**University of São Paulo
“Luiz de Queiroz” College of Agriculture**

**Identification of genes associated with intramuscular fat deposition and
composition in Nellore breed**

Aline Silva Mello Cesar

Thesis presented to obtain the degree of Doctor in
Science. Area: Animal Science and Pastures

**Piracicaba
2014**

Aline Silva Mello Cesar
Animal Scientist

**Identification of genes associated with intramuscular fat deposition and composition in
Nellore breed**

Advisor:
Prof. Dr. **LUIZ LEHMANN COUTINHO**

Thesis presented to obtain the degree of Doctor in
Science. Area: Animal Science and Pastures

**Piracicaba
2014**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Cesar, Aline Silva Mello

Identification of genes associated with intramuscular fat deposition and composition in Nellore breed / Aline Silva Mello Cesar. - - Piracicaba, 2014.
99 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2014.

1. Expressão gênica 2. Gordura 3. Gene posicional 4. GWAS 5. SNP I. Título

CDD 636.291
C421i

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte -O autor"

DEDICATION

This work is dedicated to my husband Albino and my son André for all support, patience, understanding, encouragement and love during my doctoral studies and during all time of my life.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Luiz Coutinho for his guidance, help, professionalism, dedication, commitment and friendship. I am grateful for the opportunity to work in his research team.

I would like to thank Dra. Luciana Regitano at EMBRAPA Southeast-Cattle Research Center for her suggestions, help, criticism during the conception and development of my experimental project.

To University of São Paulo - "Luiz de Queiroz" College of Agriculture (USP /ESALQ) for opportunity and support. To the Department of Animal Science for the teachings, opportunity and support provided by faculty and staff. To fellow students in the Department of Animal Science at ESALQ/USP that I had opportunity to become friends.

To the São Paulo State Research Foundation (FAPESP) for a Doctoral scholarship (Process 2011/00005-7) and a Research Abroad Internship (Process 2012/02383-1), which supports my Doctoral studies at University of São Paulo and Iowa State University.

To all researchers and staffs involved in BifeQuali project (Macroprograma 1, 01/2005) of EMBRAPA for financial support, collaborative efforts, infrastructure, teachings and opportunity.

To Dra. Marisa d'Darce for playing a very important role in my learning about food lipid chemistry, extraction, quality and nutritional importance.

To my supervisor Dr. James Reecy at Iowa State University (Ames, IA, USA) for his constant help, teaching, friendship and opportunity. To Animal Science Department at ISU for allowing me attend a classes, courses, seminars and for opening the infrastructure. To Dr. Reecy group for their reception, help and friendship.

To James Koltes and Eric Fritz-Waters for their constant help, teachings, patience and friendship during my training at Iowa State University.

To Dr. Dante Lanna and Maria Etchegaray at Animal Nutrition and Growth Laboratory at ESALQ for help and infrastructure to fatty acid analysis. To Dr. Gilberto Batista de Souza at Laboratory of Animal Nutrition at EMBRAPA Southeast-Cattle Research Center for help and infrastructure to ether extract analysis.

To Dr. Dorian Garrick and his group for their teachings, suggestions and support to genome-wide association analysis, also for allowing me to sit in his graduate classes during my research internship.

To my parents, Cesar and Madalena, my sisters Andréa, Alessandra e Andrezza, my brother, Alexandre, my nieces and nephews, and my brothers and sister-in-law for their love.

To my family-friends at Animal Biotechnology Laboratory, which I had opportunity to know and work together during my Doctoral studies. Thanks for your friendship, reception, confidence and support on the difficult and happy moments.

To all my professors for playing an important role in my personal and professional life, in special Eidiomar Angelucci, Viviane de Abreu Nunes Cerqueira Dantas, Flávio Vieira Meirelles, Paola Mariani, Robson Carlos Antunes, James Reecy and Luiz Lehmann Coutinho.

To God as an expression of my gratitude.

Aline Cesar

CONTENTS

RESUMO.....	9
ABSTRACT	11
1 INTRODUCTION	13
1.1 Hypothesis	15
1.2 Objectives	15
References.....	16
2 GENOME-WIDE ASSOCIATION STUDY FOR INTRAMUSCULAR FAT DEPOSITION AND COMPOSITION IN NELLORE CATTLE.....	19
Abstract.....	19
2.1 Introduction.....	19
2.2 Materials and Methods	20
2.2.1 Animals and Phenotypes.....	21
2.2.2 DNA Extraction and Genotypic Data	22
2.2.3 Descriptive Statistics and Heritability	22
2.2.4 Genome Wide Association Study	23
2.3 Results and discussion	25
2.3.1 Intramuscular fat deposition and composition.....	25
2.3.2 Heritability	26
2.3.3 Genome wide association studies and genomic regions identified	28
2.3.4 Saturated fatty acids.....	28
2.3.5 Monounsaturated fatty acids.....	37
2.3.6 Polyunsaturated fatty acids	41
2.4 Conclusion	44
References.....	44
3 PUTATIVE REGULATORY FACTORS INVOLVED IN MOLECULAR MECHANISMS OF INTRAMUSCULAR FAT DEPOSITION IN NELLORE STEERS	53
Abstract.....	53
3.1 Introduction.....	53
3.2 Materials and Methods	55
3.2.1 Ethics statement	55
3.2.3 Animals, samples and phenotypes	55

3.2.4 RNA extraction, quality analysis, library preparation and sequencing.....	56
3.2.5 Quality control and read alignment.....	57
3.2.6 Identification of differential expressed genes and pathway analysis	57
3.2.7 PCIT and differential hubbing network analysis	58
3.2.8 Regulatory Impact Factor (RIF) network co-expression analysis and phenotypic impact factor (PIF) scores	58
3.3 Results.....	58
3.3.1 Phenotypic groups, mapping and annotation	58
3.3.2 Differential expression analysis and differentially expressed genes	59
3.3.3 Functional enrichment and pathways.....	64
3.4 Discussion	85
3.5 Conclusion	92
References.....	92

RESUMO

Identificação de genes associados à deposição e composição da gordura intramuscular em bovinos da raça Nelore

A quantidade e composição da gordura intramuscular (GIM) pode influenciar as características sensoriais, o valor nutricional da carne bovina e na saúde humana. O perfil dos seus ácidos graxos pode se apresentar de maneira diversificada conforme a genética, o manejo e a nutrição dos animais de origem. A deposição e composição da gordura são determinadas por muitos genes que participam direta ou indiretamente da adipogênese e do metabolismo lipídico. A seleção de animais com teor e composição de gordura adequado para o consumidor é complexa pela difícil mensuração destas características, pela moderada herdabilidade e pelo desconhecimento dos genes envolvidos. Na última década, presenciamos um grande avanço na área da genômica bovina que resultou no sequenciamento completo do genoma e no desenvolvimento de chips de alta densidade de SNP. Este progresso científico, aliado aos avanços tecnológicos de equipamentos, resultou na identificação de genes responsáveis pela determinação de características quantitativas de interesse científico e comercial na bovinocultura. Este estudo teve como objetivo identificar e caracterizar genes associados à deposição e composição de gordura intramuscular em bovinos Nelore. Para este fim foi conduzido um estudo de associação genômica (*Genome-wide association studies*, GWAS) para identificar regiões genômicas associadas às características de interesse e identificar genes candidatos posicionais. Para o estudo de expressão diferencial foi conduzido um estudo do transcriptoma a partir do sequenciamento de RNA total (RNA-Seq) do músculo *Longissimus dorsi*. Foram utilizados 386 Nelores para a avaliação do teor de lipídeos total e perfil de ácidos graxos do músculo LD e, genotipagem com chip de alta densidade de SNP (*Illumina SNP800 BeadChip*). Um subconjunto de 14 animais, sendo sete animais de cada extremo para os valores genômicos estimados (GEBV) foi utilizado para o estudo de RNA-Seq. Foram encontradas 25 regiões genômicas (intervalos de 1 MB) associadas com deposição e composição de gordura intramuscular, as quais explicaram $\geq 1\%$ da variância genética. Estas regiões foram identificadas nos cromossomos 2, 3, 6, 7, 8, 9, 10, 11, 12, 17, 26 e 27, muitas destas não foram previamente detectadas em outras raças. Nestas regiões foram identificados importantes genes e podem ajudar no entendimento da base genética envolvida na deposição e composição de gordura. As regiões genômicas e genes aqui identificados e apresentados contribuem para um melhor entendimento do controle genético da deposição e composição de gordura em gado de corte e ainda podem ser aplicados em programas de seleção genética de animais que produzam carne com qualidade e com perfil de gordura saudável ao homem.

Palavras-chave: Expressão gênica; Gordura; Gene posicional; GWAS; SNP

ABSTRACT

Identification of genes associated with intramuscular fat deposition and composition in Nellore breed

The amount and composition of intramuscular fat (IMF) influence the sensory characteristics, nutritional value of beef and human health. The amount of fatty acid and its composition in beef varies by breed, nutrition, sex, age or carcass finishing level. The fat deposition and composition are determined by many genes that participate directly or indirectly in adipogenesis and lipid metabolism. The selection of animals with fat amount and composition suitable for the consumer is complex due to high cost of measurement, the moderate heritability and polygenic traits (many genes are involved with these traits). In the last decade with a great advance in bovine genomics resulted in the complete genome sequencing and the development of high-density chips of SNPs. This scientific advance jointly with technological improvement allowed the identification of genes responsible for important quantitative traits in cattle. This study aimed to identify and characterize genes associated with the deposition and composition of intramuscular fat in Nellore. A genome-wide association study (genome-wide association studies, GWAS) was performed to identify genomic regions associated with traits of interest and positional candidate genes. A total RNA sequencing (RNA-Seq) analysis was applied to transcriptome study of *Longissimus dorsi* muscle. Three hundred and eighty six Nellore steers were used for the evaluation of lipid content and fatty acid profile of LD, and genotyping with high-density chip SNP (SNP800 Illumina BeadChip). A subset of 14 animals, seven animals for each extremes of genomic estimated values (GEBV) were used to RNA-Seq analysis. Twenty-five genomic regions (1 MB window) were associated with the deposition and composition of intramuscular fat, which explained ≥ 1 % of the genetic variance. These regions were identified on chromosomes 2, 3, 6, 7, 8, 9, 10, 11, 12, 17, 26 and 27, many of these have not previously been found in other breeds and in these regions important genes were identified. Genomic regions and genes identified and presented here should be contribute to a better understanding of the genetic control of deposition and fat composition in beef cattle, and can be applied in breeding programs for animals that produce a quality and healthy beef to human consumers.

Keywords: Fat; Gene expression; GWAS; Positional gene; SNP

1 INTRODUCTION

The amount of intramuscular fat (IMF) affects the sensory characteristics and nutritional value of meat. Many consumers associate this fat with coronary heart disease, diabetes and obesity, due to the presence of cholesterol, the high concentration of saturated fatty acids and the low concentration of polyunsaturated fatty acids present in beef. However, consumption of fatty acids is necessary for human nutrition (LABORDE et al., 2001).

The fat composition influences the qualities of meat, on the sense that the saturated fats containing long-chain fatty acids that give the solid consistency and upon cooling affect the palatability. On the other hand, the unsaturated fatty acid can be easily oxidized affecting the meat flavor and odor (CASEY; VAN NIEKERK; WEBB, 2003).

The beef, as well as other animal species, is of great importance for human nutrition because it is a dense food (very low-calorie food) and high nutritional value. Meets the nutritional requirements of all ages, from children to seniors, it is a rich source of protein (essential amino acids), iron, zinc, B vitamins and essential polyunsaturated fatty acids such as linoleic and linolenic acid (LUCHIARI FILHO, 2000). Studies have shown that beef fat has a high concentration of monounsaturated fatty acids (MUFA), whose melting point is low and can reduce the concentration of bad cholesterol (LDL) in blood circulation (JAKOBSEN et al., 2008). The beef has about 48% saturated fat and 52% unsaturated fat, and is low in essential fatty acids (MOREIRA et al., 2003).

An adult animal gets to have 25-30% of your body weight compound by fats, which are very important for the proper functioning of the organism and maintaining homeostasis. They are part of cell membranes, of some hormones and vitamins; can function as cofactors enzymatic; and still are important forms of energy storage (LABORDE et al., 2001).

The set of isomers of linoleic fatty acid called conjugated linoleic acid (CLA) found in milk and meat is associated with reduced body fat, protection against cancer, fight against LDL cholesterol and control diabetes (ARITA et al., 1999). Saturated fatty acids are deposited as excess triglycerides and polyunsaturated fatty acids, specifically omega-3 phospholipids are deposited mostly as structural (PONNAMPALAM et al., 2001).

The omega-3 fatty acids have an important role in preventing cardiovascular disease and cancer (PETIT, 2002), so it is desirable in the flesh. The type of fat and deposition of this in the carcass of bovine animals vary according to nutrition, breed, sex, age or carcass finishing level (RULE et al., 1997). Results like that of (TAPIERO et al., 2002), confirm the influence of PUFAs, especially of omega-3, in reducing the serum cholesterol in prophylaxis

of atherogenesis and thrombosis.

Brazil has the largest commercial cattle herd in the world with over 193 million head in 2013, 40 million head slaughtered and from a total of 9,600 tonnes (UNITED STATES DEPARTMENT OF AGRICULTURE - USDA, 2013) of carcass equivalent 1,694 tonnes were exported presenting income of US\$ 4,835,129 (FNP, 2013).

In recent years the countries of the European Union (EU) became the largest importers of Brazilian beef, reaching a total of 21% of exports of this product in Brazil. Only Italy reported an increase of almost 45% in imports, which rose from 5,981 tons in 2008 to 13,276 tons of frozen boneless beef in 2009, according to FNP (2013).

In Brazil, cutting the animals are mostly Nellore (*Bos indicus* from Zebu origin) or crossbreeds of this breed. These animals are well adapted to tropical conditions in Brazil and are resistant to heat and parasites. The zebu animals when slaughtered in earlier ages, in relation to taurine animals, present more homogeneous distribution of subcutaneous fat (coverage) and lower amount of intramuscular fat (LUCHIARI FILHO; MOURÃO, 2006). For some markets in Continental Europe the least amount of fat in the carcass is desirable, but for others, like Asia and the UK this feature makes Brazilian beef with lower market value.

The adipogenesis (pre-adipocyte differentiation into adipocytes) and lipogenesis (fat synthesis) are controlled by transcription factors and enzymes participating in lipid metabolism. Some genes such as the acetyl coenzimaA carboxylase (ACC), esterol coenzimaA desaturase (SCD), fatty acid synthase (FAS), adiponectin (ADIPOQ) binding protein and fatty acid (FABP) (HOASHI et al., 2008) were associated with amount of body fat deposited, the profile of fatty acids synthesized and the cell differentiation of adipocytes.

Breeding programs around the world for many years, selected parents (bulls or cows) of the next generation only by phenotypes, and achieved much success. However, for some traits with low heritability (tenderness, marbling, etc.) phenotypic analysis possible only in adult animals (disease resistance, fertility, production efficiency, quality and quantity of milk) the annual genetic gain becomes limited and the selection became a challenge for improving programs (SORIA; CORVA, 2004).

With advances in biotechnology (DNA and RNA extraction, Polymerase Chain Reaction (PCR), gene expression chip, microsatellites and, more recently, high-density chips for single nucleotide polymorphisms (SNPs - single nucleotide polymorphisms) that allow the study of genome-wide association (GWAS) programs for animal breeding began to study the expression of genes of interest (OROZCO et al., 2010).

The technique of GWAS compares allele frequencies of thousands of polymorphic

markers available in unrelated individuals who possess a condition phenotypic (eg: deposition of intramuscular fat) of interest, to identify markers associated with this phenotype, 500kb to 1000kb regions are involved in GWAS and are analyzed by statistical tools and software available and appropriate for each situation (HAYES et al., 2009).

The use of panels of high-density SNPs in animal breeding programs allows the detection of significant QTLs for traits of economic importance with greater accuracy, allowing an analysis with multiple markers. These panels can be used for genomic selection from adjacent markers and to estimate the effects of these markers or haplotypes over the entire genome (HAYES et al., 2009). This enables getting all QTLs that contribute to genetic variation of a characteristic of interest.

Other interesting molecular technology is RNA-Seq that has been used in livestock to provide insight into the biological mechanisms relevant to economically important traits. For example, recent RNA-Seq studies of pigs with extreme IMF phenotypes have defined differentially expressed genes (DEG) and potential gene networks important in lipid and fatty acid metabolism in the liver (RAMAYO-CALDAS et al., 2012). In beef cattle, DEG were identified in backfat thickness from *Bos taurus* crossed steers, revealing that expression pattern depends on the genetic background (OLSON et al., 2012).

These tools of molecular biology allow a better understanding of biological processes such as deposition of IMF in animals, revealing regions of the genome, putative candidate genes and biological pathways responsible for this and other traits of economic interest to the cattle industry nationally and globally.

1.1 Hypothesis

This study was carried out with the purpose that the high-density SNP genotyping and the next-generation mRNA sequencing technologies may allow the genome-wide study, the identification of putative candidate genes and the detection of biological pathways involved with deposition and composition of intramuscular fat, specifically for Nellore steers.

1.2 Objectives

Employ a high-density single nucleotide polymorphism (SNP) chip to genotype Nellore steers, a *Bos indicus* breed and, a Bayesian approach to identify genomic regions and putative candidate genes that could be involved with deposition and composition of IMF. In

addition, identify differentially expressed genes (DEG) in the *Longissimus dorsi* (LD) muscle in steers of the Nellore breed and biological pathways associated with lipid metabolism.

References

- ARITA, Y.; KIHARA, S.; OUCHI, N.; TAKAHASHI, M.; MAEDA, K.; MIYAGAWA, J.; HOTTA, K.; SHIMOMURA, I.; NAKAMURA, T.; MIYAOKA, K.; KURIYAMA, H.; NISHIDA, M.; YAMASHITA, S.; OKUBO, K.; MATSUBARA, K.; MURAGUCHI, M.; OHMOTO, Y.; FUNAHASHI, T.; MATSUZAWA, Y. Paradoxical decrease of an adipose-specific protein, adiponectin, in obesity. **Biochemical and Biophysical Research Communications**, Orlando, v. 257, n. 1, p. 79-83, 1999.
- FNP CONSULTORIA & COMÉRCIO. **ANUALPEC 2013**: anuário da pecuária brasileira. São Paulo, 2013. 360 p.
- HAYES, B.J.; BOWMAN, P.J.; CHAMBERLAIN, A.C.; VERBYLA, K.; GODDARD, M.E. Accuracy of genomic breeding values in multi-breed dairy cattle populations. **Genetics Selection Evolution**, Paris, v. 41, n. 51, 2009. Available in: <<http://www.gsejournal.org/content/pdf/1297-9686-41-51.pdf>>. Accessed in: 22 oct. 2013.
- HOASHI, S.; HINENOYA, T.; TANAKA, A.; OHSAKI, H.; SASAZAKI, S.; TANIGUCHI, M.; OYAMA, K.; MUKAI, F.; MANNEN, H. Association between fatty acid compositions and genotypes of *FABP4* and *LXR-alpha* in Japanese Black cattle. **BMC Genetics**, London, v. 9, n. 84, 2008. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2156-9-84.pdf>>. Accessed on: 14 sept. 2013.
- JAKOBSEN, M.U.; OVERVAD, K.; DYERBERG, J.; HEITMANN, B.L. Intake of ruminant trans fatty acids and risk of coronary heart disease. **International Journal of Epidemiology**, Oxford, v. 37, n. 1, p. 173-182, 2008.
- LABORDE, F.L.; MANDELL, I.B.; TOSH, J.J.; WILTON, J.W.; BUCHANAN-SMITH, J. G. Breed effects on growth performance, carcass characteristics, fatty acid composition, and palatability attributes in finishing steers. **Journal of Animal Science**, Savoy, v. 79, n. 2, p. 355-365, 2001.
- LUCHIARI FILHO, A. **Pecuária da carne bovina**. Nova Odessa: O Autor, 2000. 147 p.
- LUCHIARI FILHO, A.; MOURÃO, G.B. **Melhoramento, raças e seus cruzamentos na pecuária de corte brasileira**. Pirassununga: A. Luchiari Filho, 2006. 142 p.
- MOREIRA, F.B.; DE SOUZA, N.E.; MATSUSHITA, M.; DO PRADO, I.N.; DO NASCIMENTO, W.G. Evaluation of carcass characteristics and meat chemical composition of *Bos indicus* and *Bos indicus* x *Bos taurus* crossbred steers finished en pasture systems. **Brazilian Archives of Biology and Technology**, Londrina, v. 46, p. 607-614, 2003.
- OROZCO, G.; BARRETT, J.C.; ZEGGINI, E. Synthetic associations in the context of genome-wide association scan signals. **Human Molecular Genetics**, Oxford, v. 19, p. 137-144, 2010.

PETIT, H.V. Digestion, milk production, milk composition, and blood composition of dairy cows fed whole flaxseed. **Journal of Dairy Science**, Champaign, v. 85, p. 1482-1490, 2002.

PONNAMPALAM, E.N.; SINCLAIR, A.J.; EGAN, A.R.; BLAKELEY, S.J.; LI, D.; LEURY, B.J. Effect of dietary modification of muscle long chain n-3 fatty acid on plasma insulin and lipid metabolites, carcass traits, and fat deposition in lambs. **Journal of Animal Science**, Savoy, v. 79, p. 895-903, 2001.

RULE, D.C.; BROUGHTON, K.S.; SHELLITO, S.M.; MAIORANO, G. Comparison of muscle fatty acid profiles and cholesterol concentrations of bison, beef cattle, elk, and chicken. **Journal of Animal Science**, Savoy, v. 80, p. 1202-1211, 2002.

TAPIERO, H.; BA, G.N.; COUVREUR, P.; TEW, K.D. Polyunsaturated fatty acids (PUFA) and eicosanoids in human health and pathologies. **Biomedicine and Pharmacotherapy**, Paris, v. 56, n. 5, p. 215-222, 2002.

UNITED STATES DEPARTMENT OF AGRICULTURE. **Livestock and poultry**: world markets and trade. Available in: <<http://usda.mannlib.cornell.edu/usda/current/livestock-poultry-ma/livestock-poultry-ma-11-08-2013.pdf>>. Accessed on: 15 Apr. 2014.

2 GENOME-WIDE ASSOCIATION STUDY FOR INTRAMUSCULAR FAT DEPOSITION AND COMPOSITION IN NELLORE CATTLE

Abstract

Meat from *Bos taurus* and *Bos indicus* breeds are an important source of nutrients for humans and intramuscular fat (IMF) influences its flavor, nutritional value and impacts human health. Human consumption of fat that contains high levels of monounsaturated fatty acids (MUFA) can reduce the concentration of undesirable cholesterol (LDL) in circulating blood. Different feeding practices and genetic variation within and between breeds influences the amount of IMF and fatty acid (FA) composition in meat. However, it is difficult and costly to determine fatty acid composition, which has precluded beef cattle breeding programs from selecting for a healthier fatty acid profile. In this study, we employed a high-density single nucleotide polymorphism (SNP) chip to genotype 386 Nellore steers, a *Bos indicus* breed and, a Bayesian approach to identify genomic regions and putative candidate genes that could be involved with deposition and composition of IMF. Twenty-three genomic regions (1-Mb SNP windows) associated with IMF deposition and FA composition that each explains $\geq 1\%$ of the genetic variance were identified on chromosomes 2, 3, 6, 7, 8, 9, 10, 11, 12, 17, 26 and 27. Many of these regions were not previously detected in other breeds. The genes present in these regions were identified and some can help explain the genetic basis of deposition and composition of fat in cattle. The genomic regions and genes identified contribute to a better understanding of the genetic control of fatty acid deposition and can lead to DNA-based selection strategies to improve meat quality for human consumption.

Keywords: Fatty acid; GWAS; *Bos indicus*; Beef; Positional candidate gene

2.1 Introduction

Many consumers associate consumption of fat from beef with coronary heart disease, diabetes and obesity, due to the presence of cholesterol, high concentration of saturated fatty acids (SFA), and low concentration of polyunsaturated fatty acids (PUFA). However, consumption of fatty acids is necessary for human nutrition (LABORDE et al., 2001). Beef has high nutritional value from children to seniors, is a rich source of protein (essential amino acids), iron, zinc, B vitamins and essential polyunsaturated fatty acids such as linoleic and linolenic acid (MCNEILL; VAN ELSWYK, 2012). Beef fat also has a high concentration of monounsaturated fatty acids (MUFA), whose melting point is low and can reduce the concentration of bad cholesterol (LDL) in blood circulation (JAKOBSEN et al., 2008).

The amount of fatty acid and its composition in beef varies by breed, nutrition, sex, age and carcass finishing level (RULE; MACNEIL; SHORT, 1997). The difficulties associated with determining intramuscular fat (IMF) deposition and composition as well as the limited knowledge on the genetic mechanisms that control these traits has limited genetic

progress in the production of healthier beef.

The development of high-density bovine genotyping (MATUKUMALLI et al., 2009) and their use in genome-wide association studies (GWAS) have allowed identification of genomic regions associated with phenotypes of interest. The technique of GWAS exploits differences in allele frequencies of thousands of polymorphic markers available in unrelated individuals who possess different phenotypes (for example, deposition and composition of intramuscular fat), and leads to the identification of markers associated with a given phenotype (HAYES et al., 2009). Bayesian approaches have been applied to GWAS to detect significant quantitative trait loci (QTL) for traits of economic importance. One such approach uses multiple regression (evaluating marker effects simultaneously), treating marker effects as random to reduce overestimation bias of significant QTL effects, generating the actual posterior distribution of QTL effects given the data which can provide richer inference than can be obtained by simply constructing p-values as well as providing an alternative to the use of p-values to avoid false positives (PETERS et al., 2012; GARRICK; FERNANDO, 2013; KIZILKAYA et al., 2013).

Brazilian beef is exported and consumed in more than 100 countries (PRAKASH; STIGLER, 2012). Purebred and crossbred Nellore cattle, which are of *Bos indicus* descent, are the predominant source of beef in Brazil. Previous research has documented that muscle and fat tissues from *Bos indicus* cattle develop in a different manner than in *Bos taurus* breeds (PICARD et al., 2002; LEHNERT et al., 2007; DU et al., 2013; DUARTE et al., 2013). However, studies documenting the genetics of fatty acid deposition and composition in *Bos indicus* breeds are limited.

In this study we performed a GWAS using high-density single nucleotide polymorphism (SNP) chips (770k) and Bayesian methods (Bayes B) to identify genomic regions associated with fat deposition and fatty acid composition (FA) in Nellore beef.

2.2 Materials and Methods

Animals were handled and managed according to Institutional Animal Care and Use Committee Guidelines from Brazilian Agricultural Research Corporation – EMBRAPA approved by the president, Dr. Rui Machado.

