GENETIC PARAMETER ESTIMATION AND GENE NETWORK DERIVATION FOR FATTY ACID TRAITS IN ANGUS BEEF CATTLE

By

JUSTIN WILLIAM BUCHANAN

Bachelor of Science in Animal Science
Oklahoma State University
Stillwater, Oklahoma
2010

Master of Science in Animal Science
Oklahoma State University
Stillwater, Oklahoma
2012

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2015

GENETIC PARAMETER ESTIMATION AND GENE NETWORK DERIVATION FOR FATTY ACID TRAITS IN ANGUS BEEF CATTLE

Dissertation Approved:

Dr. Megan Rolf

Dr. Udaya DeSilva

Dr. Peter Hoyt

Dr. Raluca Mateescu

ACKNOWLEDGEMENTS

This research was supported by Pfizer Animal Genetics.

This project is supported by USDA-NIFA Award 2012-67015-19420

The authors also acknowledge the 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

Some of the computing 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.

Name: JUSTIN WILLIAM BUCHANAN

Date of Degree: MAY, 2015

Title of Study: GENETIC PARAMETER ESTIMATION AND GENE NETWORK DERIVATION FOR FATTY ACID TRAITS IN ANGUS BEEF CATTLE

Major Field: ANIMAL BREEDING AND GENETICS

Abstract: The fatty acid profile of beef is a complex trait that affects eating quality, healthfulness attributes for the consumer, and carcass characteristics. Longissimus muscle samples were obtained from 1,833 Angus cattle to determine the intramuscular fatty acid composition for 31 lipids and lipid classes from triacylglycerol (TAG) and phospholipid (PL) fractions and classified by structure into saturated (SFA), monounsaturated (MUFA), polyunsaturated (PUFA), omega-3 (n-3), and omega-6 (n-6) fatty acids. Restricted maximum likelihood methods combined with pedigree data were used to estimate variance components. Heritability estimates ranged from 0 to 0.63 for the major classes of fatty acids. Heritability estimates differed between the TAG and PL fractions, with higher estimates for TAG up to 0.64 and lower estimates for PL that ranged up to 0.14. Phenotypic and genotypic correlations among individual fatty acids were determined for the TAG fraction as well as among carcass traits including ribeye area (REA), numerical marbling score (MARB), yield grade (YG), ether fat (EFAT), and Warner-Bratzler shear force value (WBSF). Strong negative or positive genetic correlations were observed among individual fatty acids in the TAG fraction, which ranged from -0.99 to 0.97 ($P < 0.05$). Moderate correlations between carcass traits and fatty acids from the TAG fraction ranged from -0.43 to 0.32 ($P < 0.05$). These results indicate that fatty acids prominent in beef tissues show significant genetic variation as well as genetic relationships to carcass traits. Phenotypic measures of fatty acid profile from the triacylglycerol and phospholipid fraction of longissimus muscle, pedigree information, and Illumina 54k bovine SNPchip genotypes were utilized to derive an annotated gene network underlying the fatty acid composition. The Bayes-B statistical model was utilized to perform a genome wide association study to estimate effects between 54k SNP genotypes and 39 individual fatty acid phenotypes within each fraction (TAG or PL). Effects were estimated for 1-Mb genomic windows as well as for 54k SNP genotypes. A partial correlation algorithm was used to illustrate correlated regions of the genome with a set of 1 Mb windows explaining up to 34.55% of the genetic variation in both fatty acid fractions. Annotated gene network clusters were generated by utilizing a partial correlation and information theory algorithm (PCIT) in conjunction with network scoring and visualization software to analyze correlated SNP across 39 fatty acid phenotypes to identify SNP of functional significance. Significant pathways implicated in fatty acid metabolism through network analysis included fatty acid synthesis, glycerol-phospholipid metabolism, and cell-to-cell adhesion and trafficking. A network analysis using partial correlations and annotation of significant SNP's can yield functional information about the genetic mechanisms underlying the fatty acid profile of beef.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I

LITERATURE REVIEW

INTRODUCTION

Beef is a nutritious product that provides an excellent source of protein, vitamins, minerals, and lipids in the human diet. The lipid profile of beef contributes to the overall healthfulness and palatability of the final beef product, which indicates it is a trait of interest for consumers. Previous studies have characterized the lipid profile in various beef cattle tissues under different dietary conditions, in different breeds, and at various age points. This published collection of fatty acid phenotype data indicates that lipid storage in beef cattle is a dynamic process with individual lipids exhibiting a wide range of phenotypic and genetic variance estimates. This range in observed variance can be partially explained by genetic differences among animals for lipid synthesis, desaturation and deposition, as well as by the functionality of specific lipids and lipid classes in the biological environment of muscle and fat tissues. The two major lipid depots in beef cattle tissue are represented by the triacylglycerol and phospholipid fractions of the total lipid isolated from both muscle and adipose tissue. The triacylglycerol lipid fraction captures the lipids stored as triglycerides in adipose cells.

The phospholipid fraction captures lipids contained in the more diverse phospholipid cellular membrane of both myocytes and adipocytes. The triacylglycerol and phospholipid fractions exhibit the characteristics of quantitative traits. These traits are controlled by many individual genes with many correlated individual lipids composing larger lipid classes such as saturated (SFA), monounsaturated (MUFA), and polyunsaturated (PUFA) fatty acids. There is also a significant environmental component that influences the fatty acid profile. Given these factors, the triacylglycerol and phospholipid fractions present an excellent opportunity for an analysis of genetic parameters, the identification of candidate genes, and the visualization of gene networks controlling the overall lipid profile in beef cattle tissues.

FATTY ACID PROFILE OF BEEF TISSUES

A wide range of studies have characterized the total fatty acid profile of various beef cattle tissues (Wood et al., 2008; Daley et al., 2010; Hoehne et al., 2012; Pavan and Duckett, 2013; Duckett et al., 2014). The total fatty acid fraction represents the distribution of all lipids present in a biological sample including those derived from the lipid membranes of multiple cell types other than adipocytes. Depending on factors such as diet, age, and tissue, the total fatty acid profile of beef is generally composed of approximately 40-50% of SFA, 40-60% of MUFA, and 5-15% of PUFA (Wood et al., 2008). Two important factors affecting fatty acid profile are maturity of the animal and location of the tissue sampled. Muscle type and adipose location have a significant effect on the fatty acid profile (Pavan and Duckett, 2013; Liu et al., 2015b). Also, as an animal matures and a larger proportion of excess energy is used for fatty acid synthesis there tends to be an increase in the accumulation of SFA in relation to PUFA (Warren et al., 2008). This is likely due to a shift towards lipogenesis and a

shift away from adipogenesis as an animal reaches maturity and subsequent the need for more adipocytes is reduced. During lipogenesis the primary fatty acids being produced are saturated in nature. The primary product from the major protein complex driving lipogenesis, which is fatty acid synthase (FASN), is C16:0, which explains this shift in the SFA:PUFA ratio. It is known that this shift occurs as fatty acids are stored as triglycerides in adipocytes, but it is not clear how this shift affects the phospholipid membrane in a maturing animal. This membrane is a dynamic lipid depot and seems to undergo changes in fluidity and composition as adipogenesis proceeds in a maturing tissue (Pietilainen et al., 2011). More research is needed to understand the differences between these two lipid depots and how they change under various conditions and maturity points.

A method developed by Hartman (1967) allowed the separation of polar and non-polar lipids prior to gas chromatography that yields the neutral lipid and phospholipid fatty acid fractions in separate components. The non-polar neutral lipid fraction contains the triacylglycerol, diacylglycerol, ester, and cholesterol components of the tissue. The polar fraction contains the phospholipid bilayer fatty acids which are composed of four major phospholipids in mammals: phosphatidylethanolamine, phosphatidylserine, phosphatidylcholine, and sphingomyelin (Alberts, 2002). Each of these four major phospholipids contain two nonpolar fatty acid chains which compose the individual fatty acids identified as being associated with the phospholipid fatty acid fraction through gas chromatography analysis. Separating the total fatty acid fraction allows for a more detailed analysis of the genes and biological pathways affecting adipogenesis and lipid synthesis when compared to the total fatty acid fraction. When the two fractions are combined it cannot be determined if the fatty acids in the analysis come from the triacylglycerol or the

phospholipid which likely have distinct biological origins and configurations in various cell types.

To date, there have been few studies that have presented a comprehensive separate analysis of the triacylglycerol and phospholipid fatty acid fractions in beef cattle species. Kazala et al. (1999) presented an analysis of the intramuscular fatty acid composition in crossbred Wagyu cattle in which the triacylglycerol fraction was separated from and compared to the total lipid fraction. In this study the triacylglycerol fraction was found to be very similar in composition to the total lipid with no significant difference detected in the MUFA/SFA ratio in longissimus. Dannenberger et al. (2007) presented a comprehensive analysis of the fatty acids distributed in multiple phospholipid classes in beef muscle including the phosphatidylethanolamines, phosphatidylcholines, phosphatidylinositols, cardiolipins, sphingomyolins, and lysophosphatidylethanolamines using high performance thin layer chromatography. Analysis of these individual phospholipid classes revealed that pasture feeding to finishing leads to a significant accumulation of omega-3 fatty acids in all classes when compared to finishing on a concentrate diet. The phospholipid membrane is a dynamic lipid depot and is known to undergo changes in fluidity and composition as adipogenesis proceeds in a maturing tissue (Pietilainen et al., 2011). Smith et al. (1998) presented a comprehensive analysis of distribution and saturation of triacylglycerol species in beef cattle in response to different dietary formulas. This study concluded that diet had significant effects on the distribution of saturation and the composition of triacylglycerol species. The previous studies added important data sets for fatty acid analysis, but a future research on the triacylglycerol and phospholipid fractions is needed.

Margetak et al. (2012) presented a complete comparison of the triacylglycerol and phospholipid fatty acid fractions form the pars costalis diaphragmatic muscle and subcutaneous fat in beef cattle undergoing different dietary supplementations containing sunflower and flax oils. This study found the triacylglycerol fraction of muscle tissue to contain significantly higher amounts of C14:0 and C16:0. The phospholipid fraction of muscle tissue contained higher amounts of C18:0, C16:1c9, and C18:1c9. The phospholipid fraction of subcutaneous fat was found to contain higher amounts of C14:0, C16:0, C18:0, C16:1c9, and C18:1c9. The study also concluded that oil supplementation increased the absolute amounts of elongated unsaturated fatty acids in both the triacylglycerol and phospholipid fractions. There were no other studies identified that presented a direct comparison of the triacylglycerol and phospholipid fractions in beef. Future research is needed to determine the effects of breed, maturity, muscle type, and lipid depot on the triacylglycerol and phospholipid fatty acid profiles.

GENETIC PARAMETERS ESTIMATES OF FATTY ACID TRAITS

Genetic parameter estimates for individual lipids and lipid classes in beef cattle tissues are available for the total fatty acid fraction. However, there are currently no studies available analyzing genetic parameters in triacylglycerol and phospholipid fractions. Heritability estimates for individual lipids and lipid classes range from 0 to a moderate heritability of approximately 0.6. These heritability estimates indicate the certain lipids would respond well to selection programs. Ekine-Dzivenu et al. (2014) estimated genetic parameters for fatty acids traits in 223 Angus and Charolais crossbred commercial steers. In this study heritability estimates for individual lipids ranged from approximately 0 to 0.51.

Heritability estimates for most fatty acids were low, with SFA, MUFA, and PUFA fatty acid classes having a heritabilities less than 0.15. The highest heritability estimates were found for C14:1 and C18:1 with estimates of 0.51 and 0.43, respectively. The relatively low population size likely led to an underestimation of heritability estimates for fatty acid traits in this study.

Inoue et al. (2011) also estimated heritabilities for fatty acid traits in a population of 863 Japanese Black steers. Heritability estimates in this study ranged from approximately 0 to 0.86. Traits with the highest heritability estimates included C14:1 and C14:0 with heritability estimates of 0.86 and 0.82, respectively. Total MUFA also exhibited a high heritability estimate of 0.66. The authors of this study note that the heritability estimates obtained seem high when compared to other breeds and other studies. The authors also note that this difference might be present due to differences in fatty acid synthesis and desaturation enzyme activity in Japanese Black cattle compared to other breeds. Nogi et al. (2011) presented the results of a similar study in a population of 2,275 Japanese Black cattle. Heritability estimates in this study were similar and raged from approximately 0 to 0.78. The highest heritability estimates were obtained for C14:0 and C18:1 with heritabilities of 0.70 and 0.78, respectively. The lipid classes SFA, MUFA, and PUFA had heritability estimates of 0.66, 0.68, and 0.47, respectively. These studies represent the best examples currently present in the literature for fatty acid heritability estimates in a single breed of cattle due to the large population size, complete reporting of individual lipid heritability estimates, and uniform genetic background of the animals the study.

Pitchford et al. (2002) obtained heritability estimates for fatty acid traits in a population of 1,215 animals with 7 distinct sire breeds. This study also used the percentage of the total lipid as the measurement. Heritability estimates in this population range from

approximately 0 to a moderate heritability of 0.27. Fatty classes SFA and MUFA had heritabilities of 0.27 and 0.17, respectively. The fatty acid C16:0 had one of the highest heritability estimates of all individual lipids at 0.21. In contrast to the study by Inoue et al. (2011), this study found relatively lower heritability estimates across all individual lipids and lipid classes. It is possible that using cattle from a variety of genetic backgrounds as opposed to a single breed results in lower heritability estimates.

Ahlberg et al. (2014) obtained posterior mean genomic heritability estimates for various fatty acid classes as a proportion of phenotypic variation explained by a genomic marker panel in a population of 236 crossbred steers and heifers. Heritability estimates for fatty acid classes PUFA and MUFA were 0.7 and 0.4, respectively, when measured on a percentage of total lipid basis. Heritability estimates of PUFA and MUFA were 0.7 and 0.85, respectively, when measured on the basis of mg/100 g of wet tissue.

Heritability estimates for fatty acid traits appear to be variable across the studies estimating these parameters in beef cattle. Genetic background of the animals in the study as well as the measurement system used to determine the fatty acid measurement seem to be the two factors causing the most variation in heritability estimates. In general, studies using a single breed with the percentage of total lipid measurement system yield the highest estimates of heritability for fatty acid traits. The classes SFA and MUFA as well as the individual lipids making up those classes appear to have moderate to high heritabilities. The more unsaturated lipids composing the PUFA class appear to have low to moderate heritabilities. The data gathered in these studies indicates that overall the fatty acid profile has a moderate heritability and certain fatty acids would respond to a marker assisted selection program.

GENETIC CORRELATIONS INVOLVING FATTY ACID TRAITS

Genetic correlation estimates have also been well characterized among individual lipids, lipid classes, and carcass traits for the total fatty acid fraction. Fatty acid synthesis and desaturation occurs through a pathway of related enzyme complexes to produce the many lipids and lipid classes found in mammalian tissues. The central driver of lipid synthesis in mammalian tissues is a large protein complex known as fatty acid synthase (FASN) (Alberts, 2002). The primary products of FASN synthesis are C14:0 and C16:0, which are derived by the addition of 2 carbon acetyl CoA to a growing carbon chain until the final product reaches either 14 or 16 carbons in length. Many other enzymes in addition to FASN work to lengthen individual lipids and add features such as desaturations and isomerizations after the final C14:0 or C16:0 are produced. These include the desaturase class of enzymes, such as steroyl CoA desaturase, and elongation enzymes. It is reasonable to conclude that certain fatty acids would exhibit moderate to high genetic correlations since genetic variation in these biological pathways and networks would affect all lipid products in the assembly line. It also follows that individual lipids and lipid classes should be genetically correlated to carcass traits since the fatty acid profile is known to vary at different levels of tissue maturity (Warren et al., 2008).

Both direction and strength of phenotypic and genetic correlations among fatty acids appear to be highly dependent upon the measurement system used (percent of total lipid vs. mg/100 g tissue). Using percent of total lipid calculation appears to give higher heritability estimates for the majority of lipids and lipid classes (Saatchi et al., 2013; Ahlberg et al., 2014). Phenotypic correlations between SFA and the unsaturated lipid classes MUFA and

PUFA are generally negative. SFA is the primary product of *de novo* lipid synthesis, and the newly synthesized saturated fatty acids are then used as precursors for unsaturated fatty acid products derived from that synthesis. Maturity of the animal is also known to drive this association, as fatter animals typically have higher amounts of SFA compared to unsaturated fatty acids (Warren et al., 2008). Multiple studies have identified this phenotypic association (Pitchford et al., 2002; Inoue et al., 2011; Ekine-Dzivenu et al., 2014).

Genetic correlations among fatty acids tend to be less predictable across multiple studies, but the general trend of SFA exhibiting a negative genetic correlation with other fatty acids seems to be a common association. This is likely a reflection of the general pathway of lipid elongation and desaturation that occurs as lipid synthesis and incorporation into various depots proceeds in the adipocyte. Ekine-Dzivenu et al. (2014) found SFA have a negative genetic correlation with MUFA and PUFA, with genetic correlation estimates of -0.99 and -0.41, respectively. MUFA and PUFA were found to have a weak but positive genetic correlation of 0.2. Similarly, Pitchford et al. (2002) found a negative genetic correlation between SFA and other unsaturated fatty acids. The fatty acid C14:0 had negative genetic correlations of -0.61 and -0.27 with MUFA and UFA, respectively. Inoue et al. (2011) also observed the C14:0 to have a negative genetic correlation with MUFA and UFA of -0.74 and -0.81, respectively.

Other individual lipids also exhibit predictable genetic correlations. In general, individual SFA's of different lengths tend to be positively correlated (Inoue et al., 2011; Nogi et al., 2011). A strong negative genetic correlation is also consistently observed between C18:0 and C18:1 (Inoue et al., 2011; Nogi et al., 2011), which is likely a reflection of stearoyl Co-A desaturase (SCD) variation in the catalysis of C18:0 desaturation into C18:1

(Smith et al., 2006). Individual lipids and lipid classes exhibit a wide range of genetic correlation estimations. Some individual lipids exhibit a genetic correlation of almost 1.0 or -1.0 which is likely due to the pathway being highly dependent on the products from each previous step in the synthesis, elongation, or desaturation of lipid products, or an over estimation of the parameter.

GENOMIC REGIONS OF INTEREST FOR FATTY ACID TRAITS

Multiple studies have carried out genome-wide association studies in various breeds of cattle for the total fatty acid fraction in order to identify genomic regions, markers, and genes of interest. One of the most important genes involved in *de novo* synthesis of fatty acids is FASN. This protein is a complex of multiple subunits which are transcribed from a region on chromosome 19 starting at approximately 51,384,900 base pairs (bp). Multiple studies have identified this region as having a high association with saturated fatty acids including C14:0, C16:0, and total SFA (Matsuhashi et al., 2011; Uemoto et al., 2011; Ishii et al., 2013; Saatchi et al., 2013; Hayakawa et al., 2015). There have also been multiple detailed studies of this region in relation to the fatty acid profile and there appear to be many different SNP's in the region affecting synthesis of SFA (Li et al., 2012; Oh et al., 2012; Lee et al., 2014). Saatchi et al. (2013) also estimated that markers in this region explain up to 25% of the genetic variance in saturated fatty acids with the highest genetic variance explained in *cis*-9 C18:1. This data suggests there are likely multiple causative mutations in the FASN gene that affect the fatty acid profile in multiple species of cattle (Casas et al., 2001; Casas et al., 2003; McClure et al., 2010).

However, not every species seems to have this association between SFA the FASN loci. Cesar et al. (2014) identified 8 genomic regions explaining approximately 1% of the genetic variance in SFA's, including C14:0, C16:0, and C18:0. None of these 8 regions were near or overlapped the FASN loci, but they did overlap with previously identified loci affecting marbling score, backfat thickness, and carcass and body weight in Angus cattle. These associations are likely detecting the effect of loci causing variation in carcass fatness, which directly has an effect on percentage of SFA. At different levels of carcass fatness the ratio of SFA to unsaturated fatty acids changes (Warren et al., 2008). This effect can be partially explained by the morphology of adipose cells at different maturity points. In younger animals it can be expected that adipose cells are in a stage of multiplication under conditions of excess energy intake, at which point the ratio of the lipids in the phospholipid membrane to the lipids stored as triacylglycerol is high (Graugnard et al., 2010). As the adipose tissue ages a higher proportion of the lipids synthesized and incorporated are stored in the triacylglycerol as triglycerides which are generally more saturated in nature than the phospholipid membrane (Smith et al., 1998). Given this shift in lipid storage as a tissue ages, it can be expected that loci affecting carcass fatness and adipose cell morphology would also affect the proportion of SFA.

