Genomics 112 (2020) 873–879

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Gene expression profiling and identification of hub genes in Nellore cattle with different marbling score levels

Larissa Fernanda Simielli Fonseca[a], Danielly Beraldo dos Santos Silva[a], Daniele Fernanda Jovino Gimenez[a], Fernando Baldi[a,c], Jesus Aparecido Ferro[a,c], Luis Artur Loyola Chardulo[c,d], Lucia Galvão de Albuquerque[a,b,*]

[a] São Paulo State University (Unesp), School of Agricultural and Veterinarian Sciences, Jaboticabal, SP, Brazil
[b] National Council for Scientific and Technological Development (CNPq), Brasilia, DF, Brazil
[c] CNPq Brasilia, DF, Brazil
[d] São Paulo State University (Unesp), Faculty of Veterinary Medicine and Animal Science, Botucatu, SP, Brazil

ARTICLE INFO

Keywords:
Intramuscular fat
Meat quality
RNA-Seq
Transcriptome

ABSTRACT

The marbling rate evaluation is difficult and expensive, requiring slaughter of the animal or ultrasound measurement. Thus, this trait is generally not included in animal breeding programs. The use of molecular techniques to elucidate intramuscular fat deposition may help improve this trait. In this respect, transcriptome studies and differential gene expression analysis by RNA-Seq can contribute to advances in this area. The objective of this study was to use RNA-Seq to identify differentially expressed genes (DEGs) in muscle tissue (*longissimus thoracis*) of Nellore cattle divergently ranked on marbling, in order to increase our understanding of genes involved in the expression of this trait. The results revealed 49 DEGs and three hub genes (CISH, UFM1, TSHZ1), all of them involved in insulin and diabetes mellitus metabolism. These results indicating key genes and pathways, which may help to develop strategies designed to select animals with greater marbling.

1. Background

The Nellore breed (*Bos indicus*) is known for its low amount of intramuscular fat when compared to *Bos taurus* animals. This type of fat, called marbling, is associated with sensory traits such as color, flavor and aroma. In addition, marbling, color and exudation are related to the appearance of meat [1].

Despite the importance of marbling for [1], this trait is generally not used as a selection criterion in breeding programs because of its late expression in the animal and the complexity of obtaining phenotypic data, which requires slaughter of the animal or ultrasound measurement, a method that has been found to be little efficient in pasture-raised Nellore cattle [2]. As an alternative, molecular biology techniques can be used to enhance our understanding of the biological processes related to meat quality and to help improve traits such as marbling. Large-scale RNA sequencing (RNA-Seq) can contribute to advances in this area since this tool permits the researcher to analyze transcriptome profiles, to identify differentially expressed genes, and to discover new genes and isoforms [3,4].

Previous studies using RNA-Seq have demonstrated differences in the expression of genes involved in oxidative metabolism and cell adhesion between muscle tissue and marbling fat in Hanwoo cattle (*Bos taurus coreanae*) [5]. Chen et al. [6], studying *longissimus thoracis* muscle in crossbred cattle (Fuzhou Yellow × Limousine × Wagyu) with high and low marbling levels, found a large number of genes with known and unknown functions that may help understand the mechanisms underlining the fat deposition in cattle. In castrated Nellore cattle, differentially expressed genes related to retinoic acid, IGF2 and ANKRD26 were identified in *longissimus thoracis* muscle. These molecules are important regulators of the mechanisms responsible for intramuscular fat content analyzed by chemical analysis (total lipids) [7].

Marbling is a polygenic trait regulated by various genes, which are directly or indirectly involved in adipogenesis and fatty acid metabolism [8]. However, despite the studies cited above, the biochemical mechanisms involved in intramuscular fat deposition have not been fully unraveled. In this respect, the identification of differentially

* Corresponding author at: São Paulo State University (Unesp), School of Agricultural and Veterinarian Sciences, Jaboticabal, SP; National Council for Scientific and Technological Development (CNPq), Brasilia, DF, Brazil.
E-mail addresses: lsimielli.fonseca@gmail.com (L.F.S. Fonseca), daniellyberaldo@gmail.com (D.B. dos Santos Silva), dani.jovino@yahoo.com.br (D.F.J. Gimenez), fernandobaldiuy@gmail.com (F. Baldi), jesusferro@unesp.br (J.A. Ferro), chardulo@fmvz.unesp.br (L.A.L. Chardulo), galvao.albuquerque@unesp.br (L.G. de Albuquerque).

https://doi.org/10.1016/j.ygeno.2019.06.001
Received 15 February 2019; Received in revised form 16 May 2019; Accepted 1 June 2019
Available online 03 June 2019
0888-7543/

expressed genes for marbling may contribute to the understanding of the mechanisms involved in the expression of this trait and the detection of potential genes that could be used as genetic markers in Nellore cattle to permit the early selection for marbling.

The objective of the present study was to identify differentially expressed and hub genes in muscle tissue (*longissimus thoracis*) of Nellore cattle divergently ranked on marbling by high-throughput RNA-Seq, providing data to increase our understanding of the genetic mechanisms related to this trait.

