

RESEARCH ARTICLE

# Diversity across major and candidate genes in European local pig breeds

María Muñoz<sup>1</sup>, Riccardo Bozzi<sup>2</sup>, Fabián García<sup>1</sup>, Yolanda Núñez<sup>1</sup>, Claudia Geraci<sup>3</sup>, Alessandro Crovetto<sup>2</sup>, Juan García-Casco<sup>1</sup>, Estefania Alves<sup>1</sup>, Martin Škrlep<sup>4</sup>, Rui Charneca<sup>5</sup>, Jose M. Martins<sup>5</sup>, Raquel Quintanilla<sup>6</sup>, Joan Tibau<sup>6</sup>, Goran Kušec<sup>7</sup>, Ivona Djurkin-Kušec<sup>7</sup>, Marie J. Mercat<sup>8</sup>, Juliette Riquet<sup>9</sup>, Jordi Estellé<sup>10</sup>, Christoph Zimmer<sup>11</sup>, Violeta Razmaite<sup>12</sup>, Jose P. Araujo<sup>13</sup>, Čedomir Radović<sup>14</sup>, Radomir Savic<sup>15</sup>, Danijel Karolyi<sup>16</sup>, Maurizio Gallo<sup>17</sup>, Marjeta Čandek-Potokar<sup>4</sup>, Luca Fontanesi<sup>3</sup>, Ana I. Fernández<sup>1</sup>, Cristina Óvilo<sup>1\*</sup>

**1** Departamento Mejora Genética Animal, INIA, Madrid, Spain, **2** Università degli Studi di Firenze, Firenze, Italy, **3** Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy, **4** Kmetijski inštitut Slovenije, Ljubljana, Slovenia, **5** Instituto de Ciências Agrárias e Ambientais Mediterrânicas (ICAAM), Universidade de Évora, Évora, Portugal, **6** Programa de Genética y Mejora Animal, IRTA, Barcelona, Spain, **7** University of Osijek, Faculty of Agrobiotechnical Sciences, Osijek, Croatia, **8** Institut du Porc, IFIP, Le Rheu, France, **9** Génétique Physiologie et Système d'Élevage, INRA, Castanet-Tolosan, France, **10** GABI, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France, **11** Bäuerliche Erzeugergemeinschaft Schwäbisch Hall, Schwäbisch Hall, Germany, **12** Animal Science Institute, Lithuanian University of Health Sciences, Baisogala, Lithuania, **13** Instituto Politecnico de Viana do Castelo, Viana do Castelo, Portugal, **14** Institute for Animal Husbandry-Pig Research Department, Belgrade-Zemun, Serbia, **15** University of Belgrade, Faculty of agriculture, Belgrade-Zemun, Serbia, **16** Department of animal science, Faculty of agriculture, University of Zagreb, Zagreb, Croatia, **17** Associazione Nazionale Allevatori Suini (ANAS), Roma, Italy

\* [ovilo@inia.es](mailto:ovilo@inia.es)



**OPEN ACCESS**

**Citation:** Muñoz M, Bozzi R, García F, Núñez Y, Geraci C, Crovetto A, et al. (2018) Diversity across major and candidate genes in European local pig breeds. PLoS ONE 13(11): e0207475. <https://doi.org/10.1371/journal.pone.0207475>

**Editor:** Juan J. Loor, University of Illinois, UNITED STATES

**Received:** May 18, 2018

**Accepted:** October 30, 2018

**Published:** November 20, 2018

**Copyright:** © 2018 Muñoz et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 634476 to CO (project acronym TREASURE). The content of this paper reflects only the author's view and the European Union Agency is not responsible for any use that may be made of the information it contains. MČP and MŠ were co-financed by Slovenian Research Agency (grant P4-0133). The

## Abstract

The aim of this work was to analyse the distribution of causal and candidate mutations associated to relevant productive traits in twenty local European pig breeds. Also, the potential of the SNP panel employed for elucidating the genetic structure and relationships among breeds was evaluated. Most relevant genes and mutations associated with pig morphological, productive, meat quality, reproductive and disease resistance traits were prioritized and analyzed in a maximum of 47 blood samples from each of the breeds (Alentejana, Apulo-Calabrese, Basque, Bísara, Majorcan Black, Black Slavonian (Crna slavonska), Casertana, Cinta Senese, Gascon, Iberian, Krškopolje (Krškopoljski), Lithuanian indigenous wattle, Lithuanian White Old Type, Mora Romagnola, Moravka, Nero Siciliano, Sarda, Schwäbisch-Hällisches Schwein (Swabian Hall pig), Swallow-Bellied Mangalitsa and Turopolje). We successfully analyzed allelic variation in 39 polymorphisms, located in 33 candidate genes. Results provide relevant information regarding genetic diversity and segregation of SNPs associated to production and quality traits. Coat color and morphological trait-genes that show low level of segregation, and fixed SNPs may be useful for traceability. On the other hand, we detected SNPs which may be useful for association studies as well as breeding programs. For instance, we observed predominance of alleles that might be unfavorable for disease resistance and boar taint in most breeds and segregation of many alleles involved in meat quality, fatness and growth traits. Overall, these findings provide a detailed

fundings had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist, excepting that the corresponding author Cristina Óvilo is a PLOS ONE Editorial Board member. However, this does not alter the authors' adherence to all the PLOS ONE policies.

catalogue of segregating candidate SNPs in 20 European local pig breeds that may be useful for traceability purposes, for association studies and for breeding schemes. Population genetic analyses based on these candidate genes are able to uncover some clues regarding the hidden genetic substructure of these populations, as the extreme genetic closeness between Iberian and Alentejana breeds and an uneven admixture of the breeds studied. The results are in agreement with available knowledge regarding breed history and management, although largest panels of neutral markers should be employed to get a deeper understanding of the population's structure and relationships.

## Introduction

Modern pig production is mostly based on highly-selected genotypes from lean breeds, farmed in intensive production systems for provisioning the market with fresh pork. On the other hand, the use of local examples of pig breeds is usually associated with local forms of pig husbandry and processing of their meat into high-quality products, often dry-cured. The most representative local pig breed is the Iberian raised in the southwest of the Iberian Peninsula, but there are many other breeds reared in European countries. These breeds are in general characterized by a good adaptation to specific environments, high potential for fat deposition and characteristic meat quality, mostly related to high intramuscular or intermuscular fat content, which are associated with high quality pork productions.

Conservation of animal genetic resources depends on current and future economic relevance of autochthonous breeds but also on scientific importance, the cultural and historical value, adaptation to environment and contribution to the development of local communities and marginal areas. Local breeds, although showing economic and social relevance, may be threatened by inefficient utilization [1] and the development of sustainable pork chains is a main goal. In order to help breed promotion and sustainability, a deep phenotypic and genetic characterization is needed.

Neutral markers like microsatellites have been recommended by FAO for genetic characterization of animal breeds [1]. However, the relationship between variability at neutral marker loci and productive or adaptive traits is unclear [2]. It is also known that in domestic animals some important phenotypic differences between breeds such as coat color, vertebrae number and production traits, may be due to differences at few loci [3]. Hence, sequence polymorphisms within major or candidate genes may be more useful for the purposes of genetic characterization, breed differentiation and development of new breeding procedures aiming to manage genetic information.

In this study, we have investigated, in 20 local breeds, allele frequency distribution of causative polymorphisms and polymorphisms associated with relevant traits in a total of 33 genes selected from previous studies carried out in cosmopolitan breeds. The analysis was focused on variants in candidate genes known to affect the regulation of growth and appetite, meat quality, coat color, reproduction and disease resistance. Moreover, the genotyping information was employed to assess population genetics and structure measures. This work was framed in the TREASURE project (<https://treasure.kis.si>), a multidisciplinary European Union funded project pointing toward development of sustainable pork chains in several European local pig breeds.

## Material and methods

Blood samples were obtained from 47 individuals from each one of the 20 breeds included in the study: Black Slavonian and Turopolje (Croatia), Basque and Gascon (France), Schwäbisch-

Hällisches Schwein (Germany), Apulo-Calabrese, Casertana, Cinta Senese, Mora Romagnola, Nero Siciliano and Sarda (Italy), Lithuanian indigenous wattle and Lithuanian White old type (Lithuania), Alentejana and Bísara (Portugal), Moravka and Swallow-Bellied Mangalitsa (Serbia), Krškopolje (Slovenia) and Iberian and Majorcan Black (Spain). Selection of individuals for sampling was performed avoiding highly related animals (no full- or half-sibs), balancing between sexes and prioritizing adult individuals or at least animals with adult morphology. Specialized professionals from each institution that provided animal material obtained all blood samples following standard routine monitoring procedures and guidelines, at farm or at slaughter. No animal experiment was performed within this research and blood samples were obtained as a general breeding procedure and only reused here.

The genomic DNA was extracted from leukocytes present in 8–15 mL of peripheral blood, collected in Vacutainer tubes containing 10% 0.5 M EDTA (ethylenediaminetetraacetic acid, disodium dihydrate salt) at pH 8.0. The extraction was performed using either a standardized phenol-chloroform, high-salt method or a commercial kit [4]. Most interesting candidate genes for relevant productive traits in pigs were prioritized among those prone to be genotyped with TaqMan medium-throughput OpenArray Genotyping platform (Thermo Fisher Scientific). For the SNP selection, previous evidences of association with relevant traits (morphological, productive, reproductive, meat quality and disease resistance traits) were considered, prioritizing specially those which seemed to be potential or known causal mutations. A final SNP panel of 39 SNPs (Table 1) was selected and genotyped by the aforementioned OpenArray system. The genotyping was performed in a QuantStudio 12 K Flex Real-Time PCR System (Thermo Fisher Scientific) at the Centre for Research in Agricultural Genomics (CRAG, Barcelona, Spain). In this procedure, SNPs are typed using TaqMan genotyping chemistry supported on a metal-based array. DNA samples were loaded and amplified on arrays as recommended by the manufacturer. Endpoint detection of signal intensities of allele specific fluorescent dyes was conducted by the OpenArray NT Imager, and genotypes were called using the OpenArray SNP Genotyping analysis software. All SNPs were visually examined for any clustering issues.

Melanocortin receptor 1 (*MC1R*) haplotypes were constructed from the data on individual SNPs located at the following positions on SSC6 according to Ss11.1: g.181461 (G>A; p.A243T)– g.181697 (C>T; p.A164V)– g.181818 (G>A; p.D124N)– g.181825 g.181825 (T>C; p.L102P)– g.181905 (G>A; p.V95M)– g.182126 (CC>\*), with the following definition of haplotypes: MC1R\*1: GCGTG\*\*; MC1R\*2: GCGCA\*\*; MC1R\*3: GCATG\*\*; MC1R\*4: ATGTG\*\*; MC1R\*6: GCATGCC; MC1R\*7: GCGTGCC

Genetic variability at the different loci in each population was measured with observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity and  $F_{ST}$  [5]. Population structure was evaluated using the  $F_{IS}$  statistic.  $D_S$  and  $F_{ST}$  distance measures were estimated according to Takezaki and Nei [6]. These estimates and allele frequencies were computed using the *hierfstat* library in R environment [7]. Neighbor-joining trees were constructed using  $D_S$  and  $F_{ST}$  distances with the *nj* function belonging to the *ape* library in R [8]. Bootstrap re-sampling ( $n = 1,000$ ) was performed to test the robustness of the phylogenetic tree topologies.

The population structure was also analyzed by a Discriminant Analysis of Principal Component (DAPC) through *adegenet* package v.2.0.2 [9] in R environment [10]. The optimal number of clusters was identified through the Bayesian Information Criterion (BIC) and clusters were plotted in a scatterplot of the first and second linear discriminants of DAPC. Assignment of the individuals at clusters was plotted using the results coming from two different methods:  $\alpha$ -score optimization and cross-validation. Both methods provide the optimal number of principal components (PC) to retain finding a trade-off between power of discrimination and overfitting. For the cross-validation, data were randomly divided in a training set

Table 1. SNP selection and their association with quantitative traits.

Gene	SSC	Genome position	Polymorphism	Associated Traits	References
<i>ACACA</i>	12	38624687	c.5634T>C	Fatness, meat quality, FA profile	[14]
<i>ACSL4</i>	X	89763094	c.2645G>A	Meat quality, FA profile	[15]
<i>ADIPOQ</i>	13	124643017	c.1735G>A	Fatness	[16]
<i>AHR</i>	9	86550830	c.2149G>T	Litter size	[17]
<i>CAPNS1</i>	6	45514212	c.429A>C	Meat quality	[18]
<i>CAST</i>	2	103299934	g.103299934G>A	Meat quality	[19]
<i>CYB5A</i>	1	149737752	g.149737752G>T	Meat quality, boar taint	[20,21]
<i>CYP2E1</i>	14	141690107	g.141690107C>T	Meat quality, boar taint	[22,23]
<i>ESR1</i>	1	14418777	c.669T > C	Litter size	[24]
<i>FASN</i>	12	926299	g.926299G>A	Fatness, meat quality, FA profile	[25]
<i>FTO</i>	6	31460242	g.31460242A>T	Growth, fatness	[26]
<i>FUT1</i>	6	54079560	g.54079560T>C	Disease resistance	[27]
<i>GBP5</i>	4	127301202	g.127301202G>T	Disease resistance	[28–30]
<i>IGF2</i>	2	1483832	g.1483832G>A	Growth, fatness	[31]
<i>KIT</i>	8	41488472	g.41488472C>T	Coat color	[32–34]
<i>LEP</i>	18	20111759	g.20111759A>G	Growth, fatness	[35]
<i>LEPR</i>	6	146829589	c.1987C>T	Growth, fatness	[36,37]
<i>MC1R</i>	6	181461	g.181461G>A	Coat color	[38,39]
<i>MC1R</i>	6	181697	g.181697C>T	Coat color	[38,39]
<i>MC1R</i>	6	181818	g.181818G>A	Coat color	[38,39]
<i>MC1R</i>	6	181825	g.181825T>C	Coat color	[38,39]
<i>MC1R</i>	6	181905	g.181905G>A	Coat color	[38,39]
<i>MC1R</i>	6	182126	g.182126CC>*	Coat color	[38,39]
<i>MC4R</i>	1	160773437	c.892G>A	Growth, fatness	[40]
<i>MSTN</i>	15	94629248	g.94629248C>T	Growth, fatness	[41]
<i>MTTP</i>	8	120821998	c.2573T>C	Meat quality, FA profile	[42]
<i>MUC4</i>	13	134226654	g.134226654C>G	Disease resistance	[43]
<i>NR6A1</i>	1	265347265	c.748T>C	Number of vertebrae	[44]
<i>PCK1</i>	17	57932233	c.2456A>C	Meat quality	[45]
<i>PHKG1</i>	3	16830320	g.16830320C>A	Fatness, meat quality	[46]
<i>PPARD</i>	7	31281804	g.31281804G>A	Ear size	[47]
<i>PPARGCIA</i>	8	17867068	c.1288T>A	Meat quality	[18]
<i>PRKAG3</i>	15	120863537	g.120863537G>A	meat quality	[48,49]
<i>PRKAG3</i>	15	120863533	g.120863533G>A	meat quality	[48,49]
<i>RYR1</i>	6	47357966	c.1843C>T	performance, meat quality	[50,51]
<i>SCD</i>	14	111461751	g.111461751C>T	Meat quality, FA profile	[52]
<i>TAS2R39</i>	18	7068883	g.7068883T>G	Fatness	[53]
<i>TAS2R4</i>	UN	8135115	g.8135115A>T	Fatness	[53]
<i>TYRPI</i>	1	209733431	g.209733431A>G	Coat color	[54]

Polymorphism locations are indicated according to Sus Scrofa genome annotation v11.1

<https://doi.org/10.1371/journal.pone.0207475.t001>

(90%) and a validation set (10%) with 1000 replications. Members of each group were selected in order to ensure that each population in the original data set was represented in both training and validation sets. The optimal number of PCs to retain was based on the lowest root mean square error.

