



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

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

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Fabbri M.C., Zappaterra M., Davoli R., Zambonelli P. (2020). Genome-wide association study identifies markers associated with carcass and meat quality traits in Italian Large White pigs. *ANIMAL GENETICS*, 51(6), 950-952 [10.1111/age.13013].

Availability:

This version is available at: <https://hdl.handle.net/11585/853874> since: 2022-02-08

Published:

DOI: <http://doi.org/10.1111/age.13013>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Fabbri, M.C., Zappaterra, M., Davoli, R. and Zambonelli, P. (2020), Genome-wide association study identifies markers associated with carcass and meat quality traits in Italian Large White pigs. *Anim Genet*, 51: 950-952.

The final published version is available online at:

<https://doi.org/10.1111/age.13013>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

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

3

4 Maria Chiara Fabbri^{*,§}, Martina Zappaterra^{*}, Roberta Davoli^{*}, Paolo Zambonelli^{*}

5

6 ^{*}Department of Agricultural and Food Sciences (DISTAL), Bologna University, Viale G. Fanin 46,
7 40127, Bologna, Italy.

8 [§] Department of Agriculture, Food, Environment and Forestry (DAGRI), Via delle Cascine 5,
9 50144, Firenze, Italy.

10

11 Maria Chiara Fabbri: mariachiara.fabbri@unifi.it

12 Martina Zappaterra: martina.zappaterra2@unibo.it

13 Roberta Davoli: roberta.davoli@unibo.it

14 Paolo Zambonelli: paolo.zambonelli@unibo.it

15

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

17

18

19 **Summary**

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

31

32 **Keywords**

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

34

35 **Running head**

36 SSC7 is associated with pork quality

37

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

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

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

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

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

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

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

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

86

87 **Acknowledgements**

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

90

91 The authors declare that they have no competing interests.

92

93 **Availability of data**

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

95

96 **References**

- 97 Brinster R., Köttgen A., Tayo B.O., Schumacher M., Sekula P., CKDGen Consortium. (2018)
98 Control procedures and estimators of the false discovery rate and their application in low-
99 dimensional settings: an empirical investigation. *BMC Bioinformatics* 19, 78.
- 100 Ciobanu D.C., Lonergan S.M., Huff-Lonergan E.J. (2011) Genetics of Meat Quality and Carcass
101 Traits. In: *The Genetics of the Pig* (ed. by M.F. Rothschild & A. Ruvinsky), pp. 355–389. CABI
102 Publishing, Oxfordshire, UK.
- 103 Cvorovic A., Bajic A., Zhang A., Simon M., Golic I., Sieglaff D.H., Maletic-Savatic M., Korac A.,
104 Webb, P. (2016) Ligand Independent and Subtype-Selective Actions of Thyroid Hormone
105 Receptors in Human Adipose Derived Stem Cells. *PLoS One* 11, e0164407.
- 106 Fu W., Asp P., Canter B., Dynlacht, B.D. (2014) Primary cilia control hedgehog signaling during
107 muscle differentiation and are deregulated in rhabdomyosarcoma. *PNAS* 111, 9151-9156.
- 108 Hu Z.-L., Park C.A., Reecy J.M. (2019) Building a livestock genetic and genomic information
109 knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids*
110 *Research* 47, D701–D710.
- 111 Nagamine Y., Haley C.S., Sewalem A., Visscher P.M. (2003) Quantitative trait loci variation for
112 growth and obesity between and within lines of pigs (*Sus scrofa*). *Genetics* 164, 629-635.
- 113 Nagamine Y., Pong-Wong R., Visscher P.M., Haley C.S. (2009) Detection of multiple quantitative
114 trait loci and their pleiotropic effects in outbred pig populations. *Genetics, Selection, Evolution*
115 41, 44.
- 116 Pig QTL Database, Available online <https://www.animalgenome.org/cgi-bin/QTLdb/SS/index>
117 (accessed 12.5.19).
- 118 Rattink A.P., De Koning D.J., Faivre M., Harlizius B., van Arendonk J.A., Groenen M.A. (2000)
119 Fine mapping and imprinting analysis for fatness trait QTLs in pigs. *Mammalian Genome* 11, 656-
120 661.

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

124

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

126

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	<i>CFAP20</i>	Synonymous variant
ALGA0109856	6	20,061,554	2.86E-06	0.0646	0.24	WHC	NS [0.055]	<i>CCDC113</i>	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	<i>JARID2</i>	Intergenic variant

