Genome-wide association study identifies markers associated with carcass and meat quality traits in Italian Large White pigs

Maria Chiara Fabbri[*,§], Martina Zappaterra[*], Roberta Davoli[*], Paolo Zambonelli[*]

[*]Department of Agricultural and Food Sciences (DISTAL), Bologna University, Viale G. Fanin 46, 40127, Bologna, Italy.

[§] Department of Agriculture, Food, Environment and Forestry (DAGRI), Via delle Cascine 5, 50144, Firenze, Italy.

Maria Chiara Fabbri: mariachiara.fabbri@unifi.it

Martina Zappaterra: martina.zappaterra2@unibo.it

Roberta Davoli: roberta.davoli@unibo.it

Paolo Zambonelli: paolo.zambonelli@unibo.it

To whom correspondence should be addressed: paolo.zambonelli@unibo.it, roberta.davoli@unibo.it

Summary

A GWAS was performed using the genotypes obtained by PorcineSNP60 v2 BeadChip and 11 phenotypic traits (carcass lean meat percentage; backfat thickness; Longissimus thoracis muscle thickness; lightness; backfat thickness measured with caliper at the midline; meat pH measured at about 1 h post mortem and 24 h post mortem; CIE *L**, *a**, and *b** color parameters; and water-holding capacity). Three markers were associated with three of the phenotypic traits considered: M1GA0009592 (SSC7) with backfat thickness and lean meat content, DIAS0002910 (SSC6) and ALGA0109856 (SSC6) with water-holding capacity. The marker M1GA0009592, associated with backfat thickness, lies in a QTL region near the gene *JARID2*, which is a transcription factor also involved in the regulation of adipose-derived stem cell pluripotency. The results seem to indicate a possible role of these genomic regions in the regulation of pig carcass fatness (i.e. backfat at last rib) and water-holding capacity.

Keywords

carcass traits, genetic markers, GWAS, meat quality, swine

Running head

SSC7 is associated with pork quality

Meat color, meat and carcass fat content and water-holding capacity (WHC) are parameters that strongly influence pig product organoleptic quality and have significant economic value for the meat processing industry (Ciobanu *et al*., 2011). Several candidate genes associated with color and WHC have already been reported, but to date, the knowledge of associations with QTL regions affecting these traits in the Italian Large White (ILW) pig breed is still lacking (https://www.animalgenome.org/cgi-bin/QTLdb/SS/index; Hu *et al*., 2019). A list of the significant

QTL for backfat thickness, lean meat content, meat pH, color and drip loss/WHC was reported in the Pig QTLdb for Large White/Yorkshire breeds (Table S1).

This research aimed to identify genetic markers associated with carcass and meat quality traits in a purebred population of 888 ILW pigs reared in the same environmental conditions at the Italian National Association of Pig Breeders Sib-test station.

Animal care and slaughter were performed in compliance with the European rules (Council Regulation (EC) No. 1/2005 and Council Regulation (EC) No. 1099/2009). All slaughter procedures were monitored by the veterinary team appointed by the Italian Ministry of Health. The ILW pigs were slaughtered at about 155 kg of live weight. A detailed description of the utilized method is reported in Table S2.

The genome-wide and chromosome-wide significant markers found by GWAS are reported in Table 1 and Fig. 1. The protein-coding genes detected in the region ±500 kb from each significant marker are reported in Table S3. In Table 1 are reported the False Discovery Rate (FDR)-adjusted *P*-values to correct for false positives. We decided to use FDR correction for multiple tests according to information retrieved from the literature indicating that the Bonferroni method is very restrictive in GWAS studies to correct for Type I errors (Brinster *et al*., 2018). The FDR adjustment indicated that three SNPs had FDR-adjusted *P*-values <0.10. The markers DIAS0002910 and ALGA0109856 are 105,366 bp apart, and are both located in SSC6 within a genomic region that contains the genes *Cilia and flagella associated protein 20* (*CFAP20*) and *Coiled-coil domain containing 113* (*CCDC113*). In the region spanning ±500 kb apart from the two markers associated with WHC there are 23 protein-coding genes. Interestingly, the function of some genes (*CCD113*, *CFAP20*, *KIFC3*, *KATNB1*), all located in the same genomic region (Table S3), can be related to cilia or microtubule functionality. In humans, primary cilia were recently reported to be involved in muscle development and energy homeostasis (Fu *et al*., 2014), and the expression levels of genes related to these organelles were recently found to be associated with intramuscular fat deposition in pigs (Zappaterra *et al*., 2020). In pigs, these associations are poorly known and further studies are

needed to elucidate a possible role of cilia and/or microtubules in muscle tissue development and in the cell functions related to WHC.

