GMR
The Original by FUNPEC-RP

Genomic regions and genes associated with carcass quality in Nelore cattle

M.E. Carvalho[1], F.S. Baldi[2], P.A. Alexandre[1], M.H.A. Santana[1], R.V. Ventura[3], R.S. Bueno[4], M.N. Bonin[5], F.M. Rezende[6], L.L. Coutinho[7], J.P. Eler[1] and J.B.S. Ferraz[1]

[1] Departamento de Medicina Veterinária, Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP, Brasil;
[2] Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP, Brasil;
[3] Departamento de Nutrição e Produção Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, Pirassununga, SP, Brasil;
[4] Departamento Ciências Básicas, Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, SP, Brasil;
[5] Faculdade de Medicina Veterinária e Zootecnia, Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, Brasil;
[6] Faculdade de Medicina Veterinária, Universidade Federal de Uberlândia, Uberlândia, MG, Brasil;
[7] Departamento de Ciência Animal, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, Brasil

Corresponding author: M.E. Carvalho
E-mail: minosec@gmail.com

Genet. Mol. Res. 18 (1): gmr18226
Received December 04, 2018
Accepted February 06, 2019
Published February 26, 2019
DOI http://dx.doi.org/10.4238/gmr18226

ABSTRACT. Many studies have explored variability to select cattle with high genetic potential for economic interest traits. Genetic variability is a powerful tool to improve production indexes in cattle, as it also is associated with variations in meat and carcass quality traits. We made a Genome-Wide Association Study of beef cattle of *Bos indicus* origin, in particular Nelore animals, to identify regions and genes associated with carcass quality, by examining phenotypic and genotypic data from 909 animals. Several genes in associated regions were observed to have above 1% of the portion of explained genetic variance

explained: for hot carcass weight, genes *LRGUK*, *TRIM24*, *SVOPL*, *TEX37*, *CA10*, *OXSR1*; for ribeye area, genes *TWIST2*, *SFXN1*, *CMYA5*, *CPQ* and *MRS2*; for backfat thickness, genes *OR2S2*, *5S_rRNA*, *LOC100299372*, *LOC523083*, *LOC532403*, *LOC613441*, *SNORA69* and *ITGA9*; and for marbling, genes *EMCN*, *LNX1*, *EIF5*, *SNORA28* and *DSC3*. The various genomic regions associated with small effects show the complexity of these phenotypes and that they do not depend only on the effects of a few genes to determine their variations.

Key words: *Bos indicus*; GWAS; Quality carcass; Nelore

INTRODUCTION

The meat industry is investing in research to identify predictors of quality and higher yields in a search for production with quality (Bernard et al., 2007). In addition, studies have explored variability to select animals with higher potential for genetic traits of interest, since genetic variability is useful to achieve higher production indexes, including meat and carcass quality traits. Some carcass quality traits are difficult to measure in live animals; thus, it is necessary to investigate variants that influence the genetic potential for this type of trait to select the best animals (Hocquette et al., 2007). For traditional breeding programs, traits such as carcass quality, which varies from low to medium heritability, make genetic gain slow and challenging (Corva et al., 2007). Advances in biotechnologies, such as high-density single nucleotide polymorphism (SNP) chips, allow the development of new strategies to identify genes or genomic regions responsible for traits of interest (Orozco et al., 2010). GWAS (Genome Wide Association Study) methodology, used in the analysis of high-density chip results, compares allelic frequencies of thousands of available polymorphic markers in unrelated individuals with a phenotypic condition of interest to identify markers associated to such phenotype. Regions from 500 kb to 1000 kb are involved in GWAS and are analyzed by statistical tools suitable for each situation. This application has been successfully used in breeding programs for improving cow milk production, increasing accuracy with the use of genomic EPD (Expected Progeny Differences) (Hayes et al., 2009).

Zebu cattle of *Bos indicus* origin play an important role in the meat production system and genomic predictions can improve production efficiency (Garcia et al., 2012). In a previous study, our group evaluated the association of genomic regions with meat tenderness, and also observed genes that may help the understanding of the trait for *Bos indicus* (Carvalho et al., 2017). We applied GWAS techniques in beef cattle of *Bos indicus* origin, in particular Nelore animals, to identify regions and genes associated with carcass quality.

MATERIAL AND METHODS

Animals and slaughter

Selected animals were obtained from farms belonging to the company Agropecuária CFM Ltda. This herd is located in the western region of the state of São Paulo. Carcass

traits data were obtained from the Research Support Center for Animal Breeding, Biotechnology and Transgenics (NAP-GMABT), belonging to the Department of Veterinary Medicine of the Faculty of Animal Science and Food Engineering of the University of São Paulo (FZEA / USP), in Pirassununga, São Paulo.

DNA was extracted from 909 progenies (males) from blood samples collected with EDTA and blood impregnated in FTA® (Flinders Technology Associates) cards, using the method of extraction and precipitation in NaCl, following the protocol described by Olerup and Zetterquist (1992).The progeny distribution per bull contemplated a minimum of three and a maximum of 51 individuals per bull, totaling 73 bulls represented in the evaluated population and a pedigree with 6276 individuals. Phenotypes were collected at the Marfrig Group slaughterhouse, when the animals reached a weight of 550 kg and were sent to slaughter.

Hot carcass weight, ribeye area, backfat thickness and marbling

At the end of the slaughter line, with the carcass eviscerated and cut in half, the two carcass halves were weighed to obtain the hot carcass weight phenotype. Measurements were made at the *Longissimus thoracis* muscle, between the 12^{th} and 13^{th} rib of the left half carcass. A transparent grid, in centimeters, was superimposed over the transversely cut muscle, determining the ribeye area. The backfat thickness was determined with a ruler graduated in millimeters (Luchiari Filho, 2000).

For the marbling analysis, we used the standard of the United States Department of Agriculture (USDA Quality Grade 1999) as a reference, with six assessment scales, divided into four subclasses, 0, 25, 50 and 75 for conversion of subjective values into numerical values. Marbling measurements were made after 24 h of carcass cooling.