2.2.1 Animals and Phenotypes

Nellore steers (386) bred in the Brazilian Agricultural Research Corporation (EMBRAPA/Brazil) experimental breeding herd between 2009 and 2011 were available for this study. Steers were sired by 34 unrelated sires, and were selected to represent the main genealogies used in Brazil according to the National Summary of Nellore produced by the Brazilian Association of Zebu Breeders (ABCZ) and National Research Center for Beef Cattle. Animals were raised in feedlots under identical nutrition and handling conditions until slaughter at an average age of 25 months (TIZIOTO et al., 2012). Steaks (2.54 cm thick) from the *Longissimus dorsi* muscle between the 12th and 13th ribs were collected 24 hours after slaughter.

Muscle samples (~100 g) were lyophilized and ground for IMF and FA analysis. The IMF was obtained using an Ankom XT20 extractor as described (AOCS, 2000). FA analysis was conducted as described by Hara and Radin (1978), except the hexane to propanol ratio was increased to 3:2. Approximately 4 g of LD muscle was lyophilized, ground in liquid nitrogen, mixed with 28 mL of hexane/propanol (3:2 vol/vol) and homogenized for 1 min. Samples were vacuum filtered and 12 mL sodium sulfate (67 mg mL^{-1}) solution was added and agitated for 30 s. The supernatant was transferred to a tube with 2 g of sodium sulfate and insufflated with N_2 , after which the tube was sealed and incubated at room temperature for 30 min. Subsequently, the liquid was transferred to 10 mL test tube, insufflated with N_2 , sealed and kept at -20°C until dry with N_2 for methylation. The extracted lipids were hydrolyzed and methylated as described by Christie (1983), except that hexane and methyl acetate were used instead of hexane:diethyl ether:formic acid (90:10:1). Around 40 mg of lipids were transferred to a tube containing 2 mL of hexane. Subsequently, 40 μL of methyl acetate were added, the sample agitated, and 40 μL of methylation solution (1.75 mL of methanol/0.4 mL of 5.4 mol/L of sodium methoxide) were added. This mixture was agitated for 2 min and incubated for 10 min at room temperature. Then 60 μL of finishing solution (1 g of oxalic acid/30 mL of diethyl ether) were added and the mixture was agitated for 30 s, after which 200 mg of calcium chloride was added. The sample was then mixed and incubated at room temperature for 1 h. Samples were centrifuged at 3200 rpm, for 5 min at 5°C . The supernatant was collected for determination of fatty acids. Fatty acid methyl esters were quantified with a gas chromatograph (ThermoFinnigan, Termo Electron Corp., MA, USA) equipped with a flame ionization detector and a 100 m Supelco SP-2560 (Supelco Inc., PA, USA) fused silica capillary column (100 m, 0.25 mm and 0.2 μm film thickness). The column

oven temperature was held at 70 °C for 4 min, then increased to 170 °C at a rate of 13 °C min⁻¹, and subsequently increased to 250 °C at a rate of 35 °C min⁻¹, and held at 250 °C for 5 min. The gas fluxes were 1.8 mL min⁻¹ for carrier gas (He), 45 mL min⁻¹ for make-up gas (N₂), 40 mL min⁻¹ for hydrogen, and 450 mL min⁻¹ for synthetic flame gas. One µL sample was analyzed. Injector and detector temperatures were 250 and 300 °C, respectively. Fatty acids were identified by comparison of retention time of methyl esters of the samples with standards of fatty acids butter reference BCR-CRM 164, Anhydrous Milk Fat-Producer (BCR Institute for Materials and Reference Measurements) and also with commercial standard for 37 fatty acids Supelco TM Component FAME Mix (cat 18919, Supelco, Bellefonte, PA). The nomenclature of fatty acids follow IUPAC Compendium (MCNAUGHT; WILKINSON, 1997). Fatty acids were quantified by normalizing the area under the curve of methyl esters using Chromquest 4.1 software (Thermo Electron, Italy). Fatty acids were expressed as a weight percentage (mg/mg). These analyses were performed at the Animal Nutrition and Growth Laboratory at ESALQ, Piracicaba, São Paulo, Brazil.

2.2.2 DNA Extraction and Genotypic Data

DNA was isolated from blood as described by Tizioto et al. (2012). Genotyping was performed at the Bovine Functional Genomics Laboratory ARS/USDA and Genomics Center at ESALQ, Piracicaba, São Paulo, Brazil using BovineHD 770k BeadChip (Infinium BeadChip, Illumina, San Diego, CA) according to manufacturer's protocol. Genotypes were obtained in Illumina A/B allele format and used to represent a covariate value at each locus coded as 0, 1, or 2, representing the number of B alleles. Missing genotypes, represented < 0.2% of genotypes and were replaced with the average covariate value at that locus. Initial visualization and data analysis was performed by GenomeStudio Data Analysis Software (ILLUMINA, 2012). The SNPs with call rate ≤ 95%, minor allele frequency (MAF) ≤ 5%, those located on sex chromosomes and those not mapped in the *Bos taurus* UMD 3.1 assembly were removed. After filtering, a total of 449,363 SNP were utilized in GWAS.

2.2.3 Descriptive Statistics and Heritability

Descriptive statistics for IMF and FA were estimated using PROC MEANS and normality tests were performed using PROC UNIVARIATE in SAS (Ver. 9.3; SAS Inst. Cary, NC). SAS PROC MIXED was used to test independent sources for significance. Fixed

effects included contemporary group classes (animals with the same origin, birth year and slaughter date) and hot carcass weight as a covariate. Animal and residuals were fitted as random effects. Restricted maximum likelihood was used to estimate animal and residual variance components, heritability and standard error (SE) using ASREML software (GILMOUR et al., 2009). The model used in single-trait analyses of all traits was, $\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$, where $\mathbf{y}$ is the vector of observations representing the trait of interest (dependent variable), $\mathbf{X}$ and $\mathbf{Z}$ are the design or incidence matrices for the vectors of fixed and random effects in $\mathbf{b}$ and $\mathbf{u}$, respectively, and $\mathbf{e}$ was the vector of random residuals. The variance of vector $\mathbf{u}$ was $\mathbf{G}\sigma_m^2$ for the genomic analyses where $\mathbf{G}$ is the genomic relationship matrix derived from SNP markers using allele frequencies as suggested by VanRaden (2008), with σ_m^2 being the marker-based additive genetic variance.

2.2.4 Genome Wide Association Study

Associations between SNP and phenotypes (IMF and FA) were obtained using Bayes B, which analyzed all SNP data simultaneously and assumed a different genetic variance for each SNP locus (MEUWISSEN; HAYES; GODDARD, 2001; HABIER et al., 2011). The prior genetic and residual variances were estimated using Bayes C (KIZILKAYA; FERNANDO; GARRICK, 2010), with π being 0.9997. The model equation was:

$$\mathbf{y} = \mathbf{Xb} + \sum_{j=1}^k \mathbf{a}_j \beta_j \delta_j + \mathbf{e},$$

where $\mathbf{y}$ was the vector of the phenotypic values, $\mathbf{X}$ was the incidence matrix for fixed effects, $\mathbf{b}$ was the vector of fixed effects defined above, k was the number of SNP loci (449,363), $\mathbf{a}_j$ was the column vector representing the SNP covariate at locus j coded as the number of B alleles, β_j was the random substitution effect for locus j , which conditional on σ_β^2 was assumed to be normally distributed $N(0, \sigma_\beta^2)$ when $\delta_j = 1$ but $\beta_j = 0$ when $\delta_j = 0$, with δ_j being a random 0/1 variable indicating the absence (with probability π) or presence (with probability $1-\pi$) of locus j in the model, and $\mathbf{e}$ was the vector of the random residual effects assumed normally distributed $N(0, \sigma_e^2)$. The variance σ_β^2 (or σ_e^2) was a *priori* assumed to follow a scaled inverse Chi-square with $\nu_\beta = 4$ (or $\nu_e = 10$) degrees of freedom and scale parameter S_β^2 (or S_e^2). The scale parameter for markers was derived as a function of the assumed known genetic variance of the population, based on the average SNP allele

frequency and number of SNP assumed to have nonzero effects based on parameter π being 0.9997. This procedure used GenSel software (GARRICK; FERNANDO, 2013) to obtain the posterior distributions of SNP effects using Markov chain Monte Carlo (MCMC). This comprised a burn-in period of 1,000 iterations from which results were discarded, followed by 40,000 iterations from which results were accumulated to obtain the posterior mean effect of each SNP. In the Bayesian variable selection multiple-regression models with $\pi = 0.9997$ about 100-150 SNP markers were fitted simultaneously in each MCMC iteration. Inference of associations in these multiple-regression models was based on 1-Mb genomic windows rather than on single markers (ONTERU et al., 2011; GARRICK; FERNANDO, 2013). Genomic windows were constructed from the chromosome and base-pair positions denoted in the marker map file (GARRICK; FERNANDO, 2013) based on UMD3.1 bovine assembly.

The SNP effects from every 40th post burn in iteration were used to obtain samples from the posterior distribution of the proportion of variance accounted for by each window from 1,000 MCMC samples of genomic merit for each animal following Onteru et al. (2011) and Peters et al. (PETERS et al., 2012). In the present study there were 2,527 1 Mb SNP windows across the 29 autosomes. The proportion of genetic variance explained by each window in any particular iteration was obtained by dividing the variance of window BV by the variance of whole genome BV in that iteration. The window BV was computed by multiplying the number of alleles that represent the SNP covariates for each consecutive SNP in a window by their sampled substitution effects in that iteration.

All traits were used for GWAS and genomic heritability estimate, but only the ones with genomic heritability ≥ 0.10 were reported. Fatty acids were indexed as groups of saturated, monounsaturated, polyunsaturated fatty acid, total of saturated fatty acid (SFA), total monounsaturated (MUFA), total of polyunsaturated (PUFA), total of omega 3 (n-3) and total of omega 6 (n-6). Genome windows with the highest posterior mean proportion of genetic variance $\geq 1\%$ were considered the most important regions associated with the traits, and were declared the most promising QTL regions.

Positional candidate genes were investigated for 1 Mb windows using the Cattle Genome Browser (Cattle Genome UMD3.1) and UCSC Genome Browser (MEYER et al., 2013), which allowed visualization of SNP based on *Bos taurus* genome assembly UMD 3.1. Animal QTL database (Animal QTLdb) was used to search for published QTL and trait mapping data (HU et al., 2013). Gene annotations were retrieved from Ensembl Genes 71 Database using Biomart software (HUANG; SHERMAN; LEMPICKI, 2009; FLICEK et al.,

2013). The functional classification of genes was done using DAVID (HUANG; SHERMAN; LEMPICKI, 2009 and BioGPS (WU et al., 2009) online annotation databases. Those genes reported to be involved in fatty acid and lipid metabolism were selected as positional candidate genes.

2.3 Results and discussion

2.3.1 Intramuscular fat deposition and composition

Modern consumers are concerned with their overall health and often desire to reduce their caloric intake. This has increased the demand for lean meat production with a healthier fatty acid composition, which would comprise a lower proportion of SFA and greater proportion of MUFA. The amount of fat deposition in meat represented as IMF (mean = 2.77%) reported in this feedlot-finished study was greater than for pasture-finished counterparts, as expected, but lower than normally observed in Continental and English breeds (CHARLES; JOHNSON, 1976; HOLLOWAY et al., 1990; HUERTA-LEIDENZ et al., 1996). Despite of that, IMF observed in this work was within limits of reasonable amount of fat to assure acceptable quality levels for consumers according to Nuernberg et al. (2005).

The fatty acid composition observed for the most abundant FAs were: C14:0 at 3.54%, C16:0 at 26.69%, C18:0 at 14.98%, C16:1 cis-9 at 3.31%, C18:0 at 14.98%, C18:1 cis-9 at 37.46%, C18:2 cis-9 cis 12 at 1.60%, SFA at 47.23%, MUFA at 48.34%, and PUFA at 2.87% (Table 1). The proportion of MUFA was higher than SFA, and oleic acid (C18:1 cis-9) was the most abundant single fatty acid (37.46%). The FA composition presented in this work is similar to those reported in the literature for Nellore or other *Bos indicus* breeds (PRADO et al., 2003; PADRE et al., 2006; BRESSAN et al., 2011). This population also presented, in relation to *Bos taurus* breeds, average composition of fatty acids which is in agreement with reports from the USDA (2010). However, in this study a lower quantity of PUFA was observed in general, but not for C18:2 cis-9 trans-11, consequently a lower ratio of PUFA/SFA (6.08%). Similar MUFA and PUFA results have been reported in previous studies that utilized *Bos indicus* steers (RUIZ et al., 2005; BRESSAN et al., 2011).

The omega 6: omega 3 (n-6:n-3) ratio was high (4.84), which has a beneficial effect on patients with chronic diseases. A ratio of 4:1 has been associated with a 70% decrease in total mortality in the secondary prevention of cardiovascular disease (SIMOPOULOS, 2002). Omega 6 and 3 fatty acids are essential cell membrane and tissue constituents that are

important for many biological functions. These fatty acids are also essential for synthesis of prostaglandins, thromboxane, leukotriene, hydroxyl fatty acids, and lipoxins that are involved with inflammatory response. Humans and other mammals can convert n-6 to n-3 using desaturation enzymes but this conversion is slow and there is competition between n-6 and n-3 fatty acids for the desaturation enzymes (SIMOPOULOS, 2008).

2.3.2 Heritability

Descriptive statistics and heritabilities estimated using a genomic relationship **G** matrix are in Table 1. Heritabilities estimated in this study varied from low (<0.10 for C12:0, C16:0, C18:1 cis-11, C18:1 cis-12, C18:2 cis-9 trans-11, C20:1, C20:3 n-6, C20:4 n-6, C20:5 n-3 and AI, respectively) to moderate (up to 0.29 for IMF, C14:0, C14:1 cis-9, C16:1 cis-9, C17:0, C17:1, C18:0, C18:1 cis-9, C18:1 trans-6, 7, 8, C18:1 trans-10, 11, 12, C18:2 cis 9 cis-12 n-6, C18:2 trans-11 cis-15, C18:3 n-6, C18:3 n-3, C22:5 n-3, C22:6 n-3, SFA, MUFA, PUFA, Sn-3, Sn-6 and n-6:n-3). For C15:0, C18:1 cis-13, C18:1 cis-15, C18:1 trans-16, C20:2 the heritability estimates were zero. In Angus (TAIT et al., 2007) and Japanese Black cattle (NOGI et al., 2011) estimates of heritability for IMF fat deposition and composition traits were higher than in this study. The lower values of heritability reported for this population could be explained by the reduced sample size (CASAS et al. 2001) or lower amount of genetic variation in the population (CASAS et al., 2003).

The estimates of genomic heritability from Bayes B approach (Table 2) were also low to moderate (from 0.09 to 0.46) and generally similar to values obtained in ASReml using a genomic relationship matrix.

Table 1 - Descriptive statistics, variance components and heritability by GBLUP for IMF deposition and composition in Nellore

Trait	Terminology ¹	N	Mean±SE ²	Genetic Variance	Residual Variance	Total Variance	h ² ±SE
IMF (%)	Intramuscular fat	382	2.77±0.05	0.196	0.490	0.686	0.29±0.16
C12:0 (mg/mg)	Lauric acid	374	0.07±0.001	0.000039	0.00072	0.000761	0.05±0.09
C14:0	Myristic acid	378	3.54±0.03	0.0530	0.250	0.303	0.17±0.11
C14:1 cis-9	Myristoleic acid	378	0.96±0.01	0.0076	0.041	0.0486	0.16±0.11
C15:0	Pentadecylic acid	378	0.80±0.02	0.0	0.052	0.052	0±0.06
C16:0	Palmitic acid	378	26.69±0.15	0.6070	7.068	7.675	0.08±0.10
C16:1 cis-9	Palmitoleic acid	378	3.31±0.04	0.0640	0.354	0.418	0.15±0.10
C17:0	Margaric acid	378	1.07±0.009	0.0061	0.019	0.0251	0.24±0.15
C17:1	Heptadecenoic acid	378	0.58±0.007	0.0024	0.009	0.0114	0.20±0.12
C18:0	Stearic acid	378	14.98±0.14	1.3380	5.348	6.686	0.20±0.12
C18:1 cis-9	Oleic acid	378	37.46±0.22	2.0720	10.826	12.898	0.16±0.11
C18:1 cis-11	Cis-Vaccenic acid	378	2.98±0.05	0.0850	0.357	0.442	0.02±0.09
C18:1 cis-12	Cis-12 Octadecenoic	377	0.91±0.02	0.0030	0.034	0.037	0.09±0.10
C18:1, cis-13	Cis-13 Octadecenoic	377	0.58±0.008	0.0	0.021	0.021	0±0.06
C18:1 cis-15	Cis-15 Octadecenoic	377	0.06±0.002	0.0	0.0005	0.0005	0±0.06
C18:1 trans-6, 7, 8	Trans-6,7,8 Octadecenoic	378	0.18±0.004	0.0007	0.0055	0.0062	0.11±0.09
C18:1 trans-10, 11, 12	Trans-10,11,12 Octadecenoic	378	1.07±0.02	0.0236	0.0926	0.1162	0.20±0.12
C18:1 trans-16	Trans-16 Octadecenoic	377	0.14±0.003	0.0	0.0023	0.0023	0±0.09
C18:2 cis-9 cis-12 n-6	Linoleic acid	377	1.60±0.03	0.0340	0.239	0.273	0.12±0.10
C18:2 cis-9 trans-11	Vaccenic acid	377	0.21±0.003	0.0001	0.0026	0.0027	0.04±0.09
C18:2 trans-11 cis-15	Octadecenoic acid	374	0.07±0.001	0.00007	0.0004	0.00047	0.13±0.11
C18:3 n-6	α-Linolenic acid	377	0.06±0.001	0.00008	0.0002	0.00028	0.24±0.13
C18:3 n-3	γ-Linolenic acid	376	0.16±0.005	0.00026	0.0017	0.00196	0.13±0.11
C20:1	Eicosanoic acid	374	0.11±0.002	0.00013	0.0014	0.00153	0.09±0.10
C20:2	Eicosadienoic acid	306	0.01±0.0003	0.0	0.00004	0.00004	0±0.08
C20:3 n-6	Eicosatrienoic acid	373	0.11±0.003	0.00018	0.00235	0.00253	0.07±0.09
C20:4 n-6	Arachidonic acid	377	0.36±0.008	0.00221	0.0235	0.0251	0.09±0.10
C20:5 n-3 (EPA)	Eicosapentaenoic acid	376	0.08±0.002	0.00005	0.00135	0.00139	0.04±0.09
C22:5 n-3 (DPA)	Docosapentaenoic acid	377	0.17±0.003	0.00047	0.00387	0.00434	0.11±0.10
C22:6 n-3 (DHA)	Docosahexaenoic acid	364	0.03±0.001	0.00003	0.00018	0.00021	0.13±0.10
SFA ³	Sum of saturated FA	377	47.23±0.23	1.960	15.00	16.960	0.11±0.09
MUFA ³	Sum of monounsaturated FA	377	48.34±0.23	2.367	14.963	17.33	0.14±0.10
PUFA ³	Sum of polyunsaturated FA	377	2.87±0.04	0.097	0.547	0.644	0.15±0.10
n-3 ⁴	Sum of omega-3	377	0.44±0.09	0.00034	0.0016	0.00194	0.17±0.11
n-6 ⁴	Sum of omega-6	377	2.13±0.04	0.00055	0.0031	0.00365	0.15±0.11
PUFA:SFA	Ratio of PUFA to SFA	377	6.08±0.003	0.0623	0.41	0.4723	0.13±0.10
n-6:n-3	Ratio of n-6 to n-3	377	4.84±0.11	0.00395	0.0231	0.02705	0.14±0.10
AI ⁵	Atherogenic index	377	0.82±0.05	0.03943	1.0784	1.11783	0.03±0.09

¹IUPAC Compendium of Chemical Terminology

²SE Standard error

³MUFA, PUFA and SFA were the total of all monounsaturated, polyunsaturated and saturated fatty acids, respectively.

⁴n-3 and n-6 were total of omega 3 and 6 fatty acids

⁵Atherogenic index = [12:0 + 4(14:0) + 16:0]/(SFA + PUFA)

2.3.3 Genome wide association studies and genomic regions identified

All traits were used for GWAS and genomic heritability estimate, but only the ones with genomic heritability ≥ 0.10 were reported: IMF, 33 different FA and ratios, and one FA index. Top three QTL regions associated with IMF deposition and composition traits in Nellore by Bayes B are showed in Table 3. The 1-Mb SNP window regions that explained more than 1% of the genetic variance were used to search for putative candidate genes (PCG) and are presented on Table 4. Therefore, 35 traits were used for GWAS and these represented 23 different 1-Mb genomic regions (Table 4). These regions were distributed over 12 different chromosomes: 2, 3, 6, 7, 8, 9, 10, 11, 12, 17, 26 and 27 and the corresponding PCG in these regions are reported in Table 3. Intramuscular fat was one of the traits that presented moderate heritability (0.25), however no region that explained more than 1% of the genetic variance for this trait was identified.

2.3.4 Saturated fatty acids

Eight genomic regions (1 Mb windows) explained more than 1% of genotypic variation for C12:0, C14:0, C16:0, and C18:0 (Table 4). These regions overlap with QTL previously reported for marbling score (CASAS et al., 2001), backfat thickness (CASAS et al., 2003), carcass weight and body weight in Angus cattle (MCCLURE et al., 2010). Figure 1 shows in more detail the Manhattan plot of the proportion of genetic variance explained by window across the 29 autosomes for the important saturated fatty acids for beef palatability: C12:0 (lauric acid), palmitic acid (C16:0), stearic acid (C18:0).

BTA12 at 60 Mb QTL region was associated with C12:0 fatty acid and harbors the PCG *SLIT* and *NTRK*-like family, member 6 (*SLITRK6*). This gene is related to the lipopolysaccharide receptor complex (GO:0016021). Lipopolysaccharide (endotoxin) is an important component of gram-negative cell walls, which can start the inflammatory process by binding to the *SLITRK6* complex. This complex is located at the surface of innate immune cell, which presents an important role in obesity associated with insulin resistance (SHI et al., 2006).

Table 2 - Posterior means of variance components for IMF deposition and composition in Nellore by Bayes B

Trait	Genetic Variance	Residual Variance	Total Variance	Genomic Heritability
IMF (%)	0.16	0.48	0.64	0.25
C12:0 (mg/mg)	0.0001	0.0006	0.0007	0.18
C14:0	0.06	0.23	0.29	0.20
C14:1 cis-9	0.03	0.02	0.05	0.25
C15:0	0.006	0.04	0.046	0.12
C16:0	2.42	5.47	7.89	0.31
C16:1 cis-9	0.10	0.30	0.40	0.24
C17:0	0.004	0.02	0.024	0.17
C17:1	0.002	0.01	0.012	0.17
C18:0	1.23	5.28	6.51	0.19
C18:1 cis-9	6.08	7.11	13.19	0.46
C18:1 cis-11	0.04	0.33	0.37	0.11
C18:1 cis-12	0.005	0.03	0.035	0.13
C18:1, cis-13	0.02	0.002	0.022	0.09
C18:1 cis-15	0.0009	0.0004	0.00049	0.17
C18:1 trans-6, 7, 8	0.0008	0.005	0.0058	0.13
C18:1 trans-10, 11, 12	0.02	0.09	0.11	0.16
C18:1 trans-6	0.0003	0.002	0.0023	0.16
C18:2 cis 9 cis-12 n-6	0.03	0.23	0.26	0.13
C18:2 cis9, trans-11	0.0003	0.002	0.0023	0.12
C18:2 trans-11 cis-15	0.0001	0.0004	0.0005	0.22
C18:3 n-6	0.00007	0.0003	0.00037	0.21
C18:3 n-3	0.0003	0.002	0.0023	0.14
C20:1	0.0002	0.001	0.0012	0.16
C20:2	0.00005	0.00002	0.00007	0.22
C20:3 n-6	0.0003	0.002	0.0023	0.14
C20:4 n-6	0.002	0.02	0.022	0.08
C20:5 n-3	0.0002	0.001	0.0012	0.17
C22:5 n-3	0.0007	0.003	0.0037	0.16
C22:6 n-3	0.00006	0.0002	0.00026	0.24
SFA ¹	1.36	15.54	16.90	0.08
MUFA ¹	1.67	15.18	16.85	0.10
PUFA ¹	0.09	0.54	0.63	0.14
n-3 ²	0.004	0.013	0.017	0.25
n-6 ²	0.0005	0.003	0.0035	0.15
PUFA:SFA	0.003	0.02	0.023	0.11
n-6:n-3	0.66	1.23	1.89	0.34
AI ³	0.007	0.03	0.037	0.16

¹MUFA, PUFA and SFA were the total of all monounsaturated, polyunsaturated and saturated fatty acids, respectively

²n-3 and n-6 were total of omega 3 and 6 fatty acids

³Atherogenic index = $[12:0 + 4(14:0) + 16:0]/(SFA + PUFA)$

Table 3 - Top three QTL regions associated with IMF deposition and composition traits in Nellore by Bayes B

