CrossMark

Deciphering the regulation of porcine genes influencing growth, fatness and yield-related traits through genetical genomics

Angel M. Martínez-Montes[1] · Anixa Muiños-Bühl[1] · Almudena Fernández[1] ·
Josep M. Folch[2,3] · Noelia Ibáñez-Escriche[4] · Ana I. Fernández[1]

Abstract Genetical genomics approaches aim at identifying quantitative trait loci for molecular traits, also known as intermediate phenotypes, such as gene expression, that could link variation in genetic information to physiological traits. In the current study, an expression GWAS has been carried out on an experimental Iberian×Landrace backcross in order to identify the genomic regions regulating the gene expression of those genes whose expression is correlated with growth, fat deposition, and premium cut yield measures in pig. The analyses were conducted exploiting Porcine 60K SNP BeadChip genotypes and Porcine Expression Microarray data hybridized on mRNA from *Longissimus dorsi* muscle. In order to focus the analysis on productive traits and reduce the number of analyses, only those probesets whose expression showed significant correlation with at least one of the seven phenotypes of interest were selected for the eGWAS. A total of 63 eQTL regions were identified with effects on 36 different transcripts. Those eQTLs overlapping with phenotypic QTLs on SSC4, SSC9, SSC13, and SSC17 chromosomes previously detected in the same animal material were further analyzed. Moreover, candidate genes and SNPs were analyzed. Among the most promising results, a long non-coding RNA, *ALDBSSCG0000001928*, was identified, whose expression is correlated with premium cut yield. Association analysis and in silico sequence domain annotation support *TXNRD*3 polymorphisms as candidate to regulate *ALDBSSCG0000001928* expression, which can be involved in the transcriptional regulation of surrounding genes, affecting productive and meat quality traits.

✉ Angel M. Martínez-Montes
martinez.angel@inia.es

1 Departamento de Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), 28040 Madrid, Spain

2 Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Spain

3 Plant and Animal Genomics, Centre de Recerca en Agrigenòmica (CRAG), Consorci CSIC-IRTA-UAB-UB, Campus UAB, 08193 Bellaterra, Spain

4 Departament de Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), 25198 Lleida, Spain

Introduction

Several approaches are now available in order to elucidate genetic architecture of complex traits such as growth, fat deposition, carcass composition, or meat quality in livestock species. Traditionally phenotypic QTL (QTL) mapping has been carried out using linkage analysis with limited number of microsatellite markers (Wang et al. 2002; Deng et al. 2000). Although this approach provided reliable results, further analyses to identify underlying genes or causative mutations have not been very successful, in part due to the lack of QTL position precision limited by the available markers (Würschum and Kraft 2014; Schön et al. 2004). More recently, Genome-Wide Association Study (GWAS) using high density SNP panels has emerged as a strong approach that minimizes the marker number limitation (Hill 2012; Sun et al. 2015). GWAS analyses are usually focused on the study of phenotypic traits using genomic data. Nevertheless, it is not the only possible application; genetical genomics studies (Jansen and Nap

2001; Breitling et al. 2008) can be conducted by carrying out expression GWAS (eGWAS). Genetical genomics aims at identifying QTL for molecular traits, also known as intermediate phenotypes, such as gene expression (eQTL) that could link variation in genetic information to physiological traits (Williams et al. 2007). These analyses allow us to obtain information regarding gene expression regulation, regulation paths, and interactions that could help understand the genetic architecture of complex traits (Zou et al. 2012). Studies in human have validated the successfulness of eGWAS to identify variants associated with complex human diseases, and the potential role of gene expression changes in those diseases (Kodama et al. 2012; Zou et al. 2012).

Previous studies in an Iberian×Landrace porcine experimental cross using both linkage and GWAS analyses allowed us to identify QTL with relevant effects on growth, fat deposition, and premium cut yield-related traits (Óvilo et al. 2000; Varona et al. 2002; Mercadé et al. 2005; Fernández et al. 2012, 2014). However, the identification of potential causal variants has been limited to *ELOVL6, LEPR*, and *FABP* genes (Corominas et al. 2013, 2015; Ovilo et al. 2005; Estellé et al. 2006; Pérez-Montarelo 2012). Therefore, the aim of the current study is to help in the identification of potential causal variants for these traits through a genetical genomics study. An eGWAS study has been carried out on an experimental Iberian×Landrace backcross in order to identify the genomic regions regulating the gene expression of those genes whose expression is correlated with growth, fat deposition, and premium cut yield measures. The analyses were conducted exploiting Porcine 60K SNP BeadChip genotypes and Porcine Expression Microarray (Affymetrix) data hybridized on mRNA from *Longissimus dorsi* muscle.

Materials and methods

Animals

The phenotypic information and gene expression data used in the current study belong to an experimental backcross F1 (Iberian×Landrace)×Landrace of the IBMAP population (Óvilo et al. 2000; Mercadé et al. 2005; Óvilo et al. 2005). The IBMAP F1 generation was obtained from three Iberian Guadyerbas boars and 30 Landrace sows, five of these F1 boars were mated with 25 Landrace sows obtaining 160 animals from the backcross (BC).

All animal procedures were performed according to the Spanish Policy for Animal Protection RD1201/05, which meets the European Union Directive 86/609 about the protection of animals used in experimentation.

Phenotypic data

For the study seven traits related to growth, fatness, and meat quality were recorded (Table 1). These were body weight at 150 days of mean age (BW150), backfat thickness measured at 75 kg (BFT75) and at slaughter (BFS), weights of premium cuts, such as hams (HW), shoulders (SW), and loin bone-in (LBW), and intramuscular fat content (IMF) measured in *Longissimus dorsi* samples at slaughter (Fernandez et al. 2012).

Gene expression

Gene expression data were obtained from the hybridization of mRNA samples coming from *Longissimus dorsi* of 102 backcrossed individuals with the Porcine Expression Microarray (Affymetrix) as described in Pena et al. (2013). Quality control was carried out with the microarray data using affyPLM package of the Bioconductor software (http://www.bioconductor.org/). RNA normalization was carried using BRB-Array Tools (v. 3.6.0) (http://linus.nci.nih.gov/BRB-ArrayTools.html). Expression data are expressed as the log2 of probeset signal intensity.

Correlation (phenotype and expression)

A correlation analysis was carried out between phenotypic (BW150, BFT75, BFS, HW, SW, LBW, and IMF) and expression data. Expression and phenotypic data were corrected adjusting a linear model, setting sex and batch as fixed effects, and slaughter age as random effect. Pearson correlation coefficient was calculated between the predicted values from 24,000 probesets and the predicted values of the seven phenotypic records. Genes with significant correlation levels ($r = |0.32| - |0.66|$, p value <0.001, q value <0.002) were selected for further analysis. Microarray probesets were annotated using NetAff from Affymetrix (https://www.affymetrix.com/analysis/index.affx).

Genotyping data

DNA samples from 160 backcrossed and their F1 and F0 relatives were genotyped with the PorcineSNP60 BeadChip (Illumina, Inc.), designed by Ramos et al. (2009). GenomeStudio software (Illumina, Inc.) was employed to visualize, edit, standardize quality filter, and extract genotyping data. A second process of data filtering was carried out with GenABEL software, those markers with a minimum allele frequency (MAF) $<2.5\%$, and markers deviating from Hardy–Weinberg equilibrium (FDR $<1\%$) were discarded. A total of 31,606 SNPs were considered for further analyses.