Genotyping (SNPs)

Two SNP chips were used for genotyping process. A total of 502 samples were genotyped with the GeneSeek SNP Beadchips Bovine GGP-HDi® assay (about 74K SNPs) and 407 samples with Illumina BovineHD® (777K SNPs), based on the Infinium chemistry, according to the manufacturer's protocol. The genotypes were determined using GenomeStudio software v2011.1. Calls were defined as genotypes with a GenCall Score equal to or greater than 0.15 (standard for trials with Infinium chemistry) using the cluster file provided by the manufacturer. For each sample, the overall quality of genotyping was evaluated by the genotype determination rate (Call rate), defined as the ratio between Calls and total number of markers. Samples with a Call rate below 0.9 (90% of the determined genotypes) were discarded. The animals genotyped with the 74K SNP chip were imputed using FImpute software to the Illumina BovineHD® (777K) panel. For imputation, SNPs with a P-value in a Hardy-Weinberg equilibrium z-test less than or equal to 10^{-5} and SNPs with a MAF less than 0.0001 were not considered. The accuracy of imputation was determined by cross validation for each animal in a parallel investigation, the concordance rate between the imputed and the real genotype was higher than 97.51%. For the Genome-Wide Association Study, the SNP quality control excluded SNPs of unknown genomic position and those on the sex chromosomes, monomorphic SNPs and SNPs with MAF

<0.05, markers that presented a call rate <90%, and markers with heterozygous genotype excess. After quality control 463,995 SNPs remained for association analysis.

Genomic association analysis

The ssGWAS method is a modification of BLUP with numerator relationship matrix A^{-1} matrix replaced by H^{-1} (Aguilar et al., 2010):

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix} \quad \text{(Eq. 1)}$$

where A22 is a numerator relationship matrix for genotyped animals and G is a genomic relationship matrix. The genomic matrix can be created following (VanRaden, 2008) as:

$$G = ZDZ'q \quad \text{(Eq. 2)}$$

where *Z* is a matrix of gene containing adjusted for allele frequencies, *D* is a weight matrix for SNP (initially D=I), and q is a weighting factor. The weighting factor used was according to Vitezica et al. (2011), by ensuring that the average diagonal in *G* is close to that of A_{22}. The SNP effects and weights for GWAS were derived as follows (Wang et al., 2012):

1. Let D=I in the first step.
2. Calculate G=ZDZ'qG=ZDZ'q.
3. Calculate GEBVs for the entire data set using ssGBLUP.
4. Convert GEBVs to SNP effects

$$\hat{u} = \frac{\sigma_u^2}{\sigma_a^2} DZ'G *^{-1} \hat{a}_g = DZ'[ZDZ']^{-1}\hat{a}_g \quad \text{(Eq. 3)}$$

where $\hat{a}_g$ is the GEBV of the animals which were also genotyped.

5. Calculate weight for the each SNP: $d_i = \hat{u}_i^2 2p_i(1-p_i)$, where I is the i-th SNP.
6. Normalized SNP weight to remain the total genetic variance constant.
7. Loop to 2.

The SNP weights were calculated iteratively looping through steps 4-6. Iterations increase weights of SNP with large effects and decrease those with small effects.

The percentage of genetic variance explained by the i^{-th} region was calculated as below:

$$\frac{Var(a_i)}{\sigma_a^2} = x100 = \frac{Var(\sum_{j=1}^{10} Z_j \hat{u}_j)}{\sigma_a^2} x10 \quad \text{(Eq. 4)}$$

where a_i is the genetic value of the *i*-th region, which consists of continuous 10 SNPs, σ_a^2 is the total genetic variance, Zj is vector of gene content of the j^{-th} SNP for all individuals, and $\hat{u}_j$ is marker effect of the i^{-th} SNP within the i^{-th} region.

Quantitative genetic-analysis

The contemporary groups (CGs) included animals born on the same farm and the same year, and from the same management group at slaughter. The CGs that contained less than three observations and observations that deviated ± 3 deviations from the mean of that

group were eliminated. The model used for the variance standard component estimation included random additive direct genetic effect, the fixed effect of the CG, and the animal's slaughter age as a covariable (linear and quadratic effect).

The variance components and genetic parameters were estimated using the REMLF90 program, also of the BLUPF90 family, and ssGWAS computer programs were used (Misztal et al., 2002; Aguilar et al., 2011). The statistical model can be represented by the following matrix form:

$$y = X\beta + Za + e \quad \text{(Eq. 5)}$$

where y is the vector of observations, β is the vector of fixed effects, a is the vector of direct additive genetic effects, X is the known incidence matrix, Z is the incidence matrix of the random additive direct genetic effect (associates vector β with vector y), and e is the vector of the residual effect.

Functional Analysis of SNPs

Markers located in windows associated with ribeye area, backfat thickness, marbling score and hot carcass weight were analyzed using the online tool VEP (Variant Effect Predictor - http://www.ensembl.org/info/docs/tools/vep/index.html), which predicts the functional role of SNPs based on the reference bovine genome and the position of the variants.

Prospecting genes

To determine possible regions of QTLs as a criterion to prospect genes, segments that explained values equal to or greater than 1% of the additive genetic variance were chosen. To identify the genes and their position in the selected segments in the bovine genome, a search was performed in the database resources of the National Center for Biotechnology Information (NCBI - http://www.ncbi.nlm.nih.gov) and Ensembl Genome Browser: //www.ensembl.org/index.html). These databases allowed to identify candidate genes that may be affecting the variability in expression of each trait. Identification of metabolic pathways from the genes contained in the associated regions was performed using the Panther classification system (Huaiyu et al., 2010) (Supplementary Table S2). The search for previously described QTLs was performed in CattleQTLdb using *Browse Cattle Trait Hierarchy for QTLdb* tool (Hu et al., 2016).

