LD) with one another. As mentioned in the previous sections, discovering SNPs that are in LD allows for limiting the need to have all of these included in a future genetic screening assay for advanced maturity risk in the cattle population. These results can be used to create assays and direct further research into the causation of advanced maturity traits and development of feasible mitigation strategies.

Finally, the recessive association test using the A versus B&C maturity phenotype sheet there was one additional SNP identified. For this particular SNP, the allele reported in the reference genome (Bos_taurus_UMD_3.1.1) is associated with the less desired B&C combined maturity animals. On top of that, this SNP was also identified within the haplotype block 6. This is a valuable find because its linkage with other SNPs will help create a genetic marker to identify larger populations of cattle that may still carry these SNPs associated with the advanced maturity.

In all mammals including cattle, estrogen increases the rate of skeletal ossification, and these effects are mediated directly through interaction with the estrogen receptors (Tatum, 2011). Therefore, the identification of the many SNPs within the estrogen receptor 1 gene is important because these variants are likely to validate across the larger population of animals based on the biological relevance further they will likely become useful in the context of examining mitigation strategies using genetic testing to reduce the losses associated with advanced carcass maturity grades. As animals age and reach puberty, the rate of ossification increases until replacement of cartilaginous templates is completely replaced as in the epiphyseal plates. With implants the rate of

closure at the epiphyseal plate increases. Acheson et al. (2014) did a study comparing the relationship between USDA carcass maturity attributes in steers and heifers that were under 30 MOA by dentition that showed that there were no differences in quality between the maturities and that the effectiveness of the current USDA standards in 2014 for maturity classification was challenged. This led to the eventual change to allow dentition in late 2017 however those animals that are over 30 months of age by dentition still are seen with the advanced rates of ossification. That is why the SNPs identified in this project still play a role in the grading of those over 30 month of age animals. Changes in these genes that are validated can lead to eventual changes in the production protocols for producers.

For instance, anabolic implants are used in the U.S. beef production industry to enhance growth rate and feed conversion in steers and heifers. In many commercial implants the primary active ingredients are either, estrogen, androgen or a combo of the two. Depending on the implant's dosage and the implant timing and protocol, the animal may be susceptible to the positive but also some negative effects of the particular implant composition and the effects on tissue such as bone. Implant research by Tatum (2011) states that its estrogen containing implants could greatly increase the risk of the animal producing a B maturity or a more advanced maturity grade. With knowledge of estrogen's impact on bone development, and the fact that the use of implants could be increasing the impact on bone ossification this provides an avenue for production practices to change to alleviate the increasing incidence of advanced maturities in the beef production heifers.

After unexpectedly pre-disposing cattle to faster ossification rates through genetic selection and then adding estrogenic implants into the mix, identifying the genetic changes in candidate genes such as estrogen receptor 1 is essential for changing production standards to ensure that the industry has a way to adapt and identify at risk animals and mitigate financial losses. With the 47 SNPs identified within the ESR1, assays could be developed to test larger populations for the trend and thus possibly leading to a genetic marker test that identifies cattle with the pre-disposed advanced maturity trait. This will allow for producers to change production methods including opting out the implants and sending the cattle to the processing plant sooner.

In the other candidate genes screened; estrogen receptor 2, collagen type X Alpha 1, Indian hedgehog, bone morphogenetic protein 2, growth hormone and myosin heavy chain 2A, no significant polymorphisms associated with maturity grade were identified through the characterization of this relatively small population of research heifers (n=90).

An additional challenge related to this research beyond the limited population size, some of these genes such as Indian hedgehog and bone morphogenetic protein 2 have very small mapping regions which made it more difficult to capture and amplify the region for sequencing. This resulted in limitations in minimal read depth across the gene and the inadequate sequence coverage. This ultimately results in a much smaller number of SNPs identified and called. Along with this being a smaller gene there was not enough confidence in the SNPs that were called which resulted in no significant SNPs called based this particular type of association test in the small population of heifers.

Other genes such as collagen type X alpha 1, estrogen receptor 2 and myosin heavy chain 2A that revealed no SNPs in this pilot study lacked read depth around and within

the gene region itself. This could be due to; a failure in library preparations of the gene itself, the RNA bait sets failing to capture enough region-specific sequences leading to very little SNPs calling and lack of sufficient confidence to meet the rigorous standards including false discovery rates. Also, the targeted probes may have been designed in regions with lower conservation of sequence causing greater failure rate. It is also possible that simply observations of any differences in this project small population of heifers utilized for this comparison were not detected.

As a direct example in the growth hormone gene, no SNPs were called because there was insufficient read depth resulting in the SVS software not "calling" any sequence variations within this projects pool of heifers. Despite lacking any significant outputs in this trial it is important to mention that these genes are likely still suitable candidate genes for future study as they do play roles in the ossification process and a larger population needs to be tested before determining if there are really no polymorphisms that are linked to advanced maturity within these genes.

The current research specifically utilized heifers in the three different maturities A, B and C since they tended to show advanced maturity despite their chronological age being verified as under 30 months by dentition. The heifers used received anabolic implants and were finished on a grain diet which was typical for fed-heifers in the United States and gave this project a suitable population to test. Overall the results of this study are important given that it identified the first genetic polymorphisms linked to the phenotype of advanced carcass maturity. Additionally, these results are essential for future research relating to testing a larger population for the continued association and

also development of genetic assays to screen their herds and identify specific animals at risk for advanced maturity under current production practices.

References

Acheson, R. J., D. R. Woerner, and J. D. Tatum. 2014. Effects of USDA carcass maturity on sensory attributes of beef produced by grain-finished steers and heifers classifies as less than 30 months old using dentition. J. Anim. Sci. 92:1792-1799. doi:10.2527/jas2013- 7553

Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2017). BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience,8*(2). doi:10.1093/gigascience/giy154

Karaplis, A. C., Luz, A., Glowacki, J., Bronson, R. T., Tybulewicz, V. L., Kronenberg, H. M., & Mulligan, R. C. (1994). Lethal skeletal dysplasia from targeted disruption of the parathyroid hormone-related peptide gene. *Genes & Development,8*(3), 277-289. doi:10.1101/gad.8.3.277

Li, H., and Durbin, R. (2009, July 15). Fast and accurate short read alignment with Burrows-Wheeler transform. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19451168

Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & Depristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research,20*(9), 1297-1303. doi:10.1101/gr.107524.110

Moore, M. C., G. D. Gray, D. S. Hale, C. R. Kerth, D. B. Griffin, J. W. Savell, C. R. Raines, K. E. Belk, D. R. Woerner, J. D. Tatum, J. L. Igo, D. L. VanOverbeke, G. G. Mafi, T. E. Lawrence, R. J. Delmore Jr., L. M. Christensen, S. D. Schackelford, D. A. King, T. L. Wheeler, L. R. Meadows, and M. E. O'Connor. 2012. National Beef Quality Audit-2011: In-plant survey of targeted carcass characteristics related to quality, quantity, value, and marketing of fed steers and heifers. J. Anim. Sci. 90:5143-5151.

National Center for Biotechnology Information. (n.d.). Retrieved 2019, from https://www.ncbi.nlm.nih.gov/

O'Connor, M. E., Ransom, J. R., & Feli, M. (2007, December). USDA PHYSIOLOGICAL MATURITY VALIDATION STUDY: Validating the Relationship between Chronological Age and Physiological Maturity in the U.S. Fed-Beef Population. Retrieved October 10, 2018, from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiwkJHJn_HiAhUDWq0KHb6VCvEQFjAAegQIABAC&url=http%3A%2F%2Fwww.fsc.go.jp%2Ffsciis%2FattachedFile%2Fdownload%3FretrievalId%3Dkai20080131sfc%26fileId%3D114&usg=AOvVaw0qxZm6e1z6LFq77SHEhqxc

Slatkin, M. (2008). Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics,9*(6), 477-485. doi:10.1038/nrg2361

SVS Golden Helix. (n.d.). Genotype Association Tests. Retrieved November 15, 2018, from http://goldenhelix.com/products/SNP_Variation/index.html

Tatum, J. D. 2011. Animal age, physiological maturity, and associated effects on beef tenderness. White Paper. Product Enhancement Research & Knowledge Management, NCBA.

USDA. (2014, November). Institutional Meat Purchase Specifications | Agricultural ... Retrieved December 2, 2018, from https://www.ams.usda.gov/sites/default/files/media/IMPS_100_Fresh_Beef[1].pdf

USDA. (2017). Agricultural Marketing Service Carcass Beef Livestock ... Retrieved October 30, 2018, from https://www.ams.usda.gov/sites/default/files/media/CarcassBeefStandard.pdf

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics,28*(24), 3326-3328. doi:10.1093/bioinformatics/bts606

Table 2.1: Targeted sequence capture gene regions: Genes captured during sequencing and used in SVS or GATK variant calling software describing chromosome, length and the gene's start/stop codon spans in the *Bos taurus* genome.