(continues)					
Trait	QTL Window (first and last SNP)	Number of SNPs	%Gen. Variance	Chr	Map Position (UMD 3.1 bovine)
IMF	rs110852801-rs110852801	144	0.66	10	50013104-50992412
	rs110396618-rs42932197	131	0.42	9	3018763-3997425
	rs109843584-rs137009265	243	0.4	6	28008364-28996379
C12:0	rs42924061-rs134942057	214	3.48	12	60000226-60994461
	rs135783220-rs109238095	169	0.44	5	38001237-38999839
	rs42639909-rs43332310	221	0.32	3	31001943-31992879
C14:0	rs41627556-rs110731616	332	1.43	9	36013595-36997263
	rs43328164-rs134696015	295	1.06	3	6009105-6999643
	rs134043558-rs134858042	182	0.82	2	107001640-107997035
C14:1 cis-9	rs137683417-rs29003360	232	1.86	2	26002605-26999718
	rs137754875-rs135670282	203	1.55	12	66002783-66997356
	rs110517663-rs137137362	127	1.43	11	27000446-27993515
C15:0	rs133645667-rs41591263	239	0.58	18	26001094-26997723
	rs133520595-rs135893631	199	0.45	14	75014305-75995520
	rs110569600-rs109846765	238	0.34	22	27024539-27994243
C16:0	rs109773631-rs135618512	277	1.53	12	13000697-13997298
	rs134160160-rs109942510	285	1.38	3	26002777-26998679
	rs135379047-rs136882159	173	0.58	8	20012045-20991197
C16:1 cis-9	rs133274959-rs136576856	213	1.42	3	25008982-25990012
	rs136405986-rs137105475	114	1.18	10	74001100-74957582
	rs135892505-rs109693564	184	1.04	7	80006438-80997126
C17:0	rs43471077-rs110039814	219	0.56	6	66005071-66997033
	rs137480735-rs137082211	153	0.51	4	38001142-38996557
	rs135375532-rs109902967	114	0.45	18	50001050-50870670
C17:1	rs110761991-rs110425294	83	0.83	5	55000572-55989100
	rs110775410-rs110218142	202	0.67	6	32002888-32998721
	rs133062684-rs109893445	144	0.6	6	21006361-21998823
C18:0	rs133274959-rs136576856	213	3.46	3	25008982-25990012
	rs132804279-rs109645596	102	2.08	11	105009792-105985714
	rs136405986-rs137105475	114	1.08	10	74001100-74957582
C18:1 cis-9	rs133274959-rs136576856	213	2.57	3	25008982-25990012
	rs109773631-rs135618512	277	1.91	12	13000697-13997298
	rs134160160-rs109942510	285	1.40	3	26002777-26998679
C18:1 cis-11	rs133850149-rs133020896	205	0.58	16	15004033-15997381
	rs135337759-rs43566804	192	0.51	8	76000301-76995458
	rs136472828-rs110306459	206	0.43	15	70000108-70998968
C18:1 cis-12	rs41776470-rs133872693	243	0.93	15	71015181-71997242
	rs137532692-rs134926660	186	0.44	24	49001908-49998425
	rs133496949-rs42611028	214	0.38	1	5004036-5976378
C18:1 cis-15	rs135695322-rs136588271	283	0.79	15	73014830-73997794
	rs135217991-rs133997168	260	0.48	11	13001510-13999856

Table 3 - Top three QTL regions associated with IMF deposition and composition traits in Nellore by Bayes B

(continuation)					
Trait	QTL Window (first and last SNP)	Number of SNPs	%Gen. Variance	Chr	Map Position (UMD 3.1 bovine)
C18:1 trans-6, 7, 8	rs137648958-rs134302099	161	0.41	2	32007364-32996349
	rs136989423-rs135647270	125	0.39	1	127000266-127991193
	rs134408743-rs134205414	117	0.37	9	25003343-25990361
	rs135780675-rs135123069	227	0.3	1	125001859-125996439
C18:1trans-10, 11, 12	rs136697703-rs137139554	122	0.57	5	90000699-90997254
	rs135750022-rs43138782	209	0.5	9	64000525-64992666
	rs137485825-rs133970673	110	0.49	13	58029906-58998257
C18:1 trans-16	rs133728493-rs109630757	264	2.56	6	50008629-50996673
	rs109872854-rs134277482	250	2.17	17	41001273-41991683
	rs109193786-rs109735608	215	0.62	2	65003697-65984693
C18:2cis-9 cis-12 n-6	rs134008882-rs136224281	164	0.46	9	52014336-52995340
	rs42780247-rs42794783	195	0.38	21	50005405-50997226
	rs43750576-rs42495887	212	0.31	12	62010947-62997021
C18:2 cis-9 trans-11	rs136630339-rs110298112	178	0.27	17	23004876-23997634
	rs134056892-rs377764129	222	0.26	8	27003611-27997814
	rs110518805-rs110823657	355	0.26	2	119002084-119974723
C18:2 trans-11 cis-15	rs135377389-rs133297940	267	3.49	8	68013821-68993778
	rs42404785-rs133183089	144	1.69	3	72000121-72997801
	rs137602675-rs133056879	187	1.01	26	20001875-20999270
C18:3 n-3	rs109612389-rs133661384	132	1.33	17	24002089-24987264
	rs41973816-rs109360228	142	0.43	21	21003963-21985795
	rs132831049-rs109343530	242	0.43	9	40014892-40995874
C18:3 n-6	rs42582725-rs133775322	205	1.34	7	34004443-34998434
	rs133846205-rs133219654	151	0.95	20	49025175-49991746
	rs136202159-rs134873098	235	0.91	2	5002853-5998856
C20:1	rs110940448-rs135112502	210	0.80	7	63000813-63990965
	rs134963634-rs42390829	207	0.79	20	55010178-55999450
	rs109443339-rs137113406	291	0.52	23	28001216-28998760
C20:2	rs136000962-rs136871508	213	0.26	1	63003623-63999110
	rs135902302-rs136049813	850	0.24	0	unknown
	rs42176463-rs133803377	319	0.20	29	35008674-35989729
C20:3 n-6	rs134302284-rs136338106	144	1.77	27	26000833-26975692
	rs136416983-rs134293611	373	0.38	1	69000931-69999565
	rs134477012-rs110532817	249	0.31	5	118004086-118995375
C20:5 n-3 (EPA)	rs110411459-rs137802105	202	2.19	10	29000222-29962997
	rs137693004-rs135981990	198	0.66	12	1009795-1998903
	rs137823965-rs42507195	199	0.66	10	49005824-49980174
C22:5 n-3 (DPA)	rs110411459-rs137802105	202	4.74	10	29000222-29962997
	rs132839318-rs134264692	306	3.35	3	27001604-27997941
	rs137823965-rs42507195	199	0.53	10	49005824-49980174
	rs137823965-rs42507195	199	0.70	10	49005824-49980174

Table 3 - Top three QTL regions associated with IMF deposition and composition traits in Nellore by Bayes B

Trait	QTL Window (first and last SNP)	Number of SNPs	%Gen. Variance	Chr	(conclusion) Map Position (UMD 3.1 bovine)
C22:6 n-3 (DHA)	rs109739360-rs137396104	232	0.46	7	99009734-99993114
	rs134412504-rs134289230	125	0.45	4	26004144-26999199
MUFA	rs133274959-rs136576856	213	3.24	3	25008982-26998679
	rs134160160-rs109942510	285	1.13	3	26002777-26998679
	rs133803779-rs29003226	155	0.46	3	51006901-51976646
PUFA	rs135174883-rs381709624	78	0.44	5	48109149-48993294
	rs109612389-rs133661384	132	0.42	17	24002089-24987264
	rs387618166-rs136897467	152	0.36	6	33002421-33992354
n-3	rs110411459-rs137802105	202	2.59	10	29000222-29962997
	rs137823965-rs42507195	199	2.25	10	49005824-49980174
	rs132839318-rs134264692	306	1.37	3	27001604-27997941
n-6	rs134302284-rs136338106	144	2.47	27	26000833-26975692
	rs132773171-rs42068328	123	0.77	25	33001144-33982544
	rs109016980-rs137260120	300	0.53	6	70000724-70995843
PUFA:SFA	rs42522569-rs108981640	203	0.2	18	64004373-64999126
	rs137135250-rs136369955	257	0.19	14	28001989-28998719
	rs136562362-rs43211407	315	0.18	1	2009127-2995356
n-6:n-3	rs136309202-rs43468791	216	0.74	12	58000573-58990041
	BTA-38252-rs133950976	229	0.48	16	27024889-27996193
	rs135537573-rs133828431	275	0.43	22	57004878-57970902
AI ¹	rs135990954-rs134959509	283	0.19	1	68013613-68999267
	rs42964986-rs110471760	286	0.19	28	40001693-40996451
	rs135812645-rs110103414	297	0.18	9	102003073-102995933

¹Atherogenic index = [12:0 + 4(14:0) + 16:0]/(SSFA + SPUFA)

²NS – Genomic region that explained less than 1% of genetic variance then was not studied

³NG – Genomic region with no genes associated with the specific trait

Table 4 - QTL regions associated with fatty acid composition in Nellore steers by Bayes B

(continues)

Group	Traits	QTL Window (first and last SNP)	Number of SNP /window ¹	%Variance explained SNP Window ¹	Chr	Map Position (UMD 3.1 bovine assembly)	PCG ²
Saturated Fatty Acid	C14:0	rs43328164 - rs134696015	295	1.06	3	6009105 - 6999643	- ³
	C18:0	rs133274959 - rs136576856	213	3.46	3	25008982 - 25990012	<i>HMGCS2, PHGDH, HSD3B1, HAO2</i>
	C16:0	rs134160160 - rs109942510	285	1.38	3	26002777 - 26998679	<i>WARS2</i>
	C14:0	rs41627556 - rs110731616	332	1.43	9	36013595 - 36997263	<i>GNG11, RGS5</i>
	C18:0	rs136405986 - rs137105475	114	1.08	10	74001100 - 74957582	<i>DHRS7</i>
	C18:0	rs132804279 - rs109645596	102	2.08	11	105009792 - 105985714	<i>NUP214</i>
	C16:0	rs109773631 - rs135618512	277	1.53	12	13000697 - 13997298	-
	C12:0	rs42924061 - rs134942057	214	3.48	12	60000226 - 60994461	<i>SLITRK6</i>
Monounsatur. Fatty Acid	C14:1 cis-9	rs137683417 - rs29003360	232	1.86	2	26002605 - 26999718	<i>GAD1, Sp5</i>
	C18:1 cis-9	rs134160160 - rs109942510	285	1.40	3	26002777 - 26998679	<i>WARS2</i>
	C16:1 cis-9	rs133274959-rs136576856	213	1.42	3	25008982 - 25990012	<i>HMGCS2, PHGDH, HSD3B1, HAO2</i>
	C18:1 cis-9	rs133274959-rs136576856	213	2.57	3	25008982 - 25990012	<i>HMGCS2, PHGDH, HSD3B1, HAO2</i>
	C18:1 t-16	rs133728493 - rs109630757	264	2.56	6	50008629 - 50996673	-
	C16:1cis-9	rs135892505 - rs109693564	184	1.04	7	80006438 - 80997126	-
	C16:1cis-9	rs136405986 - rs137105475	114	1.18	10	74001100 - 74957582	<i>DHRS7</i>
	C14:1cis-9	rs110517663 - rs137137362	127	1.43	11	27000446 - 27993515	<i>ABCG5</i>
	C18:1 cis-9	rs109773631 - rs135618512	277	1.91	12	13000697 - 13997298	-
	C14:1 cis-9	rs137754875 - rs135670282	203	1.55	12	66002783 - 66997356	<i>GPC6</i>
	C18:1 t-16	rs109872854 - rs134277482	250	2.17	17	41001273 - 41991683	<i>RAPGEF2, RPS27</i>

Table 4 - QTL regions associated with fatty acid composition in Nellore steers by Bayes B

(conclusion)

Group	Traits	QTL Window (first and last SNP)	Number of SNP /window ¹	%Variance explained SNP Window ¹	Chr	Map Position (UMD 3.1 bovine assembly)	PCG ²
Polyunsat. Fatty Acid	C22:5 n-3	rs132839318 - rs134264692	306	3.35	3	27001604 - 27997941	<i>SPAG7, WDR3</i>
	C18:2 t-11c-15	rs42404785 - rs133183089	144	1.69	3	72000121 - 72997801	<i>AQP7, LOXL2</i>
	C18:3 n-6	rs109612389 - rs133661384	205	1.34	7	34004443 - 34998434	-
	C18:2 t-11c-15	rs135377389 - rs133297940	267	3.49	8	68013821 - 68993778	-
	C22:5 n-3	rs110411459 - rs137802105	202	4.74	10	29000222 - 29962997	-
	C20:5 n-3	rs137823965 - rs42507195	199	2.19	10	49005824 - 49980174	<i>RORA</i>
	C18:3 n-3	rs42582725 - rs133775322	132	1.33	17	24002089 - 24987264	-
	C18:2 t-11c-15	rs137602675 - rs133056879	187	1.01	26	20001875 - 20999270	-
	C20:3 n-6	rs134302284 - rs136338106	144	1.77	27	26000833 - 26975692	-
Total of MUFA	MUFA	rs133274959 - rs136576856	213	3.24	3	25008982 - 25990012	<i>HMGCS2, PHGDH, HSD3B1, HAO2</i>
		rs134160160 - rs109942510	285	1.13	3	26002777 - 26998679	<i>WARS2</i>
Total of n-3	n-3	rs110411459 - rs137802105	202	2.59	10	29000222 - 29962997	-
		rs137823965 - rs42507195	199	2.25	10	49005824 - 49980174	-
		rs1322839318 - rs134264692	306	1.37	3	27001604 - 27997941	<i>SPAG7, WDR3</i>
Total of n-6	n-6	rs134302284 - rs136338106	144	2.47	27	26000833 - 26975692	-

¹1Mb window²Positional candidate genes³No putative candidate gene associated with the trait

Figure 1 - Manhattan plot of the genome-wide association study result for A) C12:0 (lauric acid) B) C16:0 (palmitic acid) C) C18:0 (stearic acid) in Nellore. The X-axis represents the chromosomes, and the Y-axis shows the proportion of genetic variance explained by SNP window from Bayes B analysis

BTA3 at 6 Mb and BTA9 at 36 Mb were associated with C14:0 fatty acid. In the first region, no PCG was identified. In the second region, the following PCGs were identified: guanine nucleotide binding protein (G protein), gamma 11 (*GNG11*) and regulator of G-protein signaling 5 (*RGS5*). These genes are associated with G proteins that have been implicated in the regulation of body weight and metabolic function, hyperinsulinemia, impaired glucose tolerance and resistance to insulin in mice (DENG et al., 2012).

BTA3 at 26 Mb and BTA12 at 13 Mb were associated with C16:0 fatty acid. In the first QTL region the tryptophanyl tRNA synthetase 2, mitochondrial (*WARS2*) gene was identified. In BTA12 at 13 Mb no PCG was identified. *WARS2* gene encodes an essential enzyme that catalyzes aminoacylation of tRNA with tryptophan. The *WARS2* protein contains a signal peptide for mitochondrial import, OMIM: 604733 (MCKUSICK, 2007), and is involved with regulation of fat distribution in human visceral and subcutaneous fat (SCHLEINITZ et al., 2013).

BTA3 at 25Mb, BTA10 at 74 Mb and BTA11 at 105 Mb QTL regions were associated with C18:0 fatty acid. On BTA3 at 25Mb four PCG were identified: 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2, mitochondrial (*HMGCS2*), phosphoglycerate dehydrogenase (*PHGDH*), hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1 (*HSD3B1*) and hydroxyacid oxidase 2, long chain (*HAO2*). *HMGCS2* was associated with fatty acid oxidation and ketogenesis in HepG2 cells (VILA-BRAU et al., 2011) and is directly regulated by *PPAR-α* gene that is an important key regulator of β -oxidation. On the other hand the *PPARG* expression induces lipogenesis by PXR activation in mice liver (WADA; GAO; XIE, 2009). *PHGDH* is related to the regulation of gene expression according to gene ontology (GO) terms (GO:0010468), *HSD3B1* gene is associated with steroid biosynthesis (GO:0006694) and metabolic process (GO: 0008202) and *HAO2* is associated with fatty acid oxidation (GO:0019395).

On BTA10 at 74 Mb QTL region, dehydrogenases/reductases (*SDRs*) (*DHRS7*) was identified. *DHRS7* catalyzes the oxidation/reduction of a wide range of substrates, including retinoid and steroids (HAESELEER; PALCZEWSKI, 2000) and has high expression level in adipocyte and skeletal muscle (WU et al., 2009). In addition, this gene is responsible for the final step in cholesterol production, the conversion of 7-dehydrocholesterol to cholesterol (PORTER, 2000).

On BTA11 at 105 Mb region, *NUP214* was identified. This gene is involved with intermembrane transport. The nucleoporins are characterized by phenylalanine-glycine rich

(*FG*) repeat sequences and the *FG* domains have an unfolded structure and are responsible for interaction with importin-cargo complexes that move through the pore. One of these importin-cargo complexes is sterol regulatory element binding proteins-sterol-sensing accessory factor (*SREBP-SCAP*). This factor enters the nucleus, then binds to sterol regulatory elements (*SRE*) in the promoter regions of genes, whose products mediate the synthesis of cholesterol and fatty acids (ZHOU et al., 2011).

2.3.5 Monounsaturated fatty acids

Ten genomic regions (1 Mb region) explained more than 1% of genotypic variation for monounsaturated fatty acids, which relates C14:1 cis-9, C16:1 cis-9, C18:1 cis-9, and C18:1 trans-16 (Table 4). These regions overlap with QTL reported for carcass weight, marbling score in Angus (MCCLURE et al., 2010), docosahexaenoic acid content in Charolais x Holstein crossbred cattle (GUTIERREZ-GIL et al., 2009) and palmitoleic acid in dairy cattle (MORRIS et al., 2007). Manhattan plots of the proportion of genetic variance explained by each 1-Mb window (2,527 windows) across the 29 autosomes for the most important fatty acid for beef quality and human health: myristoleic acid (C14:1 cis-9), palmitoleic acid (C16:1 cis-9), and oleic acid (C18:1 cis-9) and are in Figures 1, 2 and 3, respectively.

BTA2 at 26 Mb, BTA11 at 27 Mb, and BTA12 at 66 Mb were associated with C14:1 cis-9 fatty acid. On BTA2 at 26 Mb region, glutamate decarboxylase 1 (*GADI*) and specificity protein 5-transcription factor (*Sp5*) were identified. The *GADI* gene is involved with food behavior and insulin secretion. It was shown to be associated with morbid obesity (MEYRE et al., 2005) in humans, and body weight and daily gain in cattle (LI et al., 2012). *Sp5* transcription factors are involved in the regulation of pyruvate kinase, lactate dehydrogenase and fatty acid synthase in cancer cells (BLACK; BLACK; AZIZKHAN-CLIFFORD, 2001; ARCHER, 2011). On BTA11 at 27 Mb region, ATP-binding cassette, sub-family G (WHITE), member 5 (*ABCG5*) was identified. The *ABCG* family members were associated with cellular lipid-trafficking in macrophages and hepatocytes. This gene also presents an important role in the *PPARg* and *LXR* pathways. (SCHMITZ; LANGMANN; HEIMERL, 2001). On BTA12 at 66 Mb region, glypican 6 (*GPC6*) was identified. Glypicans family are involved with the control of cell growth and cell division. Other glypican gene, *GPC4* was associated with body fat distribution and obesity in humans (GESTA et al., 2006).

A previous GWAS using chicken population reported an association between *GPC6* and body weight, which has a stronger correlation with body fatness and obesity (GU et al., 2011).

BTA3 at 25 Mb, BTA7 at 80 Mb and BTA10 at 74 Mb were associated with C16:1 cis-9 fatty acid. Both BTA3 at 25Mb and BTA10 at 74 Mb regions were also associated with saturated fatty acids, where *HMGCs2*, *PHGDH*, *HSD3B1*, *HAO2*, and *DHRS7* were described above. On BTA7 at 80 Mb region, no PCG associated with this trait was identified.

BTA3 at 25 Mb, BTA3 at 26 Mb, and BTA12 at 13 Mb were associated with C18:1 cis-9. The first two regions were also associated with saturated fatty acids (C18:0 and C16:0, respectively) and C16:1 cis-9. On BTA12 at 13 Mb region, there were no annotated genes (Cattle Genome UMD3.1).

BTA6 at 50 Mb and BTA17 at 41 Mb were associated with C18:1 trans-16. On BTA 6 at 50 Mb region, no PCG associated with this trait was identified, and on BTA17 at 41 Mb region one PCG was found: Rap guanine nucleotide exchange factor (*RAPGEF2*), that selectively and non-covalently interacts with diacylglycerol, a diester of glycerol and two fatty acids (GO:0019992).

Four genomic regions were associated with saturated and monounsaturated fatty acids. BTA3 at 26 Mb and BTA12 at 13 Mb were associated with C16:0 and C18:1 cis-9. BTA3 at 25 Mb was associated with C18:0 and C18:1 cis-9. BTA3 at 25 and BTA10 at 74 Mb were associated with C18:0 and C16:1 cis-9. We also observed a negative phenotypic correlation between C16:0 and C18:1 ($r = -0.65$) and between C18:0 and C18:1 ($r = -0.72$). Palmitic acid (C16:0) carbon chain skeleton is a predominant source for fatty acid elongation and desaturation, which might generate palmitoleic (C16:1 cis-9), stearic (C18:0), and oleic (C18:1 cis-9) fatty acids (INAGAKI et al., 2002). Examination of the 30 markers with the larger effect in regions of BTA3 at 25 Mb (Table 5) and BTA3 at 26 Mb and BTA12 at 13 Mb (Table 6) revealed that alleles positively associated with saturated fatty acid were negatively associated with unsaturated fatty acids. This indicates that QTL in these regions are directly or indirectly involved with elongation and/or desaturation.

Table 5 - Top 30 markers effect in BTA3 at 25 Mb associated with C18:0 and C18:1 cis-9 in Nellore steers

BTA3 at 25 Mb			
SNP name	Marker Effect C18:0	Marker Effect 18:1 cis-9	Position (Chr_bp)
rs132903831	9.16E-03	-9.64E-03	3_25871241
rs109846051	7.74E-04	-1.62E-02	3_25768532
rs136685398	3.77E-04	-8.07E-04	3_25856394
rs133640984	3.10E-04	-9.15E-04	3_25724431
rs135852194	3.05E-04	-4.65E-03	3_25924712
rs135283257	1.94E-04	-3.01E-03	3_25968578
rs137571661	1.53E-04	-4.49E-03	3_25935272
rs110815636	1.31E-04	-4.17E-04	3_25899926
rs134138112	1.24E-04	-1.10E-03	3_25822926
rs133486019	1.07E-04	-1.49E-04	3_25653435
rs137726841	8.01E-05	-2.06E-04	3_25434706
rs43709927	7.45E-05	-6.09E-05	3_25361462
rs132978669	7.44E-05	-4.05E-04	3_25683554
rs134610317	6.80E-05	-6.01E-05	3_25586625
rs136526185	6.77E-05	-1.45E-03	3_25971232
rs132872603	6.60E-05	-3.28E-04	3_25760484
rs133573633	6.22E-05	-2.19E-04	3_25750133
rs110447077	5.05E-05	-6.27E-05	3_25239334
rs109380793	5.01E-05	-7.29E-05	3_25125560
rs133670803	4.94E-05	-5.21E-05	3_25187122
rs109125464	4.70E-05	-1.88E-04	3_25654265
rs133271406	4.51E-05	-9.20E-05	3_25311952
rs137417605	4.30E-05	-5.14E-05	3_25332450
rs134495781	4.23E-05	-1.58E-04	3_25013585
rs136479075	4.18E-05	-1.39E-04	3_25720847
rs109941469	4.17E-05	-6.65E-05	3_25107210
rs109791937	4.17E-05	-1.37E-05	3_25451628
rs136322611	4.04E-05	-8.51E-05	3_25185439
rs109561781	3.97E-05	-4.94E-05	3_25593196
rs134623220	3.81E-05	-2.88E-05	3_25651862

Table 6 - Top 30 markers effect in BTA3 at 26 Mb and BTA12 at 13 Mb QTL associated with C16:0 and C18:1 cis-9 in Nellore steers

BTA3 at 26 Mb				BTA12 at 13 Mb			
SNP Name	Marker Effect C16:0	Marker Effect 18:1 cis-9	Position (Chr_Bp)	SNP Name	Marker Effect C16:0	Marker Effect 18:1 cis-9	Position (Chr_bp)
rs136571109	1.49E-02	-8.42E-03	3_26426591	rs109310464	1.53E-02	-3.10E-02	12_13152910
rs110996601	1.45E-02	-6.58E-03	3_26425542	rs137546638	1.97E-03	-6.73E-03	12_13172924
rs134173038	1.40E-02	-6.05E-03	3_26427414	rs110182797	1.01E-03	-7.53E-04	12_13254936
rs137302352	4.10E-03	-2.29E-03	3_26400703	rs110951319	6.73E-04	-4.73E-04	12_13257070
rs135909241	9.74E-04	-2.55E-04	3_26446293	rs109014250	5.98E-04	-6.22E-04	12_13687406
rs109610305	7.91E-04	-1.72E-05	3_26497009	rs137052772	5.83E-04	-3.95E-04	12_13343491
rs137775131	5.12E-04	-7.50E-05	3_26501617	rs42625291	5.42E-04	-1.25E-03	12_13689345
rs109564367	5.01E-04	-8.52E-03	3_26073307	rs132782691	5.23E-04	-1.87E-04	12_13334837
rs110447932	4.30E-04	-8.35E-03	3_26072388	rs110645603	2.15E-04	-6.03E-05	12_13226418
rs109719131	3.84E-04	-9.90E-05	3_26465447	rs135746294	1.74E-04	-2.71E-05	12_13037523
rs110569793	3.62E-04	-9.64E-05	3_26476929	rs109497621	1.16E-04	-5.02E-05	12_13679288
rs133474163	3.39E-04	-1.56E-02	3_26143362	rs42625298	1.13E-04	-1.40E-04	12_13703328
rs110930607	3.22E-04	-1.90E-04	3_26332392	rs133583231	1.09E-04	-5.69E-05	12_13557660
rs134881809	3.10E-04	-2.36E-04	3_26505695	rs110721694	1.01E-04	-2.78E-03	12_13696926
rs111020805	2.70E-04	-1.92E-03	3_26108986	rs135841134	9.62E-05	-4.91E-05	12_13897680
rs134887465	2.68E-04	-2.04E-04	3_26385596	rs41667835	9.59E-05	-1.41E-04	12_13710147
rs110953593	2.61E-04	-8.47E-03	3_26075185	rs136432856	9.51E-05	3.08E-05	12_13493245
rs109658959	2.57E-04	-7.53E-03	3_26596824	rs133513235	9.45E-05	-4.21E-04	12_13251749
rs136570164	2.56E-04	-7.23E-05	3_26961906	rs137549822	8.60E-05	-6.65E-04	12_13081160
rs132696738	2.51E-04	-2.80E-04	3_26338744	rs110015470	8.45E-05	-3.18E-04	12_13682141
rs134559574	2.46E-04	-3.01E-05	3_26774155	rs133443666	8.30E-05	-1.46E-04	12_13520963
rs133573311	2.20E-04	-1.54E-04	3_26330114	rs208626835	8.10E-05	5.29E-06	12_13515819
rs135402139	1.94E-04	-1.26E-04	3_26954008	rs134142865	7.61E-05	-2.40E-04	12_13264826
rs135422840	1.71E-04	-3.00E-03	3_26020004	rs134240141	7.39E-05	-7.90E-05	12_13571494
rs136944072	1.66E-04	-2.63E-03	3_26149734	rs110659649	6.83E-05	-3.57E-05	12_13092209
rs110497471	1.56E-04	-2.50E-04	3_26576223	rs135136417	6.58E-05	-3.14E-05	12_13266029
rs137361087	1.45E-04	-3.15E-03	3_26144230	rs109064784	6.32E-05	-8.96E-05	12_13785143
rs110049045	1.37E-04	-7.04E-05	3_26975095	rs135831828	6.31E-05	-3.94E-05	12_13769919
rs133610187	1.36E-04	-2.31E-04	3_26290060	rs135423477	6.25E-05	-1.43E-04	12_13642422
rs110819650	1.35E-04	-4.69E-05	3_26979811	rs109147565	6.16E-05	-1.53E-03	12_13097096

2.3.6 Polyunsaturated fatty acids

Nine genomic regions (1 Mb window) explained > 1% of genotypic variation for group of polyunsaturated fatty acids, which relates C18:2 cis-9 cis-12 n-6, C18:2 trans-11 cis-15, C18:3 n-3, C18:3 n-6, C20:3 n-6, C20:5 n-3, and C22:5 n-3 (Table 3). These overlapped with QTL reported for marbling score, body weight in Angus cattle (MCCLURE et al., 2010), intramuscular fat, saturated fatty acid content, stearic acid in Fleckvieh bulls (BARTON et al., 2010) and RFI in half-sib families from Angus and Charolais (SHERMAN et al., 2009). Figure 2 shows in more detail the Manhattan plot of the proportion of genetic variance explained by window across the 29 autosomes for the important fatty acid to human health: γ -linolenic acid (C18:3 n-3), α -linolenic acid (C18:3 n-6) and total of n-3 (n-3).

