THESIS

IDENTIFYING SINGLE NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH BEEF

CATTLE TERRAIN-USE IN THE WESTERN UNITED STATES

Submitted by

Courtney F. Pierce

Department of Animal Sciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2019

Master's Committee:

    Advisor: Milton Thomas
    Co-advisor: Scott Speidel

    Stephen Coleman
    R. Mark Enns
    Paul Meiman

ABSTRACT

IDENTIFYING SINGLE NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH BEEF

CATTLE TERRAIN-USE IN THE WESTERN UNITED STATES

Beef cattle are drawn to areas with gentle terrain, which may result in heavy grazing near riparian zones and minimal grazing on rugged terrain. Traditional management tools to improve grazing distribution can be costly; therefore, genomic selection has been proposed as a means of improving beef cattle grazing patterns. The objective of this thesis was to identify single nucleotide polymorphisms (SNP) associated with beef cattle terrain-use in the western U.S. Variant detection using RNA-sequencing data obtained from Angus cardiovascular tissues and Brangus reproductive tissues revealed 48 potential causative mutations in five genes that were previously associated with terrain-use indices: *SDHAF3, RUSC2, SUPT20H, MAML3,* and *GRM5*. In an additional study, Bayesian multiple-regression was performed using BovineHD genotypes and global positioning system (GPS) data collected from 80 beef cows managed in Arizona, Montana, and New Mexico. Results of this analysis suggested that beef cattle terrain-use was polygenic; however, additional observations were needed to validate the quantitative trait loci (QTL) identified. Subsequent genome-wide association studies (GWAS) were performed for six terrain-use traits using BovineSNP50 genotypes and distribution data collected from a multi-breed population of cattle (n = 330) managed in the western U.S. These analyses identified 32 QTL and 29 putative candidate genes with diverse functions related to hypoxia, heat stress, feed efficiency, weight traits, energy metabolism, and lactation. In conclusion, results presented in this thesis suggested that terrain-use

is polygenic and may be improved with genetic selection; however, additional studies are needed

to further elucidate the genetic mechanisms underlying terrain-use of beef cattle.

ACKNOWLEDGEMENTS

DEDICATION

This thesis is dedicated to my grandmother Carleen who taught me the value of agriculture at a

young age and provided me with constant support throughout my academic career.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Introduction

Beef production in the United States is greatly important to rural economies and is often the top agricultural commodity for cash receipts (NASS, 2016; DelCurto et al., 2017). From a global perspective, the United States is the largest beef producer and one of the largest exporters of beef (FAS, 2018). Most of the land dedicated to beef production in the United States consists of rangeland and pastures in the mid-west/western states that are unsuited for urban development or cultivated crops (Sorensen et al., 2018). Subsequently, beef producers in the western U.S. face several challenges including the arid/semi-arid climate and rough topography (DelCurto et al., 2017).

Western rangelands support ~20% of the U.S. beef cattle inventory (DelCurto et al., 2017) and there may be opportunity for increased utilization of rangelands if beef cattle grazing distribution can be improved (Tanaka et al., 2007). Recent estimates suggest that one third of western rangelands receive minimal grazing due to physical attributes of the land that deter grazing (Bailey et al., 2017). For example, vegetation found in remote areas of pasture, further than 3.2 km from the nearest water source, often receive minimal grazing because cattle prefer to graze within 3.2 km of water (Valentine, 1947; Holechek, 1988). Improving grazing distribution may therefore increase forage harvest and reduce supplemental feeding (Tanaka et al., 2007) as well as minimize chronic heavy grazing that can negatively impact water and soil quality, vegetation, and wildlife habitat (Kauffman and Krueger, 1984; Heady and Child, 1994; CAST, 2002).

Initial management tools to improve grazing patterns of beef cattle included fencing, water development, herding, and supplementation (Bailey et al., 2018). While these methods proved

effective, they required large financial investment, which limited their implementation in most beef production systems (Tanaka et al., 2007). Genetic selection of beef cattle has been proposed as a strategy to improve grazing distribution or terrain-use and preliminary studies, using single-single nucleotide polymorphism (SNP) analyses, identified candidate SNP associated with terrain-use traits (Bailey et al., 2015). The aim of this thesis was to further examine the previously identified SNP and genes as well as identify additional quantitative trait loci (QTL) to further elucidate the role of genetics in beef cattle terrain-use. More specifically, using beef cow genotype and phenotype data, the project objectives were to: 1) identify potential causative mutations within the five genes previously associated with terrain-use (*SDHAF3*, *RUSC2*, *SUPT20H, MAML3*, and *GRM5*) using RNA-sequencing data; 2) investigate these candidate QTL using a Bayesian methodology that accounts for SNP-interactions; and 3) perform genome-wide association studies (GWAS) with new data (n = 330 cattle from 14 ranches) for six terrain-use traits using Bayesian multiple-regression.

CHAPTER 2: LITERATURE REVIEW

Livestock Grazing

History of Livestock Production Systems in the Western United States

Rangeland refers to land managed as a natural ecosystem in which the native vegetation is primarily composed of grasses, grass-like plants, forbs, or shrubs (Bedell, 1998). Rangelands comprise approximately half of earth's land surface and 31% of the total land area in the United States (761 million acres; U.S.; Havstad et al., 2009). As discussed by Derner and Jin (2012), "rangelands are the largest and among the most diverse land resources in the United States." Alpine regions, deserts, grasslands, marshes, meadows, savannas, shrublands, and tundra are all considered rangelands (Bedell, 1998). Most of the nation's rangelands are in the western U.S. where the diverse landscape includes low-elevation plains and basins as well as high-elevation rough country characterized by steep slopes and shallow/rocky soils (Figure 2.1; Havstad et al., 2009; DelCurto et al., 2017).

The western rangelands became the focal point of the cattle boom in the mid-1800's when newly constructed railroads provided access to eastern markets (Bohrer, 1975; Sayre and Fernandez-Gimenez, 2003). Financial capital enabled settlers in the western U.S. to expand their herds on credit (Bentley, 1898; Jackson, 1956; Atherton, 1972) and the open range provided vast amounts of forage (Cook and Redente, 1993). In the mid-1880's, severe drought during summer months depleted range resources and cattle market prices fell. Producers were unable to reduce their herds due to debt incurred at the time of purchase as selling animals when market prices were low caused the settlers to default on their loans (Sayre and Fernandez-Gimenez, 2003). When extreme winter conditions arrived many cattle starved to death in what would later be termed the

"tragedy of the commons" (Sayre and Fernandez-Gimenez, 2003). While the tragedy of the commons significantly reduced the number of cattle grazing on public lands, overgrazing left a lasting impact on the rangelands (Bohrer, 1975; Sayre and Fernandez-Gimenez, 2003).
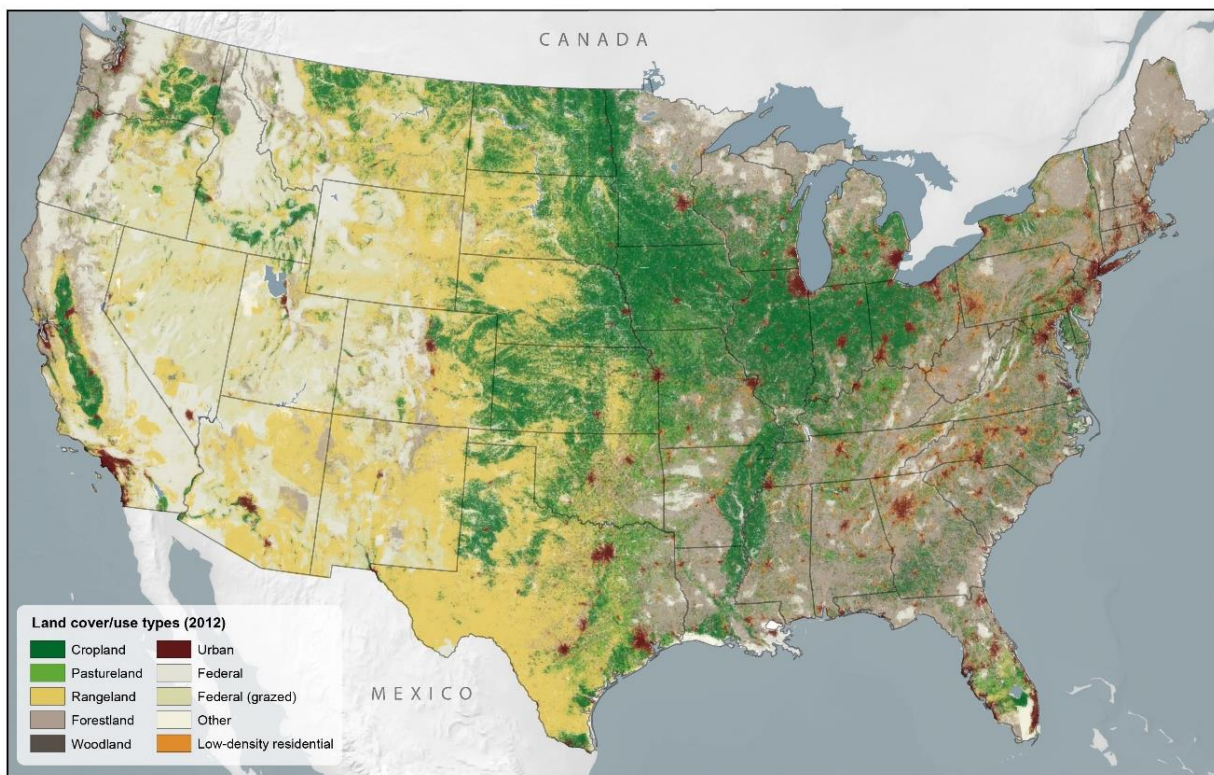
Figure 2.1. Map depicting land use in the United States in 2012 (Sorensen et al., 2018).

Jared G. Smith, H. L. Bentley, David Griffiths, E. O. Wooton, and Frederic Clements were some of the first to document the deteriorating rangelands that resulted from open access and heavy grazing (USDA Forest Service, 1944; Holechek, 1981; Sayre et al., 2012). These botanists and agronomists noted soil erosion, an increased number of poisonous plants, a higher number of woody plants, and fewer palatable plants in the rangelands spread throughout the United States (Holechek, 1981). These reports, along with many others, helped form the discipline of range science, as well as laws and policies regarding rangeland management (Sayre et al., 2012).

4

In 1898, the government started issuing permits to protect federal lands from what Sayre et al. (2012) described as "too many livestock, too often, and for too long." These permits limited the number of livestock allowed to graze on federal lands at a given time. In 1905, the Forest Service was established under the Department of Agriculture and this agency formed livestock grazing allotments as a means of coordinating grazing use and improving range conditions. During the next decade grazing laws were passed to protect national forest lands that were previously depleted by overgrazing. In 1934, the Taylor Grazing Act was passed and management responsibility for all remaining public lands was assigned by the Grazing Service which later became the Bureau of Land Management (BLM). Under the BLM, public rangelands in the western United States were separated into fenced allotments and leased to ranchers (Sayre et al., 2012).

In the years that followed, citizens became increasingly interested in the use of public lands. Rangelands previously used primarily for livestock grazing harbored untapped resources that could be exploited by the growing urban population. Consequently, laws were passed to further dictate the utilization of public lands in the western U.S. These laws include the Multiple-Use Sustained-Yield Act of 1960, which described the five major uses of national forests: range, timber, watershed, wildlife, and outdoor recreation. Comparably, the Classification and Multiple Use Act of 1964 required the BLM to classify public lands based upon equal consideration for wildlife, recreation, soil, water, range, forestry, land, and minerals. The National Environmental Policy Act of 1969 required all federal actions including the U.S. Forest Service and the BLM to evaluate the impact of federal actions, with emphasis on environmental consequences. Lastly, the Federal Land Policy and Management Act of 1976 expanded upon the 1964 act that governed how the BLM manages public lands (Hagenstein, 1971; Havstad et al., 2009).

While laws associated with the use of public lands reflected the changing needs of a growing country, these acts resulted in additional challenges for the livestock industry. Ranchers were and are faced with many challenges including the production of larger quantities of animal products on reduced land areas. Tolleson and Meiman (2015) reported that between 1982 and 2007 the total area of cropland, pastureland, and rangeland in the United States declined by 63, 12, and 9 million acres, respectively. This substantial loss in agricultural lands was a direct result of urban development and an increased interest in recreational areas, endangered species conservation, and environmental sustainability (Hendrickson, 2015; Tolleson and Meiman, 2015).

As public land use becomes increasingly controversial, it is important to acknowledge the positive aspects of livestock production and its contribution to food security. Beef cattle production provides: 1) essential amino acids and micro/macro nutrients for human consumption, 2) fertilizer, 3) products for human medicine, 4) brush control, fire prevention, and nutrient enrichment in wildlife areas, and 5) income for individual households and the nation (Tolleson and Meiman, 2015; Mottet et al., 2017). According to the National Agricultural Statistics Service (NASS) of the USDA, cattle production was the top agricultural commodity in 2015, generating $78.2 billion in cash receipts in the U.S. (cash income from commodity sales; NASS, 2016). In July of 2017, approximately 102.6 million head were reported in the cattle and calves inventory and of these 102.6 million, 32.5 million head were classified beef cows and 4.7 million were beef replacement heifers weighing more than 500 pounds (NASS, 2017).

Historically, western rangelands have supported 20% of the beef cattle inventory and "approximately 20% of the animal unit months (AUMs) for western livestock production are derived from public lands" (DelCurto et al., 2017; Drouillard, 2018). Consequently, it is crucial to implement grazing management strategies that will optimize production and appeal to the

environmentally conscious consumer who values public lands for recreation, wildlife habitat, etc. (DelCurto et al., 2017).

Effects of Grazing on Vegetation and Riparian Zones

Briske et al. (2008) summarized four major principles that must be considered when managing rangelands or forming grazing management plans: 1) plant growth and survival, 2) primary productivity, 3) forage quality, and 4) species composition. At the center of rangeland management, is the desire to have healthy vegetation which begins with the fundamental concept of plant growth and survival. Nearly all plants rely on photosynthesis as a means of deriving the chemical energy needed for growth and cellular respiration. "The process of photosynthesis requires photosynthetically active radiation, water, and carbon dioxide," all of which are obtained through the root system and the leaves (Ashton, 1998). Extended periods of heavy grazing negatively affect a plant's ability to perform photosynthesis due to continual defoliation (Briske and Richards, 1995). In turn, a chronic reduction in photosynthesis negatively impacts the number of branches as well as root mass, distribution, and longevity (Hodgkinson and Becking, 1978). In altering the structure of the plant, chronic intensive grazing restricts water and nutrient uptake and successively plant growth (Briske et al., 2008). Conversely, limiting the grazing period and providing the vegetation with a rest period has been shown to promote plant growth and re-growth (Holechek et al., 2001) which are important for long-term productivity (Huston and Pinchak, 1991).

While chronic defoliation may negatively impact plant growth, studies have shown that herbivory may positively impact plant performance under favorable conditions (McNaughton, 1979b; Paige and Whitham, 1987; Briske et al., 2008). This concept has been termed grazing optimization or grazing overcompensation (McNaughton, 1979b; Vail, 1992; Hayashi et al., 2007).

7

Grazing optimization describes how the productivity of grazed plants will surpass that of ungrazed plants when grazing intensity reaches an optimal level and then decreases (McNaughton, 1979a). For grazing optimization to occur, the grazing pattern must consist of heavy grazing early in the growing season followed by a rest period in which grazing is restricted or non-existent (Frank and McNaughton, 1993). Similarly, grazing overcompensation refers to improved plant fitness following grazing (Vail, 1992). Studies compiled by Leriche et al. (2001) suggest that herbivory may improve plant performance due to: 1) greater light availability, 2) reduction of water stress, 3) elevated nutrient cycling due to the herbivores' excretory products, 4) altered biomass reallocation, and 5) increased photosynthetic rate. Livestock and wildlife herbivory are therefore an important part of the western rangeland ecosystem.

Plant growth, survival, and productivity leads us to the third rangeland management principle, forage quality. The nutrients that can be derived from a plant are dictated by forage quality or the ratio of cell soluble contents to structural components (Briske et al., 2008). This ratio is influenced by tissue age, tissue type, plant function, cell wall components (cellulose and lignin), and plant secondary compounds (Huston and Pinchak, 1991). For example, tissue age is inversely related to the proportion of soluble cell contents. Plants with a higher ratio of cell soluble contents have a higher forage quality making them desirable to grazing animals (Briske et al., 2008). Frequent grazing positively impacts the soluble to structural component ratio or forage quality by lowering the average tissue age (Walker et al., 1989). This relationship can improve animal performance by increasing the amount of nutrients available to the animal upon consumption of the vegetation (McNaughton, 1984). On the contrary, grazing plans that involve long rest periods result in decreased forage quality due to an increase in average tissue age (Briske et al., 2008).

As discussed by Huston and Pinchak (1991), the value of a particular plant species is often determined using forage quality which is based upon chemical composition (crude protein, minerals, fat, etc.). If the goal of the rangeland beef operation is to optimize animal performance, then maximizing rangeland forage quality may appear desirable or even necessary. However, maximizing forage quality may "reduce long term secondary production by decreasing the stability of the forage resource" (Huston and Pinchak, 1991). Furthermore, lower quality plant species may better promote long term production due to longer growing seasons, greater dry matter production, or a higher tolerance for herbivory (Huston and Pinchak, 1991).

While the previous principles can be applied to an individual species or plant, species composition considers the plant community. As herbivores graze on preferred plants or groups of plants, they modify the community composition. As Briske et al. (2008) noted, "selective grazing of individual species or species groups places them at a competitive disadvantage with less severely grazed species or species groups and alters competitive interactions." Briske et al. (2008) is referring to the concept of increasers, decreasers, or invader plants. Species that decline in the presence of herbivores due to defoliation and heavy grazing are classified as decreasers. These species tend to be more palatable and are therefore consumed at a higher rate. In contrast, species that increase during grazing due to their moderate palatability or higher tolerance for defoliation are categorized as increasers. Invaders are thought of as opportunists for they appear in the plant community during periods of grazing in which dominant species are suppressed. These invader species replace more palatable species with higher forage qualities (Archer and Smeins, 1991).

Rangeland ecologists and managers used these fundamental principles of range management to develop grazing systems that promoted healthy ecosystems and optimized livestock production (DelCurto et al., 2005; Briske et al., 2011). In the 1950's, deferred-rotation

grazing systems, which promoted herbivory after seed maturation, were the preferred grazing system in the western U.S. Gradually, numerous grazing systems were implemented on the western rangelands: rest-rotation, Santa Rita, seasonal suitability, best pasture, and short duration grazing (Howery et al., 2000). While many of these grazing systems attempted to address overgrazing in riparian zones, degradation of these areas continues to trouble the western U.S. (Howery et al., 2000; DelCurto et al., 2005). Perhaps this is due to improper implementation or a lack thereof; nonetheless, degradation of riparian zones is problematic as these zones provide habitat for small mammals and aquatic species, nutrient rich forage for herbivores, and function in watershed hydrology and stream morphology (Kauffman and Krueger, 1984).

As summarized by Kauffman and Krueger (1984), overgrazing in riparian zones may negatively impact plant performance, soil and water quality, and animal biodiversity. Chronic, heavy grazing may reduce, alter, or eliminate vegetative cover which may increase soil erosion, reduce mammalian habitat, and increase water temperatures (Kauffman and Krueger, 1984; CAST, 2002). Heavy grazing may also lead to soil compaction and consequently, reduced plant growth (Heady and Child, 1994; CAST, 2002). Livestock excrement in riparian zones may reduce water quality (e.g., suspended solids) which can have negative implications for aquatic life and humans who come into contact with the water source (Kauffman and Krueger, 1984; CAST, 2002).

Abiotic and Biotic Factors that Affect Grazing Distribution

Livestock grazing distribution is defined as the dispersion of grazing animals in a pasture or allotment (Volesky et al., 1996). Issues with grazing distribution arose in the 1800's when domestic livestock were first introduced to rangelands, but concentrated grazing (i.e., large concentrations of cattle grazing in one area) continues to be a challenge for modern rangeland operations (Williams, 1954; DelCurto et al., 2005). Livestock do not uniformly disperse

throughout a management unit, but instead, often congregate in desirable areas. Uniform grazing distribution is difficult to obtain on western rangelands because numerous abiotic and biotic factors influence the grazing patterns of cattle (Bailey et al., 1996; Bailey, 2005; Lunt, 2013).

Cattle in the western United States typically graze in extensive rangeland pastures of rough terrain. In studying the grazing patterns of these beef cattle, researchers have identified an association between topographic features and site selection (Bailey et al., 1996). Research published by Valentine (1947) demonstrated that grazing intensity has an inverse relationship with distance travelled from water. Valentine (1947) reported that forage utilization was 38% greater in areas less than 0.8 km from the water source as compared to areas located approximately 3.2 to 4 km from the water. Holechek (1988) used the results of Valentine (1947) and others to formulate suggested reductions in grazing capacity that corresponded with distance from water. Areas located 1.6 to 3.2 km from water were associated with a 50% reduction in grazing capacity whereas areas greater than 3.2 km were deemed "ungrazable" (Holechek, 1988). Similarly, Roath and Krueger (1982) recognized that forage utilization nears zero for areas located ≥ 80 vertical meters above water. Results of this study suggest that vertical distance from water influences grazing patterns.

Previous studies also suggest that percent slope may limit pasture utilization. Upon studying the effect of slope on grazing distribution in a mountainous region of southwestern Montana, Mueggler (1965) observed a negative relationship between percent slope and utilization. Gillen et al. (1984) reported similar findings, cattle preferred to graze on slopes of less than 20% grade in the Malheur National Forest. As with distance from water, Holechek (1988) provided recommendations for reduction in grazing capacity given percent slope: slope 0 to 10 (no

reduction), slope 11 to 30 (30% reduction), slope 31 to 60 (60% reduction), and slope > 60 (ungrazable).

Bailey et al. (1996) described many biotic factors that affect herbivore grazing patterns including: species composition, plant morphology, and forage quality and quantity. Grazing frequency and duration for a particular area are well-documented as corresponding with nutrient availability in a given plant community (Senft et al., 1987; Bailey et al., 1996). Briefly, cattle prefer to graze in areas that contain large quantities of high quality forage (Bailey et al., 1996). As summarized by Bailey (2005), increased utilization of riparian zones is unsurprising because forage production in riparian zones can be 6x greater than forage production in the uplands and riparian forage may contain a higher crude protein.

Management Strategies to Improve Livestock Grazing Distribution

Uneven grazing distribution challenges beef production and rangeland sustainability in the western U.S. Concentrated grazing can reduce forage harvest which increases the need for supplemental feed (Vallentine, 1990; Tanaka et al., 2007), and degrades riparian areas (Kauffman and Krueger, 1984); therefore, developing tools to improve the grazing patterns of domestic livestock is necessary (Stephenson, 2015; Bailey et al., 2018). While heterogeneous pastures hinder the rapid improvement of grazing uniformity, several management methods have successfully altered the grazing patterns of range cattle (Stephenson, 2015). Bailey (2004) separated grazing management strategies into two categories: 1) methods that alter the attributes of the management unit and 2) methods that modify livestock behavior. Traditional methods such as water development, fencing, and seasonal distribution (i.e., when grazing occurs), fall within the first category whereas low-stress herding and supplementation are within the second (Bailey, 2004).

Altering pasture management to improve livestock grazing patterns is not a novel concept. In fact, many of the traditional methods listed above were implemented in the early 1900's shortly after cattle were introduced to rangelands in the United States (Williams, 1954). Ingram (1930) reported instances of producers hauling water to undesirable areas of their pastures to improve sheep grazing patterns in 1918. Pechanec and Stewart (1949) confirmed the importance of water developments on sheep rangelands in southern Idaho and Harris (1950) recommended using temporary water developments to draw cattle to ungrazed regions. Additionally, Ganskopp (2001) evaluated the utility of water developments for improving cattle grazing patterns and found that grazing patterns were drastically altered with the introduction of a new water source.

A more direct method of altering grazing distribution is fence construction. Stoddart and Smith (1943) documented the use of boundary, division, and drift fences to control grazing on rangelands. In 1954, Williams suggested using fencing to divide extensive rangelands into multiple pastures each containing a permanent water source. According to Bailey and Rittenhouse (1989), fencing homogeneous areas can improve grazing uniformity by eliminating the heterogeneity of the landscape that promotes concentrated grazing in desirable areas. For instance, cattle will graze in areas with rough terrain if the pasture does not contain gentle terrain. This provides an opportunity to protect riparian areas that are often overgrazed due to the abundance of high quality forage (Roath and Krueger, 1982; Bailey et al., 2004).

As discussed by Bailey et al. (1996), cattle are drawn to areas that have higher quantities of high quality forage in heterogenous pastures. Ranchers can use this relationship between forage quality and grazing patterns to their advantage when attempting to improve grazing distribution. In mountainous regions, high quality forage can be found in the uplands during the early summer months; however, during the late summer, the forage quality decreases significantly due to tissue

aging and drier soils. Forage within riparian areas is relatively unaffected by the changing seasons because soils in these areas have a higher moisture content (Vallentine, 1990). Parsons et al. (2003) studied the effect of season on the grazing patterns of cow/calf pairs in eastern Oregon and reported that cattle grazed in the uplands more during the early summer months than the late summer months; therefore, suggesting that early summer grazing may improve riparian conditions in mountainous regions and that livestock dispersion in areas with rugged terrain is best in the fall and early winter because all forage is dormant and of low quality.

For centuries, ranchers have used herding to move livestock from one area to another; however, its efficacy in the management of grazing distribution is controversial. Some view herding as ineffective because livestock often return to riparian areas in the absence of range rider(s); however, others describe herding as an excellent way to improve grazing uniformity on rangelands (Williams, 1954; Skovlin, 1957; Butler, 2000). Bailey et al. (2008) demonstrated that low-stress herding (i.e., a combination of pressure and release movements) can improve pasture utilization by increasing the amount of time cattle graze in the uplands and decreasing the amount of time cattle graze in riparian areas. However, as discussed by (Tanaka et al., 2007) the effectiveness of herding varies by ranch and is related to the frequency of herding, herding techniques, and natural cattle movements.

According to Bailey et al. (2008) supplementation can be combined with herding to improve pasture utilization. When cattle were herded to a site that contained both a salt block and low-moisture block (i.e., dehydrated molasses supplement), forage utilization within a 600-meter radius of the supplementation site increased by 10 to 20%. This is consistent with Bentley (1941) who reported that cattle grazed in the areas surrounding the supplementation site when salt blocks were placed 1.2 km from water. Tanaka et al. (2007) concluded that producers may use strategic

supplementation to increase grazing uniformity, lengthen the grazing season, and reduce the amount of hay fed in the fall.

While modifying pasture attributes and animal behaviour has proven to be effective, these practices are laborious and generally too costly for most beef production systems. Tanaka et al. (2007) examined the effectiveness of water development, fencing, herding, and supplementation from an economic standpoint for a 300-head cow-calf operation and found that implementing these production practices increased production costs by an estimated \$1.35 per animal unit month (AUM), \$1.55 per AUM, \$3.30 per AUM, \$6.83 per AUM, respectively. This presents a need for new management practices to be developed and implemented. Research suggests that modifying animal behavior via genomic selection may improve grazing distribution on western rangelands (Bailey et al., 2001a; Bailey et al., 2001b; Bailey, 2004; Bailey et al., 2004; Bailey, 2005; Bailey et al., 2006; Bailey et al., 2015).

Genetic Basis of Grazing Distribution

Numerous studies document beef cattle's varied terrain-use in extensive and mountainous pastures. In studying the home range of cattle in the southern Blue Mountains of Oregon, Roath and Krueger (1982) discovered that livestock distribute unevenly across pastures and vegetation types. Moreover, areas containing desirable forage were left ungrazed when surrounded by rugged terrain. Howery et al. (1996) reported terrain-use differences for cattle sharing a common home range. Time spent grazing and resting in riparian zones and upland habitats varied by animal within four distinct home ranges. These studies, along with others, suggest that individual animal selection may improve livestock dispersion and pasture utilization (Roath and Krueger, 1982; Howery et al., 1996; Howery et al., 1998).

Bailey (1999) presented four premises that must be true for selective culling to effectively improve beef cattle grazing patterns: 1) there must be substantial between-animal variation for terrain-use, 2) terrain-use phenotypes must be relatively easy to measure or predict, 3) terrain-use must be a heritable trait (differences in terrain-use must be inherent), and 4) terrain-use cannot be negatively correlated with performance traits. Bailey et al. (2001b) reported that Tarentaise cattle a breed that originated in the French Alps grazed steeper slopes and travelled to higher elevations than the Hereford cattle originating in England. Additionally, residual correlations revealed no association between terrain-use traits (slope, elevation, and distance from water) and animal performance (weight, height, and body condition score). Subsequent research by Bailey et al. (2001a) compared the terrain-use of cows sired by Angus, Charolais, Piedmontese, and Salers bulls. Daughters of Piedmontese bulls, with origins in the Italian Alps, grazed at higher elevations than daughters of Angus bulls (Scotland). These results support the findings of Bailey et al. (2001b).

In 2004, Bailey and colleagues demonstrated the utility of using global positioning system (GPS) technology to collect terrain-use data. Observations recorded by researchers on horseback were used to categorize cows as hill climbers (prefer steep slopes and high elevations) or bottom dwellers (prefer gentle terrain and riparian zones). During the following year, nine cows that exhibited extreme terrain-use (4 hill climbers and 5 bottom dwellers) were monitored using GPS tracking collars. Global positioning system coordinates reinforced the concept of bottom dwellers and hill climbers as there were no changes in terrain preference from the previous grazing season. Additionally, Bailey et al. (2006) determined that hill climbers and bottom dwellers maintain their grazing patterns even when they are moved to different pastures. Results of these studies (Bailey et al., 2001a; Bailey et al., 2001b; Bailey et al., 2004; Bailey et al., 2006) provide evidence to

suggest that individual animal selection is a suitable method for improving livestock dispersion on large-scale beef operations.

To further elucidate the role of genetics in beef cattle terrain-use, Lunt (2013) conducted a study in which embryo transfers were performed using donor/recipient cows that were categorized as hill climbers or bottom dwellers (i.e., embryos from cows classified as hill climbers were placed in cows classified as bottom dwellers, and vice versa). Donor cows were bred to one Simmental sire resulting in 39 heifer calves. Terrain-use phenotypes (slope, elevation, distance from water, distance travelled per day, and ratio index combining slope, elevation, and distance from water) were collected using GPS tracking collars when these heifer calves reached maturity (6 to 8 years old). Statistical analysis revealed no significant association between terrain-use traits and recipient status (hill climber vs bottom dweller; $P > 0.10$). Furthermore, donor status was not significantly associated with slope, elevation, distance travelled per day, or the ratio index ($P > 0.10$). On average, distance travelled from water was greater for cows whose donor dams were categorized as hill climbers ($P = 0.07$). As concluded by Lunt (2013), further studies were needed to discern the genetic factors that influence grazing patterns.

Bailey et al. (2015) conducted a study in which 158 cows were tracked using Lotek 3300 GPS collars and 80 cattle were genotyped using an Illumina BovineHD Beadchip. Average slope, elevation, and distance travelled from water were calculated for each animal and used to quantify terrain-use using two indices: rough and rolling. Single nucleotide polymorphisms (SNP) that were significantly associated with the indices were incorporated into a 50 SNP custom genotyping panel which was then used to genotype the original 80 cows as well as an additional 78 cows. A marker-trait association analysis identified 12 SNP in 5 candidate genes (*SDHAF3, SUPT20H, GRM5, MAML3,* and *RUSC2*) that may be associated with terrain-use. These SNP explained 34 to 36% of

the variation in the rolling and rough indices, which suggested that terrain-use is a moderately heritable, polygenic trait that may be improved using genomic selection.

Mercado et al. (2018) examined the repeatability of terrain-use in extensive rangeland pastures using GPS tracking data obtained from cows on 5 different ranches in the western U.S. Intra-class correlation estimates for terrain-use traits (averaged per week) varied by ranch: slope (0 to 0.60), elevation (0 to 0.71), and distance from water (0.02 to 0.77). Results of this repeated measures analysis suggest that terrain-use may be moderate to highly repeatable; however, additional records are needed to confirm the repeatability of terrain-use.

Monitoring Livestock Grazing Patterns

Initial livestock behaviour studies involved tracking animals on foot, horseback, or by vehicle. Herbel and Nelson (1966) studied the grazing patterns of beef cattle in the Chihuahuan Desert using a vehicle equipped with a spotlight. Howery et al. (1996) hiked to various locations in study pastures to document the locations and activities of crossbred cattle. Bailey et al. (2004) classified cows as hill climbers or bottom dwellers based upon locations recorded by riders on horseback. While these studies provided fundamental knowledge regarding grazing patterns and activities, the tracking methodologies provided low accuracy. Visual observations were difficult to obtain at night, unfavourable during poor weather conditions, and strenuous in areas with rugged terrain. This severely impacted the frequency and accuracy at which positions were recorded (Howery et al., 1996; Bailey et al., 2018).

In the late 1980s and early 1990s, researchers began using global positioning system technology to "collect fine-scale location data for far-ranging species" (Thomas et al., 2012). As discussed by Recio et al. (2011), GPS technology enabled researchers to increase the frequency and accuracy of their tracking as well as track animals on a 24-hour basis in diverse weather

conditions and terrain. During the past 30 years, GPS technology has been used to study livestock movement, activity, and resource utilization (Bailey, 2000; Turner et al., 2000; Ganskopp; Johnson and Ganskopp, 2008; Swain et al., 2011; Bailey et al., 2018).

Commercial GPS collars greatly improved livestock behavior research; however, this technology was costly and provided limited data-storage (Clark et al., 2006). Each GPS collar cost $1,500 to $2,000, making it difficult to simultaneously track a large number of cows (i.e., obtain a sufficient sample size; Anderson et al., 2013). In addition, many commercial collars lacked sufficient data-storage needed for high-frequency data collected over long periods of time (Clark et al., 2006). Thus, several more economical (< $1,000) GPS collars were developed (Clark et al., 2006; Allan et al., 2013; Knight, 2016; McGranahan et al., 2018). As discussed by Forin-Wiart et al. (2015) these low-cost collars must be deployed and examined to ensure adequate performance and accuracy because they may be inferior to commercial collars. Knight et al. (2018a) compared beef cow terrain-use data obtained using Lotek 3300 GPS collars ($2,000; Lotek Wireless, New Market, Ontario, Canada) and igotU Gt-120 GPS tracking collars ($250; Knight et al., 2018a) to identify major discrepancies between the two collars. No difference was observed between the terrain-use measurements obtained from the two collars ($P \geq 0.37$); however, an additional comparison of Lotek and igotU collars revealed that the igotU collars had a 13.8% lower fix rate (proportion of fix attempts that resulted in a location) and 21.7% more missed observations and 0.17% more inaccurate data points (Knight et al., 2018b). Ultimately, "the choice of which tracking collar to use is a function of functionality and cost" (Thomas et al., 2012).

Genomic Technologies

Bovine Genome

In the early 1930's, the Fort Keogh Livestock and Range Research Laboratory in Miles City, MT partnered with the Montana Agricultural Experiment Station to exploit heterosis and develop true-breeding lines of Hereford cattle. These lines were suitable for western rangelands, exhibited high fertility, and provide high-quality beef (Black and Knapp Jr, 1936; Durham, 2010). Line 1 was developed in 1934 when two sons of Advance Domino 13, Advance Domino 20 and Advance Domino 54 from Kremmling, CO, were bred to 50 cows from Miles City, MT. The female progeny produced in these matings were then bred to the paternal half-sibling of their sire. Subsequent generations of Line 1 Herefords have an average genetic relationship with Advance Domino 13 of ≥ 39% (MacNeil, 2009).

The high level of homozygosity within Line 1 Hereford made these cattle an excellent resource for genetic research (Krehbiel, 2017). More specifically, Line 1 was used to derive the first estimates of heritability and the genetic correlations for economically relevant traits (ERT) in beef cattle. Additionally, the deoxyribonucleic acid (DNA) of L1 Domino 99375 was used to create a bacterial artificial chromosome (BAC) library for future research on causal mutations associated with ERT. The most substantial contribution of the Line 1 Herefords to bovine genetics research came in 2003 when DNA from L1 Dominette 01449 was used to establish the bovine reference genome (Elsik et al., 2009; Durham, 2010).

Richard Gibbs and George Weinstock, with the collaboration of numerous international researchers, began sequencing the bovine genome in December 2003 at Baylor College of Medicine's genome sequencing center. Researchers used Sanger sequencing with 7.1-fold coverage of the bovine reference genome, physical maps (generated using genomic DNA

fragments cloned into BAC), and whole-genome shotgun sequencing to develop and release the first draft of the *Bos taurus* bovine genome (BTAU4.0) in 2009. BTAU4.0 contained 135,743 contigs (N50 contig size: 48.7 Kb) and 13,388 scaffolds, of which, 90% were mapped to the 29 bovine autosomes and the x chromosome (Burt, 2009; Zhou et al., 2015).

Shortly after BTAU4.0 was released, researchers at the University of Maryland's Center for Bioinformatics and Computational Biology developed a new bovine reference genome (UMD2) using Baylor's raw sequence data. Unlike BTAU4.0, UMD2 was assembled using paired-end BAC sequences, mapping data, and the human genome sequence. These new assembly techniques enabled researchers to map 91% of the contigs (44,433; N50 contig size: 93.56) to the 29 autosomes and sex chromosomes. In comparison, UMD2 had greater sequence coverage as well as fewer sequence gaps, misassemblies, and single nucleotide polymorphism errors than BTAU4.0 (Burt, 2009; Zhou et al., 2015).

Improvements were made to both BTAU4.0 and UMD2 (sequence gaps filled and corrected misassemblies) and new versions were released: UMD3.1 and BTAU4.6. While the updated assemblies were superior in comparison to the original bovine assemblies, inconsistencies between the UMD and BTAU assemblies continue to limit the accuracy of results from genomic analyses. Zhou et al. (2015) developed a bovine optical map (BtOM1.0) to identify discrepancies between the assemblies and provide information needed to improve UMD and BTAU. In comparing BtOM1.0 to UMD3.1.1 and BTAU4.6.1, 4,754 and 7,463 discordances were observed. Results of this analysis confirmed the need for one standard, well-constructed bovine reference genome. As discussed by Medrano (2017), "well-annotated genome assemblies in agricultural species have become essential tools to enable the understanding of phenotypic variation and practical applications of DNA technologies." Furthermore, an accurate assembly allows us to

better understand genome evolution and architecture, long-range gene regulation, polymorphisms, and pathologies associated with genome architecture (Partipilo et al., 2011).

The USDA Agricultural Research Service attempted to address this challenge by developing a bovine reference assembly using a *de-novo* assembly method and Pacific Biosciences long-read sequencing (ARS-UCD1.2). Briefly, *de-novo* assembly involves comparing raw sequence reads to identify overlapping regions that can be joined to generate a continuous sequence (Viluma, 2017). Long-read sequencing enables researchers to overcome challenges associated with assembling complex eukaryotic genomes containing repetitive DNA sequences. Unlike short-read sequencing, long-read sequencing produces reads that span long repetitive sequences reducing the number of gaps and poorly assembled repetitive regions (Berlin et al., 2015; De Bustos et al., 2016). Reference assembly statistics suggest that ARS-UCD1.2 will be superior to the current public reference assemblies (Table 2.1).

Table 2.1. Assembly statistics for bovine reference assembly ARS-UCD1.2 (Released 04/11/2018; NCBI, 2018a).