2. Methods

2.1. Animals and sample collection

Eighty non-castrated male animals with a mean age of 24 months from the Capivara Farm, which participates in the Qualitas Nellore breeding program, were studied. The animals belonged to the same contemporary group (animals of the same sex, born in the same season and maintained together until slaughter) and were slaughtered in a commercial slaughterhouse, on the same day under commercial usual process in accordance with guidelines for the Humane Slaughter of Cattle. During slaughter, the carcasses were properly identified and weighed, and *longissimus thoracis* muscle samples were collected using scissors and tweezers sterilize (RNAse free), and stored in falcon tube with RNAlater® solution (Ambion, ThermoFisher Scientific, USA) in ice dry during the transport until the lab, where it was stored in a freezer at −80 °C. Next, the carcasses were cooled for 24 h, and a *longissimus thoracis* muscle sample, was removed during deboning, between the 12th and 13th rib of each left half-carcass, for determination of the beef marbling score. None of them was aged.

2.2. Determination of marbling score

The marbling was rated in beefs colleted beetween 12th and 13th rib of each left half-carcass, and with the aid of USDA marbling score cards [9], all the 80 samples are analyzed and a score category was assigned to each of them. The categories were (1 = devoid, 2 = practically devoid, 3 = traces, 4 = slight, 5 = small, 6 = modest, 7 = moderate, 8 = slightly abundant and 9 = moderately abundant). However, in Brazil, this scale ranges from zero (devoid) to six (moderate), because of the low degree of marbling in the Brazilian herd.

Based on these results, 20 samples were selected from animals with phenotypically extreme degrees of marbling,10 with a high-marbling grade (4.80 ± 0.06) and 10 with a low-marbling grade (2.32 ± 0.07), for RNA-Seq analysis. The Student *t*-test was performed in the R environment and a significant difference was observed between the high and low marbling groups ($p < .05$).

2.3. Total RNA extraction

Total RNA was extracted from 50 mg muscle tissue, adipose tissue free, i.e., the sample used to extracts total RNA were totally separately from adipose tissue before the starting the analysis. The extraction was performed using the RNeasy Lipid Tissue Mini Kit (Qiagen, Valencia, CA, USA), according to manufacturer recommendations.

The purity of each RNA sample regarding the presence of protein or phenol was checked in a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Santa Clara, CA, USA, 2007). The RNA concentration and contamination with genomic DNA were determined in a Qubit® 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA, 2010), while the quality of the extracted RNA was evaluated in an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA, 2009).

2.4. RNA sequencing

RNA-Seq analysis of the samples were performed as follows: first, the mRNA from each sample was purified from total RNA using magnetic beads conjugated with oligo-dT. After purification, the mRNA were converted to double stranded cDNA by reverse transcriptase and DNA polymerase and then fragmented randomly to generated fragments with medium size of about 200 bp (sequencing library). The cohesive ends of the cDNA double strands were converted into blunt ends and an adenine base was added at each 3′-end of the fragments. Adapters were attached to the cDNA fragments to permit identification of the samples. The sequencing libraries were quantified by quantitative PCR and quality was validated in an Agilent 2100 Bioanalyzer®. Finally, the libraries were submitted to paired-end sequencing in an Illumina HiSeq 2500. The TruSeq PE Cluster Kit v3-cBot-HS was used to bind the libraries to complementary adapter oligos on flow cells, which were then amplified, providing clonal clusters ready to be sequenced. The TruSeq SBS Kit v3-HS (200 cycles) was used to determine the sequence of clusters according to manufacturer recommendations. Each molecule was sequenced at both ends (paired-end sequencing) and 100-bp fragments were obtained from each end (2 × 100).

2.5. Processing and alignment of the sequences

The computational analyses were performed, by individual sample, as follow: first, the quality of the sequences generated was evaluated using the Fastqc v 0.10.1 program, and low-quality reads were filtered with Trimmomatic v.0.36 [10]. The reads were mapped with the HISAT2 v.2.0.5 [11] using the bovine genome (UMD3.1 *Bos Taurus*) and chromosome Y (Btau 4.6.1), as a reference, both deposited in Ensembl [12]. Descriptive statistics for alignment rate were presented in Table 1.

The expression of the following housekeeping genes was analyzed to evaluate the quality of sequencing: beta glucuronidase (GUSB), erythrocyte hydroxymethylbilane synthase (HMBS), hypoxanthine phosphoribosyl-transferase 1 (HPRT1), phosphoglycerate kinase 1 (PGK1), and TATA box binding protein (TBP). The expression pattern of these genes was similar in the experimental groups (high- and low-marbling grade) (Additional File 1: Fig. S1).

The data were explored and visualized using the CummeRbund package [13], implemented in the R environment. A box plot was generated for each group studied using the transformed $\log_{10}$ (FPKM) values (Additional File 2: Fig. S2). The distribution of quartiles was consistent between groups, indicating the high quality of the data. The medians were similar in the groups and close to −1, demonstrating that it was possible to identify low-expressed genes ().

A plot of the principal components (PCA) revealed the formation of different groups (high- and low-marbled beef), indicating differences in the expression of genes between the groups with high and low-marbling scores (Addition File 3: Fig. S3).

Table 1
Descriptive statistics for alignment rate (means for all samples separated by low or high marbling group).

Animal group	Number of Samples	Raw reads (Mb)	Trimmed reads (Mb)	Overall alignment rate (%)	Uniquely mapped (%)	Coverage
High marbling	10	29,272,698	26,363,068	96,67	87	45×
Low marbling	10	20,550,334	17,534,936	96,64		
Mean	20	24,911,516	21,976,535	96.65		

2.6. Assembly and quantification of transcripts

The Cufflinks2 v2.1.1 program [7] was used to assemble the aligned fragments of each sample and to estimate the number of reads in fragments per kilobase of transcript per million fragments mapped (FPKM). This program normalizes the expression of transcripts to the length and total number of fragments sequenced per sample. The output file generated with Cufflinks2 for each sample contains the new transcripts, genes and isoforms found. The result of Cufflinks2, per sample, was concatenated in a single file using the Cuffmerge2 v2.1.1 program. This single file was used as a reference in the analysis of differential gene expression with the Cuffdiff2 program [14] as described below.