RESULTS AND DISCUSSION

Description of phenotypic data

Table 1 shows descriptive data of the analyzed phenotypes. The animals slaughtered were young, approximately 24 months old, reared on pasture and feedlot finished for 90 to 110 days. Except for marbling, the other traits are objective measures with metric units of value, which directly represent the phenotype. For marbling, the measurement was performed with scores that range between 400 and 800, according to the USDA - Quality Grade System. However, in the analyses, the scores were converted to numerical values,

following the scaling of the USDA system itself - Quality Grade - only for the database to contemplate numerical characters for reading the analysis programs. Thus, although the descriptive table contains average marbling values, the statistical parameter "mode", which was 400 out of 833 animals (91.6%), represents the trait distribution in the population, and is the minimum value in the marbling scale used.

Heritabilities found in this study were similar to those reported in other studies, such as Tizioto et al., (2013). However, the methodologies used in our study, both for estimation of genetic parameters and in GWAS analyses, are different from those of the Tizioto study and take into account the inclusion of pedigree in the analyses. For ribeye area and backfat thickness, Medeiros et al. (2017), using different methodologies for evaluation of the same breed, obtained higher heritability values (0.47 and 0.28) than those observed in our study. Magalhães et al., (2016) observed a lower heritability value for marbling (0.10). This difference can be attributed to several factors, including a possible difference of population lineages existing within the Brazilian Nelore herd; depth of the pedigree and also the number of animals evaluated.

Table 1. Raw means, variance components and heritability of the carcass traits.

Trait	N	Mean	Min	Max	σ_a^2	σ_e^2	h^2
Hot carcass weight (kg)	906	292.40	229.00	393.00	35.20	276.3	0.11
Ribeye area (cm^2)	900	73.58	52	101	10.98	48.08	0.18
Backfat thickness (mm)	898	4.70	1	15	0.49	3.505	0.12
Marbling	896	433.92	400	675	782.7	1492.0	0.34

N: number of animals; Min: minimum and Max: maximum.

Hot carcass weight

Hot carcass weight is an important carcass yield indicator for the meat industry and is determined by animal growth, sex, age, nutrition and genetics. Hot carcass weight is used to provide a production parameter in relation to the living animal. For this trait, 14 associated regions were observed in 10 different chromosomes with six genes identified (Table 2). The genes identified in these regions are described in the Ensembl Genome Browser, as follows: *Bos taurus leucine-rich repeats and guanylate kinase domain containing; tripartite motif containing 24; SVOP-like; testis expressed 37; carbonic anhydrase X* and *oxidative stress responsive 1*.

Among the functions of genes associated with hot carcass weight, *LRCUK* is a major determinant of microtubule structure within the male germ line (Liu et al., 2015). The gene *TRIM24* was identified as an upstream regulator; it is involved in a key part of the *TP53* mechanisms in cattle. It promotes the degradation of TP53 and is primarily involved in *TP53* induced apoptosis. *TP53* has been identified as being activated when growing follicles enter the plateau phase and initiate atresia (Mazzoni et al., 2017). *CA10* is exclusively expressed in the brain and is considered to have an essential, yet unknown, functional role because it is highly conserved across animal species (Aspatwar et al., 2014). *OXSR1* regulates downstream kinases in response to environmental stress and plays a role in ion co-transportation in kidney, reactive oxygen species formation, and migration of dendritic cells. Since stress plays a key role in meat quality, the function and localization of *OXSR1* suggest it is a candidate gene (Ponsuksili, et al., 2014). There are few descriptions of the *TEX37* and *SVOPL* genes, but the *SVOPL* gene protein encoded by this gene is

thought to be a member of solute carrier family 22, which includes transmembrane proteins that transport toxins and drugs from the body in humans (http://www.genecards.org) (Fishilevich et al., 2018).

Table 2. Genomics regions associated with hot carcass weight in Nelore cattle, percentage of additive genetic variance and candidate genes.

Genomic region	% additive genetic variance explained	Candidate genes
BTA4: 98786241 – 98827345	2.17	*LRGUK*
BTA4: 103077939 – 103103335	1.44	*TRIM24, SVOPL*
BTA11: 47406687 – 47414222	2.71	*TEX37*
BTA19: 1102575 – 1127872	1.73	*CA10*
BTA18: 11675416 – 11729922	1.27	*OXSR1*

Chromosome 3 has the region with the highest peak of additive genetic variance explained for hot carcass weight. However, no gene was identified in this region. Chromosome 4 had the largest number of regions associated with this trait and three genes were identified on that chromosome. Therefore, it may be a relevant chromosome for future studies on hot carcass weight.

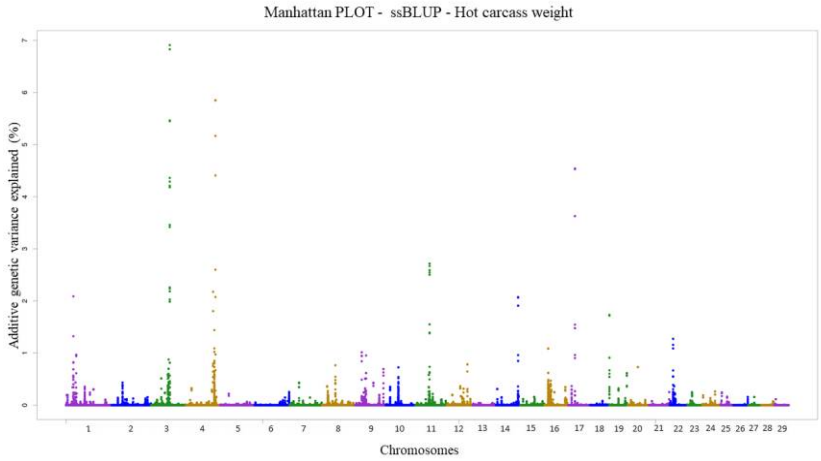

Figure 1. Manhattan plot of effects of SNPs when analyzed in windows of 10 adjacent markers for hot carcass weight.