Statistic	
Total sequence length, bp	2,715,853,792
Total assembly gap length, bp	28,162
Gaps between scaffolds	0
Number of scaffolds	2,211
Scaffold N50, bp	103,308,737
Scaffold L50, bp	12
Number of contigs	2,597
Contig N50, bp	25,896,116
Contig L50, bp	32
Total number of chromosomes and plasmids	31
Number of component sequences (WGS or clone)	2,211

Single Nucleotide Polymorphisms

As defined by Vogel and Motulsky (1987), "a polymorphism is a Mendelian or monogenic trait that exists in the population in at least two phenotypes (and presumably at least two genotypes), neither of which is rare-that is, neither of which occurs with a frequency of less than 1-2%." Single nucleotide substitutions/polymorphisms (SNP) are polymorphisms that occur during DNA replication when one base or nucleotide is substituted for another. This phenomenon is commonly referred to as a substitution mutation, point mutation, or base substitution. There are two types of nucleotide substitutions: transitions and transversions. Transitions occur when a two-ring purine is substituted for a two-ring purine or when a one-ring pyrimidine is substituted for another one-ring pyrimidine. Transversions occur when a one-ring purine is interchanged with a two-ring pyrimidine or vice versa (Vignal et al., 2002). Studies examining SNP associated with human disease suggest that the most prevalent nucleotide substitutions are as follows: C substituted for T, T substituted for C, G substituted for A, and substituted to G. Thus, transitions are more common than transversions (Antonarakis and Cooper, 2013).

Single nucleotide polymorphisms are found throughout the genome in both coding (i.e., exons) and non-coding regions (i.e., introns, intergenic regions, 5' or 3' untranslated regions, promoters, and transcription factor binding sites). As discussed by Shen et al. (1999), "the frequency of SNPs varies between genomic regions and between coding and noncoding sequences." Variants found within coding regions can be characterized as synonymous (codon encodes for the same amino acid) or non-synonymous (codon encodes for a different amino acid). Non-synonymous SNP are further classified as missense or nonsense variants. Missense SNP occur when the codon substitution alters the amino acid and nonsense SNP occur when the codon substitution results in a stop codon that prematurely stops protein translation (Klug et al., 2013).

Simply put, SNP location governs SNP function. This is important because SNP located within exons or regulatory regions can alter gene function/expression and consequently, an animal's health or performance (Ibeagha-Awemu et al., 2008). It is important to note that although synonymous SNP were previously deemed insignificant, recent studies suggest that synonymous SNP may alter protein structure, function, and expression by operating in pre-mRNA splicing, as well as mRNA stability and structure. In addition, synonymous variants may influence protein translation, and co-translational protein folding (Hunt et al., 2009).

While SNP may be multiallelic (i.e., containing three or more nucleotides), most SNP are bi-allelic containing only two nucleotides. This tendency to be bi-allelic can be attributed to the low frequency of single base pair substitutions and the higher frequency of transitions compared to transversions (Vignal et al., 2002). Another key characteristic of SNP is their high abundance in comparison to other genetic variants. In 2013, dbSNP listed 13,146,622 SNP for the bovine genome and 66,994 of these were nonsynonymous SNP (Adelson et al., 2014). Daetwyler et al. (2014) reported 26.7 million SNP identified during the 1000 bull genomes project. This high prevalence, in addition to genetic stability and the ease at which they can be analyzed with high-throughput technology, make SNP a useful tool in genomic analyses (Heaton et al., 2001).

Single Nucleotide Polymorphism Discovery and Genotyping

Single nucleotide polymorphism discovery includes scanning DNA sequences for novel SNP and genotyping animals for known SNP. Novel SNP detection is accomplished using either a global (genome-wide) approach or a targeted approach. The global approach is used to randomly detect SNP across the entire genome, whereas the targeted approach is used to identify SNP within candidate genes or a population of interest (Kwok and Chen, 2003). Genotyping cattle for known SNP is completed in two steps: 1) allelic discrimination and 2) signal detection. Allelic

discrimination describes the process of differentiating between alleles at a specific locus. This process is commonly completed using hybridization, primer extension, or enzyme cleavage (Twyman, 2005).

As described by Twyman (2005), in a hybridization-based assay allele specific oligonucleotide probes labelled with radioactive or fluorescent tags are used to detect SNP. When the probe binds to the complementary sequence, the tag is detected, enabling researchers to determine which allele is present at the locus of interest. During an allele-specific single-base extension (primer extension), primers are designed to anneal to the nucleotide that directly precedes the locus of interest. After the primer has annealed, DNA polymerase extends the primer by adding a base that is complementary to the single nucleotide polymorphism. Fluorescently labeled deoxynucleotides (dNTP) provide a detectable signal that enables researchers to determine the allele of the SNP. Restriction fragment length polymorphism (RFLP) analysis utilizes allele specific enzymatic cleavage to detect single nucleotide polymorphisms. In this analysis, the DNA sample is digested with restriction endonucleases and then the fragments are separated according to length using gel electrophoresis. When a variant occurs in the restriction site, the restriction endonuclease will fail to cut the DNA resulting in a larger fragment than expected.

The second step in SNP genotyping is signal detection, which enables researchers to identify which allele is present at the locus of interest. Three common methods of signal detection include the use of fluorescence, mass spectrometry, or pyrosequencing. Microarrays or bead arrays use direct fluorescent detection in which nucleotides are tagged with fluorescent dye. In contrast, Taqman genotyping assays and Molecular beacon assays use fluorescence resonance energy transfer (FRET) quenching in which donor and acceptor fluorophores produce a fluorescent signal. Unlike the previous methods, signal detection using mass spectrometry is based upon the

molecular weight of the DNA fragments instead of fluorescent tags. A mass spectrophotometer uses differences in molecular mass to differentiate between alleles. Lastly, pyrosequencing technology is based upon the pyrophosphate that is released when a dNTP is added to a DNA strand. Pyrophosphate is used to convert adenosine 5'-phosphosulphate (APS) to adenosine triphosphate (ATP) which results in luciferase activity that generates a chemiluminescent signal (i.e., visible light) that can be detected (Twyman, 2005; Dale et al., 2012).

Commercial genotyping panels (SNP-chips) are the most common method of genotyping cattle because they provide accurate genotyping for thousands of SNP that have been previously validated. Genotyping panels are based upon microarray technology in which hundreds or thousands of probes (oligonucleotides, amplicons, DNA fragments, or RNA fragments) are aligned on a glass or silico surface (Heller, 2002). Single stranded DNA is fragmented, tagged with a fluorescent and washed over the microarray to allow hybridization with complementary strands (Govindarajan et al., 2012). Illumina Inc. (San Diego, CA) manufactures a variety of commercial genotyping panels as well as custom genotyping panels for targeting specific regions of the genome: BovineLD, BovineSNP50, BovineHD, Infinium iSelect high definition (HD), iSelect high-throughput screening (HTS), and XT iSelect Custom BeadChips (Table 2.2).

Table 2.2. Commercial DNA microarray panels developed by Illumina Inc. (San Diego, CA) for genotyping cattle.

Microarray Product	Number of SNP	Average Distance Between SNP (Kb)	Cost	Common Applications	Reference
BovineLD	7,931	383	$40	Imputation	Boichard et al. (2012)
BovineSNP50	53,714	50.6	$100	Genome-wide association studies	Matukumalli et al. (2009)
BovineHD	777,962	3.43	$200	Genome-wide association studies	Matukumalli et al. (2011)

Single nucleotide polymorphism discovery has successfully been conducted in both beef and dairy cattle. Single nucleotide polymorphism discovery using RNA-sequencing technology has identified SNP associated with puberty, growth, development, and feed efficiency in beef cattle and lactation in Holsteins (Cánovas et al., 2010; Dias et al., 2017; Pareek et al., 2017). Stothard et al. (2011) used whole-genome resequencing to detect genetic differences (i.e. SNP and CNV) between Angus and Holstein bulls. Williams et al. (2009) implemented SNP discovery using sequence-tagged sites in European cattle to identify SNP associated with beef production and quality. Each of these studies demonstrated the utility of SNP discovery for identifying the genetic source of phenotypic variation in economically relevant traits.

Single Nucleotide Polymorphism Validation

Prior to downstream application or commercial use, newly discovered SNP need to be validated in an independent population. Advances in sequencing technology have enabled researchers to associate thousands of SNP with beef cattle trait levels; however, discovery populations are often small (< 1,000 animals) and false positives may be observed due to sequencing errors, misaligned reads, or a poorly assembled reference genome (Barendse, 2005; Van Eenennaam et al., 2007; Kumar et al., 2012). Barendse (2005) developed a 5-step process to

identify, validate, and commercialize novel SNP: 1) verify the importance of the trait and identify methods of improvement (genetic and non-genetic), 2) discover and then confirm the association between the SNP and phenotype, 3) calculate the size of the SNP effect and its economic impact, 4) examine SNP in close proximity with the candidate gene to better understand the causative mutation, and 5) design assay.

Genome-wide Association Studies

Genome-wide association studies (GWAS) refer to association studies in which high-density SNP spanning the genome are examined to identify variants significantly associated with disease or heritable quantitative traits (Hirschhorn and Daly, 2005). As defined by Bourdon (1997), a quantitative trait is "a trait in which phenotypes show continuous (numerical) expression." Designing GWAS requires careful consideration of sample size, the number of markers for genotyping, selection criteria for individuals and markers, and statistical methodology (Hirschhorn and Daly, 2005; Balding, 2006). As discussed by Mei and Wang (2016) in GWAS the number of independent variables or markers (p) often greatly exceeds the number of observations (n). This is commonly known as the "large p small n" or "fat-short data" problem. When p greatly exceeds n, it is computationally challenging to use standard regression methods to derive parameter estimates. In fact, multiple regression based on ordinary least-squares cannot simultaneously estimate all parameters (Fernando and Garrick, 2013). To overcome challenges associated with "large p small n", researchers have implemented statistical methodologies that individually estimate SNP effects.

Single-SNP analyses were the most common method of conducting GWAS and can be performed using continuous traits (linear regression), case-control outcomes (logistic regression), and ordered-categorical variables (adjacent categories regression model; Balding, 2006). Using a single-SNP approach, "simple regression models or mixed models with a fixed substitution effect

of the SNP genotype along with a polygenic effect" are used to test the association between individual markers and the phenotype (Fernando et al., 2017). The standard method of identifying significant genotype-phenotype associations is to calculate a p-value for the null hypothesis which states that none of the SNP are associated with the phenotype (Stephens and Balding, 2009).

While this methodology has proven successful in identifying genotype-phenotype associations, single-SNP testing has several limitations. During single-SNP testing thousands of variants are individually examined, through multiple comparisons, to identify associations (Johnson et al., 2010; Hong and Park, 2012; Fernando and Garrick, 2013; Yazdani and Dunson, 2015). These analyses often identify few SNPs with small effect leaving a large portion of the genetic variance unaccounted for and multiple testing correction is needed to reduce the number of false positives (genome-wide type I error rate). Linkage disequilibrium between SNP violates the assumption of independent comparisons making standard methods of controlling genome-wide type I error rate (Bonferroni correction and false discovery rate) inappropriate (Johnson et al., 2010; Yazdani, 2014). More specifically, conservative corrections applied to many SNP reduces the statistical power and increases the type II error rate (Johnson et al., 2010; Fernando and Garrick, 2013).

Bayesian approaches, in which all SNP are simultaneously fit in the model as a random effect, were originally developed for genomic selection by Meuwissen et al. (2001). However, previous studies demonstrate that Bayesian models can be applied to GWAS to alleviate single-SNP testing limitations. First, all SNP can be simultaneously tested for an association which eliminates issues associated with multiple hypothesis testing and allows researchers to account for a larger proportion of genetic variance (Fernando and Garrick, 2013). In doing so, researchers can better understand the biological system because rarely does a single SNP form the genetic basis of

a phenotype. Second, Bayesian approaches present an opportunity for multiple testing corrections that do not negatively impact statistical power (i.e., control the proportion of false positives; Fernando and Garrick, 2013).

Bayesian Inference

Bayes' theorem was first introduced in Thomas Bayes' paper, *An Essay Towards Solving a Problem in the Doctrine of Chances* (Bolstad and Curran, 2017). As described by Bolstad and Curran (2017), Bayes' theorem "showed how inverse probability could be used to calculate probability of antecedent events from the occurrence of the consequent event." Put simply, Bayes' is used to calculate the probability of an event given evidence or data. Bayes' theorem is represented mathematically as follows:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \propto P(A)P(B|A)$$

Where, A represents the unknown parameter of interest and B represents the observed data. This theorem indicates that the posterior distribution of the unknown parameter given the observed data $\left(P(A|B)\right)$ is equal to the product of the likelihood of the unknown parameter $\left(P(B|A)\right)$ given the observed data and the prior distribution of the unknown parameter $\left(P(A)\right)$ divided by the marginal distribution of the observed data $\left(P(B)\right)$. Bayes' theorem can also be written as the product of the likelihood and the prior distribution (de Vos, 2004). This theorem was later used to develop Bayesian inference, the foundation of Bayesian statistics (Bolstad and Curran, 2017). Bayesian inference is a form of statistical inference in which observed data are used to estimate the probability of a given hypothesis being true (de Vos, 2004).

As discussed in the previous section, Meuwissen et al. (2001) was the first to propose BayesA and BayesB as Bayesian estimation methods for genomic selection. Briefly, using BayesA

all available SNP are included in the model and the prior distribution of SNP effects is based upon the assumption that many QTL have a small effect while few QTL have a moderate to large effect. In comparison, a BayesB model includes only a portion of the available SNP and the prior distribution of SNP effects indicates that most QTL have no effect (π) and very few have a moderate effect (1-π). Both BayesA and BayesB assume unequal variance (van den Berg et al., 2013; Zeng, 2016). Habier et al. (2011) later developed BayesC and BayesC π which have the same prior distribution as BayesB, but these models assume constant variance. Finally, Bayes C has a fixed value for π whereas the π in BayesC π is unknown (Table 2.3; Zeng, 2016).

BayesC estimation methods are recommended when examining novel traits that lack sound prior information because prior assumptions have a lesser impact on BayesC than BayesA and BayesB (Garrick and Fernando, 2013). BayesC has been successfully applied to GWAS in beef cattle. Peters et al. (2012) used a BayesC model to identify QTL associated with growth and ultrasound traits in Brangus cattle. Similar methodology was applied by Peters et al. (2013) to associate genomic windows with first service conception and heifer pregnancy. Richardson et al. (2016) and Zeng (2016) demonstrated the utility of BayesC for estimating the residual and genetic variances of bovine tuberculosis and pulmonary arterial pressure (PAP) measurements, respectively.

As discussed by Dekkers (2012), quantitative trait locus (QTL) detection criteria in Bayesian-based GWAS vary by study. Several studies have reported important genomic regions based upon posterior inclusion probability (PIP) or the proportion of iterations in the Markov chain Monte Carlo (MCMC) chain that included a particular SNP or genomic window (van den Berg et al., 2013; Speidel et al., 2018) whereas other studies have used the proportion of variance explained by a SNP or markers within a genomic window as detection criteria (Fan et al., 2011; Sun et al.,

2011b; Onteru et al., 2012; Peters et al., 2012, 2013). Additionally, detection criteria may involve both PIP and the proportion of genetic variance explained (Wolc et al., 2012).

Table 2.3. Comparison and description of Bayesian Alphabet (Zeng, 2016).

Method	Bayes A	Bayes B	Bayes C	Bayes Cπ
Reference	Meuwissen et al. (2001)	Meuwissen et al. (2001)	Habier et al. (2011)	Habier et al. (2011)
Prior distribution[1]	t distribution	$\begin{cases} 0 \\ t \text{ distribution} \end{cases}$	$\begin{cases} 0 \\ t \text{ distribution} \end{cases}$	$\begin{cases} 0 \\ t \text{ distribution} \end{cases}$
Implication	A large number of SNP of small effect, a small proportion with moderate to large effect	π proportion of SNP with zero effect, (1-π) proportion with moderate to large effect	π proportion of SNP with zero effect, (1-π) proportion with moderate to large effect	π proportion of SNP with zero effect, (1-π) proportion with moderate to large effect
Π	NO	YES	YES	YES
Sample π	NO	NO	NO	YES
Constant variance	NO	NO	YES	YES
Sampler	Gibbs sampling	Metropolis-Hastings	Gibbs sampling	Gibbs sampling

[1]Prior marginal distribution of SNP effects

Selection Methods

  Traditional Selection Methods

Genetic improvement in the livestock industry is based upon the concept of selecting sires and dams that will produce progeny that will outperform the prior generation (Dekkers, 2012). Moreover, "the purpose of selection programs is to accelerate the rate of genetic change or selection response per unit of time, ΔG, toward a given breeding objective" (Van Eenennaam et

al., 2014). As discussed by Dekkers (2012), animals in breeding programs are selected based upon their estimated breeding values (EBV), which represent their breeding potential or genetic merit as a parent. These breeding values are estimated by summing the additive genetic effect of all loci that contribute to the trait of interest. Genetic evaluation programs generate EBV using phenotype data, pedigree information, and best linear unbiased predictions (BLUP) mixed model procedures.

Using these traditional methods, the rate of genetic change is influenced by the accuracy of the prediction, selection intensity, and generation interval, and genetic variation. More specifically, accuracy, selection intensity, and genetic variation are positively associated with the rate of genetic change while generation interval is negatively associated with the rate of genetic gain (Dekkers, 2012; Van Eenennaam et al., 2014). Traditional selection methods based upon EBV have resulted in rapid genetic gain in beef cattle traits that are moderate to highly heritable and easy to measure including: birth weight, weaning weight, yearling weight, mature cow weight, etc. (Miller, 2010; Boichard et al., 2016). In contrast, these selection methods were inefficient for lowly heritable traits or traits that were difficult to measure (Boichard et al., 2016).

Marker Assisted Selection

While traditional selection programs have proven to be effective, they are limited by our ability to collect phenotypic data and make predictions early in an animal's life. These limitations led to the development of marker assisted selection (MAS) techniques, which incorporate marker information into the EBV to increase the accuracy of prediction (Goddard and Hayes, 2007). Fernando and Grossman (1989) described MAS using BLUP as a two-step process: 1) map QTL and 2) obtain EBV using pedigree and QTL information. Although MAS generated a lot of excitement due to its potential, this approach only identified QTL with large effects due to over-conservative multiple testing correction and the limited genetic variance that was explained by the

genotypes representing QTL. Thus, its implementation in the livestock industry has been minimal and very little genetic progress has made using MAS (Dekkers, 2004; Meuwissen et al., 2016).

Genomic Selection

Following the development of marker assisted selection, Meuwissen et al. (2001) proposed a similar approach called genomic selection (Figure 2.2). In genomic selection, genotypes are obtained for a large established reference population with phenotypic data. Single nucleotide polymorphisms effects are estimated using the reference population and a prediction equation is derived. The prediction equation is applied to selection candidates who possess genotypic data but may lack phenotypic data. In applying the equation, molecular breeding values (MBV) also known as genomic estimated breeding values (GEBV) or direct genomic value (DGV) can calculated for the selection candidates (Goddard and Hayes, 2007; Van Eenennaam et al., 2014; Boichard et al., 2016; Meuwissen et al., 2016). As discussed by Van Eenennaam et al. (2014), numerous statistical models, with varying assumptions regarding the distribution of marker effects, have been developed to calculate MBV. This includes: genomic best linear unbiased prediction (GBLUP), Bayesian regression (BayesA, BayesB, BayesCπ), LASSO, Bayesian Lasso, and elastic net (Goddard, 2009; Hastie et al., 2009; Van Eenennaam et al., 2014).

Large training population
- Phenotyped
- Genotyped

Factors that improve accuracy of GS:
- Large number of animals in training population
- Number of SNP markers being genotyped
- Small effective population size
- High heritability of the trait to be improved
- Small number of large-effect QTL influencing trait
- High level of genetic relationship between training and selection candidates

Genotypic and phenotypic information from selected animals used to recalibrate marker effects to maintain prediction equation accuracy

Predict genetic merit of selection candidate within breed based on genotype

Estimate marker effects for phenotyped traits

Predict genetic merit of selection candidate within breed based on genotype

Not predictive in other breeds/strains/lines

Selection candidate population
Genotypes used to predict genetic merit

Van Eenennaam AL, et al. 2014.
Annu. Rev. Anim. Biosci. 2:105–139
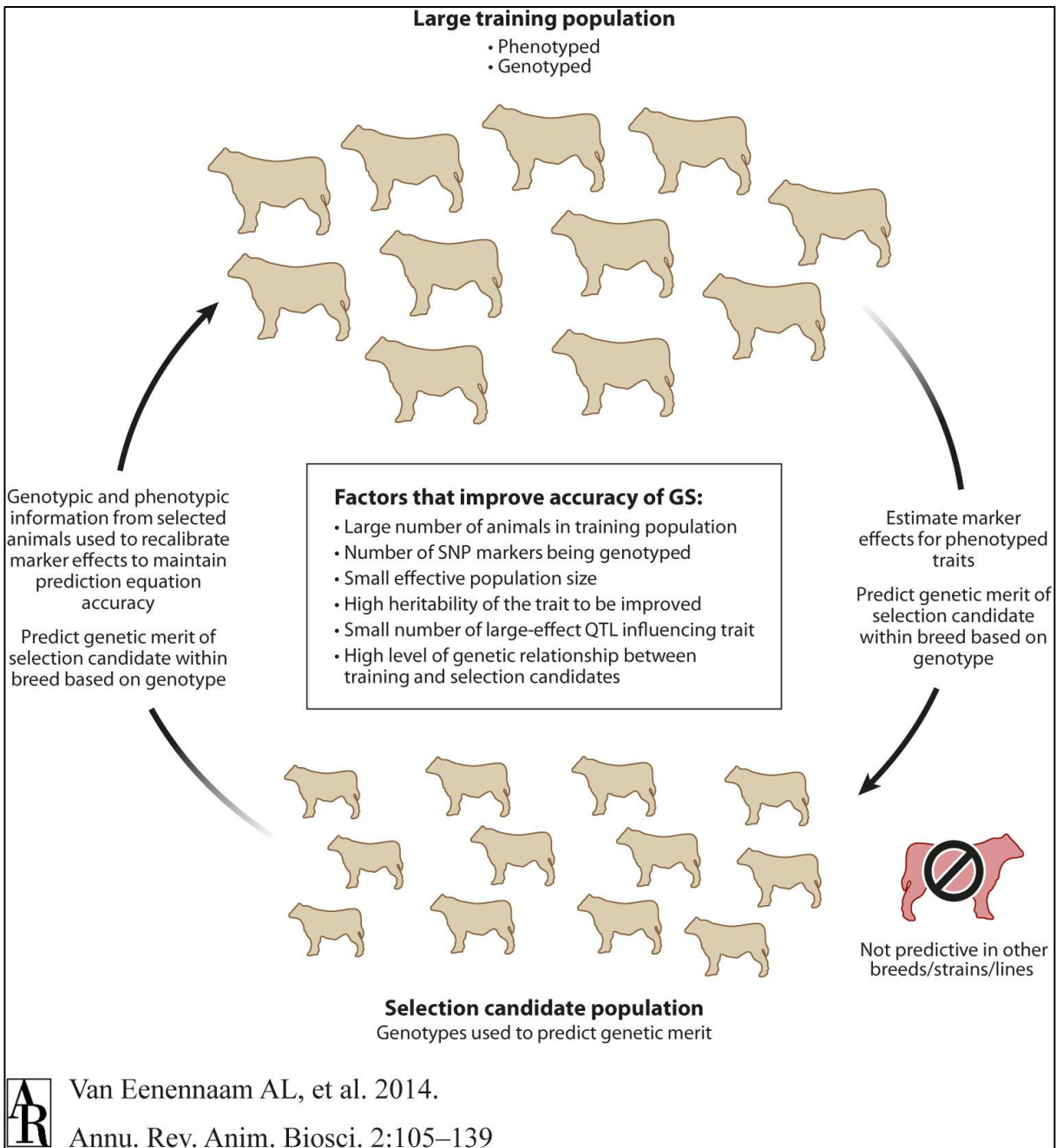
Figure 2.2. Diagram depicting the genomic selection process in livestock (Van Eenennaam et al., 2014).

Following the derivation of MBV, genomic selection can be performed using two approaches: 1) MBV can be used to select superior individuals when there is no established EBV, expected progeny difference (EPD), or indicator traits for the trait of interest (Saatchi and Garrick,

2013), or 2) MBV can be incorporated into genetic evaluations to generate genomic-enhanced estimated breeding values (GE-EBV; Van Eenennaam et al., 2014). The latter can be completed using three different methodologies including a multi-trait approach in which the MBV is fit as a correlated trait (Kachman, 2008), post-evaluation blending which combines EBV/EPD with MBV (Spangler, 2011; Spangler, 2012) and the use of a genomic relationship matrix instead of the traditional pedigree-based relationship matrix (Legarra et al., 2009).

If the population of interest (i.e., selection candidates) includes animals that lack genotype information, multiple-step genomic selection or single-step GBLUP can be used to estimate GE-EBV. In multiple-step genomic selection, genetic prediction is completed in the following manner: 1) estimated breeding values are calculated, 2) pseudo-phenotypes (i.e., phenotypes calculated using records from ungenotyped relatives) are determined for a genotyped population, 3) marker effects are estimated, and 4) total EBV is calculated using EBV and GEBV (Van Eenennaam et al., 2014; Meuwissen et al., 2016). Single-step GBLUP (SSGBLUP) eliminates the need for multiple steps because "all data are accounted for in a single estimation step" (Meuwissen et al., 2016). This method combines genotypic information, pedigree information, and phenotypic records from both genotyped and ungenotyped animals. Moreover, the relationship matrix is formed using both pedigree and genotypic information (Legarra et al., 2009; Christensen and Lund, 2010).

Incorporating genetic information into genetic evaluations may increase the accuracy of existing EBV/EPD and subsequently, increase the accuracy of selection. Moreover, generation interval may be shortened as genetic information can be collected at birth which enables the use of younger bulls (Goddard and Hayes, 2007; Spangler, 2012). As described by Armstrong et al. (1997), genetic information provides opportunity to decrease the age at first breeding and increase

the number of progeny. As previously discussed, increasing the accuracy of selection or decreasing the generation interval will increase the rate of genetic change. However, it is important to note that the accuracy of MBV depend on size of the reference population, how related the reference population is to the selection candidates, effective population size, heritability of the trait, marker density, genomic architecture for the trait, and statistical method (Goddard, 2009).

Genomic selection can have a significant impact on the rate of genetic gain for traits recorded late in an animal's life, sex-limited traits, and difficult-to-measure traits. In fact, in difficult-to-measure traits the rate of genetic gain may increase 20 to 100% following the implementation of genomic selection (van der Werf, 2013). As summarized by Meuwissen et al. (2016), genomic selection has already been incorporated into the dairy and beef industries. In the dairy industry, producers are using genomic selection to improve milk production, fertility, and somatic cell count and in the beef cattle production, researchers are focusing their attention on difficult-to-measure traits such as meat quality, feed efficiency, and disease resistance.

CHAPTER 3: SNP DISCOVERY FOR CAUSITIVE MUTATIONS IN GENES ASSOCIATED

WITH BEEF COW TERRAIN-USE USING RNA-SEQ

Introduction

Concentrated grazing (i.e., large concentrations of cattle grazing in one area) near riparian zones is common in extensive rangeland pastures with rough topography and heterogeneous vegetation. Cattle often prefer to graze lush forage on gentle terrain with minimal grazing occurring in the uplands (Senft et al., 1987; Bailey et al., 1996). Pasture attributes may be modified to promote grazing in rough terrain (e.g., water developments); however, genetic selection provides opportunity for improved pasture utilization at a lower cost than permanent infrastructure. A genome-wide association study identified twelve SNP within five candidate genes (*GRM5, MAML3*, *RUSC2*, *SDHAF3*, and *SUPT20H*) that were associated with beef cattle terrain-use indices. When combined, these SNP explained 34 to 36% of the variation in terrain-use phenotypes which suggested that terrain-use is heritable and polygenic (Bailey et al., 2015). Subsequent studies are needed to further examine these chromosomal loci and identify potential causative mutations for future association analyses.

RNA-seq analysis is a cost-effective alternative to whole genome sequencing that can be used to detect sequence variants in transcribed regions of genes (Cánovas et al., 2010; Piskol et al., 2013). Single nucleotide polymorphisms residing within transcribed exons (e.g. missense variants) can change the base pairs within codons that correspond to the amino acid sequences that form the primary structure of a protein. In altering the protein structure, SNP can alter the biological function of the protein and the resulting phenotype (Klug et al., 2013).

Variant detection using RNA sequencing data has proven successful in human studies as well as domestic livestock including: cattle, horses, sheep, goats, pigs, and chickens (Fortes et al., 2012; Park et al., 2012; Chitwood et al., 2013; Koringa et al., 2013; Zhang et al., 2013; Chen et al., 2015; Wang et al., 2015; Ghosh et al., 2016). In cattle, SNP discovery may reveal variants to be used in marker assisted selection or genomic selection programs to increase the rate of genetic improvement (Pareek et al., 2017). Recently, studies have demonstrated the utility of combining GWAS and RNA-seq analysis to identify variants and candidate genes associated with economically relevant traits in commercial crops and livestock (Fortes et al., 2012; Fortes et al., 2014; Suárez-Vega et al., 2015; Lu et al., 2016).

The objective of this study was to identify potential causative mutations in the five genes that were previously associated with terrain-use indices. Single nucleotide polymorphisms identified in this study may be used to develop a custom genotyping panel for future association analyses with terrain-use phenotypes.

Materials and Methods

All procedures involved in animal handling and management were in accordance with guidelines set forth by the Institutional Animal Care and Use Committee of New Mexico State University (protocol number: 2010-013) and Colorado State University (protocol number: 13-4111A).

Animals

The Angus and Brangus (5/8 Angus x 3/8 Brahman) cattle used in this study were from the John E. Rouse-Colorado State University Beef Improvement Center located near Riverside, Wyoming and the New Mexico State University Brangus breeding program located at the Chihuahuan Desert Rangeland Research Center northeast of Las Cruces, New Mexico. The Angus

steers were the subjects of a high-altitude tolerance study (Cánovas et al., 2016) and the Brangus heifers were from a fertility study (Cánovas et al., 2014); subsequently, animals were not measured for terrain-use traits. However, RNA-sequencing data generated from these studies were used to identify potential causative mutations within five putative candidate genes associated with beef cattle terrain-use.

In the high-altitude tolerance study, 58 tissue samples were collected from six cardiovascular regions on 10 Angus steers: left ventricle, right ventricle, pulmonary artery, aorta, *Longissimus dorsi* muscle, and lung (Cánovas et al., 2016; Table 3.1). In the fertility study, sixty-four tissues were collected from the reproductive system of eight Brangus heifers: hypothalamus, pituitary gland, liver, uterus, endometrium, ovary, adipose tissue, and *Longissimus dorsi* muscle. Two endometrium samples failed laboratory preparation leaving 62 Brangus tissues for analyses (Cánovas et al., 2014; Table 3.1).

Table 3.1. Tissue samples collected from Angus steers and Brangus heifers (Cánovas et al., 2014; Cánovas et al., 2016).

Tissue	Breed	n
Adipose tissue	Brangus	8
Aorta	Angus	10
Endometrium	Brangus	6
Hypothalamus	Brangus	8
Left ventricle	Angus	12
Liver	Brangus	8
Longissimus dorsi muscle	Angus & Brangus	16
Lung	Angus	10
Ovary	Brangus	8
Pituitary gland	Brangus	8
Pulmonary artery	Angus	6
Right ventricle	Angus	12
Uterus	Brangus	8

Ribonucleic Acid Extraction and Sequencing

Ribonucleic acid extraction was completed using Trizol (Invitrogen, Carlsbad, CA) and the TruSeq Stranded mRNA Sample Preparation kit (Illumina, California, USA; Cánovas et al., 2014; Cánovas et al., 2016). As summarized by Cánovas et al. (2014), poly-T oligo attached magnetic beads was used to purify messenger RNA (mRNA) with poly-A tails from total RNA. Messenger RNA was then fragmented and converted to complementary DNA (cDNA) using reverse transcriptase and primers. Following cDNA synthesis, adaptors were ligated to the ends of double-stranded cDNA and PCR amplification was used to generate cDNA libraries. The Illumina HiSeq 2000 analyser generated approximately 30 million single read sequences (100 bp) for each sample.

SNP Discovery using RNA-sequencing analysis

Angus and Brangus sequences were aligned to the annotated bovine reference genome (UMD3.1; release annotation 87) and analysed using CLC Genomics Workbench software (version 9.5; CLC Bio, Aarhus, Denmark) as described by Cánovas et al. (2010). After standard quality control filters were applied, 56 Angus samples and 60 Brangus samples were available for RNA-seq analysis: adipose tissue (n = 8), aorta (n = 8), endometrium (n = 6), hypothalamus (n = 6), left ventricle (n = 12), liver (n = 8), *Longissimus dorsi* muscle (n = 16), lung (n = 10), ovary (n = 8), pituitary gland (n = 8), pulmonary artery (n = 6), right ventricle (n = 12), and uterus (n = 8).

Analysis was performed for each breed using two assembly methods: 1) individual samples and 2) a pool of all samples. Sequence reads were pooled during assembly to increase the number of reads available for alignment (i.e., increase coverage) and improve variant detection (Piskol et al., 2013). Individual sample analysis was performed for each Angus tissue (n = 56) whereas individual sample analysis for Brangus tissues included: hypothalamus, pituitary gland, liver,

adipose, and *Longissimus dorsi* muscle (n = 38). The pooled sample analyses for Angus and Brangus contained all available tissues: Angus (n = 56) and Brangus (n = 60).

Following alignment, variant detection was completed using the Fixed Ploidy Variant Detection tool within CLC Genomics Workbench Software (CLC Bio, Aarhus, Denmark). As described by Dias et al. (2017), "this tool detects germline variants and discards variants when representation in reads is due to sequencing errors or mapping artifacts." Similar variant detection parameters were applied to both the individual sample assemblies and the pooled sample assemblies (Table 3.2). Only SNP located within Glutamate metabotropic receptor 5 (*GRM5),* Mastermind Like Transcriptional Coactivator 3 *(MAML3),* RUN and SH3 domain containing 2 *(RUSC2),* Suppressor of Ty 20 Homolog *(SUPT20H),* and Succinate Dehydrogenase Complex Assembly Factor 3 *(SDHAF3)* were considered due to the previously observed association between these five genes and terrain-use (Table 3.3; Bailey et al., 2015).

Table 3.2. CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) fixed ploidy variant detection parameters for RNA-sequencing data from Angus cardiovascular and Brangus puberty tissues.

Parameter	Individual Samples[1]	Pooled Samples[2]
Ploidy	2	2
Variant probability (%)	90	90
Minimum coverage (reads)	10	20
Minimum count (reads)	2	2
Minimum variant frequency (%)	5	5
Minimum central quality	20	20
Minimum neighbourhood quality	15	15
Relative read direction filter (%)	1	1

[1]Individual assembly method in which RNA sequence reads from an individual tissue are aligned to the reference assembly.
[2]Pooled assembly method in which RNA sequence reads from all tissues are aligned to the reference assembly.

Table 3.3. Description of candidate genes and single nucleotide polymorphisms associated with terrain-use indices in beef cattle.

Chr[1]	Gene	Location[2]	SNP[3]	Position[4]	Mutation	Functional Consequence
4	*SDHAF3*	14364655..14452243	rs134515496	14487987	A/G	Intergenic
8	*RUSC2*	60134638..60196456	rs43555524	60157511	G/A	Intronic
12	*SUPT20H*	24668812..24708465	-	24598260	-	-
12	*SUPT20H*	24668812..24708465	rs110062743	24593452	T/G	Intergenic
17	*MAML3*	17863364..18327762	rs133913408	18318983	A/G	Missense
17	*MAML3*	17863364..18327762	rs109619368	18299593	T/C	Intronic
29	*GRM5*	6598128..7240213	rs42921468	6598207	G/A	Intergenic
29	*GRM5*	6598128..7240213	rs42161939	7083900	C/A	Intergenic
29	*GRM5*	6598128..7240213	rs43744222	7128587	T/C	Synonymous
29	*GRM5*	6598128..7240213	rs210610001	7128668	A/G	Synonymous
29	*GRM5*	6598128..7240213	rs42162705	7240504	A/C	Downstream
29	*GRM5*	6598128..7240213	rs42162708	7241306	C/T	Downstream

[1]Chromosome number.
[2]Chromosome position in *Bos taurus* UMD3.1.1. (Ensembl genome database).
[3]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[4]SNP position according to *Bos taurus* UMD3.1.1.

The Variant Effect Predictor (VEP) tool (McLaren et al., 2016) from the Ensembl genome database (Release 92; Zerbino et al., 2017) was used to determine the functional consequence of the SNP and examine their novelty. Note that SNP may be assigned more than one functional consequence using the VEP tool. To reduce the number of false positive SNP, variants identified using the individual sample approach were compared to the variants identified in the pooled sample approach to identify concordant SNP for subsequent analyses. Similarly, results of the Angus analysis were compared to results of the Brangus analysis. A Venn diagram depicting these comparisons was created using Venny 2.1 (http://bioinfogp.cnb.csic.es/tools/venny/index.html).

Results

Single nucleotide polymorphism discovery in which Angus transcripts from cardiopulmonary tissues were individually aligned to the bovine reference genome revealed 108 SNP within *GRM5, MAML3, RUSC2, SUPT20H,* and *SDHAF3*. The greatest number of variants

(n = 37) was found within RUN and SH3 domain containing 2. In comparison, the Angus pooled sample analysis identified 560 SNP within the five candidate genes. In this analysis, Mastermind Like Transcriptional Coactivator 3 harboured the largest number of SNP (n = 155).

Eighty-six concordant SNP were revealed when Angus individual sample variants were compared to the Angus pooled sample variants. As reported by the Ensembl genome database VEP, 61 of these variants were novel whereas 25 were existing or known variants. Moreover, 79% of the SNP were intronic, 6% were downstream variants, 5% were synonymous, 4% were missense, 2% were splice acceptors, 2% were splice region variants, and 1% were in three prime untranslated regions (3'-UTR). One of the four missense SNP was in *MAML3* (rs109371446) and three were in *RUSC2* (rs467493459, rs43556445, and rs448324087).

The individual sample analysis using Brangus tissues revealed 235 SNP in the five candidate genes and the pooled sample analysis identified 1,090 SNP. In accordance with the Angus results, *RUSC2* (n = 69) and *MAML3* (n = 288) contained the largest number of variants in the individual and pooled analysis, respectively. One hundred and eighty SNP were identified using both assembly approaches. Of these 180 SNP, 78 were previously documented and 102 were considered novel. As with the Angus variants, the Brangus variants had a wide-range of functional consequences: 81% intronic, 7% synonymous, 6% downstream variants, 4% missense, 2% splice region variants, and 1% splice acceptor variants. The missense SNP were associated with *MAML3* (rs480743060, rs133913408, and rs109371446), *RUSC2* (rs43556445, rs208172401, and rs517656634), and *SUPT20H* (rs134305602).

A comparison of the Angus and Brangus results revealed 48 congruent single nucleotide polymorphisms (Figure 3.1). The greatest number of SNP were found within *GRM5* (n = 19), while *MAML3* contained the lowest number of variants (n = 2). A further examination of these

48 SNP, using the Ensembl genome database VEP (McLaren et al., 2016), suggested that 34 variants were novel and 14 had been previously identified. Functional consequences of these single nucleotide polymorphisms followed the same trend as those previously discussed. Many SNP were classified as intronic (81.25%), several SNP were considered exonic (12.5%), and three SNP had more than one functional consequence (6.25%; Table 3.4). Missense SNP were located within *MAML3* (rs109371446) and *RUSC2* (rs43556445).