2.7. Analysis of differential gene expression

The alignment (.bam) files generated were divided into two contrasting groups for beef marbling score. The FPKM values of the transcripts were compared in each sample using the Cuffdiff2 v2.1.1 program [16]. This program uses a *t*-test to calculate *p*-values. The Benjamini-Hochberg method was used to control the false discovery rate (FDR), adopting a q-value for FDR of $< 5\%$.

2.8. Co-expression network analysis

Cytoscape v.3.4 and Expression Correlation Network plugin [15] were used to generate a co-expression network using count reads from HISAT2 v.2.0.5 program. The expression profiles of differentially expressed genes were used to compute similarity matrix by Pearson correlation coefficient. A histogram was generated to choose a similarity strength threshold (cut-off of > 0.7 and < -0.7) with NetworkAnalyzer app [16]. Cytohubba app [17] were used to determine the degree of distribution and betweenness centrality, employed as a global metric to identify the top three hub genes.

2.9. Gene set enrichment analysis

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.7 [18], was used to annotate and interpret the lists of differentially expressed genes. DAVID pathway mapping was used to identify metabolic pathways in which the differentially expressed genes are involved.

3. Results

3.1. Differentially expressed genes

The analysis performed revealed 49 annotated differentially expressed genes, including 27 up-regulated and 22 down-regulated genes in relation to the group with a low-marbling score (q-value < 0.05) (Table 2). The $\log_2$ fold change was used to characterize down-regulated or up-regulated genes in relation to the low-marbling group.

The genes Pterin-4 alpha-carbinolamine dehydratase, Solute Carrier Family 16 Member 7 and Nuclear Receptor Subfamily 4 Group A Member 2 (PCBD1, SLC16A7 and NR4A2 respectively), and Solute Carrier Family 6 Member 2, Carboxypeptidase E and Beta-globin (SLC6A2, CPE and HBB respectively) were the top three genes with smaller and larger fold change related to low-marbling group, respectively.

3.2. Co-expression network analysis

The co-expression network analysis (Fig. 1) revealed 48 genes (nodes) and 299 significantly correlated gene pairs (edges). The density and the clustering coefficient were 0.041 and 0.309, respectively. The presence of hub genes were indicate when scale-free topology were analyzed ($R2 = 0.723$). The genes Cytokine Inducible SH2 Containing Protein (CISH), Ubiquitin Fold Modifier 1 (UFM1), Teashirt Zinc Finger Homeobox 1 (TSHZ1) were evidenced as hubs. The genes CISH and UFM1 were more expressed in high-marbling group, while TSHZ1 was more expressed in low-marbling group.

3.3. Gene set enrichment analysis

Biological process Gene Ontology (GO) terms were found by DAVID v6.7 database, to generate a functional annotation using all differentially expressed (up- and down-regulated) genes for meat tenderness, employing *Bos taurus* as reference. This analysis permitted the identification of 21 functional groups (annotation clusters; Additional file 4: Table S1). The genes Pellino E3 ubiquitin protein ligase, Ubiquitin fold modifier 1, Nuclear receptor subfamily 4 group A member 2, Peptidyl arginine deiminase 2 (PELI1, UFM1, NR4A2, PADI2, respectively) are related to terms response to lipid (GO:0033993) and cellular response to lipid (GO:0071396). Other GO terms were also found as response to abiotic stimulus (GO:0009628), response to stimulus (GO:0050896), apparently related to intramuscular fat deposition.

The differently expressed genes Carboxypeptidase E (CPE), MHC class II antigen (BLA-DQB), MHC Class I JSP.1 (JSP.1) and heat shock protein family A (Hsp70) member 6 (HSPA6) were involved in five pathways: type I diabetes mellitus; antigen processing and presentation; graft-versus-host disease; allograft rejection; autoimmune thyroid disease.

4. Discussion

The genes Pterin-4 alpha-carbinolamine dehydratase 1 (PCBD1), solute carrier family 16 member 7 (SLC16A7) and nuclear receptor subfamily 4 group A member 2 (NR4A2) were the top three genes that expressed more in high-marbling group. Li et al. [19] reported an association between PCBD1 gene and subcutaneous fat thickness in a genome-wide association study (GWAS) of crossbred Angus x Charolais cattle. The SLC16A7 gene is a member of Monocarboxylate transporter family and, has a broad expression in fat [20]. Berton et al. [21] studying fatty acids in Nellore cattle muscle tissue, found the SLC16A7 was upregulated for saturated fatty acids, myristic and stearic acids, and downregulated for omega 3. The NR4A2 are related to GO terms response to lipid (GO:0033993) and cellular response to lipid (GO:0071396). These terms refer to adipogenesis and adipose cell differentiation and, cellular response to triglyceride and cellular response to fatty acid, respectively.

Other three differentially expressed genes, pellino E3 ubiquitin protein ligase 1 (PELI1), Ubiquitin fold modifier 1 (UFM1) and peptidyl arginine deiminase 2 (PADI2), are related to terms response to lipid and cellular response to lipid, thus these genes act by transforming relatively unspecialized cell in a cell specialized for the synthesis and storage of fat, as an adipocyte. The response to lipid stimulus refer to any process that modifies the activity or state of a cell. The process could be a movement, secretion, enzyme production or gene expression, between others [22].