Table 1 eQTL identified in the eGWAS analyses conducted *on Longissimus dorsi* gene expression data

eQTL						Associated gene expression					Cis_Trans
Name	Chr	Start	End	Size (bp)	Num_SNPs	Probeset	Name	Chr	Gene start	Gene end	
R1	1	8,988,180	15,231,294	6,243,114	11	Ssc.1860.1.S1_at	DYNLT1	1	10,385,075	10,391,604	cis
R2	1	129,136,194	130,167,976	1,031,782	8	Ssc.25573.1.S1_at	HIVEP2	1	24,926,110	24,957,597	trans
R3	1	146,180,618	146,332,008	151,390	5	Ssc.20685.1.S1_at	BCL2	1	175,695,537	175,696,139	trans
R4	1	309,039,258	309,125,803	86,545	2	Ssc.30633.1.S1_at	MNS1	1	127,767,312	127,821,308	trans
R5	2	16,416	10,979,357	10,962,941	28	Ssc.9365.2.S1_a_at	IGF2	GL896425.1	1	2,416	trans
R6	2	251,447	10,979,357	10,727,910	22	Ssc.20525.1.S1_at	P09565	GL896425.1	1	2,416	trans
R7	2	118,359,766	120,243,915	1,884,149	9	Ssc.10952.1.S1_at	DLG1	13	142,336,267	142,629,535	trans
R8	2	162,084,552	162,298,086	213,534	2	Ssc.20525.1.S1_at	P09565	GL896425.1	1	2,416	trans
R9	2	162,084,552	162,298,086	213,534	2	Ssc.9365.2.S1_a_at	IGF2	GL896425.1	1	2,416	trans
R10	3	14,221,604	14,772,381	550,777	2	Ssc.29388.1.A1_at	SEMA3A	JH118932.1	48,007	86,320	trans
R11	3	58,059,120	62,507,762	4,448,642	19	Ssc.11208.1.S1_at	IGK	3	59,720,209	59,881,293	cis
R12	3	93,221,408	97,347,171	4,125,763	8	Ssc.7060.2.S1_at	FBXO11	3	98,538,854	98,709,802	trans
R13	3	111,541,225	112,768,775	1,227,550	11	Ssc.26316.1.S1_at	KTN1	1	205,432,796	205,534,850	trans
R14	4	13,084,536	13,330,048	245,512	2	Ssc.25555.1.S1_at	TATDN1	4	15,769,901	15,816,051	trans
R15	4	15,848,929	16,347,684	498,755	4	Ssc.25555.1.S1_at	TATDN1	4	15,769,901	15,816,051	cis
R16	4	19,818,145	22,561,896	2,743,751	13	Ssc.25555.1.S1_at	TATDN1	4	15,769,901	15,816,051	trans
R17	4	28,899,542	29,558,810	659,268	2	Ssc.25555.1.S1_at	TATDN1	4	15,769,901	15,816,051	trans
R18	4	63,841,367	63,988,841	147,474	2	Ssc.9109.1.A1_at	IDE	14	113,505,019	113,583,601	trans
R19	4	128,983,125	129,016,956	33,831	2	Ssc.7190.1.S1_at	BUB1B	1	146,310,747	146,312,591	trans
R20	5	876,762	1,864,156	987,394	9	Ssc.1790.1.S1_at	NUP50	5	1,512,925	1,526,924	cis
R21	6	48,585,961	52,336,598	3,750,637	30	Ssc.18795.1.A1_at	CNTN6	1	63,899,864	63,922,174	trans
R22	6	50,478,565	50,888,554	409,989	3	Ssc.894.1.A1_at	Mico1	6	50,803,937	50,804,291	cis
R23	6	51,775,907	52,993,696	1,217,789	6	Ssc.30633.1.S1_at	MNS1	1	127,767,312	127,821,308	trans
R24	6	80,520,240	84,333,986	3,813,746	28	Ssc.2152.1.A1_at	BSDC1	6	82,782,946	82,808,253	cis
R25	6	85,813,657	86,899,252	1,085,595	3	Ssc.11158.2.S1_at	PTP4A2	6	82,337,830	82,348,373	trans
R26	6	155,897,700	157,126,627	1,228,927	10	Ssc.21969.1.A1_at	C1orf50	6	156,177,269	156,183,027	cis
R27	7	38,189,183	39,842,612	1,653,429	2	Ssc.19634.1.S1_at	MRPL14	7	44,942,222	44,955,752	trans
R28	7	44,767,409	48,179,859	3,412,450	9	Ssc.19634.1.S1_at	MRPL14	7	44,942,222	44,955,752	cis
R29	7	50,750,507	50,781,071	30,564	2	Ssc.19634.1.S1_at	MRPL14	7	44,942,222	44,955,752	trans
R30	7	128,891,964	131,361,899	2,469,935	24	Ssc.15678.1.A1_s_at	CHCHD6	GL894452.1	10,304	10,550	trans
R31	7	128,891,964	131,323,540	2,431,576	10	Ssc.21543.1.S1_at	–	–	–	–	trans
R32	8	8,934,560	11,962,779	3,028,219	29	Ssc.10298.1.A1_at	TAPT1	8	11,018,821	11,067,248	cis
R33	8	12,476,033	12,483,876	7,843	2	Ssc.7713.1.A1_at	LAP3	6	40,305,180	40,306,031	trans
R34	8	33,235,728	33,365,922	130,194	2	Ssc.29017.1.A1_at	SETD6	6	18,327,600	18,332,824	trans
R35	9	118,021,130	120,667,155	2,646,025	5	Ssc.7190.1.S1_at	BUB1B	1	146,310,747	146,312,591	trans
R36	11	3,914,900	23,463,154	19,548,254	24	Ssc.9916.1.S1_at	SPG20	11	12,231,552	12,255,683	cis

Table 1 (continued)

eQTL						Associated gene expression					
Name	Chr	Start	End	Size (bp)	Num_SNPs	Probeset	Name	Chr	Gene start	Gene end	Cis_Trans
R37	13	19,683,797	19,750,520	66,723	2	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R38	13	19,774,122	19,859,830	85,708	2	Ssc.24608.1.S1_at	DAG1	JH118489.1	81,292	97,231	trans
R39	13	25,730,204	31,650,967	5,920,763	5	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R40	13	26,286,725	27,911,595	1,624,870	8	Ssc.24997.1.S1_at	LSAMP	13	153,563,024	153,783,293	trans
R41	13	34,526,276	36,964,009	2,437,733	6	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R42	13	46,063,345	53,877,058	7,813,713	11	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R43	13	59,740,715	73,787,637	14,046,922	56	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R44	13	78,219,919	84,277,803	6,057,884	15	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	cis
R45	13	87,702,788	89,919,771	2,216,983	10	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R46	13	98,646,754	103,194,490	4,547,736	6	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R47	13	124,746,808	126,830,429	2,083,621	5	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R48	13	134,933,329	140,308,510	5,375,181	18	Ssc.24542.1.S1_at	RGMB	2	108,921,274	108,921,825	trans
R49	13	154,043,749	154,103,781	60,032	2	Ssc.10589.1.A1_at	Ssc.10589.1.A1_at	13	80,736,233	80,736,662	trans
R50	14	31,671,713	34,956,919	3,285,206	27	Ssc.21861.1.S1_at	PYCRL	4	1,013,502	1,013,974	trans
R51	14	100,004,352	120,439,354	20,435,002	209	Ssc.9109.1.A1_at	IDE	14	113,505,019	113,583,601	cis
R52	14	135,271,348	138,328,275	3,056,927	10	Ssc.30987.1.S1_at	NHLRC2	14	135,285,101	135,358,491	cis
R53	15	12,377,279	15,286,888	2,909,609	12	Ssc.3249.1.S1_at	QSOX1	9	133,595,707	133,637,533	trans
R54	15	99,540,054	101,622,603	2,082,549	14	Ssc.16050.1.A1_at	ADAM17	3	135,176,525	135,222,600	trans
R55	15	99,540,054	101,622,603	2,082,549	14	Ssc.18795.1.A1_at	CNTN6	13	63,839,664	63,863,130	trans
R56	17	18,894,036	25,670,863	6,776,827	15	Ssc.7666.1.A1_at	PSMF1	17	38,667,378	38,711,350	trans
R57	17	23,265,918	29,684,825	6,418,907	4	Ssc.21242.1.S1_at	CTNNBL1	17	46,357,154	46,401,936	trans
R58	17	28,475,777	29,684,825	1,209,048	2	Ssc.7666.1.A1_at	PSMF1	17	38,667,378	38,711,350	trans
R59	17	33,892,898	34,888,048	995,150	3	Ssc.7666.1.A1_at	PSMF1	17	38,667,378	38,711,350	trans
R60	17	38,047,955	41,558,311	3,510,356	7	Ssc.7666.1.A1_at	PSMF1	17	38,667,378	38,711,350	cis
R61	17	40,556,315	41,558,311	1,001,996	2	Ssc.21242.1.S1_at	CTNNBL1	17	46,357,154	46,401,936	trans
R62	17	45,164,920	54,774,793	9,609,873	16	Ssc.21242.1.S1_at	CTNNBL1	17	46,357,154	46,401,936	cis
R63	17	47,981,889	56,096,189	8,114,300	14	Ssc.7666.1.A1_at	PSMF1	17	38,667,378	38,711,350	trans