The most likely number of partitions in the dataset, irrespective of breed of origin, was also determined using the algorithm implemented in STRUCTURE [11]. We carried out 20 different runs from  $K = 1$  to  $K = 25$ . All runs used a burn-in period of 50,000 iterations and a period of data collection of 200,000 iterations under an admixture model with allele frequencies correlated. The most likely number of clusters ( $K$ ) was chosen following Evanno *et al.* [12] recommendations. The similarity of the outcomes of the 20 solutions was assessed using CLUMMP software [13]; the most frequent solution was considered to be the most probable.

## Results and discussion

The analyzed SNP panel includes polymorphisms related to coat color and morphological traits, growth, fatness and carcass traits, meat quality traits, reproduction and disease resistance.

TaqMan allelic discrimination qPCR assays have been shown to be accurate and reliable [55,56]. The initial validation of the assays showed the successful genotyping of all the tested mutations. Unique clusters indicative of different genotypes were formed based on the signal intensity ratio of the employed probes for 36 out of the 39 analyzed SNPs. The remaining three SNPs were monomorphic showing only one cluster. The success rate of each genotyped marker (call rate) was higher than 98%. Regarding the samples, those with genotyping call rates lower than 80% were discarded, resulting in a successful genotyping for 95% of the analyzed samples. The number of animals successfully genotyped in each breed and allele frequencies are shown in Tables 2–5. Genotyping data is included in S1 Table.

### Markers related to morphological and coat color traits

A picture showing the characteristic phenotype and country of origin for each breed is shown in Fig 1. In general, low segregation was observed for markers in genes involved in coat color and morphology traits (Table 2), as it could be expected as these are the main selection criteria employed by breeders over the years, even in local breeds. In pigs, a large variety of coat colors and patterns, characteristics of different breeds and populations, have been produced following domestication and selection [57–59]. Among the several loci involved in pigmentation, *MC1R* and *v-kit* Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog gene (*KIT*) genes are supposed to play major roles in determining coat color variation in Mediterranean pig breeds or populations [60].

*MC1R* gene codes for a G protein-coupled receptor that is primarily expressed in melanocytes and plays a key role in melanogenesis. It is involved in regulation of eumelanin (black/brown) and pheomelanin (yellow/red) synthesis in the melanocyte and is encoded by the Extension (*E*) coat color locus. Several SNPs were genotyped for *MC1R* locus, which were used for the construction of haplotypes [38,59]. Five main haplotypes are usually considered. The wild allele ( $MC1R^*1, E^+$ ), which in many species involves synthesis of both melanin types and is present in wild boars, was present in Mora Romagnola (0.52) and fixed in Swallow-Bellied Mangalitsa. The fixation of  $MC1R^*1$  in Swallow-Bellied Mangalitsa breed has already been reported [59]. Moreover, this breed was proposed as the only domestic breed carrying this allele among a wide panel of European and Chinese commercial and local breeds [59]. Nevertheless, we also detect a very small introgression of  $MC1R^*1$  in Moravka ( $\leq 0.03$ ), Nero Siciliano and Sarda ( $\leq 0.01$ ). The  $MC1R^*2$  allele ( $E^{D1}$ ), characterized by a p.L102P aminoacid change is associated with dominant black coat. This allele has Asian origin and has been previously detected in Chinese and Large Black breeds. According to our results  $MC1R^*2$  is observed at high frequency in Black Slavonian (0.89) and with moderate presence in Moravka and Schwäbisch-Hällisches. The Black Slavonian pig comes from eastern Croatia and was



**Table 2. Allele frequencies for seven haplotypes and four polymorphisms on candidate genes for coat colour and morphological traits, in 20 European local breeds.**

Breed (n)	MC1R <sup>†</sup>						KIT		TYRP1		NR6A1		PPARD	
	1	2	3	4	6	7	C	T	A	G	T	C	G	A
Alentejana (47)	0.00	0.01	0.09	0.00	0.66	0.24	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.85	0.15	<b>1.00</b>	<b>0.00</b>
Apulo-Calabrese (44)	0.00	0.00	0.90	0.00	0.10	0.00	0.99	0.01	<b>1.00</b>	<b>0.00</b>	0.80	0.20	<b>1.00</b>	<b>0.00</b>
Basque (47)	0.00	0.01	0.98	0.00	0.01	0.00	0.88	0.12	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>
Bisara (47)	0.00	0.00	0.86	0.01	0.13	0.00	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.83	0.17
Black Slavonian (30)	0.00	0.89	0.04	0.02	0.06	0.00	0.99	0.01	<b>1.00</b>	<b>0.00</b>	0.92	0.08	0.92	0.08
Casertana (46)	0.00	0.00	0.89	0.02	0.09	0.00	0.98	0.02	0.98	0.02	0.98	0.02	0.95	0.05
Cinta Senese (46)	0.00	0.00	0.78	0.06	0.16	0.00	0.02	0.98	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>
Gascon (47)	0.00	0.00	0.98	0.00	0.02	0.00	<b>1.00</b>	<b>0.00</b>	0.99	0.01	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>
Iberian (47)	0.00	0.00	0.08	0.00	0.66	0.26	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.88	0.13	<b>1.00</b>	<b>0.00</b>
Krškopolje (36)	0.00	0.00	0.82	0.02	0.16	0.00	0.98	0.02	0.95	0.05	<b>1.00</b>	<b>0.00</b>	0.87	0.13
Lithuanian indigenous wattle (47)	0.00	0.00	0.00	0.13	0.87	0.00	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.61	0.39	0.82	0.18
Lithuanian White Old Type(47)	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.97	0.03	<b>1.00</b>	<b>0.00</b>	0.99	0.01	0.99	0.01
Majorcan Black (47)	0.00	0.02	0.92	0.01	0.05	0.00	0.88	0.12	<b>1.00</b>	<b>0.00</b>	0.95	0.05	<b>1.00</b>	<b>0.00</b>
Mora Romagnola (43)	0.52	0.00	0.00	0.48	0.00	0.00	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>
Moravka (47)	0.03	0.37	0.43	0.00	0.17	0.00	0.79	0.21	0.82	0.18	<b>1.00</b>	<b>0.00</b>	0.94	0.06
Nero Siciliano(38)	0.01	0.04	0.75	0.03	0.18	0.00	0.99	0.01	0.99	0.01	0.73	0.27	0.96	0.04
Sarda (46)	0.01	0.04	0.26	0.03	0.66	0.00	0.86	0.14	0.99	0.01	0.98	0.02	0.98	0.02
Schwäbisch-Hällisches Schwein (46)	0.00	0.16	0.84	0.00	0.00	0.00	0.83	0.17	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.96	0.04
Swallow-Bellied Mangalitsa (46)	<b>1.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	0.21	0.79	<b>1.00</b>	<b>0.00</b>	0.98	0.02	0.92	0.08
Turopolje (45)	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.57	0.43	0.99	0.01

<sup>†</sup>MC1R haplotypes: MC1R\*1: GCGTG\*\*; MC1R\*2: GCGCA\*\*; MC1R\*3: GCATG\*\*; MC1R\*4: ATGTG\*\*; MC1R\*6: GCATGCC; MC1R\*7: GCGTGCC.; Fixed alleles in bold.

<https://doi.org/10.1371/journal.pone.0207475.t002>

created at the end of the 19th century by the crossbreeding of Swallow-Bellied Mangalitsa, Berkshire, Large Black and Poland China breeds [61,62], which explains the origin of the Asian *MC1R* allele. Interestingly, MC1R\*1 allele probably provided by Swallow-Bellied Mangalitsa has not persisted, possibly due to selection of the black coat color. Since the Black Slavonian pig is the only Croatian breed with MC1R\*2 genotype, the genotyping of this locus has been already proposed to determine the purity of the Black Slavonian pigs [62], as uncontrolled crossing with modern pig breeds is usual. According to our results, some introgression of alleles 3, 4 and 6 is observed in Black Slavonian, which may indicate contamination with commercial breeds such as Duroc, Pietrain or Large White.

Haplotypes MC1R\*3 (*E<sup>D2</sup>*) and MC1R\*6 (*E<sup>P</sup>*), are widely observed in European breeds. Allele MC1R\*3 is associated with a p.D124N substitution involving black coat color. Allele MC1R\*6 causes black spotting on a red or white background and possesses, in addition to the p.D124N substitution for dominant black coloring, a two base pair insertion at codon 22. These two alleles are predominant in most of the analyzed breeds, which is in agreement with the dark coat color of most of them. MC1R\*3 mutation is the most frequent in Mediterranean and Pyrenean pig breeds whereas this variant is practically absent in Iberian Peninsula breeds. MC1R\*6 allele showed the highest frequency in Lithuanian indigenous wattle and Turopolje breeds in agreement with its characteristic patched phenotype. This allele was also fixed in Lithuanian White Old Type, in which the white coat color must be consequence of variation in the *KIT* locus.

At last, allele MC1R\*4 (*e*), associated with the recessive red coat color and characteristic of the Duroc breed was present in Mora Romagnola and detected at very low frequency in several

**Table 3. Allele frequencies for nine polymorphisms on candidate genes for meat production and carcass traits, in 20 European local breeds.**

Breed (n)	RYR 1		IGF2		MC4R		LEPR		LEP		FTO		MSTN		TAS2R39		TAS2R4	
	T	C	A	G	G	A	C	T	C	T	A	T	A	G	T	G	A	T
Alentejana (47)	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	0.98	0.02	0.02	0.98	0.16	0.84	0.32	0.68	0.80	0.20	0.01	0.99	<b>1.00</b>	<b>0.00</b>
Apulo-Calabrese (44)	0.02	0.98	0.85	0.15	0.57	0.43	0.71	0.29	0.66	0.34	0.12	0.88	0.20	0.80	0.04	0.96	<b>1.00</b>	<b>0.00</b>
Basque (47)	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	0.27	0.73	0.84	0.16	0.39	0.61	0.00	1.00	<b>1.00</b>	<b>0.00</b>
Bisara (47)	0.05	0.95	0.01	0.99	0.94	0.06	0.74	0.26	0.94	0.06	0.50	0.50	0.81	0.19	0.13	0.87	<b>1.00</b>	<b>0.00</b>
Black Slavonian (30)	<b>0.00</b>	<b>1.00</b>	0.08	0.92	0.94	0.06	0.62	0.38	0.62	0.38	0.21	0.79	0.11	0.89	0.21	0.79	<b>1.00</b>	<b>0.00</b>
Casertana (46)	0.09	0.91	0.03	0.97	0.81	0.19	0.89	0.11	0.66	0.34	0.51	0.49	0.86	0.14	0.01	0.99	<b>1.00</b>	<b>0.00</b>
Cinta Senese (46)	0.01	0.99	0.12	0.88	0.29	0.71	0.87	0.13	0.47	0.53	0.33	0.67	0.66	0.34	0.27	0.73	<b>1.00</b>	<b>0.00</b>
Gascon (47)	0.03	0.97	0.00	1.00	0.33	0.67	0.88	0.13	0.48	0.52	0.50	0.50	0.92	0.08	<b>0.00</b>	<b>1.00</b>	<b>1.00</b>	<b>0.00</b>
Iberian (47)	<b>0.00</b>	<b>1.00</b>	0.06	0.94	0.95	0.05	<b>0.00</b>	<b>1.00</b>	0.26	0.74	0.25	0.75	0.95	0.05	0.06	0.94	<b>1.00</b>	<b>0.00</b>
Krškopolje (36)	0.21	0.79	0.39	0.61	0.34	0.66	0.60	0.40	0.77	0.23	0.34	0.66	0.57	0.43	0.23	0.77	<b>1.00</b>	<b>0.00</b>
Lithuanian indigenous wattle (47)	0.09	0.91	0.10	0.90	0.38	0.63	0.41	0.59	0.85	0.15	0.04	0.96	0.34	0.66	0.01	0.99	<b>1.00</b>	<b>0.00</b>
Lithuanian White Old Type (47)	<b>0.00</b>	<b>1.00</b>	0.28	0.72	0.46	0.54	0.82	0.18	<b>1.00</b>	<b>0.00</b>	0.26	0.74	0.75	0.25	0.09	0.91	<b>1.00</b>	<b>0.00</b>
Majorcan Black (47)	<b>0.00</b>	<b>1.00</b>	0.19	0.81	0.94	0.06	0.34	0.66	0.17	0.83	0.31	0.69	0.75	0.25	0.42	0.58	<b>1.00</b>	<b>0.00</b>
Mora Romagnola (43)	<b>0.00</b>	<b>1.00</b>	0.10	0.90	0.82	0.18	<b>1.00</b>	<b>0.00</b>	0.26	0.74	0.82	0.18	0.99	0.01	0.01	0.99	<b>1.00</b>	<b>0.00</b>
Moravka (47)	0.02	0.98	0.26	0.74	0.54	0.46	0.58	0.42	0.63	0.37	0.35	0.65	0.39	0.61	0.02	0.98	<b>1.00</b>	<b>0.00</b>
Nero Siciliano (38)	0.01	0.99	0.20	0.80	0.63	0.17	0.89	0.11	0.55	0.45	0.39	0.61	0.49	0.51	0.18	0.82	<b>1.00</b>	<b>0.00</b>
Sarda (46)	0.05	0.95	0.27	0.73	0.63	0.37	0.36	0.64	0.47	0.53	0.38	0.62	0.79	0.21	0.29	0.71	<b>1.00</b>	<b>0.00</b>
Schwäbisch-Hällisches Schwein (46)	<b>0.00</b>	<b>1.00</b>	0.50	0.50	0.70	0.30	0.95	0.05	0.64	0.36	0.75	0.25	0.43	0.57	0.09	0.91	<b>1.00</b>	<b>0.00</b>
Swallow-Bellied Mangalitsa (46)	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	0.93	0.07	0.76	0.24	0.02	0.98	0.37	0.63	0.49	0.51	0.17	0.83	<b>1.00</b>	<b>0.00</b>
Turopolje (45)	<b>0.00</b>	<b>1.00</b>	<b>0.00</b>	<b>1.00</b>	0.87	0.13	<b>1.00</b>	<b>0.00</b>	0.35	0.65	0.95	0.05	0.72	0.28	<b>0.00</b>	<b>1.00</b>	<b>1.00</b>	<b>0.00</b>

Fixed alleles in bold

<https://doi.org/10.1371/journal.pone.0207475.t003>

other breeds. This allele was previously indicated to be fixed in the Mora Romagnola breed [63], which shows a solid phaeomelanin coat color. However, according to our results allele MC1R\*4 is not fixed in this Italian breed but sharing distribution with MC1R\*1 allele (allele frequencies 0.48 and 0.52, respectively). The small introgression of MC1R\*4 allele in other breeds may be caused by crossbreeding with Duroc, which is a usual strategy in various breeds to improve growth rate and feed efficiency. For example, the use of Italian Duroc has been proposed to improve the farming performance of Cinta Senese without reducing the fresh meat quality [64] and crossbreeding with Duroc has also been evaluated in Lithuanian indigenous wattle showing increased performance [65]. Although crossbreeding with Duroc is the main production system of Iberian pigs, no introgression is observed in the pure Iberian genetic pool.