The most genome-wide significant SNP was located on SSC7 (M1GA0009592) near *Jumonji, AT rich interactive domain 2* (*JARID2*) gene. The marker M1GA0009592 and *JARID2* gene are located in a genomic region with two very large QTL related to porcine backfat thickness (Pig QTLdb: ID=308, Rattink *et al*., 2000; ID=3768, Nagamine *et al*., 2003). Moreover, a refined QTL region (Pig QTLdb: ID=9901) was reported by Nagamine *et al*. (2009) on the same populations used in the previous experiment (Table S1). The gene *JARID2* is a transcription factor that was found to be expressed in human adipose-derived stem cells in response to thyroid hormone receptor actions (Cvoro *et al*., 2016).

The obtained results suggest new associations between these markers and genes that have been up to now poorly studied with respect to traits important for heavy pig production. These associations seem to indicate that molecular processes influenced by the identified genes may also have an effect on carcass fat content, lean meat production and WHC. These findings should be investigated in more depth to better understand the hypothesized effects and, if validated, these markers could be considered in pig selection.

Acknowledgements

This work was supported by PRIN 2015 national project (no. 201549TZXB001) and by Bologna University RFO funds.

The authors declare that they have no competing interests.

Availability of data

The data are available after signing a Material Transfer Agreement with the corresponding authors.

References

Brinster R., Köttgen A., Tayo B.O., Schumacher M., Sekula P., CKDGen Consortium. (2018) Control procedures and estimators of the false discovery rate and their application in low-dimensional settings: an empirical investigation. BMC Bioinformatics 19, 78.

Ciobanu D.C., Lonergan S.M., Huff-Lonergan E.J. (2011) Genetics of Meat Quality and Carcass Traits. In: The Genetics of the Pig (ed. by M.F. Rothschild & A. Ruvinsky), pp. 355–389. CABI Publishing, Oxfordshire, UK.

Cvoro A., Bajic A., Zhang A., Simon M., Golic I., Sieglaff D.H., Maletic-Savatic M., Korac A., Webb, P. (2016) Ligand Independent and Subtype-Selective Actions of Thyroid Hormone Receptors in Human Adipose Derived Stem Cells. PLoS One 11, e0164407.

Fu W., Asp P., Canter B., Dynlacht, B.D. (2014) Primary cilia control hedgehog signaling during muscle differentiation and are deregulated in rhabdomyosarcoma. PNAS 111, 9151-9156.

Hu Z.-L., Park C.A., Reecy J.M. (2019) Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Research 47, D701–D710.

Nagamine Y., Haley C.S., Sewalem A., Visscher P.M. (2003) Quantitative trait loci variation for growth and obesity between and within lines of pigs (Sus scrofa). Genetics 164, 629-635.

Nagamine Y., Pong-Wong R., Visscher P.M., Haley C.S. (2009) Detection of multiple quantitative trait loci and their pleiotropic effects in outbred pig populations. Genetics, Selection, Evolution 41, 44.

Pig QTL Database, Available online https://www.animalgenome.org/cgi-bin/QTLdb/SS/index (accessed 12.5.19).

Rattink A.P., De Koning D.J., Faivre M., Harlizius B., van Arendonk J.A., Groenen M.A. (2000) Fine mapping and imprinting analysis for fatness trait QTLs in pigs. Mammalian Genome 11, 656-661.

Zappaterra M., Gioiosa S., Chillemi G., Zambonelli P., Davoli R. (2020) Muscle transcriptome analysis identifies genes involved in ciliogenesis and the molecular cascade associated with intramuscular fat content in Large White heavy pigs. PLoS ONE 15, e0233372.

Table 1 Significant markers identified with their location, adjusted *P-value*, and SNP position relative to the nearest gene.