BTA3 at 72 Mb, BTA 8 at 68 Mb, and BTA26 at 20 Mb were associated with C18:2 trans-11 cis-15 fatty acid. On BTA3 at 72 Mb QTL region, two PCG related to lipid metabolism were observed: aquaporin 7 (*AQP7*) and lysil oxidase-like2 (*LOXL2*). *AQP7* is involved with the *PPAR* signaling pathway (Huang Da, Sherman e Lempicki, 2009), while *LOXL2* is associated with lean body mass in mouse (BAKER et al., 2012). On BTA 8 at 68 Mb and BTA26 at 20 Mb no PCG was associated with this trait.

BTA10 at 49 Mb region was associated with C20:5 n-3 fatty acid, which harbors the *RAR*-related orphan receptor A (*RORA*) gene that is related to steroid hormone receptor activity. When combined with a steroid hormone, it produces the signal within the cell to initiate a change in cell activity or function (GO:0003707).

BTA3 at 27 Mb and BTA10 at 29 Mb QTL regions explained more than 1% of genetic variance for C22:3 n-3. The first region (BTA3 at 27 Mb) harbors two PCG: *SPAG17* and *WDR3*. These genes are involved with nucleus membrane (cellular components), the lipid bilayers that surround the nucleus and that form the nuclear envelope excluding the intermembrane space (GO:0031965). In the second region no PCG involved with this trait was identified. In other regions associated with C18:3 n-3 (BTA7 at 34 Mb), C18:3 n-6 (BTA17 at 24 Mb), and C20:3 n-6 (BTA27 at 26 Mb) no annotated genes were found (HU et al., 2013).

The total of omega-3 (n-3) and omega-6 (n-6) GWAS result is consistent with individual polyunsaturated n-3 and n-6 fatty acids, where the same QTL regions were associated with these traits (BTA3 at 27 Mb and BTA 27 at 26 Mb).

Figure 2 - Manhattan plot of the genome-wide association study result for A) C18:3 n-3 (γ -linolenic acid) B) C18:3 n-6 (α -linolenic acid) C) n-3 (total of n-3) in Nellore. The X-axis represents the chromosomes, and the Y-axis shows the proportion of genetic variance explained by SNP window from Bayes B analysis

GWAS have been used to investigate complex traits in many species including livestock (HAYES et al., 2009). Deposition and composition of fat in mammals are complex traits and are influenced by many loci throughout the genome (VAN LAERE et al., 2003). GWAS provides one approach to understand the genetic variation in complex traits, by identifying regions that can be fine-mapped to identify individual loci responsible for variation (KEMPER; GODDARD, 2012). In the present GWAS two interesting QTL regions BTA3 at 25 Mb and BTA3 at 26 Mb were associated with saturated, monounsaturated, and

polyunsaturated fatty acids in *Bos indicus* cattle. These regions harbor interesting PCG, which are involved with lipid metabolism in different species.

Previous studies reported CCAAT/enhancer binding proteins (*C/EBPβ*), peroxisome proliferator-activated receptor gamma (*PPARγ*), carnitine palmitoyltransferase-1 beta (*CPT-1β*), stearoylcoenzyme A desaturase (*SCD*), AMP-activated protein kinase alpha (*AMPKα*), and G-coupled protein receptor 43 (*GPR43*) genes to be related to fat deposition and fatty acid composition in Angus (SMITH et al., 2012). In this GWAS and in a recent GWAS publication using Angus cattle, the regions that harbor these genes were not associated with fat deposition and fatty acid composition (SAATCHI et al., 2013). However, the positional candidate genes *HMGCS2*, *PHGDH*, *HSD3B1* and *HAO2* identified in the present study are involved in the *PPAR* signaling pathway in human (PATH: hsa03320), carbon metabolism (PATH: bta01200), lipid metabolism and steroid hormone biosynthesis (PATH: bta00140), carbohydrate metabolism and glycosylate and decarboxylase metabolism (PATH: bta00630) according to Kegg: Kyoto Encyclopedia of Genes and Genomes (KANEHISA et al., 2012), respectively.

This is the first GWAS for intramuscular fat deposition and composition in Nellore. GWAS results in *Bos taurus* (Angus and Japanese Black cattle) reported different genomic regions associated with fat deposition and composition than those reported herein (UEMOTO et al., 2010; ISHII et al., 2013). Previously, it has been reported that a QTL on BTA19 was associated with fatty acid composition in *Bos taurus* breeds. This region on BTA19 harbors fatty acid synthase gene (*FASN*), which is an enzyme involved with *de novo* synthesis of long-chain fatty acid in mammalian, lipogenesis. *FASN* has been suggested as a candidate gene for fat traits in beef cattle (UEMOTO et al., 2010; ISHII et al., 2013; SAATCHI et al., 2013).

Differences in SNP allele frequencies and linkage disequilibrium profile (LD between SNPs and causal variants) may explain the different marker effects between *Bos indicus* and *Bos taurus* cattle (ESPIGOLAN et al., 2013; TIZIOTO et al., 2013). This explanation has been confirmed by Bolormaa and collaborators (2013), who compared marker effects using GWAS for *Bos taurus* and *Bos indicus* animals, which demonstrated that a SNP effect depends on the origin of alleles and the QTL segregation. The QTL segregation could result from mutation lost or fixation of alleles in one of the breeds, and also that the mutations occurred after divergence of these breeds (BOLORMAA et al., 2011). Furthermore, it is possible that the differences in physiological and metabolism factors could contribute to the observed differences between different breeds (LEHNERT et al., 2007).

2.4 Conclusion

The present study using BovineHD BeadChip (770k) identified several 1-Mb SNP regions and genes within these regions that were associated with IMF and FA. The values of genomic heritabilities described in this study have not been reported before for Nellore cattle, and this information is important for breeding programs interesting in improving these traits. IMF deposition and composition are considered complex traits (polygenic) and are moderately heritable. In this study it is apparent that IMF composition are affected by many loci with small effects. Identification of several genomic regions and putative positional genes associated with lipid metabolism reported here should contribute to the knowledge of the genetic basis of IMF and FA deposition and composition in Nellore cattle (*Bos indicus* breed) and lead to selection for those traits to improve human nutrition and health.

References

- AMERICAN OIL CHEMISTS' SOCIETY. **Rapid determination of oil/fat utilizing high temperature solvent (Official Procedure Am 5-04)**. Champaign, 2000. Available in: <http://www.ankom.com/media/documents/CrudeFat_0504_013009.pdf>. Accessed on: 15 dec. 2012.
- ARCHER, M.C. Role of sp transcription factors in the regulation of cancer cell metabolism. **Genes Cancer**, New York, v. 2, n. 7, p. 712-719, 2011.
- BAKER, E.J.; JAY, J.J.; BUBIER, J.A.; LANGSTON, M.A.; CHESLER, E.J. GeneWeaver: a web-based system for integrative functional genomics. **Nucleic Acids Research**, Oxford, v. 40, p. 1067-1076, 2012.
- BARTON, L.; KOTT, T.; BURES, D.; REHÁK, D.; ZAHŘÁDKOVÁ, R.; KOTTOVÁ, B. The polymorphisms of stearoyl-CoA desaturase (SCD1) and sterol regulatory element binding protein-1 (SREBP-1) genes and their association with the fatty acid profile of muscle and subcutaneous fat in Fleckvieh bulls. **Meat Science**, London, v. 85, p. 15-20, 2010.
- BLACK, A.R.; BLACK, J.D.; AZIZKHAN-CLIFFORD, J. Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. **Journal of Cell Physiology**, Philadelphia, v. 188, p. 143-60, 2001.
- BOLORMAA, S.; HAYES, B.J.; HAWKEN, R.J.; ZHANG, Y.; REVERTER, A.; GODDARD, M.E. Detection of chromosome segments of zebu and taurine origin and their effect on beef production and growth. **Journal of Animal Science**, Savoy, v. 89, p. 2050-2060, 2011.

BOLORMAA, S.; PRYCE, J.E.; KEMPER, K.E.; HAYES, B.J.; ZHANG, Y.; TIER, B.; BARENDSE, W.; REVERTER, A.; GODDARD, M.E. Detection of quantitative trait loci in *Bos indicus* and *Bos taurus* cattle using genome-wide association studies. **Genetics Selection Evolution**, Paris, v. 45, 2013. Available in: <<http://www.gsejournal.org/content/pdf/1297-9686-45-43.pdf>>. Accessed on: 23 oct. 2013.

BRESSAN, M.C.; ROSSATO, L.V.; RODRIGUES, E.C.; ALVES, S.P.; BESSA, R.J.; RAMOS, E.M.; GAMA, L.T. Genotype x environment interactions for fatty acid profiles in *Bos indicus* and *Bos taurus* finished on pasture or grain. **Journal of Animal Science**, Savoy, v. 89, p. 221-232, 2011.

CASAS, E.; SHACKELFORD, S.D.; KEELE, J.W.; KOOHMARAIE, M.; SMITH, T.P.; STONE, R.T. Detection of quantitative trait loci for growth and carcass composition in cattle. **Journal of Animal Science**, Savoy, v. 81, n. 12, p. 2976-2983, 2003.

CASAS, E.; STONE, R.T.; KEELE, J.W.; SHACKELFORD, S.D.; KAPPES, S.M.; KOOHMARAIE, M. A comprehensive search for quantitative trait loci affecting growth and carcass composition of cattle segregating alternative forms of the myostatin gene. **Journal of Animal Science**, Savoy, v. 79, n. 4, p. 854-860, 2001.

CATTLE genome UMD3.1. Available in: <<http://www.animalgenome.org/cgi-bin/gbrowse/bovine/>>. Accessed on: 20 dec. 2013.

CHRISTIE, W.W. The composition and structure of milk lipids. **Developments in Dairy Chemistry**. New York, v. 2, p. 1-35, 1983.

DENG, W.; WANG, X.; XIAO, J.; CHEN, K.; ZHOU, H.; SHEN, D.; LI, H.; TANG, Q. Loss of regulator of G protein signaling 5 exacerbates obesity, hepatic steatosis, inflammation and insulin resistance. **PLoS ONE**, London, v. 7, n. 1, p. e30256, 2012. Available in: <<http://dx.doi.org/10.1371/journal.pone.0030256>>. Accessed on: 18 oct. 2013.

DU, M.; HUANG, Y.; DAS, A.K.; YANG, Q.; DUARTE, M.S.; DODSON, M.V.; ZHU, M.J. Manipulating mesenchymal progenitor cell differentiation to optimize performance and carcass value of beef cattle. **Journal of Animal Science**, Savoy, v. 91, p. 1419-1427, 2013.

DUARTE, M.S.; PAULINO, P.V.; DAS, A.K.; WEI, S.; SERÃO, N.V.; FU, X.; HARRIS, S.M.; DODSON, M.V.; DU, M. Enhancement of adipogenesis and fibrogenesis in skeletal muscle of Wagyu compared with Angus cattle. **Journal of Animal Science**, Savoy, v. 91, p. 2938-2946, 2013.

ESPIGOLAN, R.; BALDI, F.; BOLIGON, A.A.; SOUZA, F.R.; GORDO, D.G.; TONUSSI, R.L.; CARDOSO, D.F.; OLIVEIRA, H.N.; TONHATI, H.; SARGOLZAEI, M.; SCHENKEL, F.S.; CARVALHEIRO, R.; FERRO, J.A.; ALBUQUERQUE, L.G. Study of whole genome linkage disequilibrium in nellore cattle. **BMC Genomics**, London, v. 14, 2013. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2164-14-305.pdf>>. Accessed on: 15 sep. 2013.

FLICEK, P.; AHMED, I.; AMODE, M.R.; BARRELL, D.; BEAL, K.; BRENT, S.; CARVALHO-SILVA, D.; CLAPHAM, P.; COATES, G.; FAIRLEY, S.; FITZGERALD, S.; GIL, L.; GARCÍA-GIRÓN, C.; GORDON, L.; HOURLIER, T.; HUNT, S.; JUETTEMANN, T.; KÄHÄRI, A.K.; KEENAN, S.; KOMOROWSKA, M.; KULESHA, E.; LONGDEN, I.; MAUREL, T.; MCLAREN, W.M.; MUFFATO, M.; NAG, R.; OVERDUIN, B.; PIGNATELLI, M.; PRITCHARD, B.; PRITCHARD, E.; RIAT, H.S.; RITCHIE, G.R.; RUFFIER, M.; SCHUSTER, M.; SHEPPARD, D.; SOBRAL, D.; TAYLOR, K.; THORMANN, A.; TREVANION, S.; WHITE, S.; WILDER, S.P.; AKEN, B.L.; BIRNEY, E.; CUNNINGHAM, F.; DUNHAM, I.; HARROW, J.; HERRERO, J.; HUBBARD, T.J.; JOHNSON, N.; KINSELLA, R.; PARKER, A.; SPUDICH, G.; YATES, A.; ZADISSA, A.; SEARLE, S.M. Ensembl 2013. **Nucleic Acids Research**, Oxford, v. 41, p. 48-55, 2013.

GARRICK, D. J.; FERNANDO, R. L. Implementing a QTL detection study (GWAS) using genomic prediction methodology. **Methods in Molecular Biology**, Clifton, v. 1019, p. 275-298, 2013.

GESTA, S.; BLÜHER, M.; YAMAMOTO, Y.; NORRIS, A.W.; BERNDT, J.; KRALISCH, S.; BOUCHER, J.; LEWIS, C.; KAHN, C.R. Evidence for a role of developmental genes in the origin of obesity and body fat distribution. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 103, n. 17, p. 6676-6681, 2006.

GILMOUR, A. R.; GOGEL, B.J.; CULLIS, B.R.; THOMPSON R. **ASReml user guide release 3.0**. VSN International, 2009. Available in: <<http://vsni.de/downloads/asreml/release3/UserGuide.pdf>>. Accessed on: 26 sep. 2013.

GU, X.; FENG, C.; MA, L.; SONG, C.; WANG, Y.; DA, Y.; LI, H.; CHEN, K.; YE, S.; GE, C.; HU, X.; LI, N. Genome-wide association study of body weight in chicken F2 resource population. **PLoS One**, London, v. 6, n. 7, p. e21872, 2011. Available in: <<http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0021872&representation=PDF>>. Accessed on: 26 sep. 2013.

GUTIERREZ-GIL, B.; WILLIAMS, J.L.; HOMER, D.; BURTON, D.; HALEY, C.S.; WIENER, P. Search for quantitative trait loci affecting growth and carcass traits in a cross population of beef and dairy cattle. **Journal of Animal Science**, Savoy, v. 87, p. 24-36, 2009.

HABIER, D.; FERNANDO, R.L.; KIZILKAYA, K.; GARRICK, D.J. Extension of the Bayesian alphabet for genomic selection. **BMC Bioinformatics**, London, v. 12, 2011. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2105-12-186.pdf>>. Accessed on: 26 sept. 2013.

HAESELEER, F.; PALCZEWSKI, K. Short-chain dehydrogenases/reductases in retina. **Methods in Enzymology**, New York, v. 316, p. 372-383, 2000.

HARA, A.; RADIN, N.S. Lipid extraction of tissues with a low-toxicity solvent. In: (Ed.). **Analytical Biochemistry**, New York, v. 90, p. 420-426, 1978.

HAYES, B.J.; BOWMAN, P.J.; CHAMBERLAIN, A.C.; VERBYLA, K.; GODDARD, M.E. Accuracy of genomic breeding values in multi-breed dairy cattle populations. **Genetics Selection Evolution**, Paris, v. 41, p. 51, 2009. Available in: <<http://www.gsejournal.org/content/pdf/1297-9686-41-51.pdf>>. Accessed on: 20 oct. 2013.

HOLLOWAY, J.W.; SAVELL, J.W.; HAMBY, P.L.; BAKER, J.F.; STOUFFER, J.R. Relationships of empty-body composition and fat distribution to live animal and carcass measurements in *Bos indicus*-*Bos taurus* crossbred cows. **Journal of Animal Science**, Savoy, v. 68, n. 7, p. 1818-1826, 1990.

HU, Z.L.; PARK, C.A.; WU, X.L.; REECY, J.M. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. **Nucleic Acids Research**, Oxford, v. 41, p. 871-879, 2013.

HUANG DA, W.; SHERMAN, B.T.; LEMPICKI, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. **Nucleic Acids Research**, Oxford, v. 37, n. 1, p. 1-13, 2009.

HUERTA-LEIDENZ, N.O.; CROSS, H.R.; SAVELL, J.W.; LUNT, D.K.; BAKER, J.F.; SMITH, S.B. Fatty acid composition of subcutaneous adipose tissue from male calves at different stages of growth. **Journal of Animal Science**, Savoy, v. 74, n. 6, p. 1256-1264, 1996.

ILLUMINA. **GenomeStudio™ Software v2011.1 Release Notes v. 2013**. Available in: <http://supportres.illumina.com/documents/myillumina/1f2d00ae-b759-489b-82f9-0163ce085ae7/genomestudio_2011.1_release_notes.pdf>. Accessed on: 15 dec. 2012.

INAGAKI, K.; AKI, T.; FUKUDA, Y.; KAWAMOTO, S.; SHIGETA, S.; ONO, K.; SUZUKI, O. Identification and expression of a rat fatty acid elongase involved in the biosynthesis of C18 fatty acids. **Bioscience, Biotechnology, and Biochemistry**, Tokyo, v. 66, n. 3, p. 613-21, 2002.

ISHII, A.; YAMAJI, K.; UEMOTO, Y.; SASAGO, N.; KOBAYASHI, E.; KOBAYASHI, N.; MATSUHASHI, T.; MARUYAMA, S.; MATSUMOTO, H.; SASAZAKI, S.; MANNEN, H. Genome-wide association study for fatty acid composition in Japanese Black cattle. **Journal of Animal Science**, Savoy, v. 84, p. 675-678, 2013.

JAKOBSEN, M.U.; OVERVAD, K.; DYERBERG, J.; HEITMANN, B.L. Intake of ruminant trans fatty acids and risk of coronary heart disease. **International Journal of Epidemiology**, London, v. 37, p. 173-182, 2008.

KANEHISA, M.; GOTO, S.; SATO, Y.; FURUMICHI, M.; TANABE, M. KEGG for integration and interpretation of large-scale molecular data sets. **Nucleic Acids Research**, Oxford, v. 40, p. 109-114, 2012.

KEMPER, K.E.; GODDARD, M.E. Understanding and predicting complex traits: knowledge from cattle. **Human Molecular Genetics**, London, v. 21, p. 45-51, 2012.

KIZILKAYA, K.; FERNANDO, R.L.; GARRICK, D.J. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. **Journal of Animal Science**, Savoy, v. 88, p. 544-551, 2010.

KIZILKAYA, K.; TAIT, R.G.; GARRICK, D.J.; FERNANDO, R.L.; REECY, J.M. Genome-wide association study of infectious bovine keratoconjunctivitis in angus cattle. **BMC Genetics**, London, v. 14, 2013. Available in:

<<http://www.biomedcentral.com/content/pdf/1471-2156-14-23.pdf>>. Accessed on: 21 oct. 2013.

LABORDE, F.L.; MANDELL, I.B.; TOSH, J.J.; WILTON, J.W.; BUCHANAN-SMITH J.G. Breed effects on growth performance, carcass characteristics, fatty acid composition, and palatability attributes in finishing steers. **Journal of Animal Science**, Savoy, v. 79, n. 2, p. 355-365, 2001.

LEHNERT, S.A.; REVERTER, A.; BYRNE, K.A.; WANG, Y.; NATTRASS, G.S.; HUDSON, N.J.; GREENWOOD, P.L. Gene expression studies of developing bovine longissimus muscle from two different beef cattle breeds. **BMC Developmental Biology**, London, v. 7, 2007. Available in: <<http://www.biomedcentral.com/content/pdf/1471-213X-7-95.pdf>>. Accessed on: 18 oct. 2013.

LI, C.; ALDAI, N.; VINSKY, M.; DUGAN, M.E.; MCALLISTER, T.A. 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. **Animal Genetics**, Oxford, v. 43, n. 1, p. 93-97, 2012.

MATUKUMALLI, L.K.; LAWLEY, C.T.; SCHNABEL, R.D.; TAYLOR, J.F.; ALLAN, M.F.; HEATON, M.P.; O'CONNELL, J.; MOORE, S.S.; SMITH, T.P.; SONSTEGARD, T.S.; VAN TASSELL, C.P. Development and characterization of a high density SNP genotyping assay for cattle. **PLoS One**, London, v. 4, n. 4, p. e5350, 2009. Available in: <<http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0005350&representation=PDF>>. Accessed on: 26 sep. 2013.

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.; NORTH CUTT, S.L.; TAYLOR, J.F. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. **Animal Genetics**, London, v. 41, p. 597-607, 2010.

MCKUSICK, V.A. Mendelian inheritance in man and its online version, OMIM. **American Journal of Human Genetics**, Chicago, v. 80, n. 4, p. 588-604, 2007.

MCNAUGHT, A.D.; WILKINSON, A. **IUPAC**: compendium of chemical terminology. London: International Union of Pure and Applied Chemistry, 1997. 366 p.

MCNEILL, S.; VAN ELSWYK, M.E. Red meat in global nutrition. **Meat Science**, London, v. 92, p. 166-173, 2012.

MEUWISSEN, T.H.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, Oxford, v. 157, n. 4, p. 1819-1829, 2001.

MEYER, L.R.; ZWEIG, A.S.; HINRICHS, A.S.; KAROLCHIK, D.; KUHN, R.M.; WONG, M.; SLOAN, C.A.; ROSENBLOOM, K.R.; ROE, G.; RHEAD, B.; RANEY, B.J.; POHL, A.; MALLADI, V.S.; LI, C.H.; LEE, B.T.; LEARNED, K.; KIRKUP, V.; HSU, F.; HEITNER, S.; HARTE, R.A.; HAEUSSLER, M.; GURUVADOO, L.; GOLDMAN, M.; GIARDINE, B.M.; FUJITA, P.A.; DRESZER, T.R.; DIEKHANS, M.; CLINE, M.S.; CLAWSON, H.; BARBER, G.P.; HAUSSLER, D.; KENT, W.J. The UCSC Genome Browser database: extensions and updates 2013. **Nucleic Acids Research**, Oxford, v. 41, p. 64-69, 2013.

MEYRE, D.; BOUTIN, P.; TOUNIAN, A.; DEWEIRDER, M.; AOUT, M.; JOURET, B.; HEUDE, B.; WEILL, J.; TAUBER, M.; TOUNIAN, P.; FROGUEL, P. Is glutamate decarboxylase 2 (GAD2) a genetic link between low birth weight and subsequent development of obesity in children? **Journal of Clinical Endocrinology and Metabolism**, Baltimore, v. 90, p. 2384-2390, 2005.

MORRIS, C.A.; CULLEN, N.G.; GLASS, B.C.; HYNDMAN, D.L.; MANLEY, T.R.; HICKEY, S.M.; MCEWAN, J.C.; PITCHFORD, W.S.; BOTTEMA, C.D.; LEE, M.A. Fatty acid synthase effects on bovine adipose fat and milk fat. **Mammalian Genome**, Chapel Hill, v. 18, n. 1, p. 64-74, 2007.

NOGI, T.; HONDA, T.; MUKAI, F.; OKAGAKI, T.; OYAMA, K. Heritabilities and genetic correlations of fatty acid compositions in longissimus muscle lipid with carcass traits in Japanese Black cattle. **Journal of Animal Science**, Savoy, v. 89, p. 615-621, 2011.

NUERNBERG, K.; DANNENBERGER, D.; NUERNBERG, G.; ENDER, K.; VOIGT, J.; SCOLLAN, N.D.; WOOD, J.D.; NUTE, G.R.; RICHARDSON, R.I. Effect of grass-based and a concentrate feeding system on meat quality characteristics and fatty acid composition of longissimus muscle in different cattle breeds. **Livestock Production Science**, Amsterdam, v. 94, p. 137-147, 2005.

ONTERU, S.K.; FAN, B.; NIKKILÄ, M.T.; GARRICK, D.J.; STALDER, K.J.; ROTHSCCHILD, M.F. Whole-genome association analyses for lifetime reproductive traits in the pig. **Journal of Animal Science**, Savoy, v. 89, p. 988-995, 2011.

PADRE, R.; ARICETTI, J.A.; MOREIRA, F.B.; MIZUBUTI, I.Y.; DO PRADO, I.N.; VISENTAINER, J.V.; DE SOUZA, N.E.; MATSUSHITA, M. Fatty acid profile, and chemical composition of Longissimus muscle of bovine steers and bulls finished in pasture system. **Meat Science**, London, v. 74, p. 242-248, 2006.

PETERS, S.O.; KIZILKAYA, K.; GARRICK, D.J.; FERNANDO, R.L.; REECY, J.M.; WEABER, R.L.; SILVER, G.A.; THOMAS, M.G. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. **Journal of Animal Science**, Savoy, v. 90, p. 3398-3409, 2012.

PICARD, B.; LEFAUCHEUR, L.; BERRI, C.; DUCLOS, M.J. Muscle fibre ontogenesis in farm animal species. **Reproduction, Nutrition, Development**, Les Ulis, v. 42, n. 5, p. 415-431, 2002.

PORTER, F.D. RSH/Smith-Lemli-Opitz syndrome: a multiple congenital anomaly/mental retardation syndrome due to an inborn error of cholesterol biosynthesis. **Molecular Genetics and Metabolism**, San Diego, v. 71, p. 163-174, 2000.