We previously detected a new allele, MC1R\*7 [66], which was only present in the closely related Iberian and Alentejana breeds. This allele was characterized by the combination of the CC insertion without the p.D124N substitution. To the best of our knowledge, this combination has only been detected in a European pig with unknown phenotype [67] and in Iberian pigs with red coat color [66].

KIT gene plays key roles in melanogenesis, erythropoiesis, spermatogenesis and T-cell differentiation. Jointly with MC1R, they are the major determinants of the across-breed differences in coat color patterns in different mammals. Mutations in the porcine KIT gene (dominant white locus) have been shown to affect coat color and color distribution [66]. A large number of alleles have been reported in pigs [33,34] but their differentiation is complex as some are associated to copy number variation (CNV) of the gene, with a duplication or a triplication or even a higher number of copies of a region including the whole coding

**Table 4. Allele frequencies for 15 polymorphisms on candidate genes for meat quality, in 20 European local breeds.**

Breed (n)	SCD		ACACA		ACSL4		EASN		MTTP		CYP2E1		CYB5A		CAST		PPARGC1A		CAPNS1		ADIPOQ		PRKAG3 H199Y		PRKAG3 R200Q		PHKG1		PCK1					
	T	C	C	T	G	A	A	G	A	C	T	C	T	G	A	G	A	T	A	A	C	A	G	A	G	A	C	A	A	C				
Alentejana (47)	1.00	0.00	0.44	0.56	0.46	0.54	0.17	0.83	0.54	0.46	0.98	0.02	0.50	0.50	0.80	0.20	0.80	0.03	0.97	0.45	0.55	0.98	0.02	0.30	0.70	1.00	0.00	1.00	0.00	1.00	0.00	0.97	0.03	
Apulo-Calabrese (44)	0.74	0.26	0.63	0.37	0.71	0.29	0.53	0.47	0.47	0.53	0.83	0.17	0.99	0.01	0.68	0.32	0.63	0.63	0.37	0.19	0.81	1.00	0.00	0.50	0.50	1.00	0.00	0.50	0.50	0.67	0.33	0.33	0.67	
Basque (47)	0.97	0.03	0.81	0.19	0.99	0.01	0.34	0.66	0.07	0.93	1.00	0.00	0.91	0.09	0.35	0.65	0.68	0.32	0.13	0.88	0.98	0.02	0.71	0.29	1.00	0.00	1.00	0.00	1.00	0.00	0.50	0.50	0.50	
Bisara (47)	0.65	0.35	0.64	0.36	0.86	0.14	0.46	0.54	0.15	0.85	0.56	0.44	0.73	0.27	0.67	0.33	0.40	0.60	0.52	0.48	0.90	0.10	0.41	0.59	1.00	0.00	0.94	0.06	0.52	0.48	0.48	0.52	0.48	
Black Slavonian (30)	0.50	0.50	0.71	0.29	0.11	0.89	0.52	0.48	0.44	0.56	0.87	0.13	0.94	0.06	0.56	0.44	0.63	0.37	0.42	0.58	0.58	0.42	0.73	0.27	1.00	0.00	1.00	0.00	1.00	0.00	0.52	0.48	0.48	
Caseriana (46)	0.73	0.27	0.42	0.58	0.91	0.09	0.58	0.42	0.44	0.56	0.82	0.18	0.98	0.02	0.68	0.32	0.48	0.52	0.71	0.29	1.00	0.00	0.14	0.86	1.00	0.00	0.97	0.03	0.33	0.67	0.33	0.67	0.33	
Cinta Senese (46)	0.96	0.04	0.83	0.17	0.36	0.64	0.59	0.41	0.91	0.09	0.54	0.46	0.69	0.31	0.32	0.68	0.97	0.03	0.18	0.82	1.00	0.00	0.66	0.34	1.00	0.00	0.96	0.04	0.54	0.46	0.46	0.54	0.46	0.46
Gascon (47)	0.91	0.09	0.39	0.61	0.97	0.03	0.31	0.69	0.44	0.56	0.21	0.79	0.95	0.05	0.43	0.57	0.56	0.44	0.18	0.82	1.00	0.00	0.69	0.31	1.00	0.00	1.00	0.00	0.15	0.85	0.85	0.15		
Iberian (47)	0.98	0.02	0.48	0.52	0.65	0.35	0.43	0.58	0.77	0.23	0.95	0.05	0.51	0.49	0.14	0.86	0.06	0.94	0.61	0.39	0.96	0.04	0.36	0.64	1.00	0.00	0.95	0.05	0.96	0.04	0.96	0.04	0.96	
Krškopolje (36)	0.85	0.15	0.60	0.40	0.73	0.27	0.68	0.32	0.37	0.63	0.45	0.55	0.98	0.02	0.91	0.09	0.37	0.63	0.19	0.81	0.79	0.21	0.78	0.22	1.00	0.00	0.58	0.42	0.34	0.66	0.66	0.34	0.66	
Lithuanian indigenous wattle (47)	0.64	0.36	0.97	0.03	0.85	0.15	0.52	0.48	0.41	0.59	0.84	0.16	0.81	0.19	0.45	0.55	0.57	0.43	0.73	0.27	0.98	0.02	0.50	0.50	1.00	0.00	0.44	0.56	0.16	0.84	0.84	0.16	0.84	
Lithuanian White Old Type (47)	0.74	0.26	0.79	0.21	0.15	0.85	0.50	0.50	0.33	0.67	0.70	0.30	0.83	0.17	0.74	0.26	0.14	0.86	0.36	0.64	0.85	0.15	0.93	0.07	1.00	0.00	0.95	0.05	0.57	0.43	0.57	0.43	0.57	
Majorean Black (47)	0.95	0.05	0.47	0.53	0.60	0.40	0.35	0.65	0.60	0.40	0.90	0.10	0.82	0.18	0.80	0.20	0.27	0.73	0.60	0.40	0.99	0.01	0.56	0.44	1.00	0.00	1.00	0.00	0.98	0.02	0.89	0.11	0.89	
Mora Romagnola (43)	0.17	0.83	0.33	0.67	0.93	0.07	0.55	0.45	0.52	0.48	0.53	0.47	0.94	0.06	0.19	0.81	0.84	0.16	0.88	0.12	0.89	0.11	0.89	0.11	1.00	0.00	0.98	0.02	0.89	0.11	0.89	0.11	0.89	
Moravka (47)	0.85	0.15	0.55	0.45	0.54	0.46	0.42	0.58	0.45	0.55	0.36	0.64	0.95	0.05	0.71	0.29	0.46	0.54	0.44	0.56	0.88	0.12	0.53	0.47	1.00	0.00	0.68	0.32	0.78	0.22	0.78	0.22	0.78	
Nero Siciliano (38)	0.94	0.06	0.57	0.43	0.68	0.32	0.36	0.64	0.46	0.54	0.54	0.46	0.91	0.09	0.50	0.50	0.61	0.39	0.50	0.50	0.94	0.06	0.88	0.12	1.00	0.00	0.98	0.02	0.54	0.46	0.54	0.46	0.54	
Sarda (46)	0.94	0.06	0.65	0.35	0.65	0.35	0.58	0.42	0.37	0.63	0.65	0.35	0.83	0.17	0.69	0.31	0.64	0.36	0.47	0.53	0.92	0.08	0.61	0.39	1.00	0.00	0.58	0.42	0.55	0.45	0.55	0.45	0.55	
Schwäbisch-Hällisches Schwein (46)	0.89	0.11	0.47	0.53	0.89	0.11	0.28	0.72	0.13	0.87	0.80	0.20	1.00	0.00	0.79	0.21	0.47	0.53	0.26	0.74	0.92	0.08	0.60	0.40	1.00	0.00	1.00	0.00	0.31	0.69	0.69	0.31	0.69	
Swallow-Bellied Mangalitsa (46)	0.97	0.03	0.77	0.23	0.64	0.36	0.57	0.43	0.19	0.81	0.90	0.10	0.99	0.01	0.88	0.12	0.56	0.44	0.78	0.22	1.00	0.00	0.15	0.85	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	
Turopolje (45)	1.00	0.00	0.29	0.71	0.56	0.44	0.51	0.49	0.37	0.63	0.36	0.64	1.00	0.00	0.78	0.22	0.03	0.97	0.00	1.00	1.00	0.00	0.19	0.81	1.00	0.00	1.00	0.00	0.09	0.91	0.09	0.91	0.09	

Fixed alleles in bold

<https://doi.org/10.1371/journal.pone.0207475.t004>



**Table 5. Allele frequencies for five polymorphisms on candidate genes for disease resistance and prolificacy in 20 European local breeds.**

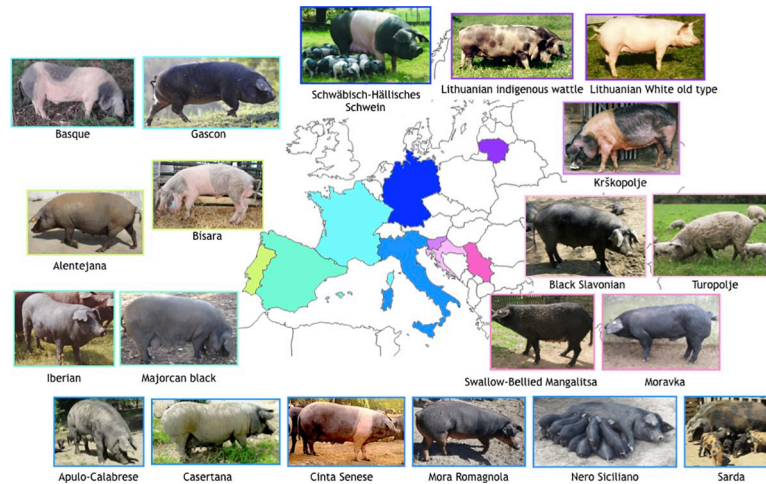
Breed (n)	GBP5		MUC4		FUT1		ESR1		AHR	
	G	T	C	G	A	G	G	A	G	T
Alentejana (47)	0.98	0.02	<b>1.00</b>	<b>0.00</b>	0.07	0.93	<b>1.00</b>	<b>0.00</b>	0.05	0.95
Apulo-Calabrese (44)	0.88	0.12	0.87	0.13	0.04	0.96	<b>1.00</b>	<b>0.00</b>	0.55	0.45
Basque (47)	<b>1.00</b>	<b>0.00</b>	0.91	0.09	0.04	0.96	<b>1.00</b>	<b>0.00</b>	0.32	0.68
Bísara (47)	0.78	0.22	0.60	0.40	0.47	0.53	<b>1.00</b>	<b>0.00</b>	0.26	0.74
Black Slavonian (30)	0.88	0.12	0.83	0.17	0.43	0.57	<b>1.00</b>	<b>0.00</b>	0.52	0.48
Casertana (46)	0.88	0.12	0.81	0.19	0.51	0.49	<b>1.00</b>	<b>0.00</b>	0.64	0.36
Cinta Senese (46)	0.49	0.51	0.94	0.06	0.12	0.88	<b>1.00</b>	<b>0.00</b>	0.49	0.51
Gascon (47)	0.82	0.18	0.82	0.18	0.48	0.52	<b>1.00</b>	<b>0.00</b>	0.27	0.73
Iberian (47)	0.92	0.08	0.95	0.05	0.06	0.94	<b>1.00</b>	<b>0.00</b>	0.07	0.93
Krškopolje (36)	0.95	0.05	0.84	0.16	0.71	0.29	<b>1.00</b>	<b>0.00</b>	0.44	0.56
Lithuanian indigenous wattle (47)	0.43	0.57	0.47	0.53	0.25	0.75	<b>1.00</b>	<b>0.00</b>	0.50	0.50
Lithuanian White Old type (47)	0.87	0.13	0.59	0.41	0.09	0.91	<b>1.00</b>	<b>0.00</b>	0.80	0.20
Majorcan Black (47)	0.91	0.09	0.88	0.12	0.25	0.75	<b>1.00</b>	<b>0.00</b>	0.11	0.89
Mora Romagnola (43)	0.58	0.42	<b>1.00</b>	<b>0.00</b>	0.09	0.91	<b>1.00</b>	<b>0.00</b>	0.13	0.88
Moravka (47)	0.93	0.07	0.89	0.11	0.09	0.91	<b>1.00</b>	<b>0.00</b>	0.42	0.58
Nero Siciliano (38)	0.84	0.16	0.89	0.11	0.12	0.88	<b>1.00</b>	<b>0.00</b>	0.24	0.76
Sarda (46)	0.73	0.27	0.77	0.23	0.18	0.82	<b>1.00</b>	<b>0.00</b>	0.48	0.52
Schwäbisch-Hällisches Schwein (46)	0.96	0.04	0.76	0.24	0.41	0.59	<b>1.00</b>	<b>0.00</b>	0.50	0.50
Swallow-Bellied Mangalitsa (46)	0.77	0.23	0.79	0.21	0.21	0.79	<b>1.00</b>	<b>0.00</b>	0.20	0.80
Turopolje (45)	0.93	0.07	0.68	0.32	0.74	0.26	<b>1.00</b>	<b>0.00</b>	0.48	0.52

Fixed alleles in bold

<https://doi.org/10.1371/journal.pone.0207475.t005>

sequence, besides a splicing mutation, a 4bp deletion and other mutations, and their multiple combinations. Thus, full characterization of *KIT* alleles was beyond the scope of this work. Nevertheless, the *KIT* g. 43597545C>T (g.41488472 in Sscrofa11.1) SNP was included in the panel, since it had been previously associated to the belted phenotype [68]. This SNP was proposed as suitable marker for differentiation of Cinta Senese pigs from other non-belted breeds and in agreement our results show a very high frequency of allele T in Cinta Senese (0.98), while most other breeds are fixed or almost fixed for the C allele. These results are similar to those reported by Ogorevc *et al.* (2017) [69] in eleven pig breeds, some of them coinciding with the ones analyzed here. Although other breeds with variable belted patterns such as Krškopolje and Schwäbisch-Hällisches Schwein are represented in the current study, our results show low frequency of the T allele, 0.02 and 0.17, respectively. These results agree with those observed by Ogorevc *et al.* [69] suggesting that unlike Cinta Senese, these SNP cannot be used to differentiate Krškopolje and Schwäbisch-Hällisches Schwein from other non-belted breeds. However, the present work evidences that Swallow-Bellied Mangalitsa breed also shows high frequency (0.79) of allele T, suggesting an introgression with Hampshire alleles due to crossbreeding.