Ribeye area

For ribeye area, a trait strongly related to muscle growth and development, 14 associated genomic regions were obtained, explaining over 1% of the additive genetic variance. These regions are distributed in 11 different chromosomes, with five genes identified (Table 3). Medeiros et al. (2017) observed associations with genomic regions different from those identified in our study. The genes identified in these regions are described in the Ensembl Genoma Browser, as follows: *Twist family bHLH transcription factor 2; sideroflexin 1; Bos taurus cardiomyopathy-associated protein 5-like; carboxypeptidase Q* and *MRS2, magnesium transporter.*

With respect to genes associated with ribeye area, some studies describe the *TWIST* gene as normally expressed in the skin and in craniofacial cartilage (Franco et al., 2011). Awasthi et al. (2017) reported that increased CNV (Copy Number Variation) in cattle might lead to an ectopic expression of *TWIST2* during neural crest development and thus show the opposite regulatory effect compared to that observed in adult skin. Accordingly, ectopically expressed *TWIST2* might decrease the number of melanoblasts and contribute to an unpigmented skin area of the belt. Another associated gene is *SFXN1*, a member of the sideroflexin family and targeted to the mitochondrial membrane. In sheep, this gene and others of the same family were identified as important in the carrier molecule and related to the regeneration of pancreatic endocrine cells (Xi et al., 2011). The *CMYA5* gene was described as correlated with meat quality in pigs (Xu et al., 2011), due to the association of SNP (A383C) with the traits drip loss and intramuscular fat. *CPQ* is a protein coding gene that may play an important role in the hydrolysis of circulating peptides. Kasvandik et al. (2015), studying cattle sperm reported that this protein appears to be secreted to the cell surface where it hydrolyses C-termini of peptides in the plasma membrane, contributing to membrane modifications during sperm maturation; using immunocytochemistry, they showed that *CPQ* is located mainly in the sperm middle and tail, suggesting its possible role in sperm tail function. The *MRS2* gene, described by Piskacek et al. (2009) as the major transporter of magnesium in humans, was also associated with antibody-mediated immune response in Holstein cows (Thompson-Crispi et al., 2014).

Table 3. Genomics regions associated with ribeye area in Nelore cattle, percentage of additive genetic variance and candidate genes.

Genomic region	% additive genetic variance explained	Candidate genes
BTA3: 118622623 – 118638999	1.66	*TWIST2*
BTA10: 5540505 – 5557246	2.32	*SFXN1*
BTA10: 10683746 – 10695569	2.51	*CMYA5*
BTA14: 69827335 – 69839646	1.97	*CPQ*
BTA23: 33069166 – 33082155	3.22	*MRS2*

Figure 2 shows the genomic regions by chromosome and the amount of additive genetic variance that is explained in percentage by windows of 10 adjacent SNPs.

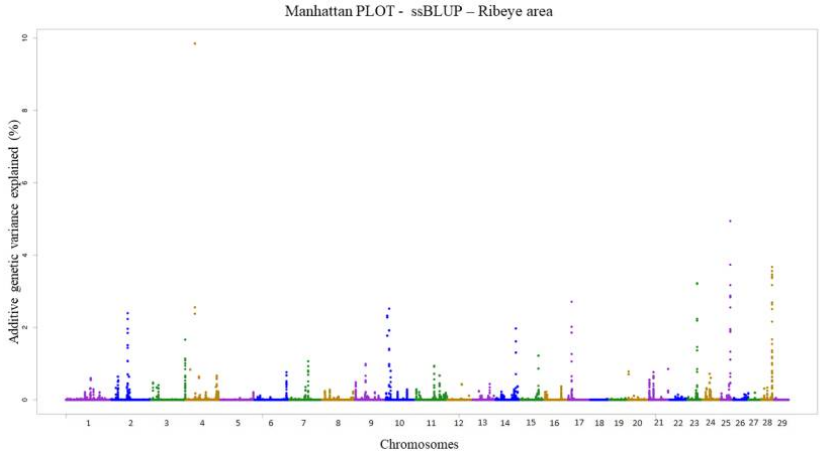

Figure 2. Manhattan plot of effects of SNPs when analyzed in windows of 10 adjacent markers for ribeye area.

In the ssGWAS analysis of ribeye area, several regions in the genome were associated with this trait; although the region of chromosome 4 explained most additive genetic variance; any candidate gene was observed in that region, which demands more studies about this region in the bovine genome.

Backfat thickness

The association of backfat thickness phenotypes, which is a trait associated with deposition and accumulation of lipids in the *Longissimus thoracis* muscle, involved seven associated genomic regions, distributed along three different chromosomes (Table 4). Medeiros et al. (2017) also observed genomic regions associated with backfat thickness, however, on different chromosomes. We identified eight genes, some of them with few information about their functions. The genes are described, according to the Ensembl Genome Browser, as follows: *integrin alpha-9 precursor; five different olfactory receptors, Olfactory receptor; Small nucleolar RNA SNORA69 e 5S ribosomal RNA.*

OR2S2 is an olfactory receptor gene as probably the LOC genes and the family of these genes are the largest in the genome (http://www.genecards.org) (Fishilevich et al., 2018). The same olfactory receptor genes associated with backfat thickness in the present study were also reported by Terakado et al., (2018) associated with weaning weight cattle of the Nelore breed. They suggest that these genes play some role in weight gain from birth to weaning. Consequently, this gene could play a role in body fat deposition along the growth curve of the animals, resulting in differences in the subcutaneous fat deposition, and, therefore, in the precocity of the animals. This is fortuitous for *Bos indicus* breeds, which are considered late finishers when compared to *Bos taurus* and could open a precedent for the selection of animals that express this gene as an alternative of gain in precocity in Zebu breeds. The *SNORA69* gene is responsible for the transcripts of small nucleolar RNAs from a group of untranslated RNA molecules of variable length that are mainly required for rRNA maturation (Dieci et. al., 2009). The *ITGA9* gene that encodes the alpha integrin of membranes (http://www.genecards.org) (Fishilevich et al., 2018), associated with fat thickness in this study, was also reported by Hou et al., (2012) because it is associated with feed efficiency in dairy cattle and found as the primary positional candidate gene in a region with large effects on mid-test metabolic weight QTL (Seabury et al., 2017).