Saatchi et al. (2013) also identified a region on chromosome 29 starting at about the 18th Mb harboring the candidate gene thyroid responsive hormone (THRSP or SPOT14) to explain the second highest amount of genetic variance in C14:0, C16:0, C16:1, cis-9 C18:1, long chain fatty acids (LCFA), and medium chain fatty acids (MCFA). This gene is known to be involved in SFA and LCFA synthesis through transcriptional activity and possibly by acting as a cofactor to FASN (Cunningham et al., 1998; LaFave et al., 2006). Other studies

have also identified an association between fatty acid traits and variation and expression of THRSP (Hudson et al., 2014; Oh et al., 2014). Variation in FASN and THRSP appear to be associated with high genetic variance in SFA and fatty acid synthesis in beef cattle tissues.

Genomic regions associated with percentage of MUFA have also been well characterized. A region on chromosome 26 starting at approximately 21,132,700 bp harbors the SCD gene which is known to be involved in lipid desaturation in mammalian tissues (Marchitelli et al., 2013; Estany et al., 2014). Multiple studies have identified this genomic region as having a significant effect on C14:1. C16:1, C18:1 and other elongated MUFA species through GWAS (Ishii et al., 2013; Saatchi et al., 2013; Cesar et al., 2014). Additional regions have been associated with MUFA containing candidate genes for fatty acid related traits. Cesar et al. (2014) identified a region on chromosome 2 in Nellore cattle near two candidate genes, glutamate decarboxylase 1 (GAD1) and specificity protein 5-transcription factor (Sp5), which are both involved in general energy metabolism, adipogenesis, and lipogenesis pathways. These studies support the hypothesis that SCD is the main candidate gene responsible for variation in MUFA species in beef cattle tissues.

Results from estimates of genetic parameters and GWAS from PUFA in beef tissues provide the least information among the three major lipid saturation classes. Heritability estimates for the PUFA lipids are the lowest for all lipid species (Inoue et al., 2011; Saatchi et al., 2013; Ekine-Dzivenu et al., 2014). This relatively low variance observed for PUFA species is likely a reflection of the biological importance of these lipids in the cell. Since the majority of these elongated and unsaturated lipids are found in the cell membrane it can be reasoned that variance in this lipid depot would be detrimental to the fluidity and function of the phospholipid bilayer. The low genetic variance estimates for these phenotypes also allow

for the discovery of fewer candidate genes explaining genetic variance using a GWAS methodology. This does not indicate that there are fewer genes involved in the synthesis or incorporation of these lipids into adipose tissue, but that it is more difficult to identify them using these methods. Another source of difficulty in identifying candidate genes involved with the PUFA species is that a number of them are not synthesized in-vivo, but are instead incorporated from dietary sources. Also, the majority of the PUFA synthesis that does occur in mammalian tissues occurs in the liver rather than in adipose. The main pathway leading to PUFA synthesis relies on the conversion of linoleic and alpha linoleic acids to arachidonic (ARA), eicosapentaenoic (EPA), and docosahexaenoic (DHA) through the activity of fatty acid elongases (ELOVLs) and fatty acid desaturases (FADS) (Jump, 2011).

Studies reporting GWAS results for PUFA species in beef cattle have not found an association with the FAD or ELOVL loci, but rather genomic regions harboring or near candidate genes involved in membrane function, membrane adhesion, adipogenesis, or cell signaling. Cesar et al. (2014 identified 9 genomic regions explaining at least 1% of the genetic variance in multiple PUFA species. Candidate genes in the regions included aquaporin 7 (AQP7), lysil oxidase-like 2 (LOXL2), and RAR-related orphan receptor. These candidate genes are involved in cellular component functions such as the PPAR signaling pathway, lean body mass determination in mice, and cellular receptor pathways but no previous association with bovine adipose tissue has been reported. Saatchi et al. (2013) reported associations between omega-3 and omega-6 fatty acids and regions on chromosomes 23, 14, 26, and 11, but no candidate genes were reported near these regions with association to lipid metabolism. There were also no regions of the genome in this study with a posterior probability of inclusion (PPI) greater than 0.9 for total PUFA or any other

individual PUFA lipids. Identification of genomic loci affecting PUFA species needs further investigation. It appears that the low phenotypic and genetic variance estimates for this trait hinder the discovery of candidate genes affecting variation.

Other candidate genes affecting fatty acid profile in livestock species have been described in the literature, but have not shown up in GWAS studies in these traits. Graugnard et al. (2010) described the PPAR gamma signaling axis as a major driver of adipogenesis in response to energy abundance under different dietary conditions. Other lipid metabolism genes implicated in through differential gene expression in this study included adiponectin (ADIPOQ), fatty acid binding protein 4 (FABP4), diglycerol acyltranferase 2 (DGAT2), and sterol regulatory element-binding transcription factor 1 (SREBF1). The role of ADIPOQ has been studied as a regulator of lipid synthesis in milk fatty acid synthesis (Singh et al., 2014; Locher et al., 2015), and it would follow that this hormone would likely have a similar role in signaling lipogenesis in adipose. The binding protein FABP4 has also exhibited differential expression in the muscle of cattle fed differing levels of soybean oil or rumen protected fat (Oliveira et al., 2014). There have been a wide variety of genes described in the literature affecting fatty acid traits in beef cattle and other species. Taken together, this set of candidate genes likely contains a large number of causal mutations contributing variation to fatty acid traits in beef cattle.

GENE NETWORK THEORY

Complex trait analysis in livestock species has been assisted by recent advances in the generation of genomic and general "omics" related data sets. However, this generation of extremely large datasets has created a need for more complex analysis systems to detect

biological phenomena and relate genotype to phenotype. Some of the primary goals in complex trait analysis using these large datasets are to identify causal genes and causal mutations, interactions among these genes and genomic regions, and to assemble these genes and interactions into networks or pathways in a meaningful way that relates to the underlying biology (Feltus, 2014). Such tasks have been the central goal of disciplines such as systems biology or systems genetics. These systems biology approaches have been developed to identify a variety of genome features such as copy number variation (Jiang et al., 2015), diagnostic features in the cancer genome (Liu et al., 2015a), and causal mutations underlying traits of interest (Hudson et al., 2009; Chen et al., 2014). Causal mutations are of particular importance to the livestock genomics industry due to the development of selection strategies based on genomic data, which can increase in accuracy when causal mutations are included in the prediction (Druet et al., 2014).

One important development in this area has been the incorporation of GWAS data into the generation of regulatory networks underlying multiple related traits of interest. A method developed by Reverter and Fortes (2013) has utilized the inclusion of SNP identified in GWAS to build gene networks highlighting genes of functional relevance to significant biological pathways, rather than just a single phenotype. The method relies on the generation of SNP networks derived from an association weight network tested for interactions by using an algorithm known as partial correlation and information theory (PCIT) (Reverter and Chan, 2008). The major principal behind this method relies on the assumption that SNP having a high impact on multiple related phenotypes are likely of high importance or contain causal mutations.

The first step in the generation of the association weight matrix is to create a matrix of SNP effects for all phenotypes used in the model. A threshold needs to be chosen that incorporates a number of SNP that have a sufficiently high effect on the phenotype of highest importance. Once the initial set of SNP are chosen, the rest of the phenotypes are populated with the same chosen SNP with effects from each respective GWAS. For Bayesian GWAS methods an appropriate threshold might include a posterior probability of association (PPA) of 0.50 to 0.95, depending on the number of SNP that fall within this range. One benefit of the method is that it can be customized to multiple types of omics data that has been generated in association with multiple related phenotypes of interest. Other methods have successfully been used to generate an association weight matrix from data such as the transcriptome (Lehnert et al., 2006; Fortes et al., 2010; Fortes et al., 2012).

The next step is to identify correlations between all SNP or data points in the association weight matrix. The PCIT algorithm was developed specifically to handle this task of identifying associations or correlations among all data points in a large matrix. The matrix consists of columns that correspond to phenotypes in the analysis, and rows that correspond to the SNP selected from GWAS results with the highest association to the most important phenotype in the analysis. The algorithm first estimates correlations between every pair of SNP in the dataset across all of the phenotypes. Next, the algorithm identifies a partial correlation between each SNP pair and every other SNP, if such a correlation exists. SNP pairs with a partial correlation of 0 to any other SNP are considered isolated and only associated with one another, and subsequently removed from the final output. The algorithm was optimized for use as an R package (Watson-Haigh et al., 2010) and also optimized to run in parallel for high performance computing applications (Koesterke et al., 2013). The final

output of the algorithm is a set of SNP pairs and their associated direct correlations which can be utilized in the final visualization of SNP networks.

Network scoring, annotation, and visualization are the final steps in the association weight matrix approach spanning multiple phenotypes. The are many software packages that can handle gene network visualization, but the Cytoscape software package (Shannon et al., 2003) is particularly useful for its ability to score highly interconnected network clusters. A plugin for the Cytoscape software called MCODE (Bader and Hogue, 2003) was developed to score highly interconnected clusters of genes. The clusters are identified by an analysis of cluster density, which is the product of the number of connections in the network and the number of SNP. Clusters with the highest network density are ranked highest in the scoring criteria. These highly interconnected clusters represent the candidate genes or SNP that have the highest impact on the overall phenotype of interest since they contain associated genes or SNP that affect all phenotypes in the model. Annotation of the final networks is necessary to determine if the captured SNP fall in or near genes of functional significance to the overall phenotype in the analysis.

Multiple studies have utilized the association weight matrix approach to analyze quantitative traits in livestock species. A study by Fortes et al. (2012) utilized transcriptome data to build an association weight matrix to analyze first service conception rates in Brangus heifers. Transcriptome data from 10 related growth and fertility traits were used in the construction of the association weight matrix. This approach identified 5 highly interconnected transcription factors hypothesized to be related to overall fertility as well as markers in multiple genes that have been previously associated with fertility traits in beef cattle.

A study by Ramayo-Caldas et al. (2014b) utilized a SNP effect based association weight matrix to analyze intramuscular fat deposition in approximately 10,000 beef cattle from 3 breeds. The study looked at 29 different traits including intramuscular fat, related fat phenotypes, feedlot performance, and various meat quality traits to identify the markers with the highest impact on fat deposition. The resulting networks produced three transcription factors as key regulators of fat deposition and carcass traits: PPARGC1A, HNF4G, and FOXP3. Multiple other markers were identified within genes of biological importance to the pathways regulating these traits of interest as well. Importantly, it was noted that the transcription factors and major genes of interest were not identified in the GWAS as markers with the highest effect associated with any one individual phenotype. Only the combination of multiple phenotypes into the association weight matrix allowed these markers to be highlighted.

Another study by Ramayo-Caldas et al. (2014a) used SNP effects to create an association weight matrix for intramuscular fatty acid composition in porcine. This study looked at 15 fatty acid phenotypes to identify key regulators of intramuscular fatty acid metabolism. The final network analysis identified the transcription factors NCOA2, FHL2, and EP300 as central regulators of fatty acid metabolism along with many other individual genes of functional significance. This study was unique in that the authors went on to validate the identified transcription factors as having differential expression at the transcriptomic level using real-time PCR. They found expression differences for extreme fatty acid phenotypes in two breeds in liver tissue for 55 genes involved in their association network, including the three identified transcription factors. Also, approximately 60% of the connections identified in the network analysis were validated at the transcriptomic level.

Creating the association weight matrix with multiple types of omics data or validating the networks through expression analysis appears to be a robust method for identifying genes of interest in pathways affecting multiple phenotypes of interest.

CONCLUSION

The fatty acid profile of beef is a complex phenotype that requires a systems biology approach to identify the genes of highest impact regulating overall lipid metabolism. The fatty acid profile is associated with economic traits of interest such as intramuscular fat and healthfulness of the final beef product. In general, individual lipids and lipids classes exhibit a wide range of heritability estimates. Lipids of higher abundance such as medium chain SFA and MUFA exhibit a moderate to high heritability, which indicates these traits would respond to a genomic selection program. Given the wide range of observed fatty acid phenotypes in various lipid depots (triacylglycerol vs. phospholipid) muscle types, breeds, and feeding programs there is a need to identify the major pathways, transcription factors, and genes responsible for variation within the overall process of lipid metabolism. The identification of the drivers of lipid metabolism has economic implications for the beef cattle industry since the value of the final beef product is highly dependent on lipogenesis during the finishing phase.

The association weight matrix approach appears to be a robust methodology that can identify the central regulators of a complex metabolic process such as the fatty acid profile of beef. The fatty acid profile has multiple related individual phenotypes that make up the overall fatty acid profile including the major lipid classes as well as the various individual lipids contained in muscle and adipose tissue. The following chapters will present an analysis

of the genetic parameters and genetic correlations associated with the intramuscular fatty acid profile from the triacylglycerol and phospholipid fatty acid fractions followed by the implementation of the association weight matrix approach to generate a network analysis of lipid metabolism in Angus beef cattle.

LITERATURE CITED

Ahlberg, C.M., Schiermiester, L.N., T, J., Howard, Calkins, C.R., and Spangler, M.L. (2014). Genome wide association study of cholesterol and poly- and monounsaturated fatty acids, protein, and mineral content of beef from crossbred cattle. *Meat Sci* 98**,** 804-814. doi: 10.1016/j.meatsci.2014.07.030.

Alberts, B. (2002). *Molecular biology of the cell.* New York: Garland Science.

Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4**,** 2.

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

Casas, E., Stone, R.T., Keele, J.W., Shackelford, S.D., Kappes, S.M., and Koohmaraie, M. (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.

Cesar, A.S., Regitano, L.C., Mourao, G.B., Tullio, R.R., Lanna, D.P., Nassu, R.T., Mudado, M.A., Oliveira, P.S., Do Nascimento, M.L., Chaves, A.S., Alencar, M.M., Sonstegard, T.S., Garrick, D.J., Reecy, J.M., and Coutinho, L.L. (2014). Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genet* 15**,** 39. doi: 10.1186/1471-2156-15-39.

Chen, J.C., Alvarez, M.J., Talos, F., Dhruv, H., Rieckhof, G.E., Iyer, A., Diefes, K.L., Aldape, K., Berens, M., Shen, M.M., and Califano, A. (2014). Identification of causal

genetic drivers of human disease through systems-level analysis of regulatory networks. *Cell* 159**,** 402-414. doi: 10.1016/j.cell.2014.09.021.

Cunningham, B.A., Moncur, J.T., Huntington, J.T., and Kinlaw, W.B. (1998). "Spot 14" protein: a metabolic integrator in normal and neoplastic cells. *Thyroid* 8**,** 815-825.

Daley, C.A., Abbott, A., Doyle, P.S., Nader, G.A., and Larson, S. (2010). A review of fatty acid profiles and antioxidant content in grass-fed and grain-fed beef. *Nutr J* 9**,** 10. doi: 10.1186/1475-2891-9-10.

Dannenberger, D., Nuernberg, G., Scollan, N., Ender, K., and Nuernberg, K. (2007). Diet alters the fatty acid composition of individual phospholipid classes in beef muscle. *J Agric Food Chem* 55**,** 452-460. doi: 10.1021/jf061793x.

Druet, T., Macleod, I.M., and Hayes, B.J. (2014). Toward genomic prediction from whole-genome sequence data: impact of sequencing design on genotype imputation and accuracy of predictions. *Heredity (Edinb)* 112**,** 39-47. doi: 10.1038/hdy.2013.13.

Duckett, S.K., Fernandez Rosso, C., Volpi Lagreca, G., Miller, M.C., Neel, J.P., Lewis, R.M., Swecker, W.S., and Fontenot, J.P. (2014). Effect of frame size and time-on-pasture on steer performance, longissimus muscle fatty acid composition, and tenderness in a forage-finishing system. *J Anim Sci* 92**,** 4767-4774. doi: 10.2527/jas.2014-7411.

Ekine-Dzivenu, C., Chen, L., Vinsky, M., Aldai, N., Dugan, M.E., Mcallister, T.A., Wang, Z., Okine, E., and Li, C. (2014). Estimates of genetic parameters for fatty acids in brisket adipose tissue of Canadian commercial crossbred beef steers. *Meat Sci* 96**,** 1517-1526. doi: 10.1016/j.meatsci.2013.10.011.

Estany, J., Ros-Freixedes, R., Tor, M., and Pena, R.N. (2014). A functional variant in the stearoyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork. *PLoS One* 9**,** e86177. doi: 10.1371/journal.pone.0086177.

Feltus, F.A. (2014). Systems genetics: a paradigm to improve discovery of candidate genes and mechanisms underlying complex traits. *Plant Sci* 223**,** 45-48. doi: 10.1016/j.plantsci.2014.03.003.

Fortes, M.R., Reverter, A., Zhang, Y., Collis, E., Nagaraj, S.H., Jonsson, N.N., Prayaga, K.C., Barris, W., and Hawken, R.J. (2010). Association weight matrix for the genetic dissection of puberty in beef cattle. *Proc Natl Acad Sci U S A* 107**,** 13642-13647. doi: 10.1073/pnas.1002044107.

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

Graugnard, D.E., Berger, L.L., Faulkner, D.B., and Loor, J.J. (2010). High-starch diets induce precocious adipogenic gene network up-regulation in longissimus lumborum of early-weaned Angus cattle. *Br J Nutr* 103**,** 953-963. doi: 10.1017/S0007114509992789.

Hartman, L. (1967). Comparison of two methods for the separation of polar from nonpolar lipids. *J Lipid Res* 8**,** 285-287.

Hayakawa, K., Sakamoto, T., Ishii, A., Yamaji, K., Uemoto, Y., Sasago, N., Kobayashi, E., Kobayashi, N., Matsuhashi, T., Maruyama, S., Matsumoto, H., Oyama, K., Mannen, H., and Sasazaki, S. (2015). The g.841G>C SNP of FASN gene is associated with fatty acid composition in beef cattle. *Anim Sci J*. doi: 10.1111/asj.12357.

Hoehne, A., Nuernberg, G., Kuehn, C., and Nuernberg, K. (2012). Relationships between intramuscular fat content, selected carcass traits, and fatty acid profile in bulls using a F2-population. *Meat Sci* 90**,** 629-635. doi: 10.1016/j.meatsci.2011.10.005.

Hudson, N.J., Reverter, A., and Dalrymple, B.P. (2009). A differential wiring analysis of expression data correctly identifies the gene containing the causal mutation. *PLoS Comput Biol* 5**,** e1000382. doi: 10.1371/journal.pcbi.1000382.

Hudson, N.J., Reverter, A., Greenwood, P.L., Guo, B., Cafe, L.M., and Dalrymple, B.P. (2014). Longitudinal muscle gene expression patterns associated with differential intramuscular fat in cattle. *Animal***,** 1-10. doi: 10.1017/S1751731114002754.

Inoue, K., Kobayashi, M., Shoji, N., and Kato, K. (2011). Genetic parameters for fatty acid composition and feed efficiency traits in Japanese Black cattle. *Animal* 5**,** 987-994. doi: 10.1017/S1751731111000012.

Ishii, A., Yamaji, K., Uemoto, Y., Sasago, N., Kobayashi, E., Kobayashi, N., Matsuhashi, T., Maruyama, S., Matsumoto, H., Sasazaki, S., and Mannen, H. (2013). Genome-wide association study for fatty acid composition in Japanese Black cattle. *Anim Sci J* 84**,** 675-682. doi: 10.1111/asj.12063.

Jiang, Y., Qin, H., and Yang, L. (2015). Using network clustering to predict copy number variations associated with health disparities. *PeerJ* 3**,** e677. doi: 10.7717/peerj.677.