Table 3.4. Functional consequences of the 48 concordant single nucleotide polymorphisms.

Functional Consequence[1]	Number of SNP
Downstream gene, Splice Acceptor	1
Intronic	39
Missense	2
Synonymous	4
Splice Region, Intronic	2

[1]Variants may be assigned more than one functional consequence.
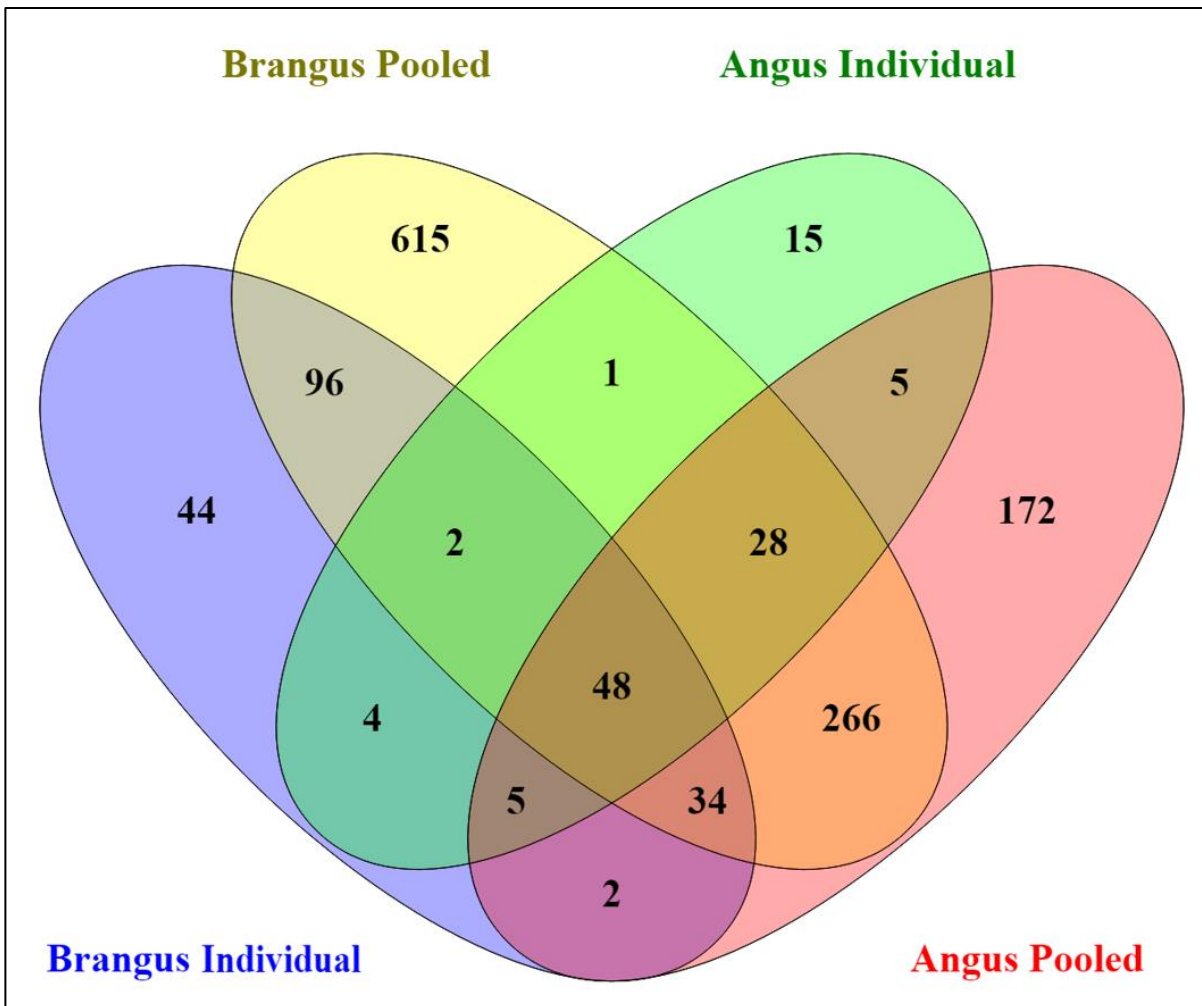
Figure 3.1. A comparison of the single nucleotide polymorphisms discovered using RNA-sequencing data from two breeds of cattle (Angus & Brangus) and two assembly approaches (Individual & Pooled).

Discussion

Numerous studies have demonstrated that RNA-seq data can be used for variant detection in domestic livestock. Moreover, performing RNA-seq analysis alongside GWAS provides opportunity for validation and functional analysis of candidate genes. In the current study, RNA-seq data obtained from Angus tissue samples revealed 108 SNP when an individual sample approach was applied and 560 SNP when a pooled sample approach was used. These results were consistent with the assertion that pooling sequence reads results in increased power for variant

detection. Single nucleotide polymorphism prevalence varied by gene which was expected given the results of previous studies (Lehne et al., 2011; Dias et al., 2017).

A comparison of the two assembly approaches revealed 86 concordant variants. Functional analysis of the 86 concordant variants revealed many intronic SNP and very few exonic SNP. Variants located in noncoding regions may affect the phenotype through alternative splicing; however, RNA splicing removes introns from precursor mRNA (pre-mRNA) prior to it becoming mRNA and so SNP discovered using RNA-seq analysis are expected to be exonic (Wang and Cooper, 2007; Klug et al., 2013). The high prevalence of intronic SNP in this RNA-seq analysis may be attributed to a multitude of factors including: poor reference assembly sequencing/annotation and isolation of pre-mRNA during mRNA extraction.

Since the release of the first bovine reference genome in 2009, numerous studies have generated results that suggest the presence of assembly and annotation errors in both UMD and BTAU (Elsik et al., 2009; Bohmanova et al., 2010; Zhou et al., 2015; Utsunomiya et al., 2016). Furthermore, assembly statistics for UMD3.1.1 included 69,281 spanned gaps and 3,193 un-spanned gaps in which sequence information is missing (NCBI, 2018b). These errors and gaps in the reference genome can have a substantial effect on RNA-seq analyses that call variants based upon sequencing reads aligned to a reference genome. Future analyses should be completed using the new bovine reference assembly ARS-UCD1.2 that contains updated locations for candidate genes (Table 3.5) and only 386 gaps (NCBI, 2018a).

Table 3.5. Comparison of candidate gene locations in bovine reference assembly UMD3.1.1 and ARS-UCD1.2.

| Candidate Gene | Chr[1] | Location[2] (bp) | |
		UMD3.1.1	ARS-UCD1.2
SDHAF3	4	14364655..14452243	14476714..14567927
RUSC2	8	60134638..60196456	59723630..59785507
SUPT20H	12	24668812..2470846	24646264..24685744
MAML3	17	17863364..18327762	17559644..18020917
GRM5	29	6598128..7240213	6557063..7201730

[1]Chromosome number.
[2]Location in *Bos taurus* reference genome according to National Center for Biotechnology Information (NCBI).

In addition to the reference assembly, the RNA extraction method may have contributed to the number of intronic SNP identified in the RNA-seq data. As discussed by Yousefi et al. (2018), intronic SNP are often identified in RNA-seq analyses because pre-mRNA (containing introns) may be captured during mRNA isolation. More specifically, RNA extraction methods that purify mRNA molecules based upon the presence of a poly-A tail, also known as poly(A) capturing protocols, may capture pre-mRNA (Piskol et al., 2013).

Variant detection using Brangus tissue samples generated results similar to those in the Angus analysis. The individual assembly method (n = 235) identified fewer SNP than the pooled assembly method (n = 1, 090) and the candidate genes containing the largest number of SNP were *RUSC2* and *MAML3*. Likewise, most of the Brangus SNP were classified as intronic which warrants subsequent analyses. One significant difference between the Angus and Brangus analyses was the total number of variants detected. The Brangus RNA-seq data contained twice as many SNP in the candidate genes as the Angus data. This result may be attributed to the breed composition of Brangus, 3/8 Brahman and 5/8 Angus. Variant detection using *Bos taurus* and *Bos indicus* cattle and candidate genes associated with growth and development revealed that Brangus

cattle had a greater number of SNP than both Angus and Brahman cattle. Moreover, Brangus cattle had both Angus derived SNP and Brahman SNP (Figure 3.2; Figure unpublished).

Forty-eight SNP were identified in the four analyses which suggested that they are true variants and not false positives due to technical error. Nine of the forty-eight SNP were considered exonic; however, only those classified as missense, splice region, or splice acceptor/donor were of interest. Synonymous variants encode the same amino acid which implies that they are not causative mutations. In contrast, nonsynonymous variants alter the amino acid sequence and the resulting protein structure, which suggests that these SNP would likely influence the phenotype (Koufariotis et al., 2014; Iso-Touru et al., 2016). As discussed by Iso-Touru et al. (2016), SNP located in splice sites or flanking regions containing regulatory elements (i.e., upstream and downstream) may also have a large influence on the phenotype.

Missense variants were located within *MAML3* and *RUSC2* while splice acceptor and splice region variants were identified in *SUPT20H* and *RUSC2,* respectively. Mastermind Like Transcriptional Coactivator 3 was found within the Notch signaling pathway, a highly-conserved pathway associated with cell fate determination during metazoan development and tissue renewal (Kopan and Ilagan, 2009). Wu et al. (2002) suggested that *MAML3* expression may influence Notch signalling and the biological processes regulated by the Notch signalling pathway: proliferation, apoptosis, neurogenesis, myogenesis, vasculogenesis, and other similar processes. Previous studies suggested that neurogenesis frequently occurs in the hippocampus of adult mammals; therefore, it has been proposed that neurogenesis is important for learning and memory (Gross, 2000). As discussed by Bailey et al. (1996), learning and memory influence diet and feeding site selection for large herbivores.

RUN and SH3 domain containing 2 codes for iporin, a protein that interacts with Rab1b and Rab1-binding protein GM130 that regulate intracellular vesicle transport (Bayer et al., 2005). Vesicle transport is vital for maintaining cell function as it enables communication between cellular compartments (Bhuin and Roy, 2014). Suppressor of Ty 20 Homolog has a role in the autophagy pathway that is stimulated during periods of cellular stress (e.g., hypoxia, nutrient deprivation; Azad et al., 2008; Gatica et al., 2015). More specifically, *SUPT20H* is involved in the activation of *ATG9* (Gatica et al., 2015) which is involved in the formation of autophagosomes (Papinski et al., 2014).
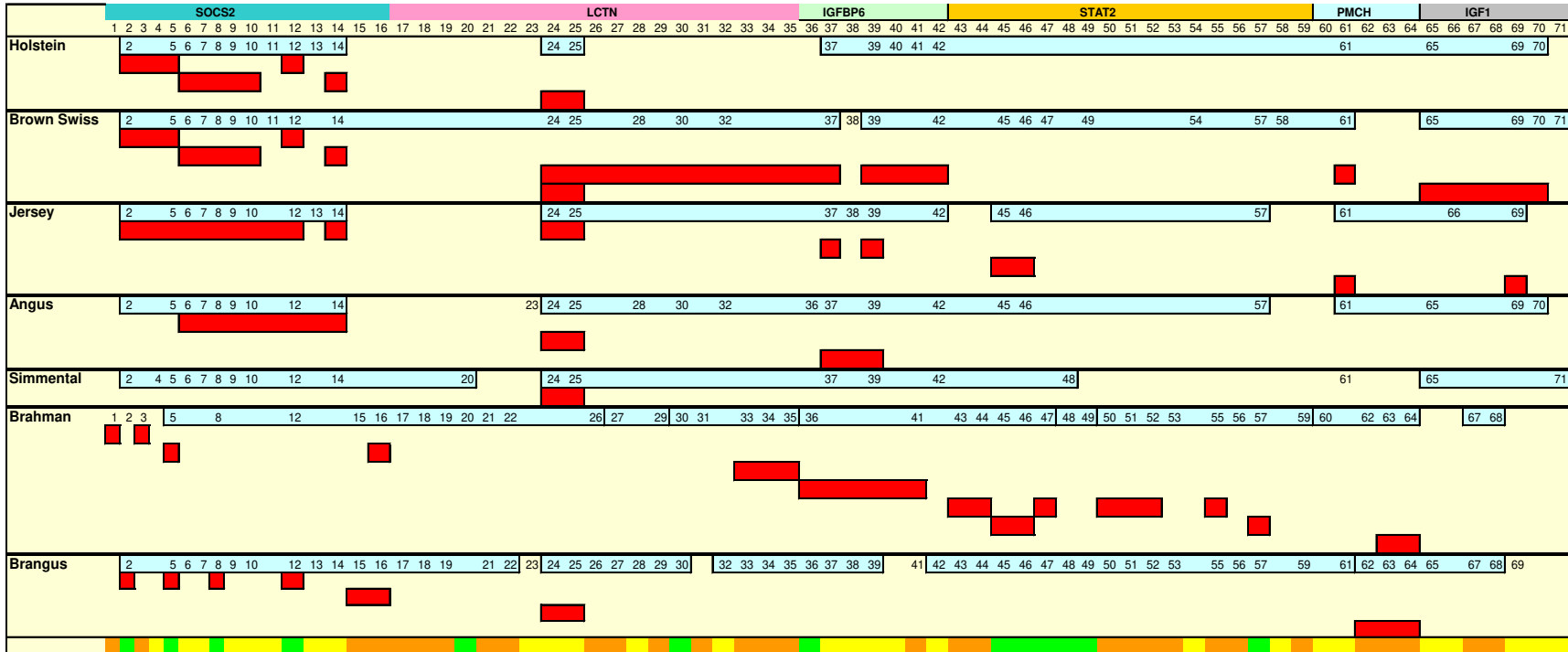
Figure 3.2. Diagram depicting single nucleotide polymorphism (SNP) discovery results for five candidate genes on BTA5 in *Bos taurus* and *Bos indicus* breeds. Blue bars indicate linkage disequilibrium (LD) given the four-gamete rule and red bars indicate LD according to Lewontin's D'. *Bos taurus* SNP represented by yellow bar, *Bos indicus* SNP represented by orange bar, and *Bos taurus/indicus* SNP represented by green bar (Figure unpublished).

Conclusion

Variant detection using RNA-seq data from Angus and Brangus cattle revealed 48 SNP in five candidate genes that were previously associated with beef cow terrain-use. Of the 48 concordant SNP, 39 (81.25%) were considered intronic and 6 (12.5%) were characterized as exonic. Missense variants were identified within Mastermind Like Transcriptional Coactivator 3; however, additional analyses using the developing annotation for the ARS-UCD1.2 assembly should be completed to confirm the functional consequences of these SNP.

CHAPTER 4: INVESTIGATION OF CANDIDATE SNP PREVIOUSLY

ASSOCIATED WITH TERRAIN-USE INDICES USING BAYESIAN-BASED GENOTYPE-

PHENOTYPE ASSOCIATIONS

Introduction

Beef cattle breeding programs were successful in implementing pedigree-based selection for easy-to-measure traits; however, these methods were inefficient for sex-limited traits, traits measured late in life, and difficult-to-measure traits. Genomic selection was developed as a means of improving artificial selection programs by incorporating single nucleotide polymorphism (SNP) data into estimated breeding value (EBV) predictions. Genomic selection may improve the rate of genetic gain for difficult-to-measure traits because phenotypic observations do not have to be collected for each animal and estimated breeding values (EBV) can be calculated for traits recorded in a reference population (Goddard and Hayes, 2007; Boichard et al., 2016; Meuwissen et al., 2016).

Terrain-use is an example of a difficult-to-measure trait that may potentially be improved using genomic selection methods. Terrain-use measurements are typically derived from global positioning system (GPS) data collected from cattle grazing in extensive rangeland pastures (Bailey et al., 2015; Bailey et al., 2018). Global positioning system collars that are used to obtain the terrain-use data are costly and therefore, make it difficult to obtain an adequate sample size (Bailey et al., 2018). Furthermore, rangeland beef operations are typically composed of commercial cattle (i.e., not registered with a breed association); therefore there is no pedigree information for use in calculation of EBV (Bailey et al., 2015).

Previous studies that examined the grazing patterns of cattle in the western United States (U.S.) suggested that terrain-use is a moderately heritable, polygenic trait (Bailey et al., 2015). Single-marker regression analyses using data from 80 beef cattle revealed four candidate genes located on *Bos taurus* autosome four, twelve, seventeen, and twenty-nine that were associated with terrain-use indices (Bailey et al., 2015). Additional studies are needed to confirm these results prior to the development of molecular breeding values (MBV) for selection.

Previous studies, using both real and simulated data, have demonstrated the utility of combining several statistical methods including linkage disequilibrium haplotype-based analysis, single-marker association analysis, Bayesian regression, and weighted single-step GBLUP to decrease the number of false positives and increase the power of detecting QTL (Legarra et al., 2015; Melo et al., 2017). Due to the polygenic nature of quantitative traits, multi-SNP models may better explain the genetic architecture underlying terrain-use than single-SNP analyses (van den Berg et al., 2013). Therefore, the objective of this study was to investigate the five SNP identified by Bailey et al. (2015) using Bayesian-based genotype-phenotype associations.

Materials and Methods

High-Density Genotypes and Global Positioning System Data

As described by Bailey et al. (2015), data were obtained from 71 mature cows and 9 yearling heifers managed on five ranches in Arizona, Montana, and New Mexico. This included: Chihuahuan Desert Rangeland Research Center managed by New Mexico State University (CDRRC; Las Cruces, NM), Corona Range and Livestock Center managed by New Mexico State University (Corona, NM), Hartley Ranch (Roy, NM), Thackeray Ranch managed by Montana State University (Havre, MT), and Todd Ranch (Willcox, AZ). Heterogenous ranches (i.e., varied topography and vegetation) were selected for the study to increase the likelihood of identifying

candidate SNP that influenced grazing patterns across varying terrain (Table 4.1). The data

included high-density genotypes (777,962 SNP) and global positioning system (GPS) data from

Lotek GPS 3300 collars (Lotek Wireless, New Market, Ontario, Canada).

Table 4.1. Terrain characteristics of the five ranches used in the study where 80 beef cattle were managed during the study period in which they were tracked using global positioning system collars (Bailey et al., 2015).

State	Ranch	n[1]	Breed[2]	Pasture Size[3], ha	Elevation, m	Slope, %	Max Dist. Water[4], km
AZ	Todd	16	Limousin	9065	1276-2010	1-130	4.8
MT	Thackeray	17	Simmental cross	336	1170-1400	0-107	1.5
NM	CDRRC	16	Brangus	3990, 2830	1250-1402	1-15	10.0
NM	Corona	22	Angus Angus cross	1601, 721	1765-1851	0-32	4.7
NM	Hartley	9	Angus Angus cross	1056	1500-1670	0-200	4.3

[1]The number of cattle tracked at each ranch.
[2]The Corona and Hartley Ranch maintained two breeds of cattle.
[3]Cattle grazed in two pastures at the CDDRC and Corona Ranch during the study period.
[4]Maximum distance cows may travel from water in the pasture.

Quality control of genotype data was applied using PLINK 1.9 (Purcell and Chang, 2015)

and standard filters: sample call rate ≥ 0.90, SNP call rate ≥ 0.90, minor allele frequency < 0.01,

and Hardy-Weinberg Equilibrium < 0.0001. After quality control, 75 animals and 728,751 SNP

were available for analysis. Genotypes were recoded from the AB format to numerical values (AA

[-1], AB [0], and BB [1]) and missing genotypes were filled with the median for that locus.

Terrain-use measurements including elevation (m) and slope (%) were derived for each

collar coordinate using USGS Digital Elevation Maps (DEMs) with 10-meter spatial resolution

and ArcGIS Spatial Analyst (Redlands, CA). Similarly, distance travelled from water (m) was

calculated using USGS DEMs with 5-meter grid resolution and the ArcGIS Euclidean Distance

tool (Redlands, CA) which calculated the distance between the GPS coordinate and the primary

water source in the pasture. The measurements for slope, elevation, and distance travelled from water were averaged, across the study period, for each cow (Table 4.2) and then incorporated into two previously developed terrain-use indices (Bailey et al., 2015). The rough index was a normalized average of slope and elevation that was calculated using the following mathematical equation:

$$\text{Rough Index} = \frac{\left[((\frac{\text{slope}_k}{\text{slope}_l}) * 100) + ((\frac{\text{elevation}_k}{\text{elevation}_l}) * 100) \right]}{2}$$

where k represented the average observation of a collared cow and l represented the average observation of all collared cows at a given ranch. The rolling index was a normalized average of slope, elevation, and distance travelled from water:

$$\text{Rolling Index} = \frac{\left[\left(\left(\frac{\text{slope}_k}{\text{slope}_l}\right) * 100 \right) + \left(\left(\frac{\text{elevation}_k}{\text{elevation}_l}\right) * 100 \right) + \left(\left(\frac{\text{distance from water}_k}{\text{distance from water}_l}\right) * 100 \right) \right]}{3}$$

where k represented the average observation of a collared cow and l represented the average observation of all collared cows at a given ranch. For both the rough and rolling index, cattle with values greater than 100 can be classified as hill climbers whereas values less than 100 suggest that the animal was a bottom dweller. When compared to their contemporaries, hill climbers utilize steeper slopes, higher elevations, and more remote areas of the pasture. In contrast, bottom dwellers prefer gentle slopes, lower elevations, and forage near water sources.

Table 4.2. Terrain-use traits, derived from global positioning system (GPS) data collected from 75 beef cows in rangeland pastures in the western U.S., averaged for each ranch.

Ranch	n	Slope[1], %		Elevation[2], m		Distance from Water[3], m	
		Mean	SD[4]	Mean	SD[4]	Mean	SD[4]
CDRRC	16	4.3	0.31	1313.2	1.2	1246.8	145.0
Corona	19	3.7	0.31	1780.4	6.8	1807.3	336.1
Hartley	8	13.6	3.68	1590.3	46.8	751.4	285.7
Thackeray	16	18.9	1.43	1283.2	7.8	470.0	40.0
Todd	16	7.6	1.82	1394.6	51.0	975.8	147.1

[1]Average slope recorded for the cow during the tracking period.
[2]Average elevation recorded for the cow during the tracking period.
[3]Average distance the cow travelled from water during the tracking period.
[4]Standard deviation.

Genotype-Phenotype Association Analyses – All Available Markers

Association analyses were performed using the BOLT software package (Release 1.2.7; http://www.thetasolutionsllc.com/bolt-software.html) and BayesC methodology described by Habier et al. (2011). Analyses were conducted for each of the candidate *Bos taurus* autosomes (BTA 4, 12, 17, 29) described by Bailey et al. (2015) and rough and rolling index values were used as phenotypes (Table 4.4). Single nucleotide polymorphisms were simultaneously fit in the following statistical model:

$$y = X\beta + \sum_{k=1}^{K} Za + e$$

where y was the vector of rough or rolling index values, X was the incidence matrix relating fixed effects to the observations in y, β was the vector of fixed effect solutions, K was the number of SNP in the analysis, Z was the vector of genotype covariates for SNP k (coded -1, 0, 1), a was the vector of random allele substitution effects, and e was the vector of residual effects. Model selection was performed using backward selection in which alpha was set at 0.05. The least

significant predictor was removed from the full model until only significant predictors ($P < 0.05$)

remained. Breed, terrain type (mountainous or rolling), and season (spring, summer, fall, winter)

were fit as fixed effects in the full model; however, these effects were linearly dependent with

ranch and therefore, were removed from the final model. Ranch was not significant for the model

containing rolling index observations ($P = 0.693$) or rough index observations ($P = 0.989$);

however, ranch was biologically significant and therefore, it was included as a fixed effect in the

models. This is in agreement with the models described by Bailey et al. (2015). Since autosomes

were examined individually, the number of makers varied per analysis (Table 4.3).

A BayesC model assumes that a proportion of SNP have no effect (π) and a proportion of

SNP have a nonzero effect ($1-\pi$). Single nucleotide polymorphism effects are normally distributed

and have a common variance denoted by σ_α^2 (Fernando et al., 2017). For this analysis, π was 0.995;

therefore, the proportion of SNP with an effect ($1-\pi$) was 0.005. As discussed by Garrick and

Fernando (2013), when examining a novel trait with BayesC, π should be selected to ensure that

the number of SNP fit in the model during each iteration does not exceed the number of

observations. The variance of SNP effects was expressed as:

$$\sigma_\alpha^2 = \frac{c\sigma_u^2}{2k(1-\pi)\overline{pq}}$$

where c was the proportion of genetic variance accounted for by the SNP effects, σ_u^2 was the

additive genetic variance, k was the number SNP in the genotype matrix, π was the proportion of

SNP with null effect, and $\overline{pq}$ was the average of the product of the p and q loci frequencies. Table

4.3 provides each of the previously discussed parameters as well as the heritability estimates from

Bailey et al. (2015) that were used to estimate the additive genetic variance of the current data.

Table 4.3. Description of parameters used to calculate the variance of marker effects for Bayesian-based association studies with two terrain-use phenotypes and four *Bos taurus* autosomes.

Parameter[1]	Rolling Index			Rough Index			
	BTA4	BTA17	BTA29	BTA4	BTA12	BTA17	BTA29
h^2	0.34	034	0.34	0.36	0.36	0.36	0.36
σ^2_P	31.27	31.27	31.27	75.11	75.11	75.11	75.11
c	1	1	1	1	1	1	1
σ^2_u	10.63	10.63	10.63	27.04	27.04	27.04	27.04
k	33,252	21,160	13,887	33,252	24,729	21,160	13,887
π	0.995	0.995	0.995	0.995	0.995	0.995	0.995
$\overline{pq}$	0.25	0.25	0.25	0.25	0.25	0.25	0.25

[1]h^2 = heritability estimate (Bailey et al., 2015); σ^2_P = phenotypic variance; c = proportion of genetic variance explained by SNP effects; σ^2_u = additive genetic variance; k = number of SNP, π = proportion of markers with null effect; $\overline{pq}$ = average of the product of the p and q loci frequencies.

Single-site Gibbs sampling with 150,000 iterations was used to obtain the posterior mean of the allele substitution effects, the posterior variance of the allele substitution effects and the number of times the marker entered the model when sampled (marker count). Pre-conditioned conjugate gradient (PCG) solutions were used as starting values for the Markov chain Monte Carlo (MCMC) sampling chain, eliminating the need for a burn-in (Golden and Garrick, 2016). Posterior inclusion probability (PIP) or "the proportion of iterations that included a specific marker in the model" was calculated for each marker by dividing the marker count by the total number of iterations (van den Berg et al., 2013). As discussed by Yi et al. (2003), SNP with relatively large effects will appear more frequently during Gibbs sampling; therefore PIP can be used to identify important markers or genomic regions. Moreover, simulations conducted by Moser et al. (2015) suggest that high PIP values may be linked to causal variants or variants associated with the casual variant. Rstudio (version 3.3.2) was used to generate Manhattan plots and the SNP with the highest PIP were identified for every analysis. These SNP were considered the "top SNP" during the study and they were compared to the candidate SNP identified by Bailey et al. (2015).

Table 4.4. Description of candidate single nucleotide polymorphisms (SNP) associated with beef cattle terrain-use indices developed by Bailey et al. (2015)

RS ID[1]	CHR[2]	Position[3]	Gene[4]	Phenotype[5]
rs134515496	4	14487987	*SDHAF3*	Rough & Rolling
rs110062743	12	24593452	*SUPT20H*	Rough
rs109619368	17	18299593	*MAML3*	Rough & Rolling
rs42161939	29	7083900	*GRM5*	Rough & Rolling
rs43744222	29	7128587	*GRM5*	Rough & Rolling

[1]Reference SNP cluster identification assigned by the National Center for Biotechnology Information (NCBI).
[2]Autosome in which the SNP was located according to *Bos taurus* UMD3.1.1.
[3]Autosome position in *Bos taurus* UMD3.1.1.
[4]Positional candidate gene associated with the SNP.
[5]Phenotype that the SNP was associated with: rough or rolling index.

Genotype-Phenotype Association Analyses – One Megabase Windows

On each autosome (BTA 4, 12, 17, 29), consecutive one megabase (Mb) genomic windows were derived and then the genomic windows containing the candidate SNP discovered by Bailey et al. (2015) were identified (Table 4.5). Genotype-phenotype associations were calculated using these one Mb windows and the rolling and rough index values. The analyses were run with BayesC methodology and the same statistical model as the GWAS using all markers ($k \geq 13,887$). Briefly, ranch was fit as a fixed effect and markers were simultaneously fit as a random effect. The MCMC algorithm was run with 150,000 iterations and $\pi = 0.995$. The variance of marker effects was calculated using the same formula; however, *k* was adjusted for each analysis to account for the reduction in markers.

Table 4.5. Description of one megabase genomic windows containing the SNP that were previously associated with beef cattle terrain-use indices.

Window[1]	RS ID[2]	Number of SNP[3]	Start (bp)[4]	End (bp)[4]
4_14	rs134515496	202	14000232	14988405
12_24	rs110062743	352	24002782	24995225
17_18	rs109619368	272	18004299	18998937
29_7	rs42161939, rs43744222	284	7002538	7996355

[1]Genomic window: autosome and n^{th} 1 Mb window on that autosome.
[2]Reference SNP cluster identification, assigned by National Center for Biotechnology Information (NCBI), for the candidate SNP within the window.
[3]Number of SNP within the window.
[4]Start of the 1 Mb window (bp).
[5]End of the 1 Mb window (bp).

Results

The association analyses, in which the rough index values were the phenotype, provided evidence to support one of the five candidate SNP that were identified by Bailey et al. (2015) . The Manhattan plot that depicts the results of these four analyses (Figure 4.1) revealed a small peak on autosome four. The single nucleotide polymorphism with the highest PIP on this peak was rs110225180 (PIP = 0.011; Table 4.6). A posterior inclusion probability of 0.011 or 1.1% indicates that the SNP was included in the model in 1,650 iterations out of the 150,000 iterations that were ran for the Gibbs sampler. The top SNP on autosome four, rs110225180, was 86.3 Mb downstream of the candidate SNP that was previously identified on BTA4 (rs134515496) which had a PIP of 0.006. Five markers were identified on autosome 12 including: rs42557694 (PIP = 0.017), rs136319514 (PIP = 0.015), rs109164448 (PIP = 0.015), rs110062743 (PIP = 0.015), and rs110450498 (PIP = 0.014). The top SNP (rs42557694) was 21.8 Mb downstream of the SNP previously associated with terrain-use (rs109164448). The candidate SNP located on autosome 17 (rs109619368) was confirmed in this study (i.e., it was the top SNP on BTA17) with a PIP of 0.015. Finally, the top three SNP on autosome 29 were rs42175994 (PIP = 0.020), rs110626028

(PIP = 0.018) and rs42161939 (PIP = 0.018). Bailey et al. (2015) reported an association between rs42161939 and both the rough and rolling indices. The second candidate SNP for this autosome, rs43744222, had a PIP of 0.016 or 0.16%.

The second analysis, in which the rolling index values were the phenotype, failed to confirm the associations between the candidate SNP and the terrain-use indices. As with the rough index analysis, the Manhattan plot for the rolling index revealed small peaks on autosome four (Figure 4.2). The SNP that was previously reported on autosome four (rs134515496) had a posterior inclusion probability of 0.008 whereas the top SNP in this study (rs110340473) had a PIP of 0.010 (Table 4.7). Results of autosome 17 were similar to those of autosome four. The candidate locus identified in 2015 (rs109619368) had a relatively small PIP (0.009) compared to the locus with the highest PIP in this study (rs41637536; PIP = 0.020). The analysis using makers on autosome 29 identified three SNP of interest (rs42190442, rs42245670, and rs135313512); however, none of these SNP were identified by Bailey et al. (2015). Furthermore, none of these markers were in close proximity with the two candidate SNP discovered in 2015 ($\geq$ 15 Mb).

The association analyses performed using the one Mb genomic windows validated the five previously identified candidate SNP (Bailey et al., 2015) and their underlying genes (Figures 4.3 and 4.4). On autosome four, rs134515496 and rs133330297 had PIP of 0.11 when associated with the rough index. Similarly, rs134515496 had the highest PIP (0.54) when rolling index values were used as the phenotype. On autosome 12, rs109164448 (PIP = 0.26), rs136319514 (PIP = 0.25), rs110450498 (PIP = 0.25), and rs110062743 (PIP = 0.23) were the top single nucleotide polymorphisms. While the SNP identified in 2015 (rs110062743) did not have the highest PIP, these four SNP were within a 0.007 Mb window which may suggest that markers are in linkage disequilibrium. As discussed by Bush and Moore (2012), "genotyped SNPs often lie in a region of

high linkage disequilibrium with an influential allele." For both the rough and rolling indices, rs109619368 had the highest PIP on autosome 17. The posterior inclusion probability for this SNP was higher for the rough index analysis (PIP = 0.61) as compared to the rolling index analysis (PIP = 0.19). The two candidate SNP previously identified on autosome 29 were part of the top three SNP for the rough index phenotype: rs42161939 (PIP = 0.74), rs134606703 (PIP = 0.21), and rs43744222 (PIP = 0.20). The rolling index association analysis revealed similar results in which rs42161939 and rs43744222 had PIP of 0.40 and 0.17, respectively.
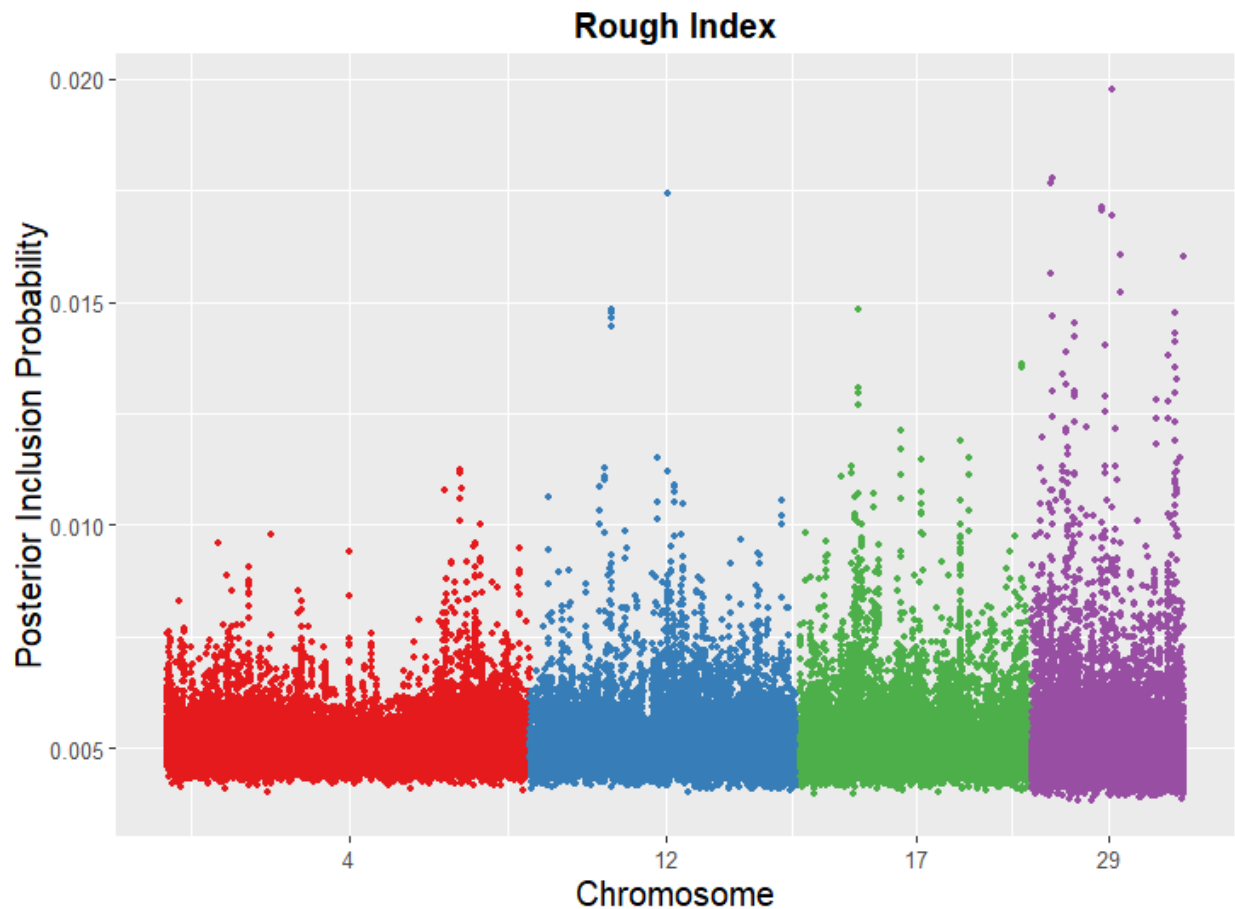

Figure 4.1. Manhattan plot of a genome-wide association study for the rough index values of 75 cows managed in Arizona, Montana, and New Mexico.
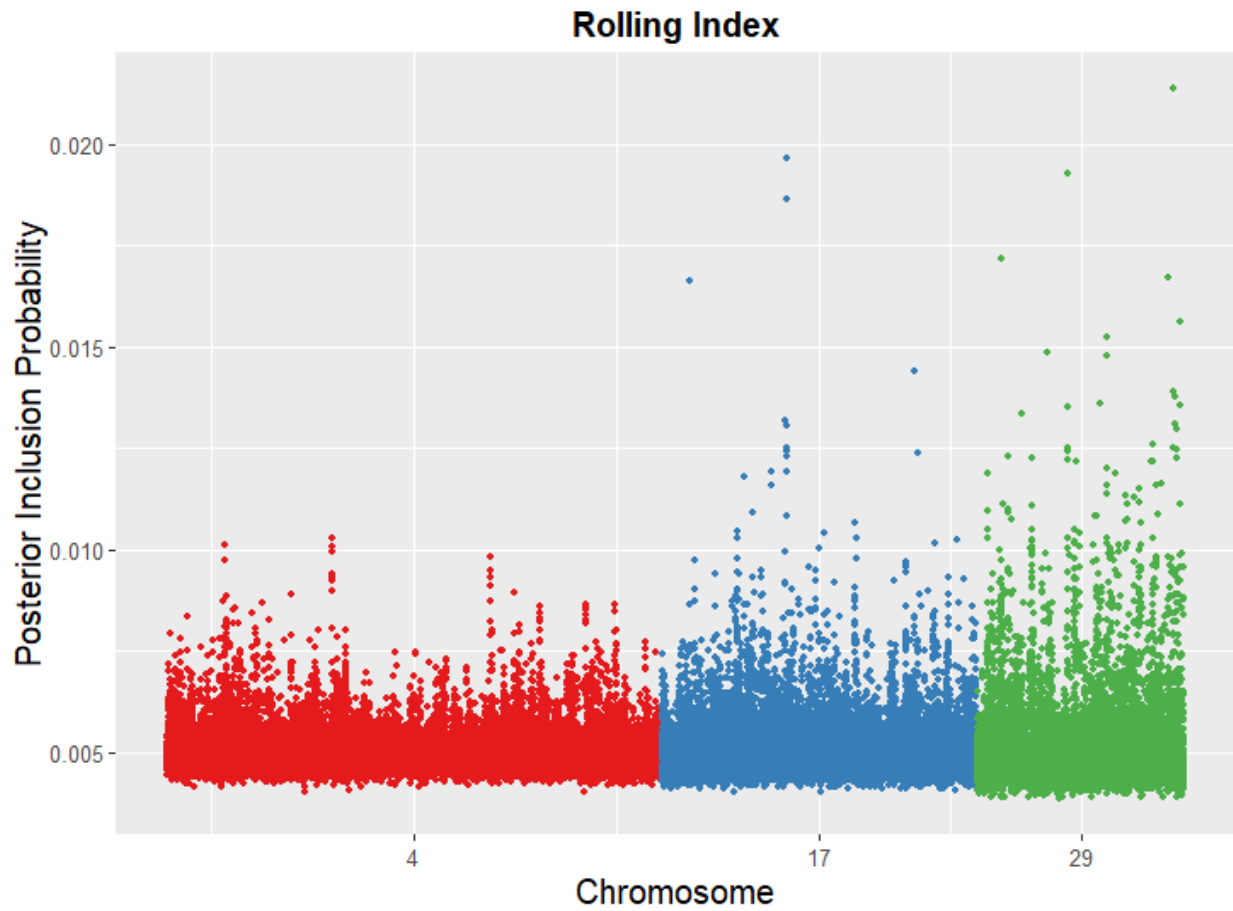
Figure 4.2. Manhattan plot of a genome-wide association study for the rolling index values of 75 cows managed in Arizona, Montana, and New Mexico.