Gene	Chromosome location	Chr	Start	Stop	Length bp
ALPL	Chr2:131791668-131857687	2	131791668	131857687	66019
BMP2	Chr13:49550050-49561304	13	49550050	49561304	11254
CNP	Chr19:42772788-42780092	19	42772788	42780092	7304
COL10A1	Chr9:34961718-34994278	9	34961718	34994278	32560
ESR1	Chr9:89967417-90256185	9	89967417	90256185	288768
ESR2	Chr10:76706017-76771172	10	76706017	76771172	65155
GH	Chr19:48768617-48772049	19	48768617	48772049	3432
GPER	Chr25:42259188-42262393	25	42259188	42262393	3205
NR3C1	Chr7:56231970-56351929	7	56231970	56351929	119959
IGF1	Chr5:66523798-66604781	5	66523798	66604781	80983
IHH	Chr2:107722665-107728939	2	107722665	107728939	6274
PTHLH	Chr5:82241708-82258858	5	82241708	82258858	17150
TNFSF11	Chr12:12741069-12782474	12	12741069	12782474	41405
VDR	Chr5: 32501968-32609170	5	32501968	32609170	107202
VEGF	Chr23:17255515-17270515	23	17255515	17270515	15000
CYP19A1	Chr10:59227858-59282939	10	59227858	59282939	55081
MYH2	Chr19:30137767-30165109	19	30137767	30165109	27342

Table 2.2: Candidate gene chromosome and location: List of candidate target genes. Gene name, bovine chromosome where gene is positioned, and chromosome location is shown. Locations are based off of the most current bovine reference genome Bos_taurus_UMD_3.1.1.

Gene	Chromosome	Location
Indian Hedgehog	2	107,722,665bp – 107,728,939bp
Alkaline Phosphatase	2	131,791,668bp – 131,857,687bp
Parathyroid Hormone Like Protein	5	82,241,708bp – 82,258,858bp
Collagen type X, Alpha 1	9	34,961,718bp – 34,994,278bp
Estrogen Receptor 1	9	89,969,586bp – 90,255,801bp
Estrogen Receptor 2	10	76,706,017bp – 76,771,172bp
Bone Morphogenetic Protein 2	13	49,550,050bp – 49 561,304bp
Growth Hormone	19	48,768,617bp – 48,772,049bp
Sarcomeric myosin heavy chain 2a	19	30,137,767bp – 30,165,109bp

Table 2.3: Alkaline phosphatase significant variants: List of significant polymorphisms identified their location, allele change, p-value, and phenotype sheet identified on.

ALPL	Allele Change	Basic Test p-value	Recessive Test	Phenotype Sheet	Location
2:131827143	G/A	0.01348		A versus BC	Intron 1
2:131837201	T/C	0.01529		A versus BC	Intron 1
2:131837038	A/G	0.02128		A versus BC	Intron 1

Table 2.4: Alkaline phosphatase major allele frequency: This table shows the SNP, its location, the reference and alternate allele. Also, the major allele and its frequency along with what maturity it was located more in.

Gene	Location	Phenotype Sheet	Variant type	Ref	Alt	Major Allele	Minor Allele	Major allele frequency A Maturity	Major allele frequency B&C Maturity
ALPL	2:131827143	A versus BC	SNV	G	A	G	A	0.95	0.76
ALPL	2:131837201	A versus BC	SNV	T	C	T	C	0.43	0.65
ALPL	2:131837038	A versus BC	SNV	A	G	A	G	0.54	0.74

Table 2.5: Estrogen receptor 1 significant variants: List of significant polymorphisms identified their location, allele change, p-value, and phenotype sheet identified on.

ESR1	Allele Change	Basic Test	Recessive Test	Phenotype Sheet	Location
9:90115650	G/A	0.002668		A versus B versus C	Intron 9
9:90097671	G/A	0.003161		A versus B versus C	Intron 9
9:90107474	C/T	0.006379		A versus B versus C	Intron 9
9:90092551	G/A	0.010002		A versus B versus C	Intron 8
9:90103841	C/T	0.011263		A versus B versus C	Intron 9
9:90025573	C/T	0.011547		A versus B versus C	Intron 7
9:90103870	G/A	0.016352		A versus B versus C	Intron 9
9:90046635	C/T	0.022058		A versus B versus C	Intron 8
9:90129949	A/C	0.033117		A versus B versus C	Intron 9
9:90106608	G/A	0.034781		A versus B versus C	Intron 9
9:90015374	C/T	0.035643		A versus B versus C	Intron 7
9:90104171	C/T	0.036602		A versus B versus C	Intron 9
9:90104191	T/C	0.038845		A versus B versus C	Intron 9
9:90115555	T/G	0.039012		A versus B versus C	Intron 9
9:90103219	G/A	0.039689		A versus B versus C	Intron 9
9:90104212	C/T	0.046735		A versus B versus C	Intron 9
9:90109970	A/C	0.011847		A versus B versus C	Intron 9
9:90097772	G/A	0.020684		A versus B versus C	Intron 9
9:90097695	T/C	0.023184		A versus B versus C	Intron 9
9:90180362	C/T	0.003098		A versus B versus C	Intron 9
9:90178812	A/G	0.011715		A versus B versus C	Intron 9
9:90118413	A/G	0.011519		A versus B versus C	Intron 9
9:90118414	C/A	0.011519		A versus B versus C	Intron 9
9:90044201	T/A	0.024167		A versus B versus C	Intron 7
9:90015095	C/T		0.003204124	A versus B versus C	Intron 7
9:90115988	T/C		0.006552151	A versus B versus C	Intron 9
9:90062953	T/C		0.018668512	A versus B versus C	Intron 8
9:90120720	G/A		0.030446979	A versus B versus C	Intron 9
9:90118245	A/G		0.045367556	A versus B versus C	Intron 9
9:90110290	A/G		0.048469522	A versus B versus C	Intron 9

9:90015128	G/C		0.048591905	A versus B versus C	Intron 7
9:90109965	A/C		0.018727089	A versus B versus C	Intron 9
9:90116049	A/C		0.005172447	A versus B versus C	Intron 9
9:90116111	T/A		0.013631376	A versus B versus C	Intron 9
9:89990110	C/T	0.00813		A versus BC	Intron 5
9:90103736	C/T	0.00904		A versus BC	Intron 9
9:90245334	C/A	0.0101		A versus BC	Intron 11
9:90103589	G/A	0.01465		A versus BC	Intron 9
9:90103593	C/G	0.01465		A versus BC	Intron 9
9:90104016	T/C	0.0194		A versus BC	Intron 9
9:90103618	A/C	0.02677		A versus BC	Intron 9
9:90103620	A/G	0.02677		A versus BC	Intron 9
9:90103630	A/G	0.02677		A versus BC	Intron 9
9:90103683	C/T	0.02677		A versus BC	Intron 9
9:90098969	A/G	0.0408		A versus BC	Intron 9
9:90103583	A/G	0.04587		A versus BC	Intron 9
9:90097591	T/C		0.03936	A versus BC	Intron 9

Table 2.6: Major allele frequency of SNPs identified on the basic association test with the A versus B versus C maturity phenotype in *ESR1:* List of significant SNPs identified in the basic allele association test with the major and minor allele for each. The major allele frequency for each phenotype group is identified above along with what trend was noticed.

Gene	Location	Phenotype Sheet	Variant type	Ref	Alt	Major Allele	Minor Allele	Major AF A Maturity	Major AF B Maturity	Major AF C Maturity	Observation
ESR1	9:90115650	AvBvC	SNP	G	A	G	A	0.75	0.52	0.46	More ref in A maturity
ESR1	9:90097671	AvBvC	SNP	G	A	G	A	0.75	0.84	0.98	More ref in C maturity
ESR1	9:90107474	AvBvC	SNP	C	T	C	T	0.38	0.52	**0.63**	More ref in C maturity
ESR1	9:90092551	AvBvC	SNP	G	A	G	A	**1**	0.95	0.87	More ref in A maturity
ESR1	9:90103841	AvBvC	SNP	C	T	C	T	0.73	0.84	0.94	More ref in C maturity
ESR1	9:90025573	AvBvC	SNP	C	T	C	T	1	0.92	0.88	More ref in A maturity
ESR1	9:90103870	AvBvC	SNP	G	A	G	A	0.74	0.84	0.94	More ref in C maturity
ESR1	9:90046635	AvBvC	SNP	C	T	C	T	0.98	0.95	0.84	More ref in A maturity
ESR1	9:90129949	AvBvC	SNP	A	C	A	C	0.7	0.67	0.88	More ref in C maturity
ESR1	9:90106608	AvBvC	SNP	G	A	G	A	0.64	0.5	0.44	More ref in A maturity
ESR1	9:90015374	AvBvC	SNP	C	T	T	C	0.57	0.59	0.77	More ref in C maturity
ESR1	9:90104171	AvBvC	SNP	C	T	C	T	0.76	0.85	0.93	More ref in C maturity
ESR1	9:90104191	AvBvC	SNP	T	C	T	C	0.76	0.83	0.93	More ref in C maturity
ESR1	9:90115555	AvBvC	SNP	T	G	T	G	0.64	0.57	0.44	More ref in A maturity
ESR1	9:90103219	AvBvC	SNP	G	A	G	A	0.76	0.85	0.92	More ref in C maturity
ESR1	9:90104212	AvBvC	SNP	C	T	C	T	0.77	0.87	0.94	More ref in C maturity
ESR1	9:90109970	AvBvC	SNP	A	C	A	C	0.7	0.55	0.46	More ref in A maturity
ESR1	9:90097772	AvBvC	SNP	G	A	G	A	0.78	0.82	0.96	More ref in C maturity
ESR1	9:90097695	AvBvC	SNP	T	C	T	C	0.83	0.84	0.98	More ref in C maturity
ESR1	9:90180362	AvBvC	SNP	C	T	C	T	0.81	0.7	0.56	More ref in A maturity
ESR1	9:90178812	AvBvC	SNP	A	G	A	G	0.77	0.79	0.56	More alt in C maturity
ESR1	9:90118413	AvBvC	SNP	A	G	A	G	0.93	0.87	0.74	More ref in A maturity
ESR1	9:90118414	AvBvC	SNP	C	A	C	A	0.93	0.87	0.74	More ref in A maturity
ESR1	9:90044201	AvBvC	SNP	T	A	T	A	0.98	0.93	0.83	More ref in A maturity

Table 2.7: Major allele frequency of SNPs identified on the recessive association test with the A versus B versus C maturity phenotype in *ESR1:* List of significant SNPs identified in the recessive association test with the major and minor allele for each. The major allele frequency for each phenotype group is identified above along with what trend was noticed.