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

128 NS: not significant.

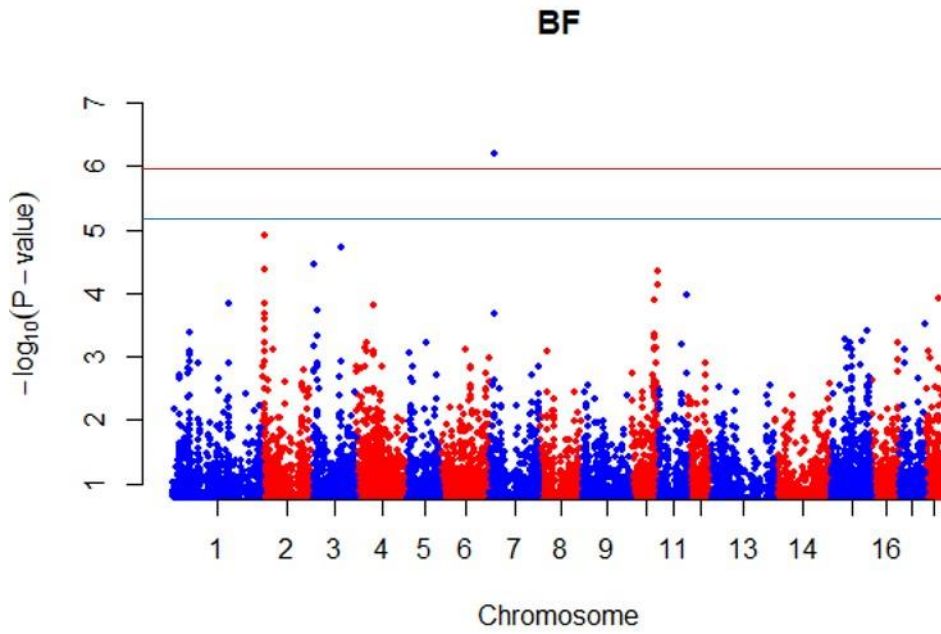
129 ¹ *Sus scrofa* chromosome.

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

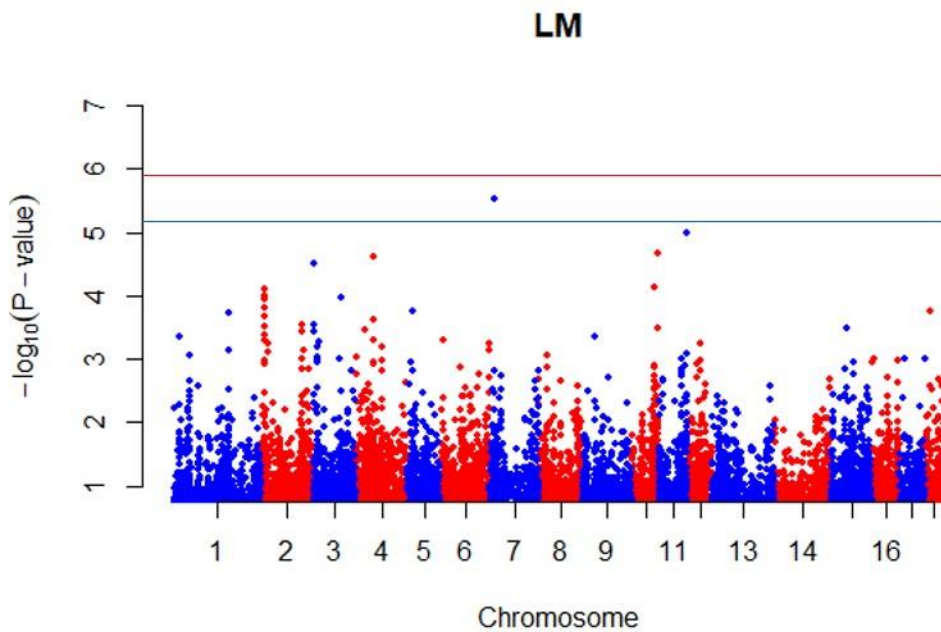
131 ³ The genes closest to the identified SNP named with the official gene symbol.

132

133 **Figure 1** Manhattan plot showing the GWAS significance for the associations between the SNPs
134 and the considered phenotypic traits. The red line indicates the genome-wide threshold of
135 significance while the blue line indicates the chromosome-wide threshold of significance calculated
136 for the chromosomes where the relevant markers map
137