Table 4.6. Single nucleotide polymorphisms with the highest posterior inclusion probabilities, on autosome 4, 12, 17, and 29, in a genome-wide association study using rough index values of beef cows managed in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Gene[4]	SNP location[5]	PIP[6]
rs110225180	4	100820951	*MTPN*	0.194 Mb	0.011
rs42557694	12	45411983	*ENSBTAG00000046942*	0.588 Mb	0.017
rs109619368	17	18299593	*MAML3*	Intron	0.015
rs42175994	29	26048742	*ZDHHC13*	Intron	0.020

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Closest annotated gene.
[5]Location of SNP within or near the gene.
[6]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.

Table 4.7. Single nucleotide polymorphisms with the highest posterior inclusion probabilities, on autosome 4, 17, and 29, in a genome-wide association study using rolling index values of beef cows managed in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Gene[4]	SNP location[5]	PIP[6]
rs110340473	4	42817580	*ENSBTAG00000022498*	0.566 Mb	0.010
rs41637536	17	30413148	*INTU*	0.008 Mb	0.200
rs42190442	29	48029098	*FADD*	Intron	0.021

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Closest annotated gene.
[5]Location of SNP within or near the gene.
[6]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
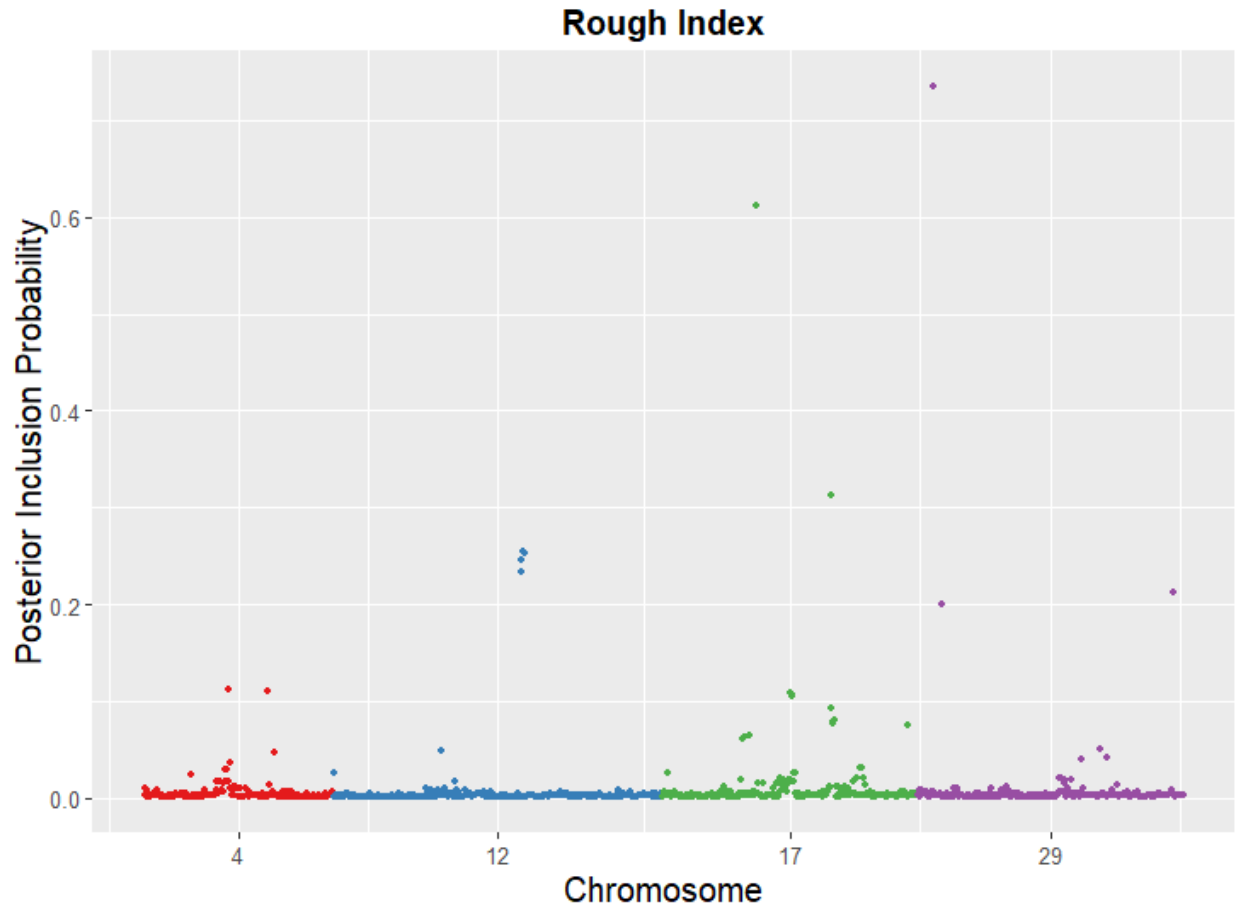
Figure 4.3. Manhattan plot depicting the results of a genome-wide association study using markers within one megabase genomic windows and rough index values of 75 cows managed in Arizona, Montana, and New Mexico.
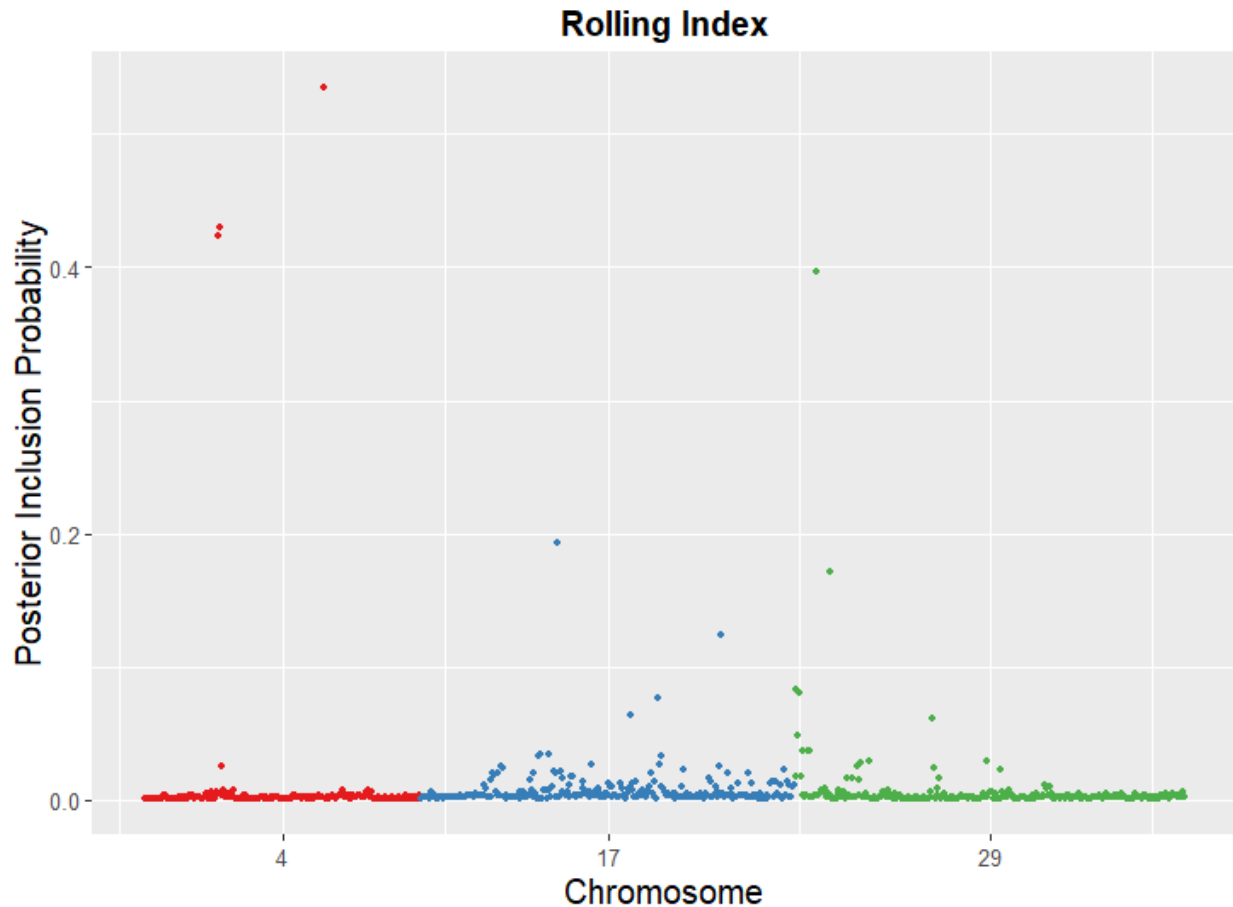
Rolling Index

Figure 4.4. Manhattan plot depicting the results of a genome-wide association study using markers within one megabase genomic windows and rolling index values of 75 cows managed in Arizona, Montana, and New Mexico.

Discussion

The association analyses using all markers ($k \geq 13{,}887$) failed to confirm four of the five candidate SNP for the rough index and all five of the candidate SNP for the rolling index. These results were unexpected given that these candidate loci explained 12 to 24% of the variation in the rough index and 11 to 36% of the variation in the rolling index when analyzed using simple linear regression. On average, SNP account for 1 to 2% of the variation in the phenotype (Bailey et al., 2015). A review of the literature suggested that accuracy of SNP detection may have been influenced by the data structure in which the number of markers greatly exceeded the number of

observations. This phenomenon is commonly referred to as "large p small n" or the "$p \gg n$ problem."

van den Berg et al. (2013) performed simulation scenarios to examine the effect of heritability, the number of quantitative trait loci (QTL), and sample size on the accuracy of QTL detection using BayesC and BayesCπ. Results of this analysis suggested that heritability and sample size were positively correlated with accuracy of detection (i.e., fewer false positives) while the number of QTL in the analysis was negatively correlated with accuracy. The $p \gg n$ problem was also described by Yazdani and Dunson (2015), "there are some clear limitations in scaling computation to very large p, as well as issues in obtaining reliable results when n is too small relative to p."

Association analyses using one megabase genomic windows ($k \leq 352$) provided evidence to suggest that $p \gg n$ influenced the accuracy of SNP detection when the GWAS was performed using all markers ($k \geq 13{,}887$). Truncating the number of markers from thousands to hundreds, to better suit the number of observations ($n = 75$), improved the accuracy of QTL detection and the top SNP identified in this study were in concordance with the simple linear regression results of Bailey et al. (2015). These results indicate that a larger sample size was needed to accurately perform QTL detection for these data using a Bayesian approach.

Conclusion

The Bayesian analyses using all markers ($k \geq 13{,}887$) confirmed the association between rs109619368 on chromosome 17 and the rough index. The remaining candidate SNP failed to meet the detection criteria with low posterior inclusion probabilities. Subsequent analyses using one megabase genomic windows ($k \leq 352$) provided evidence to support the findings of Bailey et al. (2015). More specifically, the five candidate SNP identified in 2015 had the highest PIP in this

study when the rough and rolling indices were the phenotypes. These results suggest that terrain-use may be a polygenic trait; however, a larger sample size is needed to validate the SNP associated with terrain-use indices in beef cattle.

CHAPTER 5: GENOME-WIDE ASSOCIATION STUDIES OF BEEF COW TERRAIN-USE

TRAITS USING BAYESIAN MULTIPLE-SNP REGRESSION

Introduction

Western beef producers utilize extensive rangeland pastures to maintain ~20% of the U.S. beef cattle inventory (DelCurto et al., 2017; Drouillard, 2018). Subsequently, producers face unique challenges associated with western rangeland topography including mountainous regions with rocky soils, steep slopes, and high elevation (DelCurto et al., 2017). These abiotic pasture-attributes adversely affect grazing as cattle prefer gentle terrain (Valentine, 1947; Mueggler, 1965; Roath and Krueger, 1982; Holechek, 1988). This preference often results in reduced forage harvest in the uplands (Bailey et al., 2017) and damage to riparian zones as cattle heavily graze these areas (Kauffman and Krueger, 1984). Therefore, terrain-use is an important consideration of grazing management for western beef operations.

The capital expenditure required for traditional grazing management practices (e.g., water development, fencing, supplementation) that may improve terrain-use is often very expensive (Tanaka et al., 2007); therefore, alternative practices warrant further consideration. Genetic selection has been proposed as a strategy to improve grazing patterns and Bailey et al. (2015) identified five quantitative trait loci (QTL) that may play a role in beef cattle terrain-use. While the results suggested that terrain-use is a polygenic trait that may be improved through genetic selection, these QTL and their underlying candidate genes need additional study.

The small sample size (n = 158) of the study published by Bailey et al. (2015) may have limited the detection of associations between genotypes and phenotypes and the statistical method employed (single-marker regression) did not account for single nucleotide polymorphism (SNP)

interactions. A larger sample size and a more robust statistical method may improve SNP detection for terrain-use. Furthermore, additional terrain-use phenotypes should be examined because the rolling and rough index reported by Bailey et al. (2015) were trait ratios which complicate the development of estimated breeding values for selection. More specifically, unequal selection pressure may be placed on the component traits when direct selection is applied to trait ratios. In contrast, "linear selection indices place a predetermined amount of selection pressure on the traits of interest and therefore a predictable amount of genetic change should result" (Gunsett, 1984). For example, in economic selection indices for beef cattle, larger economic weights are applied to EPD with greater economic opportunity (Lindholm and Stonaker, 1957).

The objectives of this study were to: 1) perform genome-wide association studies (GWAS) for six terrain-use traits (slope, elevation, vertical climb, distance travelled from water, rolling index, and rough index) using Bayesian methodologies and 2) survey the genome in consecutive, one megabase windows and calculate the proportion of total genetic variance explained by the markers within the windows for each terrain-use trait.

Materials and Methods

Study Sites and Cattle

Global positioning system (GPS) data and a combination of Illumina BovineHD (777,962 SNP) and BovineSNP50 genotypes (53,714 SNP) were obtained from 330 cows managed on 14 rangeland beef cattle operations located in the western United States: Carter Ranch (San Simon, AZ), Chihuahuan Desert Rangeland Research Center managed by New Mexico State University (CDRRC; Las Cruces, NM), Colorado State University Beef Improvement Center (CSU-BIC; Riverside, WY), Corona Range and Livestock Center managed by New Mexico State University (Corona, NM), Ensz Ranch (Center, CO), Evans Ranch (Silver City, NM), Fort Union Ranch (Las

71

Vegas, NM), Gund Ranch managed by the University of Nevada (Austin, NV), Hartley Ranch (Roy, NM), O RO Ranch (Prescott, AZ), Silver Spur Ranches (Encampment, WY), Thackeray Ranch managed by Montana State University (Havre, MT), Todd Ranch (Willcox, AZ), and Wilbanks Ranch (Mayhill, NM). A description of each rangeland operation, including cattle breed and pasture topography, are provided in Table 5.1.

Most of the cattle in this study were mature cows (n = 321); though, nine yearling heifers were studied at the Hartley Ranch. The cows varied in breed and physiological status (i.e., lactating verse non-lactating) across ranches (Table 5.2). Cattle were GPS monitored in the years 2011 to 2017 for approximately 3 to 19 weeks using two types of collars: Lotek 3300 GPS collars (Lotek Wireless, New Market, Ontario, Canada) and igotU Gt-120 GPS tracking collars (Knight et al., 2018a). Note that the length of GPS monitoring was dependent upon battery life. The GPS measurement interval (the time elapsed between coordinates) ranged from 5 to 15 minutes; however, the average interval was 10 minutes (Table 5.2). Assuming a 100% fix rate, a 10-minute interval would yield 144 coordinates for each cow for each day.

Table 5.1. Breed information and descriptions of pasture topography for rangeland beef operations in which cows were evaluated for terrain-use.

Ranch	Breed	Pasture size[2], ha	Elevation, m		Slope[3], %		Distance from Water, km	
			Range	Mean	Range	Mean	Max	Mean
Carter	Brangus	4184	1074 - 1424	1143	0 - 83	4	3.7	1.3
CDRRC 11/12[1]	Brangus	3994	1218 - 1427	1306	0 - 30	6	8.7	4.6
CDRRC 16[1]	Brangus	1451	1209 - 1414	1309	0 - 100+	23	4.9	2.2
Corona	Angus x Hereford Angus-cross	721, 1601	1737 - 1836	1783	0 - 36	4	4.6	2.3
Ensz	Hereford-cross	26082	2494 - 3885	3219	0 - 62	13	1.4	0.2
Evans	Angus	3610	1687 - 1954	1837	0 - 100+	12	4.7	1.7
Fort Union	Angus x Hereford	10063	2032 - 2564	2182	0 - 100+	8	3.5	1.0
Gund	Angus-cross	1404	1760 - 2527	2085	10 - 100+	37	0.8	0.3
Hartley	Angus Angus x Charolais Angus x Hereford	1056	1491 - 1766	1591	0 - 100+	11	4.4	1.1
ORO	Angus Angus x Hereford	5719	1684 - 2126	1897	0 - 100+	13	2.4	1.0
CSU-BIC	Angus	1351	2150 - 2411	2228	0 - 72	12	1.8	0.7
Silver Spur P[1]	Angus	1649	2255 - 2601	2423	0 - 100+	10	0.5	0.1
Silver Spur FS[1]	Angus	6425	2255 - 3200	2757	0 - 100+	17	0.6	0.1
Thackeray	Simmental x Hereford Simmental x Tarentaise Simmental x Hereford x Tarentaise	336	1182 - 1385	1283	0 - 100+	28	1.3	0.5
Todd	Limousin	9065	1277 - 2029	1454	0 - 100+	20	4.0	1.3
Wilbanks	Angus-cross	1201	1851 - 2112	1971	1 - 86	23	3.4	1.8

[1]Due to variation in pasture topography, pastures were classified as separate ranches: Chihuahuan Desert Rangeland Research Center pasture for cattle tracked in 2011 and 2012; Chihuahuan Desert Rangeland Research Center years pasture for cattle tracked in 2016; Silver Spur private land pasture; Silver Spur Forest Service Allotment.
[2]Size of the pasture used during the study period.
[3]Slope of 100+: the pasture contains very steep terrain that exceeds 45° (100%) slopes that cattle would rarely, if ever, use.

Table 5.2. Description of ranches, cows, and global positioning system (GPS) monitoring to evaluate terrain-use of beef cows in western U.S. mountain production systems.

Ranch	Year[1]	Number of Days Tracked[2]	Season	Interval[3], min	Cow Age[4]	Physiological Status
Carter	2011	75.75	Winter	15	Mature	Dry
CDRRC	2011	32.79	Summer	10	Mature	Lactating
CDRRC	2012	38	Winter	10	Mature	Dry
CDRRC	2016	84	Winter	10	Mature	Dry
Corona	2010	64.68	Summer	10	Mature	Lactating
Corona	2011	35.79	Summer	10	Mature	Lactating
Corona	2012	51	Summer	5	Mature	Lactating
Ensz	2016	71	Summer	10	Mature	Lactating
Evans	2012	59.42	Fall	10	Mature	Dry
Evans	2016	133	Fall	10	Mature	Dry
Fort Union	2017	85	Winter	10	Mature	Dry
Gund	2016	71	Summer	10	Mature	Lactating
Hartley	2009	122.88	Winter	15	Heifer	Dry
ORO	2017	113	Summer	10	Mature	Dry
CSU-BIC	2013	27	Summer	10	Mature	Lactating
CSU-BIC	2014	18	Summer	10	Mature	Lactating
Silver Spur	2017	119	Summer	10	Mature	Lactating
Thackeray	2011	37.63	Summer	10	Mature	Lactating
Todd	2011	90.58	Spring	15	Mature	Dry
Wilbanks	2015	89	Summer	10	Mature	Lactating

[1]Year that the GPS monitoring began.
[2]Number of days that the cows were tracked.
[3]Measurement interval for the GPS collars.
[4]Cows that were 3 to 14 years old were considered mature cows.

Phenotypes

Quality control measures applied to the GPS data and derivation of terrain-use traits (slope, elevation, and distance travelled from water) using USGS Digital Elevation Maps (DEMs) and ArcGIS software (Redlands, CA) were described by Bailey et al. (2015). Derivation of terrain-use measurements from GPS data was also described in Chapter 4. Slope and elevation were determined for each GPS coordinate using a DEM with a 10-m resolution for each pasture. Moreover, slope was calculated as the change in Y (ΔY) divided by the change in X (ΔX) multiplied by 100. Distance travelled from water was determined using DEMs with 5-m resolution

and the ArcGIS Euclidian Distance tool (Redlands, CA). Average slope, elevation, and distance travelled from water were then calculated for each cow using the GPS coordinates collected over the tracking period (Table 5.3). These trait averages where then used as three different phenotypic measures in the GWAS. In addition, slope, elevation, and distance travelled from water were incorporated into two previously developed terrain-use indices: rough index and rolling index (Bailey et al., 2015). As previously discussed in chapter three, the rough index was the normalized average of slope and elevation whereas the rolling index was the normalized average of slope, elevation, and distance travelled from water. Values generated using these indices were used as two additional phenotypes for GWAS.

Elevation was challenging to consistently describe across ranches due to variation in the location of the water source. Most of the study sites had water developments located at lower elevations in the pasture or grazing allotment; however, several ranches (CDRRC 16 and Evans) had water sources located at higher elevations. In the first scenario (A), higher elevation suggested improved grazing distribution; however, in the second scenario (B) a higher elevation suggested poor grazing distribution (Figure 5.1). To account for the variation in water source location and its effect on elevation measurements, new traits were developed to estimate the cows' use of elevation: vertical distance to water and change in elevation. Vertical distance to water was calculated using the following formula:

$$\text{Vertical distance to water} = \left| Elevation_g - Elevation_w \right|$$

where g represented the elevation of the cow for a given GPS coordinate and w represented the elevation for nearest water source. Figure 5.2 provides an example calculation for vertical distance to water. Vertical distance to water could not be calculated for all ranches due to challenges associated with flowing water (i.e., streams); therefore, change in elevation was used as an

75

alternative measurement for Carter, Corona, Ensz, Fort Union, Gund, ORO, CSU-BIC and Silver Spur ranches. Change in elevation was calculated by subtracting the lowest elevation in the study pasture from the elevation of each GPS location collected on the cow (Figure 5.3). As with the previous traits, change in elevation and vertical distance to water were averaged across all GPS coordinates collected during the study period to generate one phenotypic value per cow. For the GWAS, vertical distance to water and change in elevation were executed as one phenotype termed vertical climb (Table 5.4). Vertical climb could not be calculated for two animals due to omission of data; therefore, the sample size for vertical climb was 328 cows versus 330 for the other phenotypes (slope, elevation, distance from water, rolling index, and rough index).

Table 5.3. Descriptive statistics for terrain-use traits derived from global positioning system (GPS) measures of beef cows managed in the western U.S.

Ranch	n	Elevation[2], m				Slope[3], %				Distance from Water[4], m			
		Min	Mean	Max	SD[5]	Min	Mean	Max	SD[5]	Min	Mean	Max	SD[5]
Carter	12	1143.0	1153.2	1164.7	8.1	1.9	2.1	2.5	0.2	696.3	828.9	929.2	56.1
CDRRC 11-12[1]	32	1310.7	1315.5	1325.2	3.2	3.7	4.7	5.9	0.5	1026.9	1478.4	1864.7	260.4
CDRRC 16[1]	12	1314.7	1320.5	1327.3	4.0	5.5	6.1	6.7	0.4	1391.7	1555.8	1718.9	109.9
Corona	38	1760.6	1779.8	1792.9	7.6	3.2	3.7	4.5	0.2	1194.6	1898.6	2570.6	287.6
Ensz	16	3147.9	3252.2	3325.2	54.6	6.0	6.9	8.0	0.6	139.7	232.4	282.9	39.7
Evans	28	1732.5	1780.5	1816.6	22.9	6.7	7.9	8.7	0.5	1134.9	1522.6	1803.4	188.3
Fort Union	31	2114.4	2165.0	2363.5	47.4	4.3	6.7	11.0	1.5	434.2	724.4	1261.6	146.3
Gund	15	1919.4	1984.9	2107.0	61.3	16.0	22.6	30.2	3.5	210.7	264.5	344.0	41.4
Hartley	8	1537.7	1590.3	1655.4	46.8	9.5	13.6	18.7	3.7	390.5	751.4	1162.3	285.7
ORO	19	1792.6	1849.8	1934.0	37.2	6.6	10.3	12.7	1.4	674.8	904.6	1238.2	129.1
CSU-BIC	37	2174.4	2204.3	2233.8	15.6	6.3	8.9	10.9	1.1	345.0	444.9	592.4	76.5
Silver Spur P[1]	19	2381.2	2412.1	2442.8	19.4	5.2	7.0	8.1	0.6	71.5	80.1	89.5	6.1
Silver Spur FS[1]	9	2500.5	2583.1	2699.0	60.8	7.4	8.1	9.3	0.6	76.9	85.2	97.0	7.0
Thackeray	16	1266.3	1283.2	1293.8	7.8	14.5	18.9	20.7	1.4	388.6	470.0	530.2	40.0
Todd	16	1325.0	1394.6	1495.9	51.0	4.7	7.6	11.1	1.8	763.3	975.8	1234.8	147.1
Wilbanks	15	1929.9	1951.2	1978.9	15.3	12.9	14.6	18.4	1.9	680.5	941.0	1394.0	226.9

[1]Due to variation in pasture topography, pastures were classified as separate ranches: Chihuahuan Desert Rangeland Research Center pasture for cattle tracked in 2011 and 2012; Chihuahuan Desert Rangeland Research Center years pasture for cattle tracked in 2016; Silver Spur private land pasture; Silver Spur Forest Service Allotment.
[2]Elevation derived from each GPS coordinate; averaged across all coordinates in the study period for a given cow.
[3]Slope calculated as $(\Delta Y/\Delta X) \times 100$; averaged across all coordinates in a study period for a given cow.
[4]Distance travelled from water calculated using the nearest water source; averaged across all coordinates in a study period for a given cow.
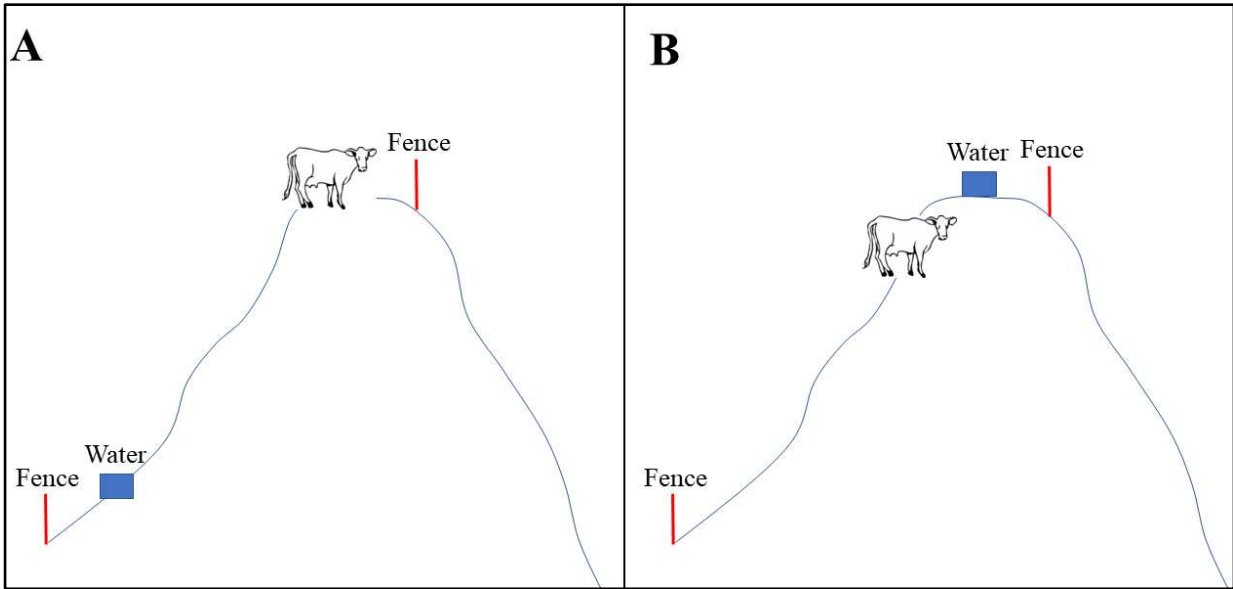[5]SD: standard deviation.

Figure 5.1. Diagram depicting the effect of water location on elevation measurements of beef cows. A) Typical scenario where higher elevation reflects improved grazing distribution; B) Unusual scenario where higher elevation reflects poor grazing distribution.
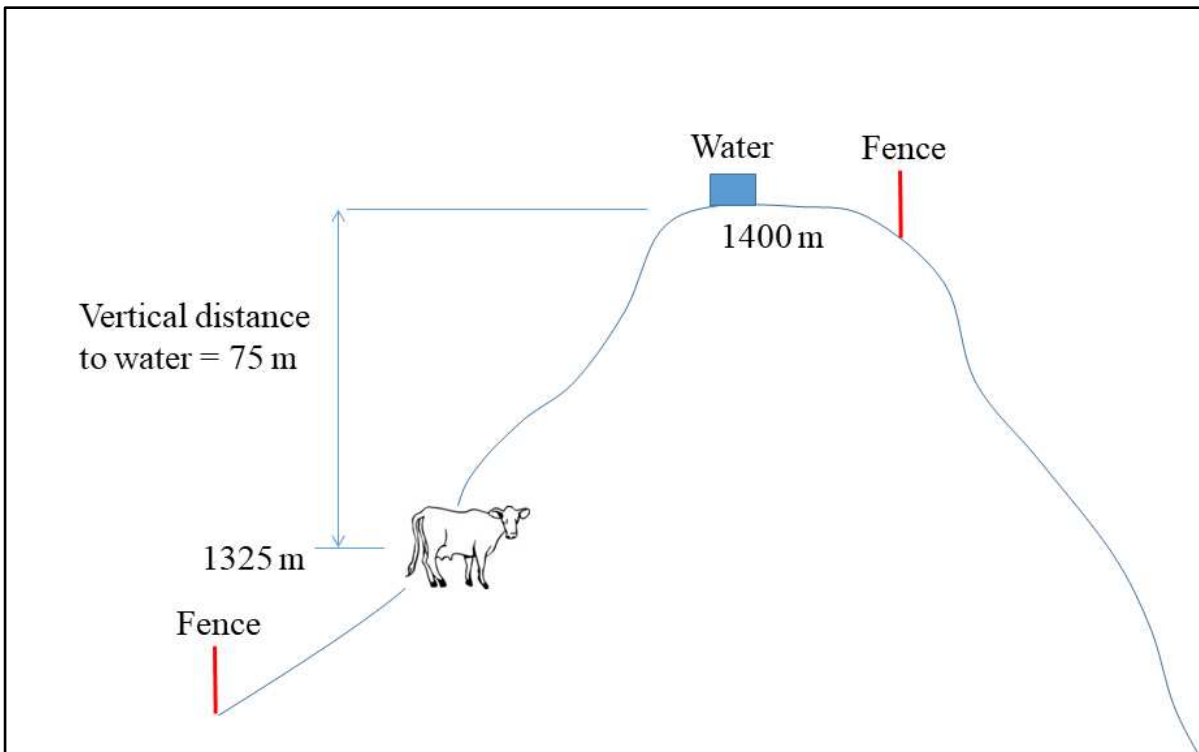

Figure 5.2. Diagram depicting the calculation of vertical distance to water for beef cows using the elevation at which the cow is located and the elevation of the nearest water source.
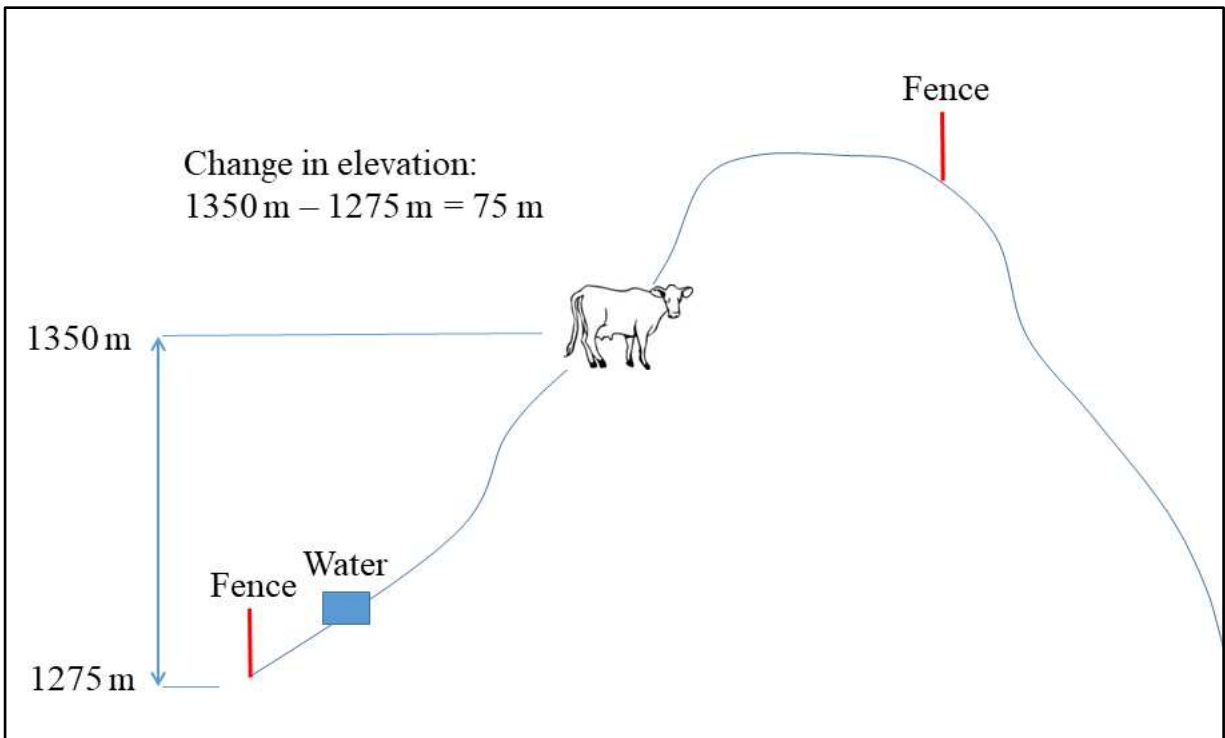
Figure 5.3. Diagram depicting the calculation of change in elevation for beef cows using the lowest elevation in the pasture and the elevation at which the cow is located.

Table 5.4. Descriptive statistics by ranch for vertical climb, derived from global positioning system (GPS) measurements of beef cows managed in extensive, rugged rangeland pastures in the western U.S.

| Ranch | n | Vertical Climb[1], m | | | |
		Min	Mean	Max	SD[2]
Carter	12	69.0	80.3	90.6	8.6
CDRRC 11-12[a]	31	10.2	15.8	32.6	5.3
CDRRC 16[b]	12	31.5	48.9	65.3	10.9
Corona	37	9.7	15.6	20.7	3.2
Ensz	16	653.9	758.3	831.2	2980.7
Evans	28	40.2	74.2	103.2	19.7
Fort Union	31	82.4	132.9	331.5	47.4
Gund	15	159.5	225.0	347.0	61.3
Hartley	8	12.8	27.0	59.8	15.8
ORO	19	108.6	165.8	250.0	37.2
CSU-BIC	37	24.4	54.2	83.8	15.6
Silver Spur P[c]	19	126.2	157.1	187.8	19.4
Silver Spur FS[d]	9	245.5	328.1	444.0	60.8
Thackeray	16	35.4	48.6	56.9	5.7
Todd	16	48.0	117.6	218.9	51.0
Wilbanks	15	24.7	45.9	68.5	13.2

[1]Due to variation in pasture topography, pastures were classified as separate ranches: Chihuahuan Desert Rangeland Research Center pasture for cattle tracked in 2011 and 2012; Chihuahuan Desert Rangeland Research Center years pasture for cattle tracked in 2016; Silver Spur private land pasture; Silver Spur Forest Service Allotment.
[2]Vertical climb calculated using vertical distance from water or change in elevation formula.
[3]SD: standard deviation.

Genotypes

BovineHD genotypes (777,962 SNP) were obtained from 293 cows in the study. The other 37 cows were genotyped with the BovineSNP50 Beadchip (53,714 SNP). The high-density SNP data were truncated to match the BovineSNP50 data to generate cohesive genotype data for analyses (n = 330). Due to differences in sample size, genotype quality control for vertical climb was performed separately from the effort for slope, elevation, distance travelled from water, rolling index and rough index. Genotype quality control was completed using PLINK 1.9 (Purcell and Chang, 2015) and standard filters: SNP call rate ≥ 0.90, minor allele frequency < 0.01, and Hardy-

Weinberg Equilibrium < 0.000001, sample call rate ≥ 0.90, and heterozygosity rate ± 3 standard deviations from the mean.