Gene	Location	Phenotype Sheet	Variant type	Ref	Alt	Major Allele	Minor Allele	Major AF A Maturity	Major AF B Maturity	Major AF C Maturity	Observation
ESR1	9:90015095	AvBvC	SNP	C	T	C	T	0.67	0.58	0.48	More ref in A maturity
ESR1	9:90115988	AvBvC	SNP	T	C	T	C	0.63	0.62	0.42	More ref in A maturity
ESR1	9:90062953	AvBvC	SNP	T	C	T	C	0.9	0.93	0.76	More ref in A maturity
ESR1	9:90120720	AvBvC	SNP	G	A	G	A	0.8	0.75	0.67	More ref in A maturity
ESR1	9:90118245	AvBvC	SNP	A	G	A	G	0.93	0.84	0.79	More ref in A maturity
ESR1	9:90110290	AvBvC	SNP	A	G	A	G	0.6	0.56	0.44	More ref in A maturity
ESR1	9:90015128	AvBvC	SNP	G	C	G	C	0.58	0.72	0.61	More ref in C maturity
ESR1	9:90109965	AvBvC	SNP	A	C	A	C	0.5	0.55	0.72	More ref in C maturity
ESR1	9:90116049	AvBvC	SNP	A	C	A	C	0.46	0.58	0.68	More ref in C maturity
ESR1	9:90116111	AvBvC	SNP	T	A	T	A	0.46	0.52	0.65	More ref in C maturity

Table 2.8: Major allele frequency of SNPs identified on the basic association test with the A versus B&C maturity phenotype in *ESR1:* List of significant SNPs identified in the basic allele association test with the A versus B&C maturity showing the major and minor allele for each. The major allele frequency for each phenotype group is identified above along with what trend was noticed.

Gene	Location	Phenotype Sheet	Variant type	Ref	Alt	Major Allele	Minor Allele	Major AF A Maturity	Major AF B&C Maturity	Observation
ESR1	9:89990110	AvBC	SNP	C	T	T	C	0.45	0.7	More Alt in less desired (B&C) maturity
ESR1	9:90103736	AvBC	SNP	C	T	C	T	0.75	0.92	More Ref in less desired (B&C) maturity
ESR1	9:90245334	AvBC	SNP	C	A	C	A	0.67	0.86	More Ref in less desired (B&C) maturity
ESR1	9:90103589	AvBC	SNP	G	A	G	A	0.75	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90103593	AvBC	SNP	C	G	C	G	0.75	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90104016	AvBC	SNP	T	C	T	C	0.73	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90103618	AvBC	SNP	A	C	A	C	0.76	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90103620	AvBC	SNP	A	G	A	G	0.76	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90103630	AvBC	SNP	A	G	A	G	0.76	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90103683	AvBC	SNP	C	T	C	T	0.76	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90098969	AvBC	SNP	A	G	A	G	0.77	0.9	More Ref in less desired (B&C) maturity
ESR1	9:90103583	AvBC	SNP	A	G	A	G	0.77	0.9	More Ref in less desired (B&C) maturity

Table 2.9 Major allele frequency of SNPs identified on the recessive association test with the A versus B&C maturity phenotype in *ESR1:* List of significant SNPs identified in the recessive association test with the A versus B&C maturity showing the major and minor allele for each. The major allele frequency for each phenotype group is identified above along with what trend was noticed.

Gene	Location	Phenotype Sheet	Variant type	Ref	Alt	Major Allele	Minor Allele	Major AF A Maturity	Major AF B&C Maturity	Observation
ESR1	9:90097591	AvBC	SNP	T	C	T	C	0.77	0.89	More Ref in less desired (B&C) maturity

Table 2.10: Significant SNPs identified using GATK variant calling: All SNPs identified by GATK. This table shows the chromosome, start/end, p-value, gene and location for each SNP identified with this software.

Chr	Start	End	P-value	Gene	SNP location
9	26110323	26110323	0.00202	NA	Intergenic Variant
7	56321628	56321628	0.0119	NR3C1	Intron 2
7	56321643	56321643	0.0119	NR3C1	Intron 2
19	30160768	30160768	0.0131	MYH2	Intron 5
5	32549362	32549362	0.0196	NA	Upstream Gene Variant
5	66596330	66596330	0.0206	IGF1	Intron 2
12	12774820	12774820	0.0206	TNFSF11	Intron 2
12	12774775	12774775	0.027	TNFSF11	Intron 2
12	52641232	52641232	0.0283	MYCBP2	Intron 23
27	42673983	42673983	0.0312	NA	Intergenic Variant
12	12774516	12774516	0.0316	TNFSF11	Intron 2
20	13731671	13731671	0.0339	NLN	Intron 5
5	82248608	82248608	0.0379	PTHLH	Intron 3
X	59080285	59080285	0.0383	NA	Intergenic Variant
9	90078489	90078489	0.0392	ESR1	Intron 8
9	90104171	90104171	0.0418	ESR1	Intron 9
9	90097792	90097792	0.0463	ESR1	Intron 9
6	10542071	10542071	0.0471	NA	Intergenic Variant

Figure 1.1: Relationship between maturity, marbling and quality grade

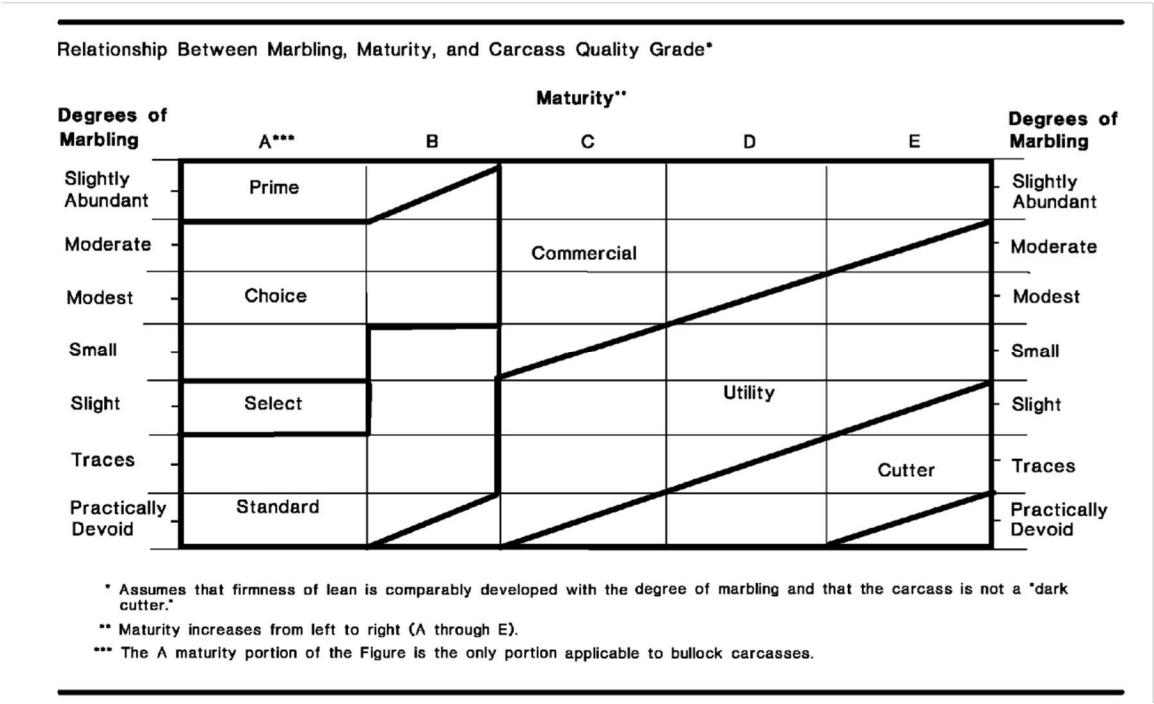
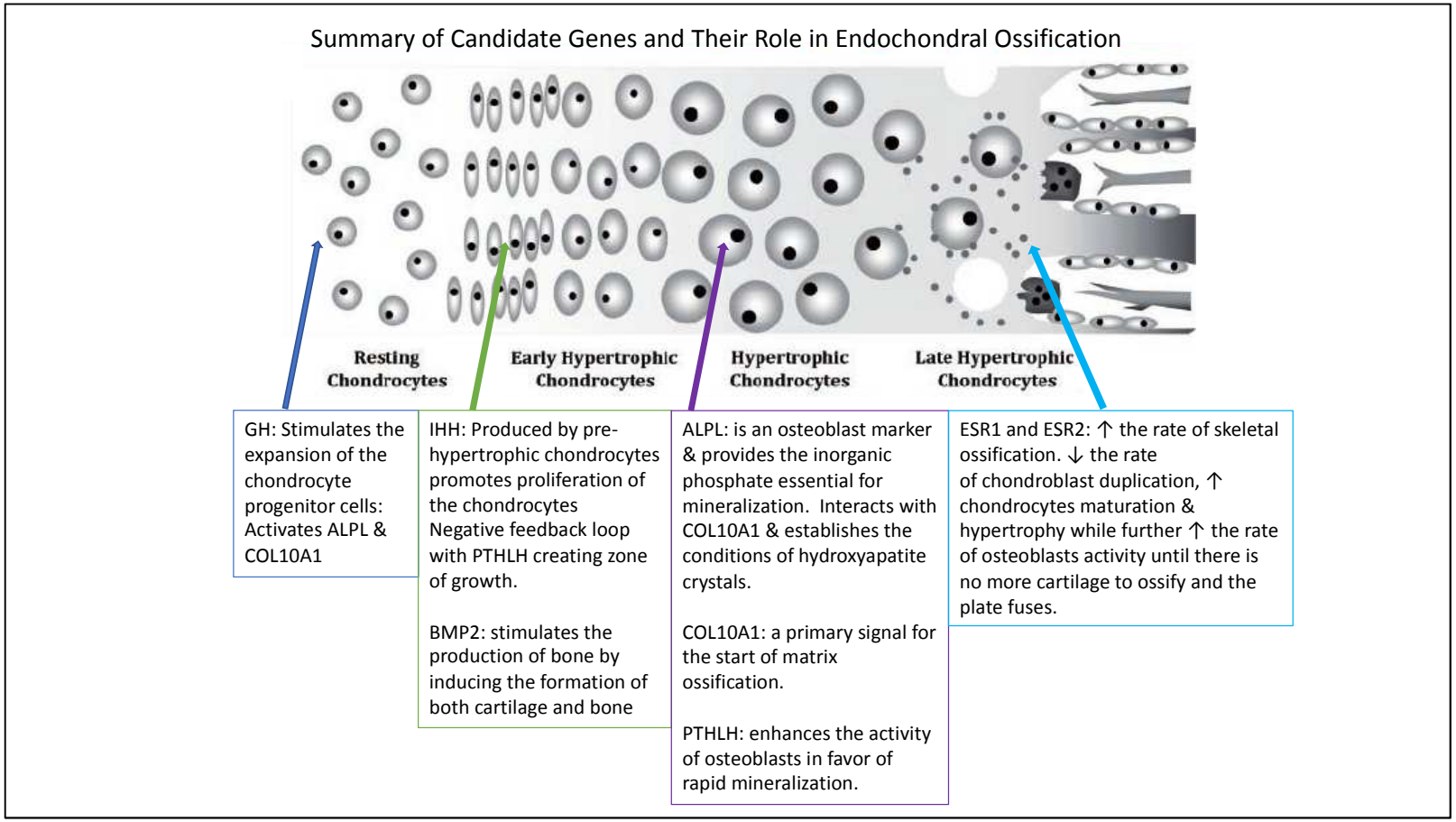