Description of the eQTL regions, including start and end positions, size, SNP number, affected probeset/gene, and cis_trans classification

eGWAS analysis

A genome-wide association study was performed using the GenABEL package (Karssen et al. 2016) in R environment. The analysis was carried on 102 individuals, those with expression and genotyping data. The genome-wide analysis was performed following the model:

$$y_{ijk} = S_i + B_j + bx_k + \sum_l \lambda_{lk} a_l + u_k + e_{ijk},$$

where y_{ijk} is the trait value of kth individual, S_i and B_j are fixed effects for sex and batch respectively, and b is the slaughter age regression coefficient. Additive effect of the SNP is a_l and λ_{lk} is the indicator related with the number of copies of the lth allele (0, 1, or 2) and u_k would be the infinitesimal effect of the kth individual, e_{ijk} is the random residual term. The same model but using carcass weight as regression coefficient showed similar results. QValue package in R was used to perform correction for multiple tests (Bass et al. 2015). Significant associations were considered for those reporting q value <0.05.

Region analysis

eQTLs were determined by two or more significantly associated SNPs within a maximum distance of 2 Mb. The genetic content of the eQTL was extracted using BioMart tool from Porcine Ensembl database. FatiGO and ReviGO online tools were used to investigate Gene Ontology enrichment and function. In order to prioritize the investigation on eQTL regions, those regions were compared with QTL regions obtained in previous studies carried out in the same material.

Candidate SNPs analyses

A candidate SNP search was done by exploiting an RNA-seq assay previously conducted on the same animal material (Martínez-Montes et al. 2016). The candidate SNPs were validated by Sanger sequencing on cDNA synthesized from mRNA. Primer pairs were designed from exon to exon, in order to avoid genomic DNA amplification (Supplemental Table 1). The PCR reactions were performed in a final volume of 25 μl, containing 4 μl of cDNA, 0.5 μl of polymerase, 2.5 μl buffer 10×, 2.5 μl of dNTPs, and 0.5 μl of each primer. Thermocycling was carried out under the following conditions: 94 °C for 5 min, 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, with a final extension of 72 °C for 10 min. The PCR reactions were carried out in a GeneAmp PCR System 9700 (Applied Biosystems, Warrington, UK). The PCR products were purified with the illustra™GFX™ PCR DNA purification kit (GE Healthcare, UK) according to the manufacturers'

protocol. PCR products were sequenced with both forward and reverse primers using the 3100 BigDye® Terminator v3.1 Matrix Standard in a 3730 DNA Analyzer (Applied Biosystems Warrington, UK). After validation, the SNPs were genotyped in the 160 backcrossed animals using different techniques: pyrosequencing using specific primers (Sup. Table 1); PCR-RFLP with restriction enzymes *Tsp451* (*ss2031475817*, GenBank ID: 100518810) and *BstUI* (*ss2031475807*, GenBank ID: 100233193); and the OpenArray platform at Servei Veterinari de Genètica Molecular (Universidad Autónoma de Barcelona, Spain).

The specific association analyses for the candidate SNPs and haplotypes, built with Phase 2.1.1 (Stephens et al. 2001), with gene expression measures were carried out using the previous quoted model using Qxpak software (Pérez-Enciso and Misztal 2011). Moreover, associations were also conducted for phenotypic traits. Bonferroni correction was applied to take into account multiple tests, setting up a p value of 0.003.

To examine the interest of the identified candidate SNPs, we analyzed in silico the potential effect of those SNPs that produce amino acid change using Predict Protein tool (Yachdav et al. 2014). Additionally, we used RegRNA (http://regrna.mbc.nctu.edu.tw/) for those synonymous SNPs, to determine the potential effects at mRNA level, stability, or gene expression regulation (Chang et al. 2013). ALDB, a domestic-animal long non-coding RNA database, was used to identify possible non-coding RNA, due to the presence of one of the SNP analyzed in this study, localized in a non-coding DNA region (Li et al. 2015). Also the MEME suite (Bailey et al. 2009) was used to identify possible motifs represented in our sequence as well as to predict potential effects of the SNPs changing the structure of a DNA motif. Multiple Em for Motif Elicitation (MEME) (Bailey and Elkan 1994), Gene Ontology for Motifs (GOMo), (Buske et al. 2010), and Motif Comparison Tool (Tomtom) (Gupta et al. 2007) tools were used to identify potential motifs in our sequence and analyze the gene ontology (GO) enrichment of this motifs by comparisons with previously described motifs.

Results

In order to focus the analysis on the productive traits and reduce the number of genes analyzed from thousands to few hundreds, only those probesets contained in the Affymetrix porcine expression microarray whose expression showed significant correlation with at least one of the seven phenotypes of interest were selected. A total of 820 probesets were selected, showing correlations between 0.32 and 0.66 (p value <0.001, q value <0.002) (Supplemental Table 2). Gene probesets were annotated using

NetAffx tool. One probeset per gene was chosen, the one showing the highest expression level. In total 776 gene-unique probesets were used for the eGWAS.

The eGWAS was carried out using GenABEL package among each of the 776 probesets with the filtered SNPs. The results revealed 954 associations between 880 SNPs with expression levels of 42 genes (eTAS) (Supplemental Table 3). These eTAS corresponded to 63 regions or eQTLs for 36 different transcripts, containing between 2 and 209 eTAS. From the total number of eQTLs 15 were *cis*-associations and 48 *trans*-associations. The eQTL regions were identified on every autosome except on SSC10, SSC12, SSC16, and SSC18 (Table 1), showing the higher association for R33-*trans* (SSC8) and R1-*cis* (SSC6). This analysis validated relevant associations such as the association between Insulin-Like Growth Factor 2 (*IGF2*) expression with R5-*trans* (28 SNPs) and R9-*trans* (2 SNPs) regions on SSC2. A total of 2630 genes were identified within these 63 eQTLS, containing candidate genes as *PTEN* (*Phosphatase and Tensin Homolog*, located on SSC14:108911519–109003081) associated with muscle development, *FADS1* (*Fatty Acid Desaturase 1*, located on SSC2:9,247,472–9,263,631), or *CTNNB1* (*Cadherin-Associated Protein, Beta 1*, located on SSC13:27,623,128–27,667,302).

We focused our further studies on those eQTLs (*cis*- and *trans*-associated) overlapping with QTLs previously detected in the same animal material (Óvilo et al. 2000; Varona et al. 2002; Mercadé et al. 2005; Fernández et al. 2012, 2014). Those were located on SSC4, SSC9, SSC13, and SSC17 (Table 2). These associations implicate four different expression probesets:

The *Ssc.7190.1.S1_at* probeset for QTL regions R19-*trans* and R35-*trans*, which corresponds with the *BUB1B* gene (ENSSSCG00000030580), SSC1:146,304,943–146,312,662. The *BUB1B* is associated with the proliferative capacity of muscle cells (Guntani et al. 2011).

The *Ssc.7666.1.A1_at* probeset for region R60-*cis*, which corresponds to *PSMF1* gene (ENSSSCG00000020887), SSC17:38,667,378–38,711,350. The PSMF1 interest lies in its interaction with *INS* (Insulin), *TGFB1* (Transforming Growth Factor Beta 1), and *CDKN1A* (Cyclin-Dependent Kinase Inhibitor 1A) which is related with myocyte terminal differentiation in muscle development (Guo et al. 1995; Qin et al. 2012).