Besides the main roles of *MC1R* and *KIT* genes on coat color, other genes have been associated to specific coat patterns. For example, several polymorphisms in the porcine Tyrosinase-related protein 1 (*TYRP1*) gene have been recently associated with the blond coat color in Liangshan pigs [54]. One of the described SNPs, TYRP1 g.209733431A>G, was included in our panel. Although individuals with blond coat color have been described for some of the studied breeds like Swallow-Bellied Mangalitsa [70], only Krškopolje (0.05) and Moravka



**Fig 1. Phenotype and geographical origin of the 20 analysed pig breeds: Black Slavonian and Turopolje (Croatia), Basque and Gascon (France), Schwäbisch-Hällisches Schwein (Germany), Apulo-Calabrese, Casertana, Cinta Senese, Mora Romagnola, Nero Siciliano and Sarda (Italy), Lithuanian indigenous wattle and Lithuanian White old type (Lithuania), Alentejana and Bisara (Portugal), Moravka and Swallow-Bellied Mangalitsa (Serbia), Krškopolje (Slovenia) and Iberian and Majorcan Black (Spain).**

<https://doi.org/10.1371/journal.pone.0207475.g001>

(0.18) breeds showed some introgression of the blond allele (G) of *TYRP1* gene even if the blond phenotype was not present in these breeds. These two breeds have known introgression of different commercial breeds. In the case of Krškopolje, German Landrace, Pietrain and Duroc breeds were commonly used as terminal sires for many years. Regarding Moravka, the creation of this breed is known to be the result of unsystematic crossing of the ancient and extinct local breed, Šumadinka, with Berkshire [71], and also crossing with Yorkshire could have taken place at the beginning of the 20<sup>th</sup> century. These influences could explain the presence of unexpected alleles.

The orphan nuclear receptor, germ cell nuclear factor (*NR6A1*) gene has been shown to affect vertebrae number in the pig. While wild boars show a fixed vertebrae number (19), most domesticated pig breeds show a vertebrae number varying between 21 and 23 [72]. Previous reports indicate that the c. 748T allele, associated with higher number of vertebrae, appeared to be fixed in most studied breeds whereas in wild boars the c. 748C allele is fixed [73]. In our results, NR6A1 T allele was predominant in agreement with previous works. Lithuanian indigenous wattle and Turopolje showed the highest introgression of the C allele (0.39 and 0.43, respectively), associated with a shorter carcass and this is in agreement with data obtained from carcass measurements within TREASURE project (personal communication; results not yet published). The highest frequency of the C allele in Turopolje breed, one of the oldest pig breeds in Europe created for production in local forests, could probably be related to a long history of contacts with wild boars, as well as, the absence of any major selection or crossbreeding program [74]. Also Alentejana, Apulo-Calabrese, Majorcan Black, Black Slavonian, Iberian and Nero Siciliano showed small introgression of the wild allele which could be originated by uncontrolled mating with wild boars (frequencies ranging from 0.05 to 0.27) as these breeds are reared mainly in free range systems. Vertebrae number is an important trait in pig production as it influences the size of important meat cuts such as the loin and also has an overall effect on carcass conformation. Moreover, segregation of markers leading to changes in carcass conformation results in increased carcass heterogeneity, thus selection of the favorable T allele would be highly advisable, especially in Lithuanian indigenous wattle, Turopolje and Nero Siciliano breeds.

A missense mutation (p.G32E) of Chinese origin in the gene encoding the peroxisome proliferator-activated receptor delta (*PPARD*) gene was shown to be a Quantitative Trait Nucleotide (QTN) affecting the ear size in pigs [47]. The *PPARD* gene has roles in skin and cartilage development besides affecting fat metabolism. The Chinese allele A, which increases ear size, is supposed to be absent in European breeds, but it has been detected in Large Black [47]. According to our results, the Chinese allele is in general scarce but present at low or moderate frequencies in several breeds including Bísara (0.17), Black Slavonian (0.08), Krškopolje (0.13), Lithuanian indigenous wattle (0.18) and Swallow-Bellied Mangalitsa (0.08). A small introgression of Large Black in local Lithuanian pigs was made before and after World War I [75], therefore ear shape and size typical for Large Black could be found among Lithuanian indigenous wattle pigs. Also, the introgression of Large Black in Black Slavonian breed is already known, as previously discussed, and could explain the presence of allele A in this breed. The presence of the Chinese allele in Bísara is also compatible with the likely crossbreeding between this and Chinese or English breeds [76].

### Markers related to meat production and carcass traits

Nine polymorphisms related to meat production and carcass traits were studied (Table 3).

Porcine stress syndrome (PSS) or Malignant hyperthermia is a condition that develops in homozygous animals carrying a single coding nucleotide substitution in the *RYR1* gene (c.1843C>T; p.R615C), upon exposure to halogenated anaesthetics [50]. The *RYR1* gene codes for the skeletal ryanodine receptor, a protein structurally involved in calcium channels controlling movements of  $Ca^{2+}$  from the sarcoplasmic reticulum to the cytosol. PSS is of concern to the worldwide swine industry from both welfare and an economic perspective as it results in losses during transport and very fast post-mortem pH decline causing denaturation of muscle proteins, water loss and formation of pale, soft, exudative meats (PSE) with reduced meat quality and commercial value [50]. Among local breeds involved in the present study, as in most commercial breeds [50], the c.1843T mutant allele is scarce since many initiatives have been carried out to eliminate this allele. However, a relevant frequency was found (0.21) in Krškopolje breed (Table 3). High incidence of this mutation in Krškopolje breed could have likely arisen due to crossing with German Landrace in times when the breed was cast out, as it was for many years the most used terminal line in the breed's recent history and it has a quite important incidence of this mutation. Although the presence of this allele in the population is known for some time, no efforts have been made by the breeders to eliminate it [50,77].

The insulin-like growth factor 2 (*IGF2*) intron 3 g.3072G>A (g.1483832G>A in Sscrofa11.1) substitution has been identified as a QTN for a paternally-expressed QTL affecting muscle growth, fat deposition, and heart size [31]. Pigs carrying the paternal A allele have higher lean growth and lower backfat thickness than those carrying the G allele [78,79]. The mutant A allele is common in breeds subjected to strong selection for lean meat content and would be of Asian origin [80], being predominant in commercial breeds such as Landrace or Large-White that had been historically introgressed with Asian breeds. This A allele is in general scarce in the breeds analyzed in the present work in agreement with the lack of selection for lean growth. However, it is present at considerable frequencies in a few breeds, such as Apulo-Calabrese who showed the highest frequency of the mutant allele (0.85), followed by Schwäbisch-Hällisches Schwein (0.50) and Krškopolje (0.39). In a pilot study of TREASURE project using modelling with InraPorc<sup>®</sup> software [81], Apulo-Calabrese and Krškopolje (Schwäbisch-Hällisches Schwein was not included in the study) showed the highest protein deposition, which could be related to the high frequencies of the mutant allele observed in

these populations. Overall, this incidence can be explained by introgression of alleles from commercial breeds.

Several genes encoding proteins with key roles in energy homeostasis were included in the SNP panel. These functionally related genes are involved in the melanocortin pathway that regulates feed intake and energy expenditure at hypothalamic level. The leptin (LEP) is the signaling molecule, produced in adipocytes, which informs the central nervous system about the increase in energy stores, unchaining an anorexigenic signal through its hypothalamic receptor (LEPR). The melanocortin receptor 4 (MC4R) is a downstream molecule involved in transmitting the LEP signal. The fat mass and obesity-associated protein (*FTO*) gene is also most likely involved in the regulation of energy balance and feed intake [82]. Thus, all these genes are candidates for a role in the variability of feed intake regulation, growth and fattening [37,83]. The SNPs selected for these genes have been proposed as useful markers for future breeding purposes, with different levels of evidence for causality. For example, the *LEPR* gene mutation has been related to several productive, fatness and meat quality traits in different genetic backgrounds [36,84–86] and it has been also associated with the hypothalamic expression of *LEPR* and downstream molecules [37], suggesting functionality of this particular variant. *LEPR*.c.1987T allele is systematically associated with higher fatness and feed intake and is fixed in the obese Iberian pig breed. Our results match this finding and moreover show that allele T is almost fixed in Alentejana (0.98), a very close breed with the same origin in the Iberian Peninsula. Besides these two close breeds, allele T is in general scarce in most of the remaining breeds except in Majorcan Black and Sarda, which show intermediate frequencies, opening selection possibilities to increase meat quality, because of its interesting association with intramuscular fat and fatty acid profile. However, effects on intake, growth and fatness should be evaluated and considered in each particular breed. It is interesting to highlight that the *LEPR* favorable T allele is completely absent in Basque, Mora Romagnola and Turopolje breeds; and at very low frequency in many other breeds, also characterized by a fat phenotype. Nevertheless, this results agrees with the polygenic nature of fatness and meat quality traits. Regarding the *MC4R* gene, different SNPs have been previously evaluated. Among them, the c.892G>A (p.D298N) is usually considered the most relevant one, being associated with variation in growth and fatness traits in most breeds and crosses, but with some discrepancies among studies [40,87,88]. In general, allele A is considered to be associated with a high feed intake and high lipid deposition, although this allele was absent (Basque) or at very low frequencies in many of our breeds (Alentejana, Bísara, Black Slavonian, Iberian, Majorcan Black and Swallow-Bellied Mangalitsa). In our study, c.892G>A SNP showed intermediate frequencies (0.30 to 0.70) in Apulo-Calabrese, Cinta Senese, Gascon, Krškopolje, both Lithuanian breeds, Moravka and Sarda, being susceptible for association studies with growth and fatness. The SNPs located in the *LEP* and *FTO* genes [35,89] are segregating at intermediate frequencies in most analyzed populations. These SNPs are most probably markers in linkage disequilibrium with the causal mutation, since their effect on production traits have not been established for all the studied breeds and thus the future validation of their effects is essential.

Myostatin (MSTN) is a member of transforming growth factor- $\beta$  (TGF- $\beta$ ) superfamily. It is a negative regulator for both embryonic development and adult muscle homeostasis. Previous works have detected and analyzed three SNPs (g.435G>A, g.447A>G and g.879T>A), located in the *MSTN* promoter region [90], concluding a relationship between these markers and muscle development. These mutations have been proposed as causal because of their location coincident with transcription factor binding sites, and their effect on mRNA expression [90]. We selected for genotyping the g.435G>A polymorphism (g.94629248C>T in Sscrofa11.1), which co-segregates with g.447A>G in several breeds [90–92]. According to previous works, the g.435A and g.447G alleles have positive effects on total muscle production and a negative effect

on fat deposition. With a few exceptions (Gascon, Iberian and Mora Romagnola, which show a very high frequency of g.435A allele responsible for muscle growth increase), intermediate frequencies were observed in most of our breeds, thus allowing further association studies to be conducted in order to support its association as in most of them the minor allele frequency was between 0.11 and 0.49.

Taste 2 Receptors (*TAS2R*) gene family encodes for receptors involved in taste perception and bitter sensing. Among them, *TAS2R39* and *TAS2R4* genes have been recently reported to be related to backfat thickness, probably as a consequence of different feed intake, due to different feed taste perception [53]. For *TAS2R39* gene, the missense mutation, p.N71T, was included in our panel. Allele G was predominant or fixed in most breeds although some others showed moderate to intermediate frequency of T allele (0.21 to 0.29), such as Majorcan Black, Cinta Senese, Black Slavonian, Krškopolje or Sarda, in which the SNP may be used in association studies. Such studies could deepen in the effects and causality of the marker and the potential mechanism for affecting fat deposition (voluntary feed intake). Finally, the marker in *TAS2R4* gene showed the A allele fixed in all breeds.

### Markers related to meat and fat quality

The fatty acid profile is one of the main determinants of animal product quality and is one of the distinguishing characteristics of high-valued Iberian meat products. Fatty acid profile has profound effects on the nutritional, sensorial and technological properties of meat [93]. Different candidate genes, mainly involved in lipogenesis, have been proposed to explain the variability of fatty acid content in animal tissues (Table 4). In most of the cases, the detected SNPs are probably neutral markers in linkage disequilibrium with the causal mutation, as it would be the case for fatty acid synthase (*FASN*), acetyl CoA carboxylase alpha (*ACACA*) and microsomal triglyceride transfer protein (*MTTP*) genes, included in our panel. The *ACACA* and *FASN* genes are involved in *de novo* lipogenesis, and mutations in these genes have been associated to MUFA and SFA content [25]. The missense c.2573T>C mutation in the porcine *MTTP* gene is highly associated with the fatty acid composition in the pig [42], influencing the percentages of oleic, palmitic and linoleic acids. The C allele (coding for Leucine) is also associated with an increased *MTTP* lipid transfer activity. For these mutations a high level of segregation is observed in most of the breeds. Association studies should be performed in order to validate the mutational effects on the different genetic backgrounds and evaluate their usefulness in breeding programs targeting the improvement of meat quality.

Mutations in the stearoyl CoA desaturase (*SCD*) and long-chain acyl-CoA synthetase (*ACSL4*) genes may be considered causal mutations of effects on meat fatty acid profile, specifically regarding MUFA content. The *SCD* is a very relevant enzyme because it catalyzes the desaturation of palmitic and stearic to palmitoleic and oleic acids. For this gene a mutation g.2228T>C (g.111461751C>T in Scrofa11.1), positioned in the core sequence of several putative transcription factor binding sites, has been repeatedly associated to MUFA content in different genetic backgrounds [52,94]. There are several plausible mechanisms by which allele T enhances 18:1/18:0 ratio and, consequently, the proportion of monounsaturated to saturated fatty acids, increasing meat quality. *ACSL4* protein catalyzes the formation of long-chain acyl-CoA from fatty acid, ATP and CoA, playing an important role in both lipid biosynthesis and fatty acid turnover. The mutation c.2645G>A SNP located in the 3' untranslated region has been associated with the percentages of oleic and MUFA. The G allele is also associated to higher *ACSL4* mRNA expression levels in liver than the A allele [15]. These SNPs may be used in breeding schemes with the purpose of improving meat quality by favoring a better fatty acid profile in terms of nutritional, organoleptic, technological and manufacturing properties. In



the case of *SCD* mutation, the favorable allele is already at high frequencies in most populations, in agreement with the high desaturation potential and favorable fatty acid profile and meat quality parameters observed in these breeds [74,94–101] but there are few as Black Slavonian and Mora Romagnola with selection possibilities. For *ACSL4*, in contrast, high level of segregation is observed in most breeds.