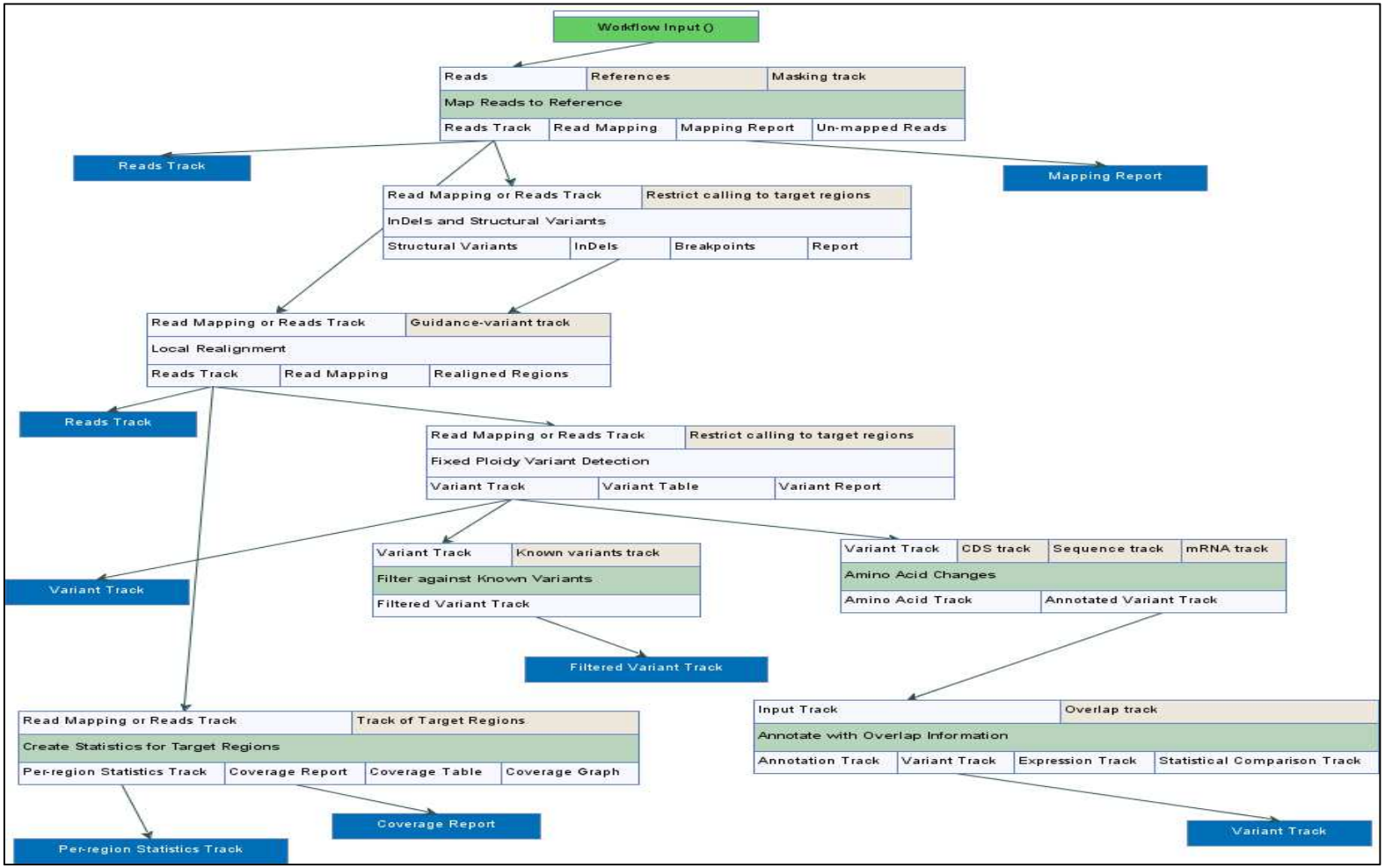
Photo courtesy of the American Meat Science Association. Reference figures not created from data obtained from the study.

Figure 1.2: Summary of candidate genes and their role in endochondral ossification



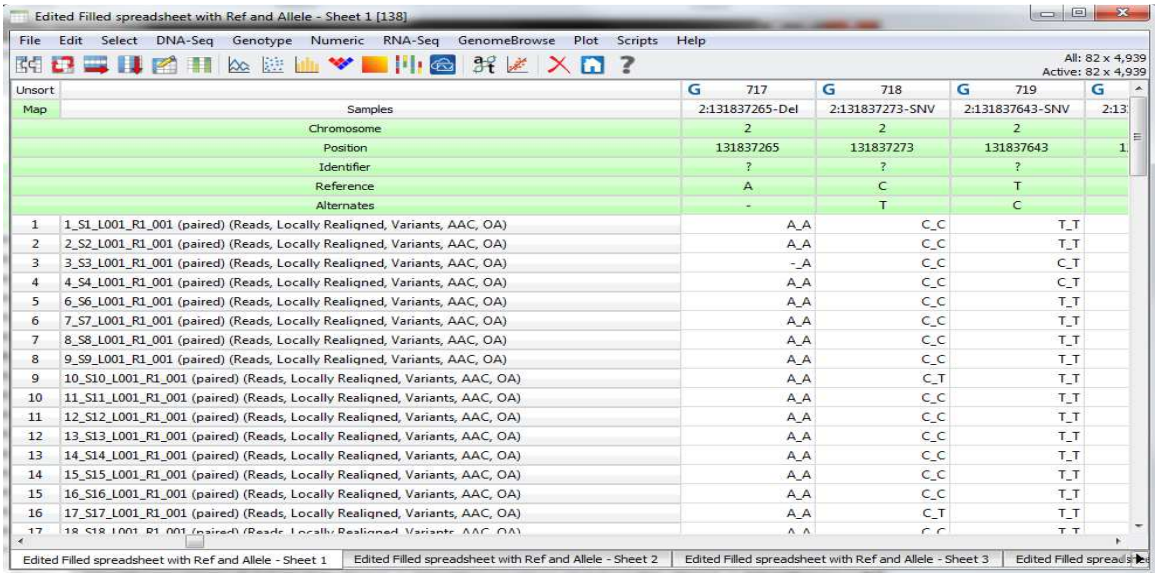
Adapted from "Modulating Endochondral Ossification of Multipotent Stromal Cells for Bone Regeneration," by Debby Gawlitta, August 2010, https://www.researchgate.net/publication/41406818_Modulating_Endochondral_Ossification_of_Multipotent_Stromal_Cells_for_Bone_Regeneration

Figure 2.1: CLC genomics bench workflow



The workflow for realigning the sequences received and the different parts to create the files for further analysis.