The genes beta-hemoglobin (HBB), solute carrier family 6 member 2 (SLC6A2) and carboxypeptidase E (CPE) were the top three genes that expressed more in low-marbling group. Analysis in David pathways revealed that all these genes have the function related to metal-binding (GO:0046872). The HBB gene are associated with heme metabolism, the main component of hemoglobin [23,24]. Hemoglobin rates may be associated with the amount of fat in the organism. This association has been observed in humans, in which normal weight subjects had higher concentrations of hemoglobin, while overweight subjects exhibited low levels of this molecule in blood [25]. The SLC6A2 gene, in a study with human, presented SNP association with total fat and monounsaturated fatty acids intake.

The CPE gene is responsible for insulin synthesis. Insulin resistances hinders the absorption of glucose from the blood, leading to diabetes. In

Table 2
Genes differentially expressed in samples with divergent marbling scores.

Gene Symbol	Formal gene name	Locus	High-marbled	Low-marbling	log2 (fold_change)[a]	*p*-Value	q_value
PCBD1	Pterin-4 alpha-carbinolamine dehydratase	28:27226793-27231724	595.951	0.318655	−422.513	0.00015	0.0366511
SLC16A7	Solute Carrier Family 16 Member 7	5:53987488-54214813	538.797	281.337	−238.449	5,00E−05	0.0137833
NR4A2	Nuclear Receptor Subfamily 4 Group A Member 2	2:39999673-40017692	651.644	18.709	−180.035	5,00E−05	0.0137833
BDH1	3-hydroxybutyrate dehydrogenase, type 1	1:72571510-72608859	489.623	16.591	−176.065	5,00E−05	0.0137833
HSPA6	heat shock protein family A (Hsp70) member 6	3:8027503-8030023	329.198	113.482	−153.649	5,00E−05	0.0137833
BLA-DQB	MHC class II antigen	23:25855144-25863045	968.061	378.306	−135.554	5,00E−05	0.0137833
METTL21C	methyltransferase like 21C	12:82906942-82918374	403.817	103.302	−135.509	5,00E−05	0.0137833
MICAL2	microtubule associated monooxygenase, calponin and LIM domain containing 2	15:40914779-41195753	997.947	242.195	−127.913	5,00E−05	0.0137833
LOC782631	60S ribosomal protein L9 pseudogene	13:59957495-59958191	235.817	105.141	−116.535	5,00E−05	0.0137833
LOC100847340	LOC100847340	20:65961543-65966689	699.273	324.921	−110.577	0.0002	0.0438816
NREP	neuronal regeneration related protein	10:2161054-2193665	636.887	100.258	−0.654612	0.0001	0.0255976
IER5	immediate early response 5	16:63689429-63691626	645.224	407.591	−0.662679	0.0001	0.0255976
CISH	Cytokine Inducible SH2 Containing Protein	22:50320204-50325618	698.641	11.49	−0.717753	0.0002	0.0438816
NIPSNAP1	Nipsnap Homolog 1	17:70869033-70884644	938.073	553.279	−0.761691	5,00E−05	0.0137833
OTUD1	OTU deubiquitinase 1	13:24654857-24657894	343.224	202.102	−0.764069	5,00E−05	0.0137833
ECHDC3	enoyl-CoA hydratase domain containing 3	13:12552171-12591181	537.079	31.128	−0.786922	5,00E−05	0.0137833
ELTD1	adhesion G protein-coupled receptor L4	3:65921130-66062041	735.223	130.005	−0.822319	0.0001	0.0255976
LOC101905399	dual specificity mitogen-activated protein kinase kinase 1 pseudogene	5:39463720-39857838	960.195	52.005	−0.884677	5,00E−05	0.0137833
PELI1	Pellino E3 ubiquitin protein ligase	11:62471600-62528075	582.145	10.933	−0.909237	5,00E−05	0.0137833
UFM1	Ubiquitin fold modifier 1	12:23685100-23717494	970.133	186.143	−0.940158	5,00E−05	0.0137833
GNPTAB	*N*-acetylglucosamine-1-phosphate transferase alpha and beta subunits	5:65912911-65996254	695.849	135.474	−0.961176	5,00E−05	0.0137833
PADI2	Peptidyl arginine deiminase 2	2:136049637-136106203	139.075	70.356	−0.983117	5,00E−05	0.0137833
EGLN3	egl-9 family hypoxia inducible factor 3	21:44805297-44834310	393.162	627.382	0.674221	0.0002	0.0438816
PDE7A	phosphodiesterase 7A	14:31801394-31907693	110.069	176.094	0.677941	5,00E−05	0.0137833
FAM134B	family with sequence similarity 134 member B	20:56709602-56758648	354.557	594.804	0.746397	5,00E−05	0.0137833
LOC104973907	13:78185912–78,193,935	13:78185912-78193935	362.693	611.116	0.752699	5,00E−05	0.0137833
CACNA2D2	calcium voltage-gated channel auxiliary subunit alpha2delta 2	22:50442292-50557414	279.679	473.345	0.759121	0.00015	0.0366511
ASB5	ankyrin repeat and SOCS box containing 5	27:6692311-6736965	210.415	356.904	0.7623	5,00E−05	0.0137833
C8H9orf41	C8H9orf41	8:51341792-51384051	113.787	194.227	0.771412	5,00E−05	0.0137833
SCUBE2	Signal Peptide, CUB Domain And EGF Like Domain Containing 2	15:44090649-44162339	557.635	966.177	0.792965	5,00E−05	0.0137833
RASA4B	RAS P21 Protein Activator 4B	25:35064266-35086722	102.019	586.068	0.7997	5,00E−05	0.0137833
JSP.1	MHC Class I JSP.1	23:28469732-28473374	135.242	751.157	0.848355	0.0002	0.0438816
MGC139164	uncharacterized LOC509649	18:61192240-61215946	29.795	539.186	0.85571	5,00E−05	0.0137833
METTL21E	RIKEN cDNA 4832428D23-like	12:83099333-83115708	167.988	311.064	0.888855	5,00E−05	0.0137833
LCP1	lymphocyte cytosolic protein 1	12:16253902-16360476	226.835	430.083	0.922973	0.0002	0.0438816
TNFRSF12A	TNF receptor superfamily member 12A	25:2435179-2441777	292.435	577.314	0.981242	5,00E−05	0.0137833
CPXM1	carboxypeptidase X, M14 family member 1	13:52844203-52850550	300.673	595.403	0.985671	5,00E−05	0.0137833
OXT	Oxytocin/neurophysin I prepropeptide	13:52575289-52576188	84.859	423.933	100.123	5,00E−05	0.0137833
ADAMTS20	ADAM metallopeptidase with thrombospondin type 1 motif 20	5:37102284-37333841	331.945	689.235	105.405	5,00E−05	0.0137833
LOC782233	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit G pseudogene	3:28623286-28623761	434.366	907.754	106.339	5,00E−05	0.0137833
C14H8orf22	chromosome 14 open reading frame, human C8orf22	14:21617721–21,622,423	104.417	218.42	106.475	5,00E−05	0.0137833
TSHZ1	Teashirt Zinc Finger Homeobox	24:3673496-3752935	254.575	603.991	124.643	5,00E−05	0.0137833
ESRRG	estrogen related receptor gamma	16:20592131-21286840	267.299	635.796	125.011	5,00E−05	0.0137833
LOC781298	phosphatidylcholine transfer protein	19:6355394-6379570	160.311	383.149	125.703	5,00E−05	0.0137833
LOC520016	multidrug resistance-associated protein 4-like	12:74453909-74520551	209.991	506.301	126.967	5,00E−05	0.0137833
RFX2	regulatory factor X2	7:19579227-19640152	220.598	542.335	129.776	5,00E−05	0.0137833
SLC6A2	Solute Carrier Family 6 Member 2	18:23941694-24013458	0.56811	158.049	147.612	5,00E−05	0.0137833
CPE	Carboxypeptidase E	17:541678-697925	0.87449	260.317	157.375	5,00E−05	0.0137833
HBB	Beta-globin	15:49022977-49024620	239.236	572.516	206.305	5,00E−05	0.0137833