Boar taint is an unpleasant odor that influences the smell and taste of cooked pork from non-castrated male pigs. This defect leads to losses in carcass value resulting in economic cost to the industry. Its main cause is the accumulation of androstenone and skatole in fat tissues. Alternatives to surgical castration for the control of this problem include selective breeding and in consequence many studies have addressed the genetic variation underlying androstenone and skatole levels. Many QTLs have been found [102] and a few candidate genes have been identified. Our panel included two genes involved in the levels of androstenone and skatole in porcine tissues. The cytochrome P450 II E1 (*CYP2E1*) gene encodes an enzyme involved in the degradation of skatole in the liver, being negatively correlated with its accumulation; the g.2412C>T (g. 141690107C>T in Sscrofa11.1) polymorphism located in the promoter of *CYP2E1* gene accounts for more than a 10% of total phenotypic variance of skatole content in backfat, the C allele being associated to higher skatole levels [22,23]. This unfavorable allele has high frequency in several of the breeds included in the current study, but intermediate or low in others such as Gascon, Moravka or Turopolje ( $\leq 0.36$ ). The cytochrome b5 (*CYB5A*) gene codify for a protein involved in androstenone biosynthesis. A c.-8T>G polymorphism was found located 8bp upstream the start codon [20] and c.-8T allele was associated with low activity of *CYB5A* protein and low fat androstenone levels in different pig populations [20,21]. According to our results, the favorable c.-8T allele is scarce, with very low frequency in most of the examined local breeds (0 to 0.31). Only Iberian and Alentejana breeds showed intermediate frequencies for this polymorphism ( $\approx 0.5$ ). Thus, in general, boar taint genes offer a wide margin for genetic selection and improvement in most of our breeds. This possibility is especially interesting in local breeds in which high fatness is characteristic and slaughter weights are usually very high, as body fat is positively associated with boar taint [103] and boar taint increases with body weight at slaughter [104].

Proteolytic enzymes participating in *postmortem* proteolysis are determinant factors in meat tenderization thus affecting eating quality. The calpastatin (*CAST*) gene codes for an inhibitor of proteases whose activity is highly related to meat tenderness in different species. More than 900 polymorphisms have been identified in this gene [19], the SNP g.49223G>A (g.103299934G>A, in Sscrofa11.1) here analyzed has been associated to tenderness and juiciness, being G the favorable allele [105]. Calpains are intracellular cysteine proteases also influencing *post-mortem* meat processes. Among them the calpain S1 (*CAPNS1*) polymorphism c.429A>C has been associated with pH, conductivity and meat color in different experimental and commercial pig populations [18]. Allele C was associated to higher pH, lower conductivity, lower cooking loss and lower lightness, thus reflecting better meat quality. Both SNPs showed a high level of segregation in local breeds, allowing the performance of association studies in all but Krškopolje and Turopolje breeds for *CAST* and *CAPNS1* genes, respectively, in which the favorable alleles are already at high frequency or even fixed.

Some genes involved in fat metabolism and deposition may also influence meat quality parameters. The peroxisome proliferator activated receptor gamma coactivator 1A (*PPARGC1A*) gene codes for a transcription coactivator involved in adipogenesis, energy metabolism and muscle fiber determination, favoring oxidative type fibers. The *PPARGC1A* c.1288T>A polymorphism was associated with pH and cooking loss in Duroc×Pietrain experimental cross and in Italian Large White and Landrace [18] showing additive and dominant effects. For this polymorphism segregation was observed in all breeds but Alentejana, Cinta



Senese, Iberian and Turopolje. The adiponectin (*ADIPOQ*) gene encodes an adipokine involved in glucose and lipid metabolism. The non-synonymous polymorphism c.1735G>A had been previously associated with fat deposition [106] and was recently shown to have significant additive and dominant effects on several sensory traits such as flavor, tenderness, juiciness and meat color, with genotype AA being the most favorable [16]. In the investigated breeds, A allele is scarce which is in agreement with previous results in commercial crossbred pigs (frequency of A allele = 0.14) and would allow the implementation of selection procedures, or even absent in some ones, such as Apulo-Calabrese, Casertana, Cinta Senese, Gascon, Swallow-Bellied Mangalitsa and Turopolje.

Glycogen content is determinant of multiple meat quality traits, such as pH, meat color, drip loss or tenderness. One of the first genes identified affecting these meat quality traits in pigs was *PRKAG3* [48]. The protein encoded by this gene is a regulatory subunit of the AMP-activated protein kinase (AMPK) which plays a key role in the regulation of glucose and energy metabolism in skeletal muscle. Two main missense mutations have been widely studied in this gene. The *PRKAG3* p.R200Q mutation appears in Hampshire breed or derived synthetic lines. The mutated dominant allele leads to high glycogen deposition in muscle, high glycolytic potential and lactate concentration, decreasing meat quality (low pH, pale meat and low yield). This mutant A allele was absent in all local porcine breeds, in agreement with the main phenotypic characteristics of our local breeds and with previous works [107]. The p.V199I mutation, also involved in several meat quality traits, had been previously shown to segregate in different breeds, with allele A being most abundant in Iberian, Celta or Bísara, and less in most breeds selected for muscularity such as Duroc, Landrace and Pietrain [108]. The A allele, coding for isoleucine, leads to lower glycogen content, being favorable for meat quality [48]. According to our results this allele has a high frequency in Alentejana, Bísara, Casertana, Iberian, Swallow-Bellied Mangalitsa and Turopolje, in agreement with previous evidences.

More recently, a splice mutation in the phosphorylase kinase catalytic subunit gamma 1 (*PHKG1*) gene has been shown to be a causal mutation leading to high glycogen content and low meat quality in pig muscle [46]. This gene was studied as a positional candidate for a QTL located in SSC3 affecting glycolytic potential. The point mutation found produces a 32 bp deletion in the ORF and induces a premature stop codon; nonsense mediated decay of the transcript and ultimately lowers protein content and activity. The mutant A allele causes a 43% increase of glycolytic potential and a >20% decrease of water-holding capacity, with consistent negative effects on meat quality in different genetic backgrounds with Duroc origin [46]. According to our results, the favorable allele is predominant, either fixed or almost fixed in most breeds although in few of them (Apulo-Calabrese, Krškopolje, Lithuanian indigenous wattle, Moravka, and Sarda) the unfavorable A allele is at a relevant frequency ranging from 0.32 to 0.56.

The phosphoenolpyruvate carboxykinase (*PCK1*) gene codes for a regulatory molecule involved in gluconeogenesis. A polymorphism in this gene (c.2456A>C, p.M139L) has been recently reported [45] as causal mutation associated to intramuscular fat content, backfat thickness and meat quality in pigs. The c.2456C allele encoding leucine has negative effects on these traits, and is present in many breeds or crosses, especially in the ones selected for high lean content, and even in wild pigs although at substantial different frequencies. The *PCK1* c.2456A allele, the one that encodes methionine and is associated to both less water losses and more favorable fat distribution, seems to be overrepresented in breeds or crosses not subjected to strong artificial selection [45]. Our results show that, in fact, a few traditional non-selected breeds, such as Iberian and Alentejana, have the favorable allele at a very high frequency, or even fixed in the case of Swallow-Bellied Mangalitsa. Conversely, the rest of analysed breeds show small (Gascon, Lithuanian indigenous wattle and Turopolje) or intermediate

frequencies. Despite the fatness of Lithuanian indigenous wattle pigs their intramuscular fat content is quite low. In these breeds, selection of this PCK1 c.2456A allele could favor the desirable phenotype in pig breeding schemes, with increased intramuscular fat content, better meat quality and reduced amounts of subcutaneous and visceral fat.

### Markers related to disease resistance

Three genes involved in disease resistance traits were explored (Table 5). The guanylate binding protein 5 (*GBP5*) gene is known to play a role in host immune response and inflammation based on studies in the knockout mouse. The rs340943904 mutation in the pig gene is a strong candidate causal mutation for a QTL on SSC4 that controls variation in host response to Porcine Respiratory and Reproductive Syndrome (PRRS) virus [29]. This intronic mutation introduces a splice acceptor site which leads to a shifted reading frame and early stop codon that truncates the 88 C-terminal amino acids of the protein. The unfavorable G allele codes for the truncated GBP5 protein which is not able to inhibit viral entry and replication as quickly as the intact GBP5 protein [28] thus leading to a poor outcome of homozygous individuals following PRRS virus infection. Moreover, the quoted SNP has been shown to be a cis eQTL [30]. With the exception of three breeds showing intermediate frequencies (Cinta Senese, Lithuanian indigenous wattle and Mora Romagnola), the unfavorable allele shows very high frequency in all breeds (>0.73), opening potential selection possibilities. Marker assisted selection to increase the frequency of T allele, corresponding to increased PRRS resistance, has been proposed to reduce PRRS viral load, thus decreasing the costs associated with PRRS by reducing its incidence, because a lower viral burden may also reduce virus shedding [109]. Besides, such selection would be expected to improve weight gain under infection. However, the amount of response is limited because the SNP explains only a portion of the genetic variance in host response. Thus, genomic selection for viral load, in combination with marker-assisted selection on the GBP5 genotype, may hold potential for improved resistance to PRRS in some of the breeds.

Enterotoxigenic *Escherichia coli* (ETEC), that expresses the F4ab and F4ac fimbriae is a major cause of diarrhea outbreaks and mortality in the pig breeding industry, infecting both newborn and weaned piglets. The mucin 4 gene (*MUC4*) SNP located in intron 7 (DQ848681: g.8227C>G, g. 134226654C>G in Sscrofa11.1) is the most extensively studied polymorphism in relation to ETEC F4ab/ac susceptibility in pigs [43]. This SNP is also associated to growth performance, with significant effects on average daily gain and backfat thickness [110]. The C allele, associated with susceptibility, dominates the resistant G allele [111]. According to this work the polymorphism is in complete linkage disequilibrium with the susceptibility phenotype and is currently used as a genetic test in the Danish pig breeding industry. Our results indicate that the unfavorable allele is majoritarian or even fixed in most populations, with the exception of both Lithuanian and Bísara breeds in which intermediate frequencies ( $\leq 0.60$ ) are observed.

Weaned piglets are also susceptible to F18-positive *Escherichia coli* (ETECF18) infections leading to post-weaning diarrhea or edema disease. The F18 receptor (F18R) plays a crucial role in this disease by mediating the binding of F18 fimbriated bacteria to the intestinal epithelium with F18R positive animals being susceptible to infection. Susceptibility to this infection appears to be dependent on the activity of the alpha-fucosyltransferase-1 (*FUT1*) gene, considered the candidate gene controlling the adhesion to F18 receptor. In this gene a G/A transition resulting in an p.A103T amino acid substitution has been discovered which is involved in susceptibility, AA individuals are resistant to ETECF18 while GG or AG are sensitive. This mutation of *FUT1* gene has been also associated to resistance to natural infection by PRRS and

*Haemophilus parasuis* virus [27], thus it might play a role in pig infection by multi-pathogens, with AA being a favorable genotype for increasing the resistance to disease. The favorable A allele has also been associated to a better growth and development, better meat quality, lower fat content and higher fetal survival [112] and has been shown to be scarce in many pig breeds [27,113]. In most of our breeds, G allele is present at high or intermediate frequency, with only Krškopolje and Turopolje showing high frequency of the favorable allele A. Marker assisted selection to increase the frequency of the A allele not only would enhance the anti-disease ability, but would also improve the growth, meat quality, carcass traits and reproductive ability.

### Markers associated with prolificacy

The estrogen receptor 1 (*ESR1*) gene polymorphism, previously associated to litter size [24], was not polymorphic in any of the analyzed breeds. The allele responsible for a higher prolificacy, with Chinese origin and predominant in Meishan and also found in Large White [114], is absent in our local breeds, with the unfavorable allele in terms of litter size being fixed in all of them (Table 5). This result is in agreement with previously published reports in some of the analyzed breeds, such as Swallow-Bellied Mangalitsa [49] and with the general low reproductive ability of autochthonous breeds. Results suggest lack or low influence from Chinese pig breeds in those tested in this work.

The aryl hydrocarbon receptor (*AHR*) gene polymorphism shows segregation at intermediate frequencies in most of tested populations. Iberian and Alentejana breeds showed the wild allele almost fixed. For this Asian-derived non-synonymous mutation the mutant G allele has been associated with a substantial increase in litter size in multiple European commercial lines [17]. The gene has also been related to fertility and embryo development in other mammals [115,116]. Again, the genotyping results agree with the characteristic low prolificacy observed in local unselected breeds.

### Genetic diversity in candidate genes

The results of diversity parameters obtained for each tested and segregating marker are shown in Table 6. The observed heterozygosity ( $H_O$ ) and expected heterozygosity ( $H_S$ ) values per locus ranged from 0.024 to 0.414 and from 0.025 to 0.415, respectively, with overall values of 0.24 for both parameters.  $F_{ST}$  is a measure of population structure due to genetic structure and its values range from 0 to 1. While a zero value indicates there is no genetic structures and the populations are in complete panmixia, a value of one implies there are not gene flux between the populations. The overall  $F_{ST}$  value from all tested loci was 0.25, pointing out that a 25% of genetic differences are due to differences between breeds and a 75% is caused by differences among individuals, therefore, the breeds studied share some genetic diversity. A similar value ( $F_{ST} = 0.27$ ) [117] was also observed in an study including 11 European breeds (Pietrain, Sortbroget, Basque, Gascon, Limousin, Normand, German Landrace, Schwäbisch-Hällisches, Great Yorkshire, Swedish Landrace, European Wild Pig). Highest  $F_{ST}$  values were obtained for *MC1R* and *KIT* loci, meaning that these markers show the highest degree of differentiation among populations, whereas the lowest values were observed for *FASN*, *PPARD* and *RYRI* polymorphisms.