Jump, D.B. (2011). Fatty acid regulation of hepatic lipid metabolism. *Curr Opin Clin Nutr Metab Care* 14**,** 115-120. doi: 10.1097/MCO.0b013e328342991c.

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

Koesterke, L., Milfield, K., Vaughn, M.W., Stanzione, D., Koltes, J.E., Weeks, N.T., Reecy, J.M. (2013). Optimizing the PCIT algorithm on stampede's Xeon and Xeon Phi processors for faster discovery of biological networks. Proceedings: XSEDE 13. doi: 10.1145/2484762.2484794.

Lafave, L.T., Augustin, L.B., and Mariash, C.N. (2006). S14: insights from knockout mice. *Endocrinology* 147**,** 4044-4047. doi: 10.1210/en.2006-0473.

Lee, J., Jin, M., Lee, Y., Ha, J., Yeo, J., and Oh, D. (2014). Gene-gene interactions of fatty acid synthase (FASN) using multifactor-dimensionality reduction method in Korean cattle. *Mol Biol Rep* 41**,** 2021-2027. doi: 10.1007/s11033-014-3050-8.

Lehnert, S.A., Byrne, K.A., Reverter, A., Nattrass, G.S., Greenwood, P.L., Wang, Y.H., Hudson, N.J., and Harper, G.S. (2006). Gene expression profiling of bovine skeletal muscle in response to and during recovery from chronic and severe undernutrition. *J Anim Sci* 84**,** 3239-3250. doi: 10.2527/jas.2006-192.

Li, C., Aldai, N., Vinsky, M., Dugan, M.E., and Mcallister, T.A. (2012). Association analyses of single nucleotide polymorphisms in bovine stearoyl-CoA desaturase and fatty acid synthase genes with fatty acid composition in commercial cross-bred beef steers. *Anim Genet* 43**,** 93-97. doi: 10.1111/j.1365-2052.2011.02217.x.

Liu, R., Guo, C.X., and Zhou, H.H. (2015a). Network-based approach to identify prognostic biomarkers for estrogen receptor-positive breast cancer treatment with tamoxifen. *Cancer Biol Ther* 16**,** 317-324. doi: 10.1080/15384047.2014.1002360.

Liu, T., Lei, Z.M., Wu, J.P., and Brown, M.A. (2015b). Fatty acid composition differences between adipose depot sites in dairy and beef steer breeds. *J Food Sci Technol* 52**,** 1656-1662. doi: 10.1007/s13197-013-1117-5.

Locher, L., Haussler, S., Laubenthal, L., Singh, S.P., Winkler, J., Kinoshita, A., Kenez, A., Rehage, J., Huber, K., Sauerwein, H., and Danicke, S. (2015). Effect of increasing body condition on key regulators of fat metabolism in subcutaneous adipose tissue depot and circulation of nonlactating dairy cows. *J Dairy Sci* 98**,** 1057-1068. doi: 10.3168/jds.2014-8710.

Marchitelli, C., Contarini, G., De Matteis, G., Crisa, A., Pariset, L., Scata, M.C., Catillo, G., Napolitano, F., and Moioli, B. (2013). Milk fatty acid variability: effect of some candidate genes involved in lipid synthesis. *J Dairy Res* 80**,** 165-173. doi: 10.1017/S002202991300006X.

Margetak, C., Travis, G., Entz, T., Mir, P.S., Wei, S., and Dodson, M.V. (2012). Fatty Acid Composition of Phospholipids and in the Central and External Positions of Triacylglycerol in Muscle and Subcutaneous Fat of Beef Steers Fed Diets Supplemented with Oil Containing n6 and n3 Fatty Acids While Undergoing One of Three 48 h Feed Withdrawal Treatments. *J Lipids* 2012**,** 543784. doi: 10.1155/2012/543784.

Matsuhashi, T., Maruyama, S., Uemoto, Y., Kobayashi, N., Mannen, H., Abe, T., Sakaguchi, S., and Kobayashi, E. (2011). Effects of bovine fatty acid synthase, stearoyl-

coenzyme A desaturase, sterol regulatory element-binding protein 1, and growth hormone gene polymorphisms on fatty acid composition and carcass traits in Japanese Black cattle. *J Anim Sci* 89**,** 12-22. doi: 10.2527/jas.2010-3121.

Mcclure, M.C., Morsci, N.S., Schnabel, R.D., Kim, J.W., Yao, P., Rolf, M.M., Mckay, S.D., Gregg, S.J., Chapple, R.H., Northcutt, S.L., and Taylor, J.F. (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. doi: 10.1111/j.1365-2052.2010.02063.x.

Nogi, T., Honda, T., Mukai, F., Okagaki, T., and Oyama, K. (2011). Heritabilities and genetic correlations of fatty acid compositions in longissimus muscle lipid with carcass traits in Japanese Black cattle. *J Anim Sci* 89**,** 615-621. doi: 10.2527/jas.2009-2300.

Oh, D., Lee, Y., La, B., Yeo, J., Chung, E., Kim, Y., and Lee, C. (2012). Fatty acid composition of beef is associated with exonic nucleotide variants of the gene encoding FASN. *Mol Biol Rep* 39**,** 4083-4090. doi: 10.1007/s11033-011-1190-7.

Oh, D.Y., Lee, Y.S., La, B.M., Lee, J.Y., Park, Y.S., Lee, J.H., Ha, J.J., Yi, J.K., Kim, B.K., and Yeo, J.S. (2014). Identification of exonic nucleotide variants of the thyroid hormone responsive protein gene associated with carcass traits and Fatty Acid composition in korean cattle. *Asian-Australas J Anim Sci* 27**,** 1373-1380. doi: 10.5713/ajas.2014.14101.

Oliveira, D.M., Chalfun-Junior, A., Chizzotti, M.L., Barreto, H.G., Coelho, T.C., Paiva, L.V., Coelho, C.P., Teixeira, P.D., Schoonmaker, J.P., and Ladeira, M.M. (2014). Expression of genes involved in lipid metabolism in the muscle of beef cattle fed

soybean or rumen-protected fat, with or without monensin supplementation. *J Anim Sci* 92**,** 5426-5436. doi: 10.2527/jas.2014-7855.

Pavan, E., and Duckett, S.K. (2013). Fatty acid composition and interrelationships among eight retail cuts of grass-feed beef. *Meat Sci* 93**,** 371-377. doi: 10.1016/j.meatsci.2012.09.021.

Pietilainen, K.H., Rog, T., Seppanen-Laakso, T., Virtue, S., Gopalacharyulu, P., Tang, J., Rodriguez-Cuenca, S., Maciejewski, A., Naukkarinen, J., Ruskeepaa, A.L., Niemela, P.S., Yetukuri, L., Tan, C.Y., Velagapudi, V., Castillo, S., Nygren, H., Hyotylainen, T., Rissanen, A., Kaprio, J., Yki-Jarvinen, H., Vattulainen, I., Vidal-Puig, A., and Oresic, M. (2011). Association of lipidome remodeling in the adipocyte membrane with acquired obesity in humans. *PLoS Biol* 9**,** e1000623. doi: 10.1371/journal.pbio.1000623.

Pitchford, W.S., Deland, M.P., Siebert, B.D., Malau-Aduli, A.E., and Bottema, C.D. (2002). Genetic variation in fatness and fatty acid composition of crossbred cattle. *J Anim Sci* 80**,** 2825-2832.

Ramayo-Caldas, Y., Ballester, M., Fortes, M.R., Esteve-Codina, A., Castello, A., Noguera, J.L., Fernandez, A.I., Perez-Enciso, M., Reverter, A., and Folch, J.M. (2014a). From SNP co-association to RNA co-expression: novel insights into gene networks for intramuscular fatty acid composition in porcine. *BMC Genomics* 15**,** 232. doi: 10.1186/1471-2164-15-232.

Ramayo-Caldas, Y., Fortes, M.R., Hudson, N.J., Porto-Neto, L.R., Bolormaa, S., Barendse, W., Kelly, M., Moore, S.S., Goddard, M.E., Lehnert, S.A., and Reverter, A. (2014b). A marker-derived gene network reveals the regulatory role of PPARGC1A, HNF4G,

and FOXP3 in intramuscular fat deposition of beef cattle. *J Anim Sci* 92**,** 2832-2845. doi: 10.2527/jas.2013-7484.

Reverter, A., and Chan, E.K. (2008). Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. *Bioinformatics* 24**,** 2491-2497. doi: 10.1093/bioinformatics/btn482.

Reverter, A., and Fortes, M.R. (2013). Breeding and Genetics Symposium: building single nucleotide polymorphism-derived gene regulatory networks: Towards functional genomewide association studies. *J Anim Sci* 91**,** 530-536. doi: 10.2527/jas.2012-5780.

Saatchi, M., Garrick, D.J., Tait, R.G., Jr., Mayes, M.S., Drewnoski, M., Schoonmaker, J., Diaz, C., Beitz, D.C., and Reecy, J.M. (2013). Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. *BMC Genomics* 14**,** 730. doi: 10.1186/1471-2164-14-730.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13**,** 2498-2504. doi: 10.1101/gr.1239303.

Singh, S.P., Haussler, S., Heinz, J.F., Akter, S.H., Saremi, B., Muller, U., Rehage, J., Danicke, S., Mielenz, M., and Sauerwein, H. (2014). Lactation driven dynamics of adiponectin supply from different fat depots to circulation in cows. *Domest Anim Endocrinol* 47**,** 35-46. doi: 10.1016/j.domaniend.2013.12.001.

Smith, S.B., Lunt, D.K., Chung, K.Y., Choi, C.B., Tume, R.K., and Zembayashi, M. (2006). Adiposity, fatty acid composition, and delta-9 desaturase activity during growth in

beef cattle. *Animal Science Journal* 77**,** 478-486. doi: DOI 10.1111/j.1740-0929.2006.00375.x.

Smith, S.B., Yang, A., Larsen, T.W., and Tume, R.K. (1998). Positional analysis of triacylglycerols from bovine adipose tissue lipids varying in degree of unsaturation. *Lipids* 33**,** 197-207.

Uemoto, Y., Abe, T., Tameoka, N., Hasebe, H., Inoue, K., Nakajima, H., Shoji, N., Kobayashi, M., and Kobayashi, E. (2011). Whole-genome association study for fatty acid composition of oleic acid in Japanese Black cattle. *Animal Genetics* 42**,** 141-148. doi: DOI 10.1111/j.1365-2052.2010.02088.x.

Warren, H.E., Scollan, N.D., Enser, M., Hughes, S.I., Richardson, R.I., and Wood, J.D. (2008). Effects of breed and a concentrate or grass silage diet on beef quality in cattle of 3 ages. I: Animal performance, carcass quality and muscle fatty acid composition. *Meat Sci* 78**,** 256-269. doi: 10.1016/j.meatsci.2007.06.008.

Watson-Haigh, N.S., Kadarmideen, H.N., and Reverter, A. (2010). PCIT: an R package for weighted gene co-expression networks based on partial correlation and information theory approaches. *Bioinformatics* 26**,** 411-413. doi: 10.1093/bioinformatics/btp674.

Wood, J.D., Enser, M., Fisher, A.V., Nute, G.R., Sheard, P.R., Richardson, R.I., Hughes, S.I., and Whittington, F.M. (2008). Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci* 78**,** 343-358. doi: 10.1016/j.meatsci.2007.07.019.

CHAPTER II

GENETIC PARAMETERS AND GENETIC CORRELATIONS AMONG TRIACYLGLYCEROL AND PHOSPHOLIPID FRACTIONS IN ANGUS CATTLE

J.W. Buchanan*, J.M. Reecy‡, D.J. Garrick‡, Q. Duan‡, D.C. Beitz†‡, R.G. Mateescu§

*Department of Animal Science, Oklahoma State University, Stillwater, OK 74078, USA

‡ Department of Animal Science, Iowa State University, Ames, IA 50011, USA

† Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA

§ Department of Animal Sciences, University of Florida, Gainesville, FL 32608, USA

This research was supported by Zoetis Animal Genetics

ABSTRACT

The objective of this study was to estimate genetic parameters for intramuscular fatty acids in beef tissue. *Longissimus* muscle samples were obtained from 1,833 Angus cattle to determine the intramuscular fatty acid composition for 31 lipids and lipid classes from triacylglycerol (TAG) and phospholipid (PL) fractions and classified by structure into saturated (SFA), monounsaturated (MUFA), polyunsaturated (PUFA), omega-3 (n-3), and omega-6 (n-6) fatty acids. An atherogenic index (AI) was also determined as a measure of the unsaturated to SFA ratio. Restricted maximum likelihood methods combined with pedigree data were used to estimate variance components with the WOMBAT software package. Heritability estimates ranged from 0 to 0.63 for the major classes of fatty acids. Heritability estimates differed between the TAG and PL fractions, with higher estimates for TAG up to 0.64 and lower estimates for PL that ranged up to 0.14. Phenotypic and genotypic correlations among individual fatty acids were determined for the TAG fraction as well as among carcass traits including ribeye area (REA), numerical marbling score (MARB), yield grade (YG), ether fat (EFAT), and Warner-Bratzler shear force value (WBSF). Strong negative or positive genetic correlations were observed among individual fatty acids in the TAG fraction, which ranged from -0.99 to 0.97 ($P < 0.05$). Moderate correlations between carcass traits and fatty acids from the TAG fraction ranged from -0.43 to 0.32 ($P < 0.05$). These results indicate that fatty acids prominent in beef tissues show significant genetic variation as well as genetic relationships to carcass traits.

INTRODUCTION

Beef is an important source of essential amino acids, minerals, vitamins and beneficial fatty acids. The nutritional value of beef has become increasingly important to specific groups of consumers some of whom are willing to pay a premium for low rather than high marbling in beef (Killinger et al., 2004). This indicates that consumer awareness of the relationship between the consumption of certain fatty acids and cardiovascular health likely affects purchasing decisions. There is interest in characterization of the genetic relationships among health traits such as fatty acid profile and mineral and vitamin contents, as well as the relationship of these traits to other carcass characteristics (Mateescu et al., 2013; Saatchi et al., 2013). Fatty acid profiles vary by breed (Wood et al., 2008), age and fatness, and together these factors indicate it is necessary to increase knowledge of how the fatty acid profile varies with carcass characteristics and in response to selection for traits such as marbling and tenderness.

Total intramuscular and subcutaneous fatty acid profile measurements can be derived from a combination of lipid fractions from the TAG and PL cellular components (Murphy, 2001). Proportions of fatty acids in each fraction are known to vary according to the relative amount of intramuscular fat (Hoehne et al., 2012). Analysis of these individual cellular components will lead to a better understanding of the genetic mechanisms behind variation in fatty acid synthesis and the relationship to carcass and adipose tissue development in beef cattle.

The objective of this study was to estimate genetic parameters for both the TAG and PL fractions in LM in Angus beef cattle. Additionally, genetic correlations were analyzed among heritable individual fatty acids and important carcass traits.

MATERIALS AND METHODS

Animals and Fatty Acid Determination

A total of 1,833 Angus-sired bulls (n = 450), steers (n = 1,022), and heifers (n = 361) representing offspring of 155 sires were used in this study. All cattle were finished on concentrate diets in Iowa (n = 908), California (n = 344), Colorado (n = 291), or Texas (n = 290). Animals were harvested at commercial facilities when they reached typical US market endpoints with an average age of 457 ± 46 days. Production characteristics and additional details of the sample collection and preparation of these cattle were reported previously (Garmyn et al., 2011). After external fat and connective tissue were removed, the 1.27-cm LM samples were analyzed for fatty acid composition. Determination of fatty acid methyl esters by gas chromatography was reported previously (Saatchi et al., 2013).

Statistical Analysis

Trait means and standard deviations were calculated using the MEANS procedure in SAS 9.3 (SAS Inst. Inc., Cary, NC). For each fatty acid or fatty acid class, restricted maximum likelihood procedures were used to estimate genetic and residual variances as well as heritability, based on a single-trait animal model fitted to the data using WOMBAT (Meyer, 2007). Restricted maximum likelihood procedures were used to estimate genetic and phenotypic covariances from a multi-trait animal model simultaneously fitted to all six traits using WOMBAT (Meyer, 2007). In matrix notation, the basic model equation was:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$$

where:

$\boldsymbol{y}$ is a vector of the observations for six traits,

$\boldsymbol{X}$ is an incidence matrix relating observations to fixed effects,

$\boldsymbol{\beta}$ is a vector of the fixed effects for each trait,

$\boldsymbol{Z}$ is an incidence matrix relating observations to random animal effects,

$\boldsymbol{u}$ is a vector containing the random genetic effects for all animals and all six traits, and

$\boldsymbol{e}$ is a vector of the random residual errors for all measured traits and animals. Contemporary groups were defined based on the three-way combination of gender at harvest (bull, heifer or steer), finishing location (California, Colorado, Iowa, Texas), and harvest date, for a total of 33 groups. Contemporary groups were fit as fixed effects in all analyses.

It is assumed that the random vectors $\boldsymbol{u}$ and $\boldsymbol{e}$ are independent and have multivariate normal distributions with mean zero so that $E[\boldsymbol{y}]= \boldsymbol{Xb}$. Variance assumptions included $Var(\boldsymbol{u}) = \boldsymbol{A} \otimes \boldsymbol{G_0}$ and $Var(\boldsymbol{e}) = \boldsymbol{I} \otimes \boldsymbol{R_0}$, where $\boldsymbol{G_0}$ =matrix of additive genetic covariances between traits and $\boldsymbol{R_0}$ = matrix of covariances between residuals on the same animal; $\boldsymbol{A}$= pedigree-based relationship matrix; $\boldsymbol{I}$ = identity matrix of order equal to the number of animals with phenotypes; and $\otimes$ = matrix direct product.

A pedigree file with 5,907 individuals including identification of all animal, sire, and dam trios for the animals with phenotypes and four ancestral generations was used to define relationships among animals in the data set. Significance of genetic correlations was obtained as $\theta \pm Z_{\alpha/2}$ (sampling error), which assumed normality of the estimator, θ.

RESULTS AND DISCUSSION

Means, standard deviations, and coefficients of variation for individual fatty acids in both the TAG and PL fractions are in **Table 2.1** on a percentage basis. Overall, the TAG

fraction contained a higher percentage of SFA and MUFA compared to the PL fraction. In contrast to TAG, the PL fraction contained a greater percentage of PUFA, which was expected as the major component of PL is the phospholipid membrane that surrounds the adipose cell. The major depot for lipid storage in mammalian tissue is white adipose tissue. The main source of lipids stored in the form of TAG are synthesized through fatty acid synthase (FASN) with the donation of an acyl-CoA. Less is known about the direct origin and synthesis of PL species, but sources include dietary incorporation into adipocytes and cellular membranes (Soriguer et al., 2000), as well as synthesis from diacylglycerol precursors (Ikeda et al., 2011). Additionally, the TAG fraction on average represents greater than 90% of the fatty acids found in beef tissues (Wood et al., 2008). In addition to the relatively greater proportion of fatty acids that originated from the TAG, Hoehne et al. (2012) report that there is also a significant correlation between increasing total muscular fat level and the TAG fraction. In contrast to the TAG fraction, lipids originating from PL appeared to have less direct correlation to total muscular fat.

Heritabilities

Phenotypic and genetic parameter estimates for total fatty acid fraction were reported previously for this population (Saatchi et al., 2013). Genetic variance and heritability estimates for individual fatty acids as well as fatty acid classes are shown in **Table 2.2** for both TAG and PL fractions. Heritability estimates for individual fatty acids from the TAG fraction ranged up to 0.64. In general, heritabilities for short to medium chain fatty acids in this fraction, which included 14:0 to 18:1 fatty acids, had the highest heritability estimates from 0.32 to 0.64. This is likely a result of direct *de-novo* fatty acid synthesis through the

FASN pathway, which produces primarily 16:0 and some 14:0 fatty acids (Wood et al., 2008). For the fatty acid classes in the TAG fraction, SFA showed the highest heritability at 0.50, with MUFA showing a similar heritability at 0.46. It should be noted that while heritabilities for TAG appeared relatively high, and there is also a relationship between the mean fraction of a particular fatty acid and heritability. Fatty acids with a higher mean fraction tended to have higher heritabilities, while lower estimates were found for fatty acids in lower abundance. Heritability estimates from the TAG fraction appear to be consistent with other studies that reported heritabilities for total fatty acid traits in Angus and Japanese Black cattle (Inoue et al., 2011; Nogi et al., 2011; Saatchi et al., 2013). These results suggest that the TAG fraction is closely representative of total fatty acid profile.