[a] The fold change estimates (relative expression) refer to the group with a low-marbling score.

the body, insulin stops the fat breakdown and avoids the breakdown of triglycerides into fatty acids. The EGL3N (Egl-9 Family Hypoxia Inducible Factor 3) gene, expressed more in low-marbling group, also act improving insulin sensitivity, and in humans, stimulating other factors and helping to stabilize the diabetes. In a study with Nellore cattle muscle tissue, the CPE gene was upregulated for myristic, palmitic, stearic and sum of saturated fatty acids, and downregulated for oleic, monounsaturated sum off fatty acids and omega 3 [21].

The CPE gene, together JSP.1 (MHC Class I JSP.1) and BoLA-BQB (MHC class II antigen - bovine leukocyte antigen), all expressed more in high-marbling group, also make part of the Type I diabetes mellitus pathway. The JSP.1 and BLA-BQB are members of Major Histocompatibility Complex family, essential for adaptive immunity [26]. The JSP.1 belongs to class I, which encodes glycoproteins expressed in almost all nucleated cells and has the function of killing cells that display specific antigens to the body by other cells of the immune system.

The BLA-BQB gene belongs to class II, which encodes glycoproteins expressed primarily on antigen-presenting cells, including macrophages that wrap foreign particles such as bacteria. Type I diabetes is an autoimmune disease, in which the immune system attacks the pancreas and ends the natural production of insulin. Because this, in humans, the patient with the disease will have to inject insulin in the body to put the circulating glucose in the cells. Therefore, the application of high doses