Figure 2.2: SVS spreadsheet



Edited Filled spreadsheet with Ref and Allele - Sheet 1 [138]

File Edit Select DNA-Seq Genotype Numeric RNA-Seq GenomeBrowse Plot Scripts Help

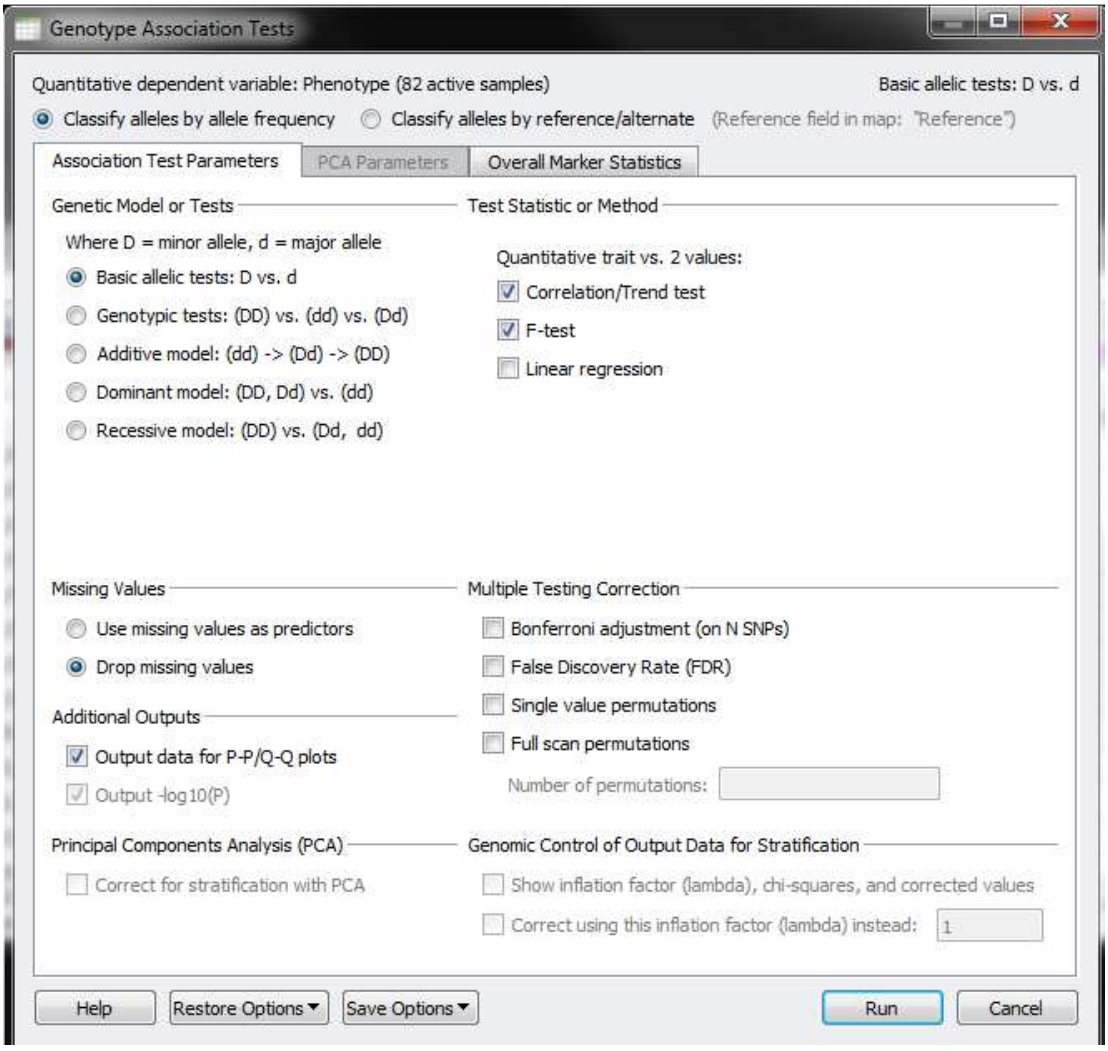
All: 82 x 4,939
Active: 82 x 4,939

Unsort		G 717	G 718	G 719	G
Map	Samples	2:131837265-Del	2:131837273-SNV	2:131837643-SNV	2:13
	Chromosome	2	2	2	
	Position	131837265	131837273	131837643	1
	Identifier	?	?	?	
	Reference	A	C	T	
	Alternates	-	T	C	
1	1_S1_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
2	2_S2_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
3	3_S3_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	-_A	C_C	C_T	
4	4_S4_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	C_T	
5	6_S6_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
6	7_S7_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
7	8_S8_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
8	9_S9_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
9	10_S10_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_T	T_T	
10	11_S11_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
11	12_S12_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
12	13_S13_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
13	14_S14_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
14	15_S15_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
15	16_S16_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	
16	17_S17_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_T	T_T	
17	18_S18_L001_R1_001 (paired) (Reads, Locally Realigned, Variants, AAC, OA)	A_A	C_C	T_T	

Edited Filled spreadsheet with Ref and Allele - Sheet 1 | Edited Filled spreadsheet with Ref and Allele - Sheet 2 | Edited Filled spreadsheet with Ref and Allele - Sheet 3 | Edited Filled spreadsh

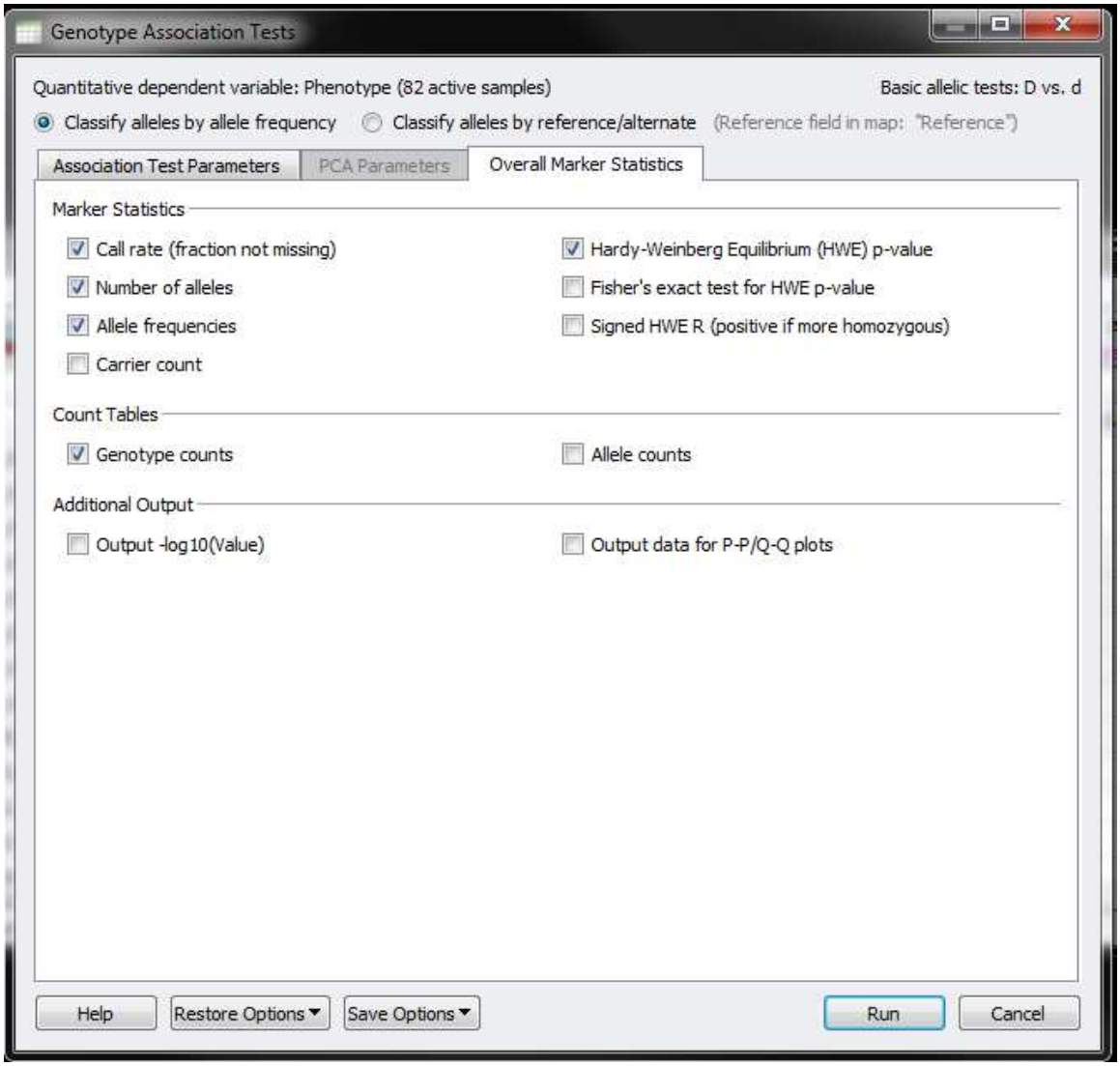
Spreadsheet format in SVS showing the sample ID, chromosome, SNP location, reference/alternate allele and the genotype of the sample.