PRADO, I.N. do; MOREIRA, F.B.; MATSUSHITA, M.; DE SOUZA, N.E. Longissimus dorsi fatty acids composition of Bos indicus and Bos indicus x Bos taurus crossbred steers finished in pasture. **Brazilian Archives of Biology and Technology**, Londrina, v. 46, n. 4, p. 601-608, 2003.

PRAKASH, A.; STIGLER, M. **FAO statistical yearbook 2012: world food and agriculture**. OECD; FAO, 2012. Available in: <<http://reliefweb.int/sites/reliefweb.int/files/resources/i2490e.pdf>>. Accessed on: 20 oct. 2013.

RUIZ, M.; MATSUSHITA, M.; VISENTAINER, J.; HERNANDEZ, J.A.; RIBEIRO, E.L. DE A.; SHIMOKOMAKI, M.; REEVES, J.J.; DE SOUZA, N.E. Proximate chemical composition and fatty acid profiles of Longissimus thoracis from pasture fed LHRH immunocastrated, castrated and intact Bos indicus bulls. **South African Journal of Animal Science**, Pretoria, v. 35, n. 1, p. 13-18, 2005.

RULE, D.C.; MACNEIL, M.D.; SHORT, R.E. Influence of sire growth potential, time on feed, and growing-finishing strategy on cholesterol and fatty acids of the ground carcass and longissimus muscle of beef steers. **Journal of Animal Science**, Savoy, v. 75, n. 6, p. 1525-1533, 1997.

SAATCHI, M.; GARRICK, D.J.; TAIT, R.G. JR.; MAYES, M.S.; DREWNOSKI, M.; SCHOONMAKER, J.; DIAZ, C.; BEITZ, D.C.; REECY, J.M. Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. **BMC Genomics**, London, v. 14, n. 1, 2013. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2164-14-730.pdf>>. Accessed on: 4 nov. 2013.

SCHERER, P.E. Adipose tissue: from lipid storage compartment to endocrine organ. **Diabetes**, New York, v. 55, p.1537-1545, 2006.

SCHLEINITZ, D.; KLÖTING, N.; LINDGREN, C.M.; BREITFELD, J.; DIETRICH, A.; SCHÖN, M.R.; LOHMANN, T.; DREBLER, M.; STUMVOLL, M.; MCCARTHY, M.I.; BLÜHER, M.; KOVACS, P. Fat depot-specific mRNA expression of novel loci associated with waist-hip ratio. **International Journal of Obesity**, London, v. 38, p. 120-125, 2013.

SCHMITZ, G.; LANGMANN, T.; HEIMERL, S. Role of ABCG1 and other ABCG family members in lipid metabolism. **Journal of Lipid Research**, Bethesda, v. 42, n. 10, p. 1513-1520, 2001.

SHERMAN, E.L.; NKRUMAH, J.D.; LI, C.; BARTUSIAK, R.; MURDOCH, B.; MOORE, S.S. Fine mapping quantitative trait loci for feed intake and feed efficiency in beef cattle. **Journal of Animal Science**, Savoy, v. 87, p. 37-45, 2009.

SHI, H.; KOKOEVA, M.V.; INOUE, K.; TZAMELI, I.; YIN, H.; FLIER, J.S. TLR4 links innate immunity and fatty acid-induced insulin resistance. **Journal of Clinical Investigation**, New York, v. 116, n. 11, p. 3015-3025, 2006.

SIMOPOULOS, A.P. The importance of the ratio of omega-6/omega-3 essential fatty acids. **Biomedicine and Pharmacotherapy**, Paris, v. 56, n. 8, p. 365-379, 2002.

_____. The omega-6/omega-3 fatty acid ratio, genetic variation, and cardiovascular disease. **Asia Pacific Journal of Clinical Nutrition**, Taipei, v. 17, suppl. 1, p. 131-134, 2008.

SMITH, S.B.; GO, G.W.; JOHNSON, B.J.; CHUNG, K.Y.; CHOI, S.H.; SAWYER, J.E.; SILVEY, D.T.; GILMORE, L.A.; GHAHRAMANY, G.; KIM, K.H. Adipogenic gene expression and fatty acid composition in subcutaneous adipose tissue depots of Angus steers between 9 and 16 months of age. **Journal of Animal Science**, Savoy, v. 90, p. 2505-2514, 2012.

TAIT, R. JR.; ZHANG, S.; KNIGHT, T.; BORMANN, J.M.; STROHBEHN, D.; BEITZ, D.; REECY, J. Heritability estimates for fatty acid quantity in Angus beef. **Journal of Animal Science**, Savoy, v. 85, suppl. 2, p. 58, 2007.

TAKAYA, J.; YAMATO, F.; KUROYANAGI, Y.; HIGASHINO, H.; KANEKO, K. Intracellular magnesium of obese and type 2 diabetes mellitus children. **Diabetes Therapy**, London, v. 1, n. 1, p. 25-31, 2010.

TIZIOTO, P.C.; MEIRELLES, S.L.; VENERONI, G.B.; TULLIO, R.R.; ROSA, A.N.; ALENCAR, M.M.; MEDEIROS, S.R.; SIQUEIRA, F.; FEIJÓ, G.L.; SILVA, L.O.; TORRES, R.A.JR.; REGITANO, L.C. A SNP in ASAP1 gene is associated with meat quality and production traits in nelore breed. **Meat Science**, London, v. 92, p. 855-857, 2012.

TIZIOTO, P.C.; DECKER, J.E.; TAYLOR, J.F.; SCHNABEL, R.D.; MUDADU, M.A.; SILVA, F.L.; MOURÃO, G.B.; COUTINHO, L.L.; THOLON, P.; SONSTEGARD, T.S.; ROSA, A.N.; ALENCAR, M.M.; TULLIO, R.R.; MEDEIROS, S.R.; NASSU, R.T.; FEIJÓ, G.L.; SILVA, L.O.; TORRES, R.A.; SIQUEIRA, F.; HIGA, R.H.; REGITANO, L.C.A. A genome scan for meat quality traits in nelore beef cattle. **Physiological Genomics**, Bethesda, v. 45, p. 1012-1020, 2013.

UEMOTO, Y.; ABE, T.; TAMEOKA, N.; HASEBE, H.; INOUE, K.; NAKAJIMA, H.; SHOJI, N.; KOBAYASHI, M.; KOBAYASHI, E. Whole-genome association study for fatty acid composition of oleic acid in Japanese Black cattle. **Animal Genetics**, Oxford, v. 42, p. 141-148, 2010.

UNITED STATES DEPARTMENT OF AGRICULTURE. **National nutrient database for standard reference, release 23**. Beltsville: USDA, Agricultural Research Service, Nutrient Data Laboratory, 2010. Available in:

<https://www.ars.usda.gov/SP2UserFiles/Place/12354500/Data/SR23/sr23_doc.pdf>.

Accessed on: 23 oct. 2013.

VAN LAERE, A.; NGUYEN, M.; BRAUNSCHWEIG, M.; NEZER, C.; COLLETTE, C.; MOREAU, L.; ARCHIBALD, A.L.; HALEY, C.S.; BUYS, N.; TALLY, M.; ANDERSSON, G.; GEORGES, M.; ANDERSSON, L. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. **Nature**, London, v. 425, p. 832-836, 2003.

VANRADEN, P.M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, Champaign, v. 91, p. 4414-4423, 2008.

VILÀ-BRAU, A.; DE SOUSA-COELHO, A.L.; MAYORDOMO, C.; HARO, D.; MARRERO, P.F. Human HMGCS2 regulates mitochondrial fatty acid oxidation and FGF21 expression in HepG2 cell line. **Journal of Biological Chemistry**, Bethesda, v. 286, n. 23, p. 20423-20430, 2011.

WADA, T.; GAO, J.; XIE, W. PXR and CAR in energy metabolism. **Trends in Endocrinology and Metabolism**, New York, v. 20, p. 273-279, 2009.

WU, C.; OROZCO, C.; BOYER, J.; LEGLISE, M.; GOODALE, J.; BATALOV, S.; HODGE, C.L.; HAASE, J.; JANES, J.; HUSS, J.W.; SU, A.I. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. **Genome Biology**, London, v. 10, n. 11, p. R130, 2009. Available in: <http://genomebiology.com/content/pdf/gb-2009-10-11-r130.pdf>. Accessed on: 18 oct. 2013.

ZHOU, T.; LI, S. ZHONG, W.; VIHervaara, T.; BÉASLAS, O.; PERTTILÄ, J.; LUO, W.; JIANG, Y.; LEHTO, M.; OLKKONEN, V.M.; YAN, D. OSBP-related protein 8 (ORP8) regulates plasma and liver tissue lipid levels and interacts with the nucleoporin Nup62. **PLoS One**, London, v. 6, n. 6, p. e21078, 2011. Available in: <http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0021078&representation=PDF>. Accessed on: 20 oct. 2013.

3 PUTATIVE REGULATORY FACTORS INVOLVED IN MOLECULAR MECHANISMS OF INTRAMUSCULAR FAT DEPOSITION IN NELLORE STEERS

Abstract

The deposition of intramuscular fat (IMF) is controlled by many genes, which participate directly, or indirectly in adipogenesis and fat metabolism. Intramuscular fat quantity is an important economic phenotype, which influences the sensorial and nutritional value of beef. Previous research has documented that muscle and fat tissues from *Bos indicus* cattle develop in a different manner than *Bos taurus* breeds. The goal of this study was to identify differentially expressed genes (DEG), biological pathways and putative regulatory factors involved with fat deposition in steers of the Nellore breed that were selected for extreme genomic estimated breeding value (GEBV) for IMF. The next generation sequencing (NGS) technology, bioinformatics tools and mathematical methods, such as partial correlation coefficient with information theory (PCIT), regulatory impact factor (RIF) and phenotypic impact factor (PIF) were utilized to better understand lipid metabolism. A total of 16,101 common expressed genes were identified in both groups (H - high and L- low GEBV) and a total of 77 differentially expressed genes (DEG) were identified between two groups. Of the 77 DEG, 41 genes were up-regulated and 36 down-regulated in the L group compared to H group. Pathway Studio software revealed 13 significant pathways that involved these DEG, which are related with lipid metabolism and other biological processes such as inflammatory system and glucose homeostasis. PCIT analyses identified genes with a significant difference in the number of gene-gene correlations between high and low group and detected putative regulatory factors involved in lipid metabolism. Interesting candidate genes identified by PCIT include: *ANKRD26*, *HOX5C* and *PPAPDC2*. RIF and PIF analyses identified several strong candidate genes, including: *GLI2* and *IGF2* (RIF1), *MPC1* and *UBL5* (RIF2) and a host of small RNAs, including miR-1281 (PIF). The present study showed the complexity of the bovine *Longissimus dorsi* transcriptome and the molecular mechanisms involved in lipid metabolism. Differential gene expression and pathway enrichment analysis revealed the transcriptome profile between Nellore steers with different percentage of IMF and indicated the stage of fat deposition of these animals. The mathematical methods PCIT, PIF and RIF allowed the identification of putative transcription factors involved in adipogenesis and lipid metabolism.

Keywords: Adipogenesis; *Bos indicus*; Marbling; *Longissimus dorsi*; RNA-Seq

3.1 Introduction

Intramuscular fat (IMF, also known as marbling) represents the amount of fat within skeletal muscle (generally measured on the *Longissimus dorsi* (LD)), which is the sum of phospholipids (present in cell membranes), and triglycerides (lipid droplets). The amount of IMF in beef cattle is related with meat quality and human health. High intramuscular fat content (marbling) has been associated with tenderness, juiciness and consumer satisfaction (O'QUINN et al., 2012). At the same time, red meat consumption or more specifically

saturated fat consumption has been associated with human diseases such as cardiovascular disease, obesity and colon cancer. This association has led for the demand of beef with low fat and high quality by consumers. In addition, consumers in some countries have different preferences of IMF amount. For example, consumers in Asia and North America desire beef with high IMF, while Europeans prefer lean beef, with low IMF (HOCQUETTE et al., 2013).

Intramuscular fat content is a polygenic trait regulated by many genes involved directly, or indirectly in adipogenesis and fat metabolism (MICHAL et al., 2006). The deposition of IMF is influenced by many factors such as sex, age, breed, nutrition, and genetics (RULE; MACNEIL; SHORT, 1997). To date, the biological and functional mechanisms involved with this trait are still unclear. Understanding the genetic control of fat deposition and composition in muscle is also important for human health. From a nutritional point of view, the ability to breed for a reduced IMF profile will allow the beef industry to produce healthier products. From the medical point of view, understanding intramuscular fat metabolism will help to prevent diseases such as obesity and type 2 diabetes.

RNA-Seq has been used in livestock to provide insight into the biological mechanisms of complex traits. For example, recent RNA-Seq studies of pigs with extreme IMF phenotypes have defined differentially expressed genes (DEG) and potential gene networks important in lipid and fatty acid metabolism in the liver (RAMAYO-CALDAS et al., 2012). In beef cattle, DEG were identified in subcutaneous fat from *Bos taurus* crossed steers, which revealed that expression pattern depends on the genetic background (JIN et al., 2012). However, there have been no published studies investigating the transcriptional control of IMF deposition in bovine species.

Insight into the biological mechanisms involved with complex traits requires analysis of genetic networks, as well as the determination of relationships between genes and networks. However, transcriptome analyses depend on computational tools (i.e. algorithms and their efficiency) and capacity (i.e. RAM and CPU hardware resources). A novel algorithm for the Partial Correlation Coefficient with Information Theory (PCIT) mathematical method was developed to determine the relationships between genes and networks. Previously, the PCIT algorithm approach identified causal regulatory changes in myostatin gene expression in beef cattle (HUDSON; REVERTER; DALRYMPLE, 2009) by co-expression network analysis.

Differential hubbing (DH) or differential connectivity (DC) is the difference in the number of significant connections, measured as partial correlations, a gene has in two different states, which is computed by PCIT algorithm. In other words, DH is the change in

the number of significant connections between two states. Previous studies discovered that the extreme values of DH when enriched show higher percentage of transcriptional regulator factors than differential expression (DE) analysis (HUDSON; REVERTER; DALRYMPLE, 2009).

DH approach is not applicable to assess the importance of each DEG on phenotype changing, so Hudson, Reverter and Dalrymple (2009) proposed a new metric approach called “phenotypic impact factor” (PIF). PIF is based on DEG numerical properties, which “weight” the contribution of the DEG on the molecular anatomy differences between two different phenotypes. These authors also proposed the “regulatory impact factor” (RIF) to account the isolated changes in regulators expression level.

In this study, the objective was to applied the NGS technology, pathway algorithms and mathematical methods as PCIT, PIF and RIF to identify DEG, potentially impacted pathways and regulators of variation in IMF, respectively, in the *Longissimus dorsi* (LD) muscle of extreme IMF GEBV Nellore steers.

3.2 Materials and Methods

3.2.1 Ethics statement

All experimental procedures involving steers were approved by the Institutional Animal Care and Use Committee Guidelines from Brazilian Agricultural Research Corporation – EMBRAPA and sanctioned by the president Dr. Rui Machado.

3.2.3 Animals, samples and phenotypes

Three hundred and ten Nellore steers from the Brazilian Agricultural Research Corporation (EMBRAPA/Brazil) experimental breeding herd between 2009 and 2011 were included in this study such as described by Cesar et al. (2014). Samples from LD muscle located between the 12th and 13th ribs was collected in two moments: at slaughter for RNA sequencing analysis, and 24 hour after slaughter for the intramuscular fat (IMF) content measurement.

Approximately 100 g samples of beef collected for IMF content analysis were lyophilized and ground to a fine powder. Five g of this ground, lyophilized tissue was used to obtain IMF. The Ankom XT20 lipids equipment was used to determine lipid content

according to the procedure of AOCS (Official Procedure Am 5-04) for IMF extraction (CESAR et al., 2014). Genomic Best Linear Unbiased Prediction (GBLUP) methodology was conducted using ASREML software (GILMOUR et al., 2009) on 310 total animals. The animal model used in this analysis was, $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$, where $\mathbf{y}$ is the vector of observations, which represented the trait of interest (dependent variable), $\mathbf{X}$ and $\mathbf{Z}$ are the design or incidence matrices for the vectors of fixed and random effects in $\mathbf{b}$ and $\mathbf{u}$, respectively, and $\mathbf{e}$ was the vector of random residuals. The expected variance of vector $\mathbf{u}$ is $\text{Var}(\mathbf{u}) = \mathbf{I} \sigma_m^2$; where σ_m^2 is the variance explained by markers, and $\mathbf{I}$ is the identity matrix. The variance of vector $\mathbf{u}$ was $\mathbf{G}\sigma_m^2$ for the genomic analyses where $\mathbf{G}$ is the genomic relationship matrix derived from SNP markers using allele frequencies as suggested by VanRaden (2008), with σ_m^2 being the marker-based additive genetic variance. A group of 14 animals were selected based on their extreme genomic estimated breeding values (GEBVs) for IMF (seven high and low).

3.2.4 RNA extraction, quality analysis, library preparation and sequencing

Total RNA was extracted from 100 mg of frozen LD muscle that was collected at slaughter using the TRIzol reagent (Life Technologies, Carlsbad, CA). RNA integrity was verified by Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). Only samples with RIN > 8 were used. A total of 2µg of total RNA from each sample was used for library preparation according to the protocol described in the TruSeq RNA Sample Preparation kit v2 guide (Illumina, San Diego, CA). Libraries average size was estimated using the Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA) and quantified using quantitative PCR with the KAPA Library Quantification kit (KAPA Biosystems, Foster City, CA, USA). Quantified, samples were diluted and pooled (three pools of six samples each). Three lanes of a sequencing flowcell, using the TruSeq PE Cluster kit v3-cBot-HS kit (Illumina, San Diego, CA, USA), were clusterized and sequenced using HiScanSQ equipment (Illumina, San Diego, CA, USA) with a TruSeq SBS Kit v3-HS (200 cycles), according to manufacturer instructions. The sequencing analyses were performed at the Genomics Center at ESALQ, Piracicaba, São Paulo, Brazil.

3.2.5 Quality control and read alignment

Sequencing adaptors and low-complexity reads were removed in an initial data-filtering step. Quality control and reads statistics were estimated with FASTQC version 0.10.1 software [<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>]. Tophat v. 1.2.0 software (TRAPNELL; PACHTER; SALZBERG, 2009) was used to map reads to the UMD3.1 *Bos taurus* reference assembly available at Ensembl [http://www.ensembl.org/Bos_taurus/Info/Index/]. A reference-guided assembly was performed using Cufflinks version 2.0.2, with a minimum alignment count per locus of 10 (LANGMEAD et al., 2009), to identify novel transcripts. A combination of novel transcripts identified by Cufflinks and those from the reference GTF file at Ensembl were used as the reference for read quantification for each transcript. The abundance (read counts) of mRNAs for all annotated genes, was calculated using HTSeq version 0.5.4 software [<http://www-huber.embl.de/users/anders/HTSeq/>] (ANDERS; HUBER, 2010). Only sequence reads that uniquely mapped to known chromosomes (excluding reads mapped to unassigned contigs) were used in this analysis.

3.2.6 Identification of differential expressed genes and pathway analysis

Differentially expressed genes were identified using the DESeq software (ANDERS; HUBER, 2010). Prior to statistical analysis, the read count data was filtered as follows: i) transcripts with zero counts were removed (unexpressed); ii) transcripts with less than 1 read per sample on average were removed (very lowly expressed); iii) transcripts that were not present in at least three samples were removed (rarely expressed). After filtering, a total of 16,101 transcripts were analyzed for differential expression using the “nbinomTest” function of DESeq to fit transcript expression level as a negative binomial distribution. Exploratory diagnostic plots were generated to check the dispersion estimates. Benjamini-Hochberg (BENJAMINI; HOCHBERG, 1995) methodology was used to control the false discovery rate (FDR) at 10%. Transcript annotations were retrieved with a perl script to query the Ensembl database using the Ensembl Perl Application Program Interface (API). Transcripts that lacked annotation information were annotated using the Genome-to-seq tool at AgBase (MCCARTHY et al., 2011). A two-tiered approach was taken to detect pathway level changes in gene ontology in the extreme IMF samples RNAseq profiles. First, enrichment analysis of curated gene ontology terms was completed with the Database for Annotation, Visualization

and Integrated Discovery (DAVID) v6.7 tool (HUANG; SHERMAN; LEMPICKI, 2009). Second, a literature based pathway enrichment analysis was performed using Pathway Studio (NIKITIN et al., 2003).

3.2.7 PCIT and differential hubbing network analysis

Expression values were normalized as number of fragments per kilobase of exon per million reads (FPKM) as reported in Cufflinks output (TRAPNELL et al., 2010) for co-expression analysis. The PCIT algorithm (REVERTER; CHAN, 2008; KOESTERKE et al., 2013) was used to identify differential hubbing (DH) for all transcripts (HUDSON; REVERTER; DALRYMPLE, 2009). A shell script pipeline was developed to summarize all DH results. Only transcripts with a direct and partial correlation ≥ 0.90 were used for the DH analysis. The DH was computed by the difference of significant connections of a transcript between the low and high IMF group. The top ten most positively and negatively DH transcripts were identified for further investigation.

3.2.8 Regulatory Impact Factor (RIF) network co-expression analysis and phenotypic impact factor (PIF) scores

The regulatory impact factor (RIF) and phenotypic impact factor (PIF) scores were calculated as described (REVERTER et al., 2010) to predict which transcripts were potential regulators of gene expression differences between the high and low IMF groups. The RIF calculations presented here were modified from the original method since all transcripts were tested as potential regulators and only transcripts with a significant partial correlation ≥ 0.90 from PCIT were included in the RIF and PIF score estimates.

3.3 Results

3.3.1 Phenotypic groups, mapping and annotation

A summary of IMF expressed as a percentage, GEBVs, and the total number of reads mapped against the *Bos taurus* UMD3.1 reference genome assembly following quality control are shown in Table 1. The GEBV for IMF values for all 310 animals were ranked and seven

animals with high GEBV for IMF (H) and seven with low (L) were selected for RNA-Seq analysis.

A total of 364.18 million (M), 100 bp paired-end reads were obtained from three lanes of Illumina HiScanSQ equipment. The average number of total reads per sample was 26 M, which should result in an estimated sequencing depth (expected mean coverage at all transcripts) of 12X coverage. The mean number of total mapped reads estimated by Tophat (TRAPNELL; PACHTER; SALZBERG, 2009) was 17 M, for an average of 65.38% paired reads mapped. A total of 16,101 expressed genes were annotated using the Ensembl Perl API, and the correlation of mean gene expression levels between both groups was high ($r = 0.99$).

Table 1 - Intramuscular fat percentage (IMF), genomic estimated breeding values (GEBV) and mapped reads for all animals within group (Low and High) based on IMF GEBV

Animal	IMF (%)	GEBV	Mapped reads (M) ¹
Low ₁ ²	1.70	-0.31	23.09
Low ₂	1.94	-0.29	11.13
Low ₃	1.86	-0.24	25.11
Low ₄	1.60	-0.59	13.32
Low ₅	1.32	-0.77	14.34
Low ₆	1.58	-0.50	14.58
Low ₇	1.62	-0.57	17.85
High ₁ ³	4.42	0.44	22.65
High ₂	4.35	0.57	17.53
High ₃	4.38	0.71	17.58
High ₄	5.27	0.85	16.31
High ₅	5.02	0.47	15.73
High ₆	4.74	0.81	15.13
High ₇	4.35	0.61	13.76
Mean Low	1.56	-0.45	17.04
Mean High	4.65	0.65	19.39

¹M million reads

3.3.2 Differential expression analysis and differentially expressed genes

The DESeq R package was used to identify DEG. This statistical package uses a parametric method, which relies on assumptions regarding the distribution of sampled data. DESeq (ANDERS; HUBER, 2010). The inference relies on estimation of the typical

relationship between the mean of gene expression levels and their data's dispersion (square of the coefficient of biological variation). Figure 1 presents the dispersion plot of the 16,101 common expressed genes identified in both groups (H and L). The black dots represent the empirical value for each gene and the red line represents a curve through the estimated dispersion values. Using this parametrical approach and extending Fisher's exact test to the data following negative binomial distribution, 77 DEG (false discovery rate (FDR) of 10%) were detected between the H and L IMF groups (Table 2 and Figure 2). The histogram of p values demonstrates the presence of DEG (Figure 3). Of the 77 DEG, 41 genes were up-regulated and 36 down-regulated in the L group compared to H group (Table 2).