The *Ssc.21242.1.S1_at* probeset for region R61-*trans*, which corresponds to *CTNNBL1* gene (ENSSSCG00000021553), is involved in basal metabolism and previously related with carcass traits in different species (Espigolan et al. 2015).

The *Ssc.10589.1.A1_at* probeset is located in a non-coding region. Current annotation indicates that this probeset corresponds to a long intergenic non-coding RNA (lncRNA), identified in the ALDB database (http://www.ibiomedical.net/aldb/) as *ALDBSSCG0000001928*, and located at SSC13:80,720,757–80,739,741. The LDB-SSCT0000003202 transcript spans 8,299 bp and two exons. This lncRNA is located in a region that overlaps with a high number of QTLs previously described on PigQTL database, as mainly associated with average daily gain (Hu et al. 2005, http://nhjy.hzau.edu.cn/kech/swxxx/jakj/dianzi/Bioinf8/Animal/Animal8.htm).

SNPs analysis

Candidate genes

A total of 44 positional and functional candidate genes for those QTLs overlapping with eQTLs were selected for candidate polymorphism search (Table 3). Polymorphism search was conducted taking the advantage of our previous SNP identification study based on an RNA-Seq assay performed on the same animal material (Martinez-Montes et al. 2016). We identified 49 SNPs in 13 of the 44 candidate genes. After validation and potential impact evaluation, a total of 20 SNPs located on 10 unique genes

Table 2 Identified eQTL overlapping with previous QTL

eQTL	Chr	Start (Mb)	End (Mb)	QTL	Trait	Start (Mb)	End (Mb)	Reference
R19	4	128,983,125	129,016,956	4	Fatness	129	134	Varona et al. (2002) and Fernández et al. (2012)
R35	9	118,021,130	120,667,155	9	Premium cut yield	120	127	Fernández et al. (2012)
R44	13	78,219,919	84,277,803	13	Premium cut yield	83	91	Fernández et al. (2012)
R45	13	87,702,788	89,919,771	13	Premium cut yield	83	91	Fernández et al. (2012)
R60	17	38,047,955	41,558,311	17	Fatness	39	44	Fernández et al. (2012)
R61	17	40,556,315	41,558,311	17	Fatness	39	44	Fernández et al. (2012)

Overlapping between eQTL identified in the current study and phenotypic QTL identified in previous studies in the IBMAP experimental population

Table 3 Positional candidate genes identified within the selected eQTLs

eQTL	Chr	Start (Mb)	End (Mb)	Candidate genes
R35	9	118,021,130	120,667,155	*LAMB1, NRCAM, LOC100623945, ENSSSCG00000025643, ZNF398, ZNF425, ZNF786, PDIA4, EZH2, CUL1*
R44	13	78,219,919	84,277,803	*LOC100738707, ACAD9, RAB7A, GATA2, MGLL, MCM7, PLXNA1, TXNRD3, CHST13, ACAD11, ICA, RYK, LOC100524398*
R45	13	87,702,788	89,919,771	*RBP2, RBP1*
R60	17	38,047,955	41,558,311	*FKBP1A, TCF15, TRIB3, LOC100519985, SOX12, ID1, BCL2L1, ENSSSCG00000022547, MYLK2, POFUT1, PLAGL2, ASXL1, COMMD7, DNMT3B, BPIFB2, BPIFB6, BPIFB4, BPIFB3*
R61	17	40,556,315	41,558,311	*POFUT1, PLAGL2, ASXL1, COMMD7, DNMT3B, BPIFB2, BPIFB6, BPIFB4, BPIFB3*

Genes located within each of the eQTLs selected by overlapping with QTLs

(Table 4) were selected for genotyping and association analyses in the backcrossed individuals (*ZNF786, ACAD11, RYK, MGLL, TRIB3, PDIA4, LAMB1, RBP1, TXNRD3*, and *ICA*):

The *MGLL* gene encodes a monoglyceride lipase that has been associated with fatty acid uptake and oxidation in pig intramuscular fatty acid composition in the *longissimus thoracic* muscle (Pena et al. 2013).

The *TXNRD3* encodes for thioredoxin reductase 3 that was shown to affect adipocyte differentiation through Wnt signaling pathway (Kipp et al. 2012).

The *ACAD11* gene that encodes an acyl-CoA dehydrogenase that was shown to be in association with variation in residual feed intake in beef cattle (Karisa et al. 2013).

The *RYK* genes encode a receptor-like tyrosine kinase that mediate muscle attachment in drosophila melanogaster via Wnt interaction (Lahaye et al. 2012).

The *RBP1* gene encodes a retinol-binding protein that regulates adipogenesis in mice (Zizola et al. 2010).

The *TRIB3* genes encode tribbles pseudokinase 3 that was shown to be in association with meat quality and production traits in Italian heavy pigs (Fontanesi et al. 2010).

Table 4 Description of candidate SNPs selected and analyzed within candidate genes

SNP ID[a]	MAF	Chr	Genomic position	Location	Predicted SNP effect
ss2031475799	0.25	9	118452493	exon 26	–
ss2031475800	0.34	9	118456942	intron (ex24–ex25)	–
ss2031475801	0.47	9	118475910	exon 17	–
ss2031475802	0.44	9	120182492	exon 3	AA change
ss2031475803	0.43	9	120183584	exon 4	miRNA Target gain
ss2031475804	0.38	9	120198443	exon 1 (5′ utr)	–
ss2031475805	0.44	9	120215042	exon 8	–
ss2031475806	0.15	13	80086134	exon 8 (3′ utr)	–
ss2031475807	0.34	13	80086582	exon 8 (3′ utr)	–
ss2031475808	0.34	13	80691943	intron (ex8–ex9)	–
ss2031475809	0.34	13	80691962	intron (ex8–ex9)	–
ss2031475810	0.22	13	80692144	intron (ex8–ex9)	ESE change
ss2031475811	0.22	13	80693764	exon 7	–
ss2031475812	0.22	13	81350505	exon 13	AA change
ss2031475813	0.49	13	82370845	exon 2	AA change
ss2031475814	0.03	13	82824584	exon 14 (3′ utr)	ESE loss
ss2031475815	0.21	13	82846842	exon 12	–
ss2031475816	0.32	13	88073665	exon 1 (5′ utr)	–
ss2031475817	0.25	17	39530285	exon 4 (3′ utr)	–
ss2031475818	0.06	17	39584558	intron (ex2–ex3)	–

Genomic localisations, minimum allele frequencies, and predicted effects are shown for candidate SNP

ESE exonic splicing enhancer

[a]Position within gene

The *LAMB1* gene encodes beta laminin 1, which is associated with skeletal muscle development in human (Wewer et al. 1997).

The *PDIA4* gene encodes a protein disulfide isomerase that is associated with HSP90 activity in muscle differentiation (Garcia de la Serrana and Johnston 2013).

The 20 SNPs were successfully genotyped in the backcrossed animals showing MAFs ranging from 0.03 for *ss2031475815* to 0.49 for *ss2031475813*, most of the SNPs showed intermediate frequencies (Table 5).

Association analysis

Most of the selected SNPs (65%) showed intermediate frequencies in the backcross population, MAF >0.25, optimal values for association analysis (Tabangin et al. 2009), and only two SNPs showed very low frequencies [*ss2031475814* (MAF = 0.03) and *ss2031475818* (MAF = 0.06)] and were discarded for the association analyses. Also linkage disequilibrium estimates were calculated for closely linked SNPs. Complete linkage was found for *ss2031475811, ss2031475810*, and *ss2031475812*, and between *ss2031475802* and *ss2031475803* polymorphisms.

Specific association analyses of each candidate SNP with the corresponding probeset expression level were conducted, in agreement with eGWAS results (Table 5). In addition, association analyses were done for the candidate SNPs with the phenotypic traits (SW, HW, BLW, BW150, IMF, BFT75, and BFS).

Significant association with gene expression measures were found for *ss2031475813, ss2031475806, ss2031475807, ss2031475816, ss2031475814, ss2031475817, ss2031475811, ss2031475809*, and *ss2031475808*. The whole results could be grouped into two different clusters conditional on the affected gene expression: Ssc.10589.1.A1_at probeset, representing ALDBSSCG0000001928 lncRNA expression and Ssc.7666.1.A1_at probeset, representing *PSMF1* gene expression.