Figure 2.3: SVS genotype association test window



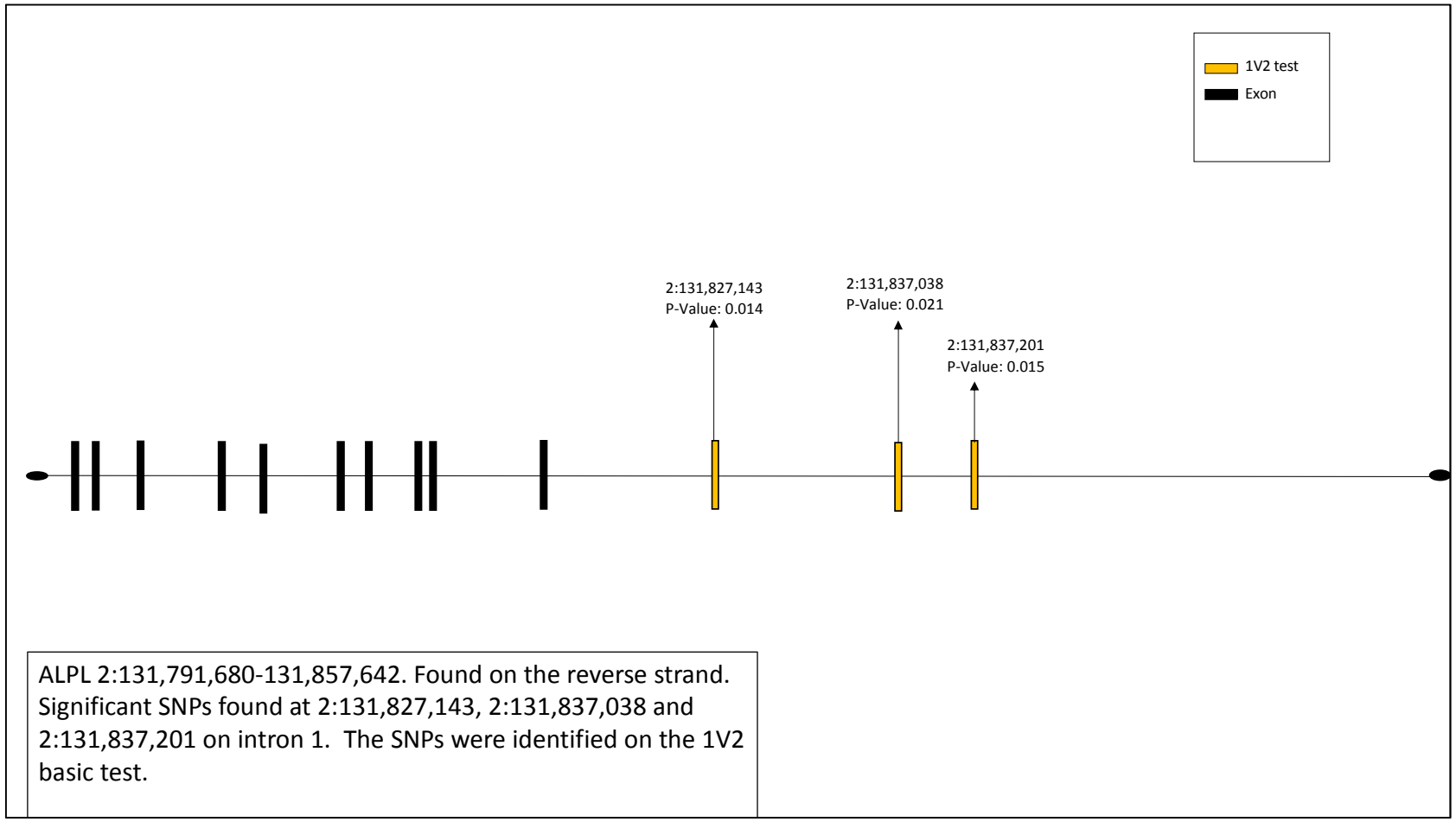
Control window for genotype association tests allowing for the selection of type of test parameters.

Figure 2.4 SVS genotype association test window



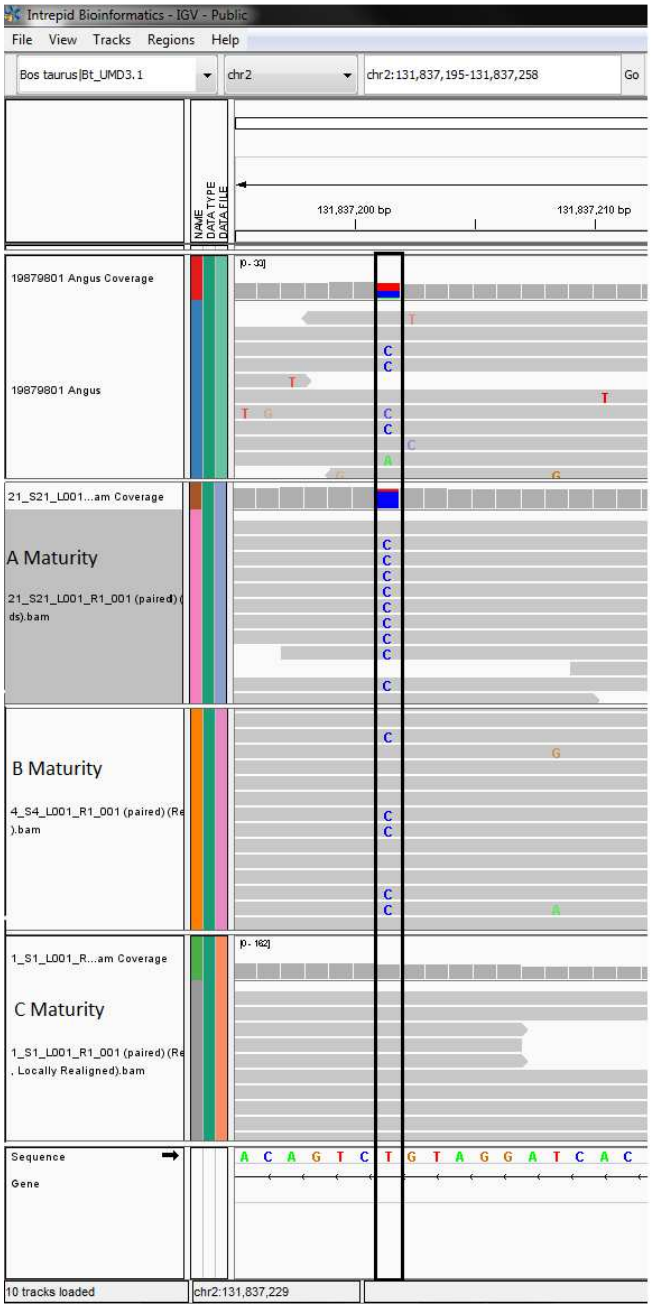
Control window for genotype association tests allowing for the selection of type of test parameters and the overall marker statistics that can be selected.

Figure 2.5: Representative image of the significant SNP locations *ALPL* gene



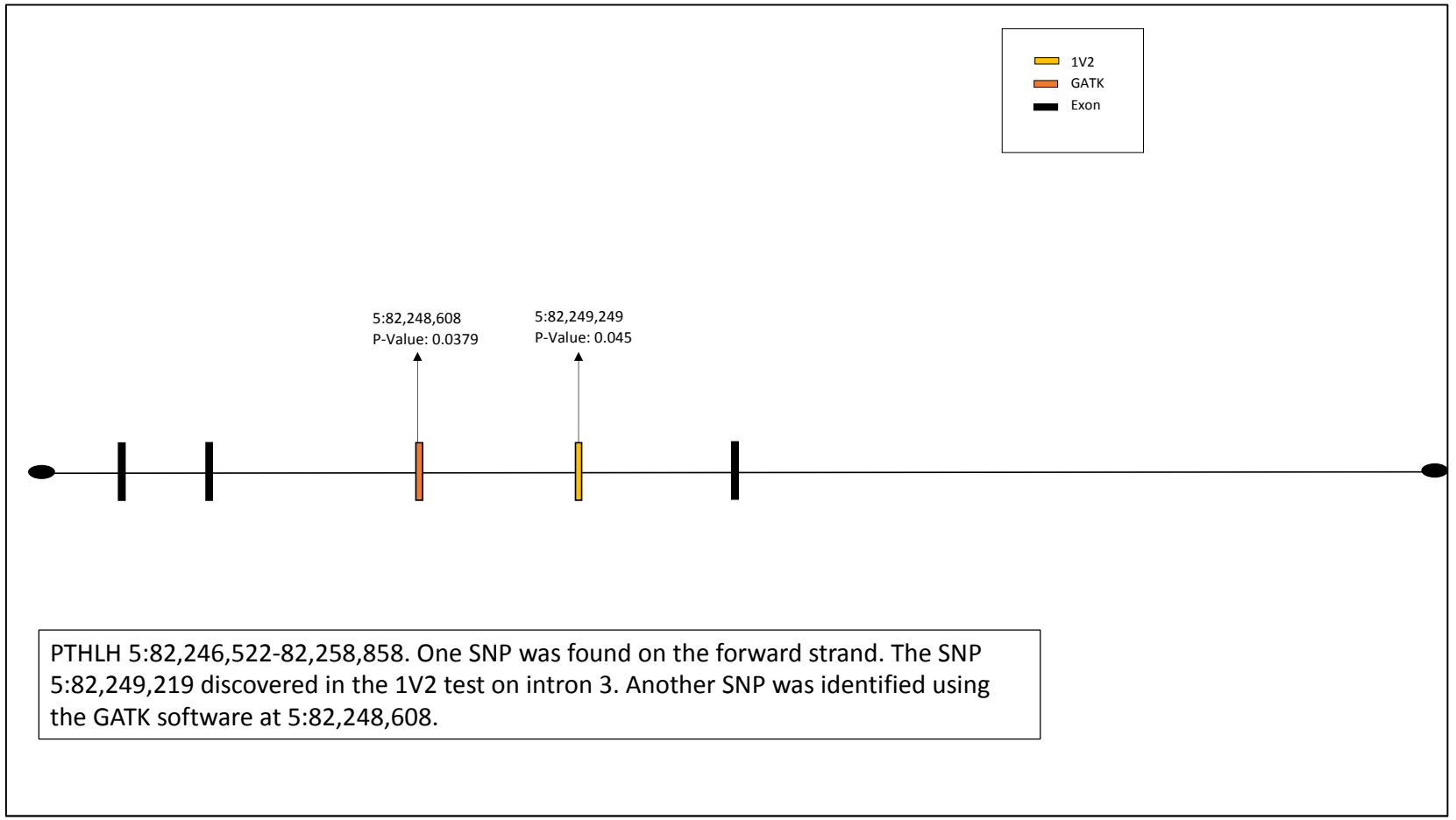
Alkaline Phosphatase gene from the 5'to 3', from start codon to end. Demonstrates the number of introns and exons in the gene and where the significant SNPs are located at.