Estimates of heritability for the PL fraction in intramuscular mammalian tissue have not been published. In the PL fraction heritabilities were significantly lower when compared to the TAG. Individual fatty acid estimates were lowly heritable and ranged up to 0.14. For the categories of fatty acids, SFA and MUFA were lowly heritable with estimates at 0.03 and 0.02, respectively. Total PUFA showed almost no heritability with the estimate at 0.001. These comparatively low heritability estimates observed in the PL fraction are likely due to very low phenotypic variation in relation to measurement errors, as well as dietary dependencies (Dannenberger et al., 2007). Low phenotypic variance in the PL is likely due to the importance of the biological function of the phospholipid membrane to adipose tissue. Overall, the heritability estimates for the TAG fraction are likely representative of the variance among individuals for *de-novo* fatty acid synthesis. These results highlight the TAG fraction as the main source of genetic variation in overall fatty acid phenotypes in muscular tissue.

Genetic Correlations among Fatty Acids

Genetic correlations among fatty acids, fatty acid classes, and carcass traits for the TAG fraction are shown in **Table 2.3**. Due to low genetic variance for PL fractions, those traits were omitted from further analysis. Genetic correlations among individual fatty acids and fatty acid classes in the TAG fraction ranged from -0.99 to 0.97 ($P < 0.05$). The genetic correlation between SFA and MUFA was -0.99 ($P < 0.05$), which indicates that the genes that cause increased levels of SFA do so at the expense of reduced levels of MUFA or vice versa. This finding is consistent with previously identified regulators of desaturation such as stearoyl-coenzyme a desaturase (SCD) (Matsuhashi et al., 2011). Significant negative genetic correlations were also observed between SFA and PUFA (-0.49, $P < 0.05$) and MUFA and PUFA (0.36, $P < 0.05$).

Genetic Correlations with Carcass Traits

Genetic correlations between fatty acids from the TAG fraction and carcass traits ranged from -0.43 to 0.32 ($P < 0.05$). MUFA was positively genetically correlated with measures of increased muscular fat (EFAT and MARB). Correlations between EFAT and MUFA were 0.13 and EFAT and MARB were 0.98 ($P < 0.05$), respectively. Increased levels of MUFA have been previously associated with increased muscular fat in beef cattle (Cecchinato et al., 2012), likely due to the relationship between an increase in the relative amount of TAG compared to PL when fattening occurs (Vernon et al., 1999). Total PUFA showed a negative genetic correlation to measures of muscular fat, with genetic correlations to MARB of -0.20, and EFAT -0.16 ($P < 0.05$). Again, these correlations emphasize the

genetic relationships between higher degree of muscular fat and lower amounts of PUFA in the TAG fraction. These relationships among fatty acids in the TAG fraction and muscular fat are consistent with Hoehne et al. (2012), which suggest overall fatty acid saturation increased in the TAG fraction with increased intramuscular fat level.

A number of fatty acids in the TAG fraction were genetically correlated with tenderness. The genetic correlation between SFA and WBSF was -0.30 ($P < 0.05$), which indicated that genetic factors promoting increased accumulation of SFA also promote a tougher LM. Following this trend, genetic correlations between PUFA and WBSF was 0.27, and between MARB and WBSF was -0.50 ($P < 0.05$). This association is likely due to the relationship of muscular fatness to tenderness, where genes contributing to increased accumulation of SFA simultaneously contribute to decreased tenderness (Garmyn et al., 2011). The genetic relationship between fatty acid desaturation and tenderness is important as selection for tenderness has the potential to alter fatty acid profile.

CONCLUSION

Heritabilities and genetic variance of fatty acids in the TAG fraction are much higher than those for PL, which essentially exhibit no genetic variation. Accordingly, heritabilities for the TAG fraction are representative of heritabilities previously estimated for total fatty acid fraction (Saatchi et al., 2013). Due to the functional nature of the membranes that are composed of the PL fraction, it was expected that fractions of certain fatty acids would remain relatively constant. Fatty acids found in the TAG fraction are of primary interest, because they are the dominant proportion of total fatty acids (greater than 90%) found in intramuscular adipose tissue. There is significant genetic association between the degree of

fatty acid saturation and measures of muscular fat and tenderness, which is important when considering selection programs designed to alter these traits.

The future direction of selection for carcass traits will have an impact on carcass healthfulness, as these results show a direct genetic association between muscular fatness and healthfulness. Current selection programs placing a premium on carcass marbling are likely to decrease healthfulness from a fatty acid standpoint, as increased muscular fatness is associated with the accumulation of higher mean fractions of short chain and saturated fatty acids. Additionally, if fatty acids were to be measured and included as a selection criterion, these data show that manual separation of fatty acid fractions is not a necessary step. The majority of the moderate to highly heritable fatty acids originate from the TAG fraction, which are able to be sufficiently estimated from the total fatty acid measurement alone.

LITERATURE CITED

Cecchinato, A., De Marchi, M., Penasa, M., Casellas, J., Schiavon, S., and Bittante, G. (2012). Genetic analysis of beef fatty acid composition predicted by near-infrared spectroscopy. *J Anim Sci* 90, 429-438. doi: 10.2527/jas.2011-4150.

Dannenberger, D., Nuernberg, G., Scollan, N., Ender, K., and Nuernberg, K. (2007). Diet alters the fatty acid composition of individual phospholipid classes in beef muscle. *J Agric Food Chem* 55, 452-460. doi: 10.1021/jf061793x.

Garmyn, A.J., Hilton, G.G., Mateescu, R.G., Morgan, J.B., Reecy, J.M., Tait, R.G., Jr., Beitz, D.C., Duan, Q., Schoonmaker, J.P., Mayes, M.S., Drewnoski, M.E., Liu, Q., and Vanoverbeke, D.L. (2011). Estimation of relationships between mineral concentration and fatty acid composition of longissimus muscle and beef palatability traits. *J Anim Sci* 89, 2849-2858. doi: 10.2527/jas.2010-3497.

Hoehne, A., Nuernberg, G., Kuehn, C., and Nuernberg, K. (2012). Relationships between intramuscular fat content, selected carcass traits, and fatty acid profile in bulls using a F2-population. *Meat Sci* 90, 629-635. doi: 10.1016/j.meatsci.2011.10.005.

Ikeda, K., Kubo, A., Akahoshi, N., Yamada, H., Miura, N., Hishiki, T., Nagahata, Y., Matsuura, T., Suematsu, M., Taguchi, R., and Ishii, I. (2011). Triacylglycerol/phospholipid molecular species profiling of fatty livers and regenerated non-fatty livers in cystathionine beta-synthase-deficient mice, an animal model for homocysteinemia/homocystinuria. *Anal Bioanal Chem* 400, 1853-1863. doi: 10.1007/s00216-011-4703-2.

Inoue, K., Kobayashi, M., Shoji, N., and Kato, K. (2011). Genetic parameters for fatty acid composition and feed efficiency traits in Japanese Black cattle. *Animal* 5, 987-994. doi: 10.1017/S1751731111000012.

Killinger, K.M., Calkins, C.R., Umberger, W.J., Feuz, D.M., and Eskridge, K.M. (2004). Consumer visual preference and value for beef steaks differing in marbling level and color. *J Anim Sci* 82, 3288-3293.

Mateescu, R.G., Garmyn, A.J., Tait, R.G., Jr., Duan, Q., Liu, Q., Mayes, M.S., Garrick, D.J., Van Eenennaam, A.L., Vanoverbeke, D.L., Hilton, G.G., Beitz, D.C., and Reecy, J.M. (2013). Genetic parameters for concentrations of minerals in longissimus muscle and their associations with palatability traits in Angus cattle. *J Anim Sci* 91, 1067-1075. doi: 10.2527/jas.2012-5744.

Matsuhashi, T., Maruyama, S., Uemoto, Y., Kobayashi, N., Mannen, H., Abe, T., Sakaguchi, S., and Kobayashi, E. (2011). Effects of bovine fatty acid synthase, stearoyl-coenzyme A desaturase, sterol regulatory element-binding protein 1, and growth hormone gene polymorphisms on fatty acid composition and carcass traits in Japanese Black cattle. *J Anim Sci* 89, 12-22. doi: 10.2527/jas.2010-3121.

Meyer, K. (2007). WOMBAT: a tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *J Zhejiang Univ Sci* B 8, 815-821. doi: 10.1631/jzus.2007.B0815.

Murphy, D.J. (2001). The biogenesis and functions of lipid bodies in animals, plants and microorganisms. *Prog Lipid Res* 40, 325-438.

Nogi, T., Honda, T., Mukai, F., Okagaki, T., and Oyama, K. (2011). Heritabilities and genetic correlations of fatty acid compositions in longissimus muscle lipid with

carcass traits in Japanese Black cattle. *J Anim Sci* 89, 615-621. doi: 10.2527/jas.2009-2300.

Saatchi, M., Garrick, D.J., Tait, R.G., Jr., Mayes, M.S., Drewnoski, M., Schoonmaker, J., Diaz, C., Beitz, D.C., and Reecy, J.M. (2013). Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. *BMC Genomics* 14, 730. doi: 10.1186/1471-2164-14-730.

Soriguer, F.J., Tinahones, F.J., Monzon, A., Pareja, A., Rojo-Martinez, G., Moreno, F., Esteva, I., and Gomez-Zumaquero, J.M. (2000). Varying incorporation of fatty acids into phospholipids from muscle, adipose and pancreatic exocrine tissues and thymocytes in adult rats fed with diets rich in different fatty acids. *Eur J Epidemiol* 16, 585-594.

Vernon, R.G., Barber, M.C., and Travers, M.T. (1999). Present and future studies on lipogenesis in animals and human subjects. *Proc Nutr Soc* 58, 541-549.

Wood, J.D., Enser, M., Fisher, A.V., Nute, G.R., Sheard, P.R., Richardson, R.I., Hughes, S.I., and Whittington, F.M. (2008). Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci* 78, 343-358. doi: 10.1016/j.meatsci.2007.07.019.

Table 2.1. Means, number of records (N), standard deviations (SD) and coefficient of variation (CV) for percent fatty acids from the triacylglycerol and phospholipid fraction in *LM* from Angus cattle.

Trait	Triacylglycerol			Phospholipid		
	Mean (N = 1833)	SD2	CV3	Mean (N = 1828)	SD	CV
Fatty Acid, %						
14:0	3.08	0.51	0.17	3.74	4.63	1.24
14:1	0.70	0.21	0.30	0.20	0.52	2.60
16:0	27.37	1.73	0.06	20.16	3.58	0.18
16:1	3.82	0.63	0.16	0.76	0.94	1.24
17:0	1.42	0.39	0.27	1.96	2.69	1.37
17:1	1.01	0.33	0.33	1.13	1.00	0.88
18:0	13.22	1.90	0.14	9.79	2.63	0.27
18:1 cis-9	40.28	2.88	0.07	19.39	7.09	0.37
18:1 cis-11/12/13	0.52	0.23	0.44	0.17	0.64	3.76
18:1, trans-10/11	3.68	1.57	0.43	0.53	1.13	2.13
18:2	2.01	0.52	0.26	25.37	6.83	0.27
18:3, n-3	0.16	0.16	1.00	0.11	0.64	5.82
18:3, n-6	0.00	0.01	0.00	0.01	0.12	12.00
20:0	0.02	0.05	2.50	0.12	0.54	4.50
20:1	0.05	0.13	2.60	0.10	0.34	3.40
20:2	0.08	0.11	1.38	0.06	0.29	4.83
20:3, n-3	0.00	0.02	0.00	0.23	0.71	3.09
20:3, n-6	0.01	0.03	3.00	0.64	1.06	1.66
20:4	0.02	0.07	3.50	8.50	2.81	0.33
22:0	0.02	0.05	2.50	0.80	1.07	1.34
22:1	0.01	0.03	3.00	0.10	0.47	4.70
23:0	0.01	0.05	5.00	0.17	0.42	2.47
24:0	0.03	0.08	2.67	0.46	0.67	1.46
Sum, %						
SFA	45.81	2.50	0.05	38.93	5.32	0.14
MUFA	51.57	2.57	0.05	23.94	8.06	0.34
PUFA	2.62	0.88	0.34	37.13	9.03	0.24
PUFA, n-3	0.33	0.40	1.21	2.11	2.05	0.97
PUFA, n-6	2.29	0.63	0.28	35.01	8.62	0.25
Ratio						
n-6:n-3	0.14	0.15	1.07	0.06	0.07	1.17
AI1	0.74	0.09	0.12	0.62	0.34	0.55

[1] Atherogenic index=weighted ratio of palmitic and myristic acids to total unsaturated fatty acids

[2] Standard deviation

[3] Coefficient of variation

Table 2.2. Genetic (σ2a) variance and heritability (h2) estimates with SE for fatty acid phenotypes (phospholipid and triacylglycerol fractions) in LM from Angus cattle obtained by single trait REML analysis.

	Triacylglycerol		Phospholipid	
Trait1	σ2a	h2 ± SE	σ2a	h2 ± SE
Fatty Acid, %				
14:0	0.130	0.587 ± 0.083	0.000	0.000 ± 0.030
14:1	0.026	0.636 ± 0.084	0.020	0.094 ± 0.056
16:0	1.443	0.537 ± 0.085	1.366	0.138 ± 0.070
16:1	0.147	0.509 ± 0.079	0.000	0.000 ± 0.044
17:0	0.025	0.550 ± 0.096	0.000	0.000 ± 0.031
17:1	0.015	0.457 ± 0.091	0.017	0.030 ± 0.037
18:0	0.595	0.437 ± 0.080	0.081	0.018 ± 0.038
18:1 cis-9	2.082	0.335 ± 0.076	2.371	0.069 ± 0.049
18:1 cis-11/12/13	0.012	0.316 ± 0.086	0.010	0.042 ± 0.037
18:1, trans-10/11	0.778	0.524 ± 0.091	0.000	0.000 ± 0.030
18:2	0.063	0.373 ± 0.074	2.490	0.082 ± 0.047
18:3, n-3	0.001	0.053 ± 0.074	0.003	0.007 ± 0.098
18:3, n-6	0.000	0.003 ± 0.052	0.000	0.009 ± 0.034
20:0	0.000	0.033 ± 0.031	0.000	0.002 ± 0.026
20:1	0.000	0.000 ± 0.020	0.004	0.039 ± 0.030
20:2	0.000	0.023 ± 0.033	0.001	0.011 ± 0.023
20:3, n-3	0.000	0.003 ± 0.035	0.137	0.059 ± 0.098
20:3, n-6	0.000	0.006 ± 0.024	0.000	0.000 ± 0.034
20:4	0.000	0.004 ± 0.028	0.422	0.087 ± 0.048
22:0	0.000	0.168 ± 0.055	0.004	0.011 ± 0.028
22:1	0.000	0.036 ± 0.028	0.000	0.001 ± 0.029
23:0	0.000	0.032 ± 0.037	0.000	0.000 ± 0.032
24:0	0.000	0.000 ± 0.032	0.000	0.001 ± 0.025
Sum, %				
SFA	2.558	0.495 ± 0.085	0.392	0.030 ± 0.040
MUFA	2.267	0.461 ± 0.082	0.773	0.020 ± 0.038
PUFA	0.070	0.151 ± 0.052	0.035	0.001 ± 0.033
PUFA, n-3	0.004	0.040 ±0 .034	0.000	0.000 ± 0.030
PUFA, n-6	0.065	0.269 ± 0.067	0.918	0.019 ± 0.036
Ratio				
n-6:n-3	0.026	0.022 ± 0.029	0.000	0.000 ± 0.028
AI2	0.004	0.546 ± 0.082	0.000	0.000 ± 0.034

[1] Additional heritability estimates for traits not listed include ribeye area (REA, 0.34 ±0.073), marbling score (MARB, 0.656 ±0.087), yield grade (YG, 0.465 ±0.091), ether fat (EFAT, 0.459 ±0.113), and Warner-Bratzler shear force (WBSF, 0.194 ±0.052).

[2] AI=atherogenic index, weighted ratio of palmitic and myristic acids to total unsaturated fatty acids

Table 2.3. Estimates of genetic (above the diagonal) and phenotypic (below the diagonal) correlations between triacylglycerol fatty acid phenotypes and carcass traits in LM from Angus cattle obtained by multiple-trait REML analysis.

Trait	14:0	16:0	18:0	18:1-c	18:1-t	18:2	SFA	MUFA	PUFA	REA	MARB	YG	EFAT	WBSF
14:0		0.64 (0.08)	-0.30 (0.13)	-0.85 (0.06)	-0.23 (0.14)	-0.09 (0.14)	0.54 (0.10)	-0.63 (0.09)	-0.10 (0.18)	-0.09 (0.14)	0.08 (0.12)	0.13 (0.13)	0.17 (0.13)	-0.06 (0.16)
16:0	0.58 (0.02)		-0.19 (0.14)	-0.77 (0.08)	-0.50 (0.13)	-0.24 (0.14)	0.81 (0.05)	-0.81 (0.05)	-0.42 (0.17)	0.01 (0.15)	0.06 (0.14)	0.13 (0.14)	0.10 (0.14)	-0.29 (0.16)
18:0	-0.23 (0.03)	-0.13 (0.03)		0.12 (0.17)	-0.35 (0.15)	-0.22 (0.15)	0.39 (0.12)	-0.37 (0.13)	-0.16 (0.20)	0.01 (0.16)	-0.31 (0.14)	0.20 (0.15)	-0.22 (0.15)	-0.14 (0.18)
18:1-c1	-0.60 (0.02)	-0.53 (0.02)	-0.17 (0.03)		0.00 (0.19)	-0.15 (0.17)	-0.70 (0.09)	0.72 (0.08)	0.01 (0.23)	0.06 (0.18)	0.19 (0.16)	-0.43 (0.14)	-0.01 (0.17)	0.05 (0.20)
18:1-t2	-0.01 (0.03)	-0.23 (0.03)	-0.13 (0.03)	-0.50 (0.02)		0.69 (0.11)	-0.70 (0.11)	0.64 (0.12)	0.70 (0.14)	0.08 (0.17)	-0.12 (0.16)	0.32 (0.15)	0.05 (0.17)	0.25 (0.19)
18:2	-0.02 (0.03)	-0.18 (0.03)	-0.11 (0.03)	-0.28 (0.03)	0.48 (0.02)		-0.34 (0.14)	0.18 (0.15)	0.97 (0.05)	0.10 (0.16)	-0.15 (0.15)	0.08 (0.15)	-0.04 (0.16)	0.03 (0.19)
SFA	0.48 (0.02)	0.75 (0.01)	0.52 (0.02)	-0.64 (0.02)	-0.25 (0.03)	-0.20 (0.03)		-0.99 (0.01)	-0.49 (0.17)	-0.04 (0.15)	-0.11 (0.14)	0.31 (0.14)	-0.03 (0.15)	-0.30 (0.17)
MUFA	-0.51 (0.02)	-0.68 (0.02)	-0.50 (0.02)	0.76 (0.01)	0.12 (0.03)	-0.02 (0.03)	-0.95 (0.00)		0.36 (0.20)	0.05 (0.16)	0.12 (0.14)	-0.37 (0.01)	0.13 (0.01)	0.24 (0.05)
PUFA	0.05 (0.03)	-0.29 (0.02)	-0.08 (0.03)	-0.26 (0.02)	0.39 (0.02)	0.70 (0.01)	-0.26 (0.02)	-0.06 (0.03)		0.11 (0.21)	-0.20 (0.20)	0.02 (0.21)	-0.16 (0.21)	0.27 (0.23)
REA3	-0.01 (0.03)	-0.05 (0.03)	-0.03 (0.03)	0.02 (0.03)	0.07 (0.03)	0.07 (0.03)	-0.07 (0.03)	0.05 (0.03)	0.05 (0.03)		0.05 (0.15)	-0.71 (0.09)	-0.07 (0.16)	-0.09 (0.19)
MARB4	0.09 (0.03)	0.05 (0.03)	-0.16 (0.03)	0.06 (0.03)	-0.04 (0.03)	-0.11 (0.03)	-0.05 (0.03)	0.06 (0.03)	-0.07 (0.03)	0.00 (0.03)		-0.26 (0.14)	0.98 (0.02)	-0.50 (0.12)
YG5	0.08 (0.03)	0.11 (0.03)	-0.02 (0.03)	-0.05 (0.03)	0.01 (0.03)	-0.10 (0.03)	0.09 (0.03)	0.13 (0.03)	-0.06 (0.03)	-0.49 (0.02)	0.16 (0.03)		-0.24 (0.16)	0.14 (0.19)
EFAT6	0.14 (0.03)	0.05 (0.03)	-0.07 (0.03)	0.01 (0.03)	-0.02 (0.03)	-0.13 (0.03)	0.01 (0.03)	0.15 (0.03)	-0.08 (0.03)	-0.07 (0.03)	0.75 (0.01)	0.25 (0.03)		-0.32 (0.17)
WBSF7	-0.05 (0.03)	-0.08 (0.03)	0.02 (0.03)	0.02 (0.03)	0.05 (0.03)	0.04 (0.03)	-0.05 (0.03)	0.18 (0.03)	0.03 (0.03)	0.04 (0.03)	-0.23 (0.02)	0.00 (0.03)	-0.19 (0.03)	