Within the first cluster, the *ss2031475813, ss2031475806, ss2031475807, ss2031475816, ss2031475814, ss2031475811, ss2031475809, ss2031475808* SNPs showed association with the Ssc.10589.1.A1_at expression levels. All eight SNPs are located in R44-*cis*, showing a decrease of expression levels between 0.437 and 0.918 (Table 5). Only *ss2031475816* SNP, which is located in R45-*trans* (trans-association), reported an increase of *Ssc.10589.1.A1_*at expression levels in 0.779 with a standard error (SE) of ±0.083. For the second cluster, only one SNP was associated with the expression levels of *PSMF1* gene, *ss2031475817*, which increases expression in 0.224 (±0.062).

Regarding the association analysis results for the production traits, the *ss2031475809*, which showed the higher association with expression levels of SSc.10589.1.A1_ at probeset, showed also the most significant effect on BW150, increasing animal weight in 2.66 kg (±1.07) (Table 5). Additionally, it also revealed effects on HW and BLW increasing weight in 156 g (±88) and 148 g (±79), respectively. Besides *ss2031475808, ss2031475814*, and *ss2031475806* SNPs showed associations, *p* value <0.05, with BW150 trait. Suggestive effects ($p < 0.10$) of *ss2031475816* on BW150, and *ss2031475817* on HW could also be reported. Here, it should be noted that the animal number is a power limitation in the identification of significant effects in the association analysis (Hong and Park 2012).

The linkage disequilibrium estimates for *TXNRD3, MGLL, ICA*, and *RBP1* SNPs (Fig. 1) revealed a significant linkage block composed by two SNPs of *MGLL* and the three SNPs of *TXNRD3* (*ss2031475806, ss2031475807, ss2031475808, ss2031475809*, and *ss2031475810*), two

Table 5 Significant association results of the analyzed candidate SNPs on gene expression and phenotypic traits

SNP_ID[a]	Gene probeset	W150D	HW	SW	BLW
ss2031475813	−0.437 (0.127)[¥]	1.860 (1.322)	−0.041 (0.107)	−0.022 (0.057)	0.000 (0.096)
ss2031475806	−0.599 (0.153)[¥]	1.365 (1.738)	0.015 (0.139)	0.000 (0.071)	0.017 (0.123)
ss2031475807	−0.781 (0.109)[¥]	2.710 (1.316)**	0.118 (0.092)	0.053 (0.056)	0.084 (0.081)
ss2031475816	0.779 (0.083)[¥]	−1.893 (1.148)*	−0.140 (0.089)	−0.054 (0.048)	−0.046 (0.081)
ss2031475814	−0.604 (0.120)[¥]	2.667 (1.316)**	−0.056 (0.109)	−0.048 (0.057)	−0.023 (0.096)
ss2031475817	0.224 (0.062)[¥]	1.878 (1.324)	0.161 (0.095)*	0.020 (0.062)	0.041 (0.086)
ss2031475810	−0.758 (0.110)[¥]	2.279 (1.365)	−0.007 (0.113)	−0.012 (0.058)	0.020 (0.099)
ss2031475809	−0.918 (0.071)[¥]	2.665 (1.072)[¥]	0.156 (0.077)**	0.0290 (0.052)	0.148 (0.070)**
ss2031475808	−0.708 (0.108)[¥]	2.651 (1.220)**	0.096 (0.088)	0.051 (0.055)	0.068 (0.079)

Results of the association analysis carried out for the 20 candidate SNPs with weight at 150 days, ham, shoulder, and bone-in-loin weights as phenotypic traits

[¥]Bonferroni correction ($q < 0.05$), **$p < 0.05$, *$p < 0.1$

[a]Position within gene
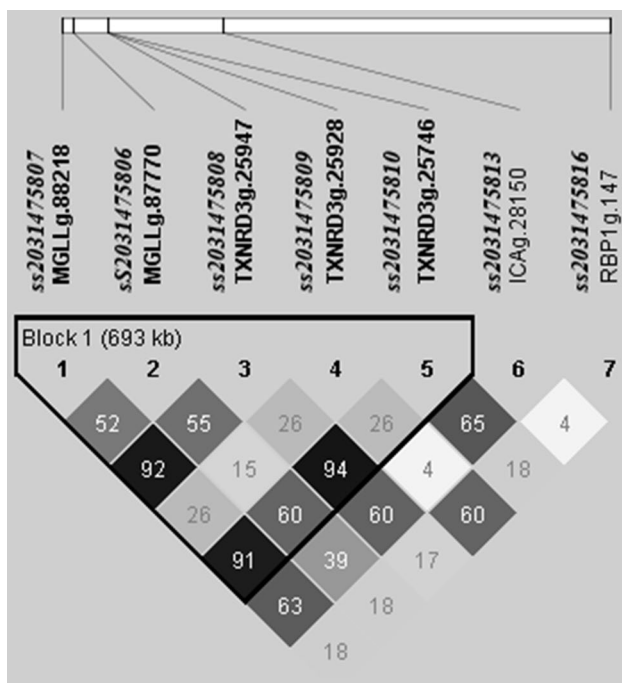
Fig. 1 Linkage disequilibrium (r^2) representation among the SNPs associated with Ssc.10589.1.A1_at expression levels

genes located very close in SSC13, at 594 kb distance. Four haplotypes were identified for these five SNPs: AAACC, AAATC, GGTTT, and GATTT. The same association analysis as those used with single SNPs were carried out for the haplotypes, trying to determine if the addition of genomic information to the analysis could better explain the effects than the individual SNPs. Nevertheless, the results were less significant. The association analysis of *ss2031475817*, the unique SNP associated with *PSMF1* gene expression, with the productive traits revealed effects on HW trait (Table 5).

In addition, few other SNPs located within these eQTL regions showing effects on production traits: *ss2031475801* showed effects on HW and BW150, *ss2031475800* on HW and SW, *ss2031475799* on HW, SW, and IMF, and *ss2031475814* on BW150 (results not shown). Some of these results may be relevant; however, due to the lack of association with probeset expression levels (the initial hypothesis of the current study), these results were not further studied.

Discussion

In the current study we focused our analyses on the identification of mutations that could affect expression levels of genes involved in porcine fat deposition and growth processes. In order to achieve this objective a genetical

genomics study was conducted using the expression levels for 776 genes selected due to the correlation between expression levels and phenotypic traits, currently known as intermediate phenotypes, correlated with fat deposition and growth-related traits, in an eGWAS. This approach allowed us to identify a total of 954 significant associations between 42 genes and 880 eTAS. Moreover, we were able to validate interesting associations between SNPs and gene expression levels, such as those identified for *Insulin-Like Growth Factor 2* (*IGF2*), R5-*trans* that contains 28 SNPs associated with expression, at SSC2:16,416–10,979,357, and R9-*trans* containing 2 SNPs SSC2:162,084,552–162,298,086, where are reported to map the causal mutation affecting *IGF2* gene expression, involved in muscle development (Van Laere et al. 2003), fatty acid composition (Hong et al. 2015), and litter size (Muñoz et al. 2010).

As expected, the identification of SNPs affecting phenotypic traits is less precise than identifying association with expression levels directly, likely due to the most direct relation between gene expression and genomic information. Gene expression seems to be regulated in a simple way if we compared it with complex phenotypic traits. Nevertheless, the interpretation and biological relevance of the associations identified here need further analyses to unravel these complex regulation mechanisms.