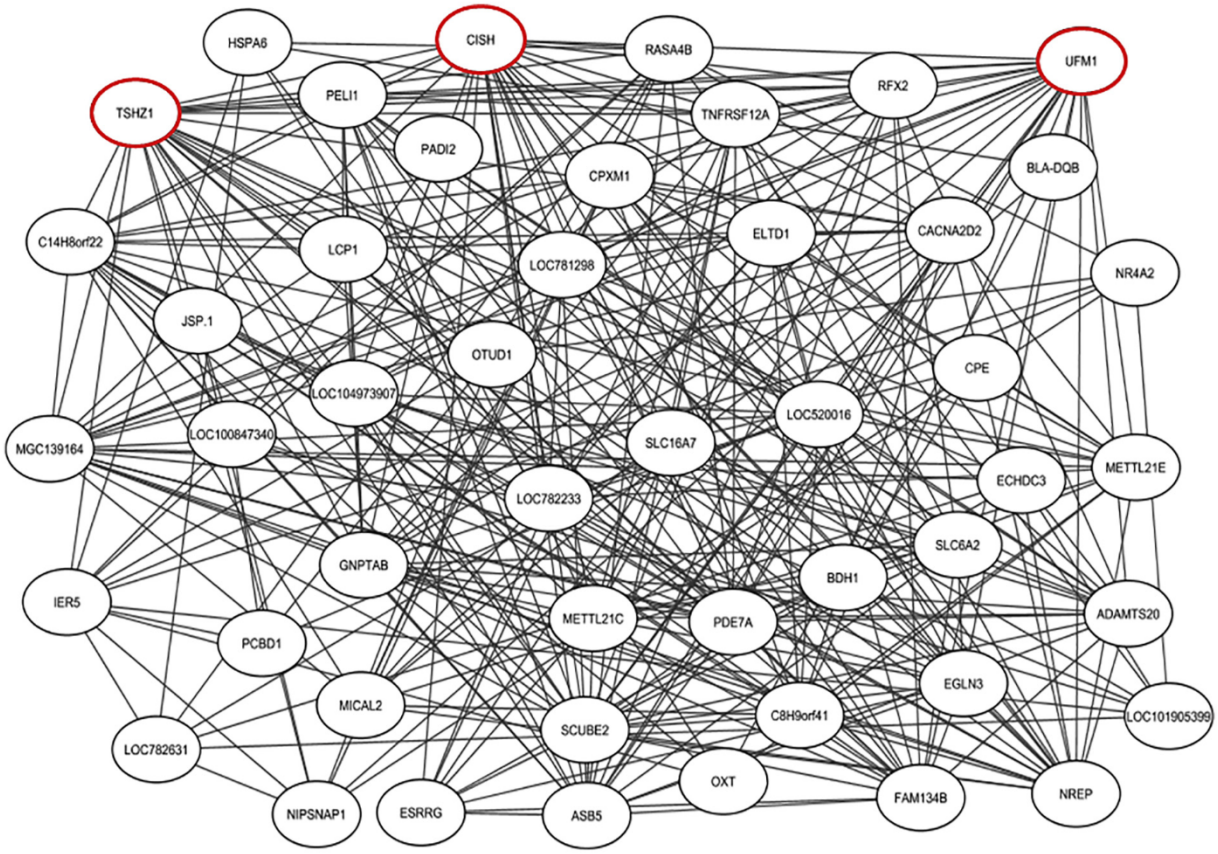

Fig. 1. Transcript correlation network for the differentially expressed genes in Nellore cattle muscle tissue divergently ranked on marbling (cut-off of > 0.7 and < -0.7). Genes circled in red were predicted as hubs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of insulin amplifies the risk of obesity and episodes of hypoglycemia. Thus, these differentially expressed genes may be related to the deposition of intramuscular fat in cattle.

The term response to abiotic stimulus (GO:0009628) refers to any process that results in a change in state or activity of a cell or an as a result of a non-living stimulus [22]. The genes that make party of this term are: Oxytocin/neurophysin I prepropeptide (OXT) and HSPA6. The heat shock protein family A (Hsp70) member 6 (*HSPA6*) transcript encodes a chaperon that is directly involved in the protection of proteins against environmental and genetic stress, contributing to the correct formation of their three-dimensional structure [27]. In a gene expression study using blood cells from obese and normal weight humans, this gene was expressed more in obese subjects [28]. In the present study, the *HSPA6* gene was expressed more in animals with a high marbling score.

The *OXT* transcript, encodes oxytocin, a hormone normally synthesized in the hypothalamus and stored in the posterior pituitary. This hormone is water-soluble and therefore circulates freely in the bloodstream. The most important and known function of oxytocin is related to lactogenesis [29]. In addition, oxytocin acts on lipid storage and energy expenditure and plays a role in processes such as lipolysis, secretory activity and plasticity [30]. In a study on Brahman cattle, De Jager et al. [31] also reported expression of this gene in *longissimus thoracis* muscle, quantified by microarray and qRT-PCR.

By means of a co-expression analysis, was possible to predict three hub genes: cytokine inducible SH2 containing protein (CISH), ubiquitin fold modifier 1 (UFM1) and Teashirt Zinc Finger Homeobox 1 (TSHZ1).

The hub genes are expected to play an important role in biology of the organism studied [32]. The CISH gene is a member of negative regulation of insulin receptor signaling pathway (GO:0046627). This pathway is responsible for any process that stops, prevents, or reduces the frequency, rate or extent of insulin receptor signaling [22].

Other important biological process that CISH gene participates is the negative regulation of JAK-STAT cascade (GO:0046426). According to Xu et al. [33], Modulation of the JAKs–STATs pathway can regulate lipid metabolism in adipocyte and lipid metabolism disorder, may result in metabolic syndrome including obesity, fatty liver and insulin resistance.

The UFM1 gene is involved in fatty acid metabolism [34] and enrichment analysis demonstrate that UFM1 is related to terms response to lipid (GO:0033993) and cellular response to lipid (GO:0071396). Other important action of this gene is in regulation of intracellular estrogen receptor signaling pathway (GO:0033146). In humans, studies demonstrate that estrogens regulate a number of steps in lipid metabolism and estrogen therapy and hormone treatment approaches can protect against fatty liver, insulin resistance and diabetes.

TSHZ1 encondes a protein member of the teashirt C2H2-type zinc-finger protein family and may be involved in transcriptional regulation of developmental processes. According to Raum et al. [35], in humans TSHZ1 is downregulated in type 2 diabetic islets. The same author identified TSHZ1 as a new regulator of endocrine cell maturation and function, that might be exploited for therapeutic benefit for diabetes.