Among the chromosomes, 8 and the 22 had the largest number of associated regions; chromosome 22 had the region with the largest portion of the genetic variance explained. Chromosome 8 showed more genes in the candidate regions, most of them related to olfactory activity.

Table 4. Genomic regions associated with backfat thickness in Nelore cattle, percentage of additive genetic variance and candidate genes.

Genomic region	% additive genetic variance explained	Candidate genes
BTA8: 60631340 – 60663277	1.04	*OR2S2, 5S_rRNA, LOC100299372, LOC523083, LOC532403,*
BTA8: 60673095 – 60694194	1.97	*LOC613441, SNORA69*
BTA22: 11039410 – 11063911	2.54	*ITGA9*

The associated genomic regions are shown in Figure 3, distributed by chromosome, and also how much of the variance is explained in percentage by windows of 10 adjacent SNPs.

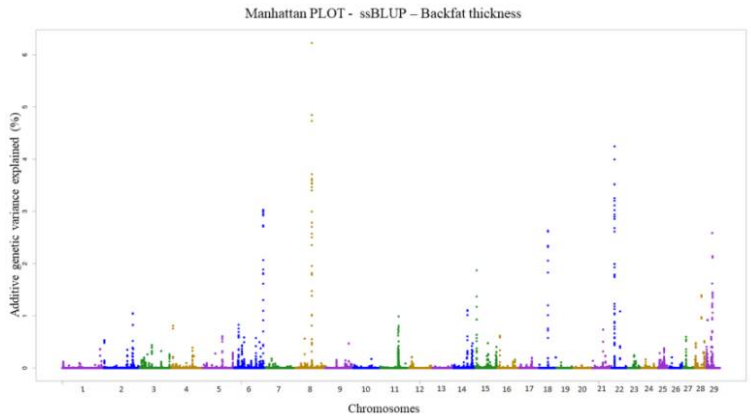

Figure 3. Manhattan plot of the effects of SNPs when analyzed in windows of 10 adjacent markers for backfat thickness.

Marbling

Marbling is a subjective evaluation of percentage of intramuscular fat deposited in adipocytes located between the bundles of the muscle fibers and is associated with meat quality for conferring tenderness, juiciness and flavor. In this study, 17 associated regions were observed, distributed on 11 chromosomes, harboring five candidate genes (Table 5). Magalhaes et al., (2016), studied genomic regions for marbling in Nelore, but the regions with association are in different chromosomes when compared to our study. Descriptions of the genes identified by the Ensembl Genome Browser were: *endomucin; ligand of numb-protein X 1, E3 ubiquitin protein ligase; eukaryotic translation initiation factor 5; Small nucleolar RNA SNORA28* and *desmocollin 3.*

Among the genes associated with additive genetic variance for marbling, *EMCN* is a mucin-like sialo glycoprotein that interferes with the assembly of focal adhesion complexes and inhibits interaction between cells and the extracellular matrix (Kinoshita et al., 2001). Kaneda et al., (2017) comparing the expression of the *EMCN* gene in parthenogenetically activated embryos with bovine embryos obtained by artificial insemination observed that in embryos of parthenogenetic origin it is practically not expressed. The *LNX1* gene encodes a membrane-bound protein that is involved in signal transduction and protein interactions. This protein may play an important role in tumorogenesis due to the location of the *LNX1* gene in the chromosome, which is frequently altered in human gliomas (Blom et al., 2008). The *DSC3* gene encodes calcium-dependent glycoprotein and mutations in this gene are a cause of hypotrichosis and recurrent skin vesicles disorder in humans. This protein can act as an autoantigen in pemphigus diseases, and it is also considered to be a biomarker for some cancers (http://www.genecards.org). The *EIF5* and *SNORA28* genes are in the chromosome 21 region associated with marbling. The *EIF5* gene plays a key role in translation initiation following recognition of the start codon (Paulin et al., 2001); we found no study of association of this gene for cattle. Gene *SNORA28* also codes a small untranslated nucleolar RNA, which is required primarily for rRNA maturation (Dieci et al., 2009).

Table 5. Genomics regions associated with marbling in Nelore cattle, percentage of additive genetic variance and candidate genes.

Genomic region	% additive genetic variance explained	Candidate genes
BTA6: 25609250 - 25621255	1.16	*EMCN*
BTA6: 70773922 – 70782550	7.56	*LNX1*
BTA21: 69639928 – 47414222	1.21	*EIF5, SNORA28*
BTA24: 26435436 – 26477292	1.01	*DSC3*

Chromosome 6 has the largest number of associated regions and the segment between 70773922 – 70782550 base pairs has the largest portion of the genetic variance explained. Even though it is a trait of fat deposition, the peaks of variances explained for marbling are not found for subcutaneous fat thickness. This may be because the deposition mechanisms of subcutaneous fat and intramuscular fat are different.

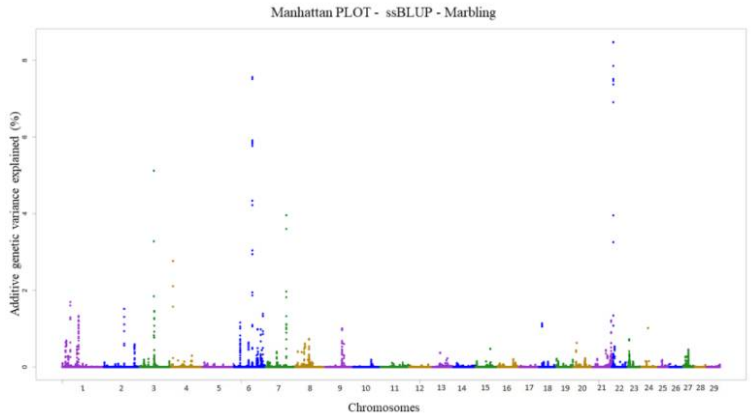

Figure 4. Manhattan Plot of the effects of SNPs when analyzed in windows of 10 adjacent markers for marbling.