Beyond the identification of associations between SNPs and gene expressions, the eGWAS has allowed us to identify 63 eQTLs. Although region size and gene content seem to be variable, a lot of information could be obtained from these genomic regions. eQTLs were identified in all autosomes except SSC10, SSC12, SSC16, and SSC18. The SSC13 showed 13 different eQTL regions associated with four different gene expression levels, ten of those were associated with *TXNRD3* expression. The regions identified on SSC13 covered almost 60% of total chromosome length, which could be explained by high linkage disequilibrium levels (Fig. 2) as previously reported (Saura et al. 2015). Positional and functional candidate genes were identified in some of these regions, allowing us to select potential genes underlying the identified eQTLs. Some of the genes are transcription factors previously associated with traits of interest such as the *FOXO1*, associated with adipogenesis in porcine preadipocytes (Yan et al. 2013), the *GATA2*, involved in adipogenesis (Szczerbal and Chmurzynska 2008), and the *RBL1* (p107), which has been proposed to regulate adipocyte differentiation (Scimè et al. 2005).

One of the challenges of this kind of studies is how to manage the great amount of results obtained from the eGWAS. Although a lot of interesting regions and genes were detected, the study was focused on those regions that overlap with phenotypic QTLs previously described in the same animal material. The eQTL regions at SSC4, SSC9, SSC13, and SSC17 overlapped with QTLs for
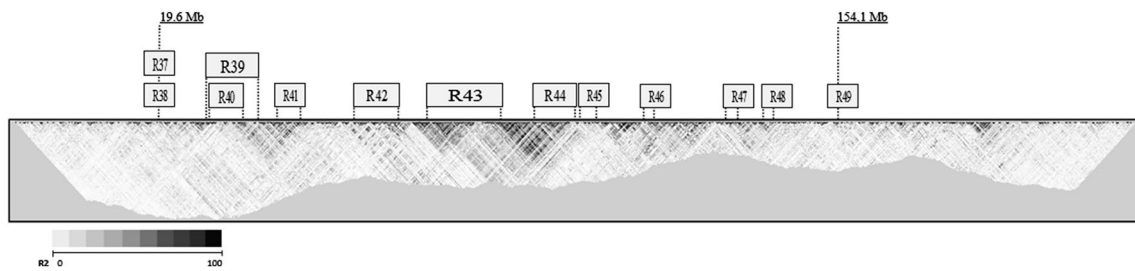
Fig. 2 Linkage disequilibrium heat map of eQTL regions on SSC13

fatness and premium cut yield (Varona et al. 2002; Fernandez et al. 2012). With this approach, some of the results remained unanalyzed but it brings interesting data for further studies.

In order to identify candidate mutations that could be underlying the selected eQTLs, SNPs identified in a previous RNA-Seq study were used (Martínez-Montes et al. 2016). This approach allowed us to select not only candidate SNPs located on those regions, but also SNPs that showed differential genotype between divergent groups for growth and fat deposition (Martínez-Montes et al. 2016). Merging both studies, SNP identification and GWAS results, we were able to focus our analysis on the identification of candidate causal mutations. The strength of this approach relies on the possibility of using different type of results in order to answer a common objective. Following this strategy, we were able to identify potential candidate genes that could be regulating the expression levels of three genes: *BUB1B*, ALDBSSCG0000001928, and *PSMF1*. Even more, candidate mutations associated with gene expression and production traits were also identified. One of the most interesting results is the association detected between *PSMF1* and *TRIB3* SNPs. Previous studies have reported the association of *TRIB3* polymorphisms with meat quality and production traits in Italian heavy pigs (Fontanesi et al. 2010) by reducing the fat levels and increasing weight. Moreover, the *PSMF1* interacts indirectly with *TRIB3* gene, via *AKT2* and *UBC* genes which have been previously associated with adipogenesis and muscle development (Pang et al. 2013; Ayuso et al. 2015). Among the most promising and novelty result is the identification of ALDBSSCG0000001928 lncRNA, whose expression seems to be associated with *TXNRD3* polymorphisms.

The analyzed *ss2031475809* could be the causal mutation affecting the expression levels of this lncRNA, and it appears to also be associated with body weight and premium cut yields. It should be also noted that several regions of SSC13 chromosome, ten different regions, are *trans*-associated with the same lncRNA expression levels (R37, R39, R41, R42, R43, R44, R45, R46, R47, and R49),

but the most significant association corresponded to the *cis*-association of R44, which includes *ss2031475809*.

The *Ssc.10589.1.A1_at* probeset was firstly annotated within *TXNRD3* gene. Nevertheless, after annotation updates and deeper sequence analysis by basic local alignment searches with BLAST, the annotation confirmed that it represents a long intergenic non-coding RNA (lncRNA) gene annotated in the domestic-animal long non-coding RNA database (ALDB) as *ALDBSSCG0000001928* gene (*ALDBSSCT0000003202* transcript). This lncRNA is located close to *TXNRD3* gene, at 3 Kb of *ss2031475809*. Annotation data showed that *ALDBSSCG000000192* gene is located within a QTL region for several productive traits such as average daily gain, body weight, and back fat weight (pigQTL database).

Long non-coding RNAs (lncRNA) have been identified as chromatin regulators in different species and act following different strategies. For instance, the *XIST* gene, which is an lncRNA upregulated in one of the female X chromosomes of mice in early development, leads to transcriptional repression and important changes in chromatin composition. Nevertheless, it acts also for dosage compensation *roX* gene in Drosophila melanogaster, increasing transcription on the single male X chromosome (Rutenberg-Schoenberg et al. 2016). But several other roles have been attributed to lncRNAs such as transcriptional regulation and post-transcriptional control (Angrand et al. 2015). In the current study, we hypothesize that *ALDBSSCG0000001928* lncRNA could be regulating expression levels, through transcriptional repression, of surrounding genes such as *MGLL* and *TXNRD3* (negative significant correlation, −0.43 and −0.42 was detected with lncRNA expression, respectively).

The potential mechanism explaining the relation between *ss2031475809* SNP and *ALDBSSCG0000001928* lncRNA was explored using Motif Comparison tool from the MEME suite. Potential motifs including or close to *ss2031475809* were identified (Fig. 3). The most relevant corresponded to the motif CAC[A/C]T[A/G]AG, which involves conservation in the *ss2031475809* position indicating its high functional relevance. Additionally,
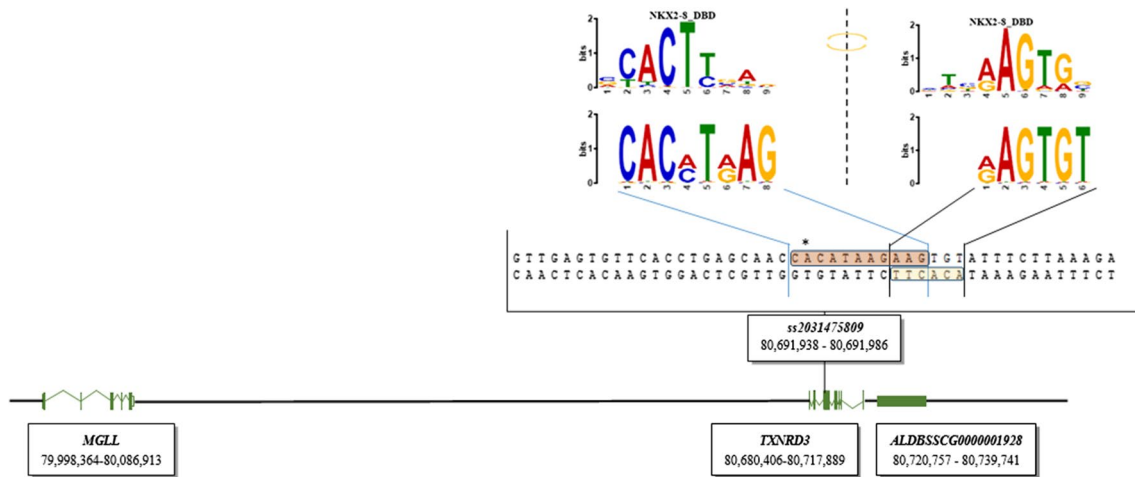
Fig. 3 Genomic organization of *MGLL, TXNRD3* (*ss2031475809*) and ALDBSSCG0000001928 lncRNA genes. Representation of NKX2-8 DBD motifs within *TXNRD3* gene sequence

two more motifs (similar to NKX2-8 DBD) were predicted (Fig. 3). Using the GOMo tool to scan promoters to determine if the identified motif is significantly associated with genes linked to one or more Genome Ontology (GO) terms, we were able to observe enrichment, among human gene catalogue, for olfactory receptor activity (GO:0004984), sensory perception of smell (GO:0007608), and sensory perception of chemical stimulus (GO:0007606). These terms involve genes such as taste receptors likely mediating growth and fatness processes (Ren et al. 2009) and olfactory receptors, which have been studied in porcine due to the possible relevance in pig over other species (Nguyen et al. 2012) and their relation with gastrointestinal tract in pigs (Priori et al. 2015). These results support *ss2031475809* as candidate mutation, to regulate *ALDBSSCG0000001928* lncRNA expression, which can be involved in the transcriptional regulation of *MGLL* and *TXNRD3*, affecting productive and meat quality traits (Pena et al. 2013; Puig-Oliveras et al. 2016).