Figure 1 - Empirical (black dots) and fitted (red lines) dispersion values ($\log_2$) plotted against the mean of the normalized counts from RNA-Seq data of *Longissimus dorsi* muscle of Nellore steers

Table 2 - The differentially expressed genes (DGE) obtained between High and Low groups based on genomic estimated breeding values (GEBV) for intramuscular fat (IMF) percentage in Nellore steers with FDR < 0.10

(continues)

Ensembl_Gene_ID	Associated Gene Name	baseMean ¹	baseMeanA ²	baseMeanB ³	foldChange ⁴	pval	padj
ENSBTAG000000021508	<i>LMOD3</i>	7,928.17	10,274.52	5,581.83	0.54	5.46x10 ⁻¹¹	8.78x10 ⁻⁷
ENSBTAG000000025210	<i>COL4A2</i>	9,959.52	7,174.80	12,744.25	1.78	7.36x10 ⁻¹⁰	5.92x10 ⁻⁶
ENSBTAG000000012849	<i>COL4A1</i>	18,562.12	13,722.93	23,401.32	1.71	7.64x10 ⁻⁹	4.10x10 ⁻⁵
ENSBTAG000000011465	<i>MYBPH</i>	98.28	44.10	152.46	3.46	5.97x10 ⁻⁸	2.40x10 ⁻⁴
ENSBTAG000000005102	<i>PHTF2</i>	1,177.63	1,464.31	890.95	0.61	1.10x10 ⁻⁷	3.53x10 ⁻⁴
ENSBTAG000000015303	<i>MPP6</i>	339.00	426.21	251.79	0.59	5.18x10 ⁻⁷	1.39x10 ⁻³
ENSBTAG000000010389	<i>STBD1</i>	1,501.74	1,940.19	1,063.28	0.55	9.24x10 ⁻⁷	2.12x10 ⁻³
ENSBTAG000000014835	<i>SPARC</i>	20,870.41	16,334.19	25,406.64	1.56	1.32x10 ⁻⁶	2.66x10 ⁻³
ENSBTAG000000014885	<i>MYOM3</i>	9,806.49	7,226.13	12,386.86	1.71	3.02x10 ⁻⁶	5.40x10 ⁻³
ENSBTAG000000001183	<i>KLHL33</i>	2,297.48	2,775.51	1,819.45	0.66	3.77x10 ⁻⁶	6.01x10 ⁻³
ENSBTAG000000009145	<i>SLC7A4</i>	110.63	152.73	68.52	0.45	4.11x10 ⁻⁶	6.01x10 ⁻³
ENSBTAG000000024555	<i>EFCAB2</i>	112.00	144.62	79.37	0.55	5.19x10 ⁻⁶	6.97x10 ⁻³
ENSBTAG000000018382	<i>FIME02</i>	564.14	449.52	678.75	1.51	1.74x10 ⁻⁵	2.15x10 ⁻²
ENSBTAG000000000828	<i>CAPN6</i>	893.42	723.26	1,063.58	1.47	2.47x10 ⁻⁵	2.72x10 ⁻²
ENSBTAG000000013419	<i>HTATIP2</i>	282.52	353.88	211.17	0.60	2.81x10 ⁻⁵	2.72x10 ⁻²
ENSBTAG000000016709	<i>NT5C3</i>	407.17	491.36	322.98	0.66	2.67x10 ⁻⁵	2.72x10 ⁻²
ENSBTAG000000004347	<i>GPR116</i>	3,229.34	2,623.20	3,835.49	1.46	2.87x10 ⁻⁵	2.72x10 ⁻²
ENSBTAG000000004036	<i>GJC1</i>	106.13	70.93	141.32	1.99	3.46x10 ⁻⁵	2.87x10 ⁻²
ENSBTAG000000004126	<i>MLF1</i>	2,367.33	2,924.83	1,809.84	0.62	3.87x10 ⁻⁵	2.87x10 ⁻²
ENSBTAG000000022989	<i>FAM174B</i>	75.66	98.74	52.58	0.53	4.07x10 ⁻⁵	2.87x10 ⁻²
ENSBTAG000000006819	<i>CTAGE5</i>	1,023.75	1,223.21	824.29	0.67	4.10x10 ⁻⁵	2.87x10 ⁻²
ENSBTAG000000018775	<i>TPX2</i>	440.69	547.59	333.78	0.61	3.29x10 ⁻⁵	2.87x10 ⁻²
ENSBTAG000000038141	<i>FIMKM4</i>	216.29	161.28	271.29	1.68	3.59x10 ⁻⁵	2.87x10 ⁻²
ENSBTAG000000019658	<i>ASB16</i>	2,684.07	3,179.60	2,188.53	0.69	4.57x10 ⁻⁵	2.97x10 ⁻²
ENSBTAG000000001514	<i>ASB11</i>	946.09	1,127.32	764.86	0.68	4.95x10 ⁻⁵	2.97x10 ⁻²
ENSBTAG000000013750	<i>B3GALNT2</i>	528.29	426.39	630.19	1.48	4.98x10 ⁻⁵	2.97x10 ⁻²
ENSBTAG000000035844	<i>HRASLS</i>	1,279.90	1,578.56	981.24	0.62	4.77x10 ⁻⁵	2.97x10 ⁻²
ENSBTAG000000014863	<i>GYPC</i>	2,021.53	2,405.81	1,637.26	0.68	5.28x10 ⁻⁵	3.30x10 ⁻²
ENSBTAG000000007211	<i>ASB12</i>	3,772.06	4,750.26	2,793.86	0.59	5.72x10 ⁻⁵	3.18x10 ⁻²
ENSBTAG000000016194	<i>FBXO32</i>	3,310.54	4,176.72	2,444.36	0.59	6.10x10 ⁻⁵	3.27x10 ⁻²
ENSBTAG000000011190	<i>FLNA</i>	6,136.61	4,791.10	7,482.11	1.56	7.46x10 ⁻⁵	3.87x10 ⁻²
ENSBTAG000000019052	<i>ANK3</i>	473.36	561.03	385.70	0.69	7.84x10 ⁻⁵	3.94x10 ⁻²
ENSBTAG000000002391	<i>TGFB11I</i>	326.41	254.27	398.56	1.57	1.00x10 ⁻⁴	4.25x10 ⁻²
ENSBTAG000000002914	<i>GALNTL4</i>	132.92	97.19	168.66	1.74	9.83x10 ⁻⁵	4.25x10 ⁻²
ENSBTAG000000011869	<i>CSRP3</i>	29,763.36	18,942.70	40,584.02	2.14	9.95x10 ⁻⁵	4.25x10 ⁻²
ENSBTAG000000009496	<i>STAT5A</i>	393.40	469.54	317.25	0.68	9.73x10 ⁻⁵	4.25x10 ⁻²
ENSBTAG000000019517	<i>ELN</i>	2,071.46	1,645.83	2,497.09	1.52	9.48x10 ⁻⁵	4.25x10 ⁻²
ENSBTAG000000012505	<i>ARHGEF17</i>	420.66	333.77	507.55	1.52	9.02x10 ⁻⁵	4.25x10 ⁻²
ENSBTAG000000016525	<i>FIMMN6</i>	767.64	615.74	919.54	1.49	1.12x10 ⁻⁴	4.64x10 ⁻²
ENSBTAG000000002126	<i>PFKFB2</i>	388.79	455.85	321.73	0.71	1.21x10 ⁻⁴	4.89x10 ⁻²

Table 2 - The differentially expressed genes (DGE) obtained between High and Low groups based on genomic estimated breeding values (GEBV) for intramuscular fat (IMF) percentage in Nellore steers with FDR < 0.10

(conclusion)

Ensembl Gene ID	Gene Symbol	baseMean ¹	baseMeanA ²	baseMeanB ³	foldChange ⁴	pval	padj
ENSBTAG00000018747	<i>PRKAA2</i>	711.69	839.81	583.57	0.69	1.29x10 ⁻⁴	4.98x10 ⁻²
ENSBTAG00000020764	<i>CNN2</i>	156.65	118.37	194.92	1.65	1.30x10 ⁻⁴	4.98x10 ⁻²
ENSBTAG00000012558	<i>ADAMTS12</i>	173.39	111.41	235.37	2.11	1.55x10 ⁻⁴	5.57x10 ⁻²
ENSBTAG00000013632	<i>GRM4</i>	36.44	50.18	22.71	0.45	1.55x10 ⁻⁴	5.57x10 ⁻²
ENSBTAG00000031134	<i>FIMPT3</i>	950.01	771.75	1,128.27	1.46	1.56x10 ⁻⁴	5.57x10 ⁻²
ENSBTAG00000004305	<i>RGS16</i>	108.28	76.15	140.40	1.84	1.77x10 ⁻⁴	6.21x10 ⁻²
ENSBTAG00000011913	<i>CKAP4</i>	178.87	140.93	216.80	1.54	1.88x10 ⁻⁴	6.40x10 ⁻²
ENSBTAG00000007632	<i>FAM151B</i>	168.02	205.93	130.11	0.63	1.91x10 ⁻⁴	6.40x10 ⁻²
ENSBTAG00000014930	<i>MYLK2</i>	18,138.96	22,394.85	13,883.08	0.62	2.16x10 ⁻⁴	7.09x10 ⁻²
ENSBTAG00000003836	<i>ADAM19</i>	1,291.98	1,086.20	1,497.77	1.38	2.35x10 ⁻⁴	7.11x10 ⁻²
ENSBTAG00000004570	<i>PRRX1</i>	272.28	201.58	342.98	1.70	2.22x10 ⁻⁴	7.11x10 ⁻²
ENSBTAG00000017233	<i>RNF213</i>	125.78	79.82	171.75	2.15	2.32x10 ⁻⁴	7.11x10 ⁻²
ENSBTAG00000016918	<i>MYOF</i>	595.62	485.32	705.92	1.45	2.38x10 ⁻⁴	7.11x10 ⁻²
ENSBTAG00000005932	<i>FAM184B</i>	31.60	42.34	20.86	0.49	2.34x10 ⁻⁴	7.11x10 ⁻²
ENSBTAG00000012450	<i>RAPGEF2</i>	955.88	784.17	1,127.59	1.44	2.45x10 ⁻⁴	7.18x10 ⁻²
ENSBTAG00000016013	<i>ASB14</i>	858.94	1,017.39	700.48	0.69	2.57x10 ⁻⁴	7.38x10 ⁻²
ENSBTAG00000021768	<i>E1BIN6</i>	336.38	420.58	252.19	0.60	2.64x10 ⁻⁴	7.45x10 ⁻²
ENSBTAG00000004259	<i>HPCALI</i>	256.20	202.46	309.94	1.53	2.70x10 ⁻⁴	7.49x10 ⁻²
ENSBTAG00000006253	<i>FLNC</i>	113,919.49	78,869.35	148,969.62	1.89	2.88x10 ⁻⁴	7.66x10 ⁻²
ENSBTAG00000016208	<i>TGM2</i>	1,430.90	1,083.10	1,778.69	1.64	2.90x10 ⁻⁴	7.66x10 ⁻²
ENSBTAG00000046162	<i>ZNF385C</i>	58.42	36.78	80.06	2.18	2.90x10 ⁻⁴	7.66x10 ⁻²
ENSBTAG00000025099	<i>RNF145</i>	470.56	388.72	552.40	1.42	3.04x10 ⁻⁴	7.90x10 ⁻²
ENSBTAG00000010940	<i>HSPB7</i>	6,297.66	4,989.43	7,605.90	1.52	3.18x10 ⁻⁴	8.13x10 ⁻²
ENSBTAG00000011010	<i>PRND</i>	38.92	15.31	62.54	4.08	3.32x10 ⁻⁴	8.21x10 ⁻²
ENSBTAG00000012720	<i>ANKRD2</i>	4,559.69	2,834.86	6,284.53	2.22	3.31x10 ⁻⁴	8.21x10 ⁻²
ENSBTAG00000005729	<i>FBXL4</i>	721.12	869.82	572.42	0.66	3.61x10 ⁻⁴	8.40x10 ⁻²
ENSBTAG00000010179	<i>COL5A3</i>	2,843.18	2,391.27	3,295.09	1.38	3.60x10 ⁻⁴	8.40x10 ⁻²
ENSBTAG00000011324	<i>EMILIN1</i>	736.34	606.71	865.98	1.43	3.55x10 ⁻⁴	8.40x10 ⁻²
ENSBTAG00000019090	<i>PLEKHH3</i>	584.98	680.38	489.57	0.72	3.66x10 ⁻⁴	8.40x10 ⁻²
ENSBTAG00000027051	<i>PTAFR</i>	56.15	39.30	73.01	1.86	3.69x10 ⁻⁴	8.40x10 ⁻²
ENSBTAG00000048210	<i>G5E6M6</i>	485.58	605.60	365.57	0.60	3.70x10 ⁻⁴	8.40x10 ⁻²
ENSBTAG00000006679	<i>MITF</i>	632.94	750.45	515.42	0.69	3.98x10 ⁻⁴	8.79x10 ⁻²
ENSBTAG00000013922	<i>MOSPD1</i>	332.11	451.56	212.66	0.47	3.96x10 ⁻⁴	8.79x10 ⁻²
ENSBTAG00000013131	<i>FAM110D</i>	59.37	40.81	77.93	1.91	4.10x10 ⁻⁴	8.92x10 ⁻²
ENSBTAG00000019011	<i>PGM1</i>	587.89	781.26	394.53	0.50	4.19x10 ⁻⁴	9.00x10 ⁻²
ENSBTAG00000020097	<i>PERP</i>	15.84	7.90	23.78	3.01	4.40x10 ⁻⁴	9.33x10 ⁻²
ENSBTAG00000022920	<i>RNF128</i>	725.41	852.37	598.45	0.70	4.47x10 ⁻⁴	9.35x10 ⁻²

¹ baseMean - mean normalised counts, averaged over all samples from both groups (H and L); ² baseMeanA mean normalised counts from high group; ³ baseMeanB - mean normalised counts from low group; ⁴ foldChange - fold change from low to high group; ⁵ pval - p value for the statistical significance of this change; ⁶ padj - p value adjusted for multiple testing with the Benjamini-Hochberg procedure (FDR)

Figure 2 - Plot of normalized mean versus log₂ fold change to contrast Low and High groups based on genomic estimated breeding values (GEBV) for intramuscular fat (IMF) percentage. Red points represent significant (10% FDR) genes

Figure 3 - Histogram of p-values from RNA-Seq data of *Longissimus dorsi* muscle of Nellore steers by DESeq program

3.3.3 Functional enrichment and pathways

A singular enrichment analysis (SEA) was performed by DAVID database to capture the major biological meaning behind the identified DEG. The SEA analysis showed (Table 3) two significant pathways: focal adhesion (p-adj < 0.02) and extracellular matrix (ECM) receptor interaction (p-adj < 0.056). A gene set enrichment analysis (GSEA) was performed for functional enrichment by Pathway Studio using a new list of DEG (FDR < 0.2). The GSEA detected 13 pathways (FDR 10%) that contained 27 DEG (Table 4). These DEG are involved with more than one pathway, which indicates that the pathways are inter-related. The biological pathways identified by Pathway Studio were: L-cysteine (Figure 4), retinoic acid (Figure 5), inflammatory cytokines (Figure 6), IL-1 β (Figure 7), protein tyrosin-kinase (Figure 8), NF-kB (Figure 9), Zn²⁺ (Figure 10), H₂O₂ (Figure 11), oxidized LDL complex (Figure 12), NO (Figure 13), ROS (Figure 14), Mg²⁺ (Figure 15), ATP.

Table 3 - Singular enrichment analysis (SEA) by DAVID database. Significant category (p < 0.10) are shown from DEG (FDR < 0.10) comparing high and low IMF GEBV animals

Category	Term	Count	%	P-Value	Benjamini
KEGG_PATHWAY	Focal adhesion	6	8.0	4.9E-4	2.0E-2
GOTERM_BP_FAT	cellular protein catabolic process	7	9.3	2.3E-4	4.6E-2
GOTERM_BP_FAT	protein catabolic process	7	9.3	3.5E-4	4.7E-2
GOTERM_BP_FAT	cellular macromolecule catabolic process	7	9.3	5.4E-4	5.3E-2
KEGG_PATHWAY	ECM-receptor interaction	4	5.3	2.7E-3	5.6E-2
SP_PIR_KEYWORDS	basement membrane	3	4.0	6.9E-4	6.0E-2
GOTERM_BP_FAT	macromolecule catabolic process	7	9.3	1.0E-3	7.9E-2
GOTERM_BP_FAT	modification-dependent protein catabolic process	6	8.0	1.2E-3	8.0E-2
GOTERM_BP_FAT	modification-dependent macromolecule catabolic process	6	8.0	1.2E-3	8.0E-2
GOTERM_BP_FAT	proteolysis involved in cellular protein catabolic process	7	9.3	2.3E-4	8.8E-2
GOTERM_BP_FAT	muscle tissue development	4	5.3	1.9E-3	9.1E-2
GOTERM_BP_FAT	striated muscle tissue development	4	5.3	1.6E-3	9.1E-2

Table 4 - Pathway enrichment using Gene Set Enrichment Analysis (GSEA) methodology in Pathway Studio software. Significant pathways ($p < 0.10$) are shown from DEG ($FDR < 0.20$) comparing high and low IMF GEBV animals

Name	Type	Total number of Neighbors	Number of Measured Neighbors	Gene Set Seed	Neighbors	padj ¹
Neighbors of L-cysteine	Pathway	366	5	L-cysteine	<i>MT3, GRM4, PGM1, TGM2, ELN</i>	0.011
Neighbors of Zn2+	Pathway	905	6	Zn2+	<i>MT3, ADAMTS12, PGM1, TGM2, SLC2A4, STAT5A</i>	0.021
Neighbors of oxidized LDL	Pathway	424	5	oxidized LDL	<i>PTAFR, COL4A2, TGM2, SERPINE2, SLC2A4</i>	0.023
Neighbors of inflammatory cytokine	Pathway	774	9	inflammatory cytokine	<i>MT3, ADAMTS12, FBXO32, CRYAB, TGM2, SERPINE2, SLC2A4, ELN, STAT5A</i>	0.025
Neighbors of H2O2	Pathway	1142	8	H2O2	<i>MT3, FBXO32, CRYAB, HTATIP2, TGM2, SLC2A4, TGFB111, ELN</i>	0.030
Neighbors of Mg2+	Pathway	562	5	Mg2+	<i>ADAMTS12, PGM1, TGM2, REM1, ELN</i>	0.033
Neighbors of ATP	Pathway	1167	9	ATP	<i>GRM4, PGM1, FBXO32, CRYAB, SLC4A4, TGM2, SLC2A4, GYPC, NRP1</i>	0.061
Neighbors of IL1B	Pathway	1191	12	IL1B	<i>MT3, CSRP3, RGS16, FBXO32, CRYAB, TGM2, SERPINE2, SLC2A4, SPARC, ELN, STAT5A, NRP1</i>	0.063
Neighbors of retinoic acid	Pathway	1605	12	retinoic acid	<i>SLC7A4, CSRP3, PTAFR, RGS16, CPM, TGM2, SLC2A4, FLNA, SPARC, ELN, STAT5A, ROCK2</i>	0.074
Neighbors of NF-kB	Pathway	1125	9	NF-kB	<i>MT3, PTAFR, RGS16, FBXO32, TGM2, SLC2A4, ELN, STAT5A, FSCN1</i>	0.085
Neighbors of protein tyrosine kinase	Pathway	823	7	protein tyrosine kinase	<i>PGM1, PTAFR, RGS16, SLC4A4, SLC2A4, STAT5A, BCR</i>	0.092
Neighbors of NO	Pathway	862	8	NO	<i>CSRP3, PTAFR, FBXO32, TGM2, SLC2A4, SPARC, ELN, STAT5A</i>	0.097
Neighbors of ROS	Pathway	787	8	ROS	<i>MT3, PTAFR, SLC4A4, TGM2, SLC2A4, TGFB111, ELN, STAT5A</i>	0.098

¹padj - p value adjusted for multiple testing with the Benjamini-Hochberg procedure, which controls false discovery rate (FDR)

Figure 4 - L-cysteine pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 5 - Retinoic acid pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 6 - Inflammatory cytokine pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 7 – Interleukin-1 beta (IL-1 β) pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 8 – Protein tyrosine kinase pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 9 – Necrosis factor kappaB pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 10 – Zinc pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($\text{FDR} \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 11 – Hydrogen peroxide pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 12 – Oxidized LDL pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 13 – Nitric oxide pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 14 – Reactive oxygen species pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 15 – Magnesium pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

Figure 16 – ATP pathway and genes identified as differentially expressed between the high and low groups for IMF GEBV are shown here ($FDR \leq 0.10$, adjusted for multiple testing using Benjamini-Hochberg method). The genes shown in red had higher expression in the low IMF group and those in blue had higher expression in animals from the High IMF group

3.3.4 PCIT and differential hubbing analysis

The PCIT algorithm was used to identify differential hubbing (DH) or the significant difference in the number of significant connections for each regulatory gene in high and low group. The top ten negative and positive DH values are showed in the Table 5 and Table 6, respectively. The PCIT algorithm allowed for the testing of all transcripts as putative regulators genes without prior knowledge (KOESTERKE et al., 2014). Among the top ten negative DH, three of them have been previously reported as associated with adipogenesis and adipose metabolism: *ANKRD26*, *HOXC5* and *PPAPDC2* (Figure 17). Among the top ten positive DH, two of them are good putative regulator genes: Zinc finger protein, friend of GATA (FOG) family member 1 (*ZFPM1*) and zinc finger protein 90 (*ZFP90*) (Figure 18).

Table 5 - Top 10 negative differentially hubbing genes comparing High and Low groups based on genomic estimated breeding values (GEBV) for intramuscular fat (IMF) percentage

Ensembl Gene ID	Ensembl Transcript ID	Gene name	DH	Transcript Biotype
ENSBTAG00000006639	ENSBTAT00000008722	<i>CSTF1</i>	-216	protein_coding
ENSBTAG00000022922	ENSBTAT00000055503	<i>ANKRD26</i>	-216	protein_coding
ENSBTAG00000018343	ENSBTAT00000024407	<i>NCR3</i>	-210	protein_coding
ENSBTAG00000006990	ENSBTAT00000009187	<i>MYRF</i>	-207	protein_coding
ENSBTAG00000025809	ENSBTAT00000036501	<i>ABHD8</i>	-204	protein_coding
ENSBTAG00000015208	ENSBTAT00000061091	<i>LRRC16B</i>	-204	protein_coding
ENSBTAG00000009769	ENSBTAT00000012885	-	-200	processed_pseudogene
ENSBTAG00000009634	ENSBTAT00000012696	<i>HOXC5</i>	-199	protein_coding
ENSBTAG00000011050	ENSBTAT00000014674	<i>PPAPDC2</i>	-198	protein_coding
ENSBTAG00000022922	ENSBTAT00000031151	<i>ANKRD26</i>	-198	protein_coding

Table 6 - Top 10 positive differentially hubbing genes comparing High and Low groups based on genomic estimated breeding values (GEBV) for intramuscular fat (IMF) percentage

Ensembl Gene ID	Ensembl Transcript ID	Gene name	DH	Transcript Biotype
ENSBTAG00000007531	ENSBTAT00000009911	<i>NCF4</i>	302	protein_coding
ENSBTAG00000015766	ENSBTAT00000047968	<i>ZFPM1</i>	297	protein_coding
ENSBTAG00000006638	ENSBTAT00000008723	<i>BCL2L12</i>	296	protein_coding
ENSBTAG00000011590	ENSBTAT00000015396	<i>GLIPR1</i>	293	protein_coding
ENSBTAG00000025398	ENSBTAT00000035669	-	292	Novel protein coding
ENSBTAG00000007488	ENSBTAT00000032517	<i>ZFP90</i>	285	protein_coding
ENSBTAG00000004206	ENSBTAT00000005510	<i>LRRC55</i>	285	protein_coding
ENSBTAG00000014491	ENSBTAT00000019267	<i>ATP6V1G2</i>	282	protein_coding
unannotated	CUFF.25603.1	-	278	-
ENSBTAG00000002823	ENSBTAT00000003653	<i>MPZL1</i>	277	protein_coding

Figure 17 - Positive differential hubbing (DH) between the high and low groups for IMF GEBV. The center spot represents the gene with high value of DH, the red edges represent the positive DH and blue edges represent the negative DH. Other spots represent the significant connections

Figure 18 - Negative differential hubbing (DH) between the high and low groups for IMF GEBV. The center spot represents the gene with high value of DH, the red edges represent the positive DH and blue edges represent the negative DH. Other spots represent the significant connections

3.3.5 Regulatory impact factor and phenotypic impact factor

A list of the top 10 RIF1 and RIF2 genes are presented in tables 7 and 8, respectively. A list of DH, RIF1, RIF2 and PIF results of all DEG are presented in Table 9. The RIF1 analysis identified several known genes and many novel or pseudo genes as potential regulators that lead to differences between the low and high IMF groups. Two particularly interesting regulators included *GLI2* and *IGF2*. The RIF2 score identified additional candidate regulators, including two strong candidates: *MPC1* and *UBL5*. The PIF analysis identified many small RNAs as having a large impact on the IMF phenotype. In fact, 9 of the top 10 most highly ranked PIF genes are small RNAs. To further investigate the functional associations of the 3,000 genes with higher scores of PIF, the functional annotation clustering

at DAVID database (<http://www.david.abcc.ncifcrf.gov>) was performed (Table 10). Several significant clustering were enriched ($p\text{-adj} < 0.10$), including GO terms of biological processes related to translation, generation of precursor metabolites and energy, ATP synthesis coupled proton transport, tricarboxylic acid cycle, acetyl-CoA catabolic process, acetyl-CoA metabolic process, posttranscriptional regulation of gene expression, protein catabolic process, mRNA metabolic process, RNA processing, RNA splicing, intracellular protein transport, intracellular transport and protein folding.