The global  $H_O$ ,  $H_S$  and  $F_{IS}$  values for each breed are shown in Table 7. The highest  $H_O$  and  $H_S$  values are observed in Moravka, Krškopolje and Sarda indicating a higher degree of genetic diversity than other breeds such as Basque, Alentejana and Iberian breeds, which exhibit the lowest values.  $F_{IS}$  estimates the departure from panmixia at the subpopulation level, and is a measure of inbreeding; negative values indicate there are less related individuals than expected by random mating and positive values indicate there are more related individuals than

**Table 6. Observed ( $H_o$ ) and expected heterozygosity ( $H_s$ ) and  $F_{ST}$  values by locus, averaged among all the breeds.**

Gene	$H_o$	$H_s$	$F_{ST}$
ACACA	0.424	0.425	0.122
ACSL4	0.198	0.331	0.260
ADIPOQ	0.113	0.115	0.137
AHR	0.379	0.394	0.158
CAPNS1	0.378	0.382	0.221
CAST	0.295	0.389	0.206
CYB5A	0.195	0.194	0.177
CYP2E1	0.331	0.329	0.232
FASN	0.660	0.471	0.053
FTO	0.371	0.382	0.220
FUT1	0.294	0.302	0.231
GBP5	0.237	0.241	0.159
IGF2	0.185	0.200	0.298
KIT	0.099	0.094	0.591
LEP	0.364	0.365	0.270
LEPR	0.223	0.274	0.384
MC1R	0.280	0.265	0.601
MC4R	0.321	0.306	0.269
MSTN	0.354	0.352	0.253
MTTP	0.414	0.415	0.150
MUC4	0.282	0.269	0.112
NR6A1	0.132	0.129	0.201
PCK1	0.330	0.343	0.304
PHKG1	0.052	0.150	0.313
PPARD	0.074	0.079	0.065
PPARGC1A	0.378	0.375	0.248
PRKAG3 I199V	0.399	0.388	0.214
RYR1	0.048	0.053	0.076
SCD	0.228	0.218	0.269
TAS2R39	0.167	0.172	0.129
TYRP1	0.024	0.025	0.103

<https://doi.org/10.1371/journal.pone.0207475.t006>

expected by random mating. According to this statistic, Casertana and Apulo Calabrese are the breeds displaying the highest inbreeding, in agreement with their endangered situation and small census [99,118,119].

Nei's genetic distances ranged from 0.007 between Alentejana and Iberian breeds to 0.240 between Black Slavonian and Turopolje. The  $F_{ST}$  pairwise estimates ranged from 0.039 between Alentejana and Iberian breeds to 0.524 between Mora Romagnola and Turopolje (S2 and S3 Tables). Fig 2A and 2B show the phylogenetic trees constructed using these distances, bootstrapping values were very low, which point out there is not enough information to solve the phylogenetic trees. However, both trees agree at least partially with geographic distribution of breeds, with those that are geographically close grouping into the same branches, as Alentejana, Iberian and Majorcan Black and the two Lithuanian breeds. Regarding the breeds coming from the Iberian Peninsula, the Portuguese Bisara breed is, according to SNP genotyping, apparently unrelated to the rest. Herrero-Medrano et al. [120] observed this separation between Iberian and Bisara breeds, which agrees with their different Iberian and Celtic origins [121].

**Table 7. Diversity parameters per breed.**

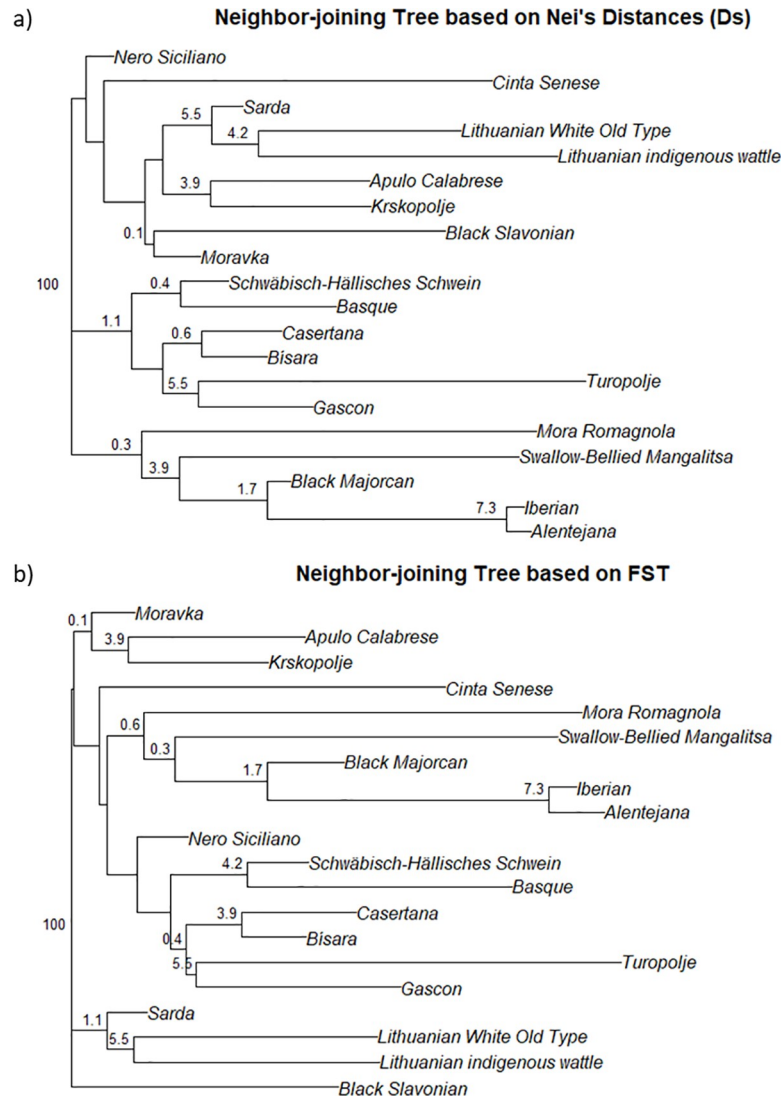
Breed	H <sub>O</sub>	H <sub>S</sub>	F <sub>IS</sub>
Alentejana	0.155	0.170	0.062
Apulo Calabrese	0.223	0.260	0.115
Basque	0.153	0.152	0.024
Bísara	0.259	0.274	0.065
Black Slavonian	0.288	0.271	-0.061
Casertana	0.216	0.239	0.131
Cinta Senese	0.247	0.249	0.016
Gascon	0.224	0.214	-0.003
Iberian	0.165	0.185	0.084
Krškopolje	0.310	0.302	-0.029
Lithuanian indigenous wattle	0.281	0.288	0.009
Lithuanian White Old Type	0.237	0.233	0.022
Majorcan Black	0.232	0.238	0.041
Mora Romagnola	0.195	0.187	-0.001
Moravka	0.310	0.315	0.007
Nero Siciliano	0.260	0.281	0.074
Sarda	0.297	0.325	0.085
Schwäbisch-Hällisches Schwein	0.250	0.250	0.006
Swallow-Bellied Mangalitsa	0.191	0.197	0.050
Turopolje	0.209	0.186	-0.096

<https://doi.org/10.1371/journal.pone.0207475.t007>

### Genetic structure of populations using DAPC and STRUCTURE

Based on the lowest BIC value found (S1 Fig) a total of 20 clusters were detected, which correspond to the number of the original local European pig breeds considered in the analysis. DAPC analysis was thus performed retaining 20 PCs, which explain more than 90% of the total variance, and seven discriminant eigenvalues. The resulting representation of the first two Linear Discriminants (Fig 3) was not able to individualize clusters formed by individuals belonging to a specific breed but it is also true that variance explained by these first two components represented only 20% of total variance. Both  $\alpha$ -score optimization (S2 Fig) and cross validation (S3 Fig) procedures identify 19 PCs to be retained, which is quite close to the value of 20 found with DAPC. Actually, the lowest RMSE for CV corresponds to 30 PCs but it was only slightly lower than the value found for 19 PCs (0.181 vs 0.191), both obtaining a mean successful assignment greater than 80%. Furthermore, the scatterplot of the DAPC cross-validation is rather flat after 15 retained PCs and thus it is not advisable retaining too many PC axes in order to not create an excessively complex model, not suitable to analyze unseen data.

Group membership based on the retained discriminant functions (Fig 4) showed a rather clear distinction among the different populations with high genetic structure for Mora Romagnola, Turopolje and Swallow-Bellied Mangalitsa pig breeds, followed by Cinta Senese, Gascon and Lithuanian indigenous wattle breeds which presented few admixed individuals within their respective clusters. For most of the other breeds, a general framework of admixed history is evident, even if no clear pattern is recognizable. On the contrary, Alentejana and Iberian breeds were basically included in two shared clusters separated from the other populations confirming the common past history of the two breeds. The proportion of successful reassignment using the 19 Linear Discriminant functions suggested by DAPC is showed in S4 Fig. Some breeds are almost fully reassigned to the original population (i.e. Turopolje, Lithuanian indigenous wattle, Mora Romagnola, Swallow-Bellied Mangalitsa, and Cinta Senese) whereas



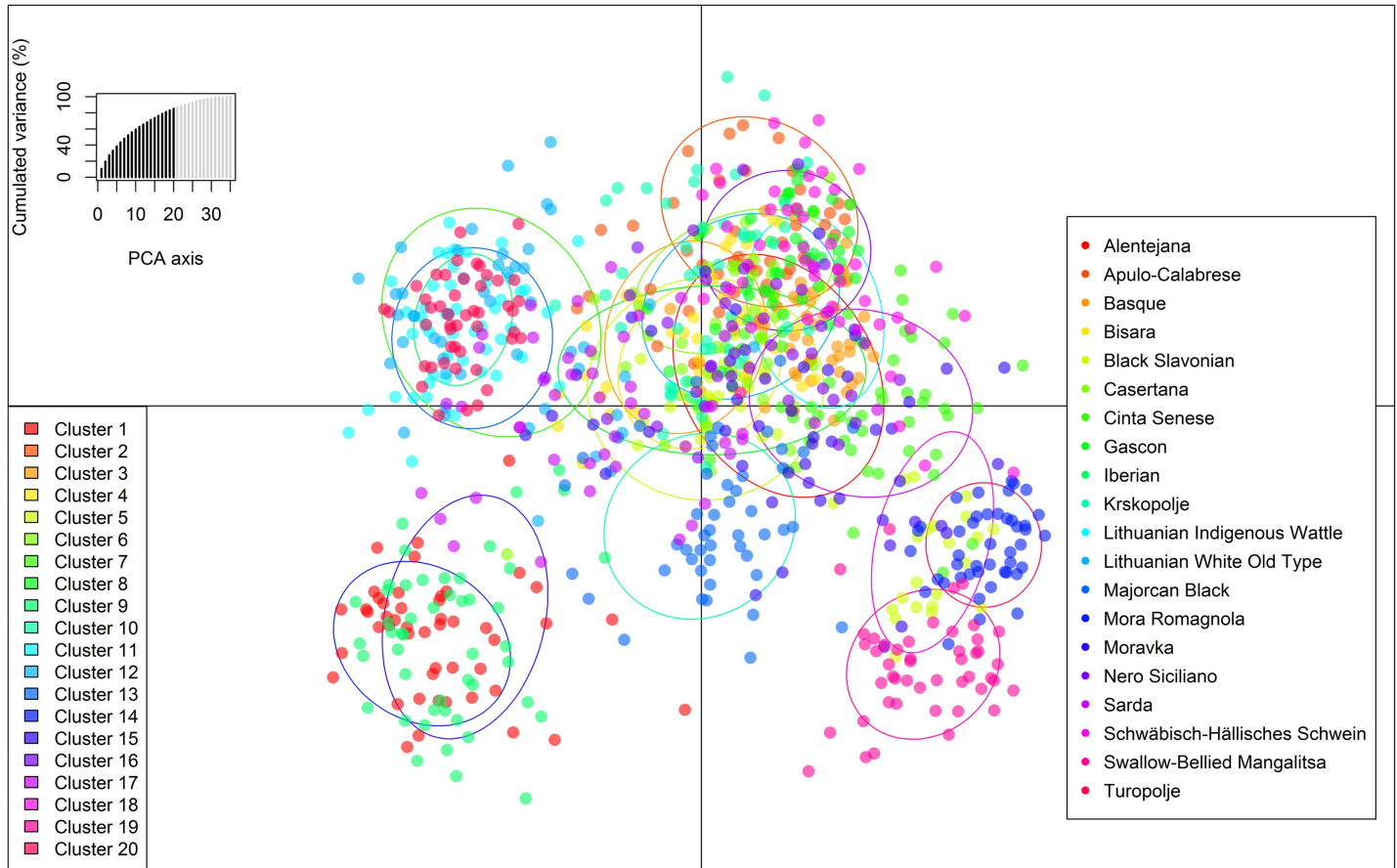
**Fig 2.** Neighbour-joining tree based on a) Nei's distances b) F<sub>ST</sub>.

<https://doi.org/10.1371/journal.pone.0207475.g002>

the lowest proportion of reassignment has been registered for Sarda, Nero Siciliano and Moravka breeds. Indeed, the breeds with a high proportion of successful reassignment are those with a clearly distinct pattern of membership assignment whereas the strongly admixed populations are those with values of correct reassignment below 80%. The same reassignments has been performed using randomized groups of individuals in order to be protected against the extraction of residual variation considered as explanatory of the underlying model structure (i.e. overfitting) (S5 Fig) and the results are quite comfortable; the highest successful proportion of reassignment was slightly higher than 30%, thus confirming the reliability of the discrimination process.