SNP (variant ID)	SSC[1]	Location (bp)[2]	Pc1df	FDR	MAF	Phenotypic traits associated with the marker	SNP additive effect	Nearest genes[3]	SNP position relative to the nearest gene
DIAS0002910	6	19,956,188	2.22E-06	0.0646	0.43	WHC	0.009	*CFAP20*	Synonymous variant
ALGA0109856	6	20,061,554	2.86E-06	0.0646	0.24	WHC	NS [0.055]	*CCDC113*	Intron variant
M1GA0009592	7	10,907,559	**8.34E-06**, 3.71E-06	0.0377, 0.1677	0.37	**BF**, LM	0.005, 0.012	*JARID2*	Intergenic variant

The trait passing the genome-wide threshold is indicated in bold.

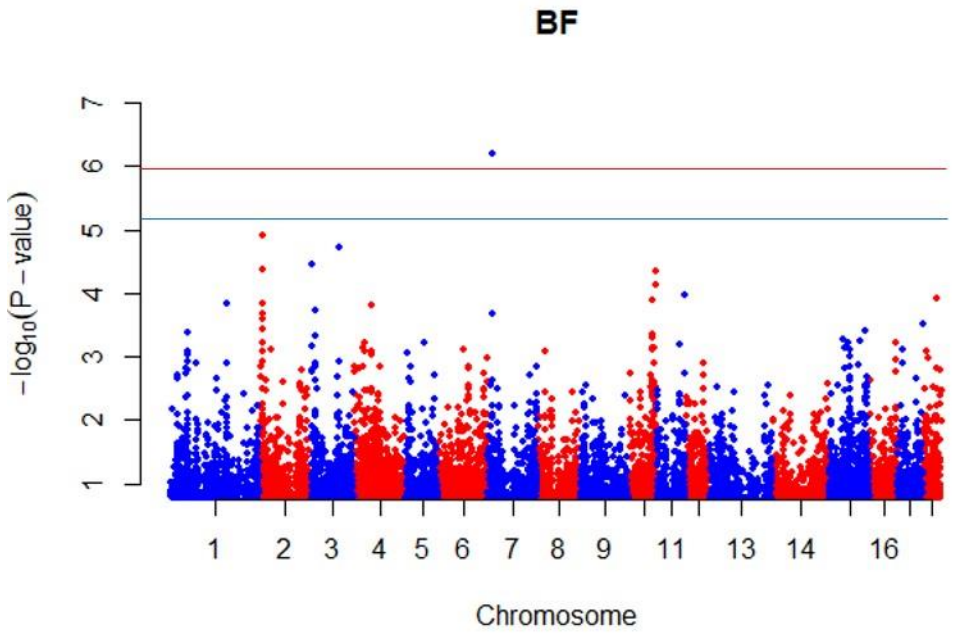
NS: not significant.

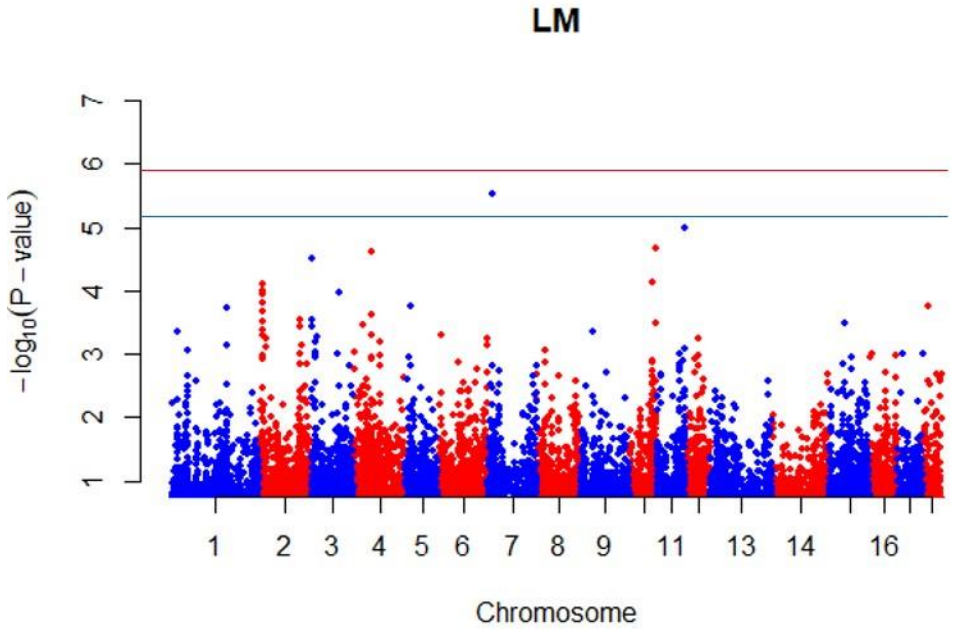
[1] *Sus scrofa* chromosome.