Table 7 - Top 10 RIF1 (High minus Low Intramuscular Fat)*

Gene ID	Transcript ID	Gene name	RIF1	Transcript Biotype
Top Positive RIF1				
ENSBTAG00000047048	ENSBTAT00000063303	Novel gene	9.27	protein coding
ENSBTAG00000010148	ENSBTAT00000013392	Novel gene	8.88	processed_pseudogene
ENSBTAG00000046926	ENSBTAT00000063044	GPX4	7.92	protein_coding
ENSBTAG00000047278	CUFF.29196.1	NDUFB11	7.71	protein_coding
ENSBTAG00000046850	ENSBTAT00000065286	Novel gene	7.64	pseudogene
Top Negative RIF1				
ENSBTAG00000011682	ENSBTAT00000015510	GLI2	-1.34	protein_coding
ENSBTAG00000001574	ENSBTAT00000002064	GPATCH2	-1.31	protein_coding
ENSBTAG00000019283	ENSBTAT00000025675	FAM186B	-1.31	protein_coding
ENSBTAG00000013066	ENSBTAT00000044139	IGF2	-1.31	protein_coding
ENSBTAG00000037440	ENSBTAT00000055187	Novel gene	-1.30	protein coding

*Note: RIF1 scores are presented as Z score normalized values

Table 8 - Top 10 RIF2 (High minus Low Intramuscular Fat)*

Gene ID	Transcript ID	Gene name	RIF2	Transcript Biotype
Top Positive RIF1				
ENSBTAG00000016997	ENSBTAT00000022603	Novel gene	7.60	protein coding
ENSBTAG00000015371	ENSBTAT00000020429	GABPB2	7.19	protein coding
ENSBTAG00000047712	ENSBTAT00000065359	Novel gene	6.75	-
ENSBTAG00000027879	ENSBTAT00000040201	MPC1	6.71	protein coding
ENSBTAG00000006398	ENSBTAT00000008388	TOMM7	6.70	protein coding
Top Negative RIF2				
ENSBTAG00000030592	ENSBTAT00000031821	UBL5	-10.27	protein coding
ENSBTAG00000012803	ENSBTAT00000017007	SRP19	-10.19	protein coding
ENSBTAG00000007199	ENSBTAT00000009472	Novel gene	-8.34	pseudogene
ENSBTAG00000047194	ENSBTAT00000063248	Novel gene	-8.30	protein coding
ENSBTAG00000019949	ENSBTAT00000026574	DBR1	-8.13	protein coding

*Note: RIF2 scores are presented as Z score normalized values

Table 9 – Differential hubbing (DH), regulatory impact factor 1 and 2 (RIF1 and RIF2), and phenotypic impact factor scores between the high and low IMF groups in Nellore steers

(continues)

Ensembl Gene ID	Gene Symbol	DH	RIF1	RIF2	PIF
ENSBTAG000000021508	<i>LMOD3</i>	-23	-0.059148705	0.268681984	10.90401594
ENSBTAG000000025210	<i>COL4A2</i>	48	0.253412595	1.210786908	11.11574964
ENSBTAG000000012849	<i>COL4A1</i>	145	0.667043993	1.75897281	13.47866165
ENSBTAG000000011465	<i>MYBPH</i>	124	-0.807657165	-0.174317789	1.24760694
ENSBTAG000000005102	<i>PHTF2</i>	-20	0.356040218	-0.175022987	3.338611618
ENSBTAG000000015303	<i>MPP6</i>	13	-0.661817232	0.243248375	1.726948145
ENSBTAG000000010389	<i>STBD1</i>	-147	-0.144997048	-0.365371389	6.800556056
ENSBTAG000000014835	<i>SPARC</i>	110	0.334214925	0.882359053	19.83211687
ENSBTAG000000014885	<i>MYOM3</i>	-79	0.51918909	-0.974868645	11.43563711
ENSBTAG000000001183	<i>KLHL33</i>	-41	-0.833075379	-0.162756585	4.251049669
ENSBTAG000000009145	<i>SLC7A4</i>	-125	-0.344939831	-0.856509326	0.217634264
ENSBTAG000000024555	<i>EFCAB2</i>	-29	-0.84776478	0.088160521	1.299559537
ENSBTAG000000018382	<i>F1ME02</i>	-43	1.192039708	-2.657097144	6.252619495
ENSBTAG000000000828	<i>CAPN6</i>	-67	0.586738765	-0.755526733	4.192593004
ENSBTAG000000013419	<i>HTATIP2</i>	-28	-0.555767117	-0.338865552	1.618205315
ENSBTAG000000016709	<i>NT5C3</i>	-93	-0.490046246	-0.255617117	2.750699405
ENSBTAG000000004347	<i>GPR116</i>	-52	-0.023425367	-0.419138621	5.752288067
ENSBTAG000000004036	<i>GJC1</i>	47	0.007123933	1.218822977	0.784542889
ENSBTAG000000004126	<i>MLF1</i>	-21	-0.49075088	-0.505289561	9.101955219
ENSBTAG000000022989	<i>FAM174B</i>	-38	-0.818805505	-0.780206886	0.680810238
ENSBTAG000000006819	<i>CTAGE5</i>	-25	-0.48783508	-0.228986849	3.17002846
ENSBTAG000000018775	<i>TPX2</i>	-61	-1.114613816	-0.390820144	1.085286324
ENSBTAG000000038141	<i>F1MKM4</i>	-28	-0.99065925	-0.092387362	1.243071548
ENSBTAG000000019658	<i>ASB16</i>	-10	-0.839831312	0.082641144	6.821270199
ENSBTAG000000001514	<i>ASB11</i>	13	-0.714125927	-0.24687957	5.151342456
ENSBTAG000000013750	<i>B3GALNT2</i>	-28	-0.189901549	-0.332713777	3.29720773
ENSBTAG000000035844	<i>HRASLS</i>	31	-0.537723244	-0.016217542	3.404517537
ENSBTAG000000014863	<i>GYPC</i>	28	0.820735073	1.354672512	10.26381326
ENSBTAG000000007211	<i>ASB12</i>	17	-0.78479825	0.022611786	11.87653723
ENSBTAG000000016194	<i>FBXO32</i>	35	-0.564251549	0.135386183	9.929328797
ENSBTAG000000011190	<i>FLNA</i>	131	-0.066303367	0.933834185	6.617062971
ENSBTAG000000019052	<i>ANK3</i>	-57	-0.99696185	-0.39532666	4.802480265
ENSBTAG000000002391	<i>TGFB1I1</i>	103	-0.14070739	0.723021685	2.570135691
ENSBTAG000000002914	<i>GALNTL4</i>	6	-0.683056496	-0.085946323	0.815799696
ENSBTAG000000011869	<i>CSRP3</i>	20	0.266094623	-1.775186352	29.94907855
ENSBTAG000000009496	<i>STAT5A</i>	99	0.078845957	0.451279624	2.08409973
ENSBTAG000000019517	<i>ELN</i>	3	-0.4273437	0.078033588	6.725402678
ENSBTAG000000012505	<i>ARHGEF17</i>	-99	0.02711418	0.063265458	0.863368248
ENSBTAG000000016525	<i>FIMMN6</i>	33	-0.209032233	-0.124309672	2.814115657
ENSBTAG000000002126	<i>PFKFB2</i>	-44	-0.001550513	0.61649961	0.150425167
ENSBTAG000000018747	<i>PRKAA2</i>	68	-0.75636163	0.281876682	3.421637097

Table 9 – Differential hubbing (DH), regulatory impact facto 1 and 2 (RIF1 and RIF2), and phenotypic impact factor scores between the high and low IMF groups in Nellore steers

(conclusion)					
Ensembl Gene ID	Gene Symbol	DH	RIF1	RIF2	PIF
ENSBTAG00000020764	<i>CNN2</i>	-17	-1.07639514	-0.404227074	2.721382158
ENSBTAG00000012558	<i>ADAMTS12</i>	83	0.140361341	1.332754395	0.318419953
ENSBTAG00000013632	<i>GRM4</i>	158	1.063772304	1.552436884	0.558589616
ENSBTAG00000031134	<i>F1MPT3</i>	-58	-0.106569454	-1.267747358	9.550186351
ENSBTAG00000004305	<i>RGS16</i>	59	-0.626854984	0.256203958	0.546258387
ENSBTAG00000011913	<i>CKAP4</i>	20	-0.476317926	0.042161971	1.318686924
ENSBTAG00000007632	<i>FAM151B</i>	-111	-0.645090785	-0.100943513	0.319846644
ENSBTAG00000014930	<i>MYLK2</i>	-44	-0.935338371	0.077127988	15.28345851
ENSBTAG00000003836	<i>ADAM19</i>	27	-0.473036758	0.557436605	2.788884348
ENSBTAG00000004570	<i>PRRX1</i>	-35	0.454805943	-0.05137304	3.206220001
ENSBTAG00000017233	<i>RNF213</i>	115	0.002955918	1.138883	0.10572983
ENSBTAG00000016918	<i>MYOF</i>	88	-0.158017221	1.083059034	1.13523625
ENSBTAG00000005932	<i>FAM184B</i>	-61	-0.685840884	0.064280483	0.150451815
ENSBTAG00000012450	<i>RAPGEF2</i>	-24	0.326513685	-0.152853198	2.065727725
ENSBTAG00000016013	<i>ASB14</i>	50	-0.731810445	-0.080398633	3.066175613
ENSBTAG000000021768	<i>E1BIN6</i>	-61	-0.619690878	-0.315811075	1.160495448
ENSBTAG00000004259	<i>HPCAL1</i>	-32	-0.444122799	-0.051675177	2.243219532
ENSBTAG00000006253	<i>FLNC</i>	-84	-0.376041385	-0.565130295	21.777812
ENSBTAG00000016208	<i>TGM2</i>	105	-0.146692725	0.960377666	4.613597736
ENSBTAG000000046162	<i>ZNF385C</i>	-11	-0.86596847	-0.001853201	0.302225751
ENSBTAG000000025099	<i>RNF145</i>	1	1.184187588	-0.170761446	1.760939256
ENSBTAG00000010940	<i>HSPB7</i>	-56	1.038258712	-1.860024384	21.35565395
ENSBTAG00000011010	<i>PRND</i>	-20	-0.144060299	0.192310476	0.146000442
ENSBTAG00000012720	<i>ANKRD2</i>	-66	0.152978201	-0.928129589	15.24075057
ENSBTAG00000005729	<i>FBXL4</i>	-42	-0.64308967	-0.135567756	2.91491186
ENSBTAG00000010179	<i>COL5A3</i>	163	0.437666474	1.718720563	4.938524544
ENSBTAG00000011324	<i>EMILIN1</i>	17	-0.356146169	0.161940073	2.797459302
ENSBTAG00000019090	<i>PLEKHH3</i>	-55	-0.576228787	-0.027451147	2.010071171
ENSBTAG000000027051	<i>PTAFR</i>	40	-0.067916229	0.377618378	0.006966391
ENSBTAG000000048210	<i>G5E6M6</i>	69	0.646201641	1.389638208	7.560260705
ENSBTAG00000006679	<i>MITF</i>	75	0.12122381	0.269352269	3.298589309
ENSBTAG00000013922	<i>MOSPD1</i>	-5	0.758552372	-0.017778297	0.883710574
ENSBTAG00000013131	<i>FAM110D</i>	-16	-0.27123537	-0.107270563	0.576546099
ENSBTAG00000019011	<i>PGM1</i>	20	-0.292110587	0.270949769	5.573008713
ENSBTAG00000020097	<i>PERP</i>	48	-0.746776946	-0.220861232	0.204877456
ENSBTAG00000022920	<i>RNF128</i>	31	-0.781872612	-0.003529397	2.597495671

Table 10 - Functional annotation clustering (FDR < 0.01) of 3,000 genes with higher phenotypic impact factor (PIF) scores comparing high and low IMF GEBV in Nellore steers

Category	Term	Genes	%	P-Value	FDR
GOTERM_BP_FAT	translation	164	5.9	1.3E-59	2.3E-56
GOTERM_BP_FAT	generation of precursor metabolites and energy	108	3.9	2.0E-32	3.5E-29
GOTERM_BP_FAT	protein catabolic process	114	4.1	2.7E-25	4.7E-22
GOTERM_BP_FAT	proteolysis involved in cellular protein catabolic process	106	3.8	6.5E-24	1.2E-20
GOTERM_BP_FAT	cellular protein catabolic process	106	3.8	9.6E-24	1.7E-20
GOTERM_BP_FAT	protein localization	156	5.6	1.2E-23	2.1E-20
GOTERM_BP_FAT	cellular macromolecule catabolic process	117	4.2	1.3E-23	2.3E-20
GOTERM_BP_FAT	macromolecule catabolic process	126	4.5	2.7E-23	4.8E-20
GOTERM_BP_FAT	protein transport	144	5.2	7.4E-23	1.3E-19
GOTERM_BP_FAT	establishment of protein localization	144	5.2	9.7E-23	1.7E-19
GOTERM_BP_FAT	modification-dependent protein catabolic process	96	3.5	2.6E-21	4.6E-18
GOTERM_BP_FAT	modification-dependent macromolecule catabolic process	96	3.5	2.6E-21	4.6E-18
GOTERM_BP_FAT	intracellular transport	104	3.8	7.4E-18	1.3E-14
GOTERM_BP_FAT	electron transport chain	48	1.7	2.6E-16	4.0E-13
GOTERM_BP_FAT	mRNA metabolic process	65	2.3	4.7E-14	8.4E-11
GOTERM_BP_FAT	cellular macromolecule localization	80	2.9	6.8E-14	1.2E-10
GOTERM_BP_FAT	cellular protein localization	79	2.8	1.6E-13	2.9E-10
GOTERM_BP_FAT	intracellular protein transport	75	2.7	2.9E-13	5.1E-10
GOTERM_BP_FAT	RNA splicing	48	1.7	7.4E-13	1.3E-9
GOTERM_BP_FAT	mRNA processing	58	2.1	8.4E-13	1.5E-9
GOTERM_BP_FAT	ubiquitin-dependent protein catabolic process	46	1.7	1.1E-12	1.9E-9
GOTERM_BP_FAT	energy derivation by oxidation of organic compounds	37	1.3	3.7E-12	6.6E-9
GOTERM_BP_FAT	RNA processing	88	3.2	6.8E-12	1.2E-8
GOTERM_BP_FAT	oxidative phosphorylation	33	1.2	2.8E-11	4.9E-8
GOTERM_BP_FAT	cellular respiration	28	1.0	2.0E-10	3.5E-7
GOTERM_BP_FAT	translational initiation	21	0.8	3.6E-9	6.4E-6
GOTERM_BP_FAT	protein modification by small protein conjugation or removal	31	1.1	6.5E-9	1.2E-5
GOTERM_BP_FAT	protein modification by small protein conjugation	28	1.0	1.1E-8	1.9E-5
GOTERM_BP_FAT	protein folding	45	1.6	2.0E-8	3.6E-5
GOTERM_BP_FAT	proton transport	25	0.9	4.5E-8	7.9E-5
GOTERM_BP_FAT	hydrogen transport	25	0.9	7.4E-8	1.3E-4
GOTERM_BP_FAT	aerobic respiration	17	0.6	8.5E-8	1.5E-4
GOTERM_BP_FAT	cell redox homeostasis	26	0.9	1.0E-7	1.8E-4
GOTERM_BP_FAT	ATP synthesis coupled proton transport	22	0.8	3.1E-7	5.6E-4
GOTERM_BP_FAT	energy coupled proton transport, down electrochemical gradient	22	0.8	3.1E-7	5.6E-4
GOTERM_BP_FAT	protein targeting	33	1.2	1.1E-6	1.9E-3
GOTERM_BP_FAT	tricarboxylic acid cycle	14	0.5	1.1E-6	2.0E-3
GOTERM_BP_FAT	acetyl-CoA catabolic process	14	0.5	1.1E-6	2.0E-3
GOTERM_BP_FAT	ion transmembrane transport	22	0.8	1.3E-6	2.4E-3
GOTERM_BP_FAT	vesicle-mediated transport	71	2.6	1.6E-6	2.9E-3
GOTERM_BP_FAT	regulation of cellular component biogenesis	23	0.8	1.7E-6	3.0E-3
GOTERM_BP_FAT	acetyl-CoA metabolic process	14	0.5	2.5E-6	4.5E-3
GOTERM_BP_FAT	posttranscriptional regulation of gene expression	32	1.2	2.6E-6	4.6E-3
GOTERM_BP_FAT	coenzyme catabolic process	14	0.5	5.3E-6	9.4E-3

3.4 Discussion

In this current study RNA-seq approach was used to better understand the molecular mechanisms involved with intramuscular fat deposition in beef cattle, more specifically in Nellore breed (*Bos indicus* species). RNA-seq approach was performed due to advantages over traditional cDNA microarray, which allows the expression level of whole transcriptome, and also due to the high correlation ($r = 0.929$) with qRT-PCR demonstrating to be an

accurate and reproducible (MARIONI et al., 2008) approach. Seven biological replicates were used for each condition in this study, which may be considered a significant number of samples for differentially expressed genes study (SIMS et al., 2014).

Seventy- seven significantly differentially expressed genes were identified between the Nellore steers with extremely high and low GEBV for IMF. The functional and pathway enrichment analysis (SEA and GSEA) identified several pathways, which are involved in molecular mechanisms of adipogenesis (Table 1 and Table 2). GSEA detected 13 pathways, which presents small molecules (L-cysteine, ROS, NO, ATP, etc.) and functional groups (tyrosine kinase and inflammatory cytosine) as the central regulators.

L-cysteine is involved with the process called acylation, which has important role on post-translational processes in eukaryotes (MILLIGAN; PARENTI; MAGEE, 1995; RESH, 2006). The acylation process is a covalent modification of intracellular polypeptides by the addition of C16 palmitic acid by a thioester linkage to cysteine residues of proteins. This process is also called palmitoylation and might result in significant alterations in protein/protein interactions, membrane binding and targeting, intracellular signaling and expression regulation (RESH, 2006). A previous study demonstrated that palmitoylation of *HMGCS2* protein, a gene that contains a peroxisome proliferator response element (PPRE) in its promoter region, autoregulates the *HMGCS2* transcription (KOSTIUK; KELLER; BERTHIAUME, 2010). The palmitoylation at internal cysteine residual of *HMGCS2* protein promotes binding to peroxisome proliferator activated receptor α (*PPAR* α) that act as ligand-activated transcription factors (QI; ZHU; REDDY, 2000). *PPAR* α is expressed in the liver, adipose tissue and skeletal muscle and is responsible for transcription of many genes involved with fatty acid metabolism (SCHOONJANS; STAELS; AUWERX, 1996).

Interestingly, the *HMGCS2* was not identified as a DEG herein, but transglutaminase 2 (*TGM2*) was, which present higher and positively correlation with *HMGCS2*. *HMGCS2* is a cellular component of mitochondria and associated with lipid deposition in white and brown adipocyte tissue (SVENSSON et al., 2011) and harbored at QTL region identified in previous GWAS analysis from the same population (CESAR et al., 2014).

The reactive oxygen species (ROS) molecules as superoxide, hydrogen peroxide (H_2O_2) and nitric oxide (NO) pathways were also identified in this study as associated with lipid metabolism. ROS shows important role in cellular signaling proteins as protein tyrosine phosphatases (PTP), which depend on the presence of redox-sensitive groups in their cysteine residues. Oxidation and reduction of cysteine groups is important for ROS responses to these proteins, which may explain L-cysteine pathway identification in this study. Other small

molecule identified herein and that is involved with redox mechanisms was ATP, which is produced by cellular respiration and used by enzymes and structural proteins in many cellular processes. Tormos et al. (2011) reported that the ROS production from mitochondrial complex III is important in activation of the adipogenic transcriptional cascade by CCAAT/enhancer binding protein (*C/EBP*) and *PPAR γ* regulation.

Redox mechanisms studies suggested that ROS may be involved in promoting both of the early differentiation and the subsequent maturation of adipose cells, and also that the adipogenesis induced by ROS on the time, intracellular location and intensity of ROS generation (KANDA et al., 2011). Herein, was observed that these mediators (ROS and H₂O₂) are involved with lipid metabolism when the gene expression level between animals with high and low GEBV for intramuscular fat was analyzed.

The micronutrients such as Magnesium (Mg), Zinc (Zn), and vitamin A (retinoic acid) play an important role in DNA metabolism, DNA repair, and regulation of gene expression (FENECH et al., 2005; GARCIA et al., 2013). Zinc is involved in lipolysis in adipocytes and plays an important role in insulin synthesis, storage and secretion (CHAUSMER, 1998). A previous study, reported that deficiency of zinc may causes insulin resistance and type 2 diabetes in the later stages of life in human (GARCIA et al., 2013). Zinc is also related with appetite regulation, which alters the hypothalamic neurotransmitter metabolism by leptin system (MANTZOROS et al., 1998).

Magnesium is an important cofactor for enzymes involved in carbohydrate metabolism and its deficiency in adipocytes is associated with insulin resistance. In erythrocytes, low magnesium levels increase membrane microviscosity, which decreases insulin binding to its receptor (TAKAYA et al., 2010). Similarly a Low-magnesium diet can affect tyrosine kinase activity, thereby decreasing muscle insulin receptor numbers (FERREIRA; RIVERA; ROMERO, 2004).

Retinoic acid (RA) is a pleiotropic activation factor, which regulates genes associated with cellular processes such as cell differentiation, cell proliferation, apoptosis, and embryonic development. Retinoic acid receptors (*RAR*) and retinoid X receptors (*RXR*) are the two families of nuclear receptor proteins that are associated with transcription and regulation of many genes. In addition, *RAR* and *RXR* can bind with other receptor proteins, including *PPAR*. The endogenous level of retinoic acid is very important because many genes are affected by RA signaling (BLOMHOFF; BLOMHOFF, 2006; MCGRANE, 2007). The nuclear receptors peroxisome proliferator-activated receptor gamma (*PPAR γ*) is essential to adipogenesis and to maintain adipocytes gene expression and function, however the

requirement of binding *PPARg/RXR* is important for it because has the enhancer function. Herein, we identified that Zn, Mg and retinoic acid (Figure 5) are involved in fatty acid metabolism, which are associated with expression regulation of DEG between two groups of animals that showed extremes GEBV for IMF deposition.

The inflammatory cytokines and tyrosine protein kinases were identified in this transcriptome study, which are related with the consumption of excess nutrients resulting in inflammatory metabolic signals from metabolic cells such as liver, adipocytes and pancreas.

Skeletal muscle is a major user of glucose and energy consumption, which contributes to glucose homeostasis (GREGOR; HOTAMISLIGIL, 2011). However, in the obese condition inflammatory cytokines and mediators from metabolic cells (ex. adipocytes and liver) can act on skeletal muscle thereby increasing insulin sensitivity (CAI et al., 2005).

Inflammatory kinases in the adipocytes inhibit insulin activity and increase the expression of proinflammatory cytokines such as NF- κ B and IL-1 β . This inflammatory signaling can also downregulate the activity of *PPARg*, which reduces adiponectin expression and the ability of adipocytes to make and store lipids and maintain insulin sensitivity (GUILHERME et al., 2008). In the liver, these inflammatory mediators increase gluconeogenesis and very-low-density lipoprotein (VLDL) secretion, which leads to hyperglycemia and high triglyceride levels in serum (GREGOR; HOTAMISLIGIL, 2011).

Other important tissues are affected by inflammatory system such as brain and pancreas. The brain is a central site of appetite regulation and energy expenditure, and pancreas is responsible to glucose homeostasis given its produce insulin and glucagon. In obese, the activation of inflammatory pathway results in an increase of food intake and body weight leading to hypothalamic insulin and leptin resistance. The main inflammatory mediator is IL-1 β activates NF- κ B pathway in pancreatic islets inducing nitric oxide (NO) production and apoptosis of this and β cells (ELDOR et al., 2006).

Adipose tissue was initially considered just a fat storage compartment, but actually it is an important endocrine organ, which can influence many processes including: inflammation, energy metabolism, cancer and infectious diseases (NAWROCKI; SCHERER, 2005). Adipokines and cytokines are the major secretory proteins of adipocyte cells and are involved with the regulation of metabolism, which affect a large number of tissues such as liver, muscle, brain, pancreas, reproductive and vasculature systems (SCHERER, 2006).

In beef cattle, adipose tissue development is of specific interest because the deposition and composition of IMF are involved with organoleptic characteristics, consumer preferences and economic gains. However, fat deposition is a consequence of increased energy intake and

decreased energy expenditure (GREGOR; HOTAMISLIGIL, 2011). In this transcriptome study, not many DEG were found between animals with extreme GEBV for IMF content, therefore revealed that most of these genes are related with lipid metabolism and other biological processes such as inflammatory system and glucose homeostasis.

Despite in this study 13 differentially pathways were identified a lower number of DEG are involved, which represent that some DEG are present in more than one pathway as *SPARC*, *MT3*, *PGM1* and *STAT5*, some exclusively in one pathway as *HTATIP2*, *ROCK2*, *BCR*, *FSCN* and *GYPC*, and others present in all pathways *TGM2*.

Secreted protein acidic and rich in cysteine (*SPARC*) was the first matricellular protein to be linked to the accumulation of WATs (Bradshaw *et al.*, 2003). *SPARC* inhibits differentiation of preadipocytes into adipocytes, by favoring osteoblastocyte differentiation. The expression level of *SPARC* was higher in the lower GEBV IMF cattle, similar to that reported by Nie and Sage (NIE; SAGE, 2009) in mice.

Metallothionein-3 (*MT3*) is a zinc-binding protein. An increased level of obesity was observed in null *MT3* male mice (BYUN; KIM; KOH, 2011). This increase in obesity was caused by reduced energy expenditure. Thus, in obese animals and high IMF GEBV cattle, *MT3* expression level was increase and differentially hubbed.

Phosphoglucomutase 1 (*PGM1*) is associated with diseases such as hypertension, obesity and cardiovascular disease. It is overexpressed in skeletal muscle from insulin resistant humans (NGUYEN *et al.*, 2006). Herein, the expression level of *PGM1* was higher in high IMF GEBV cattle, which corroborates the findings reported by NGUYEN *et al.* (2006) in human skeletal muscle from obese individuals.

During adipogenesis, the preadipocytes are stimulated by adipogenic factors (insulin and glucocorticoids), which change the chromatin conformation inducing the DNase I hypersensitivity “hotspots” (TAKAYA *et al.*, 2010). In these “hotspots” Mandrup *et al.* (2012) identified TF motifs and binding of *C/EBP* proteins, *GR* and *RXR* receptors and signal transducer and activator of transcription (*STAT5*). *STAT5* is activated by kinases associated with transmembrane receptors and play role in cell growth and division, apoptosis and inflammation (STOECKLIN *et al.*, 1999). *STAT5* mediates energy homeostasis in response to endogenous cytokines and herein showed greater expression level in group with high GEBV for IMF.

HIV-1 tat interactive protein 2 (*HTATIP2*) acts as a redox sensor, which is linked to regulation of nuclear import (GERHARD *et al.* 2004), while *GYPC*, *ROCK2*, *BCR* and *FSCN1* are associated with focal adhesion formation and cytoskeleton organization.

From the differentially expressed genes analysis and pathway enrichment described so far was possible to identify that the animals used here are in two phases of the adipogenesis. The animals that show lower GEBV for IMF presented higher expression level of genes associated with the first step of adipocyte differentiation, the transition of mesenchymal stem cells (MSC) to committed white adipocytes. It corroborates with lower IMF content measured in LD from the animals used in this work. The main genes associated with this first step of the adipogenesis are: *ROCK2*, *PAFTR*, *ELN*, *RGS16*, *TGM2*, *FLNA* and *GYPE*, which may mediate the actinomyosin cytoskeleton activation, leading to the expression of pro-adipogenic WNT (group of signal transduction pathways from outside to the inside of the cell by surface receptors) genes (CRISTANCHO; LAZAR, 2011).

The second step of adipogenesis that may be identified from this transcriptome study is the transition of committed white adipocytes to mature white adipocytes, which occurs by adipogenic stimuli as insulin and glucocorticoids. This phase of the adipogenesis may be indicated by the higher expression level of *HTATIP2* and *STAT5A* in high group, which participate in redox and inflammatory pathways (ROS, inflammatory cytokine, NO and retinoic acid).

However, in this study *PPAR α* and *FAS* genes, the major genes associated with fat deposition in *Bos taurus* breeds, were not identified as DEG. It may be explained by the fact that the animals used in the current study were in early stage of fat deposition. This findings corroborates with previous studies, which demonstrated that *Bos taurus* breeds are genetically predisposed to deposit intramuscular fat earlier than *Bos indicus* breeds and that adipogenesis and lipogenesis are associated with genetic background, nutrition program and time frame (LEHNERT et al., 2007; WANG et al., 2005; WANG et al., 2009).

To better understand the correlations between the expressed genes identified in this study and identify putative transcription regulators of the adipogenesis and fat deposition the mathematical approach as PCIT, RIF1, RIF2 and PIF was performed and revealed significant regulators.

PCIT approach identified ankyrin repeat domain 26 (*ANKRD26*) and phosphatidic acid phosphatase type 2 domain containing 2 (*PPAPDC2*), homeobox genes, such as *HOXC5*, zinc finger protein, friend of GATA (FOG) family member 1 (*ZFPM1*), and zinc finger protein 90 (*ZFP90*) as putative causal regulatory genes.

Ankyrin repeat domain 26 (*ANKRD26*) is expressed in the hypothalamus, brain, liver, adipose tissue and skeletal muscle and is located within the cell membrane. In humans and in mice, *ANKRD26* is responsible for white adipose tissue (WAT) insulin response and appetite

control (RACITI et al., 2011). Interestingly, one of the pathways enriched in this study, retinoic acid, has been demonstrated to be associated with adipose metabolism. Sahab et al. showed that the down-regulation of RAR could impact the expression of *ANKRD26*. The homeobox family of genes is a conserved family of transcription factors, which play important roles in morphogenesis and metabolism (ARCIONI et al., 1992).