The genetic structure of the different populations was also studied through a Bayesian clustering method that requires high computational time and which accuracy strictly depends on the underlying population genetic model behind the original data [122]. The combination of the modal value of the  $\Delta K$  (S6A Fig) and the mean variance (S6B Fig) of the  $\ln \Pr(G|K)$  is used as an indicator of the strength of the signal detected by the software. Our results showed the

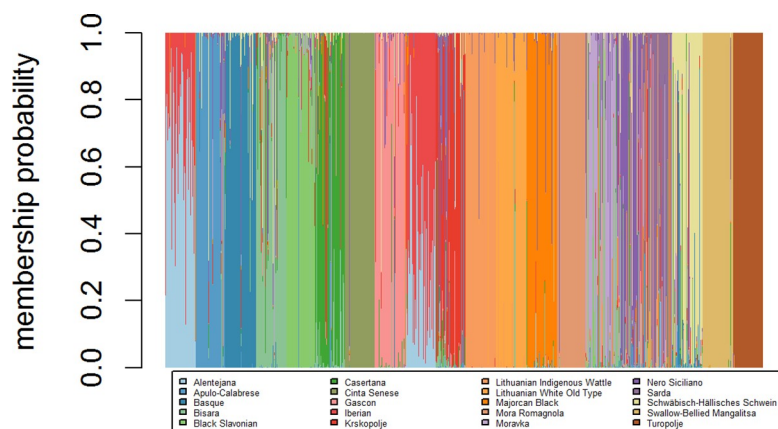




**Fig 3. Scatterplot of resulting genetic clusters after Discriminant Analysis of Principal Components for the 20 local European pig breeds.**

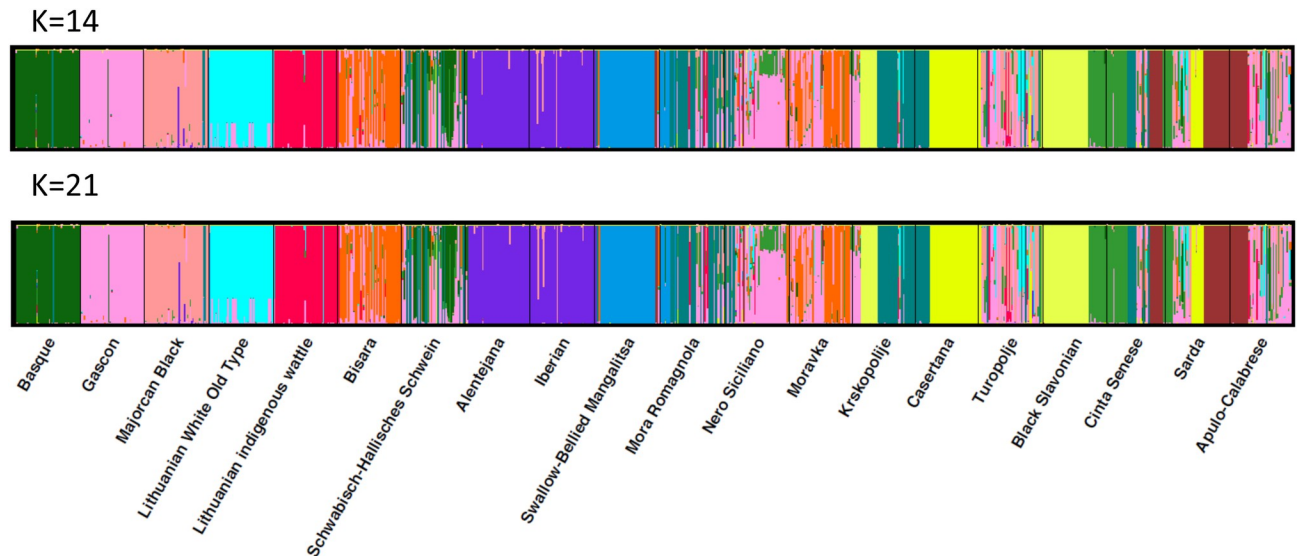
<https://doi.org/10.1371/journal.pone.0207475.g003>

strongest signal at  $k = 14$  with another significant peak at  $k = 21$  with mean and standard deviation of  $\text{LnP}(K)$  of  $-23894.17$  and  $-23124.02$ , and  $36.02$  and  $37.30$  for  $k = 14$  and  $k = 21$ , respectively.



**Fig 4. Membership assignment of the individuals to the *a priori* clusters defined with DAPC.**

<https://doi.org/10.1371/journal.pone.0207475.g004>



**Fig 5. Genetic diversity structure of the 20 pig populations.** Population memberships for each genotype is shown based on  $K = 14$  and  $K = 21$ .

<https://doi.org/10.1371/journal.pone.0207475.g005>

The level of admixture was uneven in the different populations when looking at the individual membership (Fig 5). As already stated with DAPC, Alentejana and Iberian breeds clustered together both at  $k = 14$  and  $k = 21$  but in general the software was able to better assign the individuals to specific clusters. The highest numbers of admixed individuals were detected for Bísara, Schwäbisch-Hällisches Schwein, Mora Romagnola, Nero Siciliano, Moravka and Turopolje breeds but also some other breeds (i.e. Cinta Senese, Sarda, and Apulo-Calabrese breeds) showed a certain degree of admixture among the individuals. The increase of  $k$  from 14 to 21 allows separating some breeds formerly clustered with other populations (f.i. Crna Slavonka and Casertana breeds) but it was not conclusive in terms of assignment, confirming the admixed history of the breeds. These results are not fully superimposable to those found with DAPC analysis, suggesting that the Bayesian clustering method was able to find hidden structures of the data, not easily recognizable with the previous one.

Results provide insights into the different population's genetic structure, relationships and clustering among them. Nevertheless, the genotyping data employed seems not sufficient to clarify the complex relationships in this large collection of phenotypically close pig breeds. The use of larger panels of neutral marker loci is probably advised in order to deepen in their genetic structure and to relate it with their breeding history.

### Final considerations

We report a comprehensive analysis of genetic variation and population genetic structure based on a panel of candidate genes in a wide panel of European local pig breeds. Results are in agreement with known facts of the breeds' origin and their phenotype. A clear genetic differentiation and intra-breed homogeneity was observed in some cases, while other breeds showed a high degree of admixture. Alleles with contrasted effects on production and fatness (such as *LEPR*, *FTO*, *MC4R*, *LEP* or *MSTN*), meat quality (*PCK1*, *PRKAG3*, *ACACA*, *CAST*, *MTTP*) or disease resistance (*MUC4*, *GBP5*) are segregating in many breeds, in some cases with intermediate frequencies, opening selection possibilities. Results provide an extended catalogue of fixed and segregating candidate SNPs associated with complex traits with potential usefulness for traceability purposes and for association studies and breeding programs. For a

practical application, the percentage of variance of the traits of interest explained by the polymorphisms must now be further studied in each breed. If sufficient to justify the genotyping cost, a marker assisted selection could be implemented.

## Supporting information

### **S1 Table. Genotyping data.**

(XLSX)

### **S2 Table. Nei's genetic distances computed between each pair of pig populations.**

(DOCX)

### **S3 Table. FST computed between each pair of pig populations.**

(DOCX)

### **S1 Fig. Graph of BIC value vs. number of detected clusters.**

(TIF)

### **S2 Fig. Graph of $\alpha$ -score optimization of DAPC.**

(TIF)

### **S3 Fig. Graph of cross-validation of DAPC.** x axis reports the number of PC retained in each DAPC and y axis the proportion of successful outcome prediction.

(TIF)

### **S4 Fig. Proportion of successful reassignment of the individuals to their specific breed.**

(TIFF)

### **S5 Fig. Proportion of successful reassignment using randomized groups of individuals.**

(TIFF)

### **S6 Fig. The most likely number of clusters (K) in the analyzed dataset.** a) plot of $\ln \Pr(G|K)$ values as a function of the number of clusters for each of the 20 runs carried out for each K value using the program STRUCTURE. b) plot of K as the mean of the absolute values of $L''(K)$ averaged over 20 runs divided by the standard deviation of $L(K)$ between successive K values.

(TIF)

## Acknowledgments

This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634476 (project acronym TREASURE). The content of this paper reflects only the author's view and the European Union Agency is not responsible for any use that may be made of the information it contains. MČP and MŠ were co-financed by Slovenian Research Agency (grant P4-0133).

## Author Contributions

**Conceptualization:** Luca Fontanesi, Ana I. Fernández, Cristina Óvilo.

**Data curation:** María Muñoz.

**Formal analysis:** María Muñoz.

**Funding acquisition:** Luca Fontanesi, Ana I. Fernández.

**Investigation:** Riccardo Bozzi, Yolanda Núñez, Claudia Geraci, Juan García-Casco, Estefania Alves, Martin Škrlep, Jordi Estellé, Cristina Óvilo.

**Methodology:** Fabián García, Yolanda Núñez, Claudia Geraci, Alessandro Croveti, Estefania Alves.

**Project administration:** Marjeta Čandek-Potokar, Cristina Óvilo.

**Resources:** Fabián García, Martin Škrlep, Rui Charneca, Jose M. Martins, Raquel Quintanilla, Joan Tibau, Goran Kušec, Ivona Djurkin-Kušec, Marie J. Mercat, Juliette Riquet, Christoph Zimmer, Violeta Razmaite, Jose P. Araujo, Čedomir Radović, Radomir Savić, Danijel Karolyi, Maurizio Gallo, Marjeta Čandek-Potokar, Luca Fontanesi.

**Supervision:** Riccardo Bozzi, Marjeta Čandek-Potokar, Luca Fontanesi, Cristina Óvilo.

**Writing – original draft:** María Muñoz, Cristina Óvilo.

**Writing – review & editing:** Riccardo Bozzi, Fabián García, Yolanda Núñez, Claudia Geraci, Alessandro Croveti, Juan García-Casco, Estefania Alves, Martin Škrlep, Rui Charneca, Jose M. Martins, Raquel Quintanilla, Joan Tibau, Goran Kušec, Ivona Djurkin-Kušec, Marie J. Mercat, Juliette Riquet, Jordi Estellé, Christoph Zimmer, Violeta Razmaite, Jose P. Araujo, Čedomir Radović, Radomir Savić, Danijel Karolyi, Marjeta Čandek-Potokar, Luca Fontanesi, Ana I. Fernández.

## References

1. Barker JSF. Conservation of livestock breed diversity. *Anim Genet Resour génétiques Anim généticos Anim*. Cambridge University Press; 1999; 25: 33–43.
2. Kirk H, Freeland JR. Applications and implications of neutral versus non-neutral markers in molecular ecology. *Int J Mol Sci. Molecular Diversity Preservation International*; 2011; 12: 3966–3988. <https://doi.org/10.3390/ijms12063966> PMID: 21747718
3. Wright D. The Genetic Architecture of Domestication in Animals. *Bioinform Biol Insights. Libertas Academica*; 2015; 9: 11–20. <https://doi.org/10.4137/BBI.S28902> PMID: 26512200
4. Montgomery GW, Sise JA. Extraction of dna from sheep white blood cells. *New Zeal J Agric Res*. 1990; 33: 437–441. <https://doi.org/10.1080/00288233.1990.10428440>
5. Nei M. *Molecular evolutionary genetics*. Columbia university press; 1987.
6. Takezaki N, Nei M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*. 1996; 144: 389–399. PMID: 8878702
7. Hierstat Goudet J., a package for R to compute and test hierarchical F-statistics. *Mol Ecol Resour*. 2005; 5: 184–186.
8. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*. 2004; <https://doi.org/10.1093/bioinformatics/btg412>
9. Jombart T. ADEGENET: a R package for multivariate analysis of genetic markers. *Bioinformatics*. 2008; 24: 1403–1405. <http://dx.doi.org/10.1093/bioinformatics/btn129> PMID: 18397895
10. R Foundation for Statistical Computing. R: A language and environment for statistical computing. R: A Language and Environment for Statistical Computing. 2018.
11. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155: 945–959. <https://doi.org/10.1111/j.1471-8286.2007.01758.x> PMID: 10835412
12. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol*. 2005; 14: 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> PMID: 15969739
13. Jakobsson M, Rosenberg NA. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 2007; 23: 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233> PMID: 17485429
14. Muñoz M, Fernandez AI, Benitez R, Pena RN, Folch JM, Rodriguez Mdel C, et al. Disentangling two QTL on porcine chromosome 12 for backfat fatty acid composition. *Anim Biotechnol*. 2013/06/20. 2013; 24: 168–186. <https://doi.org/10.1080/10495398.2012.763130> PMID: 23777347

15. Corominas J, Ramayo-Caldas Y, Castelló A, Muñoz M, Ibáñez-Escriche N, Folch JM, et al. Evaluation of the porcine ACSL4 gene as a candidate gene for meat quality traits in pigs. *Anim Genet.* 2012; 43: 714–720. <https://doi.org/10.1111/j.1365-2052.2012.02335.x> PMID: 22497636
16. Zhang CY, Wang Z, Bruce HL, Janz J, Goddard E, Moore S, et al. Associations between single nucleotide polymorphisms in 33 candidate genes and meat quality traits in commercial pigs. *Anim Genet.* 2014; 45: 508–516. <https://doi.org/10.1111/age.12155> PMID: 24707962
17. Bosse M, Megens HJ, Frantz LAF, Madsen O, Larson G, Paudel Y, et al. Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. *Nat Commun.* 2014;5. <https://doi.org/10.1038/ncomms5392> PMID: 25025832
18. Gandolfi G, Cinar MU, Ponsuksili S, Wimmers K, Tesfaye D, Looft C, et al. Association of PPARGC1A and CAPNS1 gene polymorphisms and expression with meat quality traits in pigs. *Meat Sci.* 2011; 89: 478–485. <https://doi.org/10.1016/j.meatsci.2011.05.015> PMID: 21680104
19. Meyers SN, Beever JE. Investigating the genetic basis of pork tenderness: Genomic analysis of porcine CAST. *Anim Genet.* 2008; 39: 531–543. <https://doi.org/10.1111/j.1365-2052.2008.01765.x> PMID: 18717668
20. Lin Z, Lou Y, Peacock J, Squires EJ. A novel polymorphism in the 5' untranslated region of the porcine cytochrome b5 (CYB5) gene is associated with decreased fat androstenone level. *Mamm genome.* Springer; 2005; 16: 367–373. PMID: 16104384
21. Peacock J, Lou Y, Lundström K, Squires EJ. The effect of a c.-8G>T polymorphism on the expression of cytochrome b5A and boar taint in pigs. *Anim Genet.* 2008; 39: 15–21. <https://doi.org/10.1111/j.1365-2052.2007.01674.x> PMID: 18162105
22. Skinner TM, Doran E, McGivan JD, Haley CS, Archibald AL. Cloning and mapping of the porcine cytochrome-p450 2E1 gene and its association with skatole levels in the domestic pig. *Anim Genet.* 2005; 36: 417–422. <https://doi.org/10.1111/j.1365-2052.2005.01342.x> PMID: 16167985
23. Mörlein D, Lungershausen M, Steinke K, Sharifi AR, Knorr C. A single nucleotide polymorphism in the CYP2E1 gene promoter affects skatole content in backfat of boars of two commercial Duroc-sired crossbred populations. *Meat Sci.* 2012; 92: 739–744. <https://doi.org/10.1016/j.meatsci.2012.06.031> PMID: 22795630
24. Rothschild MF, Larson RG, Jacobson C, Pearson P. Pvu II polymorphisms at the porcine oestrogen receptor locus ESR. *Anim Genet.* 1991; 22: 448–448. <https://doi.org/10.1111/j.1365-2052.1991.tb00715.x> PMID: 1685641
25. Muñoz G, Alves E, Fernández A, Óvilo C, Barragán C, Estellé J, et al. QTL detection on porcine chromosome 12 for fatty-acid composition and association analyses of the fatty acid synthase, gastric inhibitory polypeptide and acetyl-coenzyme A carboxylase alpha genes. *Anim Genet.* 2007; 38: 639–646. <https://doi.org/10.1111/j.1365-2052.2007.01668.x> PMID: 18052940
26. Óvilo C, C F, Y N, Benítez R, García F, Barragán C, et al. Influence of FTO and IRX3 genes on pig fatness and growth in different genetic backgrounds. IX International Symposium on Mediterranean Pig. 2016.
27. Wang SJ, Liu WJ, Yang LG, Sargent CA, Liu HB, Wang C, et al. Effects of FUT1 gene mutation on resistance to infectious disease. *Mol Biol Rep.* 2012; 39: 2805–2811. <https://doi.org/10.1007/s11033-011-1039-0> PMID: 21695432
28. Schroyen M, Easley C, Koltes JE, Fritz-Waters E, Choi I, Plastow GS, et al. Bioinformatic analyses in early host response to Porcine Reproductive and Respiratory Syndrome virus (PRRSV) reveals pathway differences between pigs with alternate genotypes for a major host response QTL. *BMC Genomics.* BioMed Central; 2016; 17: 196. <https://doi.org/10.1186/s12864-016-2547-z> PMID: 26951612
29. Koltes JE, Fritz-Waters E, Easley CJ, Choi I, Bao H, Kommadath A, et al. Identification of a putative quantitative trait nucleotide in guanylate binding protein 5 for host response to PRRS virus infection. *BMC Genomics.* 2015; 16. <https://doi.org/10.1186/s12864-015-1635-9> PMID: 26016888
30. Kommadath A, Bao H, Choi I, Reecy JM, Koltes JE, Fritz-Waters E, et al. Genetic architecture of gene expression underlying variation in host response to porcine reproductive and respiratory syndrome virus infection. *Sci Rep.* 2017; 7. <https://doi.org/10.1038/srep46203> PMID: 28393889
31. Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, et al. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature.* 2003; 425: 832–836. <https://doi.org/10.1038/nature02064> PMID: 14574411
32. Johansson A, Pielberg G, Andersson L, Edfors-Lilja I. Polymorphism at the porcine Dominant white/KIT locus influence coat colour and peripheral blood cell measures. *Anim Genet.* Wiley Online Library; 2005; 36: 288–296. <https://doi.org/10.1111/j.1365-2052.2005.01320.x> PMID: 16026338
33. Pielberg G, Olsson C, Syvänen AC, Andersson L. Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. *Genetics.* 2002; 160: 305–311. <https://doi.org/10.1038/335088a0> PMID: 11805065