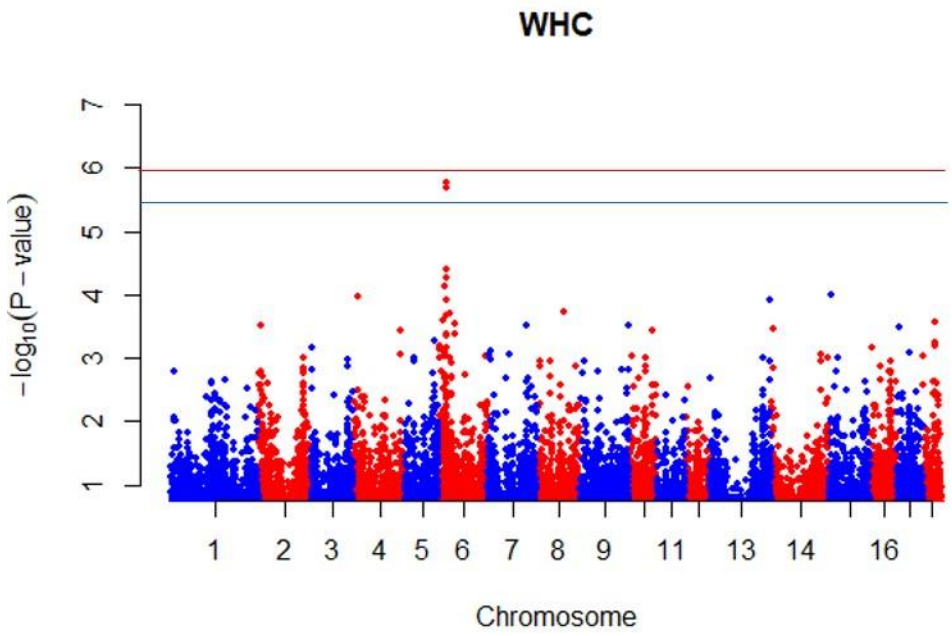
[2] SNP positions referred to *Sus scrofa* assembly Build 11.1, expressed in bp.

[3] The genes closest to the identified SNP named with the official gene symbol.

Figure 1 Manhattan plot showing the GWAS significance for the associations between the SNPs and the considered phenotypic traits. The red line indicates the genome-wide threshold of significance while the blue line indicates the chromosome-wide threshold of significance calculated for the chromosomes where the relevant markers map







BF: backfat thickness measured with FOM between the third and fourth last rib at 8 cm off the midline.

LM: carcass lean meat percentage.

WHC: water-holding capacity.