Analysis of SNPs

In total, 519 SNPs present in windows associated with ribeye area, backfat thickness, marbling score and hot carcass weight were analyzed for their functional role (Figure 5). Most of them (61%) were located in intergenic regions, which is expected since these regions represent the largest portion of the mammalian genome. SNPs in intergenic regions were classified as “modifiers” meaning that predictions of their impact are difficult or there is no evidence of impact (http://www.ensembl.org/Help/Glossary?id=535).

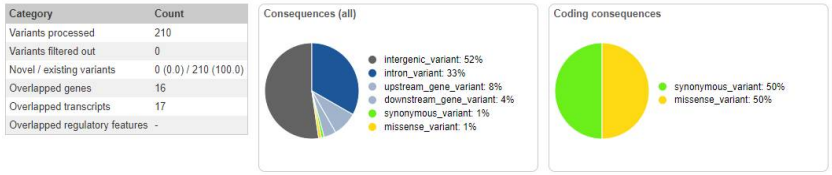

Figure 5. Summary of functional consequences of SNPs

Nevertheless, the important role of non-coding polymorphisms has been demonstrated, especially the ones surrounding or within coding genes (Zhang and Lupski, 2015). In our study, 26% of the SNPs were in intronic regions; 6% and 5% were in downstream and upstream coding genes, respectively; 3% were at 3' UTR regions. Intronic SNPs are interesting targets for study due to their many possible functions. Although not in a coding region, intronic polymorphisms harbor a variety of functional elements such as splice enhancers and silencers, trans-splicing elements and other regulatory elements that can influence the expression of the genes that host them and modulate the genotype-phenotype relationship (Cooper, 2010). More detailed results can be seen in Supplementary File S3.

Twenty-five SNPs were situated in an upstream position of genes *DBF4*, ENSBTAG00000032671, ENSBTAG00000038086, ENSBTAG00000038308, ENSBTAG00000047099, *LNX1*, *MRS2*, RF00001, RF00265, RF00400, *TEX37*, *TWIST2*. Around 80 bp upstream to a gene start site, there is a core promoter region where transcription factors (TF) bind to regulate gene expression (Barrett et al., 2012). Polymorphisms in these regions can alter the TF biding site and therefore change the ability for certain TFs to regulate gene expression (Alexandre et al., 2014). Likewise, other promoter elements such as enhancers and inhibitors can be found both upstream and downstream (Barrett et al., 2012). We identified 31 downstream variants to genes *DSC3*, *EIF5*, ENSBTAG00000015238, ENSBTAG00000016282, ENSBTAG00000031265, ENSBTAG00000032670, ENSBTAG00000038086, ENSBTAG00000043854, ENSBTAG00000047099, *OR2S2*, RF00265, RF00400, *SFXN1*, *SVOPL*, *TEX37*, *TRIM24* that could have important regulatory roles. We were also able to identify three SNPs (rs135522467, rs133489106 and rs134568613) in the 3' UTR region of *EIF5* gene, which although not affecting the translated protein itself, could have an important role in post transcriptional regulation, affecting mRNA stability, translation efficiency, polyadenylation and miRNA sites (Barrett et al., 2012). Indeed, polymorphisms in 3' UTR regions of some genes have been associated with carcass traits (Wang et al., 2013).

Finally, we found 1% of synonymous and missense variants. Synonymous variants are those that despite being in a coding region, do not change the translated amino acid but are able to cause changes in protein expression, conformation and/or function, ultimately influencing the phenotype (Sauna and Kimchi-Sarfaty, 2011). The four synonymous variants (rs109519676, rs110545249, rs110545249 and rs42672622) were located in exons 7, 5 and 1 of genes *CMYA5*, ENSBTAG00000003408 and ENSBTAG00000038086, respectively. The other four missense variants (rs133785882, rs109199928, rs134893266 and rs136525721) were located in exons 3, 15, 1 and 1 of genes *TEX37*, *TRIM24*, ENSBTAG00000038308, ENSBTAG00000039443, respectively. They cause a change in the amino acid sequence of the translated protein in a tolerated fashion according to SIFT score (Ng and Henikoff, 2003) but are classified as of "moderate" impact, meaning that they are non-disruptive variants that can change protein effectiveness.

Even though it is unclear at this point whether these variants present a direct functional effect or that they are simply in linkage disequilibrium with another functional SNP (McCauley et al., 2007), the genomic location of the SNPs in significant windows associated with carcass traits help us to prioritize variants as potential markers for animal selection.

Comparation of QTLs deposited in CattleQTLdb with regions associated with carcass quality traits

For all candidate regions that were associated with meat and carcass quality traits, QTLs were prospected in the CattleQTLdb database using the tool *Browse Cattle Trait Hierarchy for QTLdb* (Hu et al., 2016). Among the seven regions associated in the study with backfat thickness, no QTL in the same regions was observed in the database.

For ribeye area, two QTLs prospects in the database also were observed on chromosome 4, which are contained in associated regions of in our study. The two QTLs are between 32.2 - 32.4 Mbp (Yokouchi et al., 2009) and 28.3 - 38.0 Mbp (Mizoshita et al., 2004). Although these two QTLs overlap, the peaks that were observed are at different sites. In our study, no gene for ribeye area was identified in this region of chromosome 4.

In the candidate regions for marbling in our study, 4 QTLs were observed in the database in candidate regions. On chromosome 1 between 25.1 - 48.4 Mbp (McClure et al., 2012), chromosome 2 between 77.8 - 91.5 Mbp (MacNeil and Grosz, 2002), chromosome 3 between 22.5 - 58.9 Mbp (Casas et al., 2001) and chromosome 6 between 24.9 - 36.0 Mbp (McClure et al., 2012). This last QTL of chromosome 6 showed that the EMCN gene present in this region was also associated in the study on marbling. For candidate regions for hot carcass weight, three QTLs were observed in the database on chromosome 16 between 8.4 - 34.2 Mbp, (Casas et al., 2003), chromosome 17 between 22.6 - 36.6 Mbp and chromosome 22 between 11.7 - 22.1 Mbp (McClure et al., 2012). No gene was identified in searches among these regions of QTLs.