In conclusion, we were able to identify 63 eQTL regions affecting 36 transcript expressions, which overlapped with phenotypic QTLs on chromosomes SSC4, SSC9, SSC13, and SSC17. Also candidate genes on these regions, and candidate SNPs obtained from RNA-Seq data were analyzed. One of the most relevant results is the identification of *ALDBSSCG0000001928*, a long non-coding RNA, whose expression seems to be correlated with premium cut yield. In silico domain annotation and association analysis support the role of TXNRD3 polymorphisms as potential candidates to regulate *ALDBSSCG0000001928* expression. This lncRNA could be involved in the transcriptional regulation of genes surrounding it, as other lncRNA are reported to, affecting productive and meat quality traits.

References

Angrand PO, Vennin C, Le Bourhis X, Adriaenssens E (2015) The role of long non-coding RNAs in genome formatting and expression. Front Genet 6:165

Ayuso M, Fernández A, Núñez Y, Benítez R, Isabel B, Barragán C, Fernández AI, Rey AI, Medrano JF, Cánovas Á, González-Bulnes A, López-Bote C, Ovilo C (2015) Comparative analysis of muscle transcriptome between pig genotypes identifies genes and regulatory mechanisms associated to growth, fatness and metabolism. PLoS One 10:e0145162

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:202–208

Bass JD, Swcf AJ, Dabney A, Robinson D (2015). qvalue: Q-value estimation for false discovery rate control. R package version 2.2.2

Breitling R, Li Y, Tesson BM, Fu J, Wu C, Wiltshire T, Gerrits A, Bystrykh LV, de Haan G, Su AI, Jansen RC (2008) Genetical genomics: spotlight on QTL hotspots. PLoS Genet 4:e1000232

Buske FA, Bodén M, Bauer DC, Bailey TL (2010) Assigning roles to DNA regulatory motifs using comparative genomics. Bioinformatics 26:860–866

Chang TH, Huang HY, Hsu JB, Weng SL, Horng JT, Huang HD (2013) An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. BMC Bioinform 14(Suppl 2):S4

Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Pérez-Montarelo D, Noguera JL, Folch JM, Ballester M (2013) Polymorphism in the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. PLoS One 8:e53687

Corominas J, Marchesi JA, Puig-Oliveras A, Revilla M, Estellé J, Alves E, Folch JM, Ballester M (2015) Epigenetic regulation of the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. Genet Sel Evol 47:20

Deng HW, Chen WM, Recker RR (2000) QTL fine mapping by measuring and testing for Hardy–Weinberg and linkage disequilibrium at a series of linked marker loci in extreme samples of populations. Am J Hum Genet 66:1027–1045

Espigolan R, Baldi F, Boligon AA, Souza FR, Fernandes Júnior GA, Gordo DG, Venturini GC, de Camargo GM, Feitosa FL, Garcia DA, Tonhati H, Chardulo LA, Oliveira HN, Albuquerque LG (2015) Associations between single nucleotide polymorphisms and carcass traits in Nellore cattle using high-density panels. Genet Mol Res 14:11133–11144

Estellé J, Pérez-Enciso M, Mercadé A, Varona L, Alves E, Sánchez A, Folch JM (2006) Characterization of the porcine FABP5 gene and its association with the FAT1 QTL in an Iberian by Landrace cross. Anim Genet 37:589–591

Fernández AI, Pérez-Montarelo D, Barragán C, Ramayo-Caldas Y, Ibáñez-Escriche N, Castelló A, Noguera JL, Silió L, Folch JM, Rodríguez MC (2012) Genome-wide linkage analysis of QTL for growth and body composition employing the PorcineSNP60 BeadChip. BMC Genet 13:41

Fernández AI, Muñoz M, Alves E, Folch JM, Noguera JL, Enciso MP, del Rodríguez MC, Silió L (2014) Recombination of the porcine X chromosome: a high density linkage map. BMC Genet 15:148

Fontanesi L, Colombo M, Scotti E, Buttazzoni L, Bertolini F, Dall'Olio S, Davoli R, Russo V (2010) The porcine tribbles homolog 3 (TRIB3) gene: identification of a missense mutation and association analysis with meat quality and production traits in Italian heavy pigs. Meat Sci 86:808–813

Garcia de la Serrana D, Johnston IA (2013) Expression of heat shock protein (Hsp90) paralogues is regulated by amino acids in skeletal muscle of Atlantic salmon. PLoS One 8:e74295

Guntani A, Matsumoto T, Kyuragi R, Iwasa K, Onohara T, Itoh H, Katusic ZS, Maehara Y (2011) Reduced proliferation of aged human vascular smooth muscle cells–role of oxygen-derived free radicals and BubR1 expression. J Surg Res 170:143–149

Guo K, Wang J, Andrés V, Smith RC, Walsh K (1995) MyoD-induced expression of p21 inhibits cyclin-dependent kinase activity upon myocyte terminal differentiation. Mol Cell Biol 15:3823–3829

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS (2007) Quantifying similarity between motifs. Genome Biol 8:R24

Hill WG (2012) Quantitative genetics in the genomics era. Curr Genom 13:196–206

Hong J, Kim D, Cho K, Sa S, Choi S, Kim Y, Park J, Schmidt GS, Davis ME, Chung H (2015) Effects of genetic variants for the swine FABP3, HMGA1, MC4R, IGF2, and FABP4 genes on fatty acid composition. Meat Sci 110:46–51

Hu ZL, Dracheva S, Jang W, Maglott D, Bastiaansen J, Rothschild MF, Reecy JM (2005) A QTL resource and comparison tool for pigs: PigQTLDB. Mamm Genome 16:792–800

Jansen RC, Nap JP (2001) Genetical genomics: the added value from segregation. Trends Genet 17:388–391

Karisa BK, Thomson J, Wang Z, Stothard P, Moore SS, Plastow GS (2013) Candidate genes and single nucleotide polymorphisms associated with variation in residual feed intake in beef cattle. J Anim Sci 91:3502–3513

Karssen LC, van Duijn CM, Aulchenko YS (2016) The GenABEL Project for statistical genomics. F1000Res 5:914

Kipp AP, Müller MF, Göken EM, Deubel S, Brigelius-Flohé R (2012) The selenoproteins GPx2, TrxR2 and TrxR3 are regulated by Wnt signalling in the intestinal epithelium. Biochim Biophys Acta 1820:1588–1596

Kodama K, Horikoshi M, Toda K, Yamada S, Hara K, Irie J, Sirota M, Morgan AA, Chen R, Ohtsu H, Maeda S, Kadowaki T, Butte AJ (2012) Expression-based genome-wide association study links the receptor CD44 in adipose tissue with type 2 diabetes. Proc Natl Acad Sci USA 109:7049–7054

Lahaye LL, Wouda RR, de Jong AW, Fradkin LG, Noordermeer JN (2012) WNT5 interacts with the Ryk receptors doughnut and derailed to mediate muscle attachment site selection in *Drosophila melanogaster*. PLoS One 7:e32297

Li A, Zhang J, Zhou Z, Wang L, Liu Y, Liu Y (2015) ALDB: a domestic-animal long noncoding RNA database. PLoS One 10:e0124003