[1] 18:1-*cis*
[2] 18:1-*trans*
[3] Ribeye area
[4] Numerical marbling score
[5] Calculated yield grade
[6] Ether fat percentage
7 Warner-Bratzler shear force (kg)

CHAPTER III

DERIVING GENE NETWORKS FROM SNP ASSOCIATED WITH TRIACYLGLYCEROL AND PHOSPHOLIPID FATTY ACID FRACTION FROM RIBEYES OF ANGUS CATTLE

Justin W Buchanan[1], James M Reecy[2], Dorian J Garrick[2], Qing Duan[2], Don C Beitz[2], James E. Koltes[2], Mahdi Saatchi[2], Lars Koesterke[3], and Raluca G Mateescu[4]

[1]Department of Animal Science, Oklahoma State University, Stillwater, OK, USA

[2]Department of Animal Science, Iowa State University, Ames, IA, USA

[3]Texas Advanced Computing Center, University of Texas, Austin, TX, USA

[4]Department of Animal Sciences, University of Florida, Gainesville, FL, USA

ABSTRACT

The fatty acid profile of beef is a complex trait that can benefit from a gene-interaction network analysis to understand the functional relationship among loci that contribute to phenotypic variation. Phenotypic measures of fatty acid profile from the triacylglycerol and phospholipid fraction of longissimus muscle, pedigree information, and Illumina 54k bovine SNPchip genotypes were utilized to derive an annotated gene network that controls fatty acid composition in 1,833 Angus beef cattle. The Bayes-B statistical model was utilized to perform a genome wide association study to estimate associations between 54k SNP genotypes and 39 individual fatty acid phenotypes within each fraction. Effects were estimated for 1-Mb genomic windows as well as for 54k SNP. Windows that explained the majority of genetic variance in lipids from the phospholipid fraction exhibited almost no overlap with those from the triacylglycerol fraction. Partial correlations were used to identify correlated regions of the genome for that set of largest 1 Mb windows that explained up to 35% genetic variation in both fatty acid fractions. SNP within those windows were annotated based on the bovine UMD3.1 assembly. Gene network clusters were generated utilizing a partial correlation and information theory algorithm. Results were used in conjunction with network scoring and visualization software to analyze correlated SNP across 39 fatty acid phenotypes to identify SNP of functional significance. Networks derived from partial correlation analysis captured up to 67.9% of the genetic variance explained by all SNPs. Significant pathways implicated in fatty acid metabolism through GO term enrichment analysis included homeostasis of number of cells, homeostatic process, coenzyme/cofactor activity, and protein kinase activity. Network analysis using partial correlations and

annotation of significant SNPs can yield functional information about genetic mechanisms controlling associated phenotypes.

INTRODUCTION

Beef is a nutritious source of protein, fat, vitamins, and minerals when appropriately included in the human diet. A large body of research exists that suggests it is critical to maintain a balanced fatty acid intake to support a healthy blood lipid profile (Ooi et al., 2013). The synthesis of adipose tissue in beef cattle is a complex biological process controlled by numerous genetic loci as well as environmental factors. Considerable ongoing effort has been devoted to the identification of these genetic loci as well as candidate genes for fatty acid profile and adipose synthesis in various breeds of beef cattle (Barendse, 2011;Cesar et al., 2014;Kelly et al., 2014). The usefulness of these loci in DNA based beef cattle selection schemes will increase with knowledge of genomic architecture for fat deposition (Saatchi et al., 2013).

The fatty acid profile in beef cattle can be characterized by the abundance of about 40 individual lipids of varying chain lengths and degrees of saturation (Daley et al., 2010). Additionally, the total lipid fraction present in animal tissues can be separated into triacylglycerol and phospholipid fractions, which represent the two primary modes of lipid storage in cattle muscular tissue (Yen et al., 2008). The triacylglycerol fraction can typically represent from 70% to 92% of the total lipid fraction in longissimus muscle depending on age and dietary composition (Warren et al., 2008). Taken together, these data present a large number of phenotypes to consider at once. Current research into multiple-trait analyses utilizing GBLUP and Bayesian methods is ongoing (Jia and Jannink, 2012;Gao et al., 2014).

While multiple-trait methodology is being developed, there has been an interest in other computational methods that utilize single trait GWAS along with the principles of co-association across multiple phenotypes to develop clusters of SNP that have a large association to the overall trait of interest (Reverter and Chan, 2008;Fortes et al., 2011;Reverter and Fortes, 2013).

Correlations among fatty acids within and across lipid fractions are known (Hoehne et al., 2012), and genetic parameter estimates and correlations have been obtained for the data set used in this study (unpublished). According to Reverter and Fortes (2013) correlations among multiple phenotypes can be exploited to develop an association weight matrix that utilizes high-throughput data such as that obtained from GWAS. This methodology utilizes a partial correlation and information theory algorithm (PCIT) to analyze an input matrix that contains data from SNP across multiple phenotypes to generate clusters of loci that are highly associated with the overall trait of interest (Reverter and Chan, 2008). The usefulness of this analysis has been previously demonstrated through derivation of a regulatory gene network associated with puberty in beef cattle (Fortes et al., 2011). The objective of this study was to utilize the PCIT algorithm along with GWAS output from 39 different lipid classes from both the triacylglycerol and phospholipid fractions to derive gene networks associated with the fatty acid profile in Angus beef cattle.

MATERIALS AND METHODS

Animals and sample collection

A total of 1,833 offspring of 155 Angus bulls represented by bulls (n = 450), steers (n = 1,022), and heifers (n = 361) were used in this study. All cattle were finished on concentrate diets in Iowa (n = 908), California (n = 344), Colorado (n = 291), or Texas (n = 290). Animals were harvested at commercial facilities when they reached typical US market endpoints with an average age of 457 ± 46 days. Production characteristics and additional details of sample collection and preparation were reported previously (Garmyn et al., 2011). After external fat and connective tissue were removed, the 1.27-cm steaks from the longissimus muscle were analyzed for fatty acid composition at Iowa State University (Ames, IA), using methods previously described (Zhang et al., 2008).

Genome-wide association study of fatty acid fractions

Genomic DNA was extracted from the ground beef sample used for fatty acid composition and was genotyped with the Bovine SNP50 Infinium II BeadChip (Illumina, San Diego, CA). Contemporary groups were defined based on cross-classifications of gender at harvest (bull, steer or heifer), finishing location (California, Colorado, Iowa, Texas), and harvest date, for a total of 33 groups. Contemporary groups were fit as fixed effects in genomic analyses. Effects of SNP on each trait were estimated using the Bayes-B option of GenSel accessed through the BIGSGUI Version 0.9.2 (Kizilkaya et al., 2010). The Markov-chain Monte Carlo approach used to estimate the effect of each SNP involved a 1,000 iteration burn-in period followed by 40,000 iterations used to obtain posterior means of the effect of each SNP. The estimate of the proportion of genetic variation explained by each SNP and each 1 Mb window was obtained for all 39 fatty acid phenotypes for triacylglycerol and phospholipid

fractions. The windowBV yes option was used to form the posterior distribution for every 1 Mb window.

Correlations among SNP

SNP-wise correlations among 39 fatty acid phenotypes were determined using the PCIT algorithm optimized for use through the Texas Advanced Computing Center (Reverter and Chan, 2008;Koesterke et al., 2013). For the triacylglycerol fraction an initial set of 389 SNP were selected from the 17 genomic windows of size 1 Mb with the largest posterior mean for variance accounted by the window and posterior probability of association greater than 0.90 ($P < 0.05$) from the 16:0 fatty acid phenotype. A vector of posterior mean SNP effects for 389 SNP from 16:0 was augmented with the effects of all 39 fatty acid phenotypes. This 389 x 39 matrix of posterior mean SNP effects was used as the input for the PCIT algorithm to detect similar effects for any SNP across multiple fatty acids. All SNP pairs within the matrix were tested for association with at least one other SNP in order to establish network connections. SNP pairs without a significant partial correlation to at least one other SNP were removed from the dataset and not used for subsequent network association analysis since they would appear isolated.

To build a matrix of SNP effects for the phospholipid fraction SNP were selected from 20 genomic windows of size 1 Mb with the largest posterior mean for variance accounted by the window and posterior probability of association greater than 0.3 ($P < 0.05$) from the 16:0 fatty acid phenotype. The threshold for the posterior probability of association was relaxed to 0.3 in order to capture a similar number of genomic windows containing SNP for network creation. A vector of posterior mean SNP effects for 571 SNP from 16:0 was augmented with

the effects of all 39 fatty acid phenotypes. PCIT network creation and visualization proceeded identically to the methods described for the triacylglycerol fraction.

Correlations among SNP were used to visualize networks of SNP that exhibited a common effect across multiple fatty acid phenotypes. Correlation between SNP pairs with a non-zero partial correlation to another SNP were input into Cytoscape 3.0.2 (Shannon et al., 2003) software to create gene network clusters using the MCODE plugin (Bader and Hogue, 2003: Saito et al., 2012). Networks are scored and ranked by the MCODE algorithm as network density times the number of nodes. MCODE defines network density as the number of edges in a network divided by the theoretical maximum number of edges. SNP were annotated using the Bovine UMD 3.1 annotations (McLaren et al., 2010) accessed from the cattle genome analysis data repository (Koltes, 2012).

Gene Ontology Enrichment Analysis and Visualization

Gene ontology (GO) enrichment was carried out using the DAVID v6.7 Functional Annotation Tool (Huang da et al., 2009b;a) in order to identify enriched biological terms associated with genomic regions and gene networks identified in the analysis. The resulting GO biological function terms were then condensed and visualized using REVIGO (Supek et al., 2011). GO term enrichment was first carried out on all genes within 1 Mb genomic windows associated with the triacylglycerol ($PPA > 0.90$) and phospholipid ($PPA > 0.5$) fractions. Ensembl Gene ID's were extracted from 1Mb genomic regions from the *Bos taurus* UMD3.1 assembly for use in the GO term enrichment analysis. Additionally, GO term

enrichment analysis was carried out for all genes within the networks generated using the PCIT clustering algorithm.

RESULTS AND DISCUSSION

Triacylglycerol GWAS

The estimates for genetic variance explained by 1 Mb genomic windows as well as the posterior probability of association for selected lipids and lipid classes from the triacylglycerol fraction are shown in **Table 3.1**. All posterior probabilities of association for the genomic windows displayed in Table 3.1 were greater than 90% (*PPA* > 0.90), which indicates that the false discovery rate is < 10%. This means most of these genomic regions likely harbor individual loci that exhibit a very large effect on the phenotype included in the analysis. Multiple genomic windows were identified which explained between 22.13% and 34.55% of genetic variance for individual lipids found in the triacylglycerol fraction. Variance and posterior estimates for the total fatty acid fraction from which triacylglycerol and phospholipid fractions were derived for this dataset were previously described (Saatchi et al., 2013). The genomic window on chromosome 19 between 51,148,913 and 51,956,162 appears to describe a large proportion of the genetic variance across multiple fatty acids in the triacylglycerol fraction, including 14:0, 16:0, 16:1, 18:0, 18:1, SFA, and MUFA. This region harbors the candidate gene fatty acid synthase (FASN), which is known to be associated with primary lipid synthesis in adipose (Zhang et al., 2008;Abe et al., 2009). It follows that this genomic window would explain a large proportion of genetic variance

across multiple fatty acids and fatty acid classes due to the triacylglycerol functioning as the main storage site for lipids of medium chain length synthesized from FASN (Wood et al., 2008).

Other genomic regions of significance that appeared in both this data set and the total fatty acid fraction analysis presented by Saatchi et al. (2013) include windows on chromosome 29 (from 18,046,673 to 18,882,323) and a region on chromosome 26 (2 Mb from 20012464 to 21977581). The region on chromosome 29 accounts for up to 10.65% of the genetic variance in 14:0, 16:0, 16:1, 18:0, and 18:1. The region on chromosome 26 accounts for up to 5.39% of the genetic variance in 14:0, 16:1, SFA, and MUFA. These regions also harbor candidate genes related to fatty acid synthesis and metabolism including THRSP and SCD, respectively, as noted by Saatchi et al. (2013).

Several genomic regions of interest were found to explain relatively large proportions of genetic variance that were not detected in the data presented by Saatchi et al. (2013). A window on chromosome 17 from 16,003,681 to 16,985,065 accounted for 22.91% genetic variance in 16:0. This region harbors the possible candidate gene inositol polyphosphate-4-phosphatasa, type II (INPP4B). Cellular localization for INPP4B is in the cytoplasm, and the top Gene Ontology biological process entries for this gene include phospholipid metabolic process, and known associations in bovine indicate a larger role in bone remodeling (http://amigo.geneontology.org/amigo). No QTL have been reported in the Cattle QTL Database (http://www.animalgenome.org/cgi-bin/QTLdb/BT/index) in this region related to fat or fatty acid content. A previously unidentified region on chromosome 8 from 103,028,846 to 103,941,108 accounted for 6.43% genetic variance in 18:1, and a second novel region on chromosome 7 from 56,022,790 to 56,989,949 accounted for 16.26% genetic

variance in 16:0. No candidate genes were identified within these regions, nor were they reported in the Cattle QTL Database related to fatness or fatty acid metabolism.

Phospholipid GWAS

The posterior mean estimates for genetic variance explained by 1 Mb genomic windows as well as the posterior probabilities of association for selected lipids and lipid classes from the phospholipid fraction are shown in **Table 3.2**. In contrast to the triacylglycerol fraction, posterior estimates for individual lipids and lipid classes were lower, with only 10 windows displaying a posterior probability association greater than 50% (*PPA* > 0.50). **Figure 1** displays a Manhattan plot comparison of the distribution of genetic variance estimates for 1 Mb genomic windows between the triacylglycerol (plot A) and phospholipid fraction (plot B). The relatively larger proportion of genetic variance explained by fewer 1 Mb windows is visually displayed in the plots. Windows that explained the majority of genetic variance in lipids from the phospholipid fraction exhibited almost no overlap with those from the triacylglycerol fraction.

Several windows identified harbor candidate genes related to overall lipid and phospholipid metabolism. The genomic window on chromosome 16 from 4,021,893 to 4,966,340 that accounted for 1.79% of the genetic variance in 14:0 harbors the candidate gene fructose-2,6-biphosphatase 2 (PFKFB2). This gene is known to be involved in synthesis and degradation of fructose-2.,6-bisphosphate, which is a regulatory molecule involved in glycolysis in eukaryotes (Hue and Rider, 1987). A QTL that spanned this region was identified previously in Angus in relation to 12th rib fat thickness (McClure et al., 2010). Also, the genomic window on chromosome 24 from 29,013,292 to 29,942,533 accounted for

4.27% of genetic variance in 18:0 harbors the candidate gene N-cadherin (CDH2). This gene is known to be involved in cell-to-cell adhesion and has been associated with increased adipogenic proliferation in mice (Castro et al., 2004). Other novel genomic windows, including 1 Mb upstream and downstream from the identified regions, did not harbor any genes of interest related to fatty acid or membrane metabolism.

Triacylglycerol and phospholipid gene networks

Of the 389 SNP entered into the PCIT analysis for the triacylglycerol fraction, 355 were co-localized into 12 separate networks. Information detailing network scoring results from the Cytoscape MCODE plugin and proportion of genetic variance explained by each network for all SNP in 16:0 in each triacylglycerol-derived network is presented in **Table 3.3**. The two highest scoring SNP networks for triacylglycerol fraction obtained from PCIT output and subsequent visualization of nodes with Cytoscape are displayed in **Figures 2** and **3**. Networks are scored and ranked by the MCODE algorithm as network density times the number of nodes. MCODE defines network density as the number of edges in a network divided by the theoretical maximum number of edges. Nodes that were not annotated to a gene or feature were removed from the figures for visual simplicity. Location within the network indicates significance of each node, with distance from the center indicating the number of overall connections and importance to the phenotype. Each edge represents a connection, or direct correlation, identified through PCIT analysis. **Figure 2** displays the highest scoring triacylglycerol sub-network obtained by network scoring with the Cytoscape MCODE plugin. The highest scoring network contained 55 nodes and 1438 edges, or connections. The SNP captured in this network also explained up to 67.9% of the total genetic variance explained by all SNP in 16:0 in the triacylglycerol fraction. **Figure 3**

displays the second highest scoring annotated triacylglycerol network obtained with MCODE, which contained 94 nodes and 1792 edges. The SNP captured in this network also explained up to 29.8% of the total genetic variance explained by all SNP in 16:0 in the triacylglycerol fraction. The clusters of genes represent scored networks derived in combination with the PCIT algorithm that function as molecular complexes related to the input phenotype.

Candidate genes involved in fatty acid metabolism found within these networks include thyroid responsive hormone (THRSP), Acyl-CoA synthetase-5 (ACSL5), glycerol-3-phosphate acyltransferase muscle-type (GPAM), and coiled coil domain-containing 3 (CCD57). The candidate genes THRSP and GPAM have been previously identified as playing a role in lipid metabolism in beef and dairy adipose via the PPAR pathway (Graugnard et al., 2010;Ji et al., 2014;Moisa et al., 2014). ACSL5 is found primarily in cells with a high triacylglycerol synthesis activity, which suggests a likely role in development of adipose (Bu and Mashek, 2010). CCD57 is known to be involved in DNA binding and regulation of gene expression. This gene is located next to FASN on chromosome 19, and has been previously associated with 14:0 content and transcripts have been detected in excess of FASN transcripts in second-lactation dairy cattle (Bouwman et al., 2014;Canovas et al., 2014). Overall, this methodology presents strong evidence that functionally relevant genes can be co-localized with a close relationship to triacylglycerol variation and assembly.

Visualizations of the two highest scoring phospholipid networks with Cytoscape are shown in **Figures 4** and **5**. The network in **Figure 4** is the highest scoring phospholipid network containing 47 nodes and 780 edges. This network accounted for 9.42E-04% of the total genetic variance explained by all SNP in 16:0. **Figure 5** displays the second highest

scoring phospholipid network which contained 87 nodes and 1264 edges. Multiple genes identified within these networks are known to be involved in cellular trafficking and cell integrity functions associated with the phospholipid membrane. Multiple genes that are known to be involved in membrane binding and cellular trafficking were identified including myosin-IXB (MYO9B), FCH domain only protein 1 (FCHO1), and ADAM metallopeptidase domain (ADAM11). The identification of large sets of genes related to phospholipid metabolism for the phospholipid fraction is difficult due to the low variance observed in the initial phenotype. Low variance estimates in the phenotype also account for the low estimates of genetic variance explained by the genomic windows. It is known that traits with a low genetic variance should produce fewer biologically related genes responsible for genetic variation in the trait of interest, and these results reflect that relationship.