Marker filters were applied prior to individual filters to maintain a larger sample size. As discussed by Anderson et al. (2010), applying individual quality control filters prior to marker filters favors the retention of markers over individuals. With limited data for a GWAS (few hundred cows), sample size was deemed more important than the number of markers in the study. Linkage disequilibrium (LD) pruning was performed using an R^2 threshold of 0.2 and then relatedness of individuals and population structure were evaluated using identity by descent and principle component analysis (PCA). First and second degree relatives (pi-hat > 0.2; Marees et al., 2018) were retained in this study (n = 154) to maintain the sample size. A PCA plot, generated using RStudio (version 3.3.2), did not suggest population stratification (i.e., subpopulations with systematic differences in allele frequencies due to ancestry) as individuals did not form distinct clusters (Figure 5.4). Therefore, after quality control, 321 animals and 42,603 SNP were available for the vertical climb analysis and 323 animals and 42,699 SNP were available for all other analyses. Genotypes were recoded from AB format to numerical values (AA [-1], AB [0], and BB [1]) and missing genotypes were filled with the median for that locus.
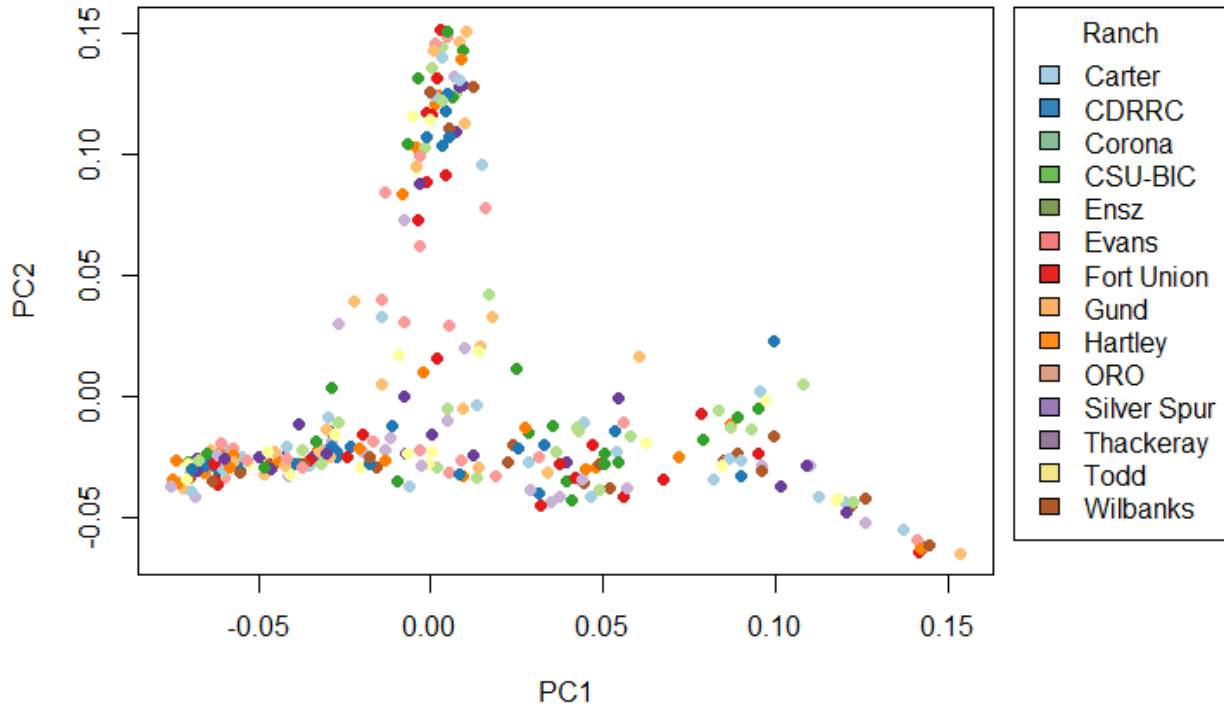
Figure 5.4. Principle component analysis (PCA) plot for a multi-breed population of beef cows (n = 323) managed on fourteen ranches across the western U.S.

Statistical Analyses and Bioinformatics

A genome wide association study was conducted for each phenotype using the BOLT software package (Release 1.2.7; http://www.thetasolutionsllc.com/bolt-software.html) and BayesC methodology developed by Habier et al. (2011). The most appropriate model for the Bayesian GWAS was as follows:

$$y = X\beta + \sum_{k=1}^{K} Za + e$$

where y was the vector of observations, X was the incidence matrix relating fixed effects to the observations in y, β was the vector of unknown fixed effect solutions, K was the number of SNP in the analysis, Z was the vector of genotype covariates for SNP k (coded -1, 0, 1), a was the vector of random allele substitution effects, and e was the vector of residual effects. Given the difficulty

of measuring terrain-use across western U.S. rangeland ranches, numerous fixed effects were examined during model selection including: ranch, breed, terrain type (mountainous or rolling), season (spring, summer, fall, winter), physiological status (lactating or dry), collar type (Lotek or igotU), and the GPS tracking start date. Backward selection with alpha 0.05 was used to identify significant predictors for each model. All predictors except collar type were linearly dependent with GPS tracking start date and therefore, these variables were removed from the models. Collar type was not significant, and thus the most appropriate model included GPS start date as a fixed effect. Fitting start date in the model was analogous to fitting a designated contemporary group to account for environmental differences among cows on ranches and pastures.

As described in Chapter 4, BayesC models assume that a proportion of SNP have no effect on the phenotype (π) and the proportion SNP with an effect (1-π) are normally distributed with a common variance denoted by σ_a^2 (Fernando et al., 2017). For these association analyses, π was 0.995 and the variance of marker effects was calculated using the following equation:

$$\sigma_\alpha^2 = \frac{c\sigma_u^2}{2k(1-\pi)\overline{pq}}$$

where c represented the proportion of genetic variance explained by SNP effects, σ_u^2 represented the additive genetic variance, k represented the number of markers, π represented the proportion of SNP with null effect, and $\overline{pq}$ represented the average of the product of the p and q loci frequencies. Parameters used to derive the variance of marker effects for each phenotype are described in Table 5.5.

Table 5.5. Parameters used to calculate the variance of marker effects for Bayesian-based genotype-phenotype association analyses using six beef cattle terrain-use traits.

Parameter[1]	Distance from Water	Elevation	Slope	Rolling Index	Rough Index	Vertical Climb
h^2	0.34	0.34	0.34	0.34	0.36	0.34
σ^2_P	30097.28	1007.85	1.95	42.74	48.71	994.95
c	1	1	1	1	1	1
σ^2_u	10233.08	342.67	0.66	14.53	17.54	338.28
k	42,699	42,699	42,699	42,699	42,699	42,603
π	0.995	0.995	0.995	0.995	0.995	0.995
$\overline{pq}$	0.25	0.25	0.25	0.25	0.25	0.25

[1]h^2 = heritability estimate (Bailey et al., 2015); σ^2_P = phenotypic variance; c = proportion of genetic variance explained by SNP effects; σ^2_u = additive genetic variance; k = number of SNP, π = proportion of markers with null effect; $\overline{pq}$ = average of the product of the p and q loci frequencies.

Allele substitution effects, variance of the allele substitution effects, and the number of times the marker was included in the model (count) were derived using a single-site Gibbs sampler with 150,000 iterations. Posterior inclusion probability (PIP) was calculated for each marker and the five SNP with the highest PIP were identified for each phenotype. These SNP were termed the candidate SNP throughout the study. Cattle QTL database (Cattle QTLdb; https://www.animalgenome.org/cgi-bin/QTLdb/BT/index) was used to determine if the QTL had been previously associated with beef cattle traits. Ensembl genome database (Release 94; Zerbino et al., 2017) was used to identify genes within one megabase of the candidate SNP and the annotated gene located nearest to SNP was deemed the putative candidate gene. Gene ontology was examined using AgBase (version 2.00; http://agbase.arizona.edu/index.html).

As summarized by Wolc et al. (2012) and Garrick and Fernando (2013), Bayesian multiple-SNP regression may fail to reveal strong associations between individual markers and the trait of interest (i.e., low PIP or small percentage of genetic variance explained) due to linkage disequilibrium; therefore, associations are often identified using genomic windows. For this study, non-overlapping one megabase genomic windows were derived using the annotation for bovine

assembly UMD3.1.1. The Markov chain Monte Carlo (MCMC) sampling chain values were used to calculate the proportion of genetic variance explained by the markers in the genomic windows and the five genomic windows explaining the largest portion of genetic variance were deemed regions of interest. The SNP with the highest PIP within the region of interest was considered the lead SNP and Ensembl genome database (Release 94; Zerbino et al., 2017) was used to identify the annotated gene located nearest to the lead SNP. Beef cattle traits previously associated with the QTL were identified using Cattle QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/BT/index) and AgBase (version 2.00; http://agbase.arizona.edu/index.html) was used to assess the gene ontology.

Results

Slope

The GWAS, in which percent slope was the phenotype, revealed QTL on chromosome 10 and 17 (Figure 5.5). Of the five candidate SNP identified in this analysis, four were located on chromosome 10 and one was located on chromosome 17 (Table 5.6). The SNP with the highest PIP was rs29013509 with a value 0.14. Therefore, rs29013509 was included in approximately 21,000 iterations out of 150,000 in the MCMC chain. In comparison, the four-other candidate SNP identified in this analysis (rs42415241, rs109097567, rs41848746, and rs29013631) had relatively low PIP ranging from 0.05 to 0.06.

Like the five candidate markers, the top five genomic windows were on chromosome 4, 10, 17, and 29 (Figure 5.6). The markers within the 20th window on chromosome 10 explained 0.0057% of the genetic variance of percent slope and the SNP located in 38th genomic window explained 0.0045% of the genetic variance (Table 5.7). Genomic windows identified on

chromosome 4, 17, and 29 explained 0.0019%, 0.0018%, and 0.0017%, respectively. In total, the proportion of genetic variance explained by the one megabase genomic windows was 0.98%.

Comparing the candidate markers identified in the GWAS to the genomic windows and their lead SNP revealed a high level of concordance. Four of the candidate markers were located within the top genomic windows and three of the candidate SNP were also considered lead SNP (rs29013509, rs42415241, rs41848746). These results provide evidence to support the importance of these QTL regarding percent slope traversed by beef cows in western rangelands.
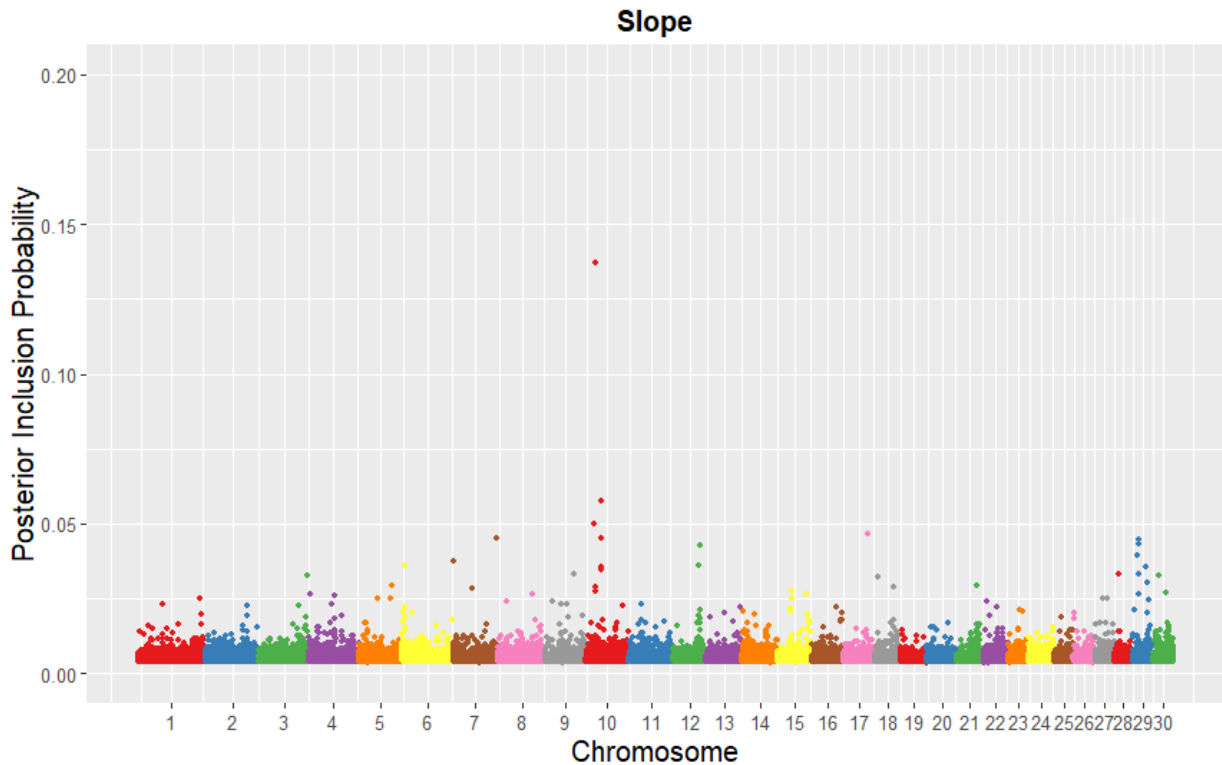
Figure 5.5. Manhattan plot of a multi-breed genome-wide association study for percent slope traversed by 323 beef cows in the western U.S.

Table 5.6. Top five single nucleotide polymorphisms (SNP) identified in the multi-breed genome-wide association study for percent slope traversed by beef cows (n = 323) in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Window[4]	Gene[5]	SNP location[6]	PIP[7]
rs29013509	10	20498087	10_20	*TBC1D21*	0.026 Mb	0.14
rs42415241	10	38986604	10_38	*EPB42*	0.448 Mb	0.06
rs109097567	10	19494672	10_19	*ADPGK*	0.085 Mb	0.05
rs41848746	17	61920415	17_61	*TBX3*	0.432 Mb	0.05
rs29013631	10	38826667	10_38	*EPB42*	0.160 Mb	0.05

[1]Reference SNP cluster identification assigned by the National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Genomic window: chromosome and n^{th} 1 Mb window on that chromosome.
[5]Closest annotated gene.
[6]Location of SNP within or near the gene.
[7]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
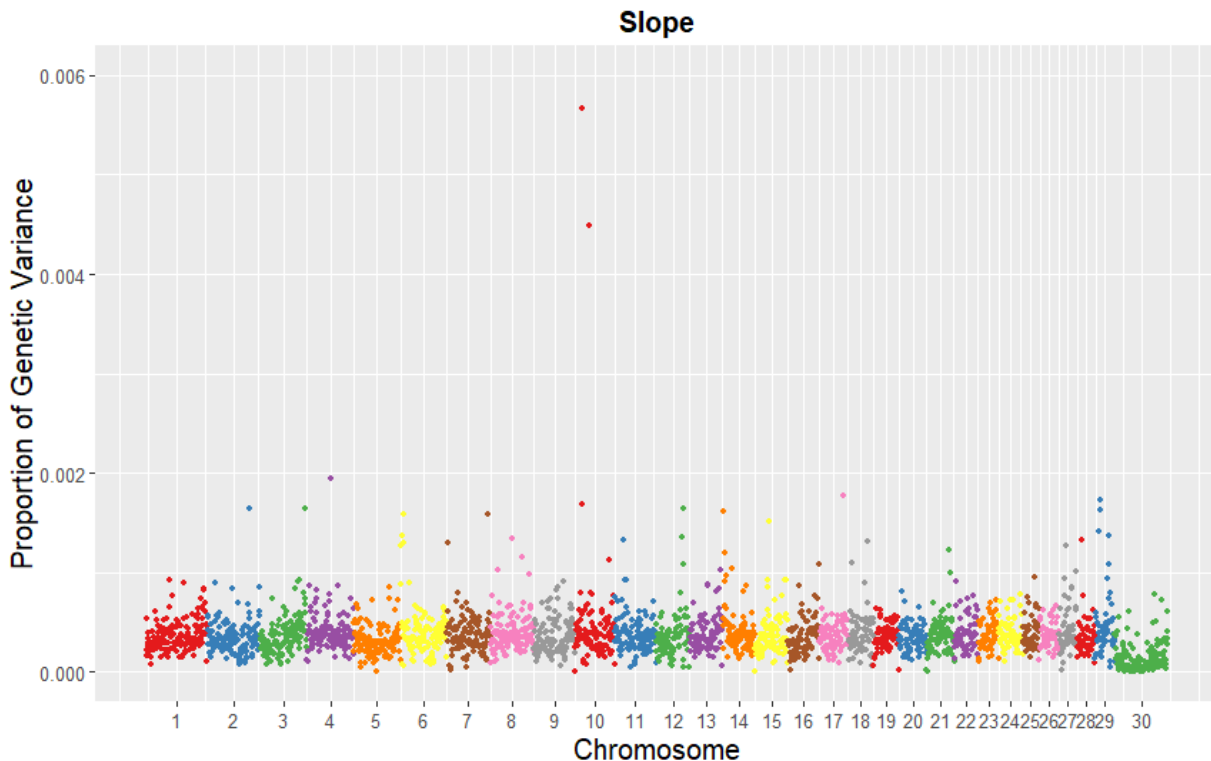

Figure 5.6. Manhattan plot of the proportion of genetic variance explained by markers in one megabase consecutive genomic windows for percent slope traversed by 323 beef cows in the western U.S.

Table 5.7. Top five genomic windows (1 Mb) identified in the multi-breed genome-wide association study for percent slope traversed by beef cows (n = 323) in the western U.S.

Window[1]	Start[2] (bp)	End[3] (bp)	SNP[4]	Var[5], %	Lead SNP[6]	Gene[7]
10_20	20008636	20932671	20	0.0057	rs29013509	*TBC1D21*
10_38	38026306	38986604	15	0.0045	rs42415241	*EPB42*
4_61	61015911	61972879	24	0.0019	rs43109323	*SEPT7*
17_61	61028494	61955493	21	0.0018	rs41848746	*TBX3*
29_14	14009294	14983423	20	0.0017	rs42477618	*ENSBTAG00000022427*

[1]Genomic window: chromosome and n[th] 1 Mb window on that chromosome.
[2]Start of the 1 Mb window (bp).
[3]End of the 1 Mb window (bp).
[4]Number of SNP within the window.
[5]Percentage of genetic variance explained by the genomic window.
[6]Lead SNP: SNP with highest PIP.
[7]Closest annotated gene.

Elevation and Vertical Climb

The Manhattan plot of the genotype to phenotype association analysis for elevation revealed several notable peaks on chromosome 4, 7, 11, 12, 23, and 24 (Figure 5.7). Further examination of the markers within the peaks, and their associated PIP, suggested that the candidate SNP for elevation were intronic and intergenic variants on chromosome 4, 11, 12, 23 and 24 (Table 5.8). This included: rs41600226 (PIP = 0.18), rs43408732 (PIP = 0.17), rs109716600 (PIP = 0.12), rs41633961 (PIP = 0.12), and rs109669554 (PIP = 0.12). Analysis using genomic windows revealed similar chromosomal regions (chromosomes 4, 7, 11, 23, and 24) related to elevation (Figure 5.8). The 47th window on chromosome 23 explained the largest proportion of genetic variance for elevation (0.0052%). Three windows explained 0.0036% of the genetic variance including: the 79th window on chromosome 4, 86th window on chromosome 7, and 18th window on chromosome 24 (Table 5.9). The fifth genomic window on chromosome 11 explained a slightly lower proportion of genetic variance (0.0030%) than the other top four windows. In total, the genomic windows explained 0.86% of the genetic variance. As with slope, three of the candidate

SNP for elevation were also considered lead SNP within the genomic windows analysis (rs41600226, rs43408732, and rs109716600).

Genome-wide association study results for vertical climb paralleled the results produced in the association analysis using the phenotype elevation. Again, important genomic regions were identified on chromosomes 4, 11, 12, and 23 (Figure 5.9) and rs41600226 had the highest rate of inclusion with a PIP of 0.16 (Table 5.10). Re-ranking occurred amongst the candidate SNP on chromosome 4, 11, and 12 and rs110978254 on chromosome 22 replaced rs110978254 on chromosome 24 in the top five SNP. In general, the PIP for the five candidate SNP were generally lower in the vertical climb analysis than the elevation analysis.

Three of the five genomic windows for vertical climb were also associated with elevation (chromosome 11, 23, and 24; Figure 5.9). The markers within 47th window on chromosome 23 explained 0.0050% of the genetic variance, the 91st window on chromosome 11 accounted 0.0034%, and the 18th window on chromosome 24 explained 0.0028% of the genetic variance in elevation (Table 5.11). The 55th window on chromosome 22 that explained 0.0028% of the genetic variance and the 70th window on chromosome 12 that explained 0.0026% of the genetic variance were unique to vertical climb. The proportion of genetic variance explained by SNP within the non-overlapping genomic windows was 0.86%. A comparison of candidate markers and genomic windows for vertical climb revealed a high level of concordance with four overlapping SNP (rs41600226, rs109502510, rs110978254, and rs41633961). These results suggested that the QTL on chromosomes 11, 12, 22, and 24 are important regions for the vertical climb of beef cattle.
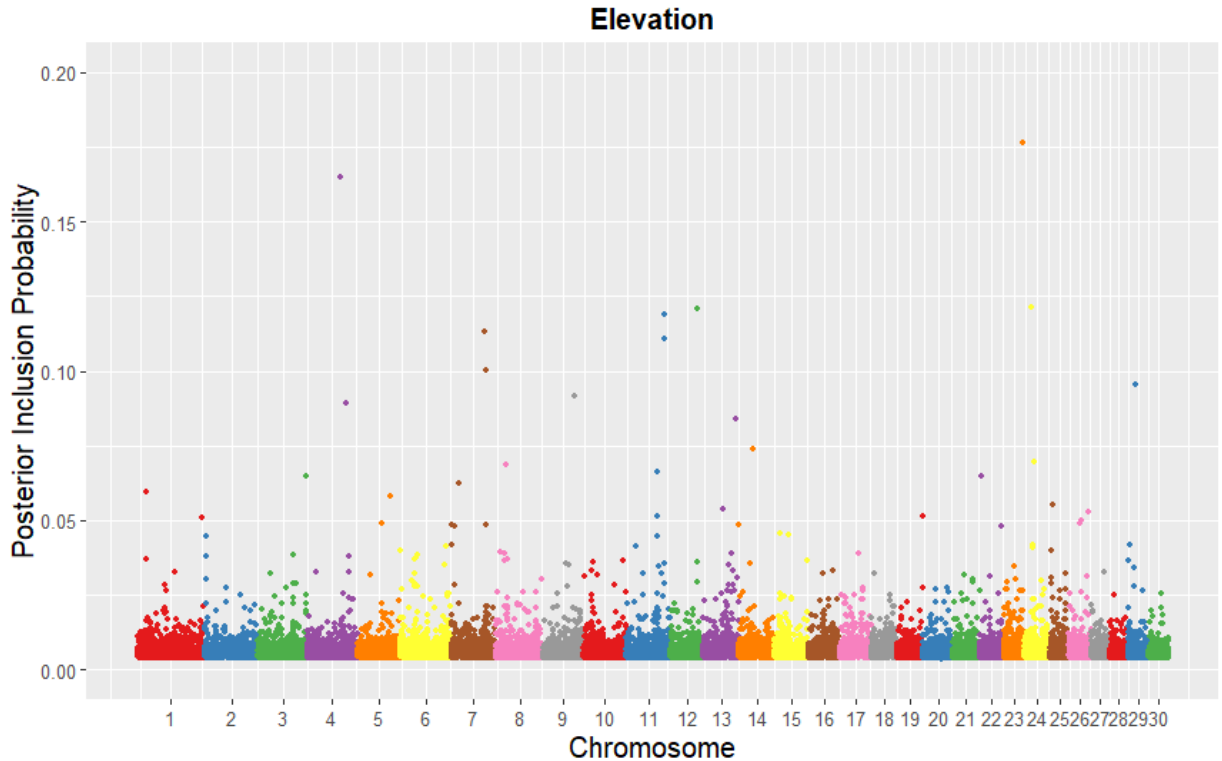
Figure 5.7. Manhattan plot of a multi-breed genome-wide association study for elevation (m) traversed by 323 beef cows in the western U.S.

Table 5.8. Top five single nucleotide polymorphisms (SNP) identified in the multi-breed genome-wide association study for elevation (m) traversed by beef cows (n = 323) in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Window[4]	Gene[5]	SNP location[6]	PIP[7]
rs41600226	23	47219946	23_47	*EEF1E1*	Intron	0.18
rs43408732	4	79996650	4_79	*INHBA*	Intron	0.17
rs109716600	24	18893792	24_18	*CELF4*	0.925 Mb	0.12
rs41633961	12	70064114	12_70	*ENSBTAG00000032603*	Intron	0.12
rs109669554	11	93142448	11_93	*MRRF*	0.004 Mb	0.12

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Genomic window: chromosome and n[th] 1 Mb window on that chromosome.
[5]Closest annotated gene.
[6]Location of SNP within or near the gene.
[7]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
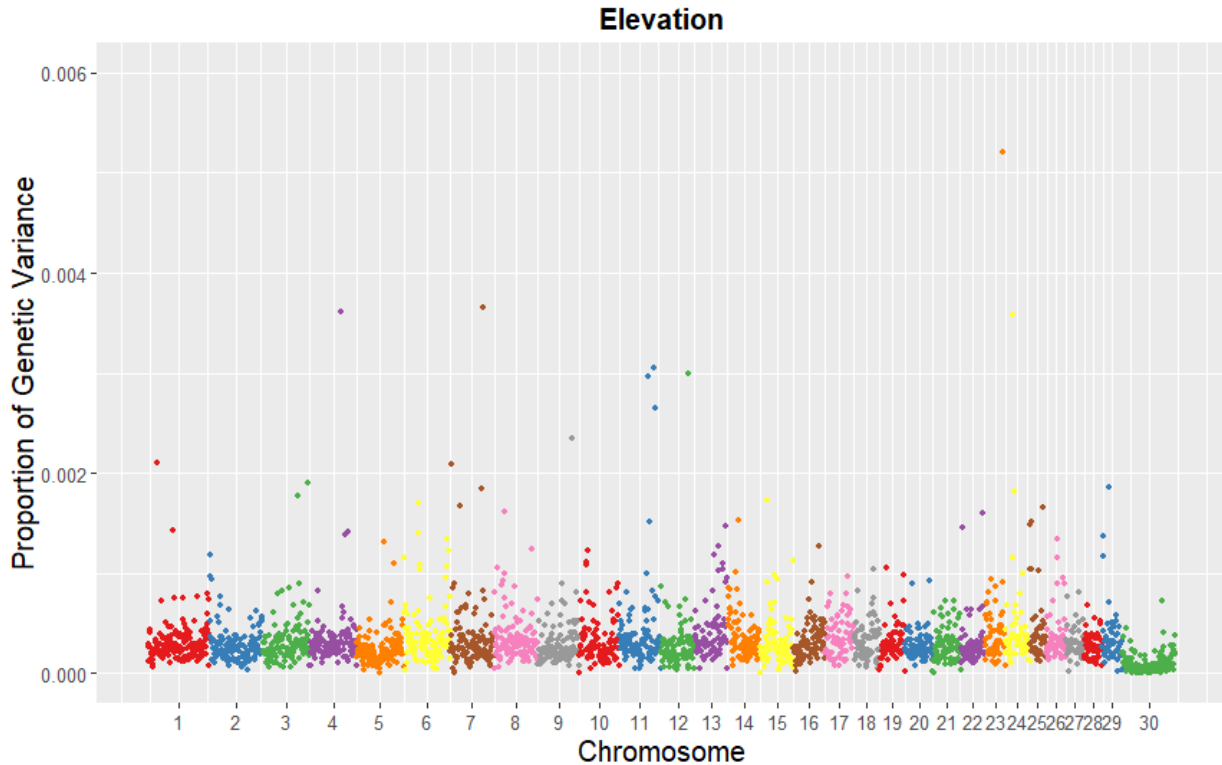
Figure 5.8. Manhattan plot of the proportion of genetic variance explained by markers in one megabase consecutive genomic windows for elevation (m) traversed by 323 beef cows in the western U.S.

Table 5.9. Top five genomic windows (1 Mb) identified in the multi-breed genome-wide association study for elevation (m) traversed by beef cows (n = 323) in the western U.S.

Window[1]	Start[2] (bp)	End[3] (bp)	SNP[4]	Var[5], %	Lead SNP[6]	Gene[7]
23_47	47005648	47953939	19	0.0052	rs41600226	*EEF1E1*
7_86	86044835	86936090	21	0.0036	rs110681394	*EDIL3*
4_79	79215287	79996650	9	0.0036	rs43408732	*INHBA*
24_18	18036453	18990335	14	0.0036	rs109716600	*CELF4*
11_91	91068648	91995272	19	0.0030	rs109502510	*ENSBTAG00000039201*

[1]Genomic window: chromosome and n[th] 1 Mb window on that chromosome.
[2]Start of the 1 Mb window (bp).
[3]End of the 1 Mb window (bp).
[4]Number of SNP within the window.
[5]Percentage of genetic variance explained by the genomic window.
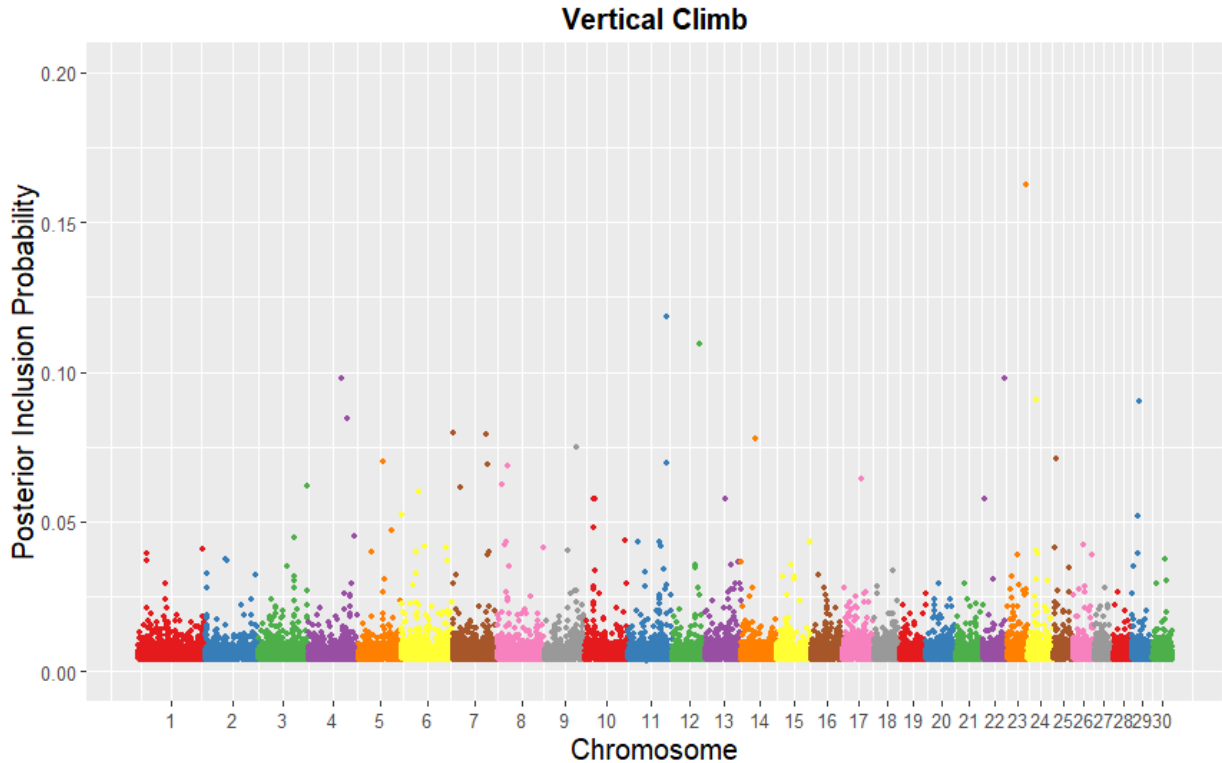[6]Lead SNP: SNP with highest PIP.
[7]Closest annotated gene.

Figure 5.9. Manhattan plot of a multi-breed genome-wide association study for vertical climb of 321 beef cows managed in the western U.S.

Table 5.10. Top five single nucleotide polymorphisms (SNP) identified in the multi-breed genome-wide association study for vertical climb of 321 beef cows managed in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Window[4]	Gene[5]	SNP location[6]	PIP[7]
rs41600226	23	47219946	23_47	*EEF1E1*	Intron	0.16
rs109502510	11	91680069	11_91	*ENSBTAG00000039201*	0.535 Mb	0.12
rs41633961	12	70064114	12_70	*ENSBTAG00000032603*	Intron	0.11
rs43408732	4	79996650	4_79	*INHBA*	Intron	0.10
rs110978254	22	55257755	22_55	*ATP2B2*	Intron	0.10

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Genomic window: chromosome and n^{th} 1 Mb window on that chromosome.
[5]Closest annotated gene.
[6]Location of SNP within or near the gene.
[7]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
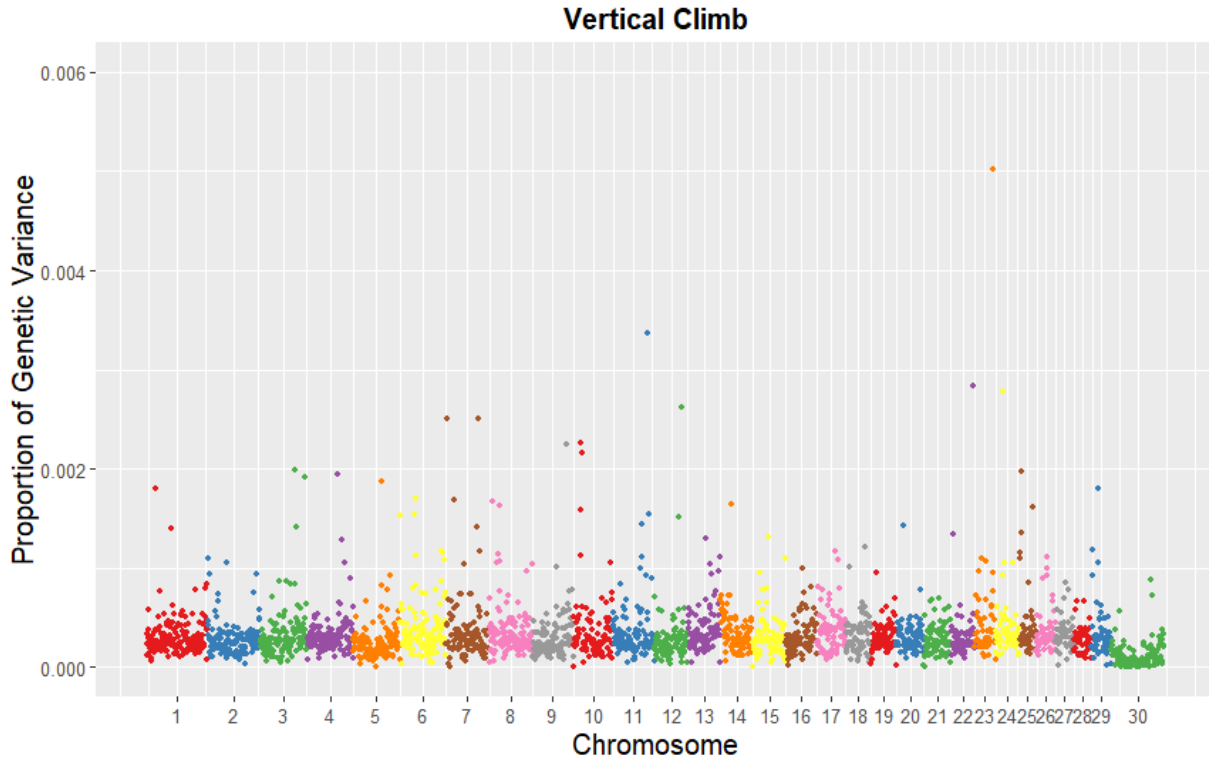
Figure 5.10. Manhattan plot of the proportion of genetic variance explained by markers in one megabase consecutive genomic windows for vertical climb of 321 beef cows managed in the western U.S.

Table 5.11. Top five genomic windows (1 Mb) identified in the multi-breed genome-wide association study for vertical climb of 321 beef cows managed in the western U.S.

Window[1]	Start[2] (bp)	End[3] (bp)	SNP[4]	Var[5], %	Lead SNP[6]	Gene[7]
23_47	47005648	47953939	19	0.0050	rs41600226	*EEF1E1*
11_91	91068648	91995272	19	0.0034	rs109502510	*ENSBTAG00000039201*
22_55	55063369	55913908	22	0.0028	rs110978254	*ATP2B2*
24_18	18036453	18990335	14	0.0028	rs109716600	*CELF4*
12_70	70000094	70290528	11	0.0026	rs41633961	*ENSBTAG00000032603*

[1]Genomic window: chromosome and nth 1 Mb window on that chromosome.
[2]Start of the 1 Mb window (bp).
[3]End of the 1 Mb window (bp).
[4]Number of SNP within the window.
[5]Percentage of genetic variance explained by the genomic window.
[6]Lead SNP: SNP with highest PIP.
[7]Closest annotated gene.

Distance Travelled from Water

Unlike slope, elevation, or vertical climb, the Manhattan plot for distance travelled from water failed to reveal any notable peaks (Figure 5.11). In addition, posterior inclusion probabilities for the markers in this GWAS were extremely low (PIP < 0.05) when compared to the other analyses. The top markers included rs41598300 on chromosome 16 (PIP = 0.03), rs109226946 on chromosome eight (PIP = 0.03), rs109175805 on chromosome 5 (PIP = 0.02), rs42599235 on chromosome 21 (PIP = 0.02) and rs43256975 on chromosome 1 (PIP = 0.02). A description of these SNP and their positional candidate genes are provided in Table 5.12.

The top five genomic windows for distance travelled from water were located on chromosome 1, 5, 8, 16 and 17 (Figure 5.12). These individual windows explained 0.0012 to 0.0014% of the genetic variation for distance travelled from water. The proportion of genetic variance explained by all of the one megabase genomic windows was 1.07%. As with the posterior inclusion probability values, these percentages were low in comparison to the proportion of genetic variance explained for the five other phenotypes. Four of the five candidate SNP were also considered lead SNP within the top five genomic windows. This included rs109226946 within the 86th window on chromosome 8, rs41598300 within the 80th window on chromosome 16, rs43256975 within the 108th window on chromosome 1, and rs109175805 the 101st window on chromosome 5 (Table 5.13).
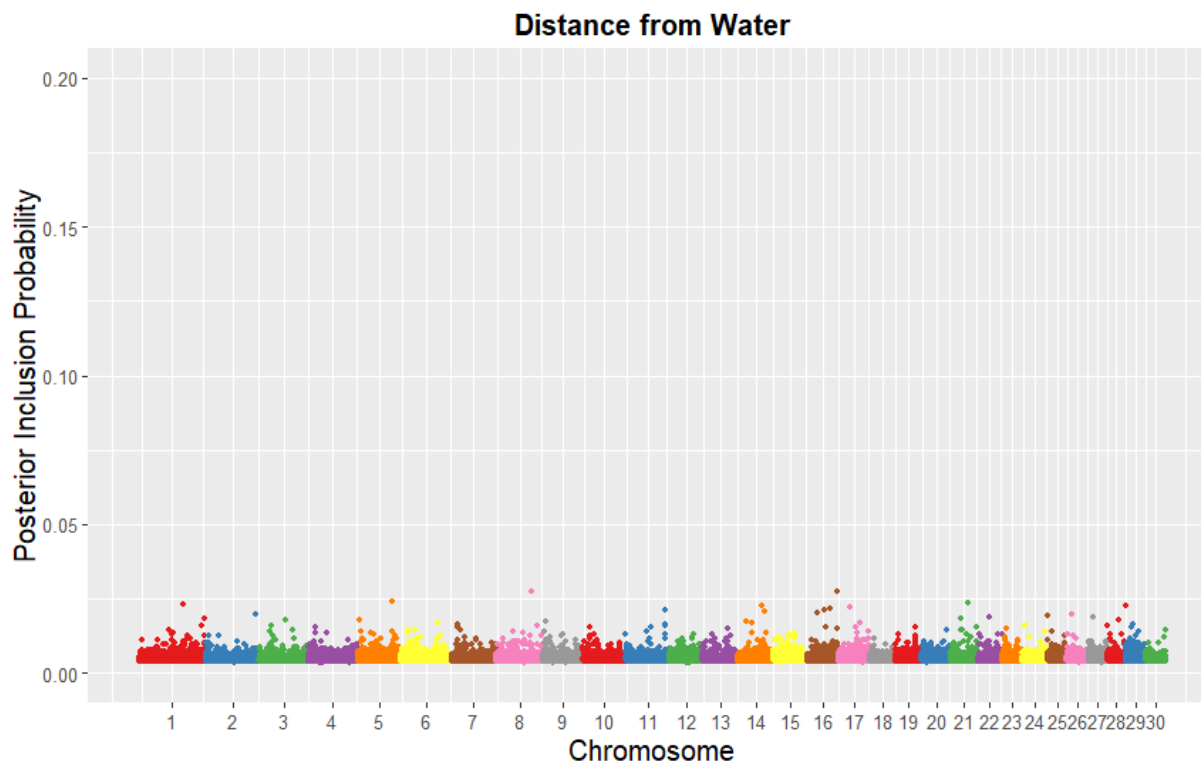
Figure 5.11. Manhattan plot of a multi-breed genome-wide association study for the distance beef cows (n = 323) travelled from water while managed on rangeland in the western U.S.