Figure 2.6: Genotype of alkaline phosphatase at position 2:131,837,201bp



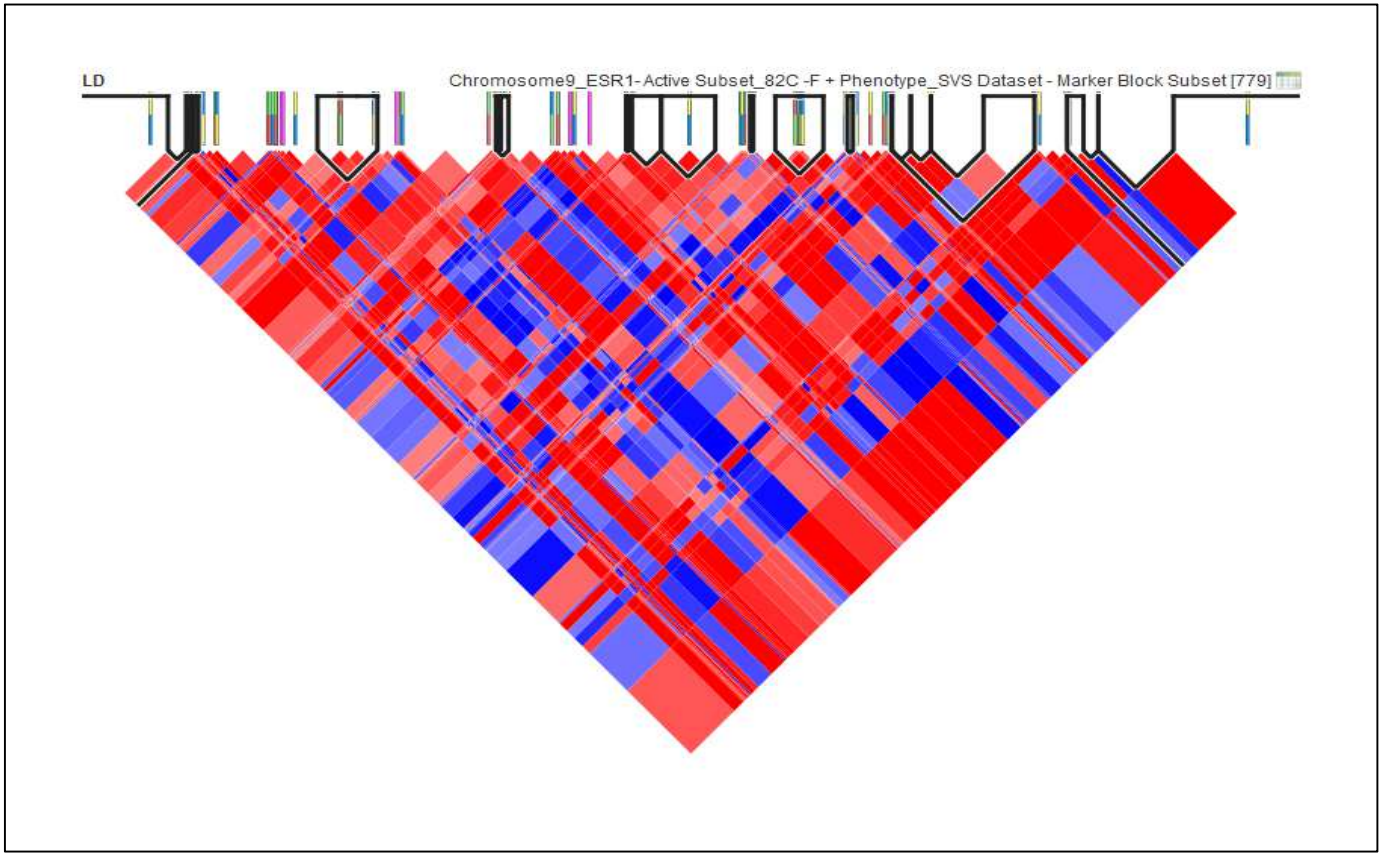
This figure exhibits a sequence alignment from a reference Angus individual in the top panel, and comparison of three animals from thus study with the different genotypes.

Figure 2.7: Parathyroid hormone like protein significant SNPs location



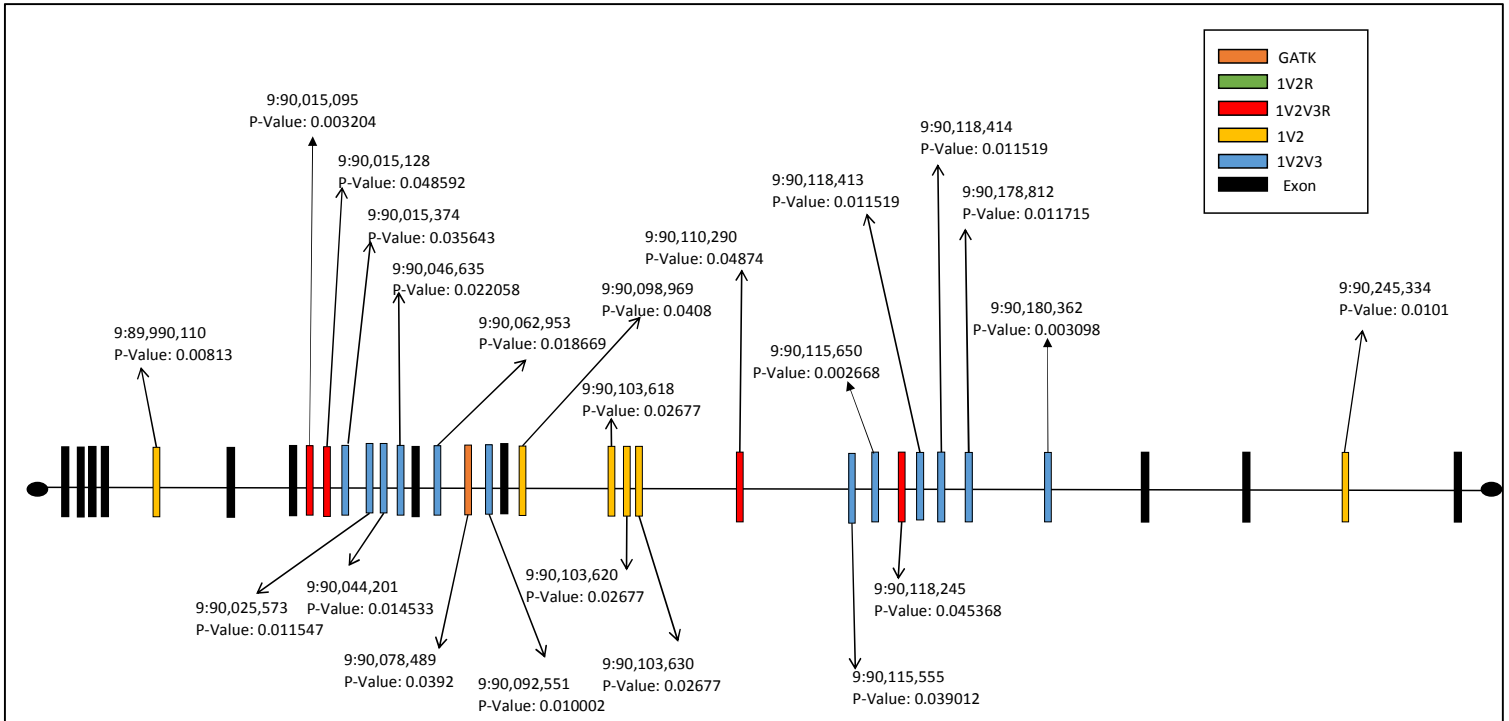
PTHLH gene running from the 5'to 3', from start codon to end. Demonstrates the number of introns and exons in the gene and where the significant SNPs are located at.