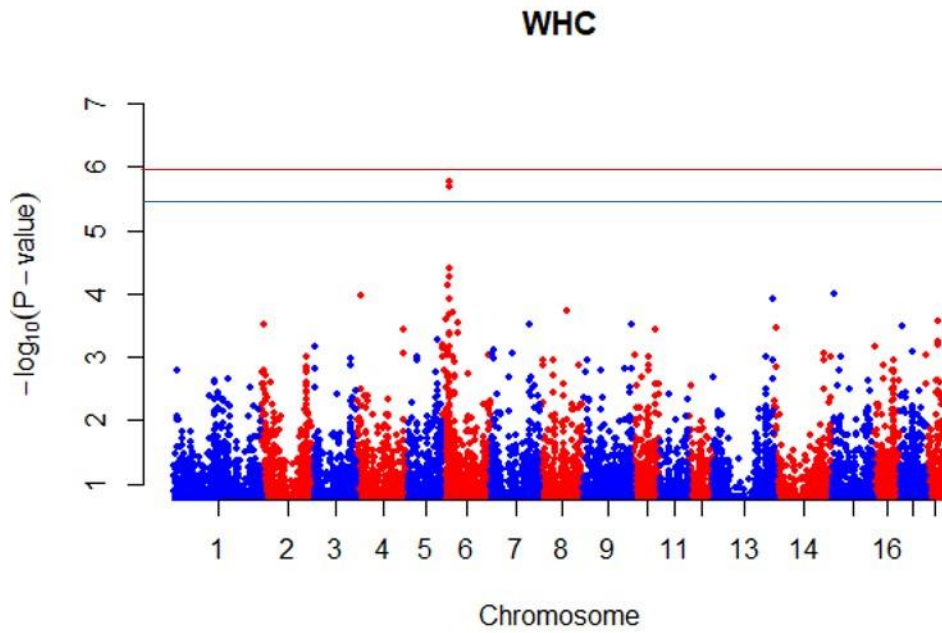


138
139



140

141



142

143

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

146 LM: carcass lean meat percentage.

147 WHC: water-holding capacity.

148

149

150 **Supporting Information**

151

152 **Figure S1.** The population structure investigated with PCA

153

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

157

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

159

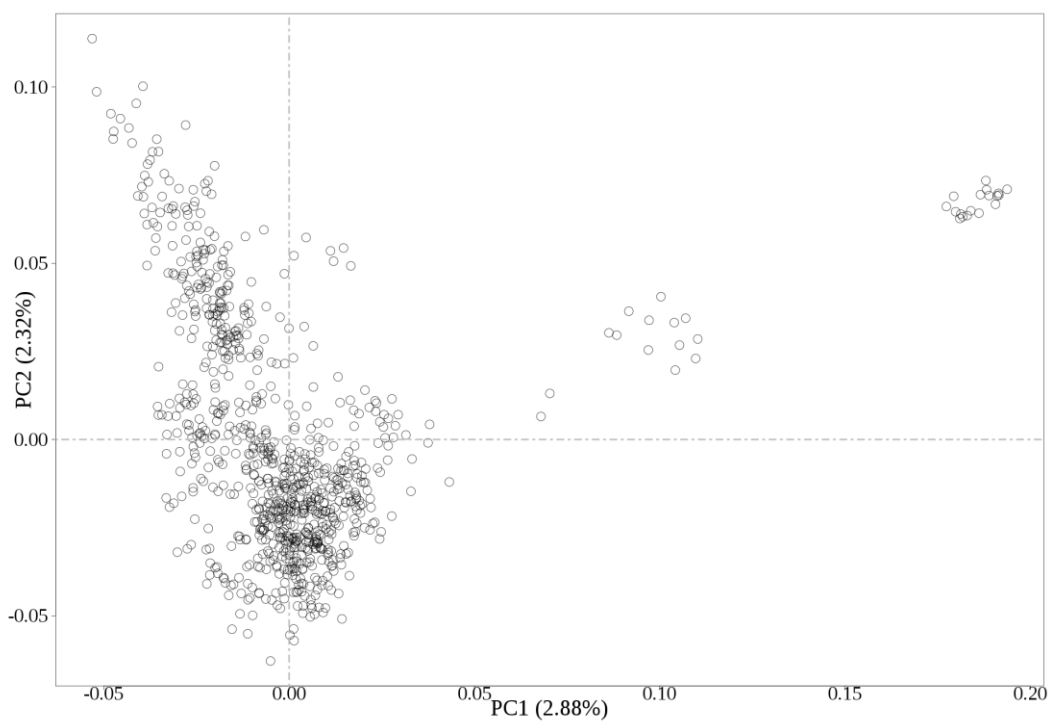
160 **Table S3.** Protein-coding genes included in the three significant chromosome regions

161

162

163 **Figure S1.** The population structure investigated with PCA

164



165

166

167

Phenotypes measurement
<p>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), <i>Longissimus thoracis</i> muscle thickness (MT), and lightness (RW). Backfat thickness was recorded with a caliber at the midline at the level of <i>Gluteus medius</i> muscle (BFT). Moreover, we also determined pH1 (meat pH measured 1 h <i>postmortem</i>) and pHu (meat pH measured 24 h <i>postmortem</i>), 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 <i>et al.</i>, 1982).</p>
DNA extraction and Genotyping
<p>DNA was isolated from <i>Semimembranosus</i> 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 <i>Sus scrofa</i> genome assembly Build 11.1.</p>
Statistical analyses
<p>After PLINK (Purcell <i>et al.</i>, 2007) filtering according to Nicolazzi <i>et al.</i> (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 <code>-pca</code> 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 <i>et al.</i>, 2007): samples with call rate <90%, SNPs with a GENO <90%, SNPs with ah Hardy-Weinberg equilibrium <i>P</i>-value <0.001, and SNPs with minor allele frequency <5% were excluded. The remaining 885</p>

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:

$$Y_i = X_i\beta_i + Z_ia_i + e_i$$

where Y_i is the observation vector for the i th 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. P_{cdf} 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.

171 **Table S3.** Protein coding genes included in the three significant chromosome regions
 172

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