Table 5.12. Top five single nucleotide polymorphisms (SNP) identified in the multi-breed genome-wide association study for the distance beef cows (n = 323) travelled from water while managed on rangeland in the western U.S

RS ID[1]	CHR[2]	Position[3]	Window[4]	Gene[5]	SNP location[6]	PIP[7]
rs41598300	16	80407423	16_80	*NR5A2*	0.344 Mb	0.03
rs109226946	8	86442366	8_86	*PHF2*	0.017 Mb	0.03
rs109175805	5	101518663	5_101	*A2ML1*	Intron	0.02
rs42599235	21	47902442	21_47	*ENSBTAG00000000655*	Intron	0.02
rs43256975	1	108154057	1_108	*C1H3orf80*	0.067 Mb	0.02

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Genomic window: chromosome and n^{th} 1 Mb window on that chromosome.
[5]Closest annotated gene.
[6]Location of SNP within or near the gene.
[7]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
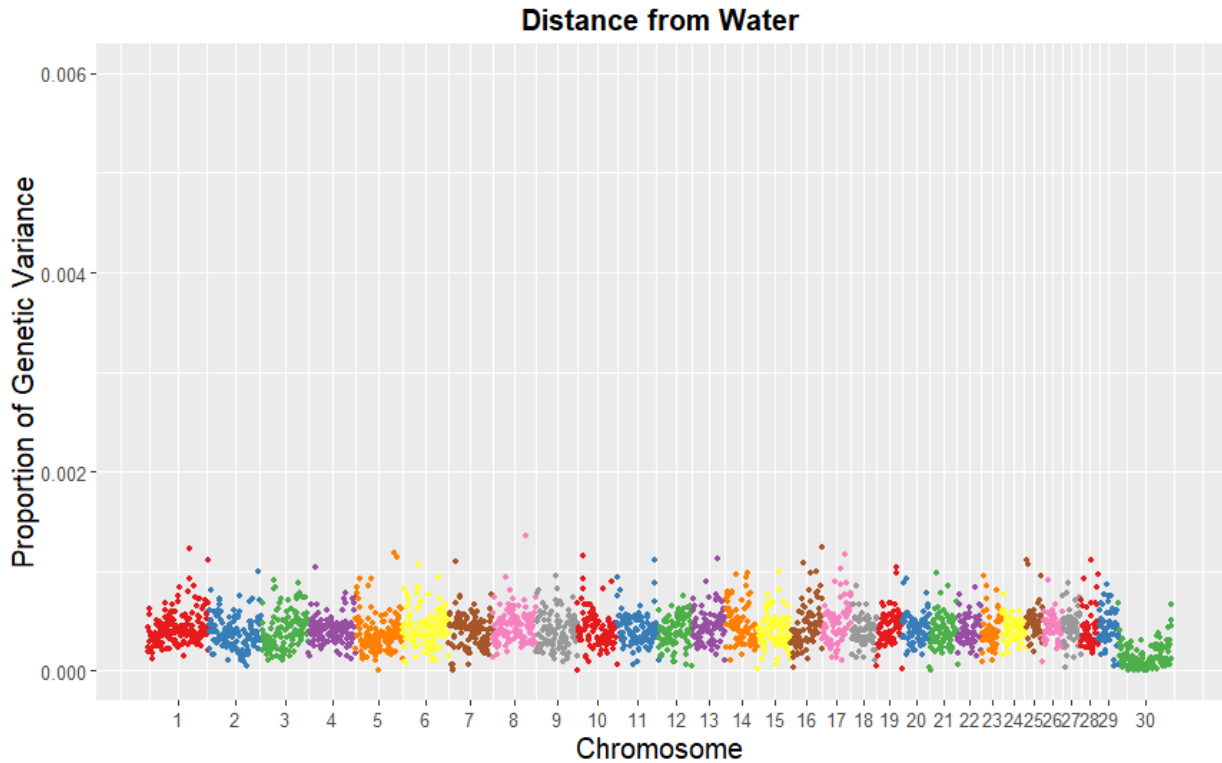
Figure 5.12. Manhattan plot of the proportion of genetic variance explained by markers in one megabase consecutive genomic windows for the distance beef cows (n = 323) travelled from water while managed in the western U.S.

Table 5.13. Top five genomic windows (1 Mb) identified in the multi-breed genome-wide association study for the distance beef cows (n = 323) travelled from water while managed in the western U.S.

Window[1]	Start[2] (bp)	End[3] (bp)	SNP[4]	Var[5], %	Lead SNP[6]	Gene[7]
8_86	86101796	86974867	18	0.0014	rs109226946	*PHF2*
16_80	80010239	80985485	22	0.0012	rs41598300	*NR5A2*
1_108	108000927	108935413	20	0.0012	rs43256975	*C1H3orf80*
5_101	101124171	101974400	15	0.0012	rs109175805	*A2ML1*
17_56	56023773	56963536	17	0.0012	rs109952637	*IFT81*

[1]Genomic window: chromosome and n^{th} 1 Mb window on that chromosome.
[2]Start of the 1 Mb window (bp).
[3]End of the 1 Mb window (bp).
[4]Number of SNP within the window.
[5]Percentage of genetic variance explained by the genomic window.
[6]Lead SNP: SNP with highest PIP.
[7]Closest annotated gene.

Rolling Index

The Bayesian GWAS using the rolling index identified candidate SNP on chromosome 10, 13, 18, and 24 (Figure 5.13). The highest PIP was associated with rs29021957 (PIP = 0.06) an intergenic variant located on chromosome 24. The second highest posterior inclusion probability (PIP = 0.05) was associated with rs110514275 on chromosome 10 (Table 5.14). The other top SNP (rs41582500, rs41576569, and rs11003025) had PIP of 0.04; therefore, 4% of samples included these SNP in the model. The top five genomic windows included the 63rd window on chromosome 13, the 55th window on chromosome eight, the 6th window on chromosome 24, the 15th window on chromosome 10, and the 45th window on chromosome 16 (Figure 5.14; Table 5.15). The SNP within the consecutive, one megabase genomic windows explained 0.99% of the genetic variance. The concordance between the candidate SNP and the lead SNP within the genomic windows suggested that four of the five candidate SNP were important for terrain-use.

The SNP identified in this study were compared to the eight candidate SNP that were previously associated with the rolling index (Bailey et al., 2015). Table 5.16 provides a description of the SNP identified in 2015 and their associated candidate genes. Unfortunately, the SNP identified using Bayesian approach did not parallel those identified by Bailey et al. (2015).
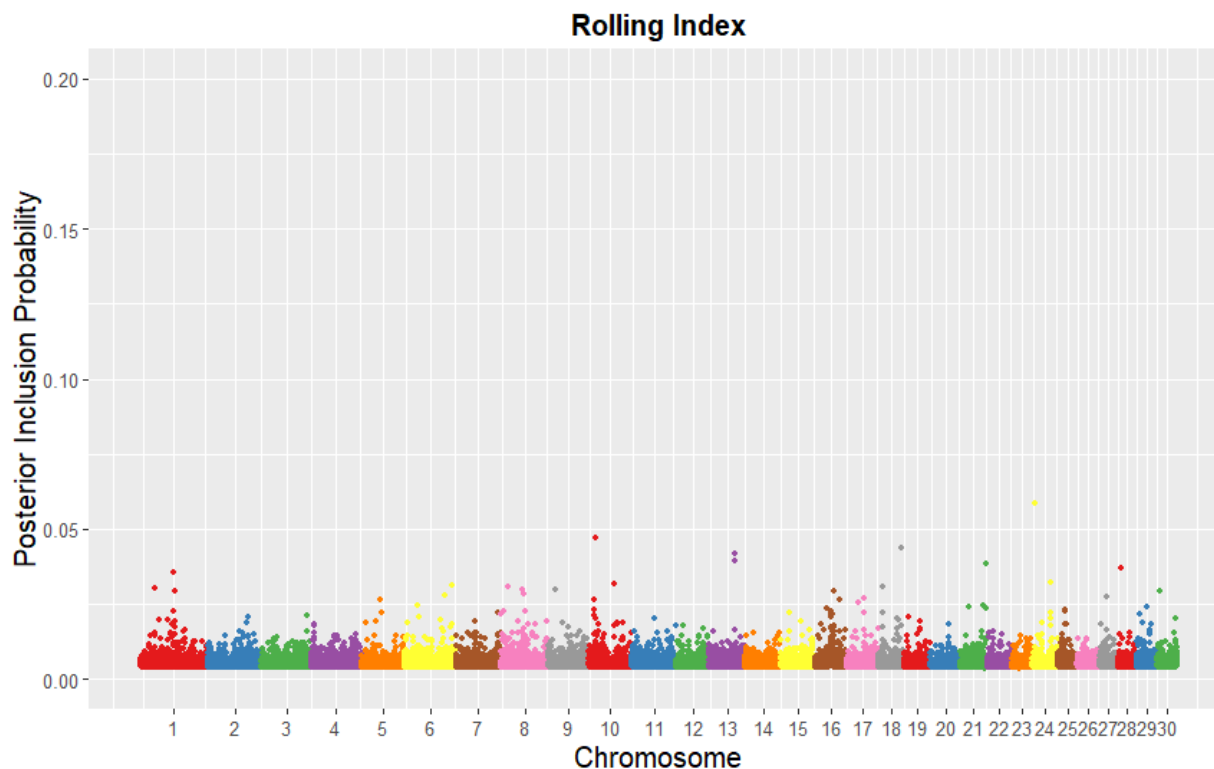
Figure 5.13. Manhattan plot of a multi-breed genome-wide association study for the rolling index values of 323 beef cows managed in the western U.S.

Table 5.14. Top five single nucleotide polymorphisms (SNP) identified in the multi-breed genome-wide association study for the rolling index values of beef cows (n = 323) managed in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Window[4]	Gene[5]	SNP location[6]	PIP[7]
rs29021957	24	6276349	24_6	*CBLN2*	.655 Mb	0.06
rs110514275	10	15202658	10_15	*ITGA11*	0.006 Mb	0.05
rs41582500	18	57115739	18_57	*LRRC4B*	Intron	0.04
rs41576569	13	63369536	13_63	*CDK5RAP1*	Intron	0.04
rs110030253	13	63391193	13_63	*CDK5RAP1*	Intron	0.04

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Genomic window: chromosome and n^{th} 1 Mb window on that chromosome.
[5]Closest annotated gene.
[6]Location of SNP within or near the gene.
[7]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
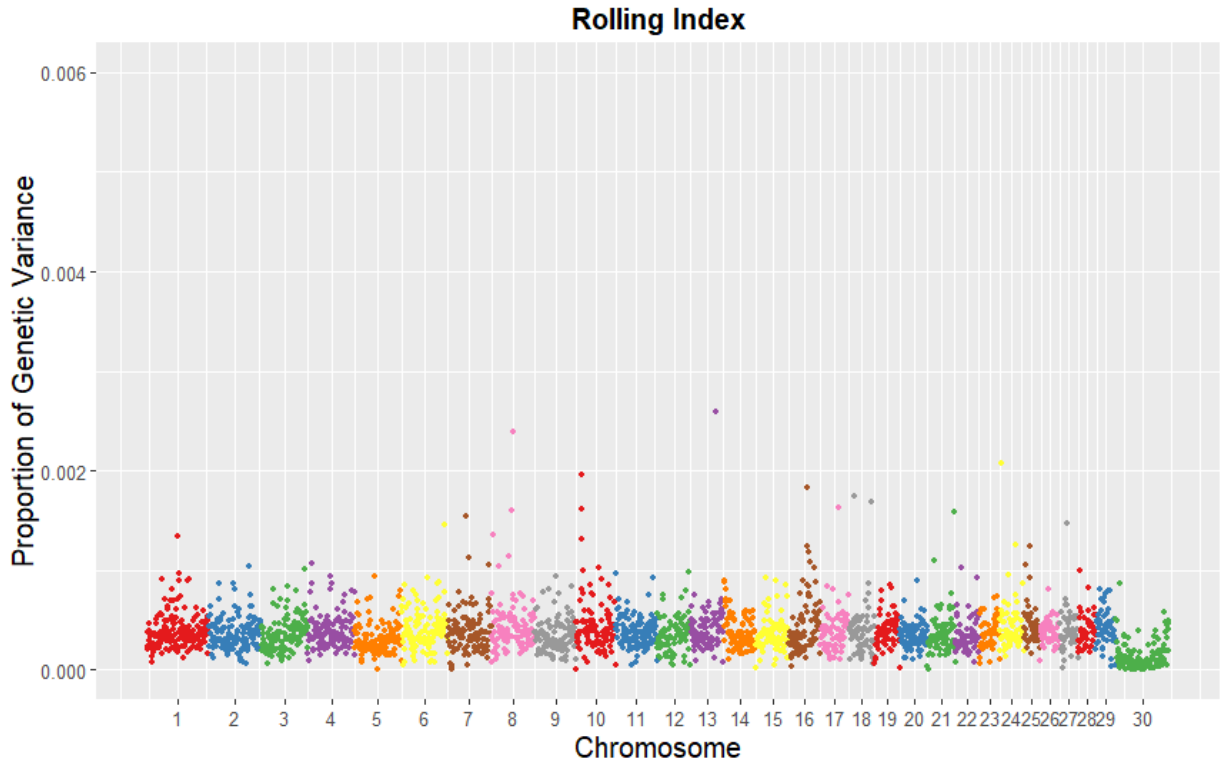
Figure 5.14. Manhattan plot of the proportion of genetic variance explained by markers in one megabase consecutive genomic windows for the rolling index values of 323 beef cows managed in the western U.S.

Table 5.15. Top five genomic windows (1 Mb) identified in the multi-breed genome-wide association study for the rolling index values of beef cows (n = 323) managed in the western U.S.

Window[1]	Start[2] (bp)	End[3] (bp)	SNP[4]	Var[5], %	Lead SNP[6]	Gene[7]
13_63	63015278	63978193	23	0.0026	rs41576569, rs110030253	*CDK5RAP1*
8_55	55004792	55996200	22	0.0024	rs41619378	*TLE4*
24_6	6047561	6914110	17	0.0021	rs29021957	*CBLN2*
10_15	15072778	15974499	21	0.0020	rs110514275	*ITGA11*
16_45	45017787	45552538	10	0.0018	rs41811366	*CA6*

[1]Genomic window: chromosome and n[th] 1 Mb window on that chromosome.
[2]Start of the 1 Mb window (bp).
[3]End of the 1 Mb window (bp).
[4]Number of SNP within the window.
[5]Percentage of genetic variance explained by the genomic window.
[6]Lead SNP: SNP with highest PIP.
[7]Closest annotated gene.

Table 5.16. Description of candidate single nucleotide polymorphisms (SNP) previously associated with the rolling index through single marker regression (Bailey et al., 2015).

RS ID[1]	CHR[2]	Position[3]	Candidate Gene[4]
rs134515496	4	14487987	*SDHAF3*
rs109619368	17	18299593	*MAML3*
-	29	6598207	*GRM5*
rs42161939	29	7083900	*GRM5*
rs43744222	29	7128587	*GRM5*
-	29	7128668	*GRM5*
-	29	7240505	*GRM5*
rs42162708	29	7241306	*GRM5*

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Closest annotated gene.

Rough Index

The Manhattan plot for the rough index revealed peaks on chromosome 10, 24, 27, and 29 (Figure 5.15). The marker identified on chromosome 10, rs29013509, was also one of the candidate markers for slope (Table 5.6). In the rough index GWAS, the PIP for rs29013509 was 0.14 whereas in the slope analysis the PIP was 0.13 (Table 5.17). The marker with the second highest PIP (0.08) was identified on chromosome 29 (rs110590993) as was the SNP with the third highest posterior inclusion probability (rs43703968; PIP = 0.07). Bailey et al. (2015) discovered two SNP on chromosome 29 that were associated with the rough index: rs42161939 and rs43744222 (Table 5.19). The markers identified in this Bayesian regression were located approximately 6.8 Mb to 10.9 Mb downstream from the SNP identified in 2015. The two-other candidate SNP identified in this analysis had PIP of 0.07 and were located on chromosome 24 (rs110959252) and 27 (rs42120868).

Three of the five genomic windows for the rough index contained lead SNP that were also considered candidate SNP (rs29013509, rs110959252, and rs42120868). The proportion of genetic variance explained by the 20th window on chromosome 10 was 0.0047% making it the top window in this analysis. The 42nd window on chromosome 24 explained 0.0041% of the genetic variance and the 17th window on chromosome 27 explained 0.0036% of the genetic variance for the rough index (Figure 5.16; Table 5.18). The cumulative proportion of genetic variance explained by the genomic windows was 0.99%. As with the individual marker results, several of the genomic windows for the rough index matched those identified for slope. The 20th and 38th window on chromosome 10 as well as the 61st window on chromosome four were part of the top five windows in both analyses. In general, these genomic windows explained a larger proportion of genetic variance for slope than for the rough index.
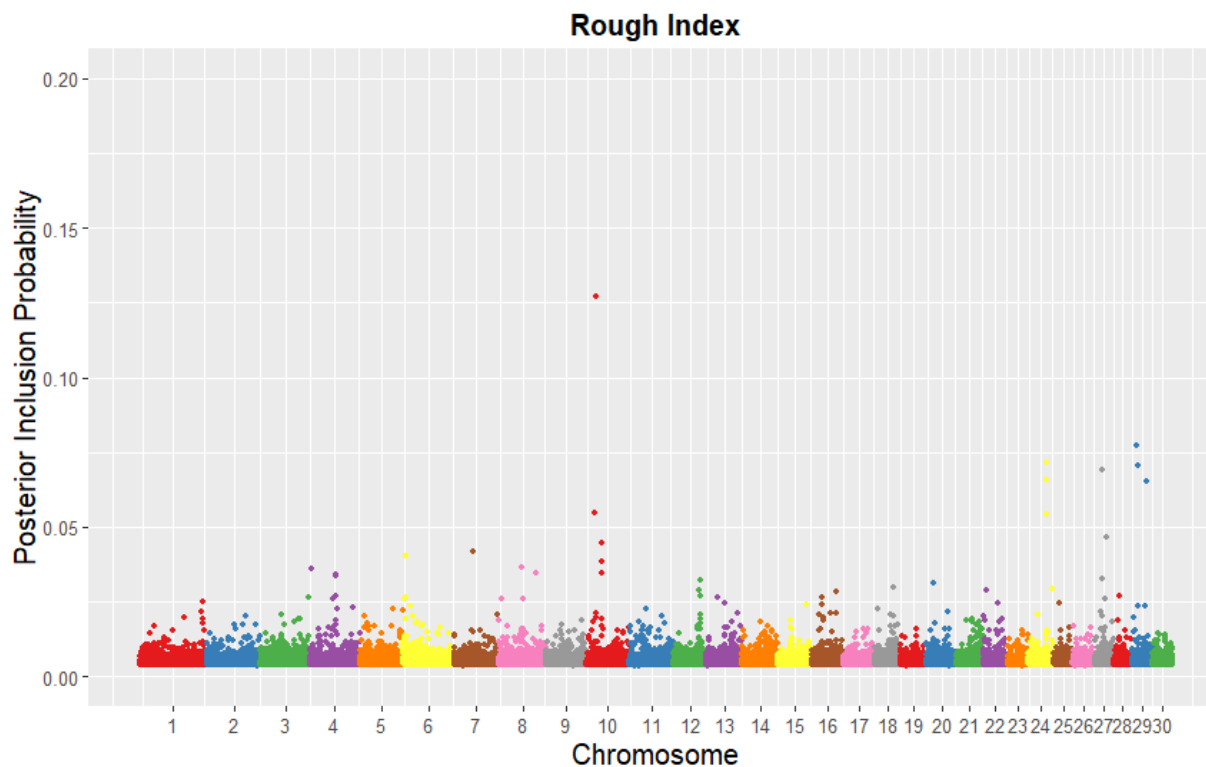
Figure 5.15. Manhattan plot of a multi-breed genome-wide association study for the rough index values of 323 beef cows managed in the western U.S.

Table 5.17. Top five single nucleotide polymorphisms (SNP) identified in the multi-breed genome-wide association study for the rough index values of beef cows (n = 323) managed in the western U.S.

RS ID[1]	CHR[2]	Position[3]	Window[4]	Gene[5]	SNP location[6]	PIP[7]
rs29013509	10	20498087	10_20	*TBC1D21*	0.026 Mb	0.13
rs110590993	29	11905442	29_11	*ENSBTAG00000027868*	0.200 Mb	0.08
rs110959252	24	42665926	24_42	*PIEZO2*	Intron	0.07
rs43703968	29	13959142	29_13	*ENSBTAG00000022427*	0.377 Mb	0.07
rs42120868	27	17956728	27_17	*FRG1*	0.288 Mb	0.07

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Genomic window: chromosome and nth 1 Mb window on that chromosome.
[5]Closest annotated gene.
[6]Location of SNP within or near the gene.
[7]Posterior inclusion probability: proportion of models (out of 1.0) in the MCMC chain that included the given SNP.
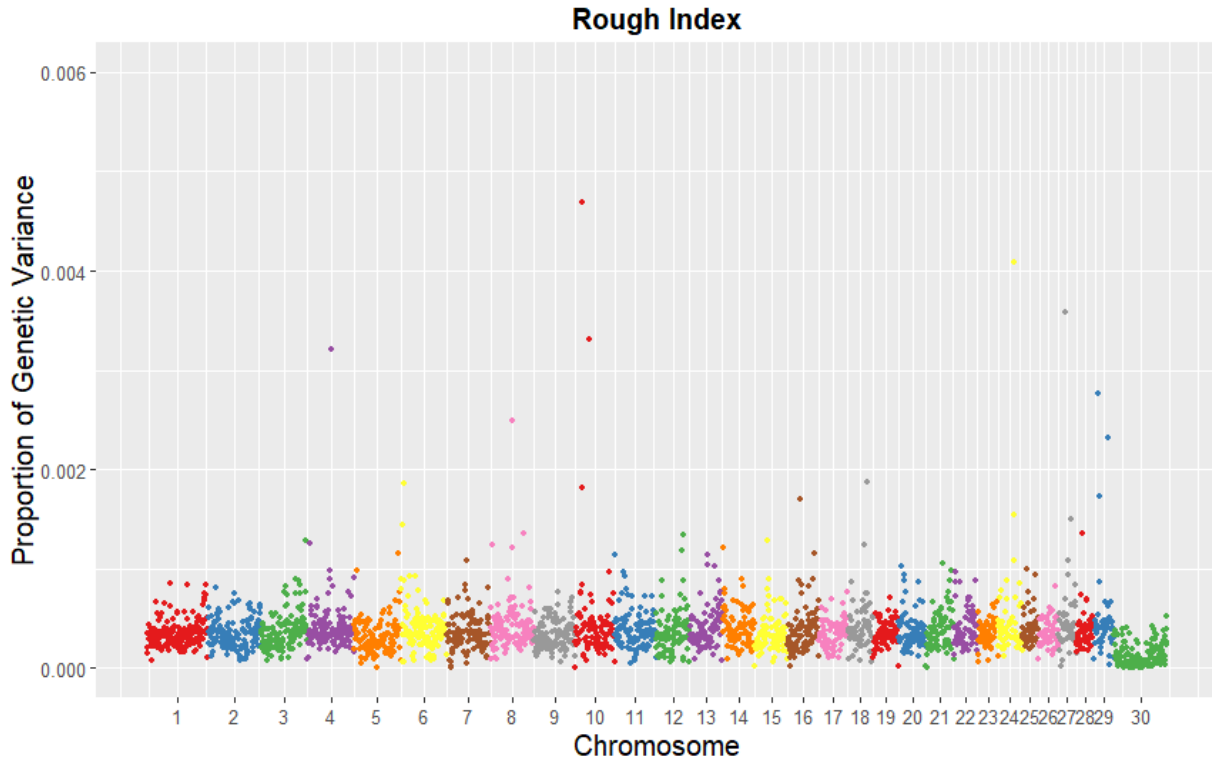
Figure 5.16. Manhattan plot of the proportion of genetic variance explained by markers in one megabase consecutive genomic windows for the rough index values of 323 beef cows in the western U.S.

Table 5.18. Top five genomic windows (1 Mb) identified in the multi-breed genome-wide association study for the rough index values of beef cows (n = 323) managed in the western U.S.

Window[1]	Start[2] (bp)	End[3] (bp)	SNP[4]	Var,[5] %	Lead SNP[6]	Gene[7]
10_20	20008636	20932671	20	0.0047	rs29013509	*TBC1D21*
24_42	42048793	42967505	17	0.0041	rs110959252	*PIEZO2*
27_17	17007677	17956728	18	0.0036	rs42120868	*FRG1*
10_38	38026306	38986604	15	0.0033	rs29013631	*EPB42*
4_61	61015911	61972879	24	0.0032	rs43109323	*SEPT7*

[1]Genomic window: chromosome and n[th] 1 Mb window on that chromosome.
[2]Start of the 1 Mb window (bp).
[3]End of the 1 Mb window (bp).
[4]Number of SNP within the window.
[5]Percentage of genetic variance explained by the genomic window.
[6]Lead SNP: SNP with highest PIP.
[7]Closest annotated gene.

Table 5.19. Description of candidate single nucleotide polymorphisms (SNP) previously associated with the rough index through single marker regression (Bailey et al., 2015).

RS ID[1]	CHR[2]	Position[3]	Candidate Gene[4]
rs134515496	4	14487987	*SDHAF3*
-	8	60157511	*RUSC2*
-	12	24598260	*SUPT20H*
rs110062743	12	24593452	*SUPT20H*
rs109619368	17	18299593	*MAML3*
rs109619368	17	18299593	*MAML3*
rs42161939	29	7083900	*GRM5*
rs43744222	29	7128587	*GRM5*

[1]Reference SNP cluster identification assigned by National Center for Biotechnology Information (NCBI).
[2]Chromosome in which the SNP is located according to *Bos taurus* UMD3.1.1.
[3]Chromosome position in *Bos taurus* UMD3.1.1.
[4]Closest annotated gene.

Discussion

Genome-wide association studies using data collected from cows managed on rangeland operations in the western United States identified QTL that may be important for terrain-use in extensive rangeland pastures. In total, 30 candidate SNP were identified for slope, elevation, vertical climb, distance travelled from water, rolling index and rough index. Four of the 30 SNP were associated with more than one terrain-use trait; therefore, 26 unique SNP, on 17 chromosomes, were discovered to have an association with at least one terrain use outcome. The posterior inclusion probabilities for the markers in the six GWAS were considered low. The maximum PIP identified for a single SNP across the six analyses was 18% (rs41600226 on chromosome 23) which meant the highest rate of inclusion was 27,000 out of 150,000 samples. Moreover, the mean PIP for elevation and vertical climb was 0.6% whereas the mean PIP for slope, distance travelled from water, rolling index and rough index was 0.5%.

Although the PIP were low, Manhattan plots for five of the six traits had notable peaks and there is no standard threshold for determining significance based on posterior inclusion probability. Furthermore, previous mammalian studies suggested that posterior inclusion probability varies greatly depending upon the trait of interest. Wolc et al. (2012) examined QTL associated with laying hen egg weight and selected candidate SNP based upon the highest PIP and largest proportion of genetic variance explained. In the study of Wolc et al. (2012), the candidate SNP had PIP ranging from 0.19 to 1.00. Speidel et al. (2018) examined the genetic architecture of heifer pregnancy and stayability in Red Angus cattle and set a PIP threshold for each trait. Speidel et al. (2018) reported lower PIP for heifer pregnancy than stayability. More specifically, the maximum PIP for heifer pregnancy was 6% whereas the max PIP for stayability was 100%. Speidel et al. (2018) hypothesized that the low inclusion of markers for heifer pregnancy was due to a lack of phenotypic data (n = 567) and the low heritability of the trait ($h^2 = 0.12$). Perhaps, the PIP for the terrain-use traits presented in this study have been influenced by similar factors. The moderate sample size of this study limited the statistical power for SNP detection and results suggested that terrain-use traits may be lowly heritable. Yet, this study was challenged by landscape diversity among ranches, breed of cattle, and GPS monitoring attributes. Higher quality data with more uniformity could improve QTL detection and SNP effect estimates.

Twenty-four genomic windows, spanning 13 chromosomes, were associated with the six terrain-use traits and six of the windows (4_61, 11_91, 10_20, 10_38, 23_47, and 24_18) were linked to multiple traits. Within the genomic windows, 25 lead SNP were identified. Parallel with the PIP, the genomic windows explained a low proportion of genetic variance for the terrain-use traits. Across the six traits, the maximum proportion of genetic variance explained by a single window was 0.0057%. Again, these estimates may have been influenced by the heritability and

marker density (i.e., the number of markers captured in the genomic window). A greater proportion of genetic variance can be explained using dense genotyping arrays as there is greater linkage disequilibrium between causative mutations and the surrounding markers (Jensen et al., 2012). Another important consideration regarding the proportion of genetic variance explained is the genetic architecture underlying quantitative traits. As discussed by Mackay (2001) and Moser et al. (2009), complex quantitative traits are influenced by many QTL with small effects (i.e., QTL explain a small percentage of total genetic variance on an individual basis). This assumption is in agreeance with Cole et al. (2009) who examined the effect of 38,416 SNP for 5,360 Holstein bulls on dairy traits.

A review of the literature suggested that the low proportion of genetic variance explained by the QTL in this study is within the realm of percentages reported for other beef cattle traits. Peters et al. (2013) reported that the highest proportion of genetic variance explained by a single window for 205-day weight and 365-d weight, which are moderately to highly heritable traits, in Brangus heifers were 0.0203% and 0.0089%. Additionally, the maximum proportion of genetic variance explained for rib fat, intramuscular fat, and longissimus muscle were, 0.0200%, 0.0167%, and 0.0156%, respectively.

A comparison of the candidate SNP to the genomic windows revealed 32 QTL and 29 putative candidate genes that may play a role in beef cow terrain-use. Thirty of the QTL had been previously documented in Cattle QTLdb and linked to numerous beef cattle production traits (Appendix I). Unfortunately, the 29 putative candidate genes identified in this analysis (Appendix II) were not concordant with the five genes identified by Bailey et al. (2015) nor did they validate the genes discovered in chapter four. Confirmation of the previously identified genes may have been hindered by differences in SNP density (high-density genotypes vs. 50k genotypes),

statistical methodology (single-SNP regression vs. Bayesian multiple-SNP regression), genomic architecture, and extreme topography variation between rangeland beef cattle operations.

Bailey et al. (2015) used BovineHD genotypes and single-SNP to identify QTL associated with beef cattle terrain-use. Significant markers identified in this analysis were used to develop a custom-genotyping panel for subsequent GWAS. The genotype data in this study was derived from a BovineSNP50 panel; therefore, the eight candidate SNP discovered by Bailey et al. (2015) were not replicated in this study. In this study, Bayesian methodology was used instead of single-SNP regression because simultaneously fitting markers accounts for all SNP in linkage disequilibrium (LD) with the QTL which decreases the proportion of unexplained genetic variance (Hayes et al., 2010; Fan et al., 2011; Dekkers, 2012; Fernando et al., 2017). In addition, Bayesian multiple-regression results in fewer false positives because population structure was explained in the model (Dekkers, 2012). Differences in the genomic architecture (i.e., allele or genotypic frequencies) between study populations most likely influenced QTL detection (Greene et al., 2009). Crawford et al. (2018) reported a similar scenario in which a linear mixed model analysis using data from cattle maintained at the CSU-BIC (n = 532) failed to confirm a previously reported association between mean pulmonary arterial pressures (mPAP) and the A allele of rs208684340 in *EPAS1* (Newman et al., 2015). Crawford et al. (2018) acknowledged the potential contribution of genetic architecture to the lack of validation between the two studies.

The putative candidate genes identified in this study were five physiological and production categories: cardiovascular system, growth traits and feed efficiency, energy metabolism, heat stress, and lactation. Seven of the putative candidate genes function in the cardiovascular system: *CBLN2*, *EDIL3*, *EEF1E1*, *ENSBTAG00000032603*, *EPB42, PIEZO2,* and *TBX3*. A GWAS for pulmonary arterial hypertension (PAH) in humans revealed an association

between Cerebellin 2 Precursor (*CBLN2*) and PAH susceptibility. More specifically, the SNP in close proximately (0.052 Mb upstream) to *CBLN2*, rs2217560, had an odds ratio of 1.97 [1.59 – 2.45] (Germain et al., 2013). EGF Like Repeats And Discoidin Domains 3 (*EDIL3)* has a role in the regulation of angiogenesis and may be involved in the development and remodeling of vessel walls (Ho et al., 2004). Previous studies in mammals suggested that angiogenesis may be stimulated under hypoxic conditions to ensure that cells are supplied with oxygen (Fong, 2008; Krock et al., 2011). EGF Like Repeats and Discoidin Domains 3 was differentially expressed in right ventricle tissues of Angus steers with high pulmonary arterial pressure (PAP) measures as compared to those with low PAP measures (unpublished data; N.F. Crawford). Eukaryotic Translation Elongation Factor 1 Epsilon 1 (*EEF1E1*) has been associated with coronary arterial calcification (CAC) in humans and CAC is predominant in individuals with coronary heart disease (Wojczynski et al., 2013; Liu et al., 2015). *ENSBTAG00000032603* was identified in a proteomics analysis that examined the effect of prenatal hypoxia, induced by highland environments, on fetal sheep heart development (Li et al., 2018). *ENSBTAG00000032603* is also a documented orthologue of *ABCC4* (ENSG00000125257)*,* a gene that may be involved in the development of Kawasaki disease in humans. Kawasaki disease causes vasculitis of arteries and can result in fatal coronary arterial aneurysm (Khor et al., 2011).

In humans, Erythrocyte Membrane Protein Band 4.2 (*EPB42)* is involved in the regulation of red blood cell shape and function. Erythrocytes rely on deformability (i.e., the ability to change in shape) to navigate small capillaries; therefore, deformation is critical for circulation and subsequently oxygen and carbon dioxide transport. Altered shape of erythrocytes may hinder deformation and negatively impact circulation (Diez-Silva et al., 2010). Piezo Type Mechanosensitive Ion Channel Component 2 (*PIEZO2*) has a role in mechanically-activated cation

channels that are key component of mechanotransduction. Mechanotransduction regulates many physiological processes including: vascular tone, blood flow, lung growth, bone and muscle homeostasis (Coste et al., 2010). Like *EDIL3, PIEZO2* was also differentially expressed in right ventricle tissues of high PAP steers as compared to low PAP steers (unpublished data N.F. Crawford). A GWAS by Levy et al. (2009) suggested that T-Box 3 (*TBX3)* may be involved in diastolic blood pressure of humans.

In this study, all of the ranches except Carter Ranch were in moderate (1,200 to 1,600 m) to high-elevation regions (≥1,600 m) as classified by Pauling et al. (2018), where a reduction in atmospheric pressure results in lower partial pressure of oxygen (PaO2; Peacock, 1998). This may help explain the association between the previously discussed candidate genes and the terrain-use traits. Cattle that are utilizing rugged terrain in high altitude regions may experience hypoxia as oxygen consumption increases in active muscle cells (Hoppeler and Weibel, 2002). In hypoxic conditions, mammals may undergo physiological changes to compensate for the lack of oxygen including angiogenesis, erythrocyte modification, and heart remodeling (Holt and Callan, 2007; Bharti et al., 2011; Krock et al., 2011). Furthermore, cattle maintained in high-elevation regions may develop high altitude disease as a result of hypoxia induced pulmonary arterial hypertension (Holt and Callan, 2007). Bailey et al. (2016) examined the relationship between terrain-use of Angus cows managed at the CSU-BIC (elevation of 2,150 to 2,411 m) and pulmonary arterial pressure (PAP) as PAP measurements are an indicator of pulmonary arterial hypertension. Results of this study suggested no significant correlation (r = 0.23) between PAP measurements and terrain-use traits; however, the authors acknowledged that the study population may have consisted of Angus beef cows adapted to high-elevation; therefore, additional studies with harsher terrain, higher elevations, and non-adapted cattle may reveal a relationship between PAP and terrain-use.

As previously discussed, three of the putative candidate genes (*INHBA, LRRC4B,* and *SEPT7*) were associated with growth traits and feed efficiency in cattle. A global gene expression profiling using liver samples revealed that Inhibin Subunit Beta A (*INHBA*) was highly expressed in samples collected from Angus bulls with low residual feed intake (RFI; Chen et al., 2011). Quantitative real-time PCR for liver samples confirmed that *INHBA* was upregulated in Angus bulls with low RFI (Chen et al., 2012). A GWAS conducted by Hardie et al. (2017) revealed an association between the genomic region (Mb 57 on BTA18) containing Leucine Rich Repeat Containing 4B and metabolic body weight in multiparous Holstein cows. As discussed by Hardie et al. (2017), "in mid-lactation dairy cows, RFI is often computed as the residual of the regression of intake on a form of ECM production, metabolic BW (MBW), and energy gained or lost in body tissues." Septin 7 (*SEPT7*) has been associated with weaning weight direct and yearling weight in Maine-Anjou cattle (Saatchi et al., 2014).

The functions of *INHBA, LRRC4B,* and *SEPT7* are particularly interesting given the potential relationship between cow-size and terrain-use and locomotion and RFI. A preliminary study by Gannon et al. (2018) suggested that larger Brangus cows (i.e., heavier mature weight, larger heart girths and greater hip heights) travelled further from water than smaller cows when grazing in extensive, rugged rangeland pasture during winter months. Herd et al. (2004) identified a positive correlation (0.32) between daily pedometer counts and RFI. Perhaps, larger body size is favorable for increased terrain-use whereas increased terrain-use, that requires greater energy expenditure, is unfavorable for RFI. Additional studies with an independent population are needed to further examination the relationship between mature cow weight/size, residual feed intake, and terrain-use traits.

Regarding energy metabolism, ADP Dependent Glucokinase (*ADPGK*) catalyzes the phosphorylation of glucose to glucose-6-phosphate using adenosine di phosphate (ADP) and this process mediates the first step of glycolysis (Ronimus and Morgan, 2004; Richter et al., 2012). During glycolysis, glucose is converted into pyruvate during which adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide plus hydrogen (NADH) are generated. Pyruvate can then be used to produce volatile fatty acids (VFA's), methane, and carbon dioxide (Cheeke and Dierenfeld, 2010). Volatile fatty acids are a major energy source for ruminants and cattle expend more energy walking up hill than they would walking on gentle terrain (Brosh et al., 2006; Freer et al., 2007). In addition, under hypoxic conditions the rate of glycolysis may increase to help compensate for a reduction in oxidative phosphorylation (Fong, 2008), a process in which ATP are synthesized (Berg et al., 2002). Again, additional studies are needed to investigate the relationship between energy metabolism and terrain-use of beef cattle in the western U.S.