As we demonstrated in this work, other differentially expressed genes are related to insulin and diabetes mellitus, suggesting the importance of the mechanisms to intramuscular fat deposition.

5. Conclusion

The transcriptome study of animals ranked divergently on marbling revealed genes principally involved in lipid metabolism, insulin

synthesis and type I diabetes mellitus. Three hub genes are identified (CISH, UFM1, TSHZ1) and all of them are involved in insulin and diabetes mellitus metabolism. These results demonstrate the complexity of the molecular mechanisms underlying intramuscular fat deposition in *longissimus thoracis* muscle, indicating key genes and pathways, principally related to insulin and diabetes mellitus metabolism, which may help to develop strategies designed to select animals with greater marbling.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2019.06.001.

Ethics statement

All experimental procedures involving animals were approved by Ethics Committee of the São Paulo State University (Unesp), School of Agricultural and Veterinarian Sciences, Jaboticabal, SP, Brazil (CEUA protocol: 18,340/2016).

Consent for publication

Not applicable.

Availability of data and materials

The dataset utilized in this study belongs to a Qualitas Nelore breeding program company, and could be available on request. The author do not have authorization to share the data.

Competing interests

The authors declare no competing interests.

Funding

This study was supported by the São Paulo Research Foundation (FAPESP, grants 2009/16118-5 and 2013/09190-7) and was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Authors' contributions

LFSF, DFJG, LGA, JAF and FB conceived and designed the experiments; LFSF, LALC and DFJG performed the experiments; LFSF, LGA and DBSS performed the analyses and interpreted the results; LFSF, LGA, FB, JAF, and DBSS drafted and revised the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgments

We thank the Nellore Qualitas breeding program for providing the tissue samples and database used in this study.