CONCLUSIONS

Several genomic regions associated with small effects demonstrate the complexity of the phenotypes related to carcass quality; they do not depend only on the effects of a few genes to determine their variations. The various genes observed in the candidate regions, though there is little description of the functions that they perform, demonstrate the need for further studies that would reveal the functions of these genes and their involvement in meat and carcass quality traits in cattle. Further studies on these traits will shed light on how they are controlled. The regions associated in our study, which are in QTL regions that have already been described, point to their importance for determining meat and carcass quality traits, which should be the target of further studies.

ACKNOWLEDGMENTS

The authors thank the University of São Paulo for structural and physical support for the study, and FAPESP (Processes: 2013/13021-6 and 2014/07566-2), CNPq (Processes: 150224/2017-6 and 478132/2013-1), and FAPEMIG (Process: APQ-01804-13) for funding.

REFERENCES

Aguilar I, Misztal I, Johnson DL, Legarra A, et al. (2010). Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93: 743–752.

Aguilar I, Misztal I, Legarra A and Tsuruta S (2011). Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. *J. Anim. Breed Genet*. 128: 422-428.

Alexandre PA, Gomes RC, Santana MHA, Silva SL, et al. (2014). Bovine NR1I3 gene polymorphisms and its association with feed efficiency traits in Nellore cattle. *Meta Gene*. 2: 206–217.

Aspatwar A, Tolvanen ME, Ortutay C and Parkkila S (2014). Carbonic anhydrase related proteins: molecular biology and evolution. *Subcell. Biochem*. 75: 135–156.

Awasthi Mishra N, Drögemüller C, Jagannathan V, Keller I, et al. (2017). A structural variant in the 5'-flanking region of the TWIST2 gene affects melanocyte development in belted cattle. *PLoS ONE*. 12: e0180170.

Barrett LW, Fletcher S and Wilton SD (2012). Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cell. Mol. Life Sci*. 69: 3613-3634.

Bernard C, Cassar-Malek I, Le Cunff M, Dubroeucq H, et al. (2007). New indicators of beef sensory quality revealed by expression of specific genes. *J. Agric. Food Chem.* 55: 5229–5237.

Blom T, Roselli A, Tanner M and Nupponen MM (2008). Mutation and copy number analysis of LNX1 and Numbl in nervous system tumors. *Cancer Genet. Cytogenet*. 186: 103-109.

Carvalho ME, Baldi FS, Santana MHA, Ventura RV, et al. (2017). Identification of genomic regions related to tenderness in Nellore beef cattle. *Adv. Anim. Biosci*. 8: 42–44.

Casas E, Shackelford SD, Keele JW, Koohmaraie M, et al. (2003). Detection of quantitative trait loci for growth and carcass composition in cattle. *J. Anim. Sci.* 81: 2976–2983.

Casas E, Stone RT, Keele JW, Shackelford SD, et al. (2001). A comprehensive search for quantitative trait loci affecting growth and carcass composition of cattle segregating alternative forms of the myostatin gene. *J. Anim. Sci*. 79: 854-860.

Cooper DN (2010). Functional intronic polymorphisms: Buried treasure awaiting discovery within our genes. *Hum Genomics*. 4: 284–288.

Corva P, Soria L, Schor A, Villarreal E, et al. (2007). Association of CAPN1 and CAST gene polymorphisms with meat tenderness in *Bos taurus* beef cattle from Argentina. *Genet. Mol. Biol*. 30: 1064 -1069.

Zerbino DR, Achuthan P, Akanni W, Amode MR, et al. Ensembl 2018, Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D754–D761. Accessed: 14 September 2018.

Database resources of the National Center for Biotechnology Information (2014). Nucleic Acids Res. 43: D6-17.

Dieci G, Preti M and Montanini B. (2009). Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* 94: 83-88.

Fishilevich S, Nudel R, Rappaport N, Hadar R, et al. GeneHancer: genome - wide integration of enhancers and target genes in GeneCards , Database(Oxford) (2017). Accessed: 2 November 2018.

Franco HL, Casasnovas J, Rodríguez-Medina JR and Cadilla CL (2011). Redundant or separate entities?—roles of Twist1 and Twist2 as molecular switches during gene transcription. *Nucleic Acids Res.* 39: 1177–1186.

Garcia JF, Carmo AS, Utsunomiya YT, Neves HHR, et al. (2012). How Bioinformatics Enables Livestock Applied Sciences in the Genomic Era. In: Advances in Bioinformatics and Computational Biology. BSB 2012. Lecture Notes in Computer Science (Souto MCP, Kann MG, eds). Springer-Verlag Berlin, Heidelberg.

Hayes BJ, Bowman PJ, Chamberlain AC, Verbyla K, et al. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet Sel Evol.* 41: 51 Available at [https://doi.org/10.1186/1297-9686-41-51].

Hocquette JF, Lehnert S, Barendse W, Cassar-Malek I, et al. (2007). Recent advances in cattle functional genomics and their application to beef quality. *Animal*. 1: 159-173.

Hou Y, Bickhart DM, Chung H, Hutchison JL, et al. (2012). Analysis of copy number variations in Holstein cows identify potential mechanisms contributing to differences in residual feed intake. *Funct. Integr. Genomics*. 12: 717–723.

Hu ZL, Park CA and Reecy JM (2016). Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Research.* 44: D827–D833.

Huaiyu M, Qing D, Anushya M, Pascale G, et al. (2010). PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. *Nucl. Acids Res.* 38: D204-D210.

Kaneda M, Takahashi M, Yamanaka KI, Saito K, et al. (2017). Epigenetic analysis of bovine parthenogenetic embryonic fibroblasts. *J. Reprod. Dev.* 63: 365-375.