Supporting Information

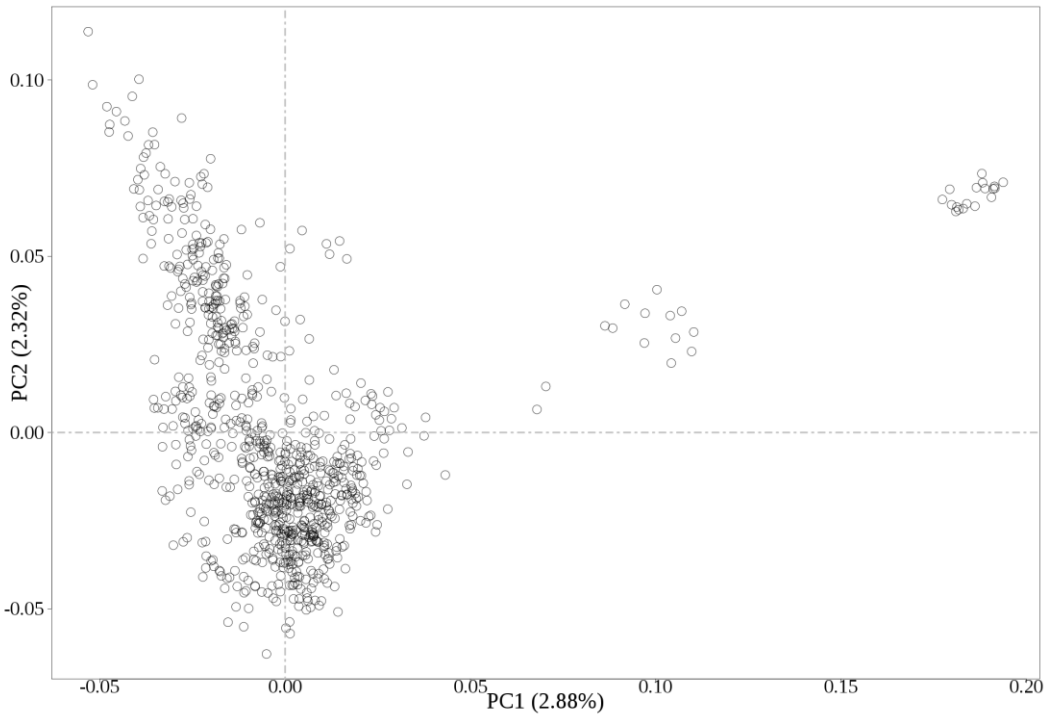
Figure S1. The population structure investigated with PCA

Table S1. List of the significant QTLs for backfat thickness, lean meat content, meat pH, color and drip loss/water holding capacity reported in Pig QTLdb for Large White/Yorkshire breeds (File TS1.xlsx)

Table S2. Supplementary Materials and methods

Table S3. Protein-coding genes included in the three significant chromosome regions

Figure S1. The population structure investigated with PCA



168 **Table S2.** Supplementary Materials and methods

169

Phenotypes measurement
We used a FAT-O meat'er (FOM) device to measure carcass lean meat percentage (LM), backfat thickness measured between the third and fourth last rib at 8 cm off the midline (BF), *Longissimus thoracis* muscle thickness (MT), and lightness (RW). Backfat thickness was recorded with a caliber at the midline at the level of *Gluteus medius* muscle (BFT). Moreover, we also determined pH1 (meat pH measured 1 h *postmortem*) and pHu (meat pH measured 24 h *postmortem*), while CIE L*, a*, b* parameters were estimated with a Chroma Meter CR-300 (Konica Minolta Sensing Inc., Osaka, Japan). Finally, WHC was calculated with filter paper press method (Hofmann *et al*., 1982).
DNA extraction and Genotyping
DNA was isolated from *Semimembranosus* muscle using the Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA). Genotyping was carried out using PorcineSNP60 v2 BeadChip (Illumina Inc., San Diego, California, USA) containing 61,565 SNPs, whose coordinates were updated to *Sus scrofa* genome assembly Build 11.1.
Statistical analyses
After PLINK (Purcell *et al*., 2007) filtering according to Nicolazzi *et al*. (2015), and after removing SNPs unmapped or located on sex chromosomes, remained 38,147 SNPs. Following the filtering carried out using PLINK a PCA was performed with the same software using the –pca flag. To visualize the results, a scatterplot of the first and second principal components has been created using "car" and "devtools" R packages. The population structure investigated with PCA showed the homogeneity of the samples (Figure S1). A further quality control was performed with the GenABEL package in the R environment (Aulchenko et al., 2007): samples with call rate <90%, SNPs with a GENO <90%, SNPs with ah Hardy-Weinberg equilibrium *P*-value <0.001, and SNPs with minor allele frequency <5% were excluded. The remaining 885

individuals and 38,111 SNPs were used to perform the Genome-Wide Association study (GWA). The following additive polygenic model was fitted with a genomic relationship matrix in GenABEL:

$Y_i = X_i\beta_i + Z_i a_i + e_i$

where Y_i is the observation vector for the ith trait; β is the vector of effects for three factors (sex: two levels for barrows and gilts; slaughtering date: 27 levels; age at slaughtering as a covariate). The random factors in the model were animal (a) and residuals (e). They were assumed to be normally distributed as $a_i \sim N(0,G\sigma_a^2)$ and $e_i \sim N(0,I\sigma_e^2)$, where G is the genomic relationship matrix and σ_a^2 and σ_e^2 the additive genomic and residual variances, respectively. Pc1df value was utilized according to Nicolazzi *et al*. (2015) and markers were considered genome-wide significant for *P*-adjusted <1.31E-06. The chromosome-wide threshold considering a *P*-adjusted <0.01 calculated for SSC6 is 4.13E-06 (2419 SNPs on SSC6). The correction for multiple tests was performed using the procedure MULTTEST using the SAS software v. 9.4 (SAS Inst., Cary, NC) and applying the False Discovery Rate (FDR) method.

The additive and dominant effects of the significant markers were calculated using the SAS software v. 9.4 (SAS Inst., Inc., Cary, NC) using General Linear Model (GLM) procedure with a model including sex, slaughtering date, age at slaughtering, and genotype as already carried out for the GenAbel analysis.

References

Aulchenko Y.S., Ripke S., Isaacs A., van Duijn C.M. (2007) GenABEL: an R library for genome-wide association analysis. Bioinforma. Oxf. Engl. 23, 1294–1296.

Hofmann K., Hamm R., Blüchel E. (1982) Neues über die Bestimmung der Wasserbindung des Fleisches mit Hilfe der Filter papier preß methode. Fleischwirtschaft 62, 87–94.

Nicolazzi E.L., Biffani S., Biscarini F., Orozco Ter Wengel P., Caprera A., Nazzicari N., Stella A. (2015) Software solutions for the livestock genomics SNP array revolution. Animal Genetics 46, 343-353.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics, 81, 559-575.

170

Table S3. Protein coding genes included in the three significant chromosome regions

SSC	SNP	Location (bp)	Gene symbol	Gene name	Gene location (from bp to bp)
6			*CCDC102A*	coiled-coil domain containing 102A	19,437,169-19,456,905
			ADGRG5	adhesion G protein-coupled receptor G5	19,463,018-19,488,665
			ADGRG1	adhesion G protein-coupled receptor G1	19,546,950-19,570,893
			ADGRG3	adhesion G protein-coupled receptor G3	19,574,514-19,599,082
			DRC7	dynein regulatory complex subunit 7	19,603,399-19,629,174
			KATNB1	katanin regulatory subunit B	19,630,444-19,667,676
			KIFC3	kinesin family member C3	19,659,987-19,730,359
			CNGB1	cyclic nucleotide gated channel subunit beta 1	19,753,036-19,825,631
			TEPP	testis, prostate and placenta expressed	19,832,473-19,838,181
			ZNF319	zinc finger protein 319	19,842,582-19,851,288
			USB1	U6 snRNA biogenesis phosphodiesterase 1	19,852,011-19,874,021
			MMP15	matrix metallopeptidase 15	19,890,884-19,918,244
	DIAS0002910	19,956,188	***CFAP20***	cilia and flagella associated protein 20	**19,854,236-19,969,347**
			CSNK2A2	casein kinase 2 alpha 2	19,976,931-20,016,741
	ALGA0109856	20,061,554	***CCDC113***	coiled-coil domain containing 113	**20,052,651-20,088,822**
			PRSS54	serine protease 54	20,083,930-20,099,885
			GINS3	GINS complex subunit 3	20,162,680-20,170,263
			NDRG4	NDRG family member 4	20,223,871-20,265,572
			SETD6	SET domain containing 6, protein lysine methyltransferase	20,266,552-20,270,694
			CNOT1	CCR4-NOT transcription complex subunit 1	20,318,946-20,372,280
			SLC38A7	solute carrier family 38 member 7	20,337,137-20,401,684
			ENSSSCG00000037660	protein coding gene	20,409,267-20,415,239
			GOT2	glutamic-oxaloacetic transaminase 2	20,407,198-20,432,068
7	M1GA0009592	10,907,559	***JARID2***	jumonji and AT-rich interaction domain containing 2	**11,357,961-11,602,104**