Figure 2.8: Estrogen receptor 1 linkage disequilibrium map



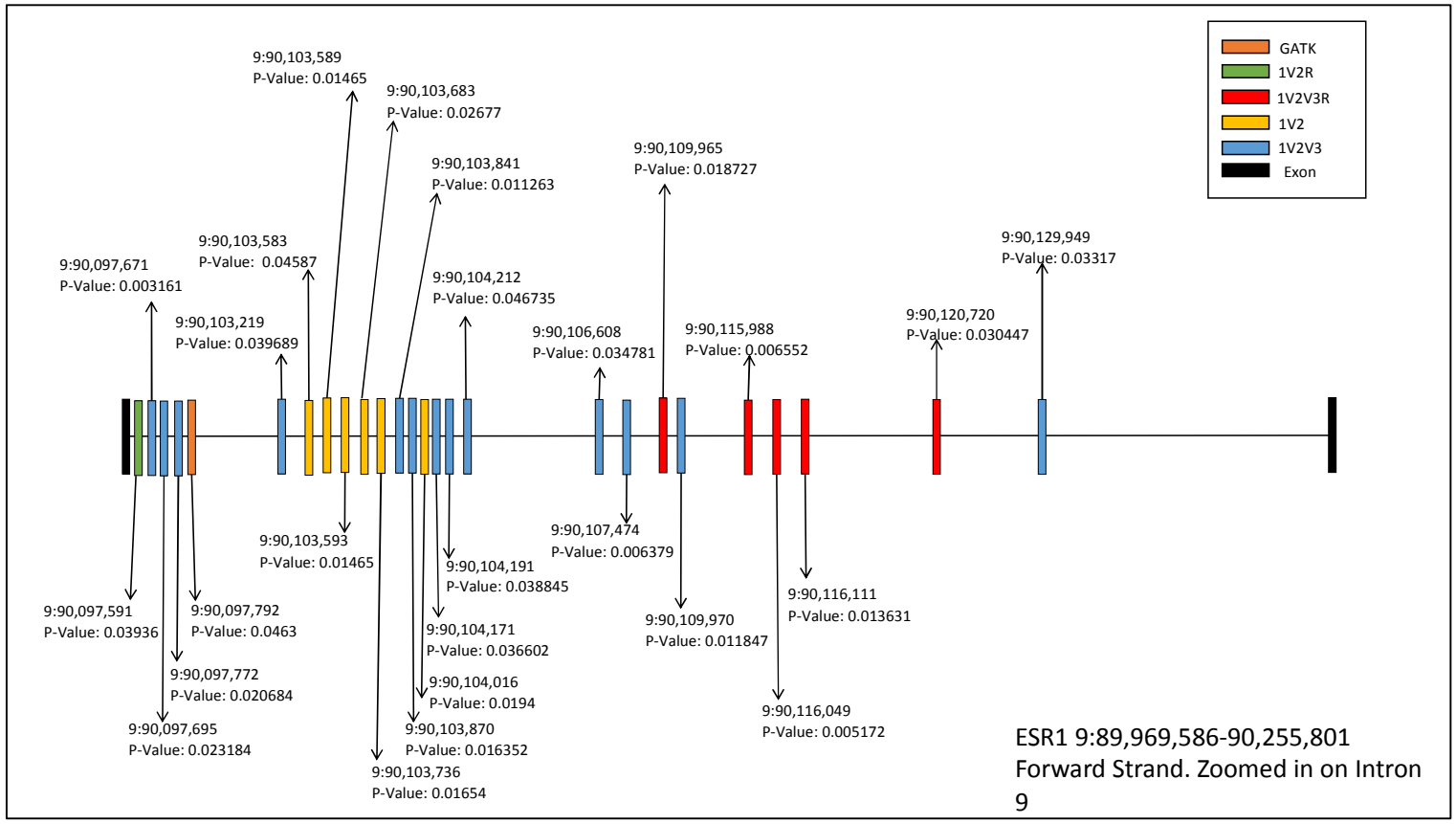
Linkage disequilibrium map of *ESR1*. The black boxes represent the haplotype blocks and the varying degrees of red are areas in high LD whereas the blue areas represent very low LD. This map shows the overall looks at the non-random association of alleles at multiple loci within a population and do not occur randomly with respect to each other.

Figure 2.9A: Estrogen receptor 1 significant SNP locations



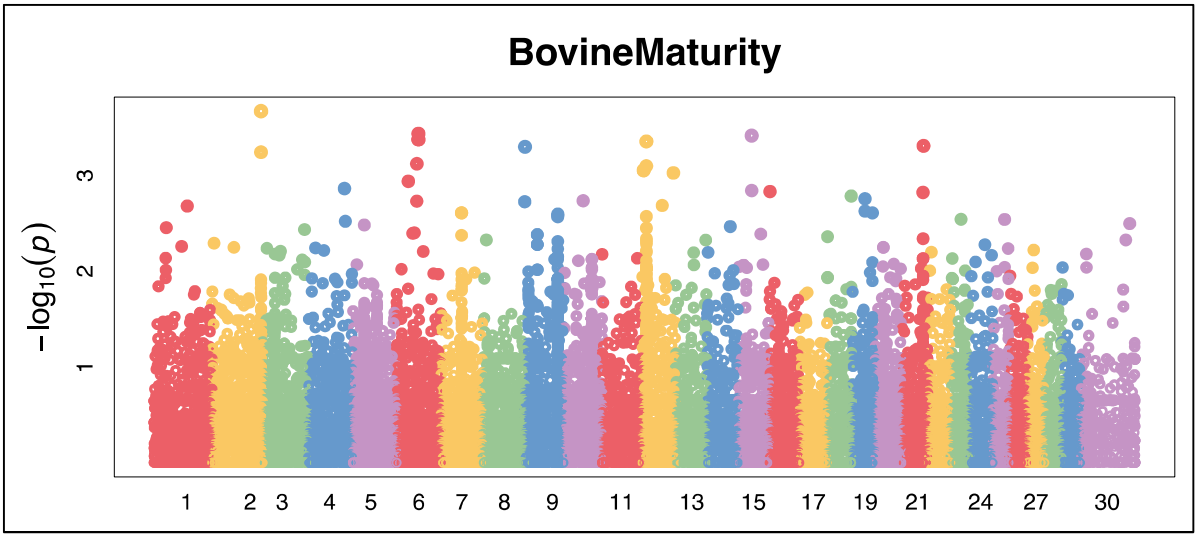
ESR1 from the 5'to 3', from start codon to end. Demonstrates the number of introns and exons in the gene and where the significant SNPs are located at.

Figure 2.9B: ESR1 significant SNP locations within Intron 9



ESR1 from the 5'to 3', from start codon to end. Demonstrates the number of introns and exons in the gene and where the significant SNPs are located at.

Figure 2.10: Manhattan Plot



Demonstrates a Manhattan plot created by the GATK software showing significant polymorphisms found across the sample population for each chromosome. X-axis represents the chromosome number.

Appendix A

Preparation of Genomic DNA from Mammalian Tissue

Materials
Liquid Nitrogen
Digestion Buffer
Phosphate-buffered saline (PBS), Ice cold
25:24:1 Phenol/chloroform/isoamyl alcohol
7.5 M ammonium acetate
TE buffer, pH 8
Sorvall H1000B rotor or equivalent

Prepare the Cells
Beginning with whole tissue

1. Take small tissue sample and quickly freeze in liquid nitrogen.

2. Grind 200mg to 1g tissue with pre-chilled mortar and pestle, or crush with hammer to fine powder. Suspend in 1.2mL digestion buffer per 100mg of tissue.

Lyse cells and extract DNA

3. Incubate samples overnight (12-18 hours) at 50°C

4. Extract samples with an equal volume of phenol/chloroform/isoamylalcohol. Centrifuge 10 min at 1700x **g** (3000 rpm). If phases do not resolve well, add another volume digestion buffer, omitting proteinase K, and repeat centrifugation. If thick white material appears at interface, repeat organic extraction. Transfer top layer (aqueous) to a new tube.

5. Add ½ vol of 7.5 M ammonium acetate and 2 vol of 100% ethanol. Centrifuge 2 min at 1700 x **g.**

6. Wash with 70% ethanol, air dry, and resuspend in TE buffer at ~1 mg/ml.

Appendix B

2018 ASAS Submitted Abstract

Over the last several decades, there has been an increase in the number of heifers categorized as B and C-maturity based upon skeletal maturity at harvest in animals that are actually under 30 months of age chronologically. While the losses associated with advanced maturity will be abated through the new USDA approved use of dentition to determine carcass age, in large commercial plants; it is important to examine the biology associated with the advanced skeletal ossification for the US beef industry and smaller plants still utilizing ossification. We hypothesize that selection for early reproductive maturity, along with faster growth rate and increased feed conversion efficiency has increased the propensity for advanced osteogenesis that physiologically mediate endochondral ossification. Specifically, we examined the genetic variation in genes that promote ossification or extracellular matrix mineralization in young heifers resulting in maturity grades more representative of heifers that are over 30 months of age. In this study, we extracted DNA from 90 heifers; 30 A-maturity, 30 B-maturity and 30 C-maturity. Following DNA library preparation we utilized targeted gene pulldown and sequencing to examine candidate genes that play an important role in ossification of cartilage for sequence variation. Sequence variations within our candidate genes, were tested for association with our A, B, and C- carcass maturity. In association with maturity grading, we have identified a significant ($p < 0.04$) single nucleotide polymorphism (SNP) in the alkaline phosphatase (ALPL) gene on chromosome two. ALPL plays a role in bone development as it is an osteoblast differentiation marker and provides the inorganic phosphate required for extracellular collagen mineralization. This represents

the first report identifying sequence variations associated with accelerated skeletal maturation in US beef heifers.