Gene Ontology Term Enrichment Analysis

Gene ontology term enrichment analysis was first carried out for all genes located in the top 1 Mb regions within **Table 3.1** and **Table 3.2** using the DAVID Functional Classification tool. Genes were obtained by extracting ENSEMBL Gene ID features from the regions of interest. Significant results for the DAVID Functional Annotation Clustering and Functional Annotation Chart results for the top GWAS regions for both fractions are located in **Tables 3.4** through **3.10**. The top GO term clusters for the top 1 Mb windows associated with the triacylglycerol fraction are shown in **Table 3.4**. DAVID Functional Annotation Clusters are considered significant above an enrichment score of 1.3, and GO terms are considered significantly enriched at a P-Value of 0.05 or less (Huang da et al., 2009b).

Two clusters were produced for the triacylglycerol fraction with an enrichment score above 1.3. The full list of GO terms is displayed in **Table 3.5**. Significant GO terms featured in the Functional Clusters and Functional Annotation Chart included homeostatic process (GO:0042592), homeostasis of number of cells (GO:0048872), dendrite development (GO:0016358), and activation of MAP kinase activity (GO:0000187). These terms appear to be associated with features relating to cellularity and energy homeostasis pathways, which have relevance to fatty acid deposition and adipogenesis. The full list of GO terms produced in the Functional Annotation Chart was reduced and visualized in **Figure 6** using the REVIGO software. Darker color indicates significance of P-value obtained through DAVID clustering. Connections indicate related biological process terms, and size represents frequency of the GO term in the Gene Ontology Annotation Database (UniProt-GOA). The link between adipose tissue cellularity and fatty acid profile has been previously established (Costa et al., 2012). These results highlight the genetic involvement of cellular homeostasis in triacylglycerol metabolism.

Functional Annotation Clustering analysis for the top GWAS regions associated with phospholipid is located in **Table 3.6** and the full list of GO terms from the Functional Annotation Chart is located in **Table 3.7**. The full list of GO terms was reduced and visualized in **Figure 7**. Only one significant GO term cluster was identified for the phospholipid regions with an annotation cluster score above 1.3. Top significant enriched GO terms included protein serine/threonine kinase activity (GO:0004674), immunoglobulin mediated immune response (GO:0002455), and plasma protein inflammatory response (GO:0002541). There is evidence for a link between prolonged protein kinase activation and a cellular signaling cascade that may result in the degradation of lipid membrane constituents

(Nishizuka, 1995). However, without further evidence it is not immediately apparent how this term enrichment relates to lipid membrane metabolism. There also appear to be a large number of enrich terms associated with immune response pathways, suggesting a possible role in unsaturated lipid signaling in these processes.

Gene ontology enrichment analysis results from the Functional Annotation Chart for the SNP captured in the triacylglycerol networks is located in **Table 3.8**. The full list of GO terms was reduced and visualized in **Figure 8**. There were no significant Functional Annotation Clusters identified for the Genes captured in the SNP networks for the triacylglycerol fraction. The significant enriched GO terms for the SNP captured in the triacylglycerol gene networks included aspartic-type endopeptidase and peptidase activity (GO:0004190, GO:0070001), and other peptidase and proteolysis terms (GO:0070011, GO:0008233, GO:0006508). A class of peptidases known as Cathepsins have been previously characterized in beef muscle tissues and are known to be involved in amino acid degradation in muscle tissues (Bolumar et al., 2014). A link between adipose accumulation and Cathepsin D-activated cell death in a human study (Eguchi and Feldstein, 2003) supports the conclusion that triacylglycerol accumulation and increased peptidase activity are associated. This association further supports the previously identified link between triacylglycerol accumulation and cellular homeostasis GO terms.

Functional Annotation Clustering analysis for the SNP captured in the phospholipid networks is located in **Table 3.9**. There was only one Functional Annotation Cluster with a significant enrichment score over 1.3 for the SNP captured in the phospholipid networks. The full list of GO terms was reduced and visualized in **Figure 9**. The significant GO term associations in **Table 3.10** from the Functional Annotation Chart included coenzyme

metabolic process (GO:0006732, GO:0009108) and cofactor biosynthetic process (GO:0051186, GO:0051188). Coenzyme activity of flavine adenine dinucleotide (FAD) is known to be critical in the synthesis of ether phospholipids from alkyldihydroxyacetonephosphate (alkyl-DHAP), which catalyzes the formation of the ether bond (Razeto et al., 2007). Ether phospholipids are known to be critical components of the phospholipid membrane in eukaryotes (Clark et al., 2014), but it is difficult to determine their abundance in phospholipid samples analyzed using gas chromatography due to their structure. It is likely that multiple coenzyme factors play a larger role in synthesis of phospholipid membrane components.

CONCLUSION

Analysis of GWAS results for triacylglycerol and phospholipid fractions appears to support the conclusion that the triacylglycerol fraction is closely representative of the total fatty acid fraction. Significant genomic windows identified highly overlapped with the results presented by (Saatchi et al., 2013). These results are supported by the fact that the phenotypic measurements of the triacylglycerol fraction mirror the percentages of the total fatty acid fraction. It follows that the genomic regions that likely harbor genes and features related to the total fatty acid fraction would also be identified when just the triacylglycerol fraction is analyzed. The triacylglycerol fraction also exhibits a much larger genetic variance when compared to the phospholipid fraction. This methodology provides a set of markers associated with intramuscular adipose accumulation as well as an enriched set of biological functions representative of fatty acid deposition in beef cattle.

An analysis of the genomic regions that affect the phospholipid fraction yielded few genes with a known biological association to lipid metabolism. Significant genomic regions identified explained lower percentages of genetic variance in comparison to the triacylglycerol. The low variation observed in the phospholipid fraction is likely due to the importance of the phospholipid membrane to biological function of the cell. Pathways prevalent in the phospholipid analysis appeared to be highly related to cell-to-cell adhesion, cellular trafficking, and coenzyme/cofactor activity. A larger dataset could possibly improve results when dealing with traits that exhibit a low phenotypic variance. In conclusion, the combination of GWAS results with the PCIT algorithm and network visualization represents a robust methodology for identifying candidate genes of interest for traits with multiple phenotypes and adequate phenotypic variance.

LITERATURE CITED

Abe, T., Saburi, J., Hasebe, H., Nakagawa, T., Misumi, S., Nade, T., Nakajima, H., Shoji, N., Kobayashi, M., and Kobayashi, E. (2009). Novel mutations of the FASN gene and their effect on fatty acid composition in Japanese Black beef. *Biochem Genet* 47**,** 397-411. doi: 10.1007/s10528-009-9235-5.

Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4**,** 2.

Barendse, W. (2011). Haplotype analysis improved evidence for candidate genes for intramuscular fat percentage from a genome wide association study of cattle. *PLoS One* 6**,** e29601. doi: 10.1371/journal.pone.0029601.

Bolumar, T., Bindrich, U., Toepfl, S., Toldra, F., and Heinz, V. (2014). Effect of electrohydraulic shockwave treatment on tenderness, muscle cathepsin and peptidase activities and microstructure of beef loin steaks from Holstein young bulls. *Meat Sci* 98, 759-765. doi: 10.1016/j.meatsci.2014.07.024.

Bouwman, A.C., Visker, M.H., Van Arendonk, J.M., and Bovenhuis, H. (2014). Fine mapping of a quantitative trait locus for bovine milk fat composition on Bos taurus autosome 19. *J Dairy Sci* 97**,** 1139-1149. doi: 10.3168/jds.2013-7197.

Bu, S.Y., and Mashek, D.G. (2010). Hepatic long-chain acyl-CoA synthetase 5 mediates fatty acid channeling between anabolic and catabolic pathways. *J Lipid Res* 51**,** 3270-3280. doi: 10.1194/jlr.M009407.

Canovas, A., Rincon, G., Bevilacqua, C., Islas-Trejo, A., Brenaut, P., Hovey, R.C., Boutinaud, M., Morgenthaler, C., Vanklompenberg, M.K., Martin, P., and Medrano,

J.F. (2014). Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. *Sci Rep* 4**,** 5297. doi: 10.1038/srep05297.

Castro, C.H., Shin, C.S., Stains, J.P., Cheng, S.L., Sheikh, S., Mbalaviele, G., Szejnfeld, V.L., and Civitelli, R. (2004). Targeted expression of a dominant-negative N-cadherin in vivo delays peak bone mass and increases adipogenesis. *J Cell Sci* 117**,** 2853-2864. doi: 10.1242/jcs.01133.

Cesar, A.S., Regitano, L.C., Tullio, R.R., Lanna, D.P., Nassu, R.T., Mudado, M.A., Oliveira, P.S., Do Nascimento, M.L., Chaves, A.S., Alencar, M.M., Sonstegard, T.S., Garrick, D.J., Reecy, J.M., and Coutinho, L.L. (2014). Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genet* 15**,** 39. doi: 10.1186/1471-2156-15-39.

Clark, J., Kay, R.R., Kielkowska, A., Niewczas, I., Fets, L., Oxley, D., Stephens, L.R., and Hawkins, P.T. (2014). Dictyostelium uses ether-linked inositol phospholipids for intracellular signalling. *EMBO J* 33, 2188-2200. doi: 10.15252/embj.201488677.

Daley, C.A., Abbott, A., Doyle, P.S., Nader, G.A., and Larson, S. (2010). A review of fatty acid profiles and antioxidant content in grass-fed and grain-fed beef. *Nutr J* 9**,** 10. doi: 10.1186/1475-2891-9-10.

Eguchi, A., and Feldstein, A.E. (2013). Lysosomal Cathepsin D contributes to cell death during adipocyte hypertrophy. *Adipocyte* 2, 170-175. doi: 10.4161/adip.24144.

Fortes, M.R., Reverter, A., Nagaraj, S.H., Zhang, Y., Jonsson, N.N., Barris, W., Lehnert, S., Boe-Hansen, G.B., and Hawken, R.J. (2011). A single nucleotide polymorphism-

derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J Anim Sci* 89**,** 1669-1683. doi: 10.2527/jas.2010-3681.

Gao, H., Zhang, T., Wu, Y., Wu, Y., Jiang, L., Zhan, J., Li, J., and Yang, R. (2014). Multiple-trait genome-wide association study based on principal component analysis for residual covariance matrix. *Heredity (Edinb)*. doi: 10.1038/hdy.2014.57.

Garmyn, A.J., Hilton, G.G., Mateescu, R.G., Morgan, J.B., Reecy, J.M., Tait, R.G., Jr., Beitz, D.C., Duan, Q., Schoonmaker, J.P., Mayes, M.S., Drewnoski, M.E., Liu, Q., and Vanoverbeke, D.L. (2011). Estimation of relationships between mineral concentration and fatty acid composition of longissimus muscle and beef palatability traits. *J Anim Sci* 89**,** 2849-2858. doi: 10.2527/jas.2010-3497.

Graugnard, D.E., Berger, L.L., Faulkner, D.B., and Loor, J.J. (2010). High-starch diets induce precocious adipogenic gene network up-regulation in longissimus lumborum of early-weaned Angus cattle. *Br J Nutr* 103**,** 953-963. doi: 10.1017/S0007114509992789.

Hoehne, A., Nuernberg, G., Kuehn, C., and Nuernberg, K. (2012). Relationships between intramuscular fat content, selected carcass traits, and fatty acid profile in bulls using a F2-population. *Meat Sci* 90**,** 629-635. doi: 10.1016/j.meatsci.2011.10.005.

Huang Da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37**,** 1-13. doi: 10.1093/nar/gkn923.

Huang Da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4**,** 44-57. doi: 10.1038/nprot.2008.211.

Hue, L., and Rider, M.H. (1987). Role of fructose 2,6-bisphosphate in the control of glycolysis in mammalian tissues. *Biochem J* 245**,** 313-324.

Ji, P., Drackley, J.K., Khan, M.J., and Loor, J.J. (2014). Overfeeding energy upregulates peroxisome proliferator-activated receptor (PPAR)gamma-controlled adipogenic and lipolytic gene networks but does not affect proinflammatory markers in visceral and subcutaneous adipose depots of Holstein cows. *J Dairy Sci* 97**,** 3431-3440. doi: 10.3168/jds.2013-7295.

Jia, Y., and Jannink, J.L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* 192**,** 1513-1522. doi: 10.1534/genetics.112.144246.

Kelly, M.J., Tume, R.K., Fortes, M., and Thompson, J.M. (2014). Whole-genome association study of fatty acid composition in a diverse range of beef cattle breeds. *J Anim Sci* 92**,** 1895-1901. doi: 10.2527/jas.2013-6901.

Kizilkaya, K., Garrick, D.J., Fernando, R.L., Mestav, B., and Yildiz, M.A. (2010). Use of linear mixed models for genetic evaluation of gestation length and birth weight allowing for heavy-tailed residual effects. *Genet Sel Evol* 42**,** 26. doi: 10.1186/1297-9686-42-26.

Koltes, J. (2012). ISU Annotation of Cattle 770k HD SNP and 54k SNP Data.

Koesterke, L., Milfield, K., Vaughn, M.W., Stanzione, D., Koltes, J.E., Weeks, N.T., Reecy, J.M. (2013). Optimizing the PCIT algorithm on stampede's Xeon and Xeon Phi processors for faster discovery of biological networks. Proceedings: XSEDE 13. doi: 10.1145/2484762.2484794.

Mcclure, M.C., Morsci, N.S., Schnabel, R.D., Kim, J.W., Yao, P., Rolf, M.M., Mckay, S.D., Gregg, S.J., Chapple, R.H., Northcutt, S.L., and Taylor, J.F. (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. doi: 10.1111/j.1365-2052.2010.02063.x.

Mclaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 26**,** 2069-2070. doi: 10.1093/bioinformatics/btq330.

Moisa, S.J., Shike, D.W., Faulkner, D.B., Meteer, W.T., Keisler, D., and Loor, J.J. (2014). Central Role of the PPARgamma Gene Network in Coordinating Beef Cattle Intramuscular Adipogenesis in Response to Weaning Age and Nutrition. *Gene Regul Syst Bio* 8**,** 17-32. doi: 10.4137/GRSB.S11782.

Nishizuka, Y. (1995). Protein kinase C and lipid signalling for sustained cellular responses. *FASEB J* 9, 484-496.

Ooi, E.M., Ng, T.W., Watts, G.F., and Barrett, P.H. (2013). Dietary fatty acids and lipoprotein metabolism: new insights and updates. Curr Opin Lipidol 24, 192-197. doi: 10.1097/MOL.0b013e3283613ba2.

Razeto, A., Mattiroli, F., Carpanelli, E., Aliverti, A., Pandini, V., Coda, A., and Mattevi, A. (2007). The crucial step in ether phospholipid biosynthesis: structural basis of a noncanonical reaction associated with a peroxisomal disorder. Structure 15, 683-692. doi:10.1016/j.str.2007.04.009

Reverter, A., and Chan, E.K. (2008). Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. *Bioinformatics* 24**,** 2491-2497. doi: 10.1093/bioinformatics/btn482.

Reverter, A., and Fortes, M.R. (2013). Association weight matrix: a network-based approach towards functional genome-wide association studies. *Methods Mol Biol* 1019**,** 437-447. doi: 10.1007/978-1-62703-447-0_20.

Saatchi, M., Garrick, D.J., Tait, R.G., Jr., Mayes, M.S., Drewnoski, M., Schoonmaker, J., Diaz, C., Beitz, D.C., and Reecy, J.M. (2013). Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. *BMC Genomics* 14**,** 730. doi: 10.1186/1471-2164-14-730.

Saito, R., Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., Lotia, S., Pico, A.R., Bader, G.D., and Ideker, T. (2012). A travel guide to Cytoscape plugins. *Nat Methods* 9**,** 1069-1076. doi: 10.1038/nmeth.2212.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13**,** 2498-2504. doi: 10.1101/gr.1239303.

Supek, F., Bosnjak, M., Skunca, N., and Smuk, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* 6, e21800. doi:10.1371/journal.pone.0021800

Warren, H.E., Scollan, N.D., Enser, M., Hughes, S.I., Richardson, R.I., and Wood, J.D. (2008). Effects of breed and a concentrate or grass silage diet on beef quality in cattle

of 3 ages. I: Animal performance, carcass quality and muscle fatty acid composition. Meat Sci 78, 256-269. doi: 10.1016/j.meatsci.2007.06.008.

Wood, J.D., Enser, M., Fisher, A.V., Nute, G.R., Sheard, P.R., Richardson, R.I., Hughes, S.I., and Whittington, F.M. (2008). Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci* 78**,** 343-358. doi: 10.1016/j.meatsci.2007.07.019.

Yen, C.L., Stone, S.J., Koliwad, S., Harris, C., and Farese, R.V., Jr. (2008). Thematic review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *J Lipid Res* 49**,** 2283-2301. doi: 10.1194/jlr.R800018-JLR200.

Zhang, S., Knight, T.J., Reecy, J.M., and Beitz, D.C. (2008). DNA polymorphisms in bovine fatty acid synthase are associated with beef fatty acid composition. *Anim Genet* 39**,** 62-70. doi: 10.1111/j.1365-2052.2007.01681.

Table 3.1. Characterization of windows that account for variation in lipids in triacylglycerol. For each lipid or lipid class the posterior mean of the genetic variance explained by the 1 Mb window is given along with window position coordinates (UMD3.1), number of SNP, and posterior probability of association (*PPA*).

Trait	Map Start	Map End	RS# Start	RS# End	# of SNP	Genetic Variance (%)	PPA
14:0							
	19_51148913	19_51956162	rs41923412	rs109147235	25	34.55	1.00
	29_18046673	29_18882323	rs42375315	rs41589183	14	10.65	1.00
	10_19017657	10_19987360	rs41647457	rs110785951	24	3.21	0.98
	19_53038373	19_53963109	rs110146710	rs41577620	25	1.86	0.92
	18_18050574	18_18997878	rs110528295	rs110871891	25	1.55	0.99
	26_21023960	26_21977581	rs109309604	rs42086690	20	1.53	0.99
	25_34025400	25_34984182	rs110966408	rs109749619	33	1.13	0.95
16:0							
	17_16003681	17_16985065	rs109550465	rs110684903	17	22.91	1.00
	7_56022790	7_56989949	rs41614823	rs42334377	17	16.26	1.00
	19_51148913	19_51956162	rs41923412	rs109147235	25	16.18	1.00
	1_80019442	1_80974985	rs43245574	rs110467946	21	10.44	1.00
	16_3035722	16_3987821	rs41790571	rs41633905	24	9.66	0.99
	29_18046673	29_18882323	rs42375315	rs41589183	14	5.87	1.00
	26_33003665	26_33962496	rs41606739	rs110568468	27	2.01	0.95
	13_31014861	13_31938349	rs29019775	rs41566146	21	1.65	0.96
16:1							
	19_51148913	19_51956162	rs41923412	rs109147235	25	15.25	1.00
	29_18046673	29_18882323	rs42375315	rs41589183	14	4.72	1.00
	26_21023960	26_21977581	rs109309604	rs42086690	20	3.23	1.00
	25_34025400	25_34984182	rs110966408	rs109749619	33	2.52	0.95
	19_35003592	19_35965938	rs109843005	rs43031950	27	1.59	0.94
	7_13002006	7_13934893	rs109277981	rs41255303	24	1.5	0.96

	25_32003625	25_32912638	rs41626335	rs110326707	22	1.47	0.94
18:0							
	29_18046673	29_18882323	rs42375315	rs41589183	14	5.95	1.00
	19_51148913	19_51956162	rs41923412	rs109147235	25	3.04	1.00
	19_35003592	19_35965938	rs109843005	rs43031950	27	2.46	0.95
18:1							
	19_51148913	19_51956162	rs41923412	rs109147235	25	22.13	1.00
	8_103028846	8_103941108	rs109285764	rs41590918	18	6.43	0.97
	16_20024651	16_20985573	rs110743197	rs42542723	23	5.88	0.99
	29_18046673	29_18882323	rs42375315	rs41589183	14	4.96	1.00
18:2							
	16_21019333	16_21956237	rs110103457	rs41583507	26	2.18	0.95
	18_43024125	18_43966013	rs41581224	rs109486478	22	1.89	0.92
SFA							
	19_51148913	19_51956162	rs41923412	rs109147235	25	15.58	1.00
	26_20012464	26_20984335	rs42981135	rs41623887	21	5.39	0.98
MUFA							
	19_51148913	19_51956162	rs41923412	rs109147235	25	16.63	1.00
	26_20012464	26_20984335	rs42981135	rs41623887	21	4.59	0.96
PUFA							
	17_51051583	17_51746326	rs109895216	rs110725997	4	2.13	0.91

Table 3.2. Characterization of windows accounting for variation in lipids in phospholipid. For each lipid or lipid class an estimate of genetic variance explained by the 1 Mb window is given along with window position coordinates, number of SNP, and posterior probability of association (*PPA*).