References

[1] L.V. Cundiff, Beef Cattle: Breeds and Genetics, in: A.W. Bell (Ed.), Pond Wg, Encyclopedia Of Animal Science. Cornell University, Ithaca, 2004, pp. 800–830.
[2] A.F. Lucchiari Filho, Pecuária da carne bovina, LinBife, São Paulo, 2000, p. 134.
[3] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, Nat. Rev. Genet. 10 (1) (2009) 57–63, https://doi.org/10.1038/nrg2484.
[4] J.H. Malone, B. Oliver, Microarrays, deep sequencing and the true measure of the transcriptome, BMC Biol. 9 (2011) 34.
[5] D.Y. Oh, Y.S. Lee, B.M. La, J.Y. Lee, Y.S. Park, J.H. Lee, et al., Identification of exonic nucleotide variants of the thyroid hormone responsive protein gene associated with carcass traits and fatty acid composition in Korean cattle, Asian Austral. J. Anim. 27 (10) (2014) 1373–1380.
[6] D. Chen, W. Li, M. Du, M. Wu, B. Cao, Sequencing and characterization of divergent marbling levels in the beef cattle (Longissimus thoracis muscle) transcriptome, Asian Australas. J. Anim. Sci. 28 (2015) 158–165.
[7] A.S. Cesar, L.C. Regitano, J.E. Koltes, E.R. Fritz-Waters, D.P. Lanna, G. Gasparin, G.B. Mourão, P.S. Oliveira, J.M. Reecy, L.L. Coutinho, Putative regulatory factors associated with intramuscular fat content, PLoS One 10 (6) (2015) 1–21.
[8] J.J. Michal, Z.W. Zhang, C.T. Gaskins, Z. Jiang, The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu × Limousin F2 crosses, Anim. Genet. 37 (2006) 400–402.
[9] USDA, Official USDA Marbling Photographs, National cattlemen's Beef Association. U.S. Depth. of Agriculture, Washington, D. C. USA, 1989.
[10] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for illumina sequence data, Bioinformatics 30 (2014) 2114–2120.
[11] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, Nat. Methods 12 (2015) 357–360.
[12] Daniel R. Zerbino, Premanand Achuthan, M. Ridwan Amode Wasiu Akanni, Daniel Barrell, Jyothish Bhai, Konstantinos Billis, Carla Cummins, Astrid Gall, Carlos García Giro'n, Laurent Gil, Leo Gordon, Leanne Haggerty, Erin Haskell, Thibaut Hourlier, Osagie G. Izuogu, Sophie H. Janacek, Thomas Juettemann, Jimmy Kiang To, Matthew R. Laird, Ilias Lavidas, Zhicheng Liu, Jane E. Loveland, Thomas Maurel, William McLaren, Benjamin Moore, Jonathan Mudge, Daniel N. Murphy, Victoria Newman, Michael Nuhn, Denye Ogeh, Chuang Kee Ong, Anne Parker, Mateus Patricio, Harpreet Singh Riat, Helen Schuilenburg, Dan Sheppard, Helen Sparrow, Kieron Taylor, Anja Thormann, Alessandro Vullo, Brandon Walts, Amonida Zadissa, Adam Frankish, Sarah E. Hunt, Myrto Kostadima, Nicholas Langridge, Fergal J. Martin, Matthieu Muffato, Emily Perry, Magali Ruffier, Dan M. Staines, Stephen J. Trevanion, Bronwen L. Aken, Fiona Cunningham, Andrew Yates, Paul Flicek, Ensembl, (2018), https://doi.org/10.1093/nar/gkx1098 (ubMed PMID: 29155950).
[13] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, Nat. Protoc. 7 (3) (2012) 562–578.
[14] C. Trapnell, D.G. Hendrickson, M. Sauvageau, L. Goff, J.L. RInn, L. Pacther, Differential analysis of gene regulation at transcript resolution with RNA-seq, Nat. Biotechnol. 31 (2013) 46–53.
[15] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504.
[16] Y. Assenov, F. Ramírez, S.E. Schelhorn, T. Lengauer, M. Albrecht, Computing topological parameters of biological networks, Bioinformatics 15 (24) (2008) 282–284.
[17] C.H. Chin, S.H. Chen, H.H. Wu, C.W. Ho, M.T. Ko, C.Y. Lin, CytoHubba: identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (Suppl. 4) (2014) S11.
[18] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources, Nat. Protoc. 4 (1) (2009) 44–57.
[19] H. Li, Genomic Variation and Whole Genome Association Studies of Economically Important Traits. Thesis submitted to the Faculty of Graduate Studies, Edmonton, Alberta, 2012.
[20] L. Fagerberg, B.M. Hallström, P. Oksvold, C. Kampf, D. Djureinovic, J. Odeberg, M. Habuka, S. Tahmasebpoor, A. Danielsson, K. Edlund, A. Asplund, E. Sjöstedt, E. Lundberg, C.A. Szigyarto, M. Skogs, J.O. Takanen, H. Berling, H. Tegel, J. Mulder, P. Nilsson, J.M. Schwenk, C. Lindskog, F. Danielsson, A. Mardinoglu, A. Sivertsson, K. von Feilitzen, M. Forsberg, M. Zwahlen, I. Olsson, S. Navani, M. Huss, J. Nielsen, F. Ponten, M. Uhlén, Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics, Mol. Cell. Proteomics 13 (2) (2014) 397–406.
[21] M.P. Berton, L.F.S. Fonseca, D.F.J. Gimenez, B.L. Utembergue, A.S.M. Cesar, L.L. Coutinho, et al., Gene expression profile of intramuscular muscle in Nellore cattle with extreme values of fatty acid, BMC Genomics 17 (1) (2016), https://doi.org/10.1186/s12864-016-3232-y.
[22] Ashburner, et al., Gene ontology: tool for the unification of biology, Nat. Genet. 25 (1) (2000) 25–29.
[23] L.E. Otterbein, M.P. Soares, K. Yamashita, F.H. Bach, Heme oxygenase-1: unleashing the protective properties of heme, Trends Immunol. 24 (2003) 449–455.
[24] S.P. Duffy, J. Shing, P. Saraon, L.C. Berger, M.V. Eiden, A. Wilde, C.S. Tailor, The Fowler syndrome-associated protein FLVCR2 is an importer of heme, Mol. Cell. Biol. 30 (2010) 5318–5532.
[25] U.V. Bagni, R.R. Luiz, G.L. Veiga, Overweight is associated with low hemoglobin levels in adolescent girls, Obes. Res. Clin. Pract. 7 (2013) 218–229.
[26] M. Wieczorek, E.T. Abualrous, J. Sticht, M. Álvaro-Benito, S. Stolzenberg, F. Noé, et al., Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation, Front. Immunol. 8 (2017) 292, , https://doi.org/10.3389/fimmu.2017.00292.
[27] M.P. Mayer, B. Bukau, Hsp70 chaperones: cellular functions and molecular mechanism, Cell. Mol. Life Sci. 62 (2005) 670–684.
[28] D. Minchenko, Expression of endoplasmic reticulum stress related genes in blood cells of obese boys with and without insulin resistance, Int. J. Biomed. 5 (2015) 24–29.
[29] D. Collier, J.P. Mcnamara, C.R. Wallace, M.H. Dehoff, A review of endocrine regulation of metabolism during lactation, Anim. Sci. J. 59 (1984) 498–510.
[30] M.H. Fonseca-Alanis, J. Takada, M.I.C. Alonso-Vale, The adipose tissue as a regulatory center of the metabolismo, Arq Bras Endocrinol. Metab. 50 (2006) 216–219.
[31] N. De Jager, N.J. Hudson, A. Reverter, H. Wang, S.H. Nagaraj, L.M. Cafe, P.L. Greenwood, R.T. Barnard, K.P. Kongsuwan, B.P. Dalrymple, Chronic exposure to anabolic steroids induces the muscle expression of oxytocin and a more than fiftyfold increase in circulating oxytocin in cattle, Physiol. Genomics 43 (2011) 467–478.

[32] P. Langfelder, P.S. Mischel, S. Horvath, When is hub gene selection better than standard meta-analysis? PLoS One 8 (4) (2013) e61505.
[33] D. Xu, C. Yin, S. Wang, Y. Xiao, JAK-STAT in lipid metabolism of adipocytes, JAKSTAT. 2 (4) (2013) e27203.
[34] Y. Zhang, M. Zhang, J. Wu, G. Lei, H. Li, Transcriptional regulation of the Ufm1 conjugation system in response to disturbance of the endoplasmic reticulum homeostasis and inhibition of vesicle trafficking, PLoS One 7 (11) (2012) e48587, , https://doi.org/10.1371/journal.pone.0048587.
[35] J.C. Raum, S.A. Soleimanpour, D.N. Groff, N. Coré, L. Fasano, A.N. Garratt, C. Dai, A.C. Powers, D.A. Stoffers, Tshz1 regulates pancreatic beta-cell maturation, Diabetes 64 (2015) 2905–2914, https://doi.org/10.2337/db14-1443 et al. (2015).