34. Fontanesi L, D'Alessandro E, Scotti E, Liotta L, Crovetto A, Chiofalo V, et al. Genetic heterogeneity and selection signature at the KIT gene in pigs showing different coat colours and patterns. *Anim Genet.* 2010; 41: 478–492. <https://doi.org/10.1111/j.1365-2052.2010.02054.x> PMID: 20477793
35. Kennes YM, Murphy BD, Pothier F, Palin MF. Characterization of swine leptin (LEP) polymorphisms and their association with production traits. *Anim Genet.* 2001; 32: 215–218. <https://doi.org/10.1046/j.1365-2052.2001.00768.x> PMID: 11531701
36. Óvilo C, Fernández A, Noguera JL, Barragán C, Letón R, Rodríguez C, et al. Fine mapping of porcine chromosome 6 QTL and LEPR effects on body composition in multiple generations of an Iberian by Landrace intercross. *Genet Res.* 2005; 85: 57–67. <https://doi.org/10.1017/S0016672305007330> PMID: 16089036
37. Óvilo C, Fernández A, Fernández AI, Folch JM, Varona L, Benítez R, et al. Hypothalamic expression of porcine leptin receptor (LEPR), neuropeptide y (NPY), and cocaine- and amphetamine-regulated transcript (CART) genes is influenced by LEPR genotype. *Mamm Genome.* 2010; 21: 583–591. <https://doi.org/10.1007/s00335-010-9307-1> PMID: 21128076
38. Kijas JMH, Wales R, Törnsten A, Chardon P, Moller M, Andersson L. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics. Genetics Soc America;* 1998; 150: 1177–1185. PMID: 9799269
39. Kijas JMH, Moller M, Plastow G, Andersson L. A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. *Genetics.* 2001; 158: 779–785. PMID: 11404341
40. Kim KS, Larsen N, Short T, Plastow G, Rothschild MF. A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm genome.* Springer; 2000; 11: 131–135. PMID: 10656927
41. Tu PA, Lo LL, Chen YC, Hsu CC, Shiau JW, Lin EC, et al. Polymorphisms in the promoter region of myostatin gene are associated with carcass traits in pigs. *J Anim Breed Genet.* 2014; 131: 116–122. <https://doi.org/10.1111/jbg.12053> PMID: 24628723
42. Estellé J, Fernández AI, Pérez-Enciso M, Fernández A, Rodríguez C, Sánchez A, et al. A non-synonymous mutation in a conserved site of the MTTP gene is strongly associated with protein activity and fatty acid profile in pigs. *Anim Genet.* 2009; 40: 813–820. <https://doi.org/10.1111/j.1365-2052.2009.01922.x> PMID: 19496770
43. Schroyen M, Stinckens A, Verhelst R, Niewold T, Buys N. The search for the gene mutations underlying enterotoxigenic Escherichia coli F4ab/ac susceptibility in pigs: a review. *Vet Res. BioMed Central;* 2012; 43: 70. <https://doi.org/10.1186/1297-9716-43-70> PMID: 23061722
44. Fontanesi L, Ribani A, Scotti E, Utzeri VJ, Veličković N, Dall'Olio S. Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the MC1R and NR6A1 genes. *Meat Sci.* 2014; 98: 781–784. <https://doi.org/10.1016/j.meatsci.2014.07.026> PMID: 25134014
45. Latorre P, Burgos C, Hidalgo J, Varona L, Carrodeguas JA, López-Buesa P. C.A2456C-substitution in Pck1 changes the enzyme kinetic and functional properties modifying fat distribution in pigs. *Sci Rep.* 2016; 6. <https://doi.org/10.1038/srep19617> PMID: 26792594
46. Ma J, Yang J, Zhou L, Ren J, Liu X, Zhang H, et al. A Splice Mutation in the PHKG1 Gene Causes High Glycogen Content and Low Meat Quality in Pig Skeletal Muscle. *PLoS Genet.* 2014; 10. <https://doi.org/10.1371/journal.pgen.1004710> PMID: 25340394
47. Ren J, Duan Y, Qiao R, Yao F, Zhang Z, Yang B, et al. A missense mutation in PPARD causes a major QTL effect on ear size in pigs. *PLoS Genet.* 2011; 7. <https://doi.org/10.1371/journal.pgen.1002043> PMID: 21573137
48. Milan D, Jeon JT, Looft C, Amarger V, Robic A, Thelander M, et al. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science (80-).* 2000; 288: 1248–1251. <https://doi.org/10.1126/science.288.5469.1248>
49. Ciobanu DC, Day a E, Nagy a, Wales R, Rothschild MF, Plastow GS. Genetic variation in two conserved local Romanian pig breeds using type 1 DNA markers. *Genet Sel Evol.* 2001; 33: 417–32. <https://doi.org/10.1186/1297-9686-33-4-417> PMID: 11559484
50. Fujii J, Otsu K, Zorzato F, de Leon S, Khanna V, Weiler J, et al. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science (80-).* 1991; 253: 448–451. <https://doi.org/10.1126/science.1862346>
51. Roberts MC, Mickelson JR, Patterson EE, Nelson TE, Jane Armstrong P, Brunson DB, et al. Autosomal dominant canine malignant hyperthermia is caused by a mutation in the gene encoding the skeletal muscle calcium release channel (RYR1). *Anesthesiology.* 2001; 95: 716–725. <https://doi.org/10.1097/0000542-200109000-00026> PMID: 11575546
52. Estany J, Ros-Freixedes R, Tor M, Pena RN. A functional variant in the stearoyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork. *PLoS One.* 2014; 9. <https://doi.org/10.1371/journal.pone.0086177> PMID: 24465944



53. Ribani A, Bertolini F, Schiavo G, Scotti E, Utzeri VJ, Dall'Olio S, et al. Next generation semiconductor based sequencing of bitter taste receptor genes in different pig populations and association analysis using a selective DNA pool-seq approach. *Anim Genet*. 2017; 48: 97–102. <https://doi.org/10.1111/age.12472> PMID: 27435880
54. Wu X, Zhang Y, Shen L, Du J, Luo J, Liu C, et al. A 6-bp deletion in exon 8 and two mutations in introns of TYRP1 are associated with blond coat color in Liangshan pigs. *Gene*. 2016; 578: 132–136. <https://doi.org/10.1016/j.gene.2015.12.011> PMID: 26680103
55. Fedick A, Su J, Jalas C, Northrop L, Devkota B, Ekstein J, et al. High-Throughput Carrier Screening Using TaqMan Allelic Discrimination. *PLoS One*. 2013; 8. <https://doi.org/10.1371/journal.pone.0059722> PMID: 23555759
56. Kerr SM, Campbell A, Murphy L, Hayward C, Jackson C, Wain L V., et al. Pedigree and genotyping quality analyses of over 10,000 DNA samples from the Generation Scotland: Scottish Family Health Study. *BMC Med Genet*. 2013; 14. <https://doi.org/10.1186/1471-2350-14-38> PMID: 23521772
57. Rothschild MF, Ruvinsky A. *The genetics of the pig*. CABI; 2011.
58. Pigs Porter V. *A handbook to the breeds of the world*. Helm Information; 1993.
59. Fang M, Larson G, Ribeiro HS, Li N, Andersson L. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genet*. Public Library of Science; 2009; 5: e1000341. <https://doi.org/10.1371/journal.pgen.1000341> PMID: 19148282
60. Fontanesi L, Russo V. Molecular genetics of coat colour in pigs. *Acta Agric Slov*. 2013; 4: 16.
61. Manson IL. *A World Dictionary of Livestock Breeds*. Types Var 4th Edn, CAB Int Wallingford, Oxon, UK. 1996; 273.
62. Margeta P, Margeta V, BudIMIR K. How black is really Black slavonian pig. *Acta Agric Slov*. 2013; 4: 25–28.
63. Marilli M, Fornarelli F, Nulle NM, Crepaldi P. Molecular analysis of MC1R gene in the italian pig breed Mora Romagnola. 56th Annual meeting of the European Association for Animal Production. Wageningen Academic Publishers; 2005. p. 95.
64. Sirtori F, Crovetto A, Zilio DM, Pugliese C, Acciaioli A, Campodoni G, et al. Effect of sire breed and rearing system on growth, carcass composition and meat traits of Cinta Senese crossbred pigs. *Ital J Anim Sci*. Taylor & Francis; 2011; 10: e47.
65. Razmaite V. Performance of Lithuanian indigenous wattle pigs in their two and three-way crosses with Durocs. *Biologija*. 2002; 3: 13–15.
66. Fernández A, Fabuel E, Alves E, Rodriguez C, Silió L, Óvilo C. DNA tests based on coat colour genes for authentication of the raw material of meat products from Iberian pigs. *J Sci Food Agric*. 2004; 84: 1855–1860. <https://doi.org/10.1002/jsfa.1829>
67. Linderholm A, Spencer D, Battista V, Frantz L, Barnett R, Fleischer RC, et al. A novel *MC1R* allele for black coat colour reveals the Polynesian ancestry and hybridization patterns of Hawaiian feral pigs. *R Soc Open Sci*. 2016; 3: 160304. <https://doi.org/10.1098/rsos.160304> PMID: 27703696
68. Fontanesi L, Scotti E, Gallo M, Nanni Costa L, Dall'Olio S. Authentication of "mono-breed" pork products: Identification of a coat colour gene marker in Cinta Senese pigs useful to this purpose. *Livest Sci*. 2016; 184: 71–77. <https://doi.org/10.1016/j.livsci.2015.12.007>
69. Ogorevc J, Zorc M, Škrlep M, Bozzi R, Petig M, Fontanesi L, et al. Is KIT Locus Polymorphism rs328592739 Related to White Belt pPhenotype in Krškopolje Pig? *Agric Conspec Sci. Agronomski fakultet Zagreb*; 2017; 82: 155–161.
70. Egerszegi I, Ratky J. Mangalica-an indigenous swine breed from Hungary (Review). *Arch fur ...* 2003; 46: 245–256. <https://doi.org/10.5194/aab-46-245-2003>
71. Petrović M, Mijatović M, Radojković D, Radović Č, Marinkov G, Stojanović L. Genetic resources in pig breeding: Moravka. *Biotechnol Anim Husband*. 2007; 23: 1–11.
72. Mikawa S, Morozumi T, Shimanuki SI, Hayashi T, Uenishi H, Domukai M, et al. Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1). *Genome Res*. 2007; 17: 586–593. <https://doi.org/10.1101/gr.6085507> PMID: 17416745
73. Burgos C, Latorre P, Altarriba J, Carrodeguas JA, Varona L, López-Buesa P. Allelic frequencies of NR6A1 and VRTN, two genes that affect vertebrae number in diverse pig breeds: A study of the effects of the VRTN insertion on phenotypic traits of a Duroc×Landrace-Large White cross. *Meat Sci*. 2015; 100: 150–155. <https://doi.org/10.1016/j.meatsci.2014.09.143> PMID: 25460119
74. Đikić M, Salajpal K, Karolyi D, Đikić D, Rupić V. Biological characteristics of turopolje pig breed as factors in renewing and preservation of population. *Stočarstvo*. 2010; 2010: 79–90.

75. Kiaulių ūkio kilimas Lietuvoje (Growth of pig production in Lithuania). In Kiaulių Kilmės Knyga I 1927–1936 (Studbook of pigs I). Kaunas: Adomavičius Printing House; 1939.
76. Silva JS, Bernardo A, da Costa JP. Genetic characterization and inventory of the Bísaro pig through visible effect genes. Their utilization in the genotypic comparison between populations and in the establishing of a nucleus for in vivo genetic conservation.
77. Čandek-Potokar M, Žlender B, Kramar Z, Šegula B, Fazarinc G, Uršič M. Evaluation of Slovene local pig breed Krškopolje for carcass and meat quality. *Czech J Anim Sci*. 2003; 48: 120–128.
78. Jungerius BJ, Van Laere AS, Te Pas MFW, Van Oost BA, Andersson L, Groenen MAM. The IGF2-intron3-G3072A substitution explains a major imprinted QTL effect on backfat thickness in a Meishan x European white pig intercross. *Genet Res*. 2004; 84: 95–101. <https://doi.org/10.1017/S0016672304007098> PMID: 15678747
79. Estellé J, Mercadé A, Noguera JL, Pérez-Enciso M, Óvilo C, Sánchez A, et al. Effect of the porcine IGF2-intron3-G3072A substitution in an outbred Large White population and in an Iberian x Landrace cross. *J Anim Sci*. 2005; 83: 2723–2728. 83:2723–2728 <https://doi.org/10.2527/2005.83122723x> PMID: 16282609
80. Oczkowicz M, Tyra M, Walinowicz K, Różycki M, Rejduch B. Known mutation (A3072G) in intron 3 of the IGF2 gene is associated with growth and carcass composition in Polish pig breeds. *J Appl Genet*. Springer; 2009; 50: 257–259.
81. Brossard L, Nieto R, Araujo JP, Pugliese C, Radović Č, Čandek-Potokar M. Modelling study with Inra-Porc to evaluate nutritional requirements of growing pigs in local breeds. 69th Annual meeting of EAAP. 2018.
82. Fan B, Lkhagvadorj S, Cai W, Young J, Smith RM, Dekkers JCM, et al. Identification of genetic markers associated with residual feed intake and meat quality traits in the pig. *Meat Sci*. 2010; 84: 645–650. <https://doi.org/10.1016/j.meatsci.2009.10.025> PMID: 20374837
83. Pérez-Montarelo D, Rodríguez MC, Fernández A, Benítez R, García