Another interesting candidate gene identified in this terrain-use study was PHD finger protein 2 (*PHF2*). In a study designed to understand high-altitude adaption of Ladakhi cattle, Verma et al. (2018) compared transcriptome signatures of peripheral blood mononuclear cells from Ladakhi cattle to the signatures of a tropically adapted breed (Sahiwal). Results suggested that *PHF2* was upregulated in the Sahiwal cattle maintained in an arid/semi-arid region. Kolli et al. (2014) reported that *PHF2* was downregulated in the leukocytes of Zebu cattle that experienced heat stress. Both feedlot cattle and those grazing in pastures can experience heat stress (Birkelo et al., 1991) and out of the 20 groups of cattle tracked during this study (Table 5.2), twelve groups were tracked during the summer months when higher temperatures can cause heat stress. As summarized by Silanikove (2000), domestic ruminants may reduce their locomotion

during periods of extreme heat and instead seek water and shade in an attempt to maintain cooler body temperatures.

One of the putative candidate genes, ATPase plasma membrane Ca^{2+} transporting 2 (*ATP2B2*), is involved in intracellular calcium homeostasis, which is critical for maintaining eukaryotic cell function (Garcia and Strehler, 1999). In mammals, ATPase plasma membrane Ca2+ transporting 2 has been associated with milk traits (Ogorevc et al., 2009) and a comparison of milk fat globule membrane proteins in milk samples collected from humans and cattle revealed the presence of *ATP2B2* in both human and cattle milk (Zhang et al., 2017). In this study, a SNP within *ATP2B2* was associated with vertical climb and approximately 56% of the cattle included in this GWAS were lactating during the study period. In a two year grazing study by Bailey et al. (2001a), nonlactating cows grazed at greater vertical distances from water than lactating cows. During the second year of the study, nonlactating cows grazed on steeper slopes than lactating cows. These results suggest that nonlactating cows utilize rugged terrain more efficiently than lactating cows.

Facioscapulohumeral muscular dystrophy (FSHD) region gene 1 (*FRG1*), which was in close proximity to a SNP (0.288 Mb) associated with the rough index, is a highly conserved gene in both invertebrates and vertebrates which suggests it has an important biological function (Grewal et al., 1998). As summarized by Sun et al. (2011a), *FRG1* is critical for the development of the muscular and vascular system. In humans, FSHD region gene 1 has been associated with facioscapulohumeral muscular dystrophy a disorder characterized by muscle weakness and atrophy (Ferri et al., 2015).

The quantitative trait loci detected in this study support the polygenic nature of complex traits; moreover, varying gene functions observed in these QTL allude to the interaction of

biological pathways regulating beef cow terrain-use traits. Detection was limited by the moderate sample size and lack of uniformity in the data. Therefore, a large independent population of beef cows, composed of one breed, grazing on uniform pastures or larger groups in variable pastures is needed to refine terrain-use measurements and further elucidate the role of genetics in beef cattle terrain-use phenotypes.

Conclusion

The objective of this study was to perform GWAS for six terrain-use traits to identify individual markers and genomic windows associated with beef cattle terrain-use in the western U.S. The Bayesian-based analyses, using data from 330 beef cows managed on 14 ranches, revealed 32 SNP and 29 putative candidate genes for terrain-use. Many of the QTL were previously associated with beef and dairy cattle health and performance traits. Four of the 29 putative genes lacked functional annotation; however, the remaining 25 genes were related to a variety of biological processes including hypoxia, feed efficiency, heat stress, and glycolysis.

REFERENCES

Adelson, D. L., Z. Qu, J. M. Raison, and S. L. Lim. 2014. Bovine genomics. In: D. Garrick and A. Ruvinsky, editors, The genetics of cattle. CABI, Boston, MA. p. 130-152.

Allan, B. M., J. P. Arnould, J. K. Martin, and E. G. Ritchie. 2013. A cost-effective and informative method of gps tracking wildlife. Wildl. Res. 40:345-348. doi:10.1071/WR13069

Anderson, C. A., F. H. Pettersson, G. M. Clarke, L. R. Cardon, A. P. Morris, and K. T. Zondervan. 2010. Data quality control in genetic case-control association studies. Nat. Protoc. 5:1564-1573. doi:https://doi.org/10.1038/nprot.2010.116

Anderson, D. M., R. E. Estell, and A. F. Cibils. 2013. Spatiotemporal cattle data—a plea for protocol standardization. Positioning. 4:115-136. doi:10.4236/pos.2013.41012

Antonarakis, S. E., and D. N. Cooper. 2013. Human gene mutation in inherited disease: Molecular mechanisms and clinical consequences. In: D. Rimoin, R. Pyeritz and K. Bruce, editors, Emery and rimoin's principles and practice of medical genetics. Academic Press, Cambridge, MA. p. 1-48.

Archer, S., and F. E. Smeins. 1991. Ecosystem - level processes. In: R. K. Heitschmidt and J. W. Stuth, editors, Grazing management: An ecological perspective. Timber Press, Portland, OR. p. 109-139.

Armstrong, D., P. Kotaras, and C. Earl. 1997. Advances in production of embryos in vitro from juvenile and prepubertal oocytes from the calf and lamb. Reprod. Fertil. Dev. 9:333-340.

Ashton, M. S. 1998. Seedling ecology of mixed-dipterocarp forest. In: S. Appanah and J. M. Turnbull, editors, A review of dipterocarps: Taxonomy, ecology, and silviculture. CIFOR, Bogor, Indonesia. p. 89-98.

Atherton, L. 1972. The cattle kings. U of Nebraska Press, Lincoln, NE.

Azad, M. B., Y. Chen, E. S. Henson, J. Cizeau, E. McMillan-Ward, S. J. Israels, and S. B. Gibson. 2008. Hypoxia induces autophagic cell death in apoptosis-competent cells through a mechanism involving bnip3. Autophagy. 4:195-204. doi:10.4161/auto.5278

Bailey, D. 2004. Management strategies for optimal grazing distribution and use of arid rangelands. J. Anim. Sci. 82:E147-E153. doi:10.2527/2004.8213_supplE147x

Bailey, D., D. Kress, D. Anderson, D. Boss, and K. Davis. 2001a. Evaluation of f1 crosses from angus, charolais, salers, piedmontese, tarentaise and hereford sires: Grazing distribution patterns. J. Anim. Sci. 52:110-113.

Bailey, D., D. Kress, D. Anderson, D. Boss, and E. Miller. 2001b. Relationship between terrain use and performance of beef cows grazing foothill rangeland. J. Anim. Sci. 79:1883-1891.

Bailey, D., M. Trotter, C. Knight, and M. G. Thomas. 2017. Use of gps tracking collars and accelerometers for rangeland livestock production research. J. Anim. Sci. 95:360. doi:https://doi.org/10.2527/asasann.2017.740

Bailey, D., H. VanWagoner, R. Weinmeister, and D. Jensen. 2008. Comparison of low-moisture blocks and salt for manipulating grazing patterns of beef cows. J. Anim. Sci. 86:1271-1277. doi:10.2527/jas.2007-0578

Bailey, D. W. 1999. Influence of species, breed and type of animal on habitat selection. Proc. Grazing Behav. Livest. Wildl. Moscow, ID. p. 101-108.

Bailey, D. W. 2000. Evaluating new approaches to improve livestock grazing distribution using gps and gis technology. Proc. 1st Nat. Conf. Grazing Lands. Las Vegas, NV. p. 91-99.

Bailey, D. W. 2005. Identification and creation of optimum habitat conditions for livestock. Rangeland Ecol. Manag. 58:109-118.

Bailey, D. W., J. E. Gross, E. A. Laca, L. R. Rittenhouse, M. B. Coughenour, D. M. Swift, and P. L. Sims. 1996. Mechanisms that result in large herbivore grazing distribution patterns. J. Range Manage. 49:386-400.

Bailey, D. W., M. R. Keil, and L. R. Rittenhouse. 2004. Research observation: Daily movement patterns of hill climbing and bottom dwelling cows. J. Range Manage. 57:20-28. doi:10.2458/azu_jrm_v57i1_bailey

Bailey, D. W., S. Lunt, A. Lipka, M. G. Thomas, J. F. Medrano, A. Cánovas, G. Rincon, M. B. Stephenson, and D. Jensen. 2015. Genetic influences on cattle grazing distribution:

Association of genetic markers with terrain use in cattle. Rangeland Ecol. Manag. 68:142-149. doi:https://doi.org/10.1016/j.rama.2015.02.001

Bailey, D. W., and L. R. Rittenhouse. 1989. Management of cattle distribution. Rangelands. 11:159-161.

Bailey, D. W., M. G. Thomas, T. N. Holt, M. B. Stephenson, R. M. Enns, and S. E. Speidel. 2016. Relationship of pulmonary arterial pressure and terrain use of angus cows grazing high-altitude foothill rangelands. Livest. Sci. 190:76-80. doi:https://doi.org/10.1016/j.livsci.2016.06.003

Bailey, D. W., M. G. Trotter, C. W. Knight, and M. G. Thomas. 2018. Use of gps tracking collars and accelerometers for rangeland livestock production research. Trans. Anim. Sci. 2:81-88. doi:10.1093/tas/txx006

Bailey, D. W., H. C. VanWagoner, and R. Weinmeister. 2006. Individual animal selection has the potential to improve uniformity of grazing on foothill rangeland. Rangeland Ecol. Manag. 59:351-358. doi:10.2111/04-165R2.1

Balding, D. J. 2006. A tutorial on statistical methods for population association studies. Nat. Rev. Genet. 7:781-791. doi:10.1038/nrg1916

Barendse, W. 2005. The transition from quantitative trait loci to diagnostic test in cattle and other livestock. Aust. J. Exp. Agric. 45:831-836. doi:10.1071/EA05067

Bayer, M., J. Fischer, J. Kremerskothen, E. Ossendorf, T. Matanis, M. Konczal, T. Weide, and A. Barnekow. 2005. Identification and characterization of iporin as a novel interaction partner for rab1. BMC Cell Biol. 6:15. doi:10.1186/1471-2121-6-15

Bedell, T. E. 1998. Glossary of terms used in range management. https://globalrangelands.org/glossary (Accessed 7 November 2018)

Bentley, H. L. 1898. Cattle ranges of the southwest: A history of the exhaustion of the pasturage and suggestions for its restoration. Govt. print. off., Washington.

Bentley, J. 1941. Automatic recording of salting and watering habits of cattle. J. For. 39:832-836.

Berg, J. M., J. L. Tymoczko, and L. Stryer. 2002. Oxidative phosphorylation, Biochemistry. W.H. Freeman and Company, New York.

Berlin, K., S. Koren, C.-S. Chin, J. P. Drake, J. M. Landolin, and A. M. Phillippy. 2015. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. Nat. Biotechnol. 33:623. doi:10.1038/nbt.3238

Bharti, V. K., D. Gogoi, G. Charan, A. Biswas, P. Kumar, and R. B. and Srivastava. 2011. Normal haematological values of changthangi sheep and goat, muzaffarnagri sheep, crossbred sheep, hfcattle, zanskar ponies, and broiler chicken at high altitude. Proc. Int. Conf. Emerging Trends Food Health Security Cold Deserts. Leh, Ladakh. p. 23-25.

Bhuin, T., and J. K. Roy. 2014. Rab proteins: The key regulators of intracellular vesicle transport. Exp. Cell Res. 328:1-19. doi:https://doi.org/10.1016/j.yexcr.2014.07.027

Birkelo, C., D. Johnson, and H. Phetteplace. 1991. Maintenance requirements of beef cattle as affected by season on different planes of nutrition. J. Anim. Sci. 69:1214-1222.

Black, W., and B. Knapp Jr. 1936. A method of measuring performance in beef cattle. J. Anim. Sci. 1936:72-77.

Bohmanova, J., M. Sargolzaei, and F. S. Schenkel. 2010. Characteristics of linkage disequilibrium in north american holsteins. BMC Genomics. 11:421. doi:10.1186/1471-2164-11-421

Bohrer, V. L. 1975. The prehistoric and historic role of the cool-season grasses in the southwest. Econ. Bot. 29:199-207.

Boichard, D., H. Chung, R. Dassonneville, X. David, A. Eggen, S. Fritz, K. J. Gietzen, B. J. Hayes, C. T. Lawley, T. S. Sonstegard, C. P. Van Tassell, P. M. VanRaden, K. A. Viaud-Martinez, G. R. Wiggans, and for the Bovine LD Consortium. 2012. Design of a bovine low-density snp array optimized for imputation. PLoS One. 7:e34130. doi:10.1371/journal.pone.0034130

Boichard, D., V. Ducrocq, P. Croiseau, and S. Fritz. 2016. Genomic selection in domestic animals: Principles, applications and perspectives. C. R. Biol. 339:274-277. doi:https://doi.org/10.1016/j.crvi.2016.04.007

Bolstad, W. M., and J. M. Curran. 2017. Introduction to bayesian statistics. John Wiley & Sons, Incorporated, Hoboken, NJ.

Bourdon, R. M. 1997. Understanding animal breeding. Prentice Hall, Englewood Cliffs, NJ.

Briske, D. D., J. D. Derner, J. R. Brown, S. D. Fuhlendorf, W. R. Teague, K. M. Havstad, R. L. Gillen, A. J. Ash, and W. D. Willms. 2008. Rotational grazing on rangelands: Reconciliation of perception and experimental evidence. Rangeland Ecol. Manag. 61:3-17. doi:10.2111/06-159r.1

Briske, D. D., and J. H. Richards. 1995. Plant responses to defoliation: A physiological, morphological and demographic evaluation. In: D. J. Bedunah and R. E. Sosebee, editors, Wildland plants: Physiological ecology and developmental morphology. Society for Range Management, Denver, CO. p. 635-710.

Briske, D. D., N. F. Sayre, L. Huntsinger, M. Fernandez-Gimenez, B. Budd, and J. D. Derner. 2011. Origin, persistence, and resolution of the rotational grazing debate: Integrating human dimensions into rangeland research. Rangeland Ecol. Manag. 64:325-334. doi:10.2111/rem-d-10-00084.1

Brosh, A., Z. Henkin, E. Ungar, A. Dolev, A. Orlov, Y. Yehuda, and Y. Aharoni. 2006. Energy cost of cows' grazing activity: Use of the heart rate method and the global positioning system for direct field estimation. J. Anim. Sci. 84:1951-1967. doi:https://doi.org/10.2527/jas.2005-315

Burt, D. W. 2009. The cattle genome reveals its secrets. J. Biol. 8:36. doi:10.1186/jbiol137

Bush, W. S., and J. H. Moore. 2012. Genome-wide association studies. PLoS Comp. Biol. 8:e1002822. doi:10.1371/journal.pcbi.1002822

Butler, P. J. 2000. Cattle distribution under intensive herded management. Rangelands Archives. 22:21-23.

Cánovas, A., R. Cockrum, D. Brown, S. Riddle, J. Neary, T. Holt, J. Medrano, A. Islas-Trejo, R. Enns, and S. Speidel. 2016. Functional genomics of high altitude disease in angus cattle: Leveraging-omics and systems biology to better understanding of the function and role of key contributing genes. J. Anim. Sci. 94:78. doi:10.2527/jas2016.94supplement478x

Cánovas, A., A. Reverter, K. L. DeAtley, R. L. Ashley, M. L. Colgrave, M. R. Fortes, A. Islas-Trejo, S. Lehnert, L. Porto-Neto, and G. Rincón. 2014. Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. PLoS One. 9:e102551. doi:10.1371/journal.pone.0102551

Cánovas, A., G. Rincon, A. Islas-Trejo, S. Wickramasinghe, and J. F. Medrano. 2010. Snp discovery in the bovine milk transcriptome using rna-seq technology. Mamm. Genome. 21:592-598. doi:10.1007/s00335-010-9297-z

CAST. 2002. Environmental impacts of livestock on u.S. Grazing lands. http://www.cast-science.org/download.cfm?PublicationID=2898&File=1e30356397f710aaf8cc325112e371c22246TR (Accessed 25 August 2018)

Cheeke, P. R., and E. S. Dierenfeld. 2010. Carbohydrate digestion and absorption. In: P. R. Cheeke and E. S. Dierenfeld, editors, Comparative animal nutrition and metabolism. CABI, Cambridge, MA. p. 95-103.

Chen, B., J. Xu, X. He, H. Xu, G. Li, H. Du, Q. Nie, and X. Zhang. 2015. A genome-wide mrna screen and functional analysis reveal foxo3 as a candidate gene for chicken growth. PLoS One. 10:e0137087. doi:https://doi.org/10.1371/journal.pone.0137087

Chen, Y., P. Arthur, I. Barchia, K. Quinn, P. Parnell, and R. Herd. 2012. Using gene expression information obtained by quantitative real-time pcr to evaluate angus bulls divergently selected for feed efficiency. Anim. Prod. Sci. 52:1058-1067. doi:https://doi.org/10.1071/AN12098

Chen, Y., C. Gondro, K. Quinn, R. M. Herd, P. F. Parnell, and B. Vanselow. 2011. Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. Anim. Genet. 42:475-490. doi:https://doi.org/10.1111/j.1365-2052.2011.02182.x

Chitwood, J. L., G. Rincon, G. G. Kaiser, J. F. Medrano, and P. J. Ross. 2013. Rna-seq analysis of single bovine blastocysts. BMC Genomics. 14:350. doi:https://doi.org/10.1186/1471-2164-14-350

Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42:2. doi:10.1186/1297-9686-42-2

Clark, P. E., D. E. Johnson, M. A. Kniep, P. Jermann, B. Huttash, A. Wood, M. Johnson, C. McGillivan, and K. Titus. 2006. An advanced, low-cost, gps-based animal tracking system. Rangeland Ecol. Manag. 59:334-340. doi:10.2111/05-162R.1

Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. J. Dairy Sci. 92:2931-2946. doi:https://doi.org/10.3168/jds.2008-1762

Cook, C. W., and E. F. Redente. 1993. Development of the ranching industry in colorado. Rangelands. 15:204-207.

Coste, B., J. Mathur, M. Schmidt, T. J. Earley, S. Ranade, M. J. Petrus, A. E. Dubin, and A. Patapoutian. 2010. Piezo1 and piezo2 are essential components of distinct mechanically-activated cation channels. Science. 330:55-60. doi:https://doi.org/10.1126/science.1193270

Crawford, N., S. Coleman, T. Holt, S. Speidel, R. Enns, R. Hamid, and M. Thomas. 2018. Allele distribution and testing for association between an oxygen dependent degradation domain snp in epas1 and pulmonary arterial pressures in yearling angus cattle. Agri Gene. 9:27-31. doi:https://doi.org/10.1016/j.aggene.2018.07.004

Daetwyler, H. D., A. Capitan, H. Pausch, P. Stothard, R. Van Binsbergen, R. F. Brøndum, X. Liao, A. Djari, S. C. Rodriguez, C. Grohs, D. Esquerré, O. Bouchez, M. Rossignol, C. Klopp, D. Rocha, S. Fritz, A. Eggen, P. Bowman, C. D, A. Chamberlain, C. Anderson, C. VanTassell, I. Hulsegge, M. Goddard, B. Guldbrandtsen, M. Lund, R. Veerkamp, D. Boichard, R. Fries, and B. Hayes. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46:858. doi:10.1038/ng.3034

Dale, J. W., M. Von Schantz, and N. Plant. 2012. From genes to genomes: Concepts and applications of DNA technology. John Wiley & Sons, Ltd., Chichester, UK.

De Bustos, A., A. Cuadrado, and N. Jouve. 2016. Sequencing of long stretches of repetitive DNA. Sci. Rep. 6. doi:10.1038/srep36665

de Vos, A. F. 2004. A primer in bayesian inference. http://personal.vu.nl/a.f.de.vos/primer/primer.pdf (Accessed 8 November 2018)

Dekkers, J. C. M. 2004. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. J. Anim. Sci. 82:E313-E328. doi:10.2527/2004.8213_supplE313x

Dekkers, J. C. M. 2012. Application of genomics tools to animal breeding. Curr. Genomics. 13:207-212. doi:https://doi.org/10.2174/138920212800543057

DelCurto, T., T. Murphy, and S. Moreaux. 2017. Demographics and long-term outlook for western us beef, sheep, and horse industries and their importance for the forage industry. Proc. 47th Calif. Alfalfa Grain Symp. Reno, NV. p. 5-16.

DelCurto, T., M. Porath, C. T. Parsons, and J. A. Morrison. 2005. Management strategies for sustainable beef cattle grazing on forested rangelands in the pacific northwest. Rangeland Ecol. Manag. 58:119-127. doi:10.2111/1551-5028(2005)58<119:MSFSBC>2.0.CO;2

Derner, J. D., and V. L. Jin. 2012. Soil carbon dynamics and rangeland management. In: M. A. Liebig, A. J. Franzluebbers and R. F. Follett, editors, Managing agricultural greenhouse gases. Elsevier Inc., London, UK. p. 79-92.

Dias, M., A. Cánovas, C. Mantilla-Rojas, D. Riley, P. Luna-Nevarez, S. Coleman, S. Speidel, R. Enns, A. Islas-Trejo, J. Medrano, S. Moore, M. Fortes, L. Nguyen, B. Venus, I. Diaz, F. Souza, L. Fonseca, F. Baldi, L. Albuquerque, M. Thomas, and H. Oliveira. 2017. Snp detection using rna-sequences of candidate genes associated with puberty in cattle. Genet. Mol. Res. 16. doi:10.4238/gmr16019522.

Diez-Silva, M., M. Dao, J. Han, C.-T. Lim, and S. Suresh. 2010. Shape and biomechanical characteristics of human red blood cells in health and disease. MRS Bull. 35:382-388. doi:https://doi.org/10.1557/mrs2010.571

Drouillard, J. S. 2018. Current situation and future trends for beef production in the united states of america — a review. Asian-australas. J. Anim. Sci. 31:1007-1016. doi:https://doi.org/10.5713/ajas.18.0428

Durham, S. 2010. Line 1 hereford cattle: More than 75 years of research. Agric. Res. 58:14.

Elsik, C. G., R. L. Tellam, and K. C. Worley. 2009. The genome sequence of taurine cattle: A window to ruminant biology and evolution. Science. 324:522-528. doi:10.1126/science.1169588

Fan, B., S. K. Onteru, Z.-Q. Du, D. J. Garrick, K. J. Stalder, and M. F. Rothschild. 2011. Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. PloS One. 6:e14726. doi:https://doi.org/10.1371/journal.pone.0014726

FAS. 2018. Livestock and poultry: World markets and trade. https://downloads.usda.library.cornell.edu/usda-esmis/files/73666448x/jm214p55j/8049g558s/livestock-poultry-ma-04-10-2018.pdf (Accessed 30 September 2018)

Fernando, R., A. Toosi, A. Wolc, D. Garrick, and J. Dekkers. 2017. Application of whole-genome prediction methods for genome-wide association studies: A bayesian approach. J. Agricult. Biol. Env. Stat. 22:172-193. doi:https://doi.org/10.1007/s13253-017-0277-6

Fernando, R. L., and D. Garrick. 2013. Bayesian methods applied to gwas. In: C. Gondro, J. van der Werf and B. Hayes, editors, Genome-wide association studies and genomic prediction. Humana Press, Totowa, NJ. p. 237-274.

Fernando, R. L., and M. Grossman. 1989. Marker assisted selection using best linear unbiased prediction. Genet. Sel. Evol. 21:467. doi:10.1186/1297-9686-21-4-467

Ferri, G., C. H. Huichalaf, R. Caccia, and D. Gabellini. 2015. Direct interplay between two candidate genes in fshd muscular dystrophy. Hum. Mol. Genet. 24:1256-1266. doi:https://doi.org/10.1093/hmg/ddu536

Fong, G.-H. 2008. Mechanisms of adaptive angiogenesis to tissue hypoxia. Angiogenesis. 11:121-140. doi:https://doi.org/10.1007/s10456-008-9107-3

Forin-Wiart, M.-A., P. Hubert, P. Sirguey, and M.-L. Poulle. 2015. Performance and accuracy of lightweight and low-cost gps data loggers according to antenna positions, fix intervals, habitats and animal movements. PLoS One. 10:e0129271. doi:10.1371/journal.pone.0129271

Fortes, M., L. Porto-Neto, K. DeAtley, A. Reverter, M. Thomas, S. Moore, S. Lehnert, and W. Snelling. 2014. Genetic markers in transcription factors of differentially expressed genes associated with post-partum anoestrus predict pregnancy outcome in an independent population of beef cattle. Proc. 10th World Congr. Genet. Appl. Livest. Prod. Vancouver, Canada. p. 250.

Fortes, M. R. S., W. M. Snelling, A. Reverter, S. H. Nagaraj, S. A. Lehnert, R. J. Hawken, K. L. DeAtley, S. O. Peters, G. A. Silver, G. Rincon, J. F. Medrano, A. Islas-Trejo, and M. G. Thomas. 2012. Gene network analyses of first service conception in brangus heifers: Use of genome and trait associations, hypothalamic-transcriptome information, and transcription factors. J. Anim. Sci. 90:2894-2906. doi:10.2527/jas.2011-4601

Frank, D. A., and S. J. McNaughton. 1993. Evidence for the promotion of aboveground grassland production by native large herbivores in yellowstone national park. Oecologia. 96:157-161.

Freer, M., H. Dove, and J. Nolan. 2007. Nutrient requirements of domesticated ruminants. CSIRO publishing, Collingwood VIC 3066, Austraila.

Gannon, M. R., D. W. Bailey, A. C. Ganguli, C. W. Knight, and M. Millward. 2018. Relationship between cattle body size and terrain use in extensive and mountainous rangeland pastures. Proc. 71st Soc. Range Mgmt. Sparks, NV.

Ganskopp, D. 2001. Manipulating cattle distribution with salt and water in large arid-land pastures: A gps/gis assessment. Appl. Anim. Behav. Sci. 73:251-262. doi:10.1016/S0168-1591(01)00148-4

Garcia, M. L., and E. E. Strehler. 1999. Plasma membrane calcium atpases as critical regulators of calcium homeostasis during neuronal cell function. Front. Biosci. 4:D869-D882.

Garrick, D. J., and R. L. Fernando. 2013. Implementing a qtl detection study (gwas) using genomic prediction methodology. In: C. Gondro, J. van der Werf and B. Hayes, editors, Genome-wide association studies and genomic prediction. Humana Press, Totowa, NJ. p. 275-298.

Gatica, D., M. Chiong, S. Lavandero, and D. J. Klionsky. 2015. Molecular mechanisms of autophagy in the cardiovascular system. Circul. Res. 116:456-467. doi:10.1161/CIRCRESAHA.114.303788.

Germain, M., M. Eyries, D. Montani, O. Poirier, B. Girerd, P. Dorfmüller, F. Coulet, S. Nadaud, S. Maugenre, C. Guignabert, W. Carpentier, A. Vonk-Noordegraaf, M. Lévy, A. Chaouat, J.-C. Lambert, M. Bertrand, A.-M. Dupuy, L. Letenneur, M. Lathrop, P. Amouyel, T. J. L. de Ravel, M. Delcroix, E. D. Austin, I. M. Robbins, A. R. Hemnes, J. E. Loyd, E. Berman-Rosenzweig, R. J. Barst, W. K. Chung, G. Simonneau, D. A. Trégouët, M. Humbert, and F. Soubrier. 2013. Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. Nat. Genet. 45:518-521. doi:https://doi.org/10.1038/ng.2581

Ghosh, M., S. S. Sodhi, N. Sharma, R. K. Mongre, N. Kim, A. K. Singh, S. J. Lee, D. C. Kim, S. W. Kim, and H. K. Lee. 2016. An integrated in silico approach for functional and structural impact of non-synonymous snps in the myh1 gene in jeju native pigs. BMC Genet. 17:35. doi:https://doi.org/10.1186/s12863-016-0341-1

Gillen, R., W. Krueger, and R. Miller. 1984. Cattle distribution on mountain rangeland in northeastern oregon. J. Range Manage. 37:549-553.

Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica. 136:245-257. doi:10.1007/s10709-008-9308-0

Goddard, M. E., and B. J. Hayes. 2007. Genomic selection. J. Anim. Breed. Genet. 124:323-330. doi:10.1111/j.1439-0388.2007.00702.x

Golden, B. L., and D. J. Garrick. 2016. An introduction to bolt. http://manual.thetasolutionsllc.com/IntroBolt.pdf (Accessed 25 July 2018)

Govindarajan, R., J. Duraiyan, K. Kaliyappan, and M. Palanisamy. 2012. Microarray and its applications. J. Pharm. Bioallied Sci. 4:S310-S312. doi:10.4103/0975-7406.100283

Greene, C. S., N. M. Penrod, S. M. Williams, and J. H. Moore. 2009. Failure to replicate a genetic association may provide important clues about genetic architecture. PloS One. 4:e5639-e5639. doi:https://doi.org/10.1371/journal.pone.0005639

Grewal, P. K., L. Carim Todd, S. van der Maarel, R. R. Frants, and J. E. Hewitt. 1998. Frg1, a gene in the fsh muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates. Gene. 216:13-19. doi:https://doi.org/10.1016/S0378-1119(98)00334-5

Gross, C. G. 2000. Neurogenesis in the adult brain: Death of a dogma. Nat. Rev. Neurosci. 1:67.

Gunsett, F. 1984. Linear index selection to improve traits defined as ratios. J. Anim. Sci. 59:1185-1193. doi:https://doi.org/10.2527/jas1984.5951185x

Habier, D., R. L. Fernando, K. Kizilkaya, and D. J. Garrick. 2011. Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics 12:186. doi:https://doi.org/10.1186/1471-2105-12-186

Hagenstein, P. R. 1971. Multiple use and environmental decisions on the public lands. J. For. 69:21-24.

Hardie, L. C., M. J. VandeHaar, R. J. Tempelman, K. A. Weigel, L. E. Armentano, G. R. Wiggans, R. F. Veerkamp, Y. de Haas, M. P. Coffey, E. E. Connor, M. D. Hanigan, C. Staples, Z. Wang, J. C. M. Dekkers, and D. M. Spurlock. 2017. The genetic and biological basis of feed efficiency in mid-lactation holstein dairy cows. J. Dairy Sci. 100:9061-9075. doi:https://doi.org/10.3168/jds.2017-12604

Harris, R. W. 1950. Observations on the starkey experimental forest and range. US Department of Agriculture, Pacific Northwest Region, US Forest Service.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. Unsupervised learning, The elements of statistical learning: Data mining, inference, and prediction. Springer, New York. p. 485-585.

Havstad, K., D. Peters, B. Allen-Diaz, J. Bartolome, B. Bestelmeyer, D. Briske, J. Brown, M. Brunson, J. Herrick, L. Huntsinger, P. Johnson, L. Joyce, R. Pieper, T. Svejcar, and J. Yao. 2009. The western united states rangelands: A major resource In: W. F. Wedin and S. L. Fales, editors, Grassland: Quietness and strength for a new american agriculture. ASA-CSSA-SSSA, Madison, WI. p. 75-94.

Hayashi, M., N. Fujita, and A. Yamauchi. 2007. Theory of grazing optimization in which herbivory improves photosynthetic ability. J. Theor. Biol. 248:367-376. doi:10.1016/j.jtbi.2007.05.018

Hayes, B. J., J. Pryce, A. J. Chamberlain, P. J. Bowman, and M. E. Goddard. 2010. Genetic architecture of complex traits and accuracy of genomic prediction: Coat colour, milk-fat percentage, and type in holstein cattle as contrasting model traits. PLoS Genet. 6:e1001139. doi:https://doi.org/10.1371/journal.pgen.1001139

Heady, H., and R. Child. 1994. Rangeland ecology and management. Boulder, co. Westview Press, Inc, Boulder, CO.

Heaton, M. P., W. M. Grosse, S. M. Kappes, J. W. Keele, C. G. Chitko-McKown, L. V. Cundiff, A. Braun, D. P. Little, and W. W. Laegreid. 2001. Estimation of DNA sequence diversity in bovine cytokine genes. Mamm. Genome. 12:32-37. doi:10.1007/s003350010223

Heller, M. J. 2002. DNA microarray technology: Devices, systems, and applications. Annu. Rev. Biomed. Eng. 4:129-153. doi:10.1146/annurev.bioeng.4.020702.153438

Hendrickson, A. W. 2015. A paradigm shift in use and management of united states public lands for livestock grazing. Anim. Front. 5:36-42. doi:10.2527/af.2015-0045

Herbel, C. H., and A. B. Nelson. 1966. Activities of hereford and santa gertrudis cattle on a southern new mexico range. J. Range Manage. 19:173-176.

Herd, R., V. Oddy, and E. Richardson. 2004. Biological basis for variation in residual feed intake in beef cattle. 1. Review of potential mechanisms. Aust. J. Exp. Agric. 44:423-430. doi:https://doi.org/10.1071/EA02220

Hirschhorn, J. N., and M. J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6:95. doi:10.1038/nrg1521

Ho, H., J. J. Jang, S. Kaji, G. Spektor, A. Fong, P. Yang, B. S. Hu, R. Schatzman, T. Quertermous, and J. P. Cooke. 2004. Developmental endothelial locus-1 (del-1), a novel angiogenic protein: Its role in ischemia. Circulation. 109:1314-1319. doi:https://doi.org/10.1161/01.CIR.0000118465.36018.2D

Hodgkinson, K., and H. B. Becking. 1978. Effect of defoliation on root growth of some arid zone perennial plants. Aust. J. Agric. Res. 29:31-42.

Holechek, J., R. D. Pieper, and C. H. Herbel. 2001. Range management: Principles and practices. Prentice Hall, Upper Saddle River, NJ.

Holechek, J. L. 1981. A brief history of range management in the united states. Rangelands. 3:16-18.

Holechek, J. L. 1988. An approach for setting the stocking rate. Rangelands. 10:10-14.

Holt, T. N., and R. J. Callan. 2007. Pulmonary arterial pressure testing for high mountain disease in cattle. Vet. Clin. North Am. Food Anim. Pract. 23:575-596. doi:https://doi.org/10.1016/j.cvfa.2007.08.001

Hong, E. P., and J. W. Park. 2012. Sample size and statistical power calculation in genetic association studies. Genom. Inform. 10:117-122. doi:10.5808/GI.2012.10.2.117

Hoppeler, H., and E. R. Weibel. 2002. Structural and functional limits for oxygen supply to muscle. Acta Physiol. Scand. 168:445-456. doi:https://doi.org/10.1046/j.1365-201x.2000.00696.x

Howery, L. D., F. D. Provenza, R. E. Banner, and C. B. Scott. 1996. Differences in home range and habitat use among individuals in a cattle herd. Appl. Anim. Behav. Sci. 49:305-320. doi:https://doi.org/10.1016/0168-1591(96)01059-3

Howery, L. D., F. D. Provenza, R. E. Banner, and C. B. Scott. 1998. Social and environmental factors influence cattle distribution on rangeland. Appl. Anim. Behav. Sci. 55:231-244.

Howery, L. D., J. E. Sprinkle, and J. E. Bowns. 2000. A summary of livestock grazing systems used on rangelands in the western united states and canada. https://extension.arizona.edu/sites/extension.arizona.edu/files/pubs/az1184-2016_0.pdf (Accessed 10 August 2018)

Hunt, R., Z. Sauna, S. Ambudkar, M. M Gottesman, and C. Kimchi-Sarfaty. 2009. Silent (synonymous) snps: Should we care about them? Methods Mol. Biol. 578:23-39. doi:10.1007/978-1-60327-411-1_2

Huston, J. E., and W. E. Pinchak. 1991. Range animal nutrition. In: R. K. Heitschmidt and J. W. Stuth, editors, Grazing management: An ecological perspective. Timber Press, Portland, OR.

Ibeagha-Awemu, E. M., P. Kgwatalala, A. E. Ibeagha, and X. Zhao. 2008. A critical analysis of disease-associated DNA polymorphisms in the genes of cattle, goat, sheep, and pig. Mamm. Genome. 19:226-245. doi:10.1007/s00335-008-9101-5

Ingram, D. C. 1930. Ranges are made usable by hauling water for livestock. USDA Yearbook of Agriculture:446-449.

Iso-Touru, T., G. Sahana, B. Guldbrandtsen, M. S. Lund, and J. Vilkki. 2016. Genome-wide association analysis of milk yield traits in nordic red cattle using imputed whole genome sequence variants. BMC Genet. 17:55. doi:10.1186/s12863-016-0363-8

Jackson, W. T. 1956. British interests in the range cattle industry. In: M. W. Frink, W. T. Jackson and A. W. Spring, editors, When grass was king. University of Colorado Press, Boulder, CO.

Jensen, J., G. Su, and P. Madsen. 2012. Partitioning additive genetic variance into genomic and remaining polygenic components for complex traits in dairy cattle. BMC Genet. 13:44. doi:https://doi.org/10.1186/1471-2156-13-44

Johnson, D. D., and D. C. Ganskopp. 2008. Gps collar sampling frequency: Effects on measures of resource use. Rangeland Ecol. Manag. 61:226-231. doi:10.2111/07-044.1

Johnson, R. C., G. W. Nelson, J. L. Troyer, J. A. Lautenberger, B. D. Kessing, C. A. Winkler, and S. J. O'Brien. 2010. Accounting for multiple comparisons in a genome-wide association study (gwas). BMC Genomics. 11:724-724. doi:10.1186/1471-2164-11-724

Kachman, S. D. 2008. Incorporation of marker scores into national genetic evaluations. 9th Genet. Predict. Workshop. Kansas City, MO. p. 92-98.

Kauffman, J. B., and W. C. Krueger. 1984. Livestock impacts on riparian ecosystems and streamside management implications... A review. J. Range Manage. 37:430-438.

Khor, C. C., S. Davila, C. Shimizu, S. Sheng, T. Matsubara, Y. Suzuki, J. W. Newburger, A. Baker, D. Burgner, W. Breunis, T. Kuijpers, V. J. Wright, M. Levin, M. L. Hibberd, and J. C. Burns. 2011. Genome-wide linkage and association mapping identify susceptibility alleles in *abcc4* for kawasaki disease. J. Med. Genet. 48:467. doi:10.1136/jmg.2010.086611

Klug, W. S., S. M. Ward, M. R. Cummings, C. A. Spencer, and M. A. Palladino. 2013. Essentials of genetics 8ed. Pearson Learning Solutions, Boston, MA.

Knight, C., D. Bailey, and D. Faulkner. 2018a. Low-cost global positioning system tracking collars for use on cattle. Rangeland Ecol. Manag. 71:506-508. doi:https://doi.org/10.1016/j.rama.2018.04.003

Knight, C. W. 2016. Intake, reproductive, and grazing activity characteristics of range cattle on semi-arid rangelands. PhD Diss. University of Arizona, Tucson.

Knight, C. W., D. W. Bailey, K. Grubb, W. Mandeville, and J. Wilker. 2018b. Data correction and analysis of spatial data using lower-cost global positioning system tracking collars. Proc. 71st Soc. Range Mgmt. Sparks, NV.

Kolli, V., R. C. Upadhyay, and D. Singh. 2014. Peripheral blood leukocytes transcriptomic signature highlights the altered metabolic pathways by heat stress in zebu cattle. Res. Vet. Sci. 96:102-110. doi:https://doi.org/10.1016/j.rvsc.2013.11.019

Kopan, R., and M. X. G. Ilagan. 2009. The canonical notch signaling pathway: Unfolding the activation mechanism. Cell. 137:216-233. doi:10.1016/j.cell.2009.03.045

Koringa, P. G., S. J. Jakhesara, V. D. Bhatt, A. B. Patel, D. Dash, and C. G. Joshi. 2013. Transcriptome analysis and snp identification in scc of horn in (bos indicus) indian cattle. Gene. 530:119-126. doi:https://doi.org/10.1016/j.gene.2013.07.061

Koufariotis, L., Y.-P. P. Chen, S. Bolormaa, and B. J. Hayes. 2014. Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle. BMC Genomics. 15:436. doi:10.1186/1471-2164-15-436

Krehbiel, B. 2017. Evaluation of population genetic structure in two british bos taurus breeds across five u.S climate zones. Master's Thesis. Colorado State University, Fort Collins.