Trait	Map Start	Map End	RS# Start	RS# End	# of SNP	Genetic Variance (%)	PPA
14:0							
	16_4021893	16_4966340	rs110257825	rs109105804	26	1.79	0.68
	19_5057128	19_5934293	rs41633989	rs109106774	17	1.11	0.50
16:0							
	21_36032144	21_36993382	rs109143576	rs42429437	22	2.54	0.72
	1_52030982	1_52979497	rs41600017	rs43711327	25	1.10	0.54
16:1							
	4_95020715	4_95969034	rs43412327	rs42421263	20	1.39	0.68
	X_5276122	X_5920404	rs109239523	rs29023191	12	0.87	0.55
18:0							
	24_29013292	24_29942533	rs110012069	rs42837712	24	4.27	0.78
SFA							
	3_114057364	3_114995892	rs110764304	rs109271147	30	4.3	0.63
	10_54012929	10_54995489	rs42291384	rs110839090	27	2.97	0.50

Table 3.3. MCODE results derived from network scoring and proportion of variance accounted for in 16:0 for triacylglycerol and phospholipid fractions.

Fraction	Network	Score	Nodes	Edges	16:0 Genetic Variance (%)
Triacylglycerol					
	1	53.26	55	1438	67.9
	2	38.54	94	1792	29.8
	3	24.42	34	403	1.73E-03
	4	18.52	67	611	1.13E-01
	5	17.38	33	278	1.95E-03
	6	11.00	11	55	2.79E-05
	7	9.14	36	160	6.18E-04
	8	3.67	7	11	8.88E-06
	9	3.50	9	14	4.18E-05
	10	3.00	3	3	9.62E-06
	11	3.00	3	3	1.63E-06
	12	3.00	3	3	3.52E-08
Phospholipid					
	1	33.91	47	780	9.42E-04
	2	29.40	87	1264	1.20
	3	26.10	97	1273	1.01E-03
	4	10.00	72	355	1.51E-05
	5	7.40	68	248	3.27E-04
	6	5.81	63	180	3.82E-05
	7	4.96	45	109	5.86E-05
	8	3.79	20	36	2.59E-05
	9	3.00	7	9	2.21E-06
	10	2.80	16	21	9.76E-05

Annotation Cluster 1	Enrichment Score: 1.717490754477256						
Category	Term	Count	%	PValue	Fold Enrichment	Benjamini	FDR
GOTERM_BP_FAT	GO:0048872~homeostasis of number of cells	5	3.47	0.00	12.02	0.39	1.08
GOTERM_BP_FAT	GO:0042592~homeostatic process	9	6.25	0.00	3.56	0.67	4.70
GOTERM_BP_FAT	GO:0030099~myeloid cell differentiation	3	2.08	0.04	9.96	0.96	41.77
GOTERM_BP_FAT	GO:0048534~hemopoietic or lymph organ dev.	4	2.78	0.05	4.61	0.97	56.44
GOTERM_BP_FAT	GO:0002520~immune system development	4	2.78	0.06	4.26	0.96	63.67
GOTERM_BP_FAT	GO:0030097~hemopoiesis	3	2.08	0.18	3.87	0.99	94.69
Annotation Cluster 2	Enrichment Score: 1.6099545079517525						
Category	Term	Count	%	PValue	Fold Enrichment	Benjamini	FDR
SP_PIR_KEYWORDS	nadp	5	3.47	0.00	8.45	0.31	3.06
SP_PIR_KEYWORDS	nad	6	4.17	0.00	5.94	0.19	3.50
INTERPRO	IPR016040:NAD(P)-binding domain	5	3.47	0.02	4.96	0.95	19.22
UP_SEQ_FEATURE	active site:Proton acceptor	6	4.17	0.02	3.67	0.96	22.54
SP_PIR_KEYWORDS	oxidoreductase	8	5.56	0.02	2.74	0.67	24.39
GOTERM_BP_FAT	GO:0055114~oxidation reduction	9	6.25	0.03	2.45	0.93	33.71
INTERPRO	IPR002347:Glucose/ribitol dehydrogenase	3	2.08	0.06	7.15	1.00	55.43
INTERPRO	IPR002198:Short dehydrogenase/reduct. SDR	3	2.08	0.08	6.30	0.99	63.77
UP_SEQ_FEATURE	binding site:Substrate	3	2.08	0.35	2.40	1.00	99.42

Table 3.4. Functional Annotation Clustering for significant 1 Mb identified through GWAS for the Triacylglycerol fraction.

Table 3.5. Functional Annotation Chart results for significant 1 Mb regions identified through GWAS for the Triacylglycerol fraction.

Category	Term	Count	%	P-Value	Fold Enrichment	FDR
GOTERM_BP_FAT	GO:0048872~homeostasis of number of cells	5	3.47	0.00	12.02	1.08
SP_PIR_KEYWORDS	nadp	5	3.47	0.00	8.45	3.06
SP_PIR_KEYWORDS	nad	6	4.17	0.00	5.94	3.50
GOTERM_BP_FAT	GO:0042592~homeostatic process	9	6.25	0.00	3.56	4.70
GOTERM_BP_FAT	GO:0016358~dendrite development	3	2.08	0.00	32.18	5.35
GOTERM_BP_FAT	GO:0000187~activation of MAPK activity	3	2.08	0.01	18.19	15.71
INTERPRO	IPR016040:NAD(P)-binding domain	5	3.47	0.02	4.96	19.22
GOTERM_CC_FAT	GO:0016607~nuclear speck	3	2.08	0.02	13.74	19.96
GOTERM_MF_FAT	GO:0000166~nucleotide binding	17	11.81	0.02	1.77	21.01
UP_SEQ_FEATURE	active site:Proton acceptor	6	4.17	0.02	3.67	22.54
GOTERM_BP_FAT	GO:0043406~positive regulation of MAP kinase activity	3	2.08	0.02	13.07	27.63
GOTERM_BP_FAT	GO:0043623~cellular protein complex assembly	4	2.78	0.02	6.56	28.45
SP_PIR_KEYWORDS	oxidoreductase	8	5.56	0.02	2.74	24.39
GOTERM_BP_FAT	GO:0055114~oxidation reduction	9	6.25	0.03	2.45	33.71
KEGG_PATHWAY	bta04130:SNARE interactions in vesicular transport	3	2.08	0.03	10.83	26.23
GOTERM_CC_FAT	GO:0070013~intracellular organelle lumen	8	5.56	0.03	2.47	33.71
GOTERM_BP_FAT	GO:0043113~receptor clustering	2	1.39	0.03	55.78	41.38
GOTERM_CC_FAT	GO:0043233~organelle lumen	8	5.56	0.03	2.46	33.94
GOTERM_BP_FAT	GO:0030099~myeloid cell differentiation	3	2.08	0.04	9.96	41.77
GOTERM_CC_FAT	GO:0005739~mitochondrion	9	6.25	0.04	2.20	38.11
GOTERM_BP_FAT	GO:0030534~adult behavior	3	2.08	0.04	9.09	47.30
GOTERM_BP_FAT	GO:0007172~signal complex assembly	2	1.39	0.04	46.48	47.33
GOTERM_CC_FAT	GO:0031974~membrane-enclosed lumen	8	5.56	0.04	2.37	39.52
SP_PIR_KEYWORDS	Chaperone	4	2.78	0.04	4.97	41.44
GOTERM_BP_FAT	GO:0043405~regulation of MAP kinase activity	3	2.08	0.04	8.72	50.00
GOTERM_CC_FAT	GO:0016604~nuclear body	3	2.08	0.05	8.51	42.22
SMART	SM00241:ZP	2	1.39	0.05	37.62	36.23
UP_SEQ_FEATURE	nucleotide phosphate-binding region:NAD	3	2.08	0.05	8.00	47.19
GOTERM_BP_FAT	GO:0048534~hemopoietic or lymphoid organ development	4	2.78	0.05	4.61	56.44
GOTERM_BP_FAT	GO:0000165~MAPKKK cascade	3	2.08	0.05	7.89	56.48

Table 3.6 Functional Annotation Clustering for significant 1 Mb regions identified through GWAS for the Phospholipid fraction.

Annotation Cluster 1	Enrichment Score: 1.7110314771797666						
Category	Term	Count	%	PValue	Fold Enrichment	Benjamini	FDR
INTERPRO	IPR000436:Sushi/SCR/CCP	3	10.71	0.00	64.13	0.03	0.71
INTERPRO	IPR016060:Complement control module	3	10.71	0.00	60.47	0.01	0.79
SMART	SM00032:CCP	3	10.71	0.00	49.59	0.01	0.62

Table 3.7. Functional Annotation Chart results for significant 1 Mb regions identified through GWAS for the Phospholipid fraction.

Category	Term	Count	%	PValue	Fold Enrichment	FDR
INTERPRO	IPR000436:Sushi/SCR/CCP	3	10.71	0.00	64.13	0.71
INTERPRO	IPR016060:Complement control module	3	10.71	0.00	60.47	0.79
SMART	SM00032:CCP	3	10.71	0.00	49.59	0.62
SP_PIR_KEYWORDS	complement pathway	2	7.14	0.02	116.65	12.64
GOTERM_BP_FAT	GO:0002455~immune response mediated by immunoglobulin	2	7.14	0.02	104.58	17.02
GOTERM_BP_FAT	GO:0006958~complement activation, classical pathway	2	7.14	0.02	104.58	17.02
UP_SEQ_FEATURE	domain:Sushi 2	2	7.14	0.02	84.74	16.64
GOTERM_BP_FAT	GO:0006956~complement activation	2	7.14	0.02	77.06	22.38
GOTERM_BP_FAT	GO:0002541~activation of plasma in inflammatory response	2	7.14	0.02	77.06	22.38
SP_PIR_KEYWORDS	sushi	2	7.14	0.02	71.00	19.93
GOTERM_MF_FAT	GO:0004674~protein serine/threonine kinase activity	3	10.71	0.03	9.91	22.91
GOTERM_BP_FAT	GO:0006959~humoral immune response	2	7.14	0.03	56.31	29.31
GOTERM_BP_FAT	GO:0016064~immunoglobulin mediated immune response	2	7.14	0.04	45.76	34.75
SP_PIR_KEYWORDS	innate immunity	2	7.14	0.04	44.14	30.08
GOTERM_BP_FAT	GO:0019724~B cell mediated immunity	2	7.14	0.04	44.37	35.62
GOTERM_BP_FAT	GO:0051605~protein maturation by peptide bond cleavage	2	7.14	0.05	38.53	39.78
GOTERM_BP_FAT	GO:0002460~immune response immunoglobulin domains	2	7.14	0.05	36.60	41.37
GOTERM_BP_FAT	GO:0002250~adaptive immune response	2	7.14	0.05	36.60	41.37
GOTERM_BP_FAT	GO:0002449~lymphocyte mediated immunity	2	7.14	0.05	36.60	41.37
GOTERM_BP_FAT	GO:0002526~acute inflammatory response	2	7.14	0.05	34.05	43.68
GOTERM_BP_FAT	GO:0002443~leukocyte mediated immunity	2	7.14	0.05	32.54	45.17

Table 3.8. Functional Annotation Chart results for genes captured in networks associated with the Triacylglycerol fraction.

Category	Term	Count	%	PValue	Fold Enrichment	FDR
SP_PIR_KEYWORDS	lyase	4	4.21	0.00	11.35	5.03
GOTERM_MF_FAT	GO:0004190~aspartic-type endopeptidase activity	3	3.16	0.02	14.45	17.92
GOTERM_MF_FAT	GO:0070001~aspartic-type peptidase activity	3	3.16	0.02	14.45	17.92
INTERPRO	IPR001594:Zinc finger, DHHC-type	2	2.11	0.04	49.84	36.61
INTERPRO	IPR011991:Winged helix repressor DNA-binding	3	3.16	0.05	8.54	42.08
GOTERM_BP_FAT	GO:0006399~tRNA metabolic process	3	3.16	0.05	8.26	48.48

Table 3.9 Functional Annotation Clustering for genes captured in networks associated with the Phospholipid fraction.

Annotation Cluster 1	Enrichment Score: 1.7110314771797666						
Category	Term	Count	%	PValue	Fold Enrichment	Benjamini	FDR
GOTERM_BP_FAT	GO:0006732~coenzyme metabolic process	4	3.54	0.01	8.69	0.99	13.33
GOTERM_BP_FAT	GO:0051186~cofactor metabolic process	4	3.54	0.02	6.81	0.99	24.15
GOTERM_BP_FAT	GO:0009108~coenzyme biosynthetic process	3	2.65	0.02	13.32	0.95	25.06
GOTERM_BP_FAT	GO:0051188~cofactor biosynthetic process	3	2.65	0.04	9.73	0.98	40.65

Table 3.10. Functional Annotation Chart results for genes captured in networks associated with the Phospholipid fraction.

Category	Term	Count	%	PValue	Fold Enrichment	FDR
GOTERM_MF_FAT	GO:0003953~NAD+ nucleosidase activity	2	1.77	0.01	205.53	10.73
GOTERM_BP_FAT	GO:0006732~coenzyme metabolic process	4	3.54	0.01	8.69	13.33
PIR_SUPERFAMILY	MARVEL domain, MAL/Bene/MAL2 types	2	1.77	0.01	141.45	8.38
GOTERM_BP_FAT	GO:0051186~cofactor metabolic process	4	3.54	0.02	6.81	24.15
GOTERM_BP_FAT	GO:0009108~coenzyme biosynthetic process	3	2.65	0.02	13.32	25.06
GOTERM_CC_FAT	GO:0005819~spindle	3	2.65	0.02	12.04	23.43
INTERPRO	IPR013295:Myelin and lymphocyte (MAL)	2	1.77	0.02	81.52	24.06
GOTERM_BP_FAT	GO:0051188~cofactor biosynthetic process	3	2.65	0.04	9.73	40.65
INTERPRO	IPR014721:Ribosomal protein S5 domain 2-type fold	2	1.77	0.05	37.05	45.43
GOTERM_CC_FAT	GO:0015630~microtubule cytoskeleton	4	3.54	0.05	4.52	44.83
GOTERM_BP_FAT	GO:0022402~cell cycle process	4	3.54	0.05	4.51	55.00

Figure 1. Manhattan plot comparison of genetic variance explained by 1 Mb genomic windows displayed for fatty acid 16:0 in longissimus muscle. Plot A is derived from the triacylglycerol fraction. Plot B is derived from the phospholipid fatty acid fraction. Genetic variance explained by each 1 Mb genomic window is displayed on the y-axis. The x-axis displays chromosomal location of each window.

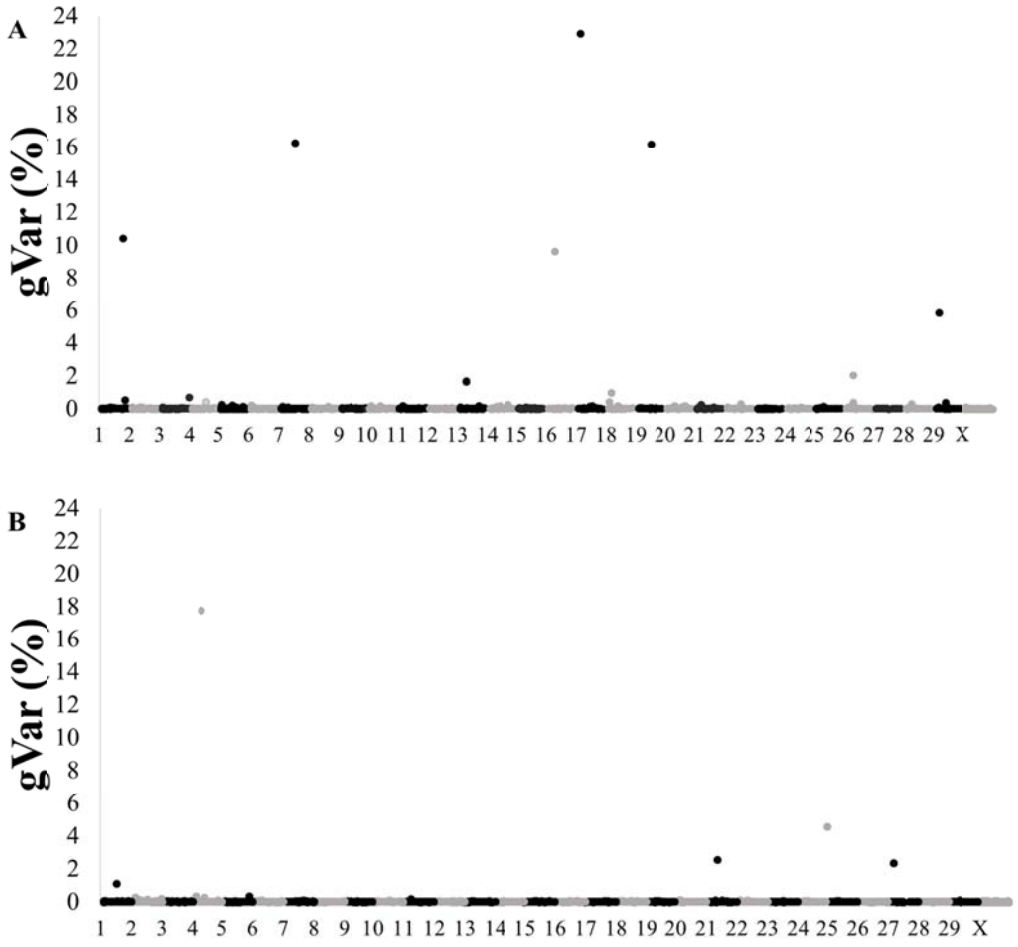


Figure 2. Highest scoring annotated SNP network from the triacylglycerol fraction derived from PCIT analysis. Location indicates significance of each node, with distance from the center indicating the number of connections. Each edge represents a correlation identified through PCIT analysis.

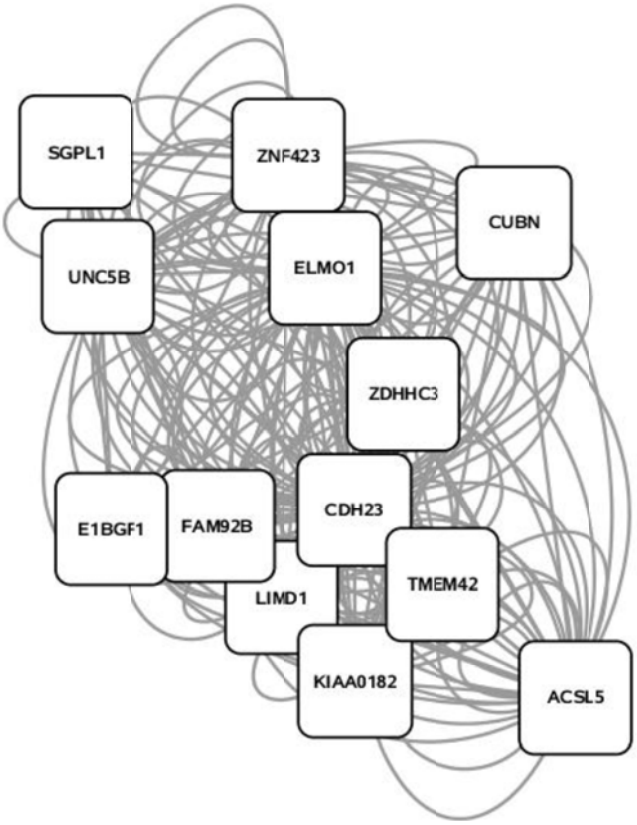


Figure 3. Second highest scoring annotated SNP network from the triacylglycerol fraction derived from PCIT analysis. Location indicates significance of each node, with distance from the center indicating the number of connections. Each edge represents a correlation identified through PCIT analysis.