Phosphatidic acid phosphatase type 2 domain containing 2 (*PPAPDC2*) is located within the endoplasmic reticulum and nuclear envelope in mammalian cells. Phosphatidic acid is a new class of lipid mediators, which are involved in cell growth, proliferation and reproduction pathways such as a regulatory factor (WANG et al., 2006). In *Saccharomyces cerevisiae*, phosphatidic acid was reported as an essential metabolic intermediate and a signaling lipid (LOEWEN et al., 2004). These findings may corroborate with the present PCIT results, which indicate that *ANKRD26*, *HOXC5*, and *PPAPDC2*.

Zinc-finger protein, friend of GATA (FOG) family member 1 (*ZFPM1*) and zinc finger protein 90 (*ZFP90*) are important putative regulator genes in lipid metabolism. GATA family proteins are zinc-finger transcription factors, which recognize GATA motifs in DNA. GATA plays an important role in transcriptional regulation of many genes expression (FUJIWARA et al., 1996). *ZFP90* is associated with RNA polymerase II core, which is involved, in negative regulation of transcription (GO:0001078). Previous study also showed the zinc-finger protein as transcriptional regulators comparing the expression profiles between adipogenic and non-adipogenic fibroblasts (CRISTANCHO; LAZAR, 2001).

The RIF1 analysis identified as putative regulators *IGF2* and *GLI2*, and RIF2 indicated *MPC1* and *UBL5*. *IGF2* identification of as a potential regulator of the degree of IMF has been suggested previously in pigs (ASLAN et al., 2012), which is involved in adipocytes proliferation (KHANDEKAR; COHEN; SPIEGELMAN, 2011). The *GLI2* gene, although not documented to play a role in adipose development, is a regulator of cellular proliferation and is known to be regulated by retinoic acid signaling (RIBES et al., 2009), which is consistent with the signaling mechanisms identified by Pathway Studio.

The *MPC1* gene is a part of a mitochondrial pyruvate carrier complex and has been implicated in the mitochondrial response to insulin and *PPAR γ* signaling (COLCA et al., 2013). The *UBL5* gene has been suggested to impact adipose deposition in both the pig and humans. The *UBL5* gene is a candidate gene for meat quality and IMF in the pig (CEPICA et al., 2012) and also associated with body fat and fat accumulation related to metabolic dysfunction and diabetes in humans (BOZAOGLU et al., 2006; JOWETT et al., 2004). Previous study also reported that *PPAR γ* proteasomal degradation is ubiquitin-dependent and

the stable overexpression of ubiquitin gene may be reduced PPAR γ protein levels and suppressed adipocyte differentiation in human cell (KIM et al., 2014). It demonstrates the importance of ubiquitin genes as UBL5 in adipogenesis process.

One of the top PIF genes was microRNA-1281, was previously identified as an adipose expressed miRNA that may regulate the lipid metabolism gene *EP300* (ROMAO, et al., 2014). The functional enrichment by DAVID database showed that the PIF approach was well suited to identify regulators involved with the phenotype studied (intramuscular fat deposition).

3.5 Conclusions

The present study showed the complexity of the bovine *Longissimus dorsi* transcriptome and the molecular mechanisms involved in lipid metabolism. Differential gene expression and pathway enrichment analysis revealed the transcriptome profile between Nellore steers with different percentage of IMF and indicated the stage of fat deposition of these animals. The mathematical methods as PCIT, PIF and RIF allowed the identification of putative transcription factors involved in adipogenesis and lipid metabolism. These findings may contribute to better understand the molecular mechanisms underlying the fat deposition and energy balance in cattle and provide humans with healthy beef choices.

References

- ANDERS, S.; HUBER, W. Differential expression analysis for sequence count data. **Genome Biology**, London, v. 11, 2010. Available in: <<http://genomebiology.com/content/pdf/gb-2010-11-10-r106.pdf>>. Accessed on: 04 aug. 2014.
- ARCIONI, L.; SIMEONE, A.; GUAZZI, S.; ZAPPAVIGNA, V.; BONCINELLI, E.; MAVILIO, F. The upstream region of the human homeobox gene HOX3D is a target for regulation by retinoic acid and HOX homeoproteins. **EMBO Journal**, London, v. 11, p. 265-277, 1992.
- ASLAN, O.; HAMILL, R.M.; DAVEY, G.; MCBRYAN, J.; MULLEN, A.M.; GISPERT, M.; SWEENEY, T. Variation in the IGF2 gene promoter region is associated with intramuscular fat content in porcine skeletal muscle. **Molecular Biology Reports**, Dordrecht, v. 39, p. 4101-4110, 2012.
- BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society Serie B (Methodological)**, London, v. 57, p. 289-300, 1995.

BLOMHOFF, R.; BLOMHOFF, H.K. Overview of retinoid metabolism and function. **Journal of Neurobiology**, New York, v. 66, p. 606-630, 2006.

BOZAOGLU, K.; CURRAN, J.E.; ELLIOTT, K.S.; WALDER, K.R.; DYER, T.D.; RAINWATER, D.L.; VANDEBERG, J.L.; COMUZZIE, A.G.; COLLIER, G.R.; ZIMMET, P.; MACCLUER, J.W.; JOWETT, J.B.; BLANGERO, J. Association of genetic variation within UBL5 with phenotypes of metabolic syndrome. **Human Biology**, Detroit, v. 78, p. 147-159, 2006.

BRADSHAW, A.D.; GRAVES, D.C.; MOTAMED, K.; SAGE, E.H. SPARC-null mice exhibit increased adiposity without significant differences in overall body weight. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 100, p. 6045-6050, 2003.

BYUN, H.R.; KIM, D.K.; KOH, J.Y. Obesity and downregulated hypothalamic leptin receptors in male metallothionein-3-null mice. **Neurobiology of Disease**, San Diego, v. 44, p. 125-132, 2011.

CAI, Z.; CHEN, H.T.; BOYLE, B.; RUPP, F.; FUNK, W.D.; DEDERA, D.A. Identification of a novel insulin-like growth factor binding protein gene homologue with tumor suppressor like properties. **Biochemical and Biophysical Research Communications**, Orlando, v. 331, p. 261-266, 2005.

CEPICA, S.; OVILO, C.; MASOPUST, M.; KNOLL, A.; FERNANDEZ, A.; LOPEZ, A.; ROHRER, G.A.; NONNEMAN, D. Four genes located on a SSC2 meat quality QTL region are associated with different meat quality traits in Landrace x Chinese-European crossbred population. **Animal Genetics**, Oxford, v. 43, p. 333-336, 2012.

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.; COUTINHO, L.L. Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. **BMC Genetics**, London, v. 15, 2014. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2156-15-39.pdf>>. Accessed on: 10 apr. 2014.

CHAUSMER, A.B. Zinc, insulin and diabetes. **Journal of the American College of Nutrition**, New York, v. 17, p. 109-115, 1998.

COLCA, J.R.; VANDERLUGT, J.T.; ADAMS, W.J.; SHASHLO, A.; MCDONALD, W.G.; LIANG, J.; ZHOU, R.; ORLOFF, D.G. Clinical proof-of-concept study with MSDC-0160, a prototype mTOT-modulating insulin sensitizer. **Clinical Pharmacology and Therapeutics**, Saint Louis, v. 93, p. 352-359, 2013.

CRISTANCHO, A.G.; LAZAR, M.A. Forming functional fat: a growing understanding of adipocyte differentiation. **Nature Reviews Molecular Cell Biology**, London, v. 12, p. 722-734, 2011.

ELDOR, R.; YEFFET, A.; BAUM, K.; DOVINER, V.; AMAR, D.; BEN-NERIAH, Y.; CHRISTOFORI, G.; PELED, A.; CAREL, J.C.; BOITARD, C.; KLEIN, T.; SERUP, P.; EIZIRIK, D.L.; MELLOUL, D. Conditional and specific NF-kappaB blockade protects pancreatic beta cells from diabetogenic agents. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 103, p. 5072-5077, 2006.

FENECH, M.; BAGHURST, P.; LUDERER, W.; TURNER, J.; RECORD, S.; CEPPI, M.; BONASSI, S. Low intake of calcium, folate, nicotinic acid, vitamin E, retinol, beta-carotene and high intake of pantothenic acid, biotin and riboflavin are significantly associated with increased genome instability: results from a dietary intake and micronucleus index survey in South Australia. **Carcinogenesis**, Oxford, v. 26, p. 991-999, 2005.

FERREIRA, A.; RIVERA, A.; ROMERO, J.R. Na⁺/Mg²⁺ exchange is functionally coupled to the insulin receptor. **Journal of Cellular Physiology**, Philadelphia, v. 19, p. 434-440, 2004.

FUJIWARA, Y.; BROWNE, C.P.; CUNNIFF, K.; GOFF, S.C. ORKIN, S.H. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 93, p. 12355-12358, 1996.

GARCIA, O.P.; RONQUILLO, D.; DEL CARMEN CAAMANO, M.; MARTINEZ, G.; CAMACHO, M.; LOPEZ, V.; ROSADO, J.L. Zinc, iron and vitamins A, C and e are associated with obesity, inflammation, lipid profile and insulin resistance in mexican school-aged children. **Nutrients**, Klybeckstrasse, v. 5, p. 5012-5030, 2013.

GERHARD, D.S.; WAGNER, L.; FEINGOLD, E.A.; SHENMEN, C.M.; GROUSE, L.H.; SCHULER, G.; KLEIN, S.L.; OLD, S.; RASOOLY, R.; GOOD, P.; GUYER, M.; PECK, A.M.; DERGE, J.G.; LIPMAN, D.; COLLINS, F.S.; JANG, W.; SHERRY, S.; FEOLO, M.; MISQUITTA, L.; LEE, E.; ROTMISTROVSKY, K.; GREENHUT, S.F.; SCHAEFER, C.F.; BUETOW, K.; BONNER, T.I.; HAUSSLER, D.; KENT, J.; KIEKHAUS, M.; FUREY, T.; BRENT, M.; PRANGE, C.; SCHREIBER, K.; SHAPIRO, N.; BHAT, N.K.; HOPKINS, R.F.; HSIE, F.; DRISCOLL, T.; SOARES, M.B.; CASAVANT, T.L.; SCHEETZ, T.E.; BROWN-STEIN, M.J.; USDIN, T.B.; TOSHIYUKI, S.; CARNINCI, P.; PIAO, Y.; DUDEKULA, D.B.; KO, M.S.; KAWAKAMI, K.; SUZUKI, Y.; SUGANO, S.; GRUBER, C.E.; SMITH, M.R.; SIMMONS, B.; MOORE, T.; WATERMAN, R.; JOHNSON, S.L.; RUAN, Y.; WEI, C.L.; MATHAVAN, S.; GUNARATNE, P.H.; WU, J.; GARCIA, A.M.; HULYK, S.W.; FUH, E.; YUAN, Y.; SNEED, A.; KOWIS, C.; HODGSON, A.; MUZNY, D.M.; MCPHERSON, J.; GIBBS, R.A.; FAHEY, J.; HELTON, E.; KETTEMAN, M.; MADAN, A.; RODRIGUES, S.; SANCHEZ, A.; WHITING, M.; MADARI, A.; YOUNG, A.C.; WETHERBY, K.D.; GRANITE, S.J.; KWONG, P.N.; BRINKLEY, C.P.; PEARSON, R.L.; BOUFFARD, G.G.; BLAKESLY, R.W.; GREEN, E.D.; DICKSON, M.C.; RODRIGUEZ, A.C.; GRIMWOOD, J.; SCHMUTZ, J.; MYERS, R.M.; BUTTERFIELD, Y.S.; GRIFFITH, M.; GRIFFITH, O.L.; KRZYWINSKI, M.I.; LIAO, N.; MORIN, R.; PALMQUIST, D.; PETRESCU, A.S.; SKALSKA, U.; SMAILUS, D.E.; STOTT, J.M.; SCHNERCH, A.; SCHEIN, J.E.; JONES, S.J.; HOLT, R.A.; BAROSS, A.; MARRA, M.A.; CLIFTON, S.; MAKOWSKI, K.A.; BOSAK, S.; MALEK, J.; MGC PROJECT TEAM. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). **Genome Research**, Cold Spring Harbor, v. 14, p. 2121-2127, 2004.

GILMOUR, A. R.; GOGEL, B.J.; CULLIS, B.R.; THOMPSON R. **ASReml user guide release 3.0**. VSN International, 2009. Available in: <http://vsni.de/downloads/asreml/release3/UserGuide.pdf>. Accessed on: 26 sept. 2013.

GREGOR, M.F.; HOTAMISLIGIL, G.S. Inflammatory mechanisms in obesity. **Annual Review of Immunology**, Palo Alto, v. 29, p. 415-445, 2011.

GUILHERME, A.; VIRBASIS, J.V.; PURI, V.; CZECH, M.P. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. **Nature Reviews Molecular Cell Biology**, London, v. 9, p. 367-377, 2008.

HOCQUETTE, J.F.; VAN WEZEMAEL, L.; CHRIKI, S.; LEGRAND, I.; VERBEKE, W.; FARMER, L.; SCOLLAN, N.D.; POLKINGHORNE, R.; RODBOTTEN, R.; ALLEN, P.; PETHICK, D.W. Modelling of beef sensory quality for a better prediction of palatability. **Meat Science**, London, v. 1307, 2013. Available in: http://ac.els-cdn.com/S0309174013003586/1-s2.0-S0309174013003586-main.pdf?_tid=96a410f6-c644-11e3-8208-00000aacb362&acdnat=1397748543_1a9a59dadd4dd9c9aabd48c4796cbb5b. Accessed on: 12 jan. 2014.

HUANG, W.; SHERMAN, B.T.; LEMPICKI, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. **Nucleic Acids Research**, Oxford, v. 37, p. 1-13, 2009.

HUDSON, N.J.; REVERTER, A.; DALRYMPLE, B.P. A differential wiring analysis of expression data correctly identifies the gene containing the causal mutation. **PLoS Computational Biology**, London, v. 5, 2009. Available in: <http://www.ploscompbiol.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pcbi.1000382&representation=PDF>. Accessed on: 12 jan. 2014.

JIN, W.; OLSON, E.N.; MOORE, S.S.; BASARAB, J.A.; BASU, U.; GUAN, L.L. Transcriptome analysis of subcutaneous adipose tissues in beef cattle using 3' digital gene expression-tag profiling. **Journal of Animal Science**, Savoy, v. 90, p. 171-183, 2012.

JOWETT, J.B.; ELLIOTT, K.S.; CURRAN, J.E.; HUNT, N.; WALDER, K.R.; COLLIER, G.R.; ZIMMET, P.Z.; BLANGERO, J. Genetic variation in BEACON influences quantitative variation in metabolic syndrome-related phenotypes. **Diabetes**, New York, v. 53, p. 2467-2472, 2004.

KANDA, Y.; HINATA, T.; KANG, S.W.; WATANABE, Y. Reactive oxygen species mediate adipocyte differentiation in mesenchymal stem cells. **Life Science**, Zhongyuan, v. 89, p. 250-258, 2011.

KHANDEKAR, M.J.; COHEN, P.; SPIEGELMAN, B.M. Molecular mechanisms of cancer development in obesity. **Nature Reviews Cancer**, London, v. 11, p. 886-895, 2011.

KOESTERKE, L.; MILFELD, K.; VAUGHN, M.W.; STANZIONE, D.; KOLTES, J.E.; WEEKS, N.T.; REECY, J.M. Optimizing the PCIT algorithm on stampede's Xeon and Xeon Phi processors for faster discovery of biological networks. In: CONFERENCE ON EXTREME SCIENCE AND ENGINEERING DISCOVERY ENVIRONMENT: GATEWAY TO DISCOVERY, 2013, San Diego. **Proceedings...** San Diego: ACM, 2013. p. 1-8.

KOSTIUK, M.A.; KELLER, B.O.; BERTHIAUME, L.G. Palmitoylation of ketogenic enzyme HMGCS2 enhances its interaction with PPARalpha and transcription at the Hmgcs2 PPRE. **FASEB Journal**, Bethesda, v. 24, p.1914-1924, 2010.

LANGMEAD, B.; TRAPNELL, C.; POP, M.; SALZBERG, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome Biology**, London, v. 10, 2009. Available in: <<http://genomebiology.com/content/pdf/gb-2009-10-3-r25.pdf>>. Accessed on: 18 Oct. 2013.

LEHNERT, S.A.; REVERTER, A.; BYRNE, K.A.; WANG, Y.; NATTRASS, G.S.; HUDSON, N.J.; GREENWOOD, P.L. Gene expression studies of developing bovine longissimus muscle from two different beef cattle breeds. **BMC Developmental Biology**, London, v. 7, 2007. Available in: <<http://www.biomedcentral.com/content/pdf/1471-213X-7-95.pdf>>. Accessed on: 18 oct. 2013.

LOEWEN, C.J.; GASPAR, M.L.; JESCH, S.A.; DELON, C.; KTISTAKIS, N.T.; HENRY, S.A.; LEVINE, T.P. Phospholipid metabolism regulated by a transcription factor sensing phosphatidic acid. **Science**, New York, v. 304, p. 1644-1647, 2004.

MANDRUP, S.; HAGER, G.L. Modulation of chromatin access during adipocyte differentiation. **Nucleus**, Lahore, v. 3, p. 12-15, 2012.

MANTZOROS, C.S.; PRASAD, A.S.; BECK, F.W.; GRABOWSKI, S.; KAPLAN, J.; ADAIR, C.; BREWER, G.J. Zinc may regulate serum leptin concentrations in humans. **Journal of the American College of Nutrition**, New York, v. 17, p. 270-275, 1998.

MARIONI, J.C.; MASON, C.E.; MANE, S.M.; STEPHENS, M.; GILAD, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. **Genome Research**, Cold Spring Harbor, v. 18, p. 1509-1517, 2008.

MCCARTHY, F.M.; GRESHAM, C.R.; BUZA, T.J.; CHOUVARINE, P.; PILLAI, L.R.; KUMAR, R.; OZKAN, S.; WANG, H.; MANDA, P.; ARICK, T.; BRIDGES, S.M.; BURGESS, S.C. AgBase: supporting functional modeling in agricultural organisms. **Nucleic Acids Research**, Oxford, v. 39, p. 497-506, 2011.

MCGRANE, M.M. Vitamin A regulation of gene expression: molecular mechanism of a prototype gene. **Journal of Nutritional Biochemistry**, Stoneham, v. 18, p. 497-508, 2007.

MICHAL, J.J.; ZHANG, Z.W.; GASKINS, C.T.; JIANG, Z. The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. **Animal Genetics**, Oxford, v. 37, p. 400-402, 2006.

MILLIGAN, G.; PARENTI, M.; MAGEE, A.I. The dynamic role of palmitoylation in signal transduction. **Trends in Biochemical Sciences**, Amsterdam, v. 20, p. 181-187, 1995.

NAWROCKI, A.R.; SCHERER, P.E. Keynote review: the adipocyte as a drug discovery target. **Drug Discovery Today**, Killington, v. 10, p. 1219-1230, 2005.

NGUYEN, L.L.; KRIKETOS, A.D.; HANCOCK, D.P.; CATERSON, I.D.; DENYER, G.S. Insulin resistance does not influence gene expression in skeletal muscle. **Journal of Biochemistry and Molecular Biology**, Seoul, v. 39, p. 457-463, 2006.

NIE, J.; SAGE, E.H. SPARC functions as an inhibitor of adipogenesis. **Journal of Cell Communication and Signaling**, London, v. 3, p. 247-254, 2009.

NIKITIN, A.; EGOROV, S.; DARASELIA, N.; MAZO, I. Pathway studio-the analysis and navigation of molecular networks. **Bioinformatics**, Oxford, v. 19, p. 2155-2157, 2003.

O'QUINN, T.G.; BROOKS, J.C.; POLKINGHORNE, R.J.; GARMYN, A.J.; JOHNSON, B.J.; STARKEY, J.D.; RATHMANN, R.J.; MILLER, M.F. Consumer assessment of beef strip loin steaks of varying fat levels. **Journal of Animal Science**, Savoy, v. 90, p. 626-634. 2012.

QI, C.,; ZHU, Y.; REDDY, J.K. Peroxisome proliferator-activated receptors, coactivators, and downstream targets. **Cell Biochemistry and Biophysics**, Totowa, v. 32, p. 187-204, 2000.

RACITI, G.A.; BERA, T.K.; GAVRILOVA, O.; PASTAN, I. Partial inactivation of Ankrd26 causes diabetes with enhanced insulin responsiveness of adipose tissue in mice. **Diabetologia**, Bristol, v. 54, p. 2911-2922, 2011.

RAMAYO-CALDAS, Y.; MACH, N.; ESTEVE-CODINA, A.; COROMINAS, J.; CASTELLO, A.; BALLESTER, M.; ESTELLE, J.; IBANEZ-ESCRICHE, N.; FERNANDEZ, A.I.; PEREZ-ENCISO, M.; FOLCH, J.M. Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. **BMC Genomics**, London, v. 13, 2012. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2164-13-547.pdf>>. Accessed on: 18 Oct. 2013.

RESH, M.D. Palmitoylation of ligands, receptors, and intracellular signaling molecules. **Science's Signal Transduction Knowledge Environment**, New York, v. 359, re14, 2006. Available in: <<http://stke.sciencemag.org/cgi/reprint/sigtrans;2006/359/re14.pdf>>. Accessed on: 18 Oct. 2013.

REVERTER, A.; CHAN, E.K. Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. **Bioinformatics**, Oxford, v. 24, p. 2491-2497, 2008.

REVERTER, A.; HUDSON, N.J.; NAGARAJ, S.H.; PEREZ-ENCISO, M.; DALRYMPLE, B.P. Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data. **Bioinformatics**, Oxford, v. 26, p. 896-904, 2010.

RIBES, V.; LE ROUX, I.; RHINN, M.; SCHUHBAUR, B.; DOLLE, P. Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. **Development**, Cambridge, v. 136, p. 665-676., 2009.

ROMAO, J.M.; JIN, W.; HE, M.; MCALLISTER, T.; GUAN LE, L. MicroRNAs in bovine adipogenesis: genomic context, expression and function. **BMC Genomics**, London, v. 15, 2014. Available in: <<http://www.biomedcentral.com/content/pdf/1471-2164-15-137.pdf>>. Accessed on: 18 Mar. 2013.

RULE, D.C.; MACNEIL, M.D.; SHORT, R.E. Influence of sire growth potential, time on feed, and growing-finishing strategy on cholesterol and fatty acids of the ground carcass and longissimus muscle of beef steers. **Journal of Animal Science**, Savoy, v. 75, n. 6, p. 1525-1533, 1997.

SAHAB, Z.J.; HALL, M.D.; ZHANG, L.; CHEEMA, A.K.; BYERS, S.W. Tumor Suppressor RARRES1 Regulates DLG2, PP2A, VCP, EB1, and Ankrd26. **Journal of Cancer**, New York, v. 1, p. 14-22, 2010.

SCHERER, P.E. Adipose tissue: from lipid storage compartment to endocrine organ. **Diabetes**, New York, v. 55, p. 1537-1545, 2006.

SCHOONJANS, K.; STAELS, B.; AUWERX, J. The peroxisome proliferator activated receptors (PPARS) and their effects on lipid metabolism and adipocyte differentiation. **Biochimica et Biophysica Acta**, Amsterdam, v. 1302, p. 93-109, 1996.

SIMS, D.; SUDBERY, I.; ILOTT, N.E.; HEGER, A.; PONTING, C.P. Sequencing depth and coverage: key considerations in genomic analyses. **Nature Reviews Genetics**, London, v. 15, p. 121-132, 2014.

STOECKLIN, E.; WISSLER, M.; SCHAETZLE, D.; PFITZNER, E.; GRONER, B. Interactions in the transcriptional regulation exerted by Stat5 and by members of the steroid hormone receptor family. **Journal of Steroid Biochemistry and Molecular Biology**, Oxford, v. 69, p. 195-204, 1999.

SVENSSON, P.A.; JERNAS, M.; SJOHOLM, K.; HOFFMANN, J.M.; NILSSON, B.E.; HANSSON, M.; CARLSSON, L.M. Gene expression in human brown adipose tissue. **International Journal of Molecular Medicine**, Cambridge, v. 27, p. 227-232, 2011.

TAKAYA, J.; YAMATO, F.; KUROYANAGI, Y.; HIGASHINO, H.; KANEKO, K. Intracellular magnesium of obese and type 2 diabetes mellitus children. **Diabetes Therapy**, London, v. 1, n. 1, p. 25-31, 2010.

TRAPNELL, C.; PACHTER, L.; SALZBERG, S.L. TopHat: discovering splice junctions with RNA-Seq. **Bioinformatics**, Oxford, v. 25, p. 1105-1111, 2009.

TRAPNELL, C.; WILLIAMS, B.A.; PERTEA, G.; MORTAZAVI, A.; KWAN, G.; VAN BAREN, M.J.; SALZBERG, S.L.; WOLD, B.J.; PACHTER, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. **Nature Biotechnology**, London, v. 28, p. 511-515, 2010.

TORMOS, K.V.; ANSO, E.; HAMANAKA, R.B.; EISENBART, J.; JOSEPH, J.; KALYANARAMAN, B.; CHANDEL, N.S. Mitochondrial complex III ROS regulate adipocyte differentiation. **Cell Metabolism**, Cambridge, v. 14, p. 537-544, 2011.

VANRADEN, P.M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, Champaign, v. 91, p. 4414-4423, 2008.

WANG, X.; DEVAIAH, S.P.; ZHANG, W.; WELTI, R. Signaling functions of phosphatidic acid. **Progress in Lipid Research**, Oxford, v. 45, p. 250-278, 2006.

WANG, Y.H.; BYRNE, K.A.; REVERTER, A.; HARPER, G.S.; TANIGUCHI, M.; MCWILLIAM, S.M.; MANNEN, H.; OYAMA, K.; LEHNERT, S. Transcriptional profiling of skeletal muscle tissue from two breeds of cattle. **Mammalian Genome**, London, v. 16, p. 201–210, 2005.

WANG, Y.H.; BOWER, N.I.; REVERTER, A.; TAN, S.H.; DE JAGER, N.; WANG, R.; MCWILLIAM, S.M.; CAFÉ, L.M.; GREENWOOD, P.L.; LEHNERT, S.A. Gene expression patterns during intramuscular fat development in cattle. **Journal of Animal Science**, Savoy, v. 87, p. 119-130, 2009.