Kasvandik S, Sillaste G, Velthut-Meikas A, Mikelsaar AV, et al. (2015). Bovine sperm plasma membrane proteomics through biotinylation and subcellular enrichment. *Proteomics*. 15: 1906-1920.

Kinoshita M, Nakamura T, Masafumi I, Haraguchi T, et al. (2001), Identification of human endomucin-1 and -2 as membrane-bound *O*-sialoglycoproteins with anti-adhesive activity1. *FEBS Letters*. 499: 121-126.

Liu Y, DeBoer K, de Kretser DM, O'Donnell L, et al. (2015). LRGUK-1 Is Required for Basal Body and Manchette Function during Spermatogenesis and Male Fertility. *PLoS Genet.* 11(3): e1005090.

Luchiari Filho A (2000). Pecuária da Carne Bovina. 1st edn. Albino Luchiari Filho, São Paulo.

McCauley JL, Kenealy SJ, Margulies EH, Schnetz-Boutaud N, et al. (2007). SNPs in Multi-Species Conserved Sequences (MCS) as useful markers in association studies: a practical approach. *BMC Genomics*. 8: 266.

MacNeil MD and Grosz MD (2002). Genome-wide scans for QTL affecting carcass traits in Hereford x composite double backcross populations. *J. Anim. Sci*. 80: 2316-2324.

Magalhães AF, de Camargo GM, Fernandes GA, Gordo DG, et al. (2016). Genome-Wide Association Study of Meat Quality Traits in Nellore Cattle. *PLoS One*.11: e0157845.

Mazzoni G, Salleh SM, Freude K, Pedersen HS, et al. (2017). Identification of potential biomarkers in donor cows for in vitro embryo production by granulosa cell transcriptomics. *PLoS ONE*. 12(4): e0175464.

Medeiros de OSR, Bonvino SN de O, Fragomeni B, Miguel F de CG, et al. (2017) Genome-Wide Association Study for Carcass Traits in an Experimental Nelore Cattle Population. *PLoS One*. 12: e0169860.

Misztal I, Tsuruta S, Strabel T, Auvray B, et al. (2002). BLUPF90 and related programs (BGF90). Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France.

Mizoshita K, Watanabe T, Hayashi H, Kubota C, et al. (2004). Quantitative trait loci analysis for growth and carcass traits in a half-sib family of purebred Japanese Black (Wagyu) cattle. *J. Anim. Sci*. 82: 3415-3420.

McClure MC, Ramey HR, Rolf MM, McKay SD, et al. (2012). Genome-wide association analysis for quantitative trait loci influencing Warner–Bratzler shear force in five taurine cattle breeds. *Anim. Genet*. 43: 662-673.

Ng PC and Henikoff S (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 31: 3812-3814.

Olerup O and Zetterquist H (1992). HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. *Tissue Antigens*. 39: 225-235.

Orozco G, Barrett JC and Zeggini E (2010). Synthetic associations in the context of genome-wide association scan signals. *Hum. Mol. Genet.* 19: 137-144.

Paulin FE, Campbell LE, O'Brien K, Loughlin J, et al. (2001). Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. *Curr. Biol*. 11: 55–59.

Piskacek M, Zotova L, Zsurka G and Schweyen RJ (2009). Conditional knockdown of hMRS2 results in loss of mitochondrial Mg(2+) uptake and cell death. *J. Cell. Mol. Med*. 13: 693-700.

Ponsuksili S, Murani E, Trakooljul N, Schwerin M, et al. (2014). Discovery of Candidate Genes for Muscle Traits Based on GWAS Supported by eQTL-analysis. *Int. J. Biol. Sci*. 10: 327–337.

Sauna ZE and Kimchi-Sarfaty C (2011). Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet*. 12: 683-691.

Seabury CM, Oldeschulte DL, Saatchi M, Beever JE, et al. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* 18*:* 386.

Terakado APN, Costa, RB, De Camargo GMF, Irano N, et al. (2018). Genome-wide association study for growth traits in Nelore cattle. *Animal* 12: 1358-1362.

Thompson-Crispi KA, Sargolzaei M, Ventura R, Abo-Ismail M, et al. (2014). A genome-wide association study of immune response traits in Canadian Holstein cattle. *BMC Genomics* 15: 559. Available at [https://doi.org/10.1186/1471-2164-15-559].

Tizioto PC, Decker JE, Taylor JF, Schnabel RD, et al. (2013). Genome scan for meat quality traits in Nelore beef cattle. *Physiol. Genomics*. 45: 1012-1020.

USDA (1999). Official United States Standards for Grades of Carcass Beef. Washington, D.C.: Agriculture Marketing Service.

VanRaden PM (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci*. 91: 4414–4423.

Vitezica ZG, Aguilar I, Misztal I and Legarra A (2011). Bias in genomic predictions for populations under selection. *Genet. Res*. 93: 357-366.

Xi D, He Y, Sun Y, Gou X, et al. (2011) Molecular cloning, sequence identification and tissue expression profile of three novel genes Sfxn1,Snai2 and Cno from Black-boned sheep (Ovis aries). *Mol. Biol. Rep*. 38: 1883–1887

Xu X, Yin Q, Sun L, Liu B, et al. (2011). The molecular characterization and associations of porcine cardiomyopathy associated 5 (CMYA5) gene with carcass trait and meat quality. *Mol. Biol. Rep*. 38: 2085–2090.

Wang H, Misztal I, Aguilar I, Legarra A, et al. (2012). Genome wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* 94: 73–83.

Wang X, Li T, Zhao HB and Khatib H (2013). Short communication: A mutation in the 3′ untranslated region diminishes microRNA binding and alters expression of the OLR1 gene. *J. Dairy Sci*. 96: 6525-6528.

Yokouchi K, Mizoguchi Y, Watanabe T, Iwamoto E, et al. (2009). Identification of a 3.7-Mb region for a marbling QTL on bovine chromosome 4 by identical-by-descent and association analysis. *Anim. Genet*. 40: 945–951.

Zhang F and Lupski JR (2015). Non-coding genetic variants in human disease. *Hum. Mol. Genet*. 24: 102-110.