Martínez-Montes AM, Fernández A, Pérez-Montarelo D, Alves E, Benítez RM, Nuñez Y, Óvilo C, Ibañez-Escriche N, Folch JM, Fernández AI (2016) Using RNA-Seq SNP data to reveal potential causal mutations related to pig production traits and RNA editing. Anim Genet. doi:10.1111/age.12507

Mercadé A, Estellé J, Noguera JL, Folch JM, Varona L, Silió L, Sánchez A, Pérez-Enciso M (2005) On growth, fatness, and form: a further look at porcine chromosome 4 in an Iberian×Landrace cross. Mamm Genome 16:374–382

Muñoz M, Fernández AI, Ovilo C, Muñoz G, Rodriguez C, Fernández A, Alves E, Silió L (2010) Non-additive effects of RBP4, ESR1 and IGF2 polymorphisms on litter size at different parities in a Chinese-European porcine line. Genet Sel Evol 42:23

Nguyen DT, Lee K, Choi H, Choi MK, Le MT, Song N, Kim JH, Seo HG, Oh JW, Lee K, Kim TH, Park C (2012) The complete swine olfactory subgenome: expansion of the olfactory gene repertoire in the pig genome. BMC Genom 13:584

Ovilo C, Pérez-Enciso M, Barragán C, Clop A, Rodríquez C, Oliver MA, Toro MA, Noruera JL (2000) A QTL for intramuscular fat and backfat thickness is located on porcine chromosome 6. Mamm Genome 11:344–346

Óvilo C, Fernández A, Noguera JL, Barragán C, Letón R, Rodríguez C, Mercadé A, Alves E, Folch JM, Varona L, Toro MA (2005) Fine mapping of porcine chromosome 6 QTL and LEPR effects on body composition in multiple generations of an Iberian by Landrace intercross. Genet Res 85:57–67

Pang W, Wang Y, Wei N, Xu R, Xiong Y, Wang P, Shen Q, Yang G (2013) Sirt1 inhibits akt2-mediated porcine adipogenesis potentially by direct protein–protein interaction. PLoS One 8:e71576

Hong EP1, Park JW (2012) Sample size and statistical power calculation in genetic association studies. Genom Inform 10:117–122

Pena RN, Noguera JL, Casellas J, Díaz I, Fernández AI, Folch JM, Ibáñez-Escriche N (2013) Transcriptional analysis of intramuscular fatty acid composition in the longissimus thoracis muscle of Iberian×Landrace back-crossed pigs. Anim Genet 44:648–660

Pérez-Enciso M, Misztal I (2011) Qxpak.5: old mixed model solutions for new genomics problems. BMC Bioinforma 12:202

Pérez-Montarelo D, Fernández A, Folch JM, Pena RN, Ovilo C, Rodríguez C, Silió L, Fernández AI (2012) Joint effects of porcine leptin and leptin receptor polymorphisms on productivity and quality traits. Anim Genet 43:805–809

Priori D, Colombo M, Clavenzani P, Jansman AJ, Lallès JP, Trevisi P, Bosi P (2015) The olfactory receptor OR51E1 is present along the gastrointestinal tract of pigs, co-localizes with enteroendocrine cells and is modulated by intestinal microbiota. PLoS One 10:e0129501

Puig-Oliveras A, Revilla M, Castelló A, Fernández AI, Folch JM, Ballester M (2016) Expression-based GWAS identifies variants, gene interactions and key regulators affecting intramuscular fatty acid content and composition in porcine meat. Sci Rep 6:31803

Qin LL, Li XK, Xu J, Mo DL, Tong X, Pan ZC, Li JQ, Chen YS, Zhang Z, Wang C, Long (2012) QM Mechano growth factor (MGF) promotes proliferation and inhibits differentiation of porcine satellite cells (PSCs) by down-regulation of key myogenic transcriptional factors. Mol Cell Biochem 370:221–230

Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu ZL, Kerstens HH, Law AS, Megens HJ, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TP, Schnabel RD, Van Tassell CP, Taylor JF, Wiedmann RT, Schook LB, Groenen MA (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS One 4:e6524

Ren X, Zhou L, Terwilliger R, Newton SS, de Araujo IE (2009) Sweet taste signaling functions as a hypothalamic glucose sensor. Front Integr Neurosci 3:12

Rutenberg-Schoenberg M, Sexton AN, Simon MD (2016) The properties of long noncoding RNAs that regulate chromatin. Annu Rev Genom Hum Genet 17:69–94

Saura M, Tenesa A, Woolliams JA, Fernández A, Villanueva B (2015) Evaluation of the linkage-disequilibrium method for the estimation of effective population size when generations overlap: an empirical case. BMC Genom 16:922

Schön CC, Utz HF, Groh S, Truberg B, Openshaw S, Melchinger AE (2004) Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167:485–498

Scimè A, Grenier G, Huh MS, Gillespie MA, Bevilacqua L, Harper ME, Rudnicki MA (2005) Rb and p107 regulate preadipocyte differentiation into white versus brown fat through repression of PGC-1alpha. Cell Metab 2:283–295

Stephens M, Smith N, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Sun R, Chang Y, Yang F, Wang Y, Li H, Zhao Y, Chen D, Wu T, Zhang X, Han Z (2015) A dense SNP genetic map constructed using restriction site-associated DNA sequencing enables detection of QTLs controlling apple fruit quality. BMC Genom 16:747

Szczerbal I, Chmurzynska A (2008) Chromosomal localization of nine porcine genes encoding transcription factors involved in adipogenesis. Cytogenet Genome Res 121:150–154

Tabangin ME, Woo JG, Martin LJ (2009) The effect of minor allele frequency on the likelihood of obtaining false positives. BMC Proc 3(Suppl 7):S41

Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, Andersson G, Georges M, Andersson L (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature 425:832–836

Varona L, Ovilo C, Clop A, Noguera JL, Pérez-Enciso M, Coll A, Folch JM, Barragán C, Toro MA, Babot D, Sánchez A (2002) QTL mapping for growth and carcass traits in an Iberian by Landrace pig intercross: additive, dominant and epistatic effects. Genet Res 80:145–154

Wang D, Lemon WJ, You M (2002) Linkage disequilibrium mapping of novel lung tumor susceptibility quantitative trait loci in mice. Oncogene 21:6858–6865

Wewer UM, Thornell LE, Loechel F, Zhang X, Durkin ME, Amano S, Burgeson RE, Engvall E, Albrechtsen R, Virtanen I (1997) Extrasynaptic location of laminin beta 2 chain in developing and adult human skeletal muscle. Am J Pathol 151:621–631

Williams RB, Chan EK, Cowley MJ, Little PF (2007) The influence of genetic variation on gene expression. Genome Res 17:1707–1716

Würschum T, Kraft T (2014) Cross-validation in association mapping and its relevance for the estimation of QTL parameters of complex traits. Heredity (Edinb) 112:463–468

Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Hönigschmid P, Schafferhans A, Roos M, Bernhofer M (2014) PredictProtein—an open resource for online prediction of protein structural and functional features. Nucleic Acids Res 42(Web Server issue):W337–W343

Yan X, Weijun P, Ning W, Yu W, Wenkai R, Gongshe Y (2013) Knockdown of both FoxO1 and C/EBPβ promotes adipogenesis in porcine preadipocytes through feedback regulation. Cell Biol Int 37:905–916

Zizola CF, Frey SK, Jitngarmkusol S, Kadereit B, Yan N, Vogel S (2010) Cellular retinol-binding protein type I (CRBP-I) regulates adipogenesis. Mol Cell Biol 30:3412–3420

Zou F, Chai HS, Younkin CS, Allen M, Crook J, Pankratz VS, Carrasquillo MM, Rowley CN, Nair AA, Middha S, Maharjan S, Nguyen T, Ma L, Malphrus KG, Palusak R, Lincoln S, Bisceglio G, Georgescu C, Kouri N, Kolbert CP, Jen J, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA, Schellenberg GD; Alzheimer's Disease Genetics Consortium, Petersen RC, Graff-Radford NR, Dickson DW, Younkin SG, Ertekin-Taner N (2012) Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. PLoS Genet 8:e1002707