Krock, B. L., N. Skuli, and M. C. Simon. 2011. Hypoxia-induced angiogenesis: Good and evil. Genes Cancer. 2:1117-1133. doi:https://doi.org/10.1177/1947601911423654

Kumar, S., T. W. Banks, and S. Cloutier. 2012. Snp discovery through next-generation sequencing and its applications. Int. J. Plant Genomics. 2012:187-212. doi:10.1155/2012/831460

Kwok, P.-Y., and X. Chen. 2003. Detection of single nucleotide polymorphisms. Curr. Issues Mol. Biol. 5:43-60.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656-4663. doi:10.3168/jds.2009-2061

Legarra, A., P. Croiseau, M. P. Sanchez, S. Teyssèdre, G. Sallé, S. Allais, S. Fritz, C. R. Moreno, A. Ricard, and J.-M. Elsen. 2015. A comparison of methods for whole-genome qtl mapping using dense markers in four livestock species. Genet. Sel. Evol. 47:6. doi:10.1186/s12711-015-0087-7

Lehne, B., C. M. Lewis, and T. Schlitt. 2011. From snps to genes: Disease association at the gene level. PLoS One. 6:e20133. doi:10.1371/journal.pone.0020133

Leriche, H., X. LeRoux, J. Gignoux, A. Tuzet, H. Fritz, L. Abbadie, and M. Loreau. 2001. Which functional processes control the short-term effect of grazing on net primary production in grasslands? Oecologia. 129:114-124.

Levy, D., G. B. Ehret, K. Rice, G. C. Verwoert, L. J. Launer, A. Dehghan, N. L. Glazer, A. C. Morrison, A. D. Johnson, T. Aspelund, Y. Aulchenko, T. Lumley, A. Köttgen, R. S. Vasan, F. Rivadeneira, G. Eiriksdottir, X. Guo, D. E. Arking, G. F. Mitchell, F. U. S. Mattace-Raso, A. V. Smith, K. Taylor, R. B. Scharpf, S.-J. Hwang, E. J. G. Sijbrands, J. Bis, T. B. Harris, S. K. Ganesh, C. J. O'Donnell, A. Hofman, J. I. Rotter, J. Coresh, E. J. Benjamin, A. G. Uitterlinden, G. Heiss, C. S. Fox, J. C. M. Witteman, E. Boerwinkle, T. J. Wang, V. Gudnason, M. G. Larson, A. Chakravarti, B. M. Psaty, and C. M. van Duijn. 2009. Genome-wide association study of blood pressure and hypertension. Nat. Genet. 41:677. doi:https://doi.org/10.1038/ng.384

Li, J., J. Hu, Y. Liu, X. Wang, S. Tang, X. Chen, M. Niu, N. Waili, Y. Bai, and Y. Wei. 2018. Effects of prenatal hypoxia on fetal sheep heart development and proteomics analysis. Int. J. Clin. Exp. Pathol. 11:1909-1922.

Lindholm, H. B., and H. H. Stonaker. 1957. Economic importance of traits and selection indexes for beef cattle. J. Anim. Sci. 16:998-1006. doi:https://doi.org/10.2527/jas1957.164998x

Liu, W., Y. Zhang, C.-M. Yu, Q.-W. Ji, M. Cai, Y.-X. Zhao, and Y.-J. Zhou. 2015. Current understanding of coronary artery calcification. J. Geriatr. Cardiol. 12:668-675. doi:https://doi.org/10.11909/j.issn.1671-5411.2015.06.012

Lu, K., Z. Xiao, H. Jian, L. Peng, C. Qu, M. Fu, B. He, L. Tie, Y. Liang, X. Xu, and J. Li. 2016. A combination of genome-wide association and transcriptome analysis reveals candidate genes controlling harvest index-related traits in brassica napus. Sci. Rep. 6:36452. doi:10.1038/srep36452

Lunt, S. T. 2013. Modifying individual grazing distribution patterns of cows in extensive rangeland pastures through genetic selection. Master's thesis. New Mexico State University, Las Cruces.

Mackay, T. F. C. 2001. The genetic architecture of quantitative traits. Annu. Rev. Genet. 35:303-339. doi:https://doi.org/10.1146/annurev.genet.35.102401.090633

MacNeil, M. D. 2009. Invited review: Research contributions from seventy-five years of breeding line 1 hereford cattle at miles city, montana. J. Anim. Sci. 87:2489-2501. doi:10.2527/jas.2009-1909

Marees, A. T., H. Kluiver, S. Stringer, F. Vorspan, E. Curis, C. Marie-Claire, and E. M. Derks. 2018. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int. J. Methods Psychiatr. Res. 27:e1608. doi:https://doi.org/10.1002/mpr.1608

Matukumalli, L., S. Schroeder, S. DeNise, T. Sonstegard, C. Lawley, M. Georges, W. Coppieters, K. Gietzen, J. Medrano, and G. Rincon. 2011. Analyzing ld blocks and cnv segments in cattle: Novel genomic features identified using the bovinehd beadchip. Illumina Inc., San Diego, CA.

Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'Connell, S. S. Moore, T. P. L. Smith, T. S. Sonstegard, and C. P. Van Tassell. 2009. Development and characterization of a high density snp genotyping assay for cattle. PLoS One. 4:e5350. doi:10.1371/journal.pone.0005350

McGranahan, D. A., J. W. Spiess, and B. A. Geaumont. 2018. Our low-cost, open-source livestock gps collars work well, provide insight into spatial data collection. Proc. 71st Soc. Range Mgmt. Sparks, NV.

McLaren, W., L. Gil, S. E. Hunt, H. S. Riat, G. R. Ritchie, A. Thormann, P. Flicek, and F. Cunningham. 2016. The ensembl variant effect predictor. Genom. Bio. 17:122. doi:10.1186/s13059-016-0974-4

McNaughton, S. 1979a. Grazing as an optimization process: Grass-ungulate relationships in the serengeti. Am. Nat. 113:691-703.

McNaughton, S. J. 1979b. Grazing as an optimization process: Grass-ungulate relationships in the serengeti. Am. Nat. 113:691-703.

McNaughton, S. J. 1984. Grazing lawns: Animals in herds, plant form, and coevolution. Am. Nat. 124:863-886.

Medrano, J. F. 2017. The new bovine reference assembly and its value for genomic research. Proc. 22nd Assoc. Advmt. Anim. Breed. Genet. Rydges Southbank Townsville, Queensland. p. 161-166.

Mei, B., and Z. Wang. 2016. An efficient method to handle the 'large p, small n'problem for genomewide association studies using haseman–elston regression. J. Genet. 95:847-852. doi:10.1007/s12041-016-0705-3

Melo, T., B. G. Neto, M. Fortes, L. Albuquerque, and R. Carvalheiro. 2017. Agreement among gwas results from different statistical methods as a strategy to increase the power of qtl detection. Proc. 22nd Assoc. Advmt. Anim. Breed. Genet. Rydges Southbank Townsville, Queensland. p. 425-428.

Mercado, T. J., D. W. Bailey, M. G. Thomas, R. M. Enns, and S. E. Speidel. 2018. Repeatability of terrain use by cattle in rugged rangeland pastures. Proc. 71st Soc. Range Mgmt. Sparks, NV.

Meuwissen, T., B. Hayes, and M. Goddard. 2016. Genomic selection: A paradigm shift in animal breeding. Anim. Front. 6:6-14. doi:10.2527/af.2016-0002

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 157:1819-1829.

Miller, S. 2010. Genetic improvement of beef cattle through opportunities in genomics. Rev. Bras. Zoot. 39:247-255. doi:10.1590/S1516-35982010001300027

Moser, G., S. H. Lee, B. J. Hayes, M. E. Goddard, N. R. Wray, and P. M. Visscher. 2015. Simultaneous discovery, estimation and prediction analysis of complex traits using a bayesian mixture model. PLoS Genet. 11:e1004969. doi:10.1371/journal.pgen.1004969

Moser, G., B. Tier, R. E. Crump, M. S. Khatkar, and H. Wf. Raadsma. 2009. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide snp markers. Genet. Sel. Evol. 41:56-56. doi:https://doi.org/10.1186/1297-9686-41-56

Mottet, A., C. de Haan, A. Falcucci, G. Tempio, C. Opio, and P. Gerber. 2017. Livestock: On our plates or eating at our table? A new analysis of the feed/food debate. Glob. Food. Sec. 14:1-8. doi:https://doi.org/10.1016/j.gfs.2017.01.001

Mueggler, W. F. 1965. Cattle distribution on steep slopes. J. Range Manage. 5:255-257.

NASS. 2016. Overview of the united states cattle industry. http://usda.mannlib.cornell.edu/usda/current/USCatSup/USCatSup-06-24-2016.pdf (Accessed 8 November 2018)

NASS. 2017. Cattle. http://usda.mannlib.cornell.edu/usda/nass/Catt//2010s/2017/Catt-07-21-2017.pdf (Accessed 25 September 2018)

NCBI. 2018a. Ars-ucd1.2. https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1/ (Accessed 30 June 2018)

NCBI. 2018b. Bos_taurus_umd_3.1.1. https://www.ncbi.nlm.nih.gov/assembly/GCF_000003055.6/ (Accessed 30 June 2018)

Newman, J. H., T. N. Holt, J. D. Cogan, B. Womack, J. A. Phillips III, C. Li, Z. Kendall, K. R. Stenmark, M. G. Thomas, and R. D. Brown. 2015. Increased prevalence of epas1 variant in cattle with high-altitude pulmonary hypertension. Nat. Commun. 6:6863. doi:https://doi.org/10.1038/ncomms7863.

Ogorevc, J., T. Kunej, A. Razpet, and P. Dovc. 2009. Database of cattle candidate genes and genetic markers for milk production and mastitis. Anim. Genet. 40:832-851. doi:https://doi.org/10.1111/j.1365-2052.2009.01921.x

Onteru, S. K., B. Fan, Z.-Q. Du, D. J. Garrick, K. J. Stalder, and M. F. Rothschild. 2012. A whole-genome association study for pig reproductive traits. Anim. Genet. 43:18-26. doi:doi:10.1111/j.1365-2052.2011.02213.x

Paige, K. N., and T. G. Whitham. 1987. Overcompensation in response to mammalian herbivory: The advantage of being eaten. Amer. Nat. 129:407-416.

Papinski, D., M. Schuschnig, W. Reiter, L. Wilhelm, C. A. Barnes, A. Maiolica, I. Hansmann, T. Pfaffenwimmer, M. Kijanska, I. Stoffel, S. S. Lee, A. Brezovich, J. H. Lou, B. E. Turk, R. Aebersold, G. Ammerer, M. Peter, and C. Kraft. 2014. Early steps in autophagy depend on direct phosphorylation of atg9 by the atg1 kinase. Mol. Cell. 53:471-483. doi:10.1016/j.molcel.2013.12.011

Pareek, C. S., P. Błaszczyk, P. Dziuba, U. Czarnik, L. Fraser, P. Sobiech, M. Pierzchała, Y. Feng, H. N. Kadarmideen, and D. Kumar. 2017. Single nucleotide polymorphism discovery in bovine liver using rna-seq technology. PLoS One. 12:e0172687. doi:10.1371/journal.pone.0172687

Park, K.-D., J. Park, J. Ko, B. C. Kim, H.-S. Kim, K. Ahn, K.-T. Do, H. Choi, H.-M. Kim, and S. Song. 2012. Whole transcriptome analyses of six thoroughbred horses before and after exercise using rna-seq. BMC Genomics. 13:473. doi:https://doi.org/10.1186/1471-2164-13-473

Parsons, C. T., P. A. Momont, T. Delcurto, M. McInnis, and M. L. Porath. 2003. Cattle distribution patterns and vegetation use in mountain riparian areas. J. Range Manage. 56:334-341.

Partipilo, G., P. D'Addabbo, G. M. Lacalandra, G. E. Liu, and M. Rocchi. 2011. Refinement of bos taurus sequence assembly based on bac-fish experiments. BMC Genomics. 12:639-639. doi:10.1186/1471-2164-12-639

Pauling, R. C., S. E. Speidel, M. G. Thomas, T. N. Holt, and R. M. Enns. 2018. Evaluation of moderate to high elevation effects on pulmonary arterial pressure measures in angus cattle. J. Anim. Sci. 96:3599-3605. doi:https://doi.org/10.1093/jas/sky262

Peacock, A. J. 1998. Abc of oxygen: Oxygen at high altitude. BMJ. 317:1063-1066.

Pechanec, J. F., and G. Stewart. 1949. Grazing springfall sheep ranges of southern idaho. U.S. Dept. Agric.

Peters, S., K. Kizilkaya, D. Garrick, R. Fernando, J. Reecy, R. Weaber, G. Silver, and M. Thomas. 2012. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in brangus heifers. J. Anim. Sci. 90:3398-3409. doi:10.2527/jas.2012-4507

Peters, S., K. Kizilkaya, D. Garrick, R. Fernando, J. Reecy, R. Weaber, G. Silver, and M. Thomas. 2013. Heritability and bayesian genome-wide association study of first service conception and pregnancy in brangus heifers. J. Anim. Sci. 91:605-612. doi:https://doi.org/10.2527/jas.2012-5580

Piskol, R., G. Ramaswami, and Jin B. Li. 2013. Reliable identification of genomic variants from rna-seq data. Am. J. Hum. Genet. 93:641-651. doi:10.1016/j.ajhg.2013.08.008

Purcell, S., and C. Chang. 2015. Plink 1.9. https://www.cog-genomics.org/plink2 (Accessed 15 January 2018)

Recio, M. R., R. Mathieu, P. Denys, P. Sirguey, and P. J. Seddon. 2011. Lightweight gps-tags, one giant leap for wildlife tracking? An assessment approach. PLoS One. 6:e28225. doi:10.1371/journal.pone.0028225

Richardson, I. W., D. P. Berry, H. L. Wiencko, I. M. Higgins, S. J. More, J. McClure, D. J. Lynn, and D. G. Bradley. 2016. A genome-wide association study for genetic susceptibility to

mycobacterium bovis infection in dairy cattle identifies a susceptibility qtl on chromosome 23. Genet. Sel. Evol. 48:19. doi:10.1186/s12711-016-0197-x

Richter, S., J. P. Richter, S. Y. Mehta, A. M. Gribble, A. J. Sutherland-Smith, K. M. Stowell, R. S. Ronimus, and W. R. Wilson. 2012. Expression and role in glycolysis of human adp-dependent glucokinase. Mol. Cell. Biochem. 364:131-145. doi:https://doi.org/10.1007/s11010-011-1212-8

Roath, L. R., and W. C. Krueger. 1982. Cattle grazing and behavior on a forested range. J. Range Manage. 35:332-338.

Ronimus, R. S., and H. W. Morgan. 2004. Cloning and biochemical characterization of a novel mouse adp-dependent glucokinase. Biochem. Biophys. Res. Commun. 315:652-658.

Saatchi, M., and D. J. Garrick. 2013. Accuracies of genomic prediction in beef cattle. In: Animal industry report. Rep. No. Asl r2771. Iowa state univ., ames. P. 22.

Saatchi, M., R. D. Schnabel, J. F. Taylor, and D. J. Garrick. 2014. Large-effect pleiotropic or closely linked qtl segregate within and across ten us cattle breeds. BMC Genomics. 15:442. doi:https://doi.org/10.1186/1471-2164-15-442

Sayre, N., and M. Fernandez-Gimenez. 2003. The genesis of range science, with implications for current development policies. Proc. Int. Rangelands Congr. Durban, South Africa. p. 1976-1985.

Sayre, N. F., W. deBuys, B. T. Bestelmeyer, and K. M. Havstad. 2012. "The range problem" after a century of rangeland science: New research themes for altered landscapes. Rangeland Ecol. Manag. 65:545-552. doi:10.2111/REM-D-11-00113.1

Senft, R. L., M. B. Coughenour, D. W. Bailey, L. R. Rittenhouse, O. E. Sala, and D. M. Swift. 1987. Large herbivore foraging and ecological hierarchies. Bioscience. 37:789-799. doi:10.2307/1310545

Shen, L. X., J. P. Basilion, and V. P. Stanton. 1999. Single-nucleotide polymorphisms can cause different structural folds of mrna. Proc. Natl. Acad. Sci. U.S.A. p. 7871-7876.

Silanikove, N. 2000. Effects of heat stress on the welfare of extensively managed domestic ruminants. Livest. Prod. Sci. 67:1-18. doi:https://doi.org/10.1016/S0301-6226(00)00162-7

Skovlin, J. M. 1957. Range riding-the key to range management. J. Range Manage. 10:269-271.

Sorensen, A. A., J. Freedgood, J. Dempsey, and D. M. Theobald. 2018. Farms under threat: The state of america's farmland. https://www.farmlandinfo.org/sites/default/files/AFT_Farms_Under_Threat_May2018%20maps%20B_0.pdf (Accessed 8 November 2018)

Spangler, M. 2011. Implementation of marker assisted epds. Proc. Range Beef Cow Symp. XXII. Mitchell, NE. p. 297.

Spangler, M. L. 2012. Incorporating DNA information into epds for angus cattle and potential for other breeds. 65th Annu. Recip. Meat Conf. Fargo, ND. p. 1-4.

Speidel, S. E., B. A. Buckley, R. J. Boldt, R. M. Enns, J. Lee, M. L. Spangler, and M. G. Thomas. 2018. Genome-wide association study of stayability and heifer pregnancy in red angus cattle. J. Anim. Sci. 96:846-853. doi:https://doi.org/10.1093/jas/sky041

Stephens, M., and D. J. Balding. 2009. Bayesian statistical methods for genetic association studies. Nat. Rev. Genet. 10:681-690. doi:10.1038/nrg2615.

Stephenson, M. 2015. Influencing livestock grazing distribution. https://grassland.unl.edu/documents/Stephenson%20edited.pdf (Accessed 8 July 2018)

Stoddart, L. A., and A. D. Smith. 1943. Range management, 547. McGraw-Hill Book Co. Inc, New York.

Stothard, P., J.-W. Choi, U. Basu, J. M. Sumner-Thomson, Y. Meng, X. Liao, and S. S. Moore. 2011. Whole genome resequencing of black angus and holstein cattle for snp and cnv discovery. BMC Genomics. 12:559. doi:10.1186/1471-2164-12-559

Suárez-Vega, A., B. Gutiérrez-Gil, J. Benavides, V. Perez, G. Tosser-Klopp, C. Klopp, S. J. Keennel, and J. J. Arranz. 2015. Combining gwas and rna-seq approaches for detection of the causal mutation for hereditary junctional epidermolysis bullosa in sheep. PLoS One. 10:e0126416. doi:10.1371/journal.pone.0126416

Sun, C.-Y. J., S. van Koningsbruggen, S. W. Long, K. Straasheijm, R. Klooster, T. I. Jones, M. Bellini, L. Levesque, W. M. Brieher, S. M. van der Maarel, and P. L. Jones. 2011a. Facioscapulohumeral muscular dystrophy region gene 1 is a dynamic rna-associated and

actin-bundling protein. J. Mol. Biol. 411:397-416. doi:https://doi.org/10.1016/j.jmb.2011.06.014

Sun, X., D. Habier, R. L. Fernando, D. J. Garrick, and J. C. Dekkers. 2011b. Genomic breeding value prediction and qtl mapping of qtlmas2010 data using bayesian methods. BMC Proc. 5:S13-S13. doi:10.1186/1753-6561-5-S3-S13

Swain, D. L., M. A. Friend, G. J. Bishop-Hurley, R. N. Handcock, and T. Wark. 2011. Tracking livestock using global positioning systems – are we still lost? Anim. Prod. Sci. 51:167-175. doi:https://doi.org/10.1071/AN10255

Tanaka, J. A., N. R. Rimbey, L. A. Torell, D. T. Taylor, D. Bailey, T. DelCurto, K. Walburger, and B. Welling. 2007. Grazing distribution: The quest for the silver bullet. Rangelands. 29:38-46. doi:https://doi.org/10.2458/azu_rangelands_v29i4_tanaka

Thomas, B., J. D. Holland, and E. O. Minot. 2012. Wildlife tracking technology options and cost considerations. Wildl. Res. 38:653-663. doi:10.1071/WR10211

Tolleson, D., and P. Meiman. 2015. Global effects of changing land-use on animal agriculture. Anim. Front. 5:14-23. doi:10.2527/af.2015-0042

Turner, L. W., M. C. Udal, B. T. Larson, and S. A. Shearer. 2000. Monitoring cattle behavior and pasture use with gps and gis. Can. J. Anim. Sci. 80:405-413. doi:10.4141/A99-093

Twyman, R. M. 2005. Single nucleotide polymorphism (snp) genotyping techniques—an overview. In: J. Fuchs and M. Podda, editors, Encyclopedia of diagnostic genomics and proteomics. Marcel Dekker, Inc., New York, New York. p. 1202-1207.

USDA Forest Service. 1944. The history of western range research. Agric. Hist. 18:127-143.

Utsunomiya, A. T. H., D. J. A. Santos, S. A. Boison, Y. T. Utsunomiya, M. Milanesi, D. M. Bickhart, P. Ajmone-Marsan, J. Sölkner, J. F. Garcia, R. da Fonseca, and M. V. G. B. da Silva. 2016. Revealing misassembled segments in the bovine reference genome by high resolution linkage disequilibrium scan. BMC Genomics. 17:705. doi:10.1186/s12864-016-3049-8

Vail, S. 1992. Selection for overcompensatory plant responses to herbivory: A mechanism for the evolution of plant-herbivore mutualism. Am. Nat. 139:1-8. doi:10.1086/285309

Valentine, K. A. 1947. Distance from water as a factor in grazing capacity of rangeland. J. For. 45:749-754.

Vallentine, J. F. 1990. Grazing management. Academic Press, Inc., San Diego, CA.

van den Berg, I., S. Fritz, and D. Boichard. 2013. Qtl fine mapping with bayes c(π): A simulation study. Genet. Sel. Evol. 45:19. doi:10.1186/1297-9686-45-19

van der Werf, J. 2013. Genomic selection in animal breeding programs. In: C. Gondro, J. van der Werf and B. Hayes, editors, Genome-wide association studies and genomic prediction. Humana Press, Totowa, NJ. p. 543-561.

Van Eenennaam, A. L., J. Li, R. M. Thallman, R. L. Quaas, M. E. Dikeman, C. A. Gill, D. E. Franke, and M. G. Thomas. 2007. Validation of commercial DNA tests for quantitative beef quality traits. J. Anim. Sci. 85:891-900. doi:10.2527/jas.2006-512

Van Eenennaam, A. L., K. A. Weigel, A. E. Young, M. A. Cleveland, and J. C. Dekkers. 2014. Applied animal genomics: Results from the field. Annu. Rev. Anim. Biosci. 2:105-139. doi:10.1146/annurev-animal-022513-114119

Verma, P., A. Sharma, M. Sodhi, K. Thakur, R. S. Kataria, S. K. Niranjan, V. K. Bharti, P. Kumar, A. Giri, S. Kalia, and M. Mukesh. 2018. Transcriptome analysis of circulating pbmcs to understand mechanism of high altitude adaptation in native cattle of ladakh region. Sci. Rep. 8:7681. doi:https://doi.org/10.1038/s41598-018-25736-7

Vignal, A., D. Milan, M. SanCristobal, and A. Eggen. 2002. A review on snp and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol. 34:275-305. doi:10.1186/1297-9686-34-3-275

Viluma, A. 2017. Applying massive parallel sequencing technologies to unravel the genomic complexity in domestic species. PhD Diss. Swedish University of Agricultural Sciences, Uppsala.

Vogel, F., and A. G. Motulsky. 1987. Human genetics: Problems and approaches. Springer-Verlag, New York.

Volesky, J., W. H Schacht, and S. Waller. 1996. Proper livestock grazing distribution on rangeland.

http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=2313&context=extensionhist (Accessed 8 November 2018)

Walker, J. W., R. K. Heitschmidt, E. A. De Moraes, M. M. Kothmann, and S. L. Dowhower. 1989. Quality and botanical composition of cattle diets under rotational and continuous grazing treatments. J. Range Manage. 42:239-242.

Wang, G.-S., and T. A. Cooper. 2007. Splicing in disease: Disruption of the splicing code and the decoding machinery. Nat. Rev. Genet. 8:749. doi:10.1038/nrg2164

Wang, L., Y. Zhang, M. Zhao, R. Wang, R. Su, and J. Li. 2015. Snp discovery from transcriptome of cashmere goat skin. Asian-australas. J. Anim. Sci. 28:1235-1243. doi:10.5713/ajas.15.0172

Williams, J., S. Dunner, A. Valentini, R. Mazza, V. Amarger, M. Checa, A. Crisa, N. Razzaq, D. Delourme, and F. Grandjean. 2009. Discovery, characterization and validation of single nucleotide polymorphisms within 206 bovine genes that may be considered as candidate genes for beef production and quality. Anim. Genet. 40:486-491. doi:10.1111/j.1365-2052.2009.01874.x

Williams, R. E. 1954. Modern methods of getting uniform use of ranges. J. Range Manage. 7:77-81.

Wojczynski, M. K., M. Li, L. F. Bielak, K. F. Kerr, A. P. Reiner, N. D. Wong, L. R. Yanek, L. Qu, C. C. White, L. A. Lange, J. F. Ferguson, J. He, T. Young, T. H. Mosley, J. A. Smith, B. G. Kral, X. Guo, Q. Wong, S. K. Ganesh, S. R. Heckbert, M. E. Griswold, D. H. O'Leary, M. Budoff, J. J. Carr, H. A. Taylor, D. A. Bluemke, S. Demissie, S.-J. Hwang, D. N. Paltoo, J. F. Polak, B. M. Psaty, D. M. Becker, M. A. Province, W. S. Post, C. J. O'Donnell, J. G. Wilson, T. B. Harris, M. Kavousi, L. A. Cupples, J. I. Rotter, M. Fornage, L. C. Becker, P. A. Peyser, I. B. Borecki, and M. P. Reilly. 2013. Genetics of coronary artery calcification among african americans, a meta-analysis. BMC Med. Genomics. 14:75. doi:https://doi.org/10.1186/1471-2350-14-75

Wolc, A., J. Arango, P. Settar, J. E. Fulton, N. P. O'Sullivan, R. Preisinger, D. Habier, R. Fernando, D. J. Garrick, W. G. Hill, and J. C. M. Dekkers. 2012. Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens. Anim. Genet. 43:87-96. doi:https://doi.org/10.1111/j.1365-2052.2012.02381.x

Wu, L., T. Sun, K. Kobayashi, P. Gao, and J. D. Griffin. 2002. Identification of a family of mastermind-like transcriptional coactivators for mammalian notch receptors. Mol. Cell. Biol. 22:7688-7700.

Yazdani, A. 2014. Statistical approaches in genome-wide association studies. PhD Diss. Universit`a degli Studi di Padova, Padua.

Yazdani, A., and D. B. Dunson. 2015. A hybrid bayesian approach for genome-wide association studies on related individuals. Bioinformatics. 31:3890-3896. doi:10.1093/bioinformatics/btv496

Yi, N., V. George, and D. B. Allison. 2003. Stochastic search variable selection for identifying multiple quantitative trait loci. Genetics. 164:1129-1138.

Yousefi, S., T. Abbassi-Daloii, T. Kraaijenbrink, M. Vermaat, H. Mei, P. van't Hof, M. van Iterson, D. V. Zhernakova, A. Claringbould, and L. Franke. 2018. A snp panel for identification of DNA and rna specimens. BMC Genomics. 19:90. doi:10.1186/s12864-018-4482-7

Zeng, X. 2016. Angus cattle at high altitude: Pulmonary arterial pressure, estimated breeding value and genome-wide association study. PhD Diss. Colorado State University, Fort Collins.

Zerbino, D. R., P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, and C. G. Girón. 2017. Ensembl 2018. Nucleic Acids Res. 46:D754-D761. doi:https://doi.org/10.1093/nar/gkx1098.

Zhang, C., G. Wang, J. Wang, Z. Ji, Z. Liu, X. Pi, and C. Chen. 2013. Characterization and comparative analyses of muscle transcriptomes in dorper and small-tailed han sheep using rna-seq technique. PLoS One. 8:e72686. doi:https://doi.org/10.1371/journal.pone.0072686

Zhang, L., A. D. J. van Dijk, and K. Hettinga. 2017. An interactomics overview of the human and bovine milk proteome over lactation. Proteome Sci. 15:1. doi:https://doi.org/10.1186/s12953-016-0110-0

Zhou, S., S. Goldstein, M. Place, M. Bechner, D. Patino, K. Potamousis, P. Ravindran, L. Pape, G. Rincon, and J. Hernandez-Ortiz. 2015. A clone-free, single molecule map of the domestic cow (bos taurus) genome. BMC Genomics. 16:644. doi:10.1186/s12864-015-1823-7

APPENDIX I: QUANTITATIVE TRAIT LOCI AND ASSOCIATED TRAITS

Table A.1. Quantitative trait loci (QTL) associated with terrain-use in the western United States and beef cattle traits that were previously associated with the QTL as reported by the Cattle QTL Database[1].

RS ID[2]	CHR[3]	Position[4]	Trait[5]
rs43256975	1	108154057	milk palmitoleic acid content
rs43109323	4	61702691	GL, FP, parasites mean of natural logarithm, SC, WW, YH, YW
rs43408732	4	79996650	-
rs109175805	5	101518663	milk capric acid content, milk caproic acid content, milk caprylic acid content, milk decenoic acid content, milk myristic acid content, milk myristoleic acid content, milk oleic acid content, milk palmitoleic acid content
rs110681394	7	86485159	BW, COLDT, DMI
rs41619378	8	55927704	adhesion, BW, CE, CW, FA, HT, LM, MS, milk butyric acid content, milk lauroleic acid content, muscle calcium content, SB, SS, TWIN
rs109226946	8	86442366	angularity, BW, CW, milk capric acid content, milk myristic acid content, MY
rs110514275	10	15202658	BD, BW, WW, CE, GIT weight, HT, LM, milk alpha-casein percentage, FY, milk protein percentage, PY, muscle nitrogen content, muscle pH, social separation (vocalization), strength, SUBFAT, TL, UA
rs109097567	10	19494672	BD, BW, WW, YW, CE, CW, long-chain fatty acid content, LM, medium-chain fatty acid content, milk alpha-casein percentage, FY, PY, muscle nitrogen content, muscle pH, myristic acid content, myristoleic acid content, palmitoleic acid content, social separation (vocalization), strength, SUBFAT, TL, TWIN, UA
rs29013509	10	20498087	ADG, BD, BW, WW, YW, CE, CW, LM, milk alpha-casein percentage, FY, PY, muscle nitrogen content, muscle pH, social separation (vocalization), strength, SUBFAT, TL, TWIN, UA
rs29013631	10	38826667	YH, MEATP, milk alpha-casein percentage, FY, PY, muscle nitrogen content, muscle pH, non-return rate, SF, SCC, SUBFAT, TL, tick resistance, UA
rs42415241	10	38986604	FY, SF, muscle pH, muscle nitrogen content, non-return rate, SCC, UA, FY, tick resistance, SUBFAT, MEATP, milk alpha-casein percentage, YH
rs109502510	11	91680069	milk beta-lactoglobulin protein content, stearic acid content

rs109669554	11	93142448	milk beta-lactoglobulin protein content, milk butyric acid content, milk caproic acid content, RFI, stearic acid content
rs41633961	12	70064114	muscle iron content
rs41576569	13	63369536	DMI, milk capric acid content, milk caproic acid content, milk caprylic acid content, milk myristoleic acid content, milk palmitoleic acid content, SC, TL
rs110030253	13	63391193	DMI, milk capric acid content, milk caproic acid content, milk caprylic acid content, milk myristoleic acid content, milk palmitoleic acid content, SC, TL
rs41811366	16	45309651	ADG, WW, bone percentage, CW, fat thickness at the 12th rib, juiciness, SC, social separation (vocalization)
rs41598300	16	80407423	-
rs109952637	17	56310258	milk myristoleic acid content, milk palmitoleic acid content, SF, trans-16-C18:1 fatty acid content
rs41848746	17	61920415	milk myristoleic acid content, milk palmitoleic acid content, SF, trans-16-C18:1 fatty acid content
rs41582500	18	57115739	birth index, calf size, CE, palmitic acid content, retail product yield, SB
rs42599235	21	47902442	abomasum displacement, CW, gastrointestinal nematode burden, LM, SCC
rs110978254	22	55257755	bovine tuberculosis susceptibility, MSPD, MY
rs41600226	23	47219946	infectious bovine keratoconjunctivitis susceptibility, milk palmitoleic acid content
rs29021957	24	6276349	BD, body form composite, CE, CW, LM, oleic acid content, strength
rs109716600	24	18893792	angularity, BD, BFCI, BW, CE, CW, fat thickness at the 12th rib, feet and leg conformation, immunoglobulin G level, interval to first estrus after calving, LM, MS, FP (EBV), PCVM, oleic acid content, percentage decrease in PCV up to day 150 after challenge, strength
rs110959252	24	42665926	feed conversion ratio, gastrointestinal nematode burden, MS, milk myristic acid content, oleic acid content, UA, UCI, udder depth
rs42120868	27	17956728	BW, MW, WW, CW, dairy form, dystocia, fat thickness at the 12th rib, HT, MS, FP, PY, MY, non-return rate, PCV variance, PCVI minus PCVM, percentage decrease in PCV up to day 100 after challenge, percentage decrease in PCV up to day 150 after challenge, SF
rs110590993	29	11905442	BW, MW, WW, BWF scaled by BWI, CW, fat thickness at the 12th rib, FA, GL, HT, CONCEPT, interval to first estrus after calving, LM, margaric acid content, PY, MSPD, myristic acid content, paired testes volume, paired testes weight, parasites natural logarithm of mean number, PTA type, RFI, rump angle, SS, TPL, temperament, TEND, UA, UCI

| rs43703968 | 29 | 13959142 | BWF scaled by BWI, MW, WW, CW, fat thickness at the 12th rib, FA, GL, HT, LM, margaric acid content, PY, MY, paired testes volume, paired testes weight, parasites natural logarithm of mean number, PTA type, RFI, rump angle, SS, TPL, TEND, UA, UCI, UH, UW |
| rs42477618 | 29 | 14009294 | BWF scaled by BWI, MW, WW, CW, fat thickness at the 12th rib, FA, GL, HT, margaric acid content, LM, PY, MY, paired testes volume, paired testes weight, parasites natural logarithm of mean number, PTA type, RFI, rump angle, SS, TPL, TEND, UA, UCI, UH, UW |

[1]https://www.animalgenome.org/cgi-bin/QTLdb/BT/index.

[2]Reference SNP cluster identification assigned by the National Center for Biotechnology Information (NCBI).

[3]Chromosome in which the SNP was located according to *Bos taurus* UMD3.1.1.

[4]Position (bp) in which the SNP was located according to *Bos taurus* UMD3.1.1.

[5]ADG: average daily gain, BD: body depth, BFCI: body form composite index, BW: birth weight, BWF: final body weight, BWI: initial body weight, CE: calving ease, COLDT: cold tolerance, CONCEPT: inseminations per conception, CW: carcass weight, DMI: dry matter intake, EBV: estimated breeding value, FA: foot angle, FP: milk fat percentage, FY: milk fat yield, GIT: gastrointestinal tract, GL: gestation length, HT: mature height, LM: Longissimus muscle area, MEATP: meat percentage, MS: marbling score, MSPD: milking speed, MW: mature weight, MY: milk yield, PCV: packed red blood cell volume, PCVI: initial packed red blood cell volume, PCVM: minimum packed red blood cell volume, PTA: predicted transmitting ability, PY: milk protein yield, RFI: residual feed intake, SB: stillbirth, SC: scrotal circumference, SCC: somatic cell count, SF: shear force, SS: structural soundness, SUBFAT: subcutaneous fat, TEND: tenderness score, TL: teat length, TPL: teat placement, TWIN: twinning, UA: udder attachment, UCI: udder composite index, UH: udder height, UW: udder width, WW: weaning weight, YH: yearling height, YW: yearling weight.

Table A.2. Twenty-nine genes associated with grazing distribution traits of beef cows managed in the western U.S. and their gene ontology according to AgBase[1].

Gene Symbol	Gene Name	Biological Process	Molecular Function	Cellular Component
A2ML1	Alpha-2-macroglobulin like 1	regulation of endopeptidase activity	peptidase inhibitor activity	extracellular space
ADPGK	ADP Dependent Glucokinase	glucose metabolic process	ADP-specific glucokinase activity	endoplasmic reticulum
ATP2B2	ATPase plasma membrane Ca2+ transporting 2	calcium ion transmembrane transport	calcium-transporting ATPase activity	integral component of membrane
C1H3orf80	Chromosome 1 C3orf80 homolog	-	-	integral component of membrane
CA6	Carbonic Anhydrase 6	one-carbon metabolic process	carbonate dehydratase activity	extracellular space
CBLN2	Cerebellin 2 Precursor	positive regulation of synapse assembly	-	extracellular space
CDK5RAP1	CDK5 Regulatory Subunit Associated Protein 1	tRNA modification	transferase activity	cytoplasm
CELF4	CUGBP Elav-Like Family Member 4	regulation of alternative mRNA splicing, via spliceosome	RNA binding	nucleus
EDIL3	EGF Like Repeats And Discoidin Domains 3	positive regulation of cell-substrate adhesion	integrin binding	extracellular vesicle
EEF1E1	Eukaryotic translation elongation factor 1 epsilon 1	translational elongation	translation elongation factor activity	nucleus

ENSBTAG00000000655	ENSBTAG00000000655	-	-	-
ENSBTAG00000022427	ENSBTAG00000022427	-	-	-
ENSBTAG00000027868	ENSBTAG00000027868	-	-	-
ENSBTAG00000032603	ENSBTAG00000032603	-	-	-
ENSBTAG00000039201	ENSBTAG00000039201	-	-	-
EPB42	Erythrocyte Membrane Protein Band 4.2	peptide cross-linking	protein-glutamine gamma-glutamyltransferase activity	-
FRG1	FSHD region gene 1	-	actin filament binding	nucleolus
IFT81	Intraflagellar Transport 81	tubulin binding	intraciliary transport	intraciliary transport particle B
INHBA	Inhibin Subunit Beta A	G1/S transition of mitotic cell cycle	signaling receptor binding	inhibin A complex
ITGA11	Integrin subunit alpha 11	substrate-dependent cell migration	collagen binding	focal adhesion
LRRC4B	Leucine Rich Repeat Containing 4B	positive regulation of synapse assembly	signaling receptor binding	cerebellar mossy fiber
MRRF	Mitochondrial Ribosome Recycling Factor	translation	ribosomal large subunit binding	mitochondrion
NR5A2	Nuclear receptor subfamily 5 group A member 2	regulation of transcription, DNA-templated	transcriptional activator activity, RNA polymerase II proximal promoter sequence-specific DNA binding	nucleus
PHF2	PHD finger protein 2	protein demethylation	transcription coactivator activity	nucleolus
PIEZO2	Piezo type mechanosensitive ion channel component 2	-	mechanosensitive ion channel activity	integral component of membrane
SEPT7	Septin 7	-	GTP binding	septin complex
TBC1D21	TBC1 domain family member 21	intracellular protein transport	GTPase activator activity	intracellular

TBX3	T-Box 3	regulation of transcription, DNA-templated	DNA-binding transcription factor activity	nucleus
TLE4	Transducin Like Enhancer of Split 4 2	regulation of transcription, DNA-templated	-	nucleus

[1]http://agbase.arizona.edu/index.html