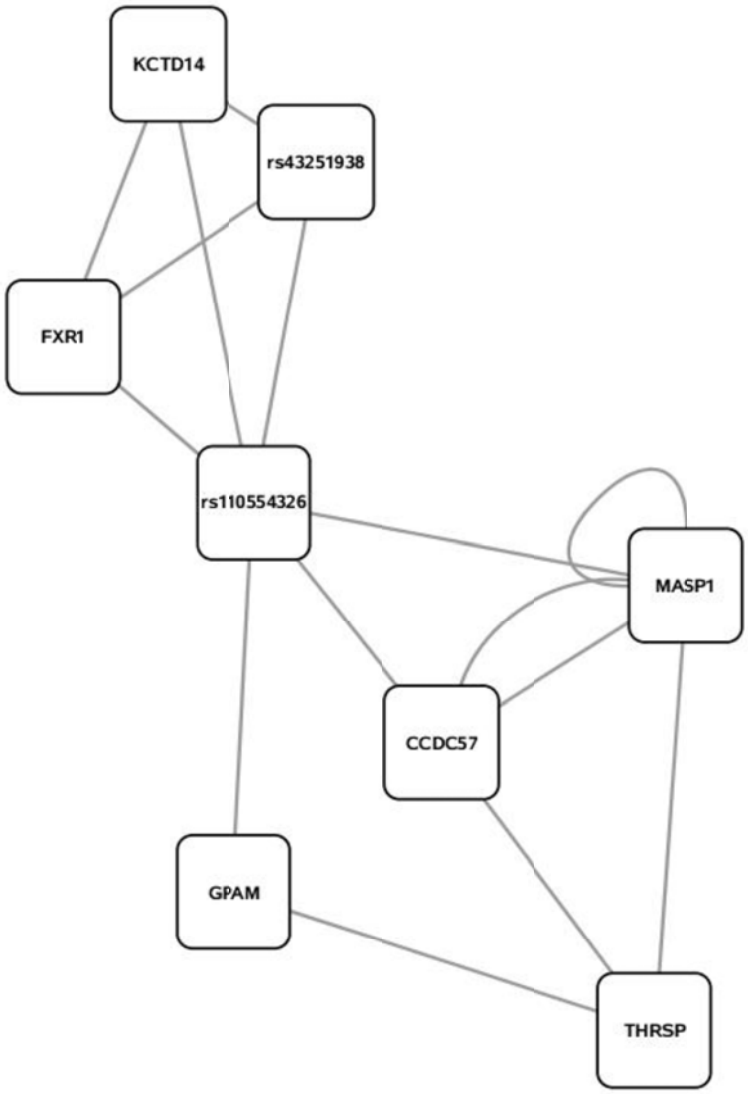


Figure 4. Highest scoring annotated SNP network from the phospholipid fraction derived from PCIT analysis. Location indicates significance of each node, with distance from the center indicating the number of connections. Each edge represents a correlation identified through PCIT analysis.

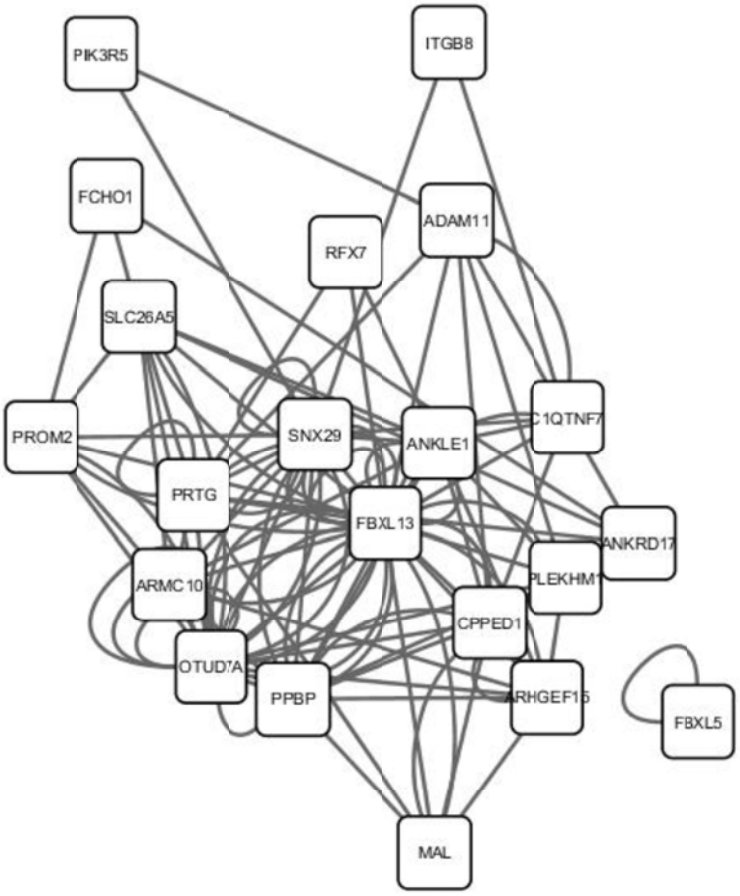


Figure 5. Second highest scoring annotated SNP network from the phospholipid fraction derived from PCIT analysis. Location indicates significance of each node, with distance from the center indicating the number of connections. Each edge represents a correlation identified through PCIT analysis.

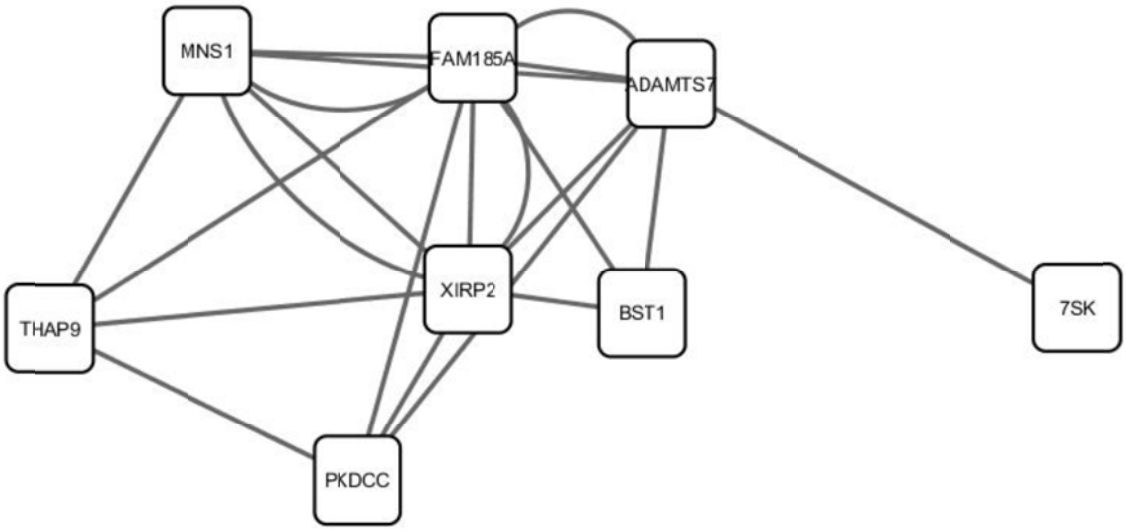


Figure 6. REVIGO visualization of DAVID Functional Annotation Clustering of genes from 1 Mb windows associated with triacylglycerol. Darker color indicates significance of P-value obtained through DAVID clustering. Connections indicate related biological process terms, and size represents frequency of the GO term in the GOA database.

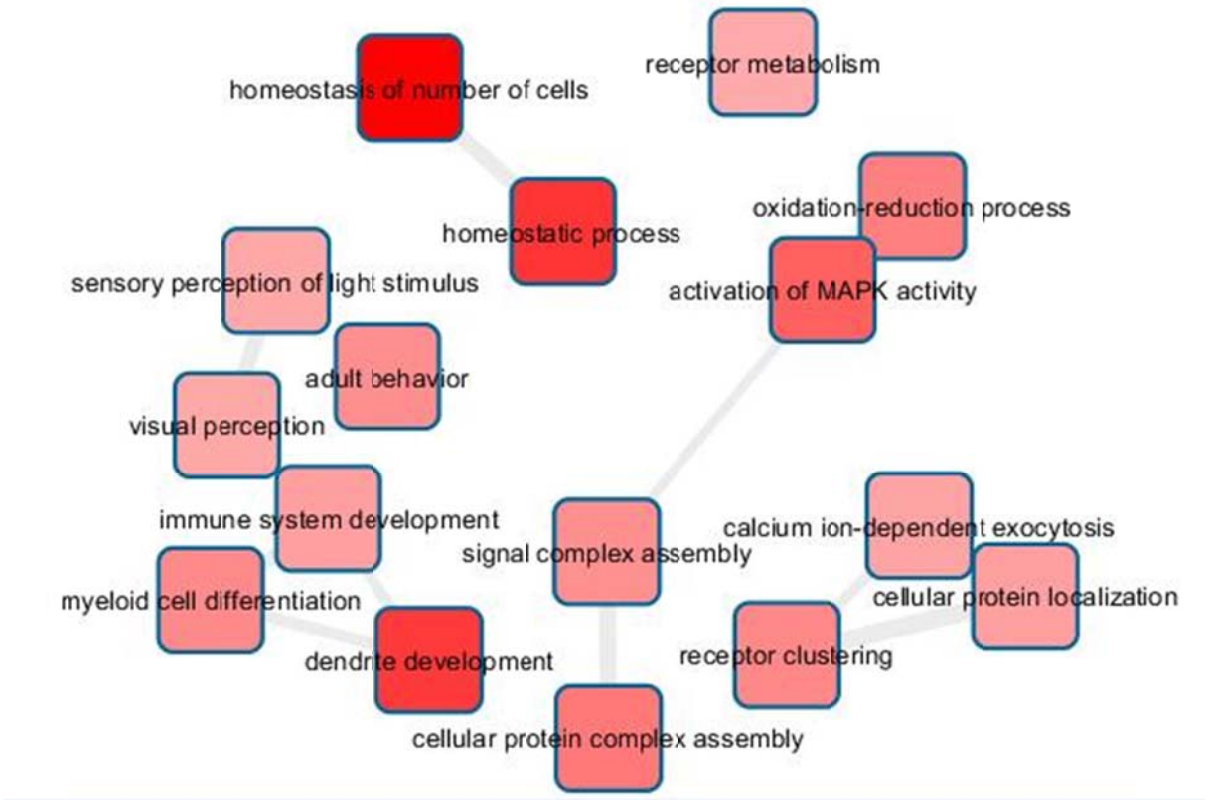


Figure 7. REVIGO visualization of DAVID Functional Annotation Clustering of genes from 1 Mb windows associated with phospholipid. Darker color indicates significance of P-value obtained through DAVID clustering. Connections indicate related biological process terms, and size represents frequency of the GO term in the GOA database.

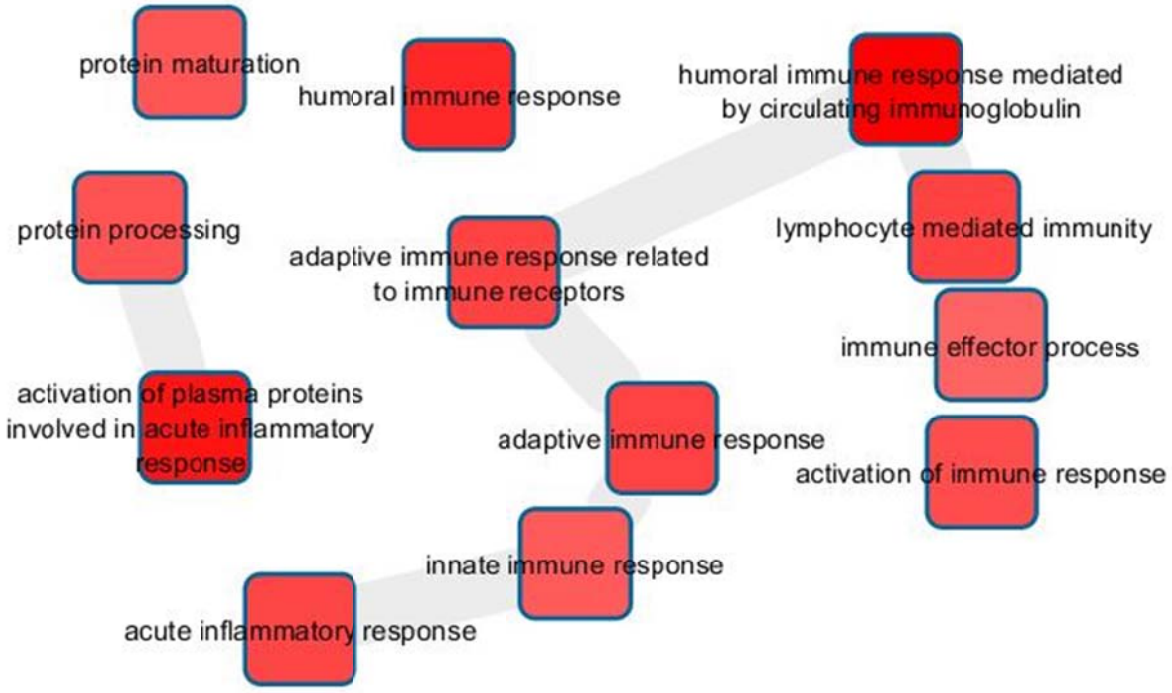


Figure 8. REVIGO visualization of DAVID Functional Annotation Clustering of genes from PCIT derived gene networks associated with triacylglycerol. Darker color indicates significance of P-value obtained through DAVID clustering. Connections indicate related biological process terms, and size represents frequency of the GO term in the GOA database.

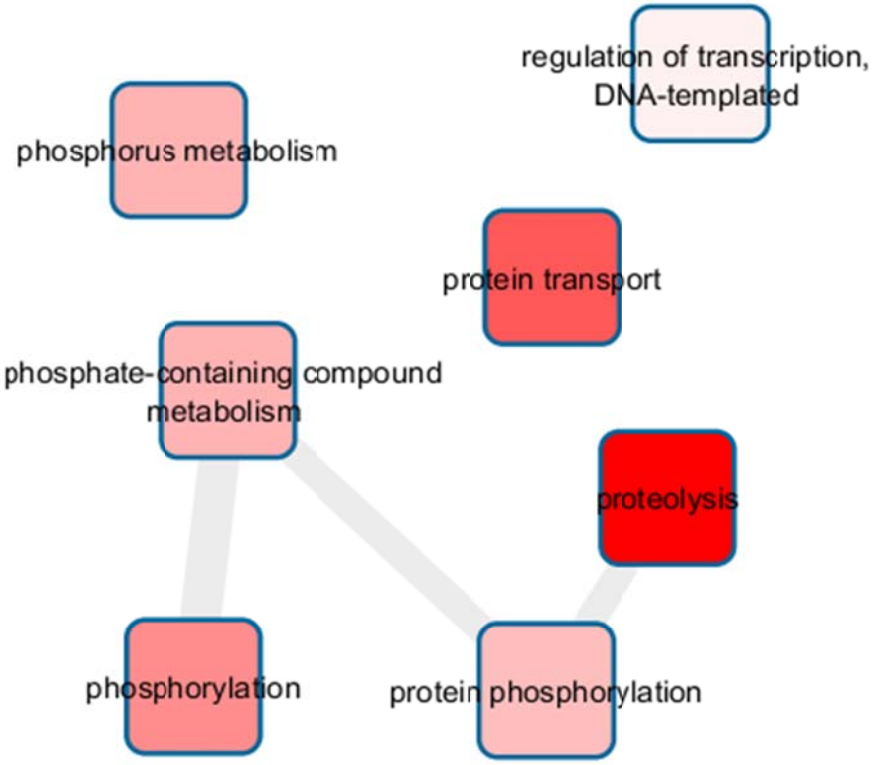
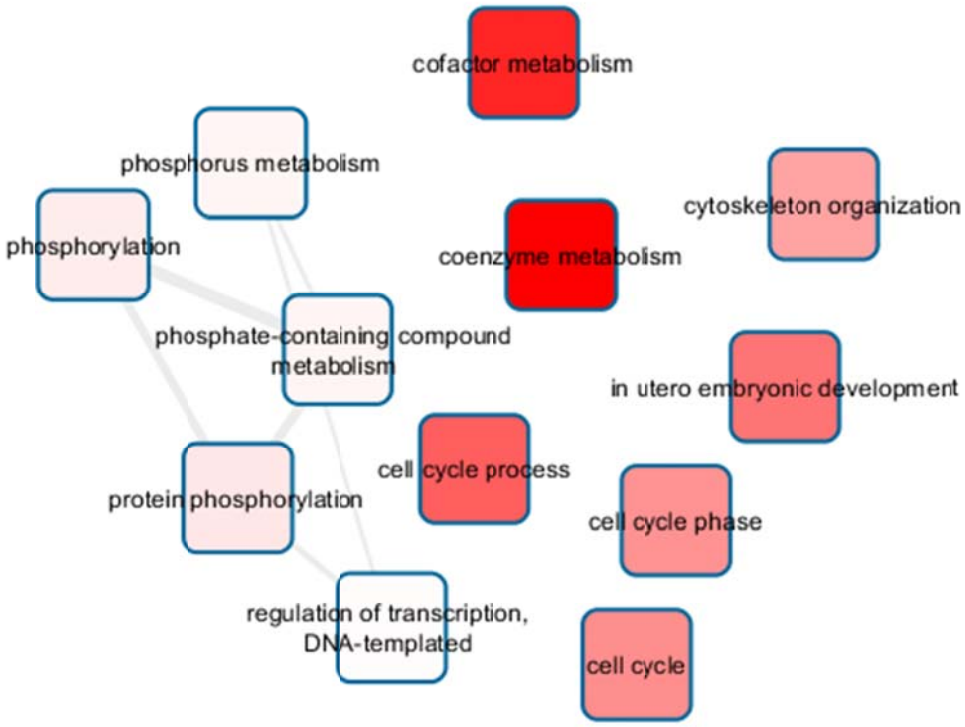


Figure 9. REVIGO visualization of DAVID Functional Annotation Clustering of genes from PCIT derived gene networks associated with phospholipid. Darker color indicates significance of P-value obtained through DAVID clustering. Connections indicate related biological process terms, and size represents frequency of the GO term in the GOA database.



VITA

Justin William Buchanan

Candidate for the Degree of

Doctor of Philosophy

Dissertation: GENETIC PARAMETER ESTIMATION AND GENE NETWORK DERIVATION FOR FATTY ACID TRAITS IN ANGUS BEEF CATTLE

Major Field: Animal Breeding and Genetics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Animal Breeding and Genetics at Oklahoma State University, Stillwater, Oklahoma in May, 2015.

Completed the requirements for the Master of Science in Animal Science at Oklahoma State University, Stillwater, Oklahoma in 2012.

Completed the requirements for the Bachelor of Science in Animal Science at Oklahoma State University, Stillwater, Oklahoma in 2010.

Experience:

Graduate Research Assistant, 2012 – Present
Dr. Megan Rolf & Dr. Raluca Mateescu, OSU Department of Animal Science, Stillwater, OK

Graduate Research Assistant, 2010 – 2012
Dr. Raluca Mateescu, OSU Department of Animal Science, Stillwater, OK

Animal Science Research Scholar, 2007-2010
Dr. Raluca Mateescu, OSU Department of Animal Science, Stillwater, OK

Professional Memberships:

American Society of Animal Science
Animal Science Graduate Student Association
Gamma Sigma Delta, The Honor Society of Agriculture
Phi Kappa Phi Professional Honor Society