USE OF ULTRA-HIGH-DENSITY SNP DATA FOR GENETIC EVALUATION OF PREDICTED FEED INTAKE, FEED EFFICIENCY, CARCASS, AND GROWTH TRAITS IN HEREFORD CATTLE

By

KIMBERLY BRANHAM

Bachelor of Science in Animal Science

Oklahoma State University

Stillwater, Oklahoma

2013

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 2015

USE OF ULTRA-HIGH-DENSITY SNP DATA FOR GENETIC EVALUATION OF PREDICTED FEED INTAKE, FEED EFFICIENCY, CARCASS, AND GROWTH TRAITS IN HEREFORD CATTLE

Thesis Approved:

Megan Rolf

Thesis Adviser

Michelle Calvo-Lorenzo

Udaya DeSilva

ACKNOWLEDGEMENTS

I would like to first thank my Lord and Savior Jesus Christ for the opportunities He has blessed me with. I would not be where I am today without His guiding hand in my life. Next, I would like to thank my dad, my sisters, and my extended family. They have been there for me through all of the highs and lows in graduate school and I would be lost without their support and constant encouragement. I also want to give a special thanks to my dad, Keith, for helping me fund my undergraduate and graduate degrees so that I could focus on following my dreams without having to worry about finances. In addition, I would like to thank my friends who have helped me get through school whether it by being a shoulder to cry on or someone to laugh with. I am blessed to have met some great friends along this road- Andrea English, Samantha Lowman, Emily Andreini, Corbit Bayliff, Kristi Allwardt, and Jarrod Cole. A big thanks to my best friends Madeline Nehls and Daniela Hiatt who have not been close geographically, but they have always been a phone call away whenever I needed to vent or get encouragement. There is also a multitude of faculty in the department that have helped me succeed. I want to thank my advisor, Dr. Megan Rolf, for pushing me and helping me become a more independent researcher. I also want to thank Drs. Michelle Calvo-Lorenzo and Udaya DeSilva who not only served on my committee, but they were also a source of encouragement and assistance that was much appreciated. I would also like to thank Justin Buchanan who was critical in helping and teaching me in various aspects of my project. In addition, I would like to thank Dr. Jerry Fitch who has helped me succeed in school since I was an undergraduate. He has allowed me to both laugh and cry in his office, and I was blessed to have him on my side all these years. Furthermore, I would like to thank Dr. Bob Kropp for encouraging me to step out of my comfort zone and challenge myself in a difficult field within animal science. Lastly, I would not be where I am today if it had not been for my sweet angel mother, Cynthia. Although she passed away from colon cancer in 2012, she has taught me so much even after she has passed. I have learned that no matter what is going on in life or with school, I am blessed and I have so much to be thankful for. She was the greatest support system and encourager, and even though I miss her every day, I know that she is proud as she watches me fulfill my dreams.

Name: KIMBERLY BRANHAM

Date of Degree: JULY, 2015

Title of Study: USE OF ULTRA-HIGH-DENSITY SNP DATA FOR GENETIC EVALUATION OF PREDICTED FEED INTAKE, FEED EFFICIENCY, CARCASS, AND GROWTH TRAITS IN HEREFORD CATTLE

Major Field: ANIMAL SCIENCE

Abstract: Keeping input costs low and producing the best quality product are two important goals of beef cattle producers. One way to reach these goals is to select for feed efficient animals to lower input costs and to select for the animals with the highest carcass quality traits to increase revenue. This study used ultra-high-density single nucleotide polymorphism (SNP) data to evaluate genetic merit for these traits in beef cattle. To find a more economical method for producers to select for feed efficiency independent of average daily gain (ADG), this study analyzed whether predicted dry matter intake (pDMI) phenotypes could be utilized as an indicator trait for efficiency. A genome-wide association study (GWAS) was performed to identify genomic regions that are important to predicting genetic merit in feed efficiency traits. A subset of the largest effect SNPs for each trait were compared to large effect SNP regions of other phenotypes to identify regions of the genome that impact feed efficiency but are independent of gain. The largest effect SNPs were analyzed to identify genes and biological pathways that could be directly linked to these regions. Carcass traits were analyzed using the same procedure. The direct genomic value (DGV) accuracies for ADG, dry matter intake (DMI), pDMI, and residual feed intake (RFI) ranged from 0.27 to 0.51. There were seven QTL regions in common between pDMI and DMI that were independent of ADG QTL regions. A gene clustering tool, Partial Correlation coefficient with Information Theory (PCIT), identified various genes in these QTL regions that are linked to obesity and weight loss in mice. In addition, the accuracies for carcass phenotypes varied from 0.47 to 0.60.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

LITERATURE REVIEW

INTRODUCTION

In the beef cattle industry, one of the main goals of producers is to yield the greatest amount of product possible while controlling costs. One way this goal is achieved is by selecting for animals that are feed efficient. Since feed costs are a large part of the expenses of raising cattle (Archer et al., 1999; Anderson et al., 2005), decreasing monetary inputs while maximizing outputs is critical to increasing profitability and sustainability of beef cattle operations.

Utilizing genetics to help select for more efficient animals is a feasible and economical method for producers to employ. A study by Jenkins and Ferrell (1994) compared different breeds of cattle and showed there is evidence for a genetic component of feed efficiency. To develop genetic selection tools for feed efficiency, feed intake data on large populations of animals must be obtained. Since feed intake data is difficult and expensive to measure, finding the genomic regions responsible for feed efficiency may provide an opportunity to utilize genomic selection methodologies to improve feed efficiency on a large scale while minimizing phenotyping costs for producers. One major challenge to this approach is assembling vast repositories of feed efficiency phenotypes, which could be alleviated by identifying suitable indicator traits for actual feed intake data.

FEED EFFICIENCY- PAST, PRESENT, AND FUTURE

In the beef cattle industry, an important goal for producers is to yield the greatest amount

of product possible while minimizing input costs. One way this goal is achieved is selecting for animals that are feed efficient, because feed costs are a large part of the expenses of raising cattle (Archer et al., 1999; Anderson et al., 2005). Decreasing monetary inputs while maximizing product outputs is critical to increasing profitability and sustainability of beef cattle operations. In the past, selection decisions in the beef cattle industry have been based largely on increasing outputs, such as weaning weight or rib eye area, in order to increase income. The other option was to reduce input costs. Since feed costs are a majority of the input costs for producers, being able to decrease the intake of animals without decreasing production levels would be a practical way to improve profits. The most effective way to do this is to focus on both input traits and output traits (like feed intake and average daily gain (ADG)) simultaneously to increase the efficiency of production.

Utilizing genetics to help select for more efficient animals is a feasible and economical method for producers to employ. To develop genetic selection tools for feed efficiency, feed intake data on large populations of animals must be obtained. Because feed intake data is difficult and expensive to measure, identifying the genomic regions responsible for feed efficiency may provide a more economical method for producers to utilize in order to decrease their feed input costs.

Background of feed efficiency

Selection for feed efficiency involves identifying animals that eat less feed than other animals that have similar or greater levels of productivity. Feed efficient animals need less feed for growth and maintenance, which reduces feed costs. Since feed can account for up to 70% of total input costs (Moore et al., 2006), even small improvements in feed efficiency can dramatically impact profitability.

One factor that is responsible for differences in feed efficiency in beef cattle is the breed of the animal (Archer et al., 1999). Frisch and Vercoe (1984) showed that Hereford x Shorthorn

bulls needed 20% more feed than Brahman bulls to be able to maintain their body condition. This variation can be explained by differences in biological mechanisms, such as digestion efficiency, microbial populations in the rumen, activity level, and heat dissipation (Veerkamp et al., 2013).

Traditional methods of selecting for feed efficiency

Feed efficiency is a difficult trait to measure and evaluate for a variety of reasons, including expense and level of technology necessary to measure exact feed intake values for individual animals. The process of measuring feed intakes is even more challenging when cattle are in a grazing environment because dry matter intake values are even more difficult to accurately measure in such an environment. Since measuring traits for feed efficiency can be costly and difficult, some possible solutions are to identify phenotypes that can be used as indicator traits for feed efficiency genetic evaluation and to utilize genomics to maximize the return on phenotyping of animals.

Some researchers define feed efficiency as the ratio of feed inputs to production outputs. Feed efficiency is not something that can be measured directly, but it can be inferred from measuring feed intake and weight gain. Dry matter intake (DMI) is a component trait of efficiency that represents system inputs. Average daily gain (ADG), or the amount of daily gain achieved by an animal during a specific feeding period, is another important part of the feed efficiency equation.

Many traditional measures of feed efficiency are based on ADG. For example, feed:gain ratio (F:G), also known as feed conversion ratio (FCR), is calculated as the ratio of daily DMI to body weight. While calculating FCR for feed efficiency is useful, it is only a gross measurement; therefore, FCR does not specify between feed consumed for growth and maintenance requirements (Carstens and Tedeschi, 2006). A trait very similar to this one, gain:feed ratio (G:F), has also been used and is calculated as the ratio of weight gain to feed consumption (Koch et al., 1963). When FCR and G:F is used for selection decisions, emphasis will be placed on ADG,

which can lead to larger animals (Archer et al., 1999). This increase in mature size can lead to greater production costs in the cowherd due to higher feed requirements within the system. Another ratio trait used to calculate feed efficiency is referred to as maintenance efficiency. It is the ratio of feed consumed that is used for maintenance per unit of metabolic body weight ($BW^{0.75}$) (Carstens and Tedeschi, 2006). Maintenance efficiency is calculated as the actual intake of an animal minus the predicted intake that an animal requires for growth per unit of metabolic body weight (Carstens and Tedeschi, 2006). In addition to this trait, partial efficiency of growth (PEG) is another ratio trait that is used to predict feed efficiency. This method involves calculating average daily gain (ADG) per unit of feed partitioned for growth requirements (Carstens and Tedeschi, 2006). Using PEG may be a valuable alternative to FCR because it is more strongly correlated with feed intake than FCR is. Because selecting for feed efficiency utilizing ratios based on growth can lead to an increase in mature cow size and, therefore, an increase in the amount of feed required for maintenance and growth, it is important to identify selection tools that allow producers to select for genetic variation in feed intake that is not linked to gain through pleiotropy.

Koch et al. (1963) was the first to suggest calculating feed efficiency as residual feed intake (RFI). The RFI is expressed as the deviation from the animal's predicted feed intake and is calculated as follows:

$$RFI = ADMI - eDMI$$

$$eDMI = b_0 + b_1ADG + b_2MMWT^{0.75}$$

The calculation for RFI involves subtracting the expected dry matter intake (eDMI) from the observed dry matter intake (ADMI). The expected feed intake values are calculated from a regression of feed intake on ADG and metabolic mid-weight (MMWT). Residual feed intake is a ratio trait that is useful because the calculation used to derive the value forces it to be independent of average daily gain. Although since it is a ratio trait, an animal that eats less feed and gains less can have the same RFI as an animal that eats more and gains more. An animal with a negative

RFI is considered more efficient because they consume less feed than was predicted. RFI uses an animal's weight and growth rate to account for maintenance and growth requirements separately (Carstens and Tedeschi, 2006). Accounting for these requirements separately allows RFI to be a trait that is phenotypically independent of growth rate (Kennedy et al., 1993).

There are conflicting views between scientists on the use of calculating RFI to predict feed efficiency. Arthur et al. (2001a) suggests that it is one of the most effective ways to express feed efficiency because it is a better alternative to ratio traits and it is strongly correlated with feed conversion ratio. Because of the correlation between RFI and its component traits, there is likely to be no new genetic information gained from focusing on RFI that is not gained from selection on the component traits alone. A trait similar to RFI is residual body weight gain (RG) which regresses ADG on feed intake and body weight to identify animals that do not require to be on feed for as long while still having a feed intake value that is less than other animals (Berry and Crowley, 2012). An animal with a higher RG grows faster than other animals but it is not associated with feed intake values (Berry and Crowley, 2012). The most important difference between RFI and RG is that RG is not phenotypically independent of ADG since ADG can be an important trait to consider when selecting for feed efficiency. The same selection responses in feed efficiency are likely to be seen by focusing on the component traits of feed intake, gain, and body weight which is why the use of a selection index may be the best approach for selection of feed efficiency. Although there are many options used to estimate feed efficiency, there is currently no universally-accepted method.

Feed efficiency prediction models

In an attempt to alleviate the need to collect individual-animal feed intake data, simulation models have been created that can estimate the predicted DMI of cattle. One such model is the Cornell/Cattle Value Discovery System (CVDS), which allocates feed fed to groups of animals to individual animals based on growth, body weight, and carcass measurements (Tedeschi and Fox, 2006; Tedeschi et al., 2006). Fox et al. (2004) demonstrated that CVDS

under-predicted intake by only 2%. Another study reported that DMI of steers was under-predicted by 0.91% while heifers were over-predicted by 0.89% (Fox et al., 2001). Another similar model, the Decision Evaluator for the Cattle Industry (DECI), is able to predict DMI (pDMI) required for an animal to reach a certain performance level (Williams et al., 2006). Williams et al. (2006) compared DECI and CVDS to test the accuracy of these two models in predicting individual DMI. The heritability of DMI for CVDS and DECI were almost equal at 0.32 and 0.33, respectively. When observed daily DMI values were compared to model-predicted values, CVDS under-predicted DMI by 3.4% on average while DECI over-predicted DMI by 0.4% on average. In addition, the phenotypic correlations for observed DMI with predicted DMI using CVDS and DECI were 0.785 and 0.798, respectively. The genetic correlations for observed DMI with predicted DMI using CVDS and DECI were 0.95 ± 0.07 and 0.96 ± 0.07, respectively (Williams et al., 2006). These results indicate that both models could be useful to predict DMI for genetic evaluation of feed efficiency. Ration, gain information, and carcass information is routinely collected in the feedlot sector, and model-predicted intakes could provide a methodology to incorporate these pen data into individual-animal evaluation systems.

Heritability and genetic and phenotypic correlations between feed efficiency phenotypes

Indicator traits that are most likely to be effective for selecting for feed efficiency are traits that have at least a moderate heritability. Genetic variation exists for DMI, RFI, and ADG, which is necessary for developing selection tools for feed efficiency, and estimates of heritability for these traits are moderate and indicate that genetic improvement is possible (Table 1). Another important aspect of the genetics of indicator traits when selecting for feed efficiency is the genetic correlations of the traits. The presence of genetic correlations indicates that pleiotropy exists and may cause correlated responses to selection. Phenotypic correlations for RFI and DMI range from 0.58 to 0.72, and genotypic correlations range from 0.59 to 0.79 (Arthur et al., 2001a; Crowley et al., 2010; Herd and Bishop, 2000; Arthur et al., 2001b). These correlations are moderate to high

because the calculation of RFI uses DMI. Phenotypic correlations for RFI and ADG range from -0.06 to 0.09, and the genotypic correlations range from -0.10 to 0.01 (Arthur et al., 2001a; Crowley et al., 2010; Herd and Bishop, 2000; Arthur et al., 2001b) signifying that these two traits are independent of each other. The correlation between RFI and ADG is expected to be zero because RFI is phenotypically independent of gain due to its inclusion in the regression model to estimate expected intake. The phenotypic correlations for ADG and DMI range from 0.38 to 0.47, and the genotypic correlations range from 0.39 to 0.55 (Arthur et al., 2001a; Crowley et al., 2010; Arthur et al., 2001b) which is a low to moderate positive correlation signifying that as DMI increases, ADG often increases as well. The phenotypic correlation between observed DMI and predicted DMI was found to be 0.785 and the genetic correlation was 0.95 ± 0.07 (Williams et al., 2006), as previously mentioned. Because the genetic correlation is very high, predicted DMI is a good candidate for an indicator trait for feed intake.

Utilizing genomic information to better select for feed efficiency

The use of direct genomic values (DGV) are especially useful for traits that are traditionally difficult to measure. These traits may be difficult to measure because of the expense, such as feed intake, or because of the practicality of obtaining the measurement, such as disease resistance. The DGVs are calculated as the sum of the individual marker effects identified in QTL studies (Weber et al., 2012). In order for a DGV to be calculated, a training population which possesses dense phenotype and genotype data is used to generate genomic prediction equations. In research studies, these predictions are then evaluated in a validation population that has also been phenotyped and genotyped to determine the accuracy of the predictions. The DGVs can also be utilized by breed associations to generate genomic-enhanced EPDs for producers to evaluate the genetic potential of an animal in their breeding program. When DGV are utilized within the industry, they are typically trained utilizing all of the available information/animals, so the un-phenotyped population where DGV are utilized is referred to as an implementation population.

There are many QTL regions identified in the literature for RFI, ADG, and DMI. In some cases, the genomic regions overlap (Rolf et al., 2012; Barendse et al., 2007; Nkrumah et al., 2007b). However, there is often low concordance of QTL regions between studies (Saatchi et al., 2014). Some possible reasons for this are that different cattle populations experience different genotype x environment interactions (Sherman et al., 2010), QTL may be specific to a particular breed or population, or an association may be spurious. Saatchi et al. (2014) suggested that another reason why QTL studies tend to be population-specific is due to differences between breeds, including differences in linkage disequilibrium between breeds. Before extrapolating the results of one QTL study across different cattle breeds, and even within breeds, the QTL need to be validated in independent populations.

QTL studies which have been completed have been critical in identifying important regions of the genome related to traits of interest. These studies have increased the efficiency of SNP technology and have advanced the understanding of the genes responsible for the traits of interest. Although identifying QTL regions is important to understand the biological underpinnings of a trait, genomic selection processes are transforming to examine all the variation in the entire genome simultaneously.

CARCASS QUALITY TRAITS

Selecting for carcass traits in beef cattle

Superior carcass merit contributes to consumer satisfaction with the beef-eating experience, and is an important suite of traits for genetic improvement. Some of these critical carcass traits are rib eye area (REA), hot carcass weight (HCW), back fat thickness (BFAT), yield grade (YG), and marbling score (MARB). Ultrasound measurement of carcass traits contributes to genetic evaluation on young animals and those that will not have their own phenotype available for carcass traits, such as bulls and replacement females. However, these measures can vary from carcass data that is obtained at an abattoir because the ultrasound data is typically

collected before the animal is finished (Drake, 2004). The comparison of the ultrasound data and the data collected at slaughter can vary between traits and between sexes, but overall the correlations between the two data types were moderate to high indicating ultrasound scanning can provide important information for genetic improvement (Reverter et al., 2000).

Yield grade is estimated based on expected retail yield of a carcass (Drake, 2004). The YG is calculated based on BFAT, the kidney, pelvic, and heart fat percentage, HCW, and REA. The BFAT trait is important because there should be an adequate layer of fat to prevent the carcass from drying out, but there should not be too much fat so as to take away from the amount of meat that can be obtained from each carcass (Drake, 2004). Marbling score is a measure of intramuscular fat and is important in determining the quality grade of carcasses.
It is important for beef producers to be able to produce a consistent, quality product that fits consumer demands and all of these traits interact and contribute towards the eating experience a consumer encounters.

Breed Effects for Carcass Traits

Breed effects are an important consideration in the study of carcass traits (Retallick et al., 2013). Some general relationships between carcass traits in different breeds can guide producers in breed selection decisions, most notably when utilizing breed complementarity in crossbreeding systems. Angus and Hereford cattle have the greatest BFAT and Angus have the greatest marbling score when compared to Pinzgauer and Brahman (Crouse et al., 1989). According to the 2006 Germplasm Evaluation Program Report No. 23, Angus and Hereford had higher marbling scores than Brangus and Beefmaster, but Beefmaster had a higher BFAT than Hereford. This is a trend that has remained fairly constant over the years with similar results reported in the 1974 Germplasm Evaluation Progress Report No. 1 (United States Department of Agriculture-Agricultural Research Service), which shows that Hereford and Angus are above the average of all breeds tested for both BFAT and marbling score. In contrast, Charolais, Brahman, and

Limousin cattle have been reported to have the lowest BFAT and marbling scores (Marshall, 1994). Angus cattle have been found to have less desirable YGs compared to other breeds, possibly due to high fat levels in the carcasses (Retallick et al, 2013). This statement is confirmed by the Germplasm Evaluation Progress Report No. 1 and No. 23, which lists Angus as having the highest YG of any other breed listed in almost every comparison. The breeds with the highest HCW are Simmental, who also had the highest REA, and Charolais, who also have the highest growth rate of any other breed (Marshall, 1994; Schenkel et al., 2004; Gregory et al., 1995). British breeds tend to have greater marbling scores and BFAT than Continental breeds (Johnson et al., 1988). British and Continental breeds also tend to have a heavier HCW than Brahman-type breeds with Continental breeds being the heaviest (Johnson et al., 1988). One large issue that producers of Brahman-type cattle breeds face is decreased tenderness and lower marbling scores. While many producers who focus on carcass traits may avoid these breeds all together, having an animal with a portion of Brahman in their genetic background is important for other critical traits such as their ability to cope with high temperatures (Crouse et al., 1989; Marshall, 1994).

In addition to breed differences, there are also sex differences in carcass traits between heifers, steers and bulls. For example, heifers generally have a lower HCW than bulls or steers, because heifers are typically smaller than steers and bulls (Hassen et al., 1999). Generally, bulls are leaner than steers (Johnson et al., 1988; Hassen et al., 1999). Steers have a higher back fat thickness and marbling score than bulls, but bulls have a higher REA than both steers and heifers. (Johnson et al., 1988; Hassen et al., 1999). It is important to keep these differences in mind when doing genetic analyses on carcass traits and comparing the results between sexes.

Differences in carcass quality among animals due to genetics

Genetic and phenotypic correlations between carcass traits are important tools that allow identification of unfavorable responses in performance due to the selection of correlated traits. Genetic correlations between marbling score (MARB) and REA range from -0.02 to 0.44 with phenotypic correlations from -0.05 to 0.12 (Riley et al., 2002; Gregory et al., 1995). These correlations tend to be zero or negative, which indicates these traits are often not dependent on each other or as one increases the other decreases. Genetic correlations for REA and HCW varied from 0.52 to 0.66 with phenotypic correlations from 0.40 to 0.44 (Riley et al., 2002; Gregory et al., 1995). These correlations indicate a favorable relationship between REA and HCW which could enhance selection for both traits. For YG and MARB, genetic correlations range from 0.31 to 0.45 with phenotypic correlations from 0.26 to 0.40 (Riley et al., 2002; Gregory et al., 1995). These moderate correlations are what would be expected because the more fat on a carcass, the higher the YG. In addition, MARB and HCW had genetic correlations ranging from 0.31 to 0.39 with phenotypic correlations ranging from 0.13 to 0.17 (Riley et al., 2002; Gregory et al., 1995). The genetic correlations indicate that as HCW increases, MARB tends to increase concomitantly. The genetic correlations for BFAT and REA ranged from -0.14 to 0.02 and phenotypic correlations from -0.06 to 0.13 (Robinson and Oddy, 2004; Riley et al., 2002; Gregory et al., 1995), which indicates these traits have an unfavorable relationship. Furthermore, the genetic correlation for YG and HCW ranged from 0.25 to 0.56 with a phenotypic correlation ranging from 0.30 to 0.48 (Wheeler et al., 2005; Riley et al., 2002). The same study found the genetic correlation of YG and REA to be -0.26 and a phenotypic correlation of -0.30 (Riley et al., 2002). This negative correlation is what would be expected since YG is calculated by subtracting the REA, so a larger REA would result in a lower YG. The genetic correlations for BFAT with YG and BFAT with HCW were 0.93 and 0.60, respectively, and the phenotypic correlations were 0.81 and 0.43, respectively (Riley et al., 2002). The correlations for BFAT and YG are very high because BFAT is utilized in calculation of YG. The higher the amount of backfat, the higher the YG will be. In addition, the high correlation between BFAT and HCW suggests that an animal

with more fat will have a heavier carcass. Lastly, the genetic correlation for MARB and BFAT range from 0.44 to 0.56 and phenotypic correlations range from 0.25 to 0.30 (Gregory et al., 1995; Riley et al., 2002). Heritabilities for carcass quality traits are generally high. Literature estimates for a variety of carcass traits are summarized in Table 2. The trait with the highest heritability is MARB (Mao et al., 2013), while the trait with the lowest heritability estimate is BFAT (Mao et al., 2013).

Genetic Selection for Carcass Quality

Throughout the genome, single base pair changes, also known as single nucleotide polymorphisms (SNP), are responsible for a portion of the variation that is observed in phenotypes. A genome-wide association study (GWAS) examines the effect each of these SNPs has on the overall phenotype and can identify quantitative trait loci (QTL) regions that presumably harbor causal mutations. Currently, as many as 770,000 SNPs can be tested simultaneously utilizing bead arrays, which provides dense coverage (approximately 1 SNP per 3900 base pairs) in the genome for fine-scale QTL mapping (Pryce et al., 2012; Pryce et al., 2014). While carcass trait QTL regions are very useful for understanding the biological processes behind carcass performance and validating SNP associations utilizing biological information, genomic selection for carcass traits will be more useful than single-marker approaches for most traits. Identifying QTL regions and development of genomic selection models can be accomplished concomitantly, which enables genomic selection approaches while simultaneously providing information on biological processes and pathways that contribute to variation in phenotype. The use of genomics also allows for the calculation of a direct genomic value (DGV) which gives producers an idea of how an animal will perform for a specific trait. The accuracy of a DGV calculates the correlation between the actual and predicted genetic ability of an animal (Saatchi et al., 2011). For example, the DGV accuracies for HCW, BFAT, MARB, and REA have been estimated at 0.471, 0.603, 0.690, 0.601, respectively (Saatchi et al., 2011).

Genome-wide association studies allow for the identification of QTL regions linked to carcass traits by looking at the predicted SNP effects. A study by Yuan et al. (2013) found that diacylglycerol O-acyltransferase (DGAT1), which is a critical gene in the formation of triglyceride synthesis, could be a candidate gene for fat-related carcass traits in beef cattle, such as MARB and BFAT. In addition, another study found that SNPs within the gene family MYOD, which is important for muscle development, had a significant association with HCW (Bhuiyan et al., 2009). A study by Marques et al. (2009) found two candidate genes that have associations with BFAT and MARB. The two genes, 2,4 dienoyl CoA reductase 1 (DECR1) and core binding factor (CBFA2T1), have been previously associated with lipid metabolism in other species. The number of candidate genes will continue to increase as additional studies are performed and validated in independent populations.

Table 1. Literature estimates of heritability (± SE) of dry matter intake (DMI), residual feed intake (RFI), and average daily gain (ADG)

	DMI	RFI	ADG	Breed*
Archer et al. (2002)	0.28	0.23	0.33	AN, HP, & SP
Arthur et al. (2001a)	0.39 ± 0.03	0.38 ± 0.06	0.28 ± 0.04	AN
Arthur et al. (2001b)	0.48 ± 0.06	0.43 ± 0.06	0.41 ± 0.06	CH
Barendse et al. (2007)	----	0.26 ± 0.07	----	AN, BR, BE, HP, MG, SG, & SP
Bolormaa et al. (2011)	0.16 ± 0.13	0.18 ± 0.13	0.24 ± 0.14	AN, MG, SP, & HP
Bouquet et al. (2010)	0.48	0.45	----	LM
Crowley et al. (2010)	0.49 ± 0.06	0.45 ± 0.06	0.30 ± 0.06	AN, CH, HP, LM, & SM
Herd and Bishop (2000)	0.31 ± 0.08	0.16 ± 0.08	0.38 ± 0.10	HE
Koch et al. (1963)	0.64 ± 0.12	0.28 ± 0.11	----	HP, AN, SP
Lu et al. (2013)	0.28 ± 0.04	0.22 ± 0.05	0.29 ± 0.06	AN, CH, PI, & XB
MacNeil et al. (1991)	----	----	0.38 ± 0.16	AN, HP, SM, PZ, & RP
Mao et al. (2013)	0.54 ± 0.13 0.39 ± 0.10	0.68 ± 0.14 0.47 ± 0.12	0.54 ± 0.13 0.38 ± 0.12	CH AN
Mujibi et al. (2011)	0.41 ± 0.12	0.29 ± 0.12	0.28 ± 0.11	AN, CH, HP, & XB
Nkrumah et al. (2007b)	0.54 ± 0.15	0.21 ± 0.12	0.59 ± 0.17	AN, CH, HP, & XB
Robinson & Oddy (2004)	0.27 ± 0.06	0.18 ± 0.06	----	BR, BE, SG, AN, HP, MG, & SP
Rolf et al. (2010)	0.16	0.21	0.00	AN
Rolfe et al. (2011)	0.40	0.52	0.26	HP, AN, SM, CH, LM, GV
Saatchi et al. (2014)	0.35 0.35 0.41 0.27	0.49 0.21 0.45 0.32	0.30 0.19 0.27 0.23	XB AN HP SM & AN
Schenkel et al. (2004)	0.44 ± 0.06	0.38 ± 0.07	0.35 ± 0.03	CH, LM, SM, HP, AN
Williams et al. (2006)	0.27 ± 0.12	----	0.35 ± 0.13	HP, AN, & SP

*Breed abbreviations: AN- Angus; CH- Charolais; HP- Hereford; LM- Limousin; SM- Simmental; SP- Shorthorn; RP- Red Poll; PZ- Pinzgauer; BR- Brahman; SG- Santa Gertrudis; BE- Belmont Red; MG- Murray Grey; PI- Piedmontese; GV- Gelbvieh; XB- Beef crossbreeds

Table 2. Literature estimates of heritability (± SE) of rib eye area (REA), marbling score (MARB), back fat thickness (BFAT), calculated yield grade (CYG), and hot carcass weight (HCW)

	REA	MARB	BFAT	CYG	HCW	Breed*
Benyshek (1981)	0.40 ± 0.04	0.47 ± 0.04	0.52 ± 0.04	----	0.48 ± 0.04	HP
Crews and Kemp (2001)	0.54 ± 0.19	0.55 ± 0.19	0.46 ± 0.18	----	0.38 ± 0.16	XB- (AN, CH, SM, LM, HP & SP)
Cundiff et al. (1971)	0.41 ± 0.19	0.31 ± 0.17	0.50 ± 0.19	----	0.56 ± 0.21	AN, HP &SP
Eriksson et al. (2003)			0.37		0.37	SM
			0.46		0.48	HP
	----	----	0.34	----	0.5	CH
Gregory et al. (1995)	0.22 ± 0.08	0.48 ± 0.09	0.25 ± 0.08	----	0.23 ± 0.08	RP, BV, HP, AN, SM, LM, CH & XB
Kim et al. (2003)	----	----	0.41	----	0.45	AN, BR &XB
Koch et al. (2004)	0.33	0.36	0.41	----	0.26	HP
Lamb et al. (1990)	0.28	0.33	----	0.24	0.31	HP
Mao et al. (2013)	0.64 ± 0.15	0.74 ± 0.14	0.17 ± 0.11		0.29 ± 0.06	CH
	0.49 ± 0.14	0.37 ± 0.11	0.35 ± 0.12	----	0.23 ± 0.08	AN
Nkrumah et al. (2007b)	0.45 ± 0.15	0.49 ± 0.16	0.51 ± 0.15	0.58 ± 0.18	0.33 ± 0.14	XB- (AN, CH, HP)
Reverter et al. (2000)	0.38		0.27		0.54	HP
	0.26		0.28		0.31	AN
Riley et al. (2002)	0.44	0.47	0.63	0.71	0.55	BR
Veseth et al. (1993)	0.51	0.31	----	----	0.38	HP
Wilson et al. (1976)	0.42	.33	0.41	----	0.21	HP & AN

*Breed abbreviations: AN- Angus; CH- Charolais; HP- Hereford; LM- Limousin; SM- Simmental; SP- Shorthorn; RP- Red Poll; BR- Brahman; Braunvieh- BV; XB- Beef crossbreeds

CHAPTER II

COMPARISON OF ACTUAL VERSUS PREDICTED FEED INTAKE PHENOTYPES FOR GENETIC EVALUATION OF FEED EFFICIENCY IN BEEF CATTLE

Abstract

Feed efficiency is expensive to measure in beef cattle because of the technology it requires to measure individual animal dry matter intakes (DMI). However, genetic correlations between predicted DMI (pDMI), which utilizes pen feed intake, and actual DMI indicate that pDMI may be useful as an indicator trait for genetic evaluation. Therefore, the objective of this study is to evaluate whether quantitative trait loci (QTL) mapping approaches identify the same regions of the genome for pDMI and DMI. Because average daily gain (ADG) is a primary driver of the prediction models and residual feed intake (RFI) is a popular metric for feed efficiency, the overlap of pDMI and DMI QTL regions with QTL for ADG and RFI will also be evaluated. To achieve these objectives, individual animal feed intake, weight, and carcass data was obtained on 849 Hereford steers and heifers fed within a GrowSafe (GrowSafe Systems Ltd.) feed intake system. The Cattle Value Discovery System (CVDS) growth and carcass data model was utilized to obtain pDMI from DMI pooled within pens and reallocated to individual animals. Phenotypic correlations were 0.64 ($P < 0.0001$) and 0.56 ($P < 0.001$) between pDMI and DMI and pDMI and ADG, respectively. Phenotypic correlations for RFI and ADG were zero ($P > 0.35$), as expected, and 0.38 ($P < 0.001$) and -0.08 ($P < 0.02$) between RFI and DMI and RFI and pDMI, respectively. Genotypes were assayed using the Illumina BovineHD Beadchip assay. After filtering for quality control, a final dataset of 648,625 single nucleotide polymorphisms (SNP)

was available for analysis. The SNP effects for RFI, ADG, pDMI, and DMI were estimated utilizing a BayesB0 model in GenSel. The 5-SNP windows surrounding the 100 largest SNP effects for each phenotype were compared to determine overlap between QTL regions. Seven of the QTL regions in common between pDMI and DMI were independent of ADG QTL regions. A pathway analysis of genes found in the top 50 QTL regions identified pathways implicated in human obesity. These results show that there is concordance between genomic regions for pDMI and feed intake and efficiency traits independent of the model drivers (ADG).

Introduction

An important goal of beef cattle producers is to increase the efficiency of production by either decreasing input costs or increasing outputs, such as pounds of product produced. Feed costs are a logical input to reduce because feed can contribute up to 70% of total production costs (Moore et al., 2006). One way to decrease feed costs is to raise cattle that are feed efficient. An animal that is feed efficient will eat less feed than other animals but will still gain weight at the same level or even faster compared to other animals with the same feed resources because they are more metabolically efficient.

In the past, selection for efficiency has largely involved selecting for increased outputs, including average daily gain (ADG), because of the ease of phenotype collection. Gain:feed and feed:gain ratio, or feed conversion ratio (FCR), which are ratios of daily dry matter intake (typically calculated based on pen feed intake) and body weight, have historically been popular metrics for feed efficiency. As with selecting for ADG directly, selection for FCR focuses on weight gain (Archer et al., 1999). Although improvement in feed efficiency was achieved by selecting for gain, selecting for ADG often leads to a larger mature size (Archer et al., 1999).

Residual feed intake (RFI) expresses an animal's deviation from a predicted intake of the animal based on gain and weight (Koch et al., 1963). Residual feed intake has gained popularity because it is phenotypically independent of growth rate (Kennedy et al., 1993). One drawback

includes the fact that RFI is a ratio trait, which can be misleading because an animal that eats less feed and gains less weight can have a similar RFI to an animal that eats more and gains more. Another drawback of RFI is that it is not genetically independent of its component traits (Kennedy et al., 1993). Another drawback to using RFI is that it requires individual animal feed intake values which can be expensive to collect. The reason genomic selection approaches for feed efficiency are desirable in the beef cattle industry is because it allows for the expansion of data collected in smaller discovery populations to be extrapolated effectively to populations that do not have phenotype data. This approach is particularly effective in traits that are expensive to measure on large herds of animals.

A multitude of studies have identified quantitative trait loci (QTL) regions linked to feed efficiency (i.e. Rolf et al., 2012; Saatchi et al., 2014; Barendse et al., 2007; Nkrumah et al., 2007a). While the identification of these QTL are critical, animal breeding and selection practices have also focused on genomic selection approaches, such as the estimation of direct genomic values (DGV) (Weber et al., 2012; Saatchi et al., 2013a). However, the development of DGV concomitantly enables the identification of QTLs for those traits. DGVs are especially useful for measuring traits, such as feed efficiency, that are traditionally difficult to measure due to cost and the labor it requires (Weber et al., 2012).

Feed efficiency evaluation in beef cattle is hampered by the fact that dry matter intake (DMI) phenotypes are expensive to collect and, although it is expanding rapidly, there is limited infrastructure for the collection of this data on a large scale. The development of a cost-effective indicator trait, such as predicted feed intake values based on animal performance and pen intake data, could enhance the ability of the industry to utilize already-existing data to supplement feed efficiency evaluations. These models, such as the Cattle Value Discovery System (CVDS), already exist and have been characterized in the literature. The CVDS program was first published by Fox and Black (1984) with later modifications detailed in Fox et al. (1988), Fox et al. (1992), Tylutki et al. (1994), and Fox and Tylutki (1998). The CVDS system predicts

individual DMI values for cattle by allocating pen fed intake data to individual animals based on each animal's performance. Fox et al. (2004) reported that CVDS predictions were only 2% lower than the actual intakes. In addition, another study discovered that CVDS under-predicted the intake of steers by only -0.91% and over-predicted for heifers by 0.89% (Fox et al., 2001).

An indicator trait for genetic evaluation should ideally be less expensive to measure, and exhibit a high genetic correlation with the trait of economic interest. Williams et al. (2006) reported phenotypic correlations for pDMI and DMI of approximately 0.785. However, the genetic correlation was considerably higher at 0.95 ± 0.07, as seen in Table 3. These results suggest that pDMI may be useful in supplementing actual DMI for genetic evaluation. The purpose of this study was to generate DGV predictions for a variety of feed efficiency related phenotypes in Hereford cattle and utilize the resulting QTL regions to evaluate concordance rates between predicted feed intake and other feed efficiency traits.

Materials and Methods

Animals and phenotypic data

Phenotypic records for a total of 870 Hereford steers and heifers fed between 2009 and 2011 were obtained, which were a subset of records from Saatchi et al. (2014). A total of 8 contemporary groups were formed based on calving season, calving year, number of days on feed, and sex, ranging in size from 31 to 205 animals. In addition, since sex was an identifying characteristic for assigning contemporary groups, one animal was in a contemporary group by itself so this animal was removed.

Feed intake records were collected using a GrowSafe System (Airdrie, Alberta, Canada) which can measure individual feed intakes on large numbers of pen-fed animals. Additional phenotypes that were collected included bi-weekly body weights, birth weight (BW), weaning weight (WW), yearling weight (YW), metabolic mid-weight (MMWT), and average daily gain

(ADG). Table 4 indicates each growth phenotype and how many records there were for each trait. Residual feed intake was calculated as follows:

$$RFI = ADMI - eDMI$$

$$eDMI = b_0 + b_1ADG + b_2MMWT^{0.75}$$

The calculation for RFI involves subtracting the expected dry matter intake (eDMI) from the average dry matter intake (ADMI). The expected feed intake is calculated from a regression of ADG and MMWT. Descriptive statistics for feed efficiency traits are provided in Table 4.

Carcass data was also collected, including hot carcass weight (HCW), marbling score (MARB), rib eye area (REA), back fat thickness (BFAT), and yield grade (YG) and is summarized in Table 5.

The phenotypic correlations and correlations between SNP effects for each trait were calculated using SAS© software, version 9.3 of the SAS System for Windows (SAS Institute Inc., Cary, NC, USA).

Predicted Feed Intakes

Growth and carcass data were utilized to calculate pDMI within CVDS. The "Individual Management" application within the program was used to allocate pen feed intake data to individual animals. All default settings for this program were used. This program utilizes pen feed intake records, ration information, periodic body weights, YG, HCW, BFAT, MARB, and REA data to predict the dry matter intake required for an animal to achieve the their specified level of performance. Specific ration information was unavailable for these animals, so a standard feedlot ration fed at the Willard Sparks Beef Research Center (Oklahoma State University, Stillwater, Oklahoma) was utilized. This ration, on a dry matter basis, consisted of approximately 17% cracked corn grain, 22% prairie hay, 43.5% sweet bran, 13% distiller's grain, and 4.5% B-273 supplement. Depending on the contemporary group of the animals, they were either on feed for 70, 71, or 140 days.

Genotypic data

All animals with genetic information (n=849) was genotyped using the Illumina Bovine HD BeadChip (770K) (Illumina Inc., San Diego, CA; Rincon et al., 2011). Animals which were in single-animal contemporary groups, or that did not possess both phenotypes and genotyped were excluded from further analysis. The dataset was filtered for extreme deviations from Hardy-Weinberg Equilibrium (HWE) and for a minor allele frequency (MAF) < 0.01. The missing genotypes in the dataset were imputed using BEAGLE (Browning and Browning, 2009). Files were filtered for marker heterozygosity (HETA) if the value was greater than 0.40. The data was also filtered for Kleinfelter males, call rate and incorrect chromosome assignment. After all filtering was complete, 648,625 SNPs were available for analysis on 849 animals.

Data analysis

Direct Genomic Value (DGV) Predictions

DGV were predicted utilizing a Bayesian framework implemented within GENSEL (Fernando and Garrick, 2009), which was accessed through the iPlant Collaborative website (Goff et al., 2011). Because the genomic architecture of pDMI is unknown, a BayesB analysis was utilized due to its ability to estimate allele substitution effect-specific variances rather than the common variances utilized in BayesC analyses. However, BayesB analyses are particularly sensitive to the starting values (Rolf et al., 2015), so BayesC analyses where $\pi = 0.95$ were utilized to estimate variance components to be utilized in subsequent BayesB analyses. Starting values for all BayesC analyses were the average of literature estimates of genetic and residual variance components, and are reported in Table 6. No previous literature estimates for the variance components of pDMI were found, so the variance components for actual DMI were also used for pDMI.

BayesB analyses where $\pi = 0$ were utilized to generate DGV, and a five-fold cross validation was used to ensure that reported DGV accuracies for each animal were generated when

they were independent of the training population. Training and validation populations for the cross validation were created by randomly allocating animals from each contemporary group to one of five populations to ensure an equal representation of each contemporary group in each training and validation group so that accurate estimates of contemporary group effects would be produced in each training run. Any animals remaining after initial assignment to contemporary groups (numbers not divisible by five) were combined, and randomly assigned to one of the five populations irrespective of their contemporary group.

Training populations were formed as a combination of four of the five groups, and the validation group was specified as the one remaining population so that all animals received an independent prediction and accuracy, as seen in Table 7. Accuracies and correlations for all five analyses are also reported in Table 7. Correlations are between the direct genomic value (DGV) and the phenotype, as de-regressed breeding values were not available for analysis in this population. In addition, realized accuracy for each trait was calculated as $\frac{r_{\hat{g},y}}{\sqrt{h^2}}$ to account for the fact that phenotypes, rather than de-regressed breeding values, were utilized for the calculation of the correlations (Rolf et al., 2015).

The additive genetic and residual variance components obtained from the BayesC results, as seen in Table 8, were utilized as starting values for all BayesB analyses. Each trait was analyzed in GENSEL using 60,000 Markov chain Monte Carlo (MCMC) iterations with a burn-in value of 1,000.

Fixed effects in this analysis consisted of the phenotypic means and contemporary group, therefore, the model for the Bayes B0 analysis was:

$$y_i = \mu + Xb + \sum_{j=1}^{k} z_{ij}u_j + e_i$$

where:

y_i= feed efficiency phenotypes for each trait,

μ= sample mean,

X= matrix corresponding to fixed effects vector b,

b= vector of fixed effects for contemporary groups and phenotypic mean,

k= number of genetic markers in the analysis,

z_{ij}= allelic state (AA=-10, AB=0, BB=10) of animal i at marker j,

u_j= random effect for marker j, and

e_i= residual effect.

Genome-wide association study

Development of DGV has the added benefit of simultaneously providing results for genome-wide association analyses. If SNPs were within 250,000 base pairs of each other, they were collapsed together. Then the SNP windows were collapsed together if the minimum or maximum genomic position of one SNP overlapped with the position of another SNP window until the top 100 regions were identified. The average size of each SNP window was 22,000 with a SNP density of approximately one SNP per 4,317 base pairs. Regions comprised of five-SNP windows around the top 100 largest effect SNP regions were compared between traits to identify the concordance rate between pDMI and the other feed efficiency traits in the analysis. Using such a small window size is a highly conservative approach for assessing the concordance rate, and the actual concordance rate would likely be higher if larger regions, such as those that approximate the range of linkage disequilibrium within the genome, were utilized. The 50 largest effect SNP regions were retained for further analysis. The top 50 largest effect SNPs were also compared to results found in the CattleQTLdb (release 26; Hu et al., 2013; Hu et al., 2010; Hu and Reecy, 2007; Hu et al., 2005) to identify whether the QTL regions identified in this study could be validated independently through the scientific literature.

Pathway analysis

The top 50 SNP regions with the largest effect for each trait were used for gene annotation clustering analyses. The 50 SNP regions with the largest effect were identified and a 500 Kb window was again formed around each SNP so that each SNP within 250 Kb on either side of the primary SNP of interest was included in the list. This list was then imported into Ensembl to identify genes within these regions (Cunningham et al., 2015). These lists of genes for each trait were then utilized for Functional Annotation Clustering within the Database for Annotation, Visualization, and Integrated Discovery (DAVID) version 6.7 (Huang et al., 2009a, Huang et al., 2009b) to group these SNPs based on their biological function. The use of DAVID also allows for the identification of SNPs involved in gene pathways through utilization of the KEGG pathway database in an effort to identify pathways that are critical for feed efficiency.

Because of the range of linkage disequilibrium in the bovine genome (McKay et al. 2007), 500 kb windows were formed around the top 25 SNPs with the largest effect for each trait. Twenty-five SNPs rather than 50 were utilized due to the SNP density and the amount of information that can be visualized within a gene network. All SNPs within these windows and their effects for each trait were analyzed with a Partial Correlation and Information Theory (PCIT) algorithm (Koesterke et al., 2014; Reverter and Chan, 2008) developed at Iowa State University and implemented within the Texas Advanced Computing Center (TACC) at the University of Texas. The PCIT algorithm utilizes the SNP effect information and calculates direct and partial correlations between various SNPs in the model to evaluate their association with one another (Watson-Haigh et al., 2010). These partial correlations can then be visualized in the Cytoscape software which allows for the visualization of gene networks (Shannon et al., 2003).

Results and Discussion

Phenotypic correlations

A plot of observed and predicted DMI is presented in Figure 1. The phenotypic correlation found in this study (0.64), as reported in Table 9, was lower than previous literature estimates (0.785; Williams et al. 2006). One possible explanation for the lower correlation is the lack of availability of ration information for the pDMI predictions. The phenotypic correlation for RFI and DMI was also lower in this study (0.38) compared to previous estimates of 0.602 (Arthur et al., 2001b), 0.72 (Arthur et al., 2001a), and 0.70 (Herd and Bishop, 2000). The correlation between RFI and ADG was not different from zero (0.03; $P < 0.357$) as expected due to the properties of the calculation of RFI. The correlations for pDMI with ADG and RFI were 0.57 and -0.08, respectively.

Heritabilities

Heritabilities for each trait are reported in Table 10. The trait with the highest heritability was DMI (0.40), which is intermediate to previous literature estimates (0.27 to 0.64; Robinson and Oddy, 2004; Koch et al., 1963). Literature estimates of the heritability for RFI range from 0.16 to 0.68 (Herd and Bishop, 2000; Mao et al., 2013), which is also consistent with our heritability estimate of 0.38. However, the heritability estimate for ADG in this study (0.20) was lower than previous literature estimates which ranged from 0.28 to 0.59 (Mujibi et al., 2011; Nkrumah et al., 2007b), possibly because intense selection has been placed on gain. The heritability estimate for pDMI in this study was 0.23 which is slightly lower than the only other literature estimate of pDMI (0.32; Williams et al., 2006).

Direct Genomic Values

DGV correlations for all traits are provided in Table 10 and were 0.22, 0.13, 0.32 and 0.28 for ADG, pDMI, DMI, and RFI, respectively. The range of correlations and accuracies for each training and validation analysis can be seen in Table 8. The analyses on pDMI brought the range down and increased the variation because it is a prediction and so it was not as accurate as

the actual observations. The correlations for ADG and RFI were lower than previous literature estimates (0.414 and 0.402, respectively; Mujibi et al., 2011) while DMI was higher than previous estimates (0.27; Mujibi et al., 2011). A possible explanation as to why ADG and RFI correlations were lower than previous estimates could be due to a difference in breed since the study by Mujibi et al. (2011) used crossbred animals with only a subset of the animals having a proportion of Hereford lineage. The calculated DGV realized accuracies for ADG, pDMI, DMI, and RFI were 0.49, 0.27, 0.51, and 0.46, respectively. These accuracies, provided in Table 10, are moderate to high given the sample size compared with other studies (Rolf et al., 2010; Pryce et al., 2012) although pDMI did exhibit the lowest accuracy. This study utilized different breeds compared to studies by Rolf et al. (2010), which used Angus, and Pryce et al. (2012), which used Holstein. In addition, the sample size for this study was smaller which is important because the variation in the phenotype would not be as noticeable in the results of a large sample size since all of the other records average out the outliers.

Association Analyses

Manhattan plots for each of the phenotypes are provided in Figures 2, 3, 4, and 5 for ADG, DMI, pDMI, and RFI, respectively. Overall, the SNP effects for ADG were universally small and tended to follow the infinitesimal model. Plots for pDMI followed this same trend, but RFI and DMI had SNPs that had similar effect sizes, and the largest effect SNPs were not substantially larger than other SNPs, unlike in pDMI and ADG. The correlations between the SNP effects between traits were calculated and can be seen in Table 11. The correlation was highest for ADG and pDMI with a correlation of 0.7583, which is not surprising since ADG is used to calculate the value for pDMI. The correlations were lowest between ADG with RFI and pDMI with RFI, which had correlations of 0.054 and 0.0623, respectively. These correlations indicate there were different regions of the genome that were responsible for the phenotypes. The

low correlation between ADG and RFI is what would be expected since RFI is independent of gain.

The QTLs validated for the top 50 largest effect SNP regions for each trait are reported in Table 12. For ADG, two QTLs could be validated from a study that included 698 Angus cattle (Rolf et al., 2012). One of these QTLs was found within a 2 Mb region of chromosome 10 (27,034,490 to 29,073,969 bp). Within this region there are coding regions for the olfactory receptor gene, lysophosphatidylcholine acyltransferase 4 (LPCAT4), ribonucleoprotein homolog (NOP10). According to Ensembl, the NOP10 gene has been associated with liver and gut development in zebrafish. The top 50 SNP effect regions for pDMI had two SNP regions that matched with previous QTLs detailed in the Cattle QTLdb (Nkrumah et al., 2007b; Veerkamp et al., 2012). One of the QTLs was on chromosome 15 at 14,318,486 to 14,440,400 bp and was identified by Nkrumah et al. (2007b), but there were no genes found to be associated with this region. A region on chromosome 3 at location 85,304,450 to 86,538,180 bp in Holstein cattle was also validated in our study (Veerkamp et al. 2012). In addition, QTLs for DMI were validated in this study. The QTL that was found for DMI was the same QTL that was found for pDMI, as was previously mentioned, on chromosome 15 (Nkrumah et al., 2007b). The QTLs for DMI which were validated in the literature include regions on chromosomes 2 and 22 (Martinez et al., 2010; Lu et al., 2013). A study by Sherman et al. (2009) identified two QTL for RFI on chromosome 3 at 68,923,918 to 69,125,116 bp in crossbred beef cattle that were also identified in this study. In addition, two different QTLs in this study matched with QTLs identified for RFI in a study by Nkrumah et al. (2007b) located on chromosomes 8 and 19. Furthermore, two associations were discovered by Lu et al. (2013) who identified the same area of the genome as in this study for RFI on chromosome 22 at chromosome position 51,303,323 to 51,388,333 bp. Besides the genes identified for QTL regions for ADG, no other QTL regions were found to have genes associated with these genome regions.

Concordance between Genomic Regions

Concordance rates between the top 100 QTL regions for all feed efficiency traits are reported in Table 13. The concordance rates between the traits are important because two traits that exhibit a large genetic correlation should share a large number of QTL regions. The largest effect SNP for ADG can be seen in Figure 2 as a peak on chromosome 8 at 764,159 bp with an effect size of 0.000333. This same peak is also identified for pDMI, as can be seen in Figure 4. Closer inspection of these specific regions, as seen in Figures 6 and 7, reveals that the SNPs in the peaks are the same between the two traits. The SNPs denoted with a circle identify SNPs within the same SNP window on chromosome 8 for both traits. The same SNPs are identified in Figures 8 and 9 for DMI and RFI, respectively. Another large-effect SNP region for ADG was on chromosome 7 centered at a genomic position of 93,206,020 bp with an effect size of 0.000213, as seen represented by a triangle in Figure 6 and for other traits in Figures 7, 8, and 9 for pDMI, DMI, and RFI, respectively. Overall, this SNP region also seems to have a large effect among all of the traits that were tested with the lowest effect listed for this SNP being 0.00006 for RFI. This also happens to be the largest effect SNP for pDMI with an effect size of 0.00206, as seen in Figure 4. The largest effect SNP region for RFI was on chromosome 13 centered at a genomic position of 27,686,629 with an effect size of 0.000158, although it was not in the top 100 SNP regions for any other trait. The largest effect SNP for DMI (Figure 3) was on chromosome 3 centralized at a genome position of 70,094,743. This SNP region also has a large effect for pDMI which can be seen in Figures 10 and 11 (denoted by a triangle). Another large effect SNP that was common between DMI and pDMI was located on chromosome 3 (Figures 10 and 11; designated by a circle).

The concordance rate was highest between pDMI and ADG, likely because the CVDS model is largely driven by weight gain. The lowest concordance rate was between ADG and RFI (2%), which is expected because the calculation for RFI forces the trait to be phenotypically independent of gain. The concordance rates for RFI with pDMI and DMI were 6% and 15%,

respectively. Sixteen regions were in common between ADG, pDMI, and DMI, indicating that some of the variation in these traits is shared. Residual feed intake has the lowest concordance rate when compared with all of the other feed efficiency traits. In addition, the concordance rates for DMI compared to ADG and pDMI was 19% and 26%, respectively. While the concordance rate for DMI and pDMI was not as high as the genetic correlations suggested, the approach utilized was very conservative. Compared to the concordance between all of the traits in this study, pDMI and DMI do show considerable overlap, which indicates that it's utility as an indicator trait should be explored further. Seven of the regions in common between pDMI and DMI are not shared with ADG, which indicates that although pDMI does exhibit significant overlap with ADG, it also identifies some regions that are independent of gain. Because of this, it is possible that pDMI could be utilized in selection decisions, provided that the component traits are well understood and efforts to mitigate increases in mature size were in place to prevent increases in maintenance costs within the herd.

Gene network Analyses

The 25 largest-effect SNP regions for each trait and all the SNPs within 250,000 base pairs on either side of that SNP were analyzed through PCIT and the results were viewed in Cytoscape (Figure 12). The cluster was very large with 206,583 edges (correlations) and 8,024 nodes (SNPs). The clustering coefficient for the network was 0.450 and the average number of neighbors, which is the number of correlations a SNP has to other SNPs, was 51.49. The network density (0.006), was very close to 0 which indicates that most of the nodes were isolated and did not group strongly with each other, although no nodes were completely isolated. This data suggests that many of the nodes are associated with one another, due to a high average number of neighbors, but few nodes are densely clustered together.

Annotation clusters obtained from DAVID were considered significant with an enrichment score of 1.3, and were considered suggestive if the enrichment score was 1. Enriched

clusters are summarized in Table 14. The majority of genes in the enriched cluster associated with pDMI were a type of Rho guanine nucleotide exchange factor (ARHGEF). None of the genes had phenotypes identified in cattle, but the International Mouse Phenotyping Consortium (Brown and Moore, 2012) discovered ARHGEF4 has a link to decreased body weight in mice. In addition, ARHGEF11 has been associated with insulin resistance in humans (Ma et al., 2007). The other significant gene cluster was related to RFI and it had an enrichment score of 1.56. Some of the genes that were included in this cluster were insulin, insulin-like growth factor-2, gastric inhibitory peptide, cathepsin D, and TIMP metallopeptidase inhibitor-2. It is known that insulin is important in the absorption of glucose and improper absorption of glucose can lead to obesity in humans (Guilherme et al., 2008). In addition, the gastric inhibitory peptide (GIP) and TIMP metallopeptidase inhibitor-2 (TIMP2) genes in this cluster have also been linked to diabetes and obesity in mice and humans (Miyawaki et al., 2002; Jaworski et al., 2011). Another cluster that fell just short of the suggestive threshold (enrichment score of 0.99) was related to ADG and also included ARHGEF11. A cluster related to DMI had an enrichment score of only 0.77, but it was worth looking into because it contained the suppressor of cytokine signaling-3 gene (SOC3) which has been associated with leptin sensitivity in humans and mice, which can lead to obesity if the gene is not functioning properly (Howard et al., 2004). While the link between these genes and weight gain in cattle has not been directly identified, the function of these orthologous genes in other species provides evidence of their likely importance for feed efficiency and its component traits in beef cattle.

In addition to gene clustering, DAVID also allowed for the identification of genes that were involved in important pathways within the KEGG database. Overall, the pathways that were identified were not enriched and are not likely to have a critical part in a biological pathway that would have a direct impact on any of the phenotypes. An enriched KEGG pathway is one that has a P-value less than 0.05. For example, a pathway that was identified for ADG was the regulation of the actin cytoskeleton. The genes fibroblast growth factor-23 and fibroblast growth factor-6

were identified in the actin cytoskeleton pathway which was suggestive of being enriched ($P < 0.075$) and can be denoted by the red stars in Figure 13. In addition, a pathway identified for pDMI was related to O-glycan biosynthesis which has been shown to be important in immune function and lipid metabolism (Tian et al., 2009; Figure 14). The other phenotype that had a connection to a KEGG pathway was DMI which is linked to ubiquitin mediated proteolysis. Ubiquitins are critical in a variety of functions, including targeting proteins responsible for growth modulation (Ciechanover et al., 2000). The genes that are associated with this pathway are represented by red stars in Figure 15.

Conclusion

Realized accuracies for feed efficiency traits (ADG, RFI, pDMI, and DMI) ranged from 0.27 to 0.51. The GWAS analysis also identified regions of the genome important for predicting genetic merit in these traits, and identified biological pathways in which these genes are involved. Concordance rates for regions of the genome that are important in the prediction of pDMI were moderate, and they were lower than the genetic correlations between pDMI and DMI would have suggested. Concordance between pDMI and ADG were especially high, likely due to the fact that ADG is one of the main drivers of the CVDS model. It is possible that pDMI phenotypes could be utilized to enhance DMI information for genetic evaluation, provided that care is taken to ensure that selection emphasis is placed on DMI-related variation rather than gain to prevent unwanted increases in mature size.

Supplementary Materials

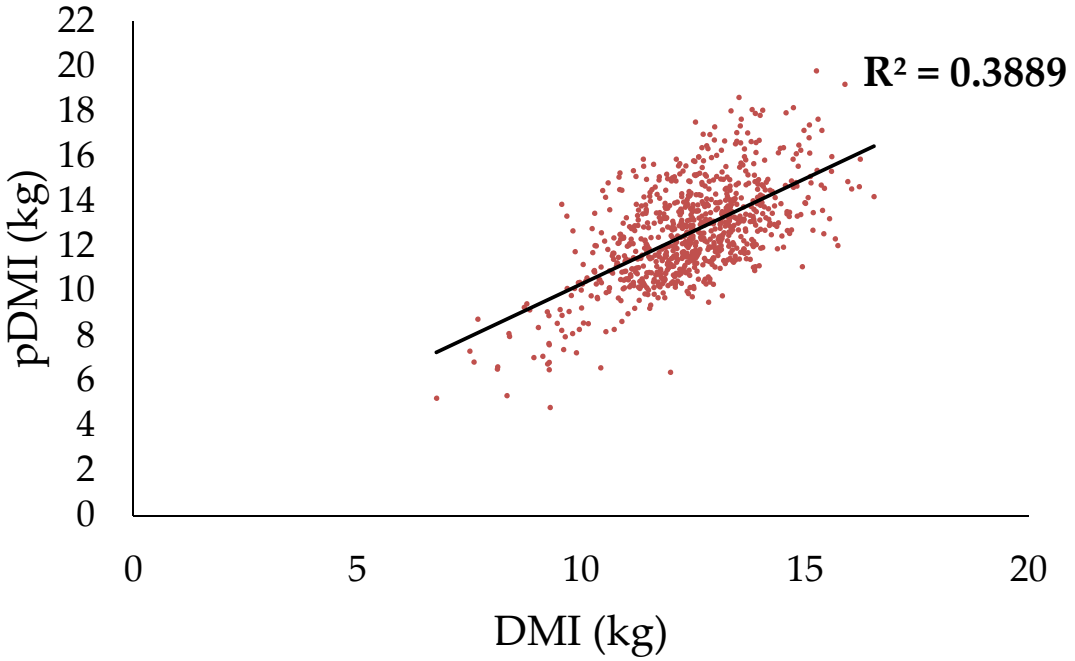

Figure 1- Phenotype plot for DMI vs pDMI data with an R^2 value (coefficient of determination) and a line of best fit

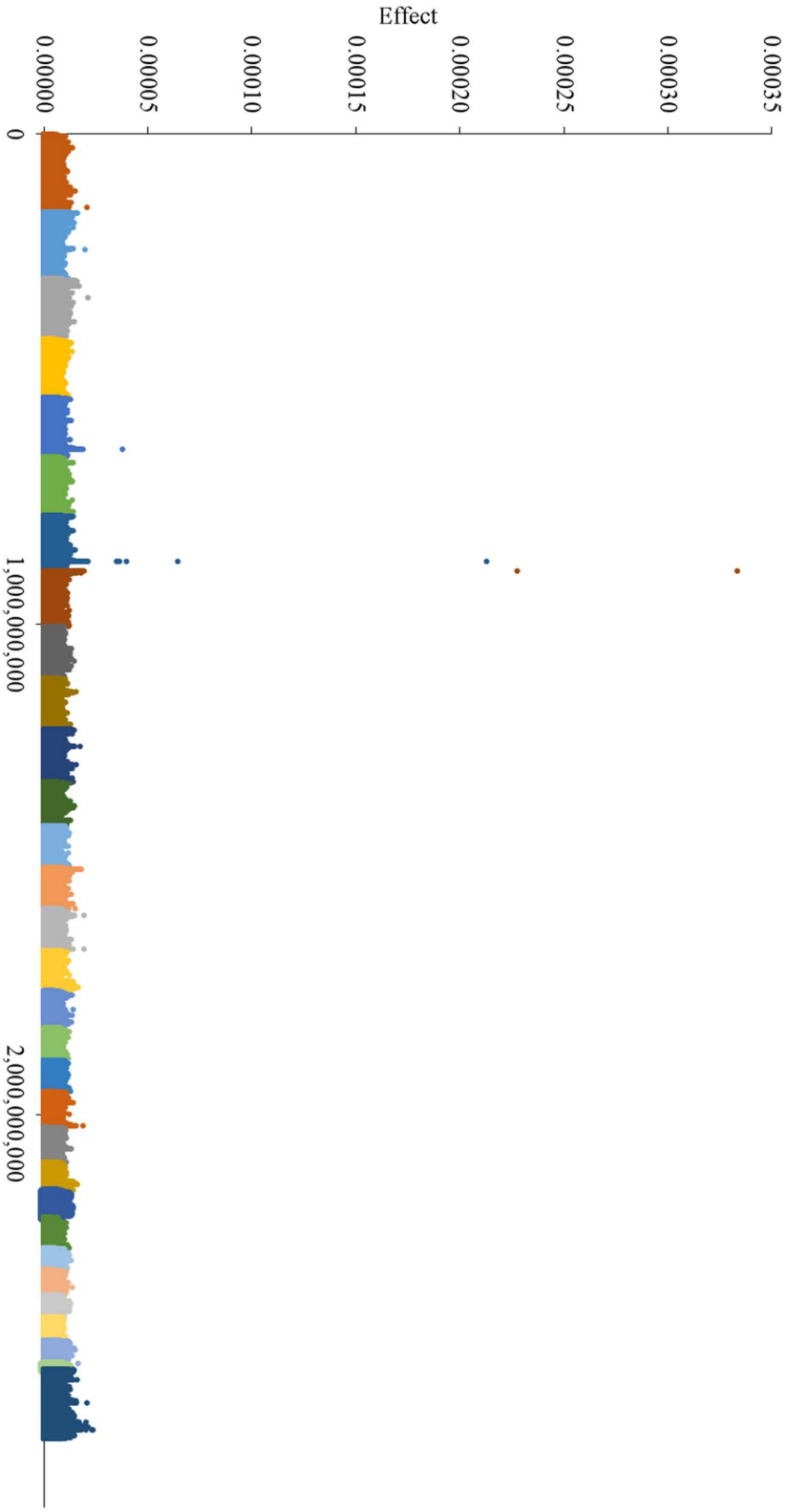

Figure 2- Manhattan plot for SNP effects for average daily gain from GenSel analysis in Hereford cattle. Each different color series represents a chronological chromosome.

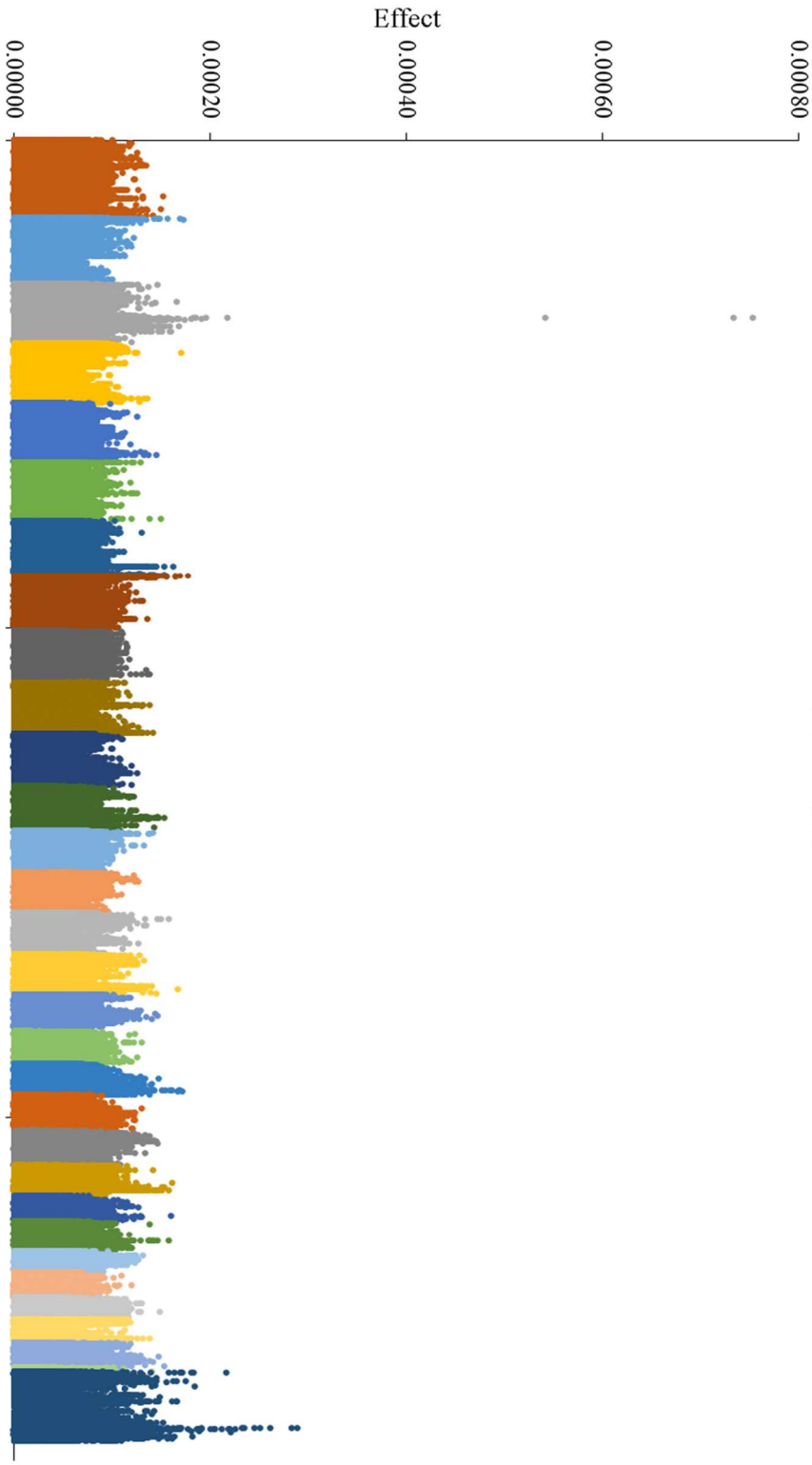

Figure 3- Manhattan plot for SNP effects for dry matter intake from GenSel analysis in Hereford cattle. Each different color series represents a chronological chromosome.

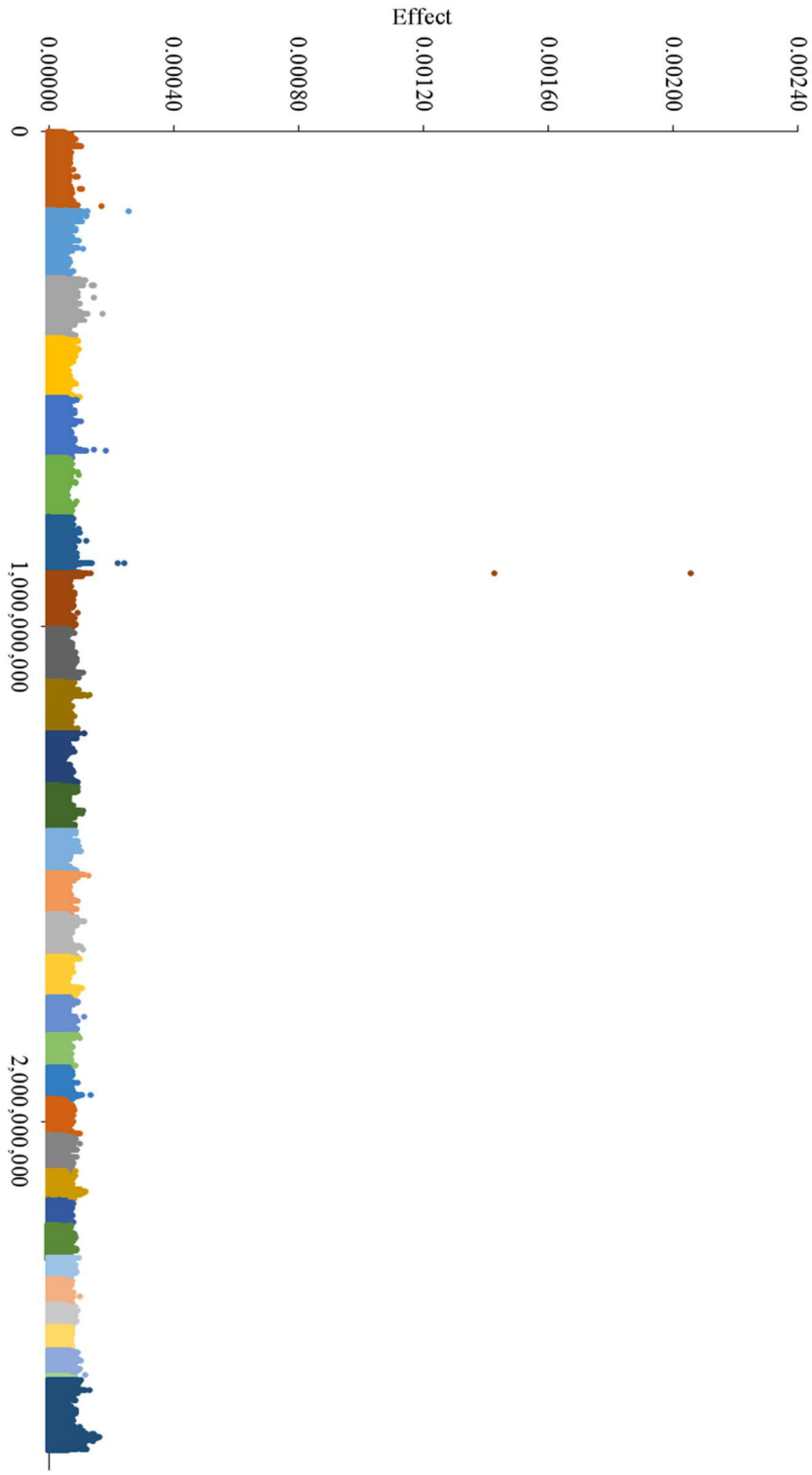

Figure 4- Manhattan plot for SNP effects for predicted dry matter intake from GenSel analysis in Hereford cattle. Each different color series represents a chronological chromosome.

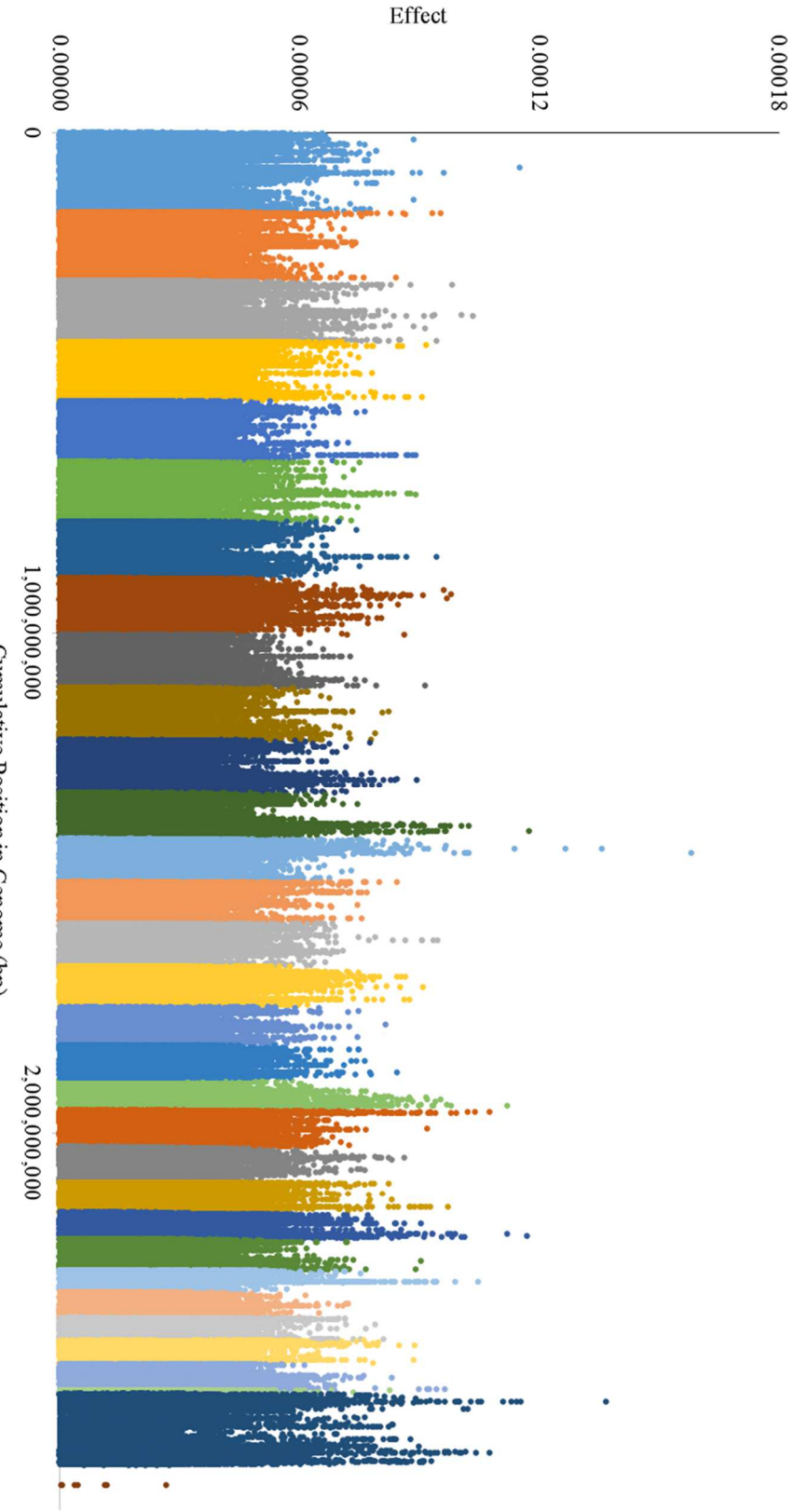

Figure 5- Manhattan plot for SNP effects for residual feed intake from GenSel analysis in Hereford cattle. Each different color series represents a chronological chromosome.

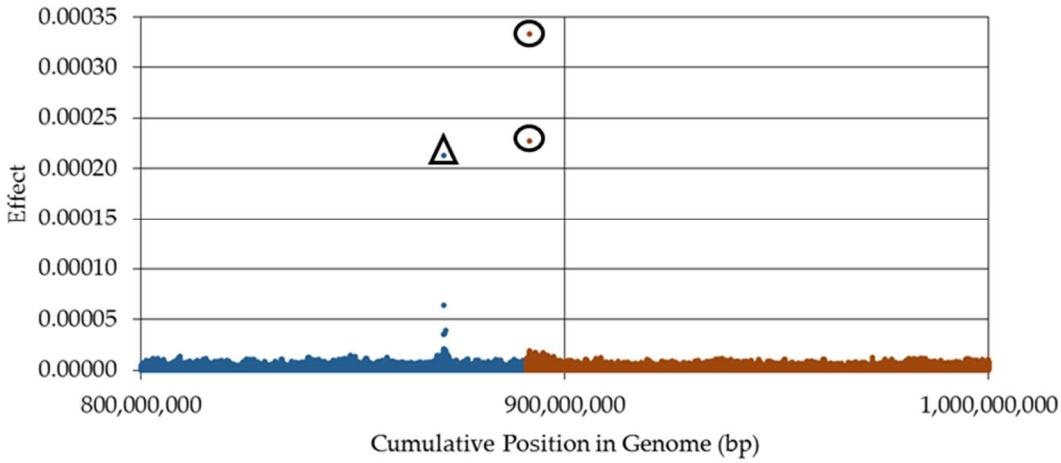

Figure 6- Selected concordance region (chromosomes 7 and 8) for ADG SNP Effects

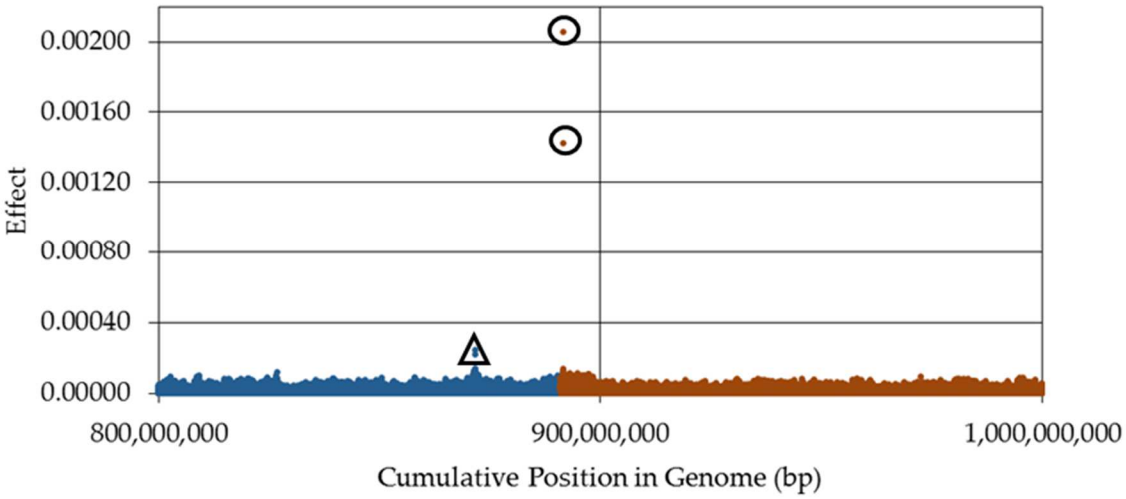

Figure 7- Selected concordance region (chromosomes 7 and 8) for pDMI SNP Effects

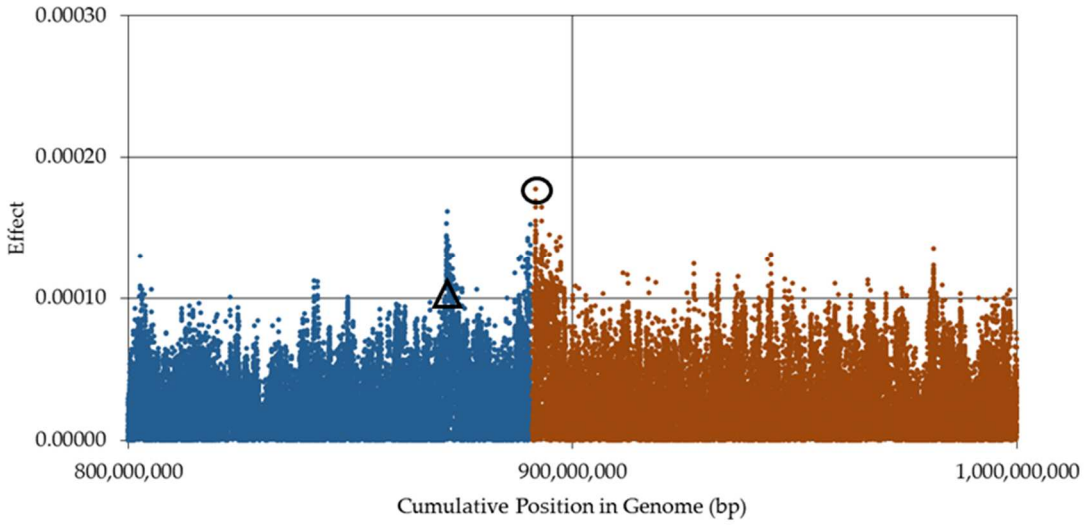

Figure 8- Selected concordance region (chromosomes 7 and 8) for DMI SNP Effects

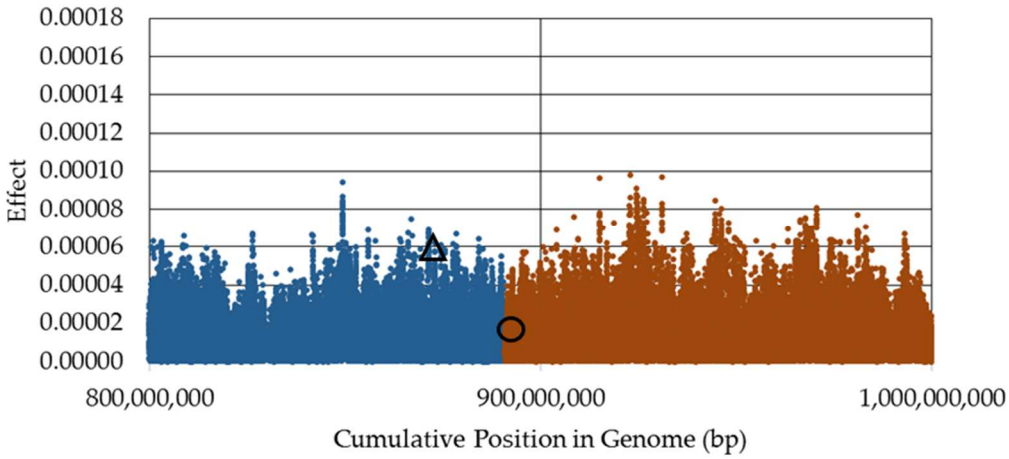

Figure 9- Selected concordance region (chromosomes 7 and 8) for RFI SNP Effects

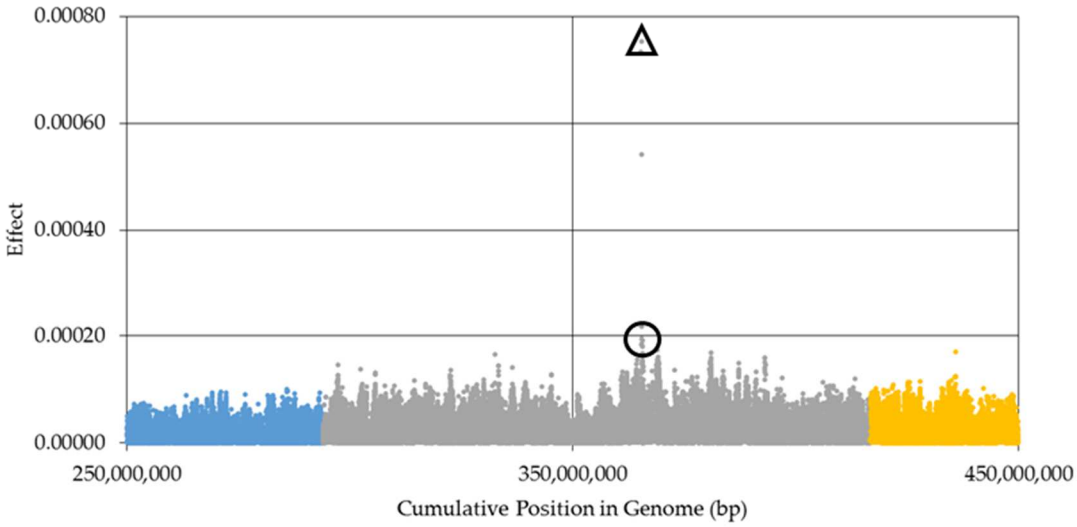

Figure 10- Selected concordance region (centered on chromosome 3) for DMI SNP Effects

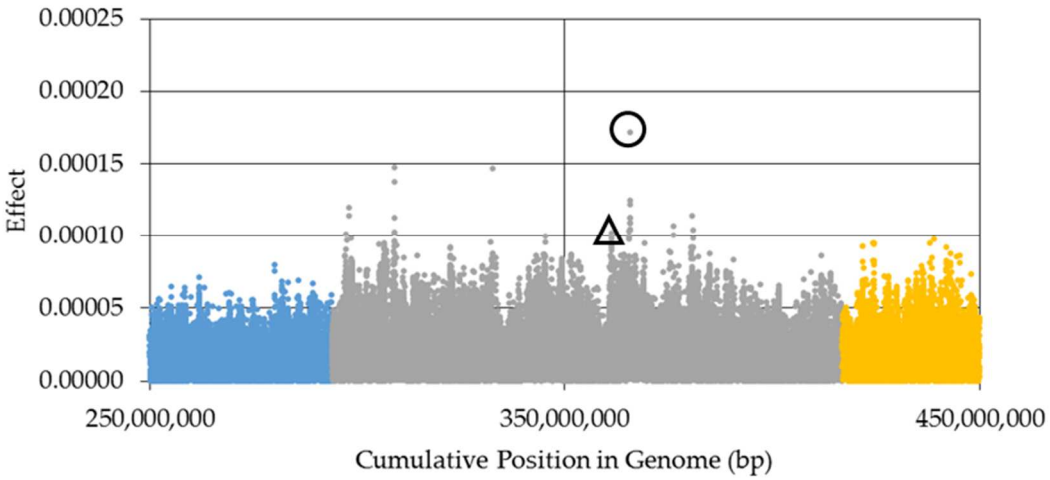

Figure 11- Selected concordance region (centered on chromosome 3) for pDMI SNP Effects

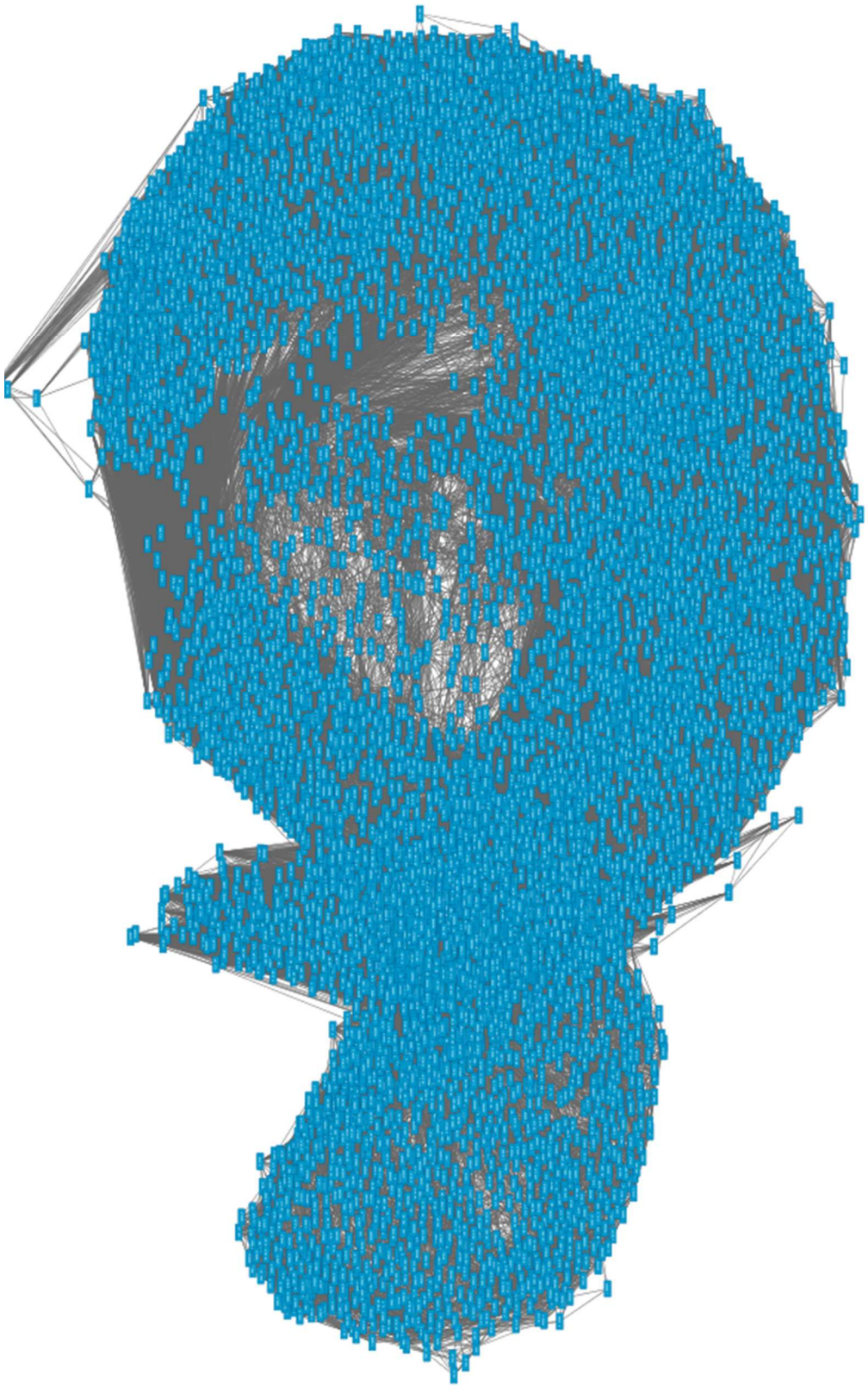

Figure 12- PCIT gene network results visualized in Cytoscape. Blue boxes represent SNPs and gray lines represent correlations.

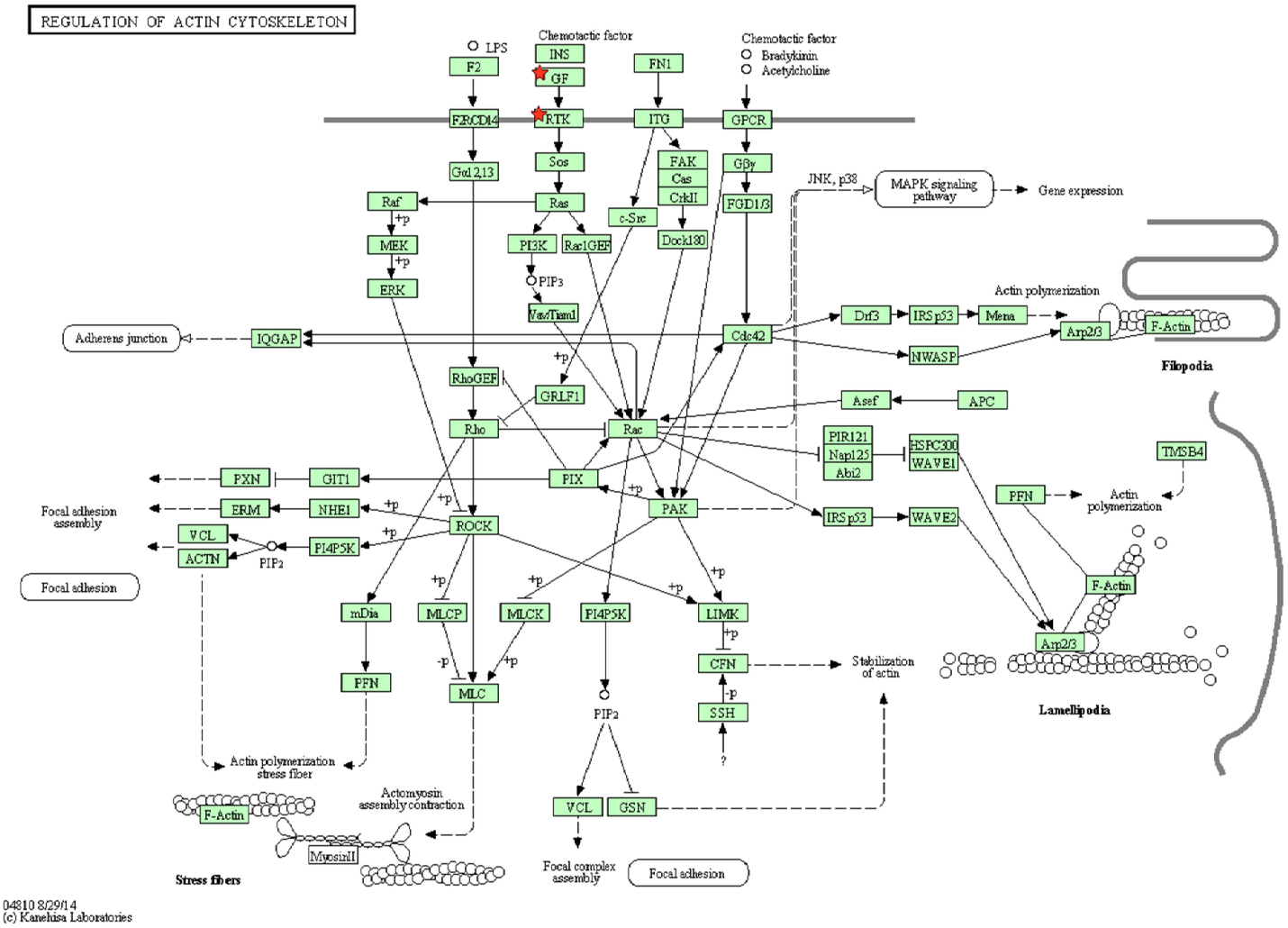

Figure 13. KEGG pathway for regulation of actin cytoskeleton from clustering analysis for average daily gain. Red stars indicate the genes in our study that were identified in this pathway.

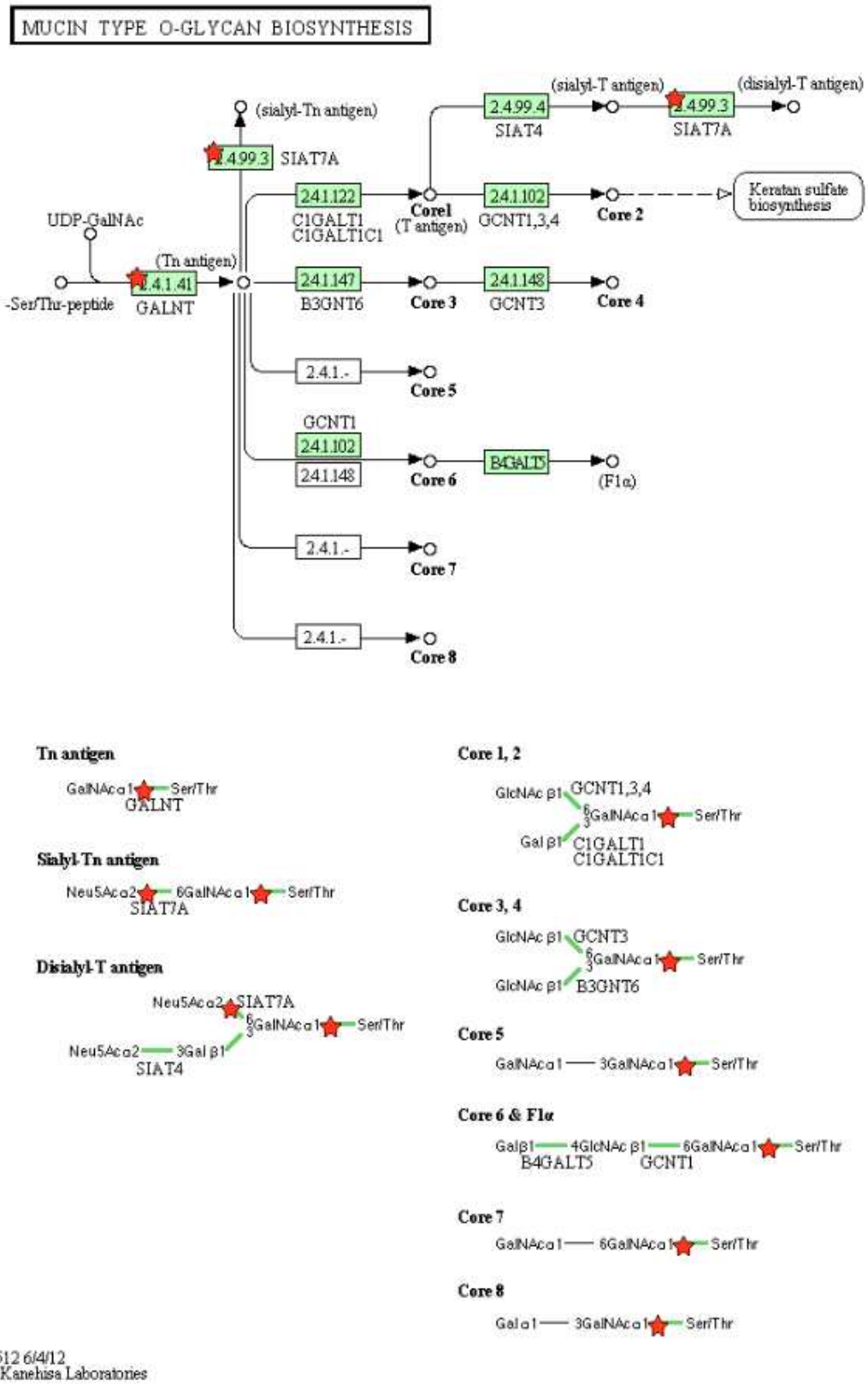

Figure 14. KEGG pathway for mucin type O-glycan biosynthesis from clustering analysis for predicted dry matter intake. Red stars indicate the genes in our study that were identified in this pathway.

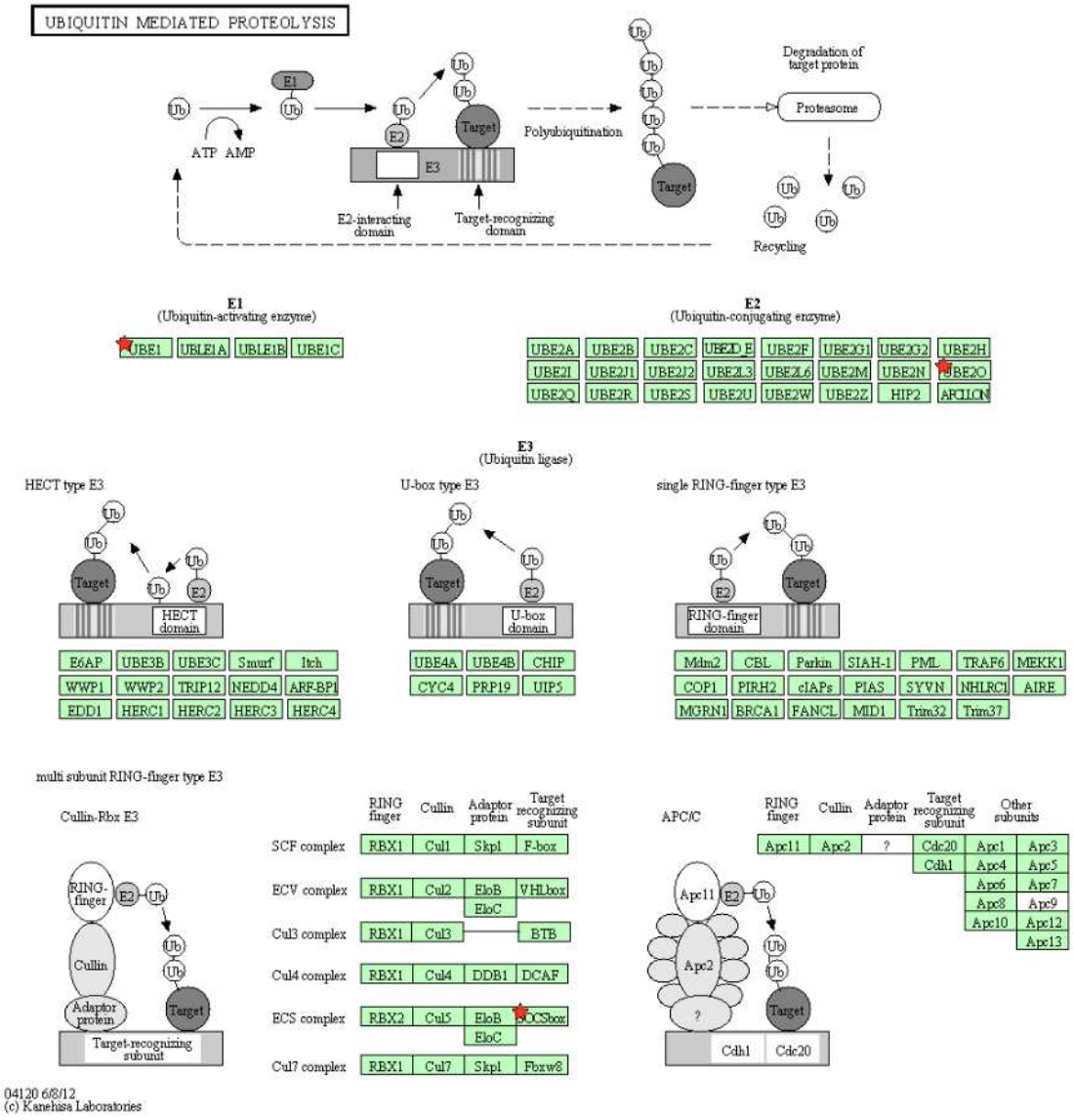

Figure 15. KEGG pathway for ubiquitin mediated proteolysis from clustering analysis for dry matter intake. Red stars indicate the genes in our study that were identified in this pathway.

Table 3- Literature estimates of phenotypic and genetic correlations* for dry matter intake (DMI), predicted dry matter intake (pDMI), average daily gain (ADG), and residual feed intake (RFI)

Trait	DMI	pDMI	ADG	RFI
DMI	1.0	0.785[1]	0.47[2]	0.602[2]
pDMI	0.95 ± 0.07[1]	1.0	0.90[1]	--
ADG	0.39 ± 0.08[2]	0.90 ± 0.06[1]	1.0	0.01(ns)[2]
RFI	0.79±0.04[2]	--	-0.10±0.13[2]	1.0

*Phenotypic correlations are above the diagonal, genotypic correlations are below the diagonal.
[1]Williams et al., 2006. [2]Arthur et al., 2001. (ns = not significant)

Table 4- Growth phenotypes collected and number of records, minimum, maximum, mean, and standard deviation for Hereford cattle*

Phenotype	BW	WW	YW	DMI	ADG	MMWT	RFI
Number of Records	820	824	844	846	849	849	846
Minimum	27	107	203	6.54	0.51	57	-3.73
Maximum	51	280	745	16.53	3.06	140	3.50
Mean	38.1	189.6	342.6	12.41	2.11	106.6	0.26
Standard Deviation	3.98	27.36	73.5	1.42	0.30	15.44	1.00

*BW- birth weight; WW- weaning weight; YW- yearling weight; DMI- dry matter intake; ADG- average daily gain; MMWT- metabolic mid-weight; RFI- residual feed intake

Table 5- Carcass phenotypes collected and number of records, minimum, maximum, mean, and standard deviation for Hereford cattle*

Phenotype	HCW	MARB	REA	BFAT	YG
Number of Records	849	849	840	836	775
Minimum	238	410	48.4	0	1.85
Maximum	498	870	109.4	3	5.81
Mean	407.1	562.4	85.8	1.51	3.72
Standard Deviation	36.00	75.24	8.51	0.45	0.64

*HCW- hot carcass weight; MARB- marbling score; REA- rib eye area; BFAT- backfat thickness; YG- yield grade

Table 6- Previous literature estimates of residual (σ^2_e) and genetic (σ^2_a) variance components and the average of each used in the BayesC95 analysis for Hereford cattle*

Phenotype	Source 1	Source 2	Source 3	Average
ADG (kg/d)	σ^2_e= 0.0195 σ^2_a= 0.0076 Arthur et al. (2001a)	σ^2_e= 0.013 σ^2_a=0.008 Herd et al. (2000)	σ^2_e=0.0115 σ^2_a= 0.008 Arthur et al. (2001b)	σ^2_e= 0.015 σ^2_a= 0.008
DMI (kg)	σ^2_e= 1.117 σ^2_a= 0.843 Nkrumah et al. (2007)	σ^2_e= 1.191 σ^2_a= 1.399 Herd et al. (2000)	σ^2_e= 1.71 σ^2_a= 0.64 Robinson & Oddy (2004)	σ^2_e= 1.34 σ^2_a= 0.96
RFI (kg/d)	σ^2_e= 0.628 σ^2_a= 0.139 Robinson & Oddy (2004)	σ^2_e= 0.608 σ^2_a= 0.162 Nkrumah et al. (2007)	σ^2_e= 0.363 σ^2_a= 0.274 Arthur et al. (2001b)	σ^2_e= 0.533 σ^2_a= 0.19
pDMI (kg)	σ^2_e= 1.117 σ^2_a= 0.843 Nkrumah et al. (2007)	σ^2_e= 1.191 σ^2_a= 1.399 Herd et al. (2000)	σ^2_e= 1.71 σ^2_a= 0 Robinson & Oddy (2004)	σ^2_e= 1.34 σ^2_a= 0.96

*DMI- dry matter intake; ADG- average daily gain; MMWT- metabolic mid-weight; RFI- residual feed intake

Table 7- Structure for analyses using randomized training and validation populations with range of DGV realized accuracies and correlations from GenSel GWAS analysis

Analysis #	**Randomly Assigned Populations**					**Accuracy**
1	Validation	Training	Training	Training	Training	0.31-0.50
2	Training	Validation	Training	Training	Training	0.35-0.63
3	Training	Training	Validation	Training	Training	0.33-0.51
4	Training	Training	Training	Validation	Training	0.10-0.56
5	Training	Training	Training	Training	Validation	0.28-0.72
Correlation	0.15-0.32	0.17-0.39	0.16-0.26	0.05-0.35	0.13-0.45	

Table 8- Residual (σ^2_e) and genetic (σ^2_a) variance starting values for BayesB0 GWAS analysis in GenSel*

Trait	σ^2_e	σ^2_a
ADG (kg/d)	0.0522	0.0126
DMI (kg)	0.960	0.643
RFI (kg/d)	0.43	0.26
pDMI (kg)	1.96	0.572

*DMI- dry matter intake; ADG- average daily gain; MMWT- metabolic mid-weight; RFI- residual feed intake

Table 9- Phenotypic correlations for dry matter intake (DMI), predicted dry matter intake (pDMI), average daily gain (ADG), and residual feed intake (RFI) in Hereford cattle

	pDMI	**DMI**	**ADG**	**RFI**
pDMI	n=832*	0.64[a]	0.57[a]	-0.08[b]
DMI	--	n=846*	0.49[a]	0.38[a]
ADG	--	--	n=849*	0.03[ns]
RFI	--	--	--	N=846*

[a]P<.001 [b]P<.05 ns=not significant
*Number of animals in each analysis.

Table 10- Heritabilities, correlations and accuracies, and final residual (σ^2_e) and genetic (σ^2_a) variance values from BayesB0 analysis for the feed efficiency phenotypes from GWAS analysis in GenSel*

	ADG	**pDMI**	**DMI**	**RFI**
Heritability	0.20	0.23	0.40	0.38
Average Correlation	0.22	0.13	0.32	0.28
Average Accuracy	0.49	0.27	0.51	0.46
$\boldsymbol{\sigma^2_e}$	0.052	1.95	0.947	0.435
$\boldsymbol{\sigma^2_a}$	0.136	0.610	0.657	0.249

*DMI- dry matter intake; ADG- average daily gain; MMWT- metabolic mid-weight; RFI- residual feed intake

Table 11- Correlations* of SNP effects between dry matter intake (DMI), predicted dry matter intake (pDMI), average daily gain (ADG), and residual feed intake (RFI) from ultra-high-density SNP data used in GWAS analysis

	ADG	**pDMI**	**DMI**	**RFI**
ADG	1.0	0.7583	0.3363	0.0540
pDMI	<0.0001	1.0	0.4177	0.0623
DMI	<0.0001	<0.0001	1.0	0.4632
RFI	<0.0001	<0.0001	<0.0001	1.0

*Correlations are above the diagonal, P-values for the correlations are below the diagonal

Table 12- Comparison of QTL results from this study to previously published studies for dry matter intake (DMI), predicted dry matter intake (pDMI), average daily gain (ADG), and residual feed intake (RFI)

Trait	QTL ID	Chr	Position (Mbp)	Association or Linkage	Validation Source
ADG	20941-20944	10	27.0-29.1	Association	Rolf et al. (2012)
	20981	10	27.0-29.1	Association	Rolf et al. (2012)
pDMI	22646	3	85.3-86.5	Association	Veerkamp et al. (2012)
	4368	15	14.3-14.4	Linkage	Nkrumah et al. (2007)
DMI	11872	2	3.6-3.9	Association	Martinez et al. (2010)
	4368	15	14.3-14.4	Linkage	Nkrumah et al. (2007)
RFI	5322-5333	3	68.9-69.1	Linkage	Sherman et al. (2009)
	4355-4356	8	40.1-40.3	Linkage	Nkrumah et al. (2007)
	4453	19	41.3-41.4	Linkage	Nkrumah et al. (2007)
	23912	22	51.3-51.4	Association	Lu et al. (2013)
	23913	22	51.3-51.4	Association	Lu et al. (2013)

Table 13- Concordance rates* of genome regions between dry matter intake (DMI), predicted dry matter intake (pDMI), average daily gain (ADG), and residual feed intake (RFI) from ultra-high-density SNP data used in GWAS analysis

	ADG	**pDMI**	**DMI**	**RFI**
ADG	1	50%	19%	2%
pDMI	--	1	26%	6%
DMI	--	--	1	15%
RFI	--	--	--	1

*Percentage of concordant regions out of 100 regions total

Table 14- Functional annotation clustering results from DAVID for feed efficiency phenotypes

Cluster #	Trait	Function	Enrichment score	Significant/ Suggestive	# of genes included
1	RFI	Vesicle	1.01	Suggestive	3
		Cytoplasmic membrane-bounded vesicle			3
		Pigment granule			3
		Melanosome			3
2	RFI	Glycosylation site: N-linked	1.56	Significant	6
		Extracellular region part			4
		Glycoprotein			8
		Signal peptide			11
		Extracellular region			9
		Disulfide bond			9
		Signal			9
3	DMI	mRNA metabolic process	1.2	Suggestive	4
		Nucleus			3
		RNA splicing/processing			3
		mRNA splicing/processing			3
4	pDMI	Intracellular signaling cascade	2.06	Significant	5
		Nuleoside-triphosphatase regulator activity			4
		Small GTPase regulator activity			4
		Regulation of Rho/Ras protein signal transduction			4
		Rho/Ras guanyl-nucleotide exchange factor activity			4

Acknowledgements

The authors would like to thank Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper. URL: http://www.tacc.utexas.edu. In addition, the GENSEL data analysis was utilized through the iPlant Collaborative website which is based upon work supported by the National Science Foundation under Award Numbers DBI-0735191 and DBI-1265383. URL: www.iplantcollaborative.org. Furthermore, some of the data computation for this project was performed at the OSU High Performance Computing Center at Oklahoma State University supported in part through the National Science Foundation grant OCI–1126330. Lastly, we would like to thank the following grant for access to the data: Agriculture and Food Research Initiative Competitive Grant no. 2011-68004-30214 from the USDA National Institute of Food and Agriculture (J.E. Beever, D.B. Faulkner, S.C. Fahrenkrug, H.L Neibergs, K.A. Johnson, C.M. Seabury, D.J. Garrick, D.D. Loy, S.L. Hansen, H.C. Freetly, M.L. Spangler, J.F. Taylor).

CHAPTER III

GENOME-WIDE ASSOCIATION ANALYSIS OF CARCASS TRAITS IN HEREFORD CATTLE USING ULTRA-HIGH DENSITY SNP DATA

Abstract

While ultrasound technology is useful for genetic evaluation of carcass traits, genomic technologies promise another avenue for genetic improvement. The objective of this study was to use ultra-high density single nucleotide polymorphism (SNP) data to develop direct genomic values (DGV) for carcass traits and identify genomic regions critical for predicting carcass merit in beef cattle. Carcass data was collected on a population of Hereford cattle (n=849) and animals were genotyped on the Illumina BovineHD Beadchip, which assays 770,000 SNPs within the genome. The DGVs were estimated and a genome-wide association study (GWAS) was performed using GenSel to identify regions of the genome important for prediction of genetic merit in carcass quality. The largest effect SNPs were examined to identify genes and biological pathways associated with carcass merit. The DGV realized accuracies for the carcass traits varied from 0.47 to 0.60 with heritabilities ranging from 0.44 to 0.72. A total of five QTL regions were independently validated within the literature. Many genes were identified within these regions that could be looked at further for possibly being included as candidate genes for some of the carcass phenotypes.

Introduction

Carcass traits play an important role in the breeding objectives of cattle producers and contribute to the consumer beef eating experience. The gold standard for genetic evaluation of carcass traits is individual-animal carcass data collected at harvest. However, ultrasound data can function as an indicator trait and thus contribute towards genetic evaluation because collection of actual carcass data can often be difficult when producers do not retain ownership of their calves. However, ultrasound data can vary from the actual data that is obtained at an abattoir because the ultrasound data is typically collected before the animal is fed a finishing diet (Drake, 2004).

Carcass traits generally have a moderate to high heritability which makes selection for carcass traits very effective. Expected progeny differences (EPD) have been the foundation for selection for carcass quality for decades. More recently, incorporation of genomic data has provided the opportunity to make even faster progress through the utilization of genomic-enhanced EPDs. Since carcass traits can only be measured post-mortem, the prediction of genetic merit for potential parents is critical for a successful breeding program. Several quantitative trait loci (QTL) studies have identified regions of the genome that are critical for carcass quality (Stone et al., 1999; Casas et al., 2000; Casas et al., 2001; Casas et al., 2003; Casas et al., 2004; Kim et al., 2003; Abe et al., 2008; Gutierrez-Gil et al., 2009; Morris et al., 2009; McClure et al., 2010; Esmailizadeh et al., 2011; Nalaila et al., 2012; Peters et al., 2012; Lu et al., 2013; Abo-Ismail et al., 2014).

The objective of this study is to utilize ultra-high density single nucleotide polymorphism (SNP) data to predict direct genomic values (DGV) on a large population of Hereford steers and identify and validate QTL regions for a variety of carcass traits. In addition, these QTL regions will be evaluated to identify new candidate genes that are linked to carcass traits and identify biological pathways related to the traits of interest.

Materials and Methods

Animals and phenotypic data

A total of 870 Hereford steers and heifers were fed a feedlot ration between 2009 and 2011. Our study population is a subset of animals from Saatchi et al. (2014). The animals were divided into 8 contemporary groups ranging in size from 31 to 205 animals and were formed based on calving season and year, number of days on feed, and sex. The number of days on feed was either 70, 71, or 140 days.

Phenotypic records for a variety of carcass traits were obtained, including hot carcass weight (HCW), marbling score (MARB), rib eye area (REA), back fat thickness (BFAT), and yield grade (YG). Summary statistics for phenotypic records are provided in Table 15.

Genotypic data

All of the animals (n=850) were genotyped using the Illumina Bovine HD BeadChip (770K) (Illumina Inc., San Diego, CA; Rincon et al., 2011). One animal was in a contemporary group consisting of only themselves so that animal was excluded from further analyses. Filtering of the dataset removed any data that was an extreme deviation from Hardy-Weinberg Equilibrium (HWE). In addition, the data was filtered for a minor allele frequency (MAF) < 0.01. The dataset was also filtered for call rate, marker heterozygosity (HETA) greater than 0.40, and incorrect chromosome assignment. Any genotypes that were missing in the dataset were imputed using BEAGLE (Browning and Browning, 2009). Upon the completion of data filtering, there were 648,625 SNPs available for analysis on 849 animals.

Data analysis

Direct Genomic Value (DGV) Predictions

The GENSEL application (Fernando and Garrick, 2009) implemented via the iPlant Collaborative website (Goff et al., 2011), was employed to derive DGV predictions using a

Bayesian framework. Because BayesC utilizes only a single common variance for each allele substitution effect, a BayesB analysis was used. While BayesB analyses are useful because of their ability to handle large effects SNPs better than a BayesC analysis since a BayesB does not regress towards the mean. BayesB are also sensitive to starting values for the variances as noted in Rolf et al. (2015). Due to this issue, BayesC analyses where $\pi = 0.95$ were run initially to estimate variance components that were then used in subsequent BayesB analyses. A list of previous literature estimates of genetic and residual variance components was created and the average of these variances were calculated, as seen in Table 16. The average variances were used as the starting values for all BayesC analyses. Analyses were run for 60,000 Markov chain Monte Carlo (MCMC) iterations with a burn-in of 1,000.

Since animals were placed into contemporary groups and the mean was not previously adjusted before the analysis, the model for the Bayes B0 analysis was:

$$y_i = \mu + Xb + \sum_{j=1}^{k} z_{ij}u_j + e_i$$

where:

y_i= feed efficiency phenotypes for each trait,

μ= sample mean,

X= matrix corresponding to fixed effects vector b,

b= vector of fixed effects for contemporary groups and phenotypic mean,

k= number of genetic markers in the analysis,

z_{ij}= allelic state (AA=-10, AB=0, BB=10) of animal i at marker j,

u_j= random effect for marker j, and

e_i= residual effect.

The means of the posterior distributions for the genetic and residual variance components obtained from the BayesC results, as seen in Table 17, were utilized for BayesB analyses where π

= 0. The BayesB analyses were utilized to generate DGVs using a five-fold cross validation. Training and validation populations were created to ensure that predictions could be made in a training population and the DGV accuracies could be calculated for each animal independently from the training group in a validation group. Training and validation populations were formed by randomly assigning animals within each contemporary group for each trait into one of five groups. This random assignment ensures that each contemporary group is equally represented across the populations. If there was not enough animals to evenly fill each contemporary group (groups not divisible by five), then these animals were grouped together and re-randomized to then be randomly allocated to one of the five groups.

A classic cross-validation approach was used, whereby four of the groups were used as the training population with the fifth group being used as the validation population, as seen in Table 18. Correlations were calculated between the DGV and the phenotype, because de-regressed breeding values were not available for analysis in this population. In addition, the realized accuracy for each trait was calculated as $\frac{r_{\hat{g},y}}{\sqrt{h^2}}$ which accounts for the fact that environment is included in the phenotype, but not the DGV (Rolf et al., 2015).

Genome-wide association study

In the process of estimating DGV, useful information regarding the distribution of SNP effects in the genome is obtained, which allows one to simultaneously perform a genome-wide association analysis. Windows were formed around the top 50 SNP regions by collapsing SNPs within the range of linkage disequilibrium (LD; +/250,000 base pairs (bp) of the largest-effect SNP within that region) to compare to previous literature estimates using the CattleQTLdb (release 26; Hu et al., 2013; Hu et al., 2010; Hu and Reecy, 2007; Hu et al., 2005). The CattleQTLdb allows for the comparison of SNP regions in this study to see if they validate QTL regions that have been found in other studies.

Biological network analysis

Gene annotation clustering analyses were completed for the top 50 largest-effect SNP regions. A 500 Kb window was formed centered around the top 50 SNP. The Ensembl browser (release 80; Cunningham et al., 2015) was utilized to obtain gene identifiers for coding regions within these windows. The list of resulting genes were imported into the Database for Annotation, Visualization, and Integrated Discovery (DAVID; version 6.7; Huang et al., 2009a; Huang et al., 2009b) for Functional Annotation Clustering, groups genes based on their involvement in biological pathways. The Functional Annotation Clustering tool makes it possible to identify genomic regions from association analyses that are critical in KEGG (Kyoto Encyclopedia of Genes and Genomes) biological pathways database.

Results and Discussion

Heritabilities

Heritabilities for each carcass trait are provided in Table 19. The trait with the highest heritability was MARB (0.72), which is within the bounds of previous literature estimates (0.31 to 0.74; Cundiff et al., 1971; Mao et al., 2013). The heritability for REA (0.52) agreed with previous literature estimates which range from 0.22 to 0.64 (Gregory et al., 1995; Mao et al., 2013). In addition, literature estimates of the heritability for HCW range from 0.23 to 0.56 (Gregory et al., 1995; Cundiff et al., 1971), which is consistent with our estimate of 0.47. The heritability estimate for YG (0.46) was slightly lower than previous literature estimates (0.58 and 0.71; Nkrumah et al., 2007b; Riley et al., 2002), but fewer estimates exist for YG within the literature. Backfat thickness had the lowest heritability (0.44), but was consistent with previous literature estimates (0.17 to 0.63; Mao et al., 2013; Riley et al., 2002).

Direct Genomic Values

The DGV correlations for all carcass traits are provided in Table 19 and were 0.34, 0.41, 0.43, 0.34, and 0.35 for BFAT, HCW, MARB, REA, and YG, respectively. The DGV correlations for all of the traits in this study were substantially lower than the study by Saatchi et al. (2011), with correlations ranging from 0.68 to 0.80 for these same traits. The correlations in this study are likely smaller than estimates from Saatchi (2011) because this study had substantially fewer animals than the study by Saatchi et al. (2011), which had 3668 animals. While these correlations are smaller than previous literature estimates, the carcass correlations are higher than feed efficiency correlations that have been reported in a study by Mujibi et al. (2011) which consisted of a similar population size (n=728), possibly because carcass traits have higher heritabilities than feed efficiency phenotypes. Calculated DGV realized accuracies for BFAT, HCW, MARB, REA, and YG were 0.51, 0.60, 0.51, 0.47, and 0.52, respectively, and are reported in Table 19. A study by Saatchi et al. (2013b), which also looked at Hereford cattle but had a larger population size (n=1081), estimated DGV accuracies that were slightly lower for BFAT, MARB, and REA. In addition, another study estimated DGV accuracies that were both higher and lower than some of the traits in this study with accuracies for BFAT, HCW, MARB, and REA of 0.603, 0.471, 0.690, and 0.60, respectively (Saatchi et al., 2011). Comparing the results of this study with Saatchi et al. (2011), BFAT, MARB, and REA were higher in the other study, but HCW was higher in our study. The reason the accuracies in Saatchi et al. (2011) were more similar to this study compared to our very low correlations compared to their study is due to the differences in heritability between the studies, and they used deregressed breeding values instead of phenotype values like our study. Differences in correlations and DGV accuracies can be seen between our study and other studies based on the number of animals in the population, the heritability values, and whether deregressed breeding values were available for the analysis.

Association analyses

Manhattan plots for each phenotype are provided in Figures 16, 17, 18, 19, and 20 for BFAT, HCW, MARB, REA, and YG, respectively. Many of the phenotypes had SNPs with large peaks, indicating large effect SNPs. One trait which had a large effect SNP was REA which had an effect size of 0.30 that was located on chromosome 7 at 93,172,661 to 93,221,590 bp. There were no bovine genes annotated within this region, but according to the UCSC Genome Browser there was a human gene (LUCAT1) related to lung cancer and mouse long non-coding RNA found in this region. In addition, there were large effect SNPs for REA on chromosome X (3,500,270 to 3,520,961 bp) and chromosome 17 (64,626,684 to 64,639,054 bp), but neither of these regions had genes within these areas for bovine. The region on chromosome X that was identified had no bovine genes identified in this region, but in mice there is a gene in this region related to a kinase anchor protein (AKAP17B) which is involved in functions related to cyclic compound binding (NCBI BioSystems Database; Geer et al., 2010). In addition, the region on chromosome 17 was associated with transforming growth factor beta-1 (TGFB1I1) in mice which is related to the regulation of actin cytoskeleton (NCBI BioSystems Database; Geer et al., 2010). Since this was a region for REA and the gene has known functions related to tissue development, TGFB1I1 is a gene that could possibly explain variation of REA in cattle. A SNP on chromosome 29 at position 49,350,297 to 49,377,731 bp had an effect of 0.001466 and is within the coding region for cyclin-dependent kinase inhibitor 1C (CDKN1C) and solute carrier family 22 (SLC22A18) in bovine. CDKN1C has functions related to endochondral ossification, or bone formation, in bovine. The phenotype with the next highest SNP effect (0.18) was HCW, located on chromosome 7 at 93,222,309 to 93,229,001 bp. While there were no genes identified in this region, upstream at position 93,240,416 to 93,253,094 bp is the bovine gene arrestin domain containing 3 (ARRDC3) which has functions related to metal ion binding. Other large effect SNPs for HCW were located on chromosomes 8 (2,937,703 to 2,987,221 bp) and chromosome X (133,326,608 to 133,331,145 bp). There were no genes in this region on chromosome 8, but

upstream of this region there was a region at 3,816,704 to 5,330,615 bp for the bovine gene polypeptide N-acetylgalactosaminyltransferase-like 6 (GALNTL6), which has functions related to macromolecule glycosylation (NCBI BioSystems Database; Geer et al., 2010). The previously mentioned region on chromosome X has no known associated genes in this region for bovine, but the gene NHS for Nance-Horan syndrome, which causes cataracts and dental problems, has been found in this region for humans, mice, and rats. One of the highest effect SNPs for BFAT was on chromosomes 8 (6475070 to 6488197 bp), which has an association with the centrosomal protein-44 (CEP44) gene in humans which is responsible for microtubule organization in cells (NCBI BioSystems Database; Geer et al., 2010). In addition, another region for BFAT was on chromosome 11 (4,511,327 to 4,520,723 bp), and this region has an association to REV1 in humans, mice, and rats. According to the NCBI Reference Sequence Database (RefSeq Release 70; Tatusova et al., 2014), REV1 is a protein that contains a BRCT domain and acts as a recruiter for DNA polymerases. A large effect region for BFAT at chromosome 19 (36,610,816 to 36,621,137 bp) is in a gene coding region for ankyrin repeat domain 40 (ANKRD40) in bovine which has functions related to protein binding. In addition, there was also a large effect SNP for YG on chromosome 8 (7,437,831 to 7,449,420 bp) which had a gene in this region for farnesyl-diphosphate farnesyltransferase-1 (FDFT1). This gene codes for a protein which is responsible for steroid biosynthesis (NCBI BioSystems Database; Geer et al., 2010). Since steroids play a critical role in the body for a variety of different processes, such as bovine growth hormone (GH) and its effect on muscle anabolism (Liu et al., 2013), this gene could be a critical gene in the variation that is seen in the YG phenotype. For example, a study by discovered that cattle implanted with the steroids trenbolone acetate and estradiol had Lastly, the largest effect SNP for MARB was 0.035 on chromosome 11 at position 47,863,101 to 47,899,487 bp. This SNP is in a genome region that codes for an uncharacterized protein and lysine-rich coiled-coil 1 (KRCC1), which has functions related to the regulation of mitochondrial degradation (NCBI BioSystems

Database; Geer et al., 2010), which could be critical to fat content due to the mitochondria's function as an energy supplier for the cell.

The top 50 largest-effect 500 Kbp regions were compared to QTLs identified in previous studies, which are summarized in Table 20. An association study completed by Kim et al. (2013) identified the same QTL as this study for BFAT on chromosome 1. The QTL region that was validated had several genes within it, such as U6 spliceosomal RNA (U6), ATP synthase H+ transporting, mitochondrial F1 complex (ATP5O), transmembrane protein 50B (TMEM50B), and DNAJ HSP40 homolog (DNAJC28). ATP5O has functions related to oxidative phosphorylation and in metabolism in general (NCBI BioSystems Database; Geer et al., 2010). Because of ATP5O's functions, this could be a plausible gene for the variation that is seen since metabolism is a critical aspect in weight loss and weight gain. The genes U6, TMEM50B, and DNAJC28 generally have intracellular functions within the cytoplasm (NCBI BioSystems Database; Geer et al., 2010), and are not obvious candidates for having direct impact on BFAT. In addition, there were multiple regions on chromosome 7 and 8 for HCW that were validated utilizing previous studies. These QTL IDs from the database (24625, 24626, 24627, and 10817) are located where the previously mentioned large peaks were for HCW on chromosomes 7 and 8, around 93.0 to 93.9 Mbp and 1.13 to 4.78 Mbp, respectively. There are several genes within the QTL region on chromosome 8 including chloride channel 3 (CLCN3), GrpE-like 2, and acetylgalactosaminyltransferase-like 6 (GALNTL6). These genes have functions related to transmembrane transport, mitochondrial protein import, and biosynthesis of the N-glycan precursor, respectively (NCBI BioSystems Database; Geer et al., 2010). In addition, an association study completed by Tong et al. (2014) identified a QTL on chromosome 5 for MARB that was validated in this study, and a bovine gene for solute carrier family 38 (SLC38A2) is in this region. This gene codes for proteins with functions related to endochondral ossification, similar to the CDKN1C discussed previously. Lastly, the largest peak on chromosome 7 for REA was consistent with QTL identified by Saatchi et al. (2014). Although it is important to note that

the study by Saatchi et al. (2014) used the same dataset as this study, and this study was not an independent validation. Bovine genes related to microRNA-2464 (MIR2464) and arrestin domain containing 3 (ARRDC3) reside within this region of the genome, which has functions related to heat generation and biological regulation (NCBI BioSystems Database; Geer et al., 2010).

Biological network analyses

Functional annotation clusters obtained from DAVID were considered significant if they had an enrichment score of 1.3 and were considered suggestive if the enrichment score was 1.0 (Huang et al., 2009b). Enriched clusters are summarized in Table 21. One of the significant clusters was for YG, with an enrichment score of 1.82. This cluster consisted of genes, such as cathepsin B, that are critical to peptidase activity. Another significant cluster for YG had genes with links to cation binding and zinc ion binding, which involves the interaction and binding of positively charged atoms. The highest scoring cluster was for MARB, which had a score of 3.26 and contained fibroblast growth factor receptor binding and interleukin-1 receptor antagonist activity. The genes that were identified in this cluster for MARB were two different members of the interleukin-1 family which is very closely associated to the innate immune response (Dinarello, 2009). DAVID also has the ability to identify genes that are in biological pathways on the KEGG Pathway Database (Kanehisa and Goto, 2000; Kanehisa et al., 2014). A pathway identified for MARB was primary immunodeficiency, seen in Figure 21, and antigen processing and presentation, as seen in Figure 22. A pathway is considered enriched if it has a P-value less than 0.05, and is suggestive of enrichment if $P < 0.10$. The pathway for primary immunodeficiency was significantly enriched ($P < 0.044$) while the antigen processing and presentation was only suggestive ($P < 0.091$). The red stars in Figures 21 and 22 indicate genes in pathways that were within the top 50 largest effect 500 Kbp regions. Furthermore, there was a suggestive cluster with a score of 1.09 for REA that had genes related to ion channel activity and ion transport. Ion transport is a function that could be critical to REA because of the importance

of calcium transport in muscle tissue.

Conclusion

Realized accuracies for a variety of carcass quality traits (BFAT, HCW, MARB, REA, and YG) were estimated and ranged from 0.47 to 0.60 and correlations ranged from 0.34 to 0.43. A GWAS analysis was also performed and regions of the genome important for predicting genetic merit for carcass traits were identified. In addition, this information was used to identify biological pathways that are associated with these genome regions and identify candidate genes that may harbor causal mutations. Additionally, five QTL regions identified in this study were validated utilizing results from the scientific literature. The regions identified in this study should be examined further utilizing imputed full sequence data to determine if any harbor causal mutations for these phenotypes, which should dramatically increase power of DGV predictions for these traits.

Supplementary Materials

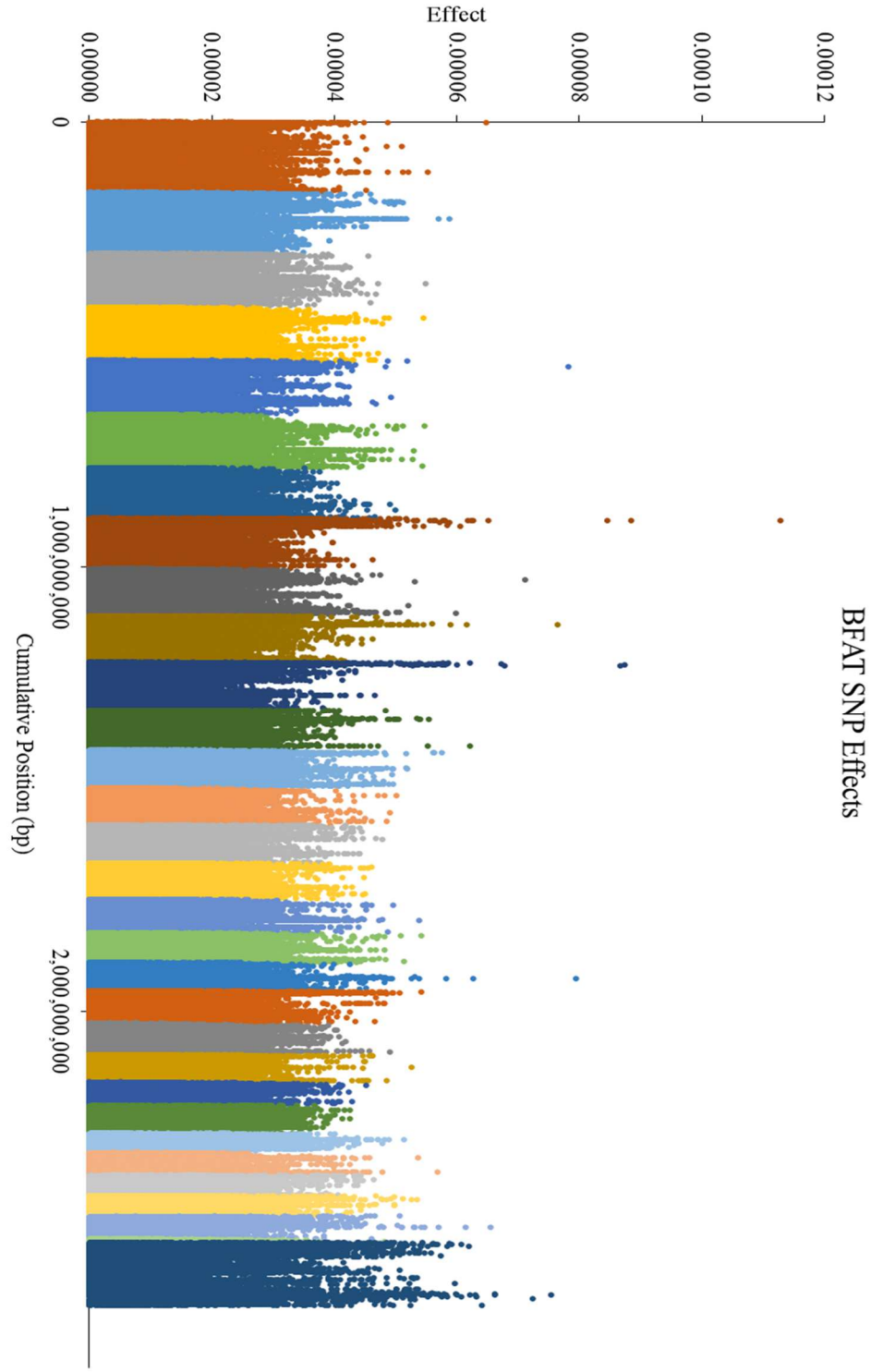

Figure 16. Manhattan plot of SNP effects for backfat thickness from GenSel analysis in Hereford cattle. Each color represents a chronological chromosome.

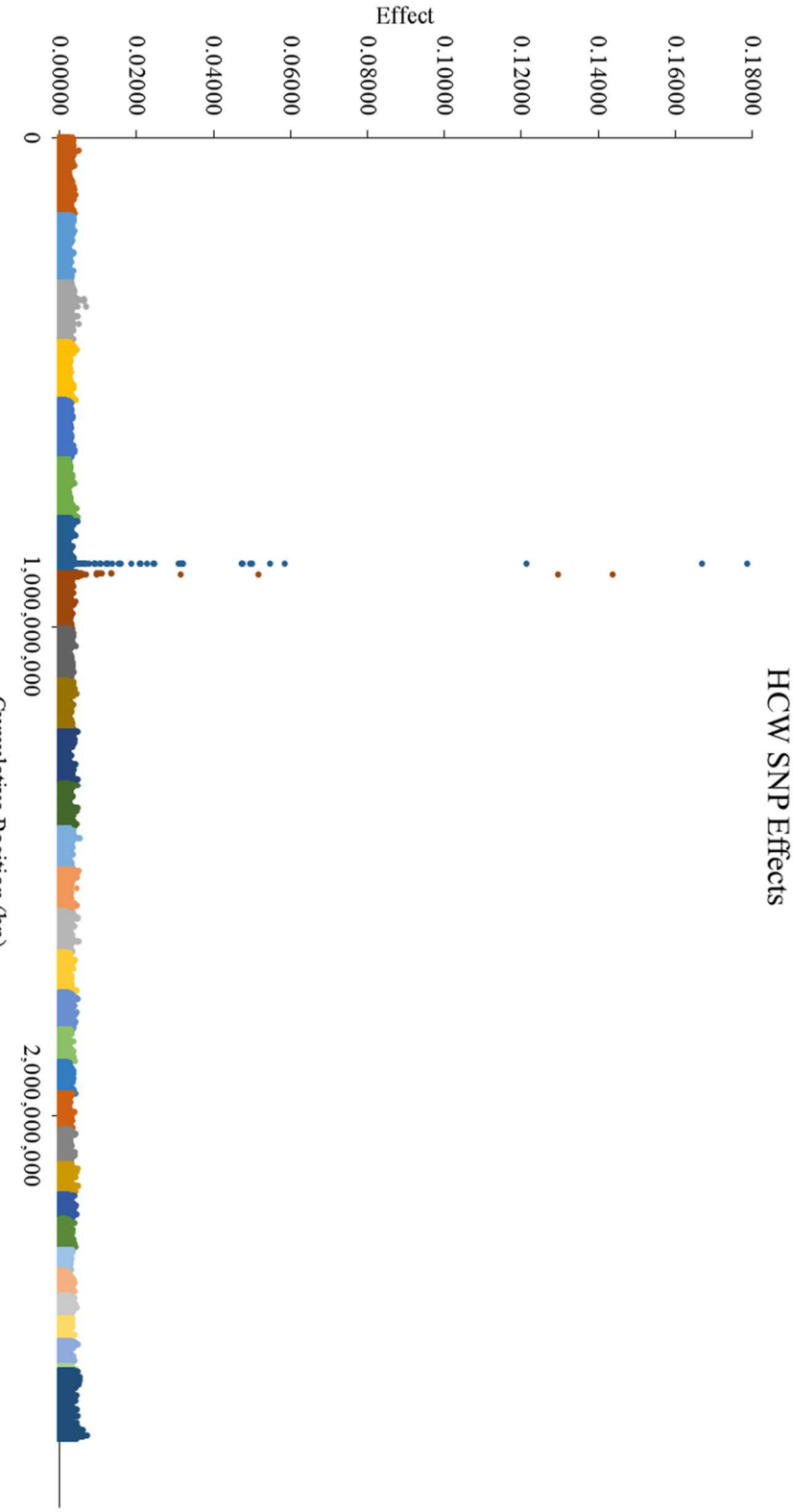

Figure 17. Manhattan plot of SNP effects for hot carcass weight from GenSel analysis in Hereford cattle. Each color represents a chronological chromosome.

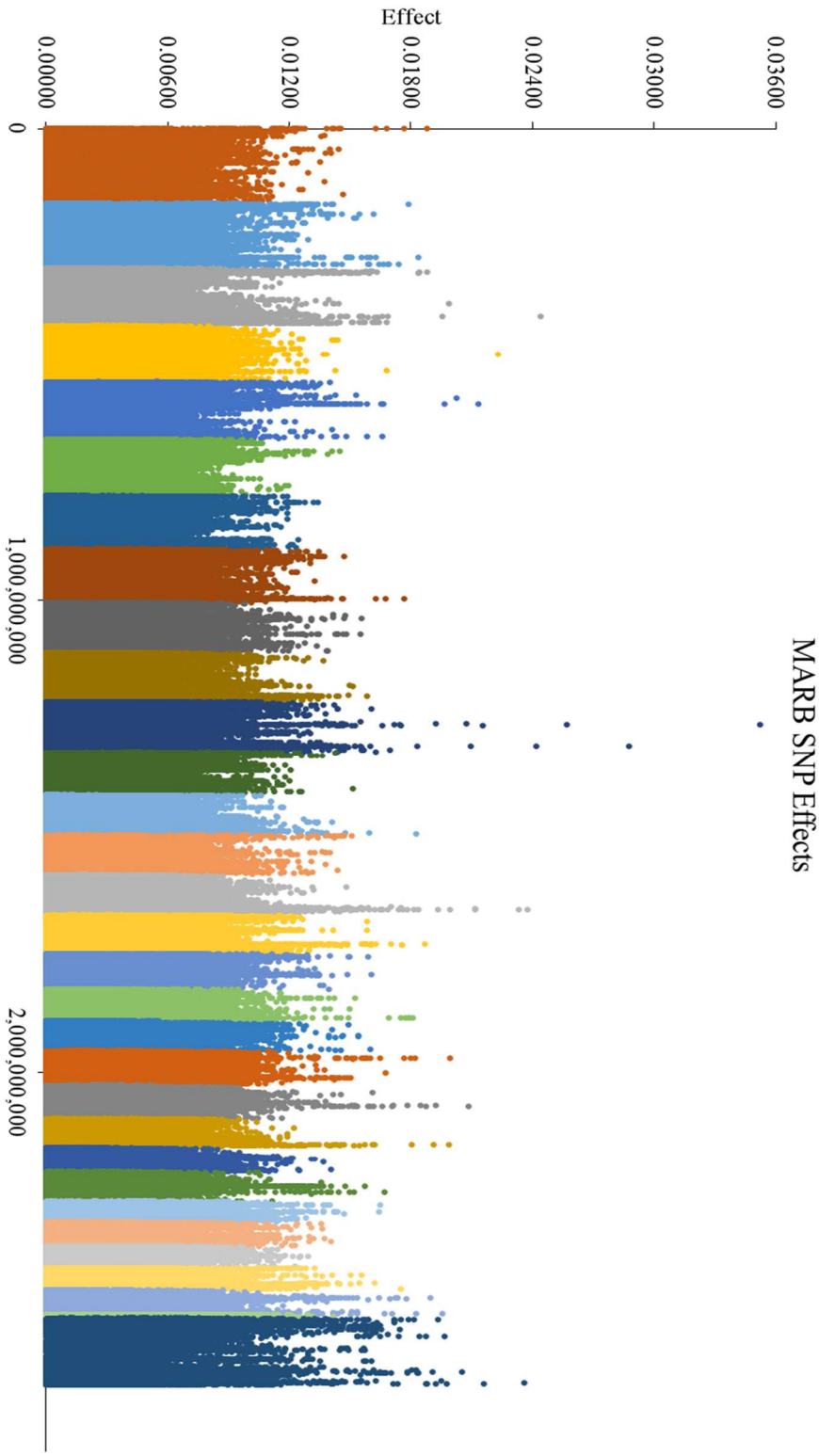

Figure 18. Manhattan plot of SNP effects for marbling score from GenSel analysis in Hereford cattle. Each color represents a chronological chromosome.

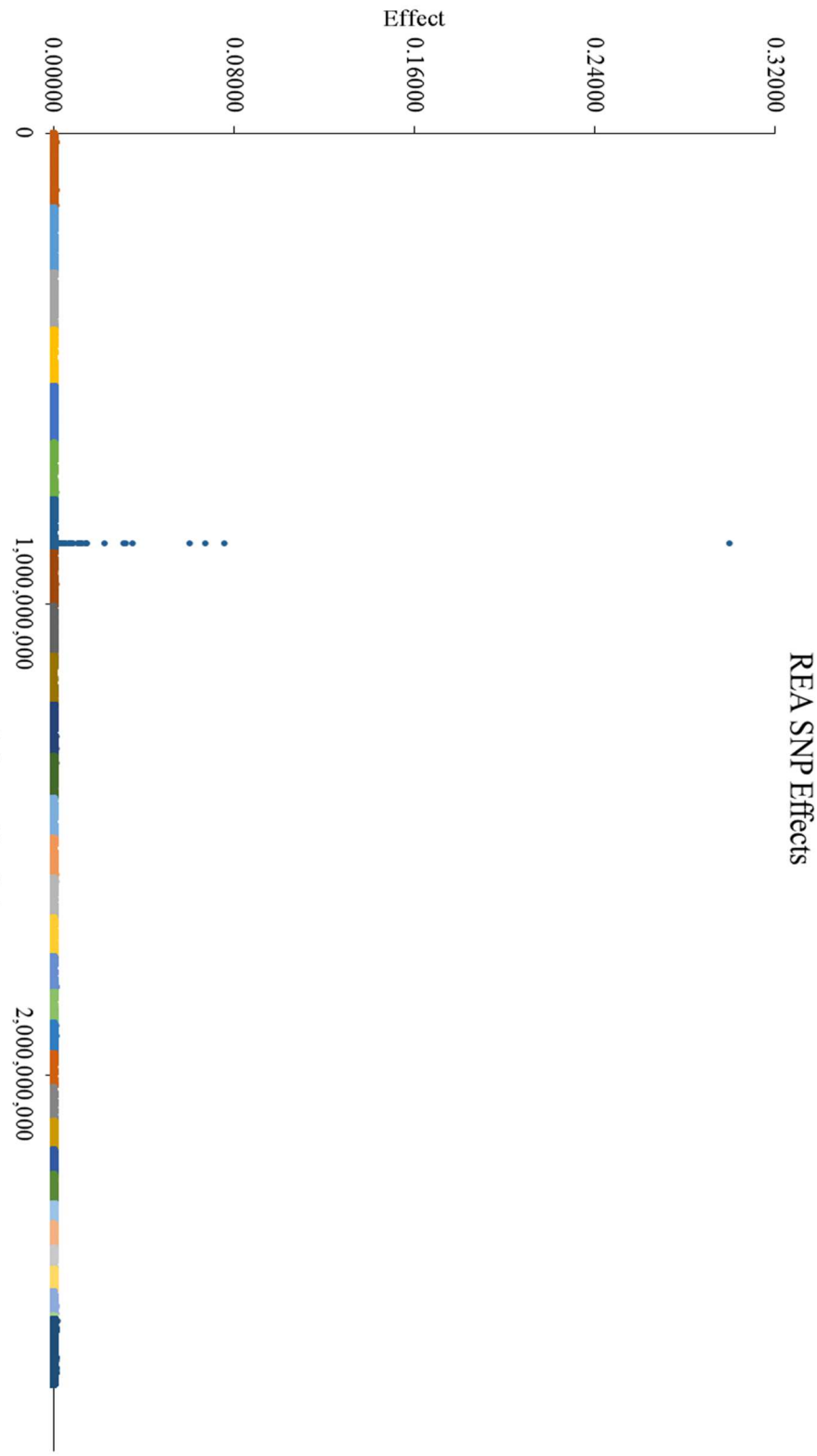

Figure 19. Manhattan plot of SNP effects for rib eye area from GenSel analysis in Hereford cattle. Each color represents a chronological chromosome.

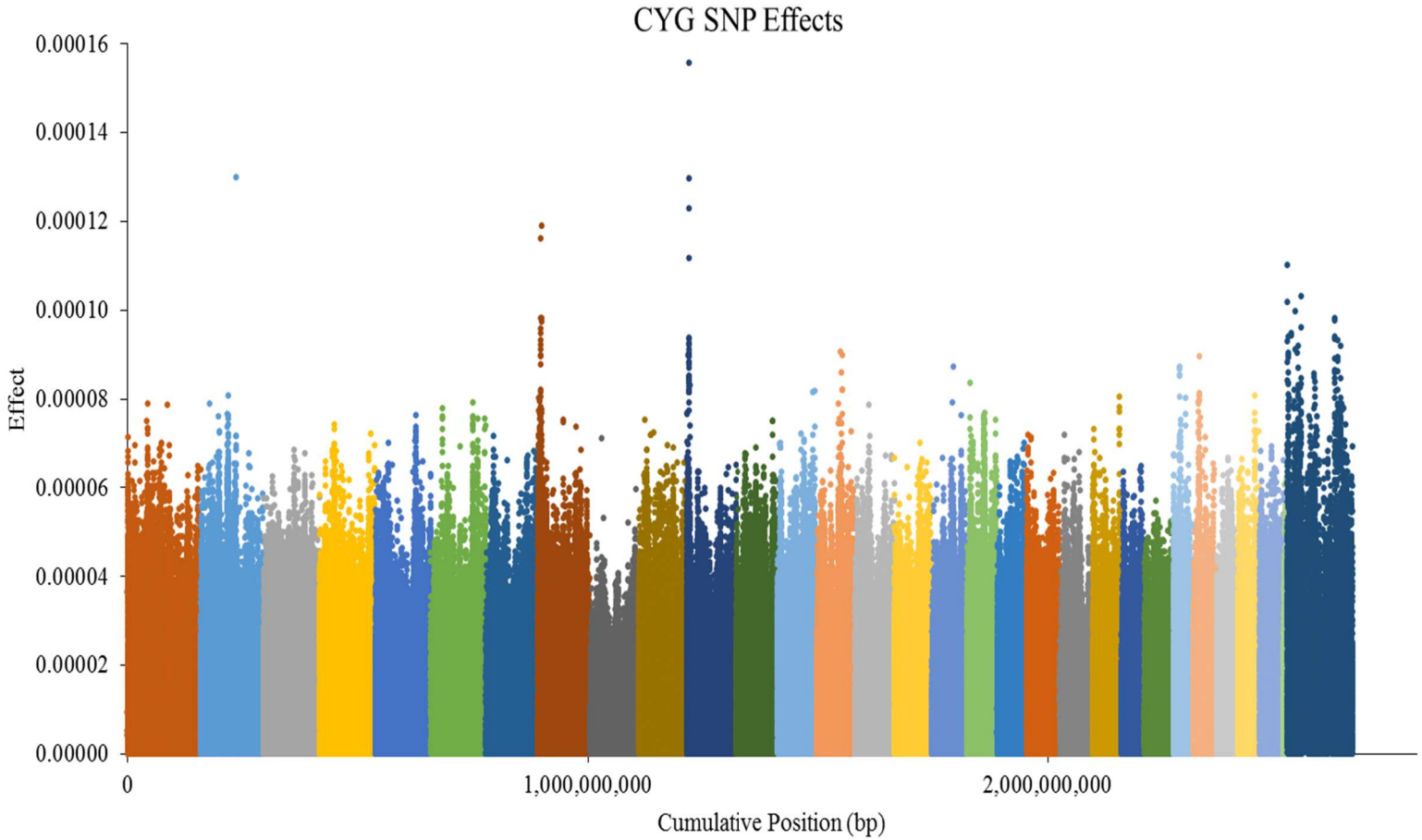

Figure 20. Manhattan plot of SNP effects for calculated yield grade from GenSel analysis in Hereford cattle. Each color represents a chronological chromosome.

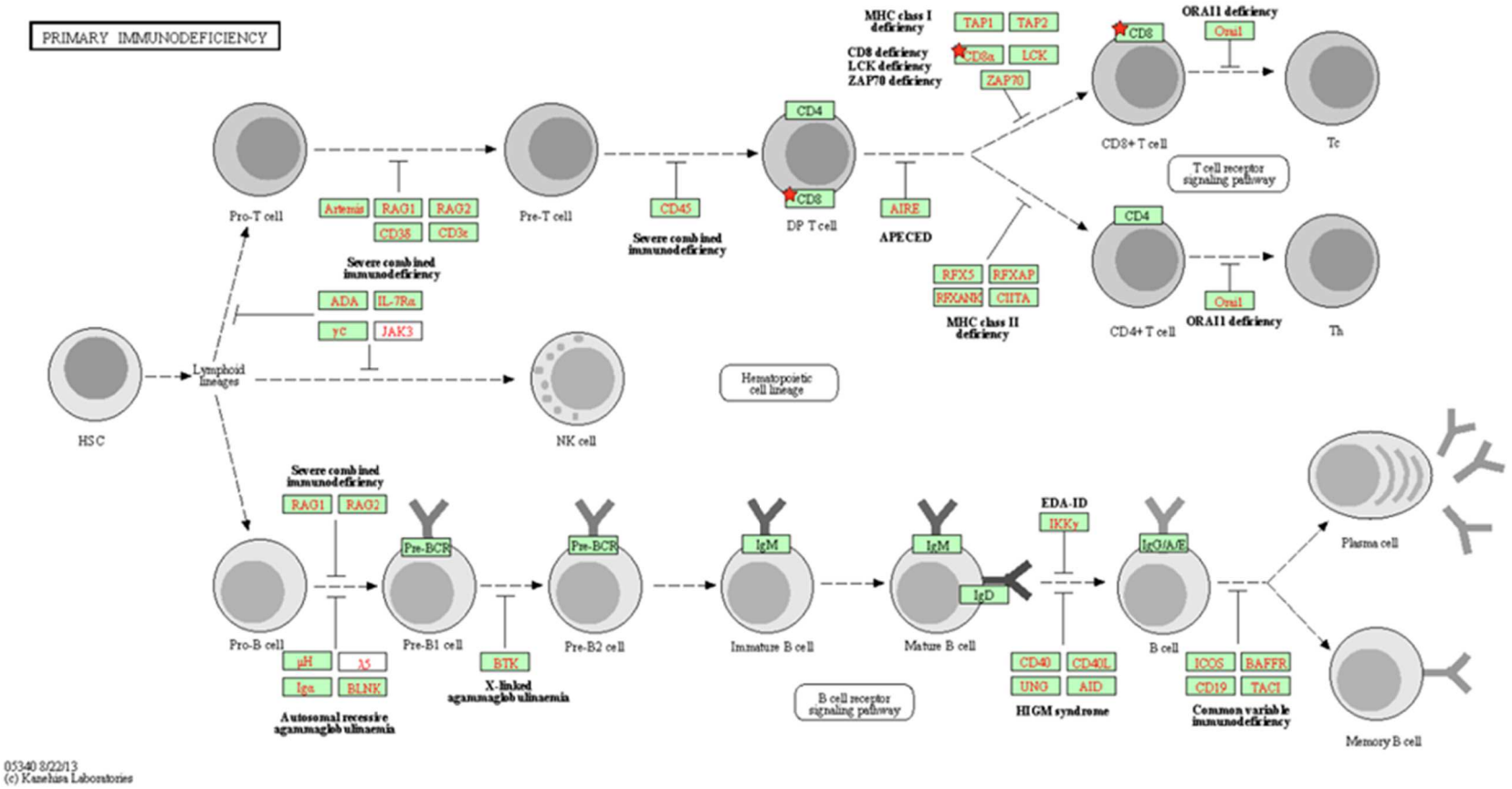

Figure 21. KEGG pathway for primary immunodeficiency from DAVID clustering analysis for marbling in Hereford cattle. Red stars indicate the genes in our study that were identified in this pathway.

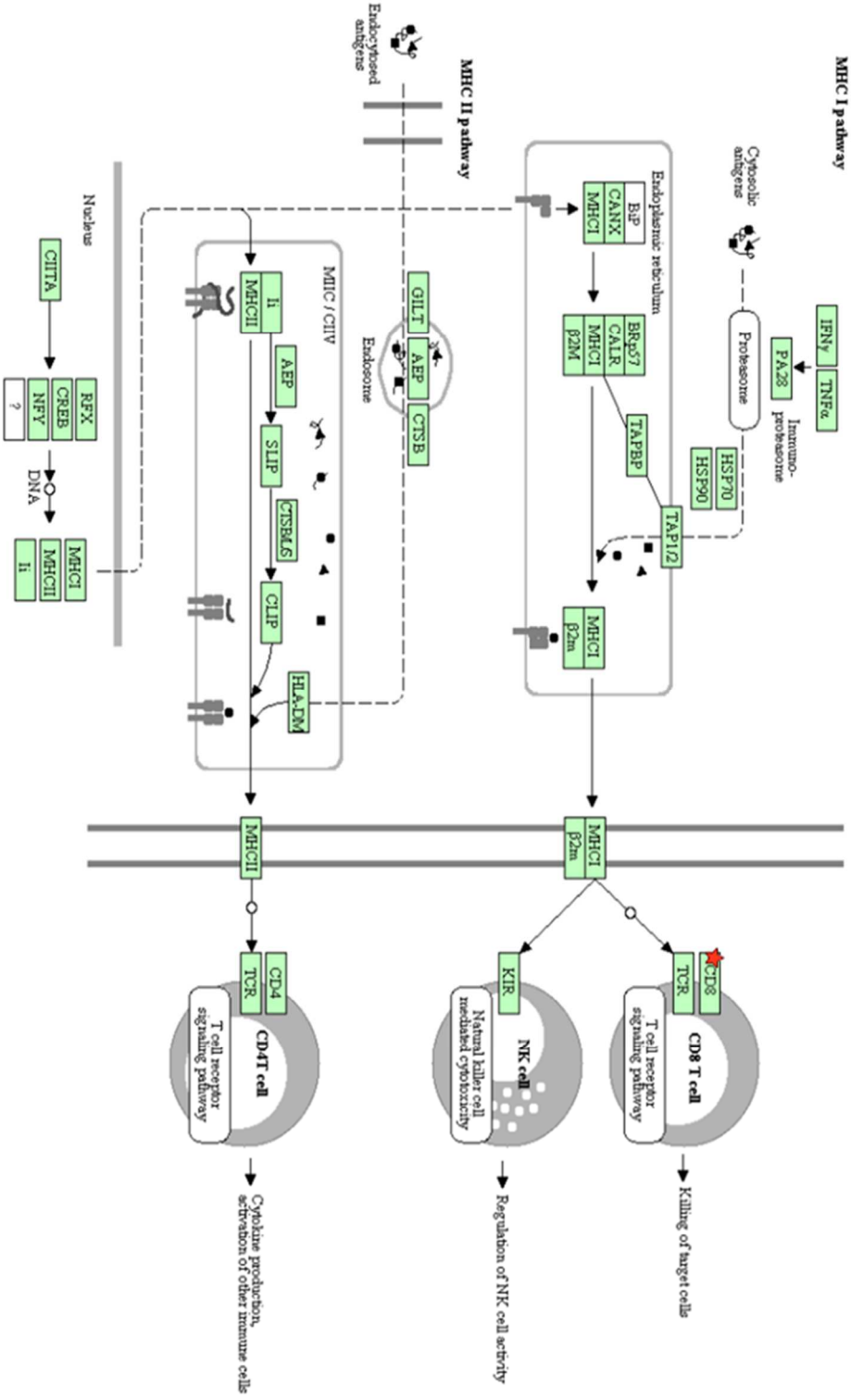

Figure 22. KEGG pathway for antigen processing and presentation from DAVID clustering analysis for marbling in Hereford cattle. Red stars indicate the genes in our study that were identified in this pathway.

Table 15- Data for hot carcass weight (HCW), marbling score (MARB), rib eye area (REA), backfat thickness (BFAT), and yield grade (YG) collected on Hereford cattle fed for 70, 71, and 140 days and number of records, minimum, maximum, mean, and standard deviation

Phenotype	**HCW**	**MARB**	**REA**	**BFAT**	**YG**
Number of Records	849	849	840	836	775
Minimum	238	410	48.4	0	1.85
Maximum	498	870	109.4	3	5.81
Mean	407.1	562.4	85.8	1.51	3.72
Standard Deviation	36.00	75.24	8.51	0.45	0.64

Table 16- Previous literature estimates of residual (σ^2_e) and genetic (σ^2_a) variance components and the average of each used in the BayesC95 analysis for the GenSel analysis*

Phenotype	Source 1	Source 2	Source 3	Average
BFAT (mm)	σ^2_e= 4.14 σ^2_a=3.06 Robinson & Oddy (2004)	σ^2_e= 4.09 σ^2_a=3.49 Crews & Kemp (2001)	σ^2_e=5.98 σ^2_a= 2.32 Davis & Simmen (2000)	σ^2_e=4.73 σ^2_a= 2.96
HCW (kg)	σ^2_e= 445 σ^2_a= 273 Crews & Kemp (2001)	σ^2_e= 538 σ^2_a= 253 Davis & Simmen (2000)	σ^2_e= 385 σ^2_a= 385 Splan et al. (1998)	σ^2_e= 456 σ^2_a= 304
MARB	σ^2_e= 1556 σ^2_a= 1207 Riley et al. (2002)	σ^2_e= 1324 σ^2_a= 1755 Wheeler et al. (2001)	----	σ^2_e= 1440 σ^2_a= 1481
REA (cm^2)	σ^2_e= 23.28 σ^2_a= 8.61 Arthur et al. (2001a)	σ^2_e= 16.71 σ^2_a= 31.58 Wheeler et al. (2001)	σ^2_e= 21.43 σ^2_a= 16.82 Riley et al. (2002)	σ^2_e= 20.47 σ^2_a= 19.00
YG	σ^2_e= 0.144 σ^2_a= 0.096 Davis & Simmen (2000)	σ^2_e= 0.080 σ^2_a= 0.197 Riley et al. (2002)	σ^2_e= 0.045 σ^2_a= 0.303 Wheeler et al. (2001)	σ^2_e= 0.090 σ^2_a= 0.199

*HCW- hot carcass weight; MARB- marbling score; REA- rib eye area; BFAT- backfat thickness; YG- yield grade

Table 17- Residual (σ^2_e) and genetic (σ^2_a) variance starting values for BayesB0 analysis in the GenSel analysis*

Trait	σ^2_e	σ^2_a
BFAT (mm)	0.165	0.078
HCW (kg)	704	546
MARB	1030	4228
REA (cm^2)	29.6	30.0
YG	0.212	0.180

*HCW- hot carcass weight; MARB- marbling score; REA- rib eye area; BFAT- backfat thickness; YG- yield grade

Table 18- Structure for analyses using randomized training and validation populations with DGV correlation and realized accuracy ranges from the analyses

Analysis #	**Randomly Assigned Populations**					**Accuracy**
1	Validation	Training	Training	Training	Training	0.35-0.73
2	Training	Validation	Training	Training	Training	0.33-0.61
3	Training	Training	Validation	Training	Training	0.17-0.64
4	Training	Training	Training	Validation	Training	0.38-0.68
5	Training	Training	Training	Training	Validation	0.54-0.63
Correlation	0.25-0.50	0.22-0.45	0.12-0.38	0.26-0.45	0.36-0.49	

Table 19- Heritabilities, correlations and accuracies, and final residual (σ^2_e) and genetic (σ^2_a) variance values from BayesB0 analysis for hot carcass weight (HCW), marbling score (MARB), rib eye area (REA), backfat thickness (BFAT), and yield grade (YG)

	BFAT	**HCW**	**MARB**	**REA**	**YG**
Heritability	0.44	0.47	0.72	0.52	0.46
Average Correlation	0.34	0.41	0.43	0.34	0.35
Average Accuracy	0.51	0.60	0.51	0.47	0.52
σ^2_e	0.0989	666.5	1518.5	29.0	0.213
σ^2_a	0.0783	602.7	3820.8	31.0	0.181

Table 20- Carcass QTLs externally validated utilizing previously published studies*

Trait	QTL ID	Chr	Position (Mbp)	Association or Linkage	Validation Source
BFAT	1317	1	0.50-1.80	Association	Kim et al. (2013)
HCW	24625-24627	7	93.0-93.9	Association	Saatchi et al. (2014)[1]
	10817	8	1.13-4.78	Linkage	McClure et al. (2010)
MARB	28698	5	34.2-34.4	Association	Tong et al. (2014)
REA	24699-24702	7	93.0-93.9	Association	Saatchi et al. (2014)[1]

*HCW- hot carcass weight; MARB- marbling score; REA- rib eye area; BFAT- backfat thickness; YG- yield grade; Chr- chromosome number

[1] This study used the same dataset as this study, but this study was not an independent validation.

Table 21- Functional annotation clustering results from DAVID for hot carcass weight (HCW), marbling score (MARB), rib eye area (REA), backfat thickness (BFAT), and yield grade (YG)

Cluster #	Trait	Function	Enrichment score	Significant/ Suggestive	# of genes included
1	MARB	Fibroblast growth factor receptor binding	3.26	Significant	3
		Interleukin-1 receptor antagonist activity			3
		Receptor inhibitor/regulator activity			3
		Extracellular region part			3
2	MARB	Amino acid transport	2.4	Significant	3
		Organic acid transport			4
3	YG	Metalloendopeptidase activity	1.82	Significant	3
		Proteolysis			7
4	YG	Metalloendopeptidase activity	1.57	Significant	7
		Metal ion bonding			4
		Cation binding			3
5	HCW	Voltage-gated channel activity	1.61	Significant	3
		Ion channel activity			3
		Ion transport			3
6	REA	Ion channel activity	1.09	Suggestive	3
		Passive transmembrane transporter activity			3
		Ion transport			4

Acknowledgements

The authors would like to thank the iPlant Collaborative team as the GENSEL data analysis was utilized through the iPlant Collaborative website which is based upon work supported by the National Science Foundation under Award Numbers DBI-0735191 and DBI-1265383. URL: www.iplantcollaborative.org. Furthermore, some of the data computation for this project was performed at the OSU High Performance Computing Center at Oklahoma State University supported in part through the National Science Foundation grant OCI–1126330. Lastly, we would like to thank the following grant for access to the data: Agriculture and Food Research Initiative Competitive Grant no. 2011-68004-30214 from the USDA National Institute of Food and Agriculture (J.E. Beever, D.B. Faulkner, S.C. Fahrenkrug, H.L Neibergs, K.A. Johnson, C.M. Seabury, D.J. Garrick, D.D. Loy, S.L. Hansen, H.C. Freetly, M.L. Spangler, J.F. Taylor).

REFERENCES

Abe, T., J. Saburi, H. Hasebe, T. Nakagawa, T. Kawamura, K. Saito, T. Nade, S. Misumi, T. Okumura, and K. Kuchida. 2008. Bovine quantitative trait loci analysis for growth, carcass, and meat quality traits in an F population from a cross between Japanese Black and Limousin. J Anim Sci., 86: 2821-2832.

Abo-Ismail, M. K., G. Vander Voort, J. J. Squires, K. C. Swanson, I. B. Mandell, X. P. Liao, P. Stothard, S. Moore, G. Plastow, and S. P. Miller. 2014. Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. BMC Genetics, 15.

Anderson, R., R. Rasby, T. Klopfenstein, and R. Clark. 2005. An evaluation of production and economic efficiency of two beef systems from calving to slaughter. J Anim Sci., 83: 694-704.

Archer, J., E. Richardson, R. Herd, and P. Arthur. 1999. Potential for selection to improve efficiency of feed use in beef cattle: a review. Austr J Agric Res., 50: 147-162.

Archer, J., A. Reverter, R. Herd, D. Johnston, and P. Arthur. 2002. Genetic variation in feed intake and efficiency of mature beef cows and relationships with postweaning measurements. In: World Congress on Genet App to Livest Prod. p 221-224.

Arthur, P. F., J. A. Archer, D. J. Johnston, R. M. Herd, E. C. Richardson, and P. F. Parnell. 2001a. Genetic and phenotypic variance and covariance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. J Anim Sci.,, 79: 2805-2811.

Arthur, P. F., G. Renand, and D. Krauss. 2001b. Genetic and phenotypic relationships among different measures of growth and feed efficiency in young Charolais bulls. Livest Prod Sci., 68: 131-139.

Barendse, W., A. Reverter, R. Bunch, B. Harrison, W. Barris, and M. Thomas. 2007. A validated whole-genome association study of efficient food conversion in cattle. Genetics, 176: 1893-1905.

Benyshek, L. L. 1981. Heritabilities for Growth and Carcass Traits Estimated from Data on Herefords under Commercial Conditions. J Anim Sci., 53: 49-56.

Berry, D. P., and J. J. Crowley. 2012. Residual intake and body weight gain: a new measure of efficiency in growing cattle. J Anim Sci., 90: 109-115.

Bhuiyan, M. S. A., N. K. Kim, Y. M. Cho, D. Yoon, K. S. Kim, J. T. Jeon, and J. H. Lee. 2009. Identification of SNPs in MYOD gene family and their associations with carcass traits in cattle. Livest Sci., 126: 292-297.

Bolormaa, S., B. J. Hayes, K. Savin, R. Hawken, W. Barendse, P. F. Arthur, R. M. Herd, and M. E. Goddard. 2011. Genome-wide association studies for feedlot and growth traits in cattle. J Anim Sci., 89: 1684-1697.

Bouquet, A., M. N. Fouilloux, G. Renand, and F. Phocas. 2010. Genetic parameters for growth, muscularity, feed efficiency and carcass traits of young beef bulls. Livest Sci., 129: 38-48.

Brown S.D., and M.W. Moore. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. Mamm Genome. 2012 Oct;23(9-10):632-640.

Browning, B., and S. Browning. 2009. A unified approach to genotype imputation and haplotype phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet., 84: 210-223.

Carstens, G. E., and L. O. Tedeschi. 2006. Defining feed efficiency in beef cattle. In: Proc Beef Impr Fed 38th Ann Res Symp and Ann Meet., Apr 18-21, Choctaw, MS.

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

Casas, E., R. Stone, J. Keele, S. Shackelford, S. Kappes, and M. Koohmaraie. 2001. A comprehensive search for quantitative trait loci affecting growth and carcass composition of cattle segregating alternative forms of the myostatin gene. J Anim Sci., 79: 854-860.

Casas, E., S. D. Shackelford, J. W. Keele, M. Koohmaraie, T. P. L. Smith, and R. T. Stone. 2003. Detection of quantitative trait loci for growth and carcass composition in cattle12. J. Anim. Sci., 81: 2976-2983.

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

Ciechanover, A., A. Orian, and A. L. Schwartz. 2000. Ubiquitin-mediated proteolysis: biological regulation via destruction. Bioessays, 22: 442-451.

Crews, D. H., Jr., and R. A. Kemp. 2001. Genetic parameters for ultrasound and carcass measures of yield and quality among replacement and slaughter beef cattle. J Anim Sci., 79: 3008-3020.

Crowley, J., M. McGee, D. Kenny, D. Crews, R. Evans, and D. Berry. 2010. Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance-tested beef bulls. J Anim Sci., 88: 885-894.

Crouse, J. D., L. V. Cundiff, R. M. Koch, M. Koohmaraie, and S. C. Seideman. 1989. Comparisons of Bos Indicus and Bos Taurus Inheritance for Carcass Beef Characteristics and Meat Palatability. J Anim Sci., 67: 2661-2668.

Cundiff, L. V., K. E. Gregory, R. M. Koch, and G. E. Dickerson. 1971. Genetic relationships among growth and carcass traits of beef cattle. J Anim Sci., 33: 550-555.

Cunningham, F., M. R. Amode, D. Barrell, K. Beal, K. Billis, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fitzgerald, L. Gil, C. G. Girón, L. Gordon, T. Hourlier, S. E.

Hunt, S. H. Janacek, N. Johnson, T. Juettemann, A. K. Kähäri, S. Keenan, F. J. Martin, T. Maurel, W. McLaren, D. N. Murphy, R. Nag, B. Overduin, A. Parker, M. Patricio, E. Perry, M. Pignatelli, H. S. Riat, D. Sheppard, K. Taylor, A. Thormann, A. Vullo, S. P. Wilder, A. Zadissa, B. L. Aken, E. Birney, J. Harrow, R. Kinsella, M. Muffato, M. Ruffier, S. M. J. Searle, G. Spudich, S. J. Trevanion, A. Yates, D. R. Zerbino, and P. Flicek. 2015. Ensembl 2015. Nucleic Acids Res., 43: D662-D669.

Davis, M. E., and R. C. Simmen. 2000. Genetic parameter estimates for serum insulin-like growth factor-I concentration and carcass traits in Angus beef cattle. J Anim Sci., 78: 2305-2313.

Dinarello, C. A. 2009. Immunological and Inflammatory Functions of the Interleukin-1 Family. Annu Rev Immunol., 27: 519-550.

Drake, D. 2004. Understanding and improving beef cattle carcass quality. UCANR Publications, Publication 8130.

Eriksson, S., A. Näsholm, K. Johansson, and J. Philipsson. 2003. Genetic analyses of field-recorded growth and carcass traits for Swedish beef cattle. Livest Prod Sci., 84: 53-62.

Esmailizadeh, A. K., C. A. Morris, N. G. Cullen, Z. A. Kruk, D. S. Lines, S. M. Hickey, P. M. Dobbie, C. D. Bottema, and W. S. Pitchford. 2011. Genetic mapping of quantitative trait loci for meat quality and muscle metabolic traits in cattle. Anim Genet., 42: 592-599.

Fernando, R., and D. Garrick. 2009. GenSel - User manual for a portfolio of genomic selection related analyses. BioMed Central, 3rd ed.

Fox, D. G., and J. R. Black. 1984. A system for Predicting Body Composition and Performance of Growing Cattle. J Anim Sci., 58: 725-739.

Fox, D. G., C. J. Sniffen, and J. D. O'Connor. 1988. Adjusting Nutrient Requirements of Beef Cattle for Animal and Environmental Variations. J Anim Sci., 66: 1475-1495.

Fox, D. G., C. J. Sniffen, J. D. O'Connor, J. B. Russell, and P. J. Van Soest. 1992. A net carbohydrate and protein system for evaluating cattle diets: III. Cattle requirements and diet adequacy. J Anim Sci., 70: 3578-3596.

Fox, D. G., and T. P. Tylutki. 1998. Accounting for the Effects of Environment on the Nutrient Requirements of Dairy Cattle. J Dairy Sci., 81: 3085-3095.

Fox, D., L. Tedeschi, and P. Guiroy. 2001. Determining feed intake and feed efficiency of individual cattle fed in groups. Proc Beef Impr Fed 33rd Ann Res Symp Annu Meet., 33: 80-98.

Fox, D. G., L. O. Tedeschi, and M. J. Baker. 2004. Identifying differences in efficiency in beef cattle. Animal Science Dept. Mimeo No. 225, Cornell Univ., Ithaca, NY.

Frisch, J., and T. Vercoe. 1984. An analysis of growth of different cattle genotypes reared in different environments. J Agr Sci., 103: 137-153.

Geer, L. Y., A. Marchler-Bauer, R. C. Geer, L. Han, J. He, S. He, C. Liu, W. Shi, and S. H. Bryant. 2010. The NCBI BioSystems database. Nucleic Acids Research, 38: D492-D496.

Goff, S. A., M. Vaughn, S. McKay, E. Lyons, A. E. Stapleton, D. Gessler, N. Matasci, L. Wang, M. Hanlon, A. Lenards, A. Muir, N. Merchant, S. Lowry, S. Mock, M. Helmke, A. Kubach, M. Narro, N. Hopkins, D. Micklos, U. Hilgert, M. Gonzales, C. Jordan, E. Skidmore, R. Dooley, J. Cazes, R. McLay, Z. Lu, S. Pasternak, L. Koesterke, W. H. Piel, R. Grene, C. Noutsos, K. Gendler, X. Feng, C. Tang, M. Lent, S.-j. Kim, K. Kvilekval, B. S. Manjunath, V. Tannen, A. Stamatakis, M. Sanderson, S. M. Welch, K. Cranston, P. Soltis, D. Soltis, B. O'Meara, C. Ane, T. Brutnell, D. J. Kleibenstein, J. W. White, J. Leebens-Mack, M. J. Donoghue, E. P. Spalding, T. J. Vision, C. R. Myers, D. Lowenthal, B. J. Enquist, B. Boyle, A. Akoglu, G. Andrews, S. Ram, D. Ware, L. Stein, and D. Stanzione. 2011. The iPlant Collaborative: Cyberinfrastructure for Plant Biology. Frontiers in Plant Science, 2.

Gregory, K. E., L. V. Cundiff, and R. M. Koch. 1995. Genetic and phenotypic (co)variances for growth and carcass traits of purebred and composite populations of beef cattle. J Anim Sci., 73: 1920-1926.

Guilherme, A., J. V. Virbasius, V. Puri, and M. P. Czech. 2008. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. Nat Rev: Mol Cell Biol., 9: 367-377.

Gutierrez-Gil, B., J. Williams, D. Homer, D. Burton, C. Haley, and P. Wiener. 2009. Search for quantitative trait loci affecting growth and carcass traits in a cross population of beef and dairy cattle. J Anim Sci., 87: 24-36.

Hassen, A., D. Wilson, and G. Rouse. 1999. Evaluation of carcass, live, and real-time ultrasound measures in feedlot cattle: I. Assessment of sex and breed effects. J Anim Sci., 77: 273-282.

Herd, R., and S. Bishop. 2000. Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. Livest Prod Sci., 63: 111-119.

Howard, J. K., B. J. Cave, L. J. Oksanen, I. Tzameli, C. Bjorbaek, and J. S. Flier. 2004. Enhanced leptin sensitivity and attenuation of diet-induced obesity in mice with haploinsufficiency of Socs3. Nat Med., 10: 734-738.

Hu, Z.-L., S. Dracheva, W. Jang, D. Maglott, J. Bastiaansen, M. Rothschild, and J. Reecy. 2005. A QTL resource and comparison tool for pigs: PigQTLDB. Mamm Genome, 16: 792-800.

Hu, Z.-L., C. A. Park, X.-L. Wu, and J. M. Reecy. 2010. QTLdb: A comprehensive database tool building bridges between genotypes and phenotypes. In: 9th World Congress Genet Appl Livest Prod.: Aug. 1-6., Leipzig, Germany.

Hu, Z.-L., C. A. Park, X.-L. Wu, and J. M. Reecy. 2013. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. Nucleic Acids Res., 41: D871-D879.

Hu, Z.-L., and J. Reecy. 2007. Animal QTLdb: beyond a repository. Mamm Genome, 18: 1-4.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res., 37: 1-13.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc., 4: 44-57.

Jaworski, D. M., O. Sideleva, H. M. Stradecki, G. D. Langlois, A. Habibovic, B. Satish, W. G. Tharp, J. Lausier, K. LaRock, T. L. Jetton, M. Peshavaria, and R. E. Pratley. 2011. Sexually Dimorphic Diet-Induced Insulin Resistance in Obese Tissue Inhibitor of Metalloproteinase-2 (TIMP-2)-Deficient Mice. Endocrinology, 152: 1300-1313.

Jenkins, T. G., and C. L. Ferrell. 1994. Productivity through weaning of nine breeds of cattle under varying feed availabilities: I. Initial evaluation. J Anim Sci., 72: 2787-2797.

Johnson, D. D., D. K. Lunt, J. W. Savell, and G. C. Smith. 1988. Factors Affecting Carcass Characteristics and Palatability of Young Bulls. J. Anim. Sci., 66: 2568-2577.

Kanehisa, M., and S. Goto. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res., 28: 27-30.

Kanehisa, M., S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res., 42: D199-205.

Kennedy, B., J. van der Werf, and T. Meuwissen. 1993. Genetic and statistical properties of residual feed intake. J Anim Sci., 71: 3239-3250.

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

Koch, R., L. Swiger, D. Chambers, and K. Gregory. 1963. Efficiency of feed use in beef cattle. J Anim Sci., 22: 486-494.

Koch, R. M., L. V. Cundiff, K. E. Gregory, and L. D. Van Vleck. 2004. Genetic response to selection for weaning weight or yearling weight or yearling weight and muscle score in Hereford cattle: efficiency of gain, growth, and carcass characteristics. J Anim Sci., 82: 668-682.

Koesterke, L., J. E. Koltes, N. T. Weeks, K. Milfeld, M. W. Vaughn, J. M. Reecy, and D. Stanzione. 2014. Discovery of biological networks using an optimized partial correlation coefficient with information theory algorithm on Stampede's Xeon and Xeon Phi processors. Concurrency and Computation- Pract Ex., 26: 2178-2190.

Lamb, M., O. Robison, and M. Tess. 1990. Genetic parameters for carcass traits in Hereford bulls. J Anim Sci., 68: 64-69.

Liu, W., S. G. Thomas, S. L. Asa, N. Gonzalez-Cadavid, S. Bhasin, and S. Ezzat. 2003. Myostatin Is a Skeletal Muscle Target of Growth Hormone Anabolic Action. The Journal of Clinical Endocrinology & Metabolism, 88: 5490-5496

Lu, D., S. Miller, M. Sargolzaei, M. Kelly, G. Vander Voort, T. Caldwell, Z. Wang, G. Plastow, and S. Moore. 2013. Genome-wide association analyses for growth and feed efficiency traits in beef cattle. J Anim Sci., 91: 3612-3633.

Ma, L., R. L. Hanson, L. N. Que, A. M. G. Cali, M. Fu, J. L. Mack, A. M. Infante, S. Kobes, t. I. T. D. q. Consortium, C. Bogardus, A. R. Shuldiner, and L. J. Baier. 2007. Variants in ARHGEF11, a Candidate Gene for the Linkage to Type 2 Diabetes on Chromosome 1q, Are Nominally Associated With Insulin Resistance and Type 2 Diabetes in Pima Indians. Diabetes, 56: 1454-1459.

MacNeil, M. D., D. R. Bailey, J. J. Urick, R. P. Gilbert, and W. L. Reynolds. 1991. Heritabilities and genetic correlations for postweaning growth and feed intake of beef bulls and steers. J Anim Sci., 69: 3183-3189.

Mao, F., L. Chen, M. Vinsky, E. Okine, Z. Wang, J. Basarab, D. H. Crews, Jr., and C. Li. 2013. Phenotypic and genetic relationships of feed efficiency with growth performance, ultrasound, and carcass merit traits in Angus and Charolais steers. J Anim Sci., 91: 2067-2076.

Marques, E., J. D. Nkrumah, E. L. Sherman, and S. S. Moore. 2009. Polymorphisms in positional candidate genes on BTA14 and BTA26 affect carcass quality in beef cattle. J Anim Sci., 87: 2475-2484.

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

Martinez, A., N. Aldai, R. Celaya, and K. Osoro. 2010. Effect of breed body size and the muscular hypertrophy gene in the production and carcass traits of concentrate-finished yearling bulls. J Anim Sci., 88: 1229-1239.

McClure, M., N. Morsci, R. Schnabel, J. Kim, P. Yao, M. Rolf, S. McKay, S. Gregg, R. Chapple, and S. Northcutt. 2010. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim Genet., 41: 597-607.

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

Miyawaki, K., Y. Yamada, N. Ban, Y. Ihara, K. Tsukiyama, H. Y. Zhou, S. Fujimoto, A. Oku, K. Tsuda, S. Toyokuni, H. Hiai, W. Mizunoya, T. Fushiki, J. J. Holst, M. Makino, A. Tashita, Y. Kobara, Y. Tsubamoto, T. Jinnouchi, T. Jomori, and Y. Seino. 2002. Inhibition of gastric inhibitory polypeptide signaling prevents obesity. Nat Med., 8: 738-742.

Moore, S. S., D. H. Crews, Jr., and J. D. Nkrumah. 2006. Multiple and candidate gene approaches to genetic evaluation of feed efficiency in beef cattle. In: 8th World Congr Genet Appl Livest Prod., Belo Horizonte, MG, Brazil.

Morris, C., W. Pitchford, N. Cullen, A. Esmailizadeh, S. Hickey, D. Hyndman, K. Dodds, R. Afolayan, A. Crawford, and C. Bottema. 2009. Quantitative trait loci for live animal and carcass composition traits in Jersey and Limousin back-cross cattle finished on pasture or feedlot. Anim Genet., 40: 648-654.

Mujibi, F., J. Nkrumah, O. Durunna, J. Grant, J. Mah, Z. Wang, J. Basarab, G. Plastow, D. Crews, and S. Moore. 2011. Associations of marker panel scores with feed intake and efficiency traits in beef cattle using preselected single nucleotide polymorphisms. J Anim Sci., 89: 3362-3371.

Nalaila, S. M., P. Stothard, S. S. Moore, C. Li, and Z. Wang. 2012. Whole-genome QTL scan for ultrasound and carcass merit traits in beef cattle using Bayesian shrinkage method. J Anim Breed Genet., 129: 107-119.

Nkrumah, J., E. Sherman, C. Li, E. Marques, D. Crews, R. Bartusiak, B. Murdoch, Z. Wang, J. Basarab, and S. Moore. 2007a. Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. J Anim Sci., 85: 3170-3181.

Nkrumah, J. D., J. A. Basarab, Z. Wang, C. Li, M. A. Price, E. K. Okine, D. H. Crews, and S. S. Moore. 2007b. Genetic and phenotypic relationships of feed intake and measures of efficiency with growth and carcass merit of beef cattle. J Anim Sci., 85: 2711-2720.

Peters, S. O., K. Kizilkaya, D. J. Garrick, R. L. Fernando, J. M. Reecy, R. L. Weaber, G. A. Silver, and M. G. 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.

Pryce, J. E., J. Arias, P. J. Bowman, S. R. Davis, K. A. Macdonald, G. C. Waghorn, W. J. Wales, Y. J. Williams, R. J. Spelman, and B. J. Hayes. 2012. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. J Dairy Sci., 95: 2108-2119.

Pryce, J. E., J. Johnston, B. J. Hayes, G. Sahana, K. A. Weigel, S. McParland, D. Spurlock, N. Krattenmacher, R. J. Spelman, E. Wall, and M. P. L. Calus. 2014. Imputation of genotypes from low density (50,000 markers) to high density (700,000 markers) of cows from research herds in Europe, North America, and Australasia using 2 reference populations. J Dairy Sci., 97: 1799-1811.

Retallick, K. M., D. B. Faulkner, S. L. Rodriguez-Zas, J. D. Nkrumah, and D. W. Shike. 2013. The effect of breed and individual heterosis on the feed efficiency, performance, and carcass characteristics of feedlot steers. J Anim Sci., 91: 5161-5166.

Reverter, A., D. J. Johnston, H. U. Graser, M. L. Wolcott, and W. H. Upton. 2000. Genetic analyses of live-animal ultrasound and abattoir carcass traits in Australian Angus and Hereford cattle. J Anim Sci., 78: 1786-1795.

Reverter, A., and E. K. F. Chan. 2008. Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. Bioinformatics, 24: 2491-2497.

Riley, D. G., C. C. Chase, A. C. Hammond, R. L. West, D. D. Johnson, T. A. Olson, and S. W. Coleman. 2002. Estimated genetic parameters for carcass traits of Brahman cattle. J Anim Sci., 80: 955-962.

Rincon, G., K. Weber, A. Eenennaam, B. Golden, and J. Medrano. 2011. Hot topic: performance of bovine high-density genotyping platforms in Holsteins and Jerseys. J Dairy Sci., 94: 6116 - 6121.

Robinson, D. L., and V. H. Oddy. 2004. Genetic parameters for feed efficiency, fatness, muscle area and feeding behaviour of feedlot finished beef cattle. Livest Prod Sci., 90: 255-270.

Rolf, M., J. Taylor, R. Schnabel, S. McKay, M. McClure, S. Northcutt, M. Kerley, and R. Weaber. 2010. Impact of reduced marker set estimation of genomic relationship matrices on genomic selection for feed efficiency in Angus cattle. BMC Genet., 11: 24.

Rolf, M. M., J. F. Taylor, R. D. Schnabel, S. D. McKay, M. C. McClure, S. L. Northcutt, M. S. Kerley, and R. L. Weaber. 2012. Genome-wide association analysis for feed efficiency in Angus cattle. Anim Genet., 43: 367-374.

Rolf, M. M., D. J. Garrick, T. Fountain, H. R. Ramey, R. L. Weaber, J. E. Decker, E. J. Pollak, R. D. Schnabel, and J. F. Taylor. 2015. Comparison of Bayesian models to estimate direct genomic values in multi-breed commercial beef cattle. Genet Sel Evol., 47: 23.

Rolfe, K. M., W. M. Snelling, M. K. Nielsen, H. C. Freetly, C. L. Ferrell, and T. G. Jenkins. 2011. Genetic and phenotypic parameter estimates for feed intake and other traits in growing beef cattle, and opportunities for selection. J Anim Sci., 89: 3452-3459.

Saatchi, M., M. C. McClure, S. D. McKay, M. M. Rolf, J. Kim, J. E. Decker, T. M. Taxis, R. H. Chapple, H. R. Ramey, and S. L. Northcutt. 2011. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. Genet Sel Evol., 43:40.

Saatchi, M., J. Ward, and D. Garrick. 2013a. Accuracies of direct genomic breeding values in Hereford beef cattle using national or international training populations. J Anim Sci., 91: 1538-1551.

Saatchi, M., and D. Garrick. 2013b. Accuracies of genomic predictions in US beef cattle. In: Proc Assoc Adv Anim Breed Genet Conf., p 207-210.

Saatchi, M., J. Beever, J. Decker, D. Faulkner, H. Freetly, S. Hansen, H. Yampara-Iquise, K. Johnson, S. Kachman, M. Kerley, J. Kim, D. Loy, E. Marques, H. Neibergs, E. Pollak, R. Schnabel, C. Seabury, D. Shike, W. Snelling, M. Spangler, R. Weaber, D. Garrick, and J. Taylor. 2014. QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. BMC Genomics, 15: 1004.

Schenkel, F. S., S. P. Miller, and J. W. Wilton. 2004. Genetic parameters and breed differences for feed efficiency, growth, and body composition traits of young beef bulls. Canadian J Anim Sci., 84: 177-185.

Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res., 13: 2498-2504.

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

Sherman, E. L., J. D. Nkrumah, and S. S. Moore. 2010. Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. J Anim Sci., 88: 16-22.

Splan, R. K., L. V. Cundiff, and L. D. Van Vleck. 1998. Genetic parameters for sex-specific traits in beef cattle. J Anim Sci., 76: 2272-2278.

Stone, R. T., J. W. Keele, S. D. Shackelford, S. M. Kappes, and M. Koohmaraie. 1999. A primary screen of the bovine genome for quantitative trait loci affecting carcass and growth traits. J Anim Sci., 77: 1379-1384.

Tatusova T., S. Ciufo, B. Fedorov, K. O'Neill, and I. Tolstoy. 2014. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Res., 42(1):D553-9.

Tedeschi, L., and D. Fox. 2006. Using mathematical nutrition models to improve beef cattle efficiency. Beef Cattle Research in Texas. Dept. of Animal Science, Texas A&M University, College Station, Texas: 95-103.

Tedeschi, L. O., D. G. Fox, M. J. Baker, and D. P. Kirschten. 2006. Identifying differences in feed efficiency among group-fed cattle. J Anim Sci., 84: 767-776.

Tian, E., and K. G. Ten Hagen. 2009. Recent insights into the biological roles of mucin-type O-glycosylation. Glycoconjugate J., 26: 325-334.

Tong, B., S. Sasaki, Y. Muramatsu, T. Ohta, H. Kose, H. Yamashiro, T. Fujita, and T. Yamada. 2014. Association of a single-nucleotide polymorphism in myosin-binding protein C, slow-type (MYBPC1) gene with marbling in Japanese Black beef cattle. Anim Genet., 45: 611-612.

Tylutki, T. P., D. G. Fox, and R. G. Anrique. 1994. Predicting net energy and protein requirements for growth of implanted and nonimplanted heifers and steers and nonimplanted bulls varying in body size. J Anim Sci., 72: 1806-1813.

Veerkamp, R. F., M. P. Coffey, D. P. Berry, Y. de Haas, E. Strandberg, H. Bovenhuis, M. P. L. Calus, and E. Wall. 2012. Genome-wide associations for feed utilisation complex in primiparous Holstein-Friesian dairy cows from experimental research herds in four European countries. Animal, 6: 1738-1749.

Veerkamp, R. F., J. E. Pryce, D. Spurlock, D. Berry, M. Coffey, P. Lovendahl, R. van der Linde, J. Bryant, F. Miglior, Z. Wang, M. Winters, N. Krattenmacher, N. Charfeddine, J. Pedersen, and Y. De Haas. 2013. Selection on feed intake or feed efficiency: A position paper from gDMI breeding goal discussions Interbull Bulletin No. 47.

Veseth, D. A., W. L. Reynolds, J. J. Urick, T. C. Nelsen, R. E. Short, and D. D. Kress. 1993. Paternal half-sib heritabilities and genetic, environmental, and phenotypic correlation estimates from randomly selected Hereford cattle. J Anim Sci., 71: 1730-1736.

Watson-Haigh, N. S., H. N. Kadarmideen, and A. Reverter. 2010. PCIT: an R package for weighted gene co-expression networks based on partial correlation and information theory approaches. Bioinformatics, 26: 411-413.

Weber, K. L., D. J. Drake, J. F. Taylor, D. J. Garrick, L. A. Kuehn, R. M. Thallman, R. D. Schnabel, W. M. Snelling, E. J. Pollak, and A. L. Van Eenennaam. 2012. The accuracies of DNA-based estimates of genetic merit derived from Angus or multibreed beef cattle training populations. J Anim Sci., 90: 4191-4202.

Wheeler, T. L., L. V. Cundiff, S. D. Shackelford, and M. Koohmaraie. 2001. Characterization of biological types of cattle (Cycle V): carcass traits and longissimus palatability. J Anim Sci., 79: 1209-1222.

Wheeler, T., L. Cundiff, S. Shackelford, and M. Koohmaraie. 2005. Characterization of biological types of cattle (Cycle VII): Carcass, yield, and longissimus palatability traits. J Anim Sci., 83: 196-207.

Williams, C. B., G. L. Bennett, T. G. Jenkins, L. V. Cundiff, and C. L. Ferrell. 2006. Using simulation models to predict feed intake: Phenotypic and genetic relationships between observed and predicted values in cattle. J Anim Sci., 84: 1310-1316.

Wilson, L. L., J. R. McCurley, J. H. Ziegler, and J. L. Watkins. 1976. Genetic Parameters of Live and Carcass Characters from Progeny of Polled Hereford Sires and Angus-Holstein Cows. J Anim Sci., 43: 569-576.

Yuan, Z., J. Li, J. Li, X. Gao, H. Gao, and S. Xu. 2013. Effects of DGAT1 gene on meat and carcass fatness quality in Chinese commercial cattle. Mol Biol Rep., 40: 1947-1954.

VITA

Kimberly Ann Branham

Candidate for the Degree of

Master of Science

Thesis: USE OF ULTRA-HIGH-DENSITY SNP DATA FOR GENETIC EVALUATION OF PREDICTED FEED INTAKE, FEED EFFICIENCY, CARCASS, AND GROWTH TRAITS IN HEREFORD CATTLE

Major Field: Animal Science

Biographical:

Education:

Completed the requirements for the Master of Science in Animal Science at Oklahoma State University, Stillwater, Oklahoma in July, 2015.

Completed the requirements for the Bachelor of Science in Animal Science at Oklahoma State University, Stillwater, Oklahoma in 2013.

Experience:

Undergraduate Research Scholar (Oklahoma State University, 2012-2013)
Graduate Research Assistant (Oklahoma State University, 2013-2015)

Professional Memberships:

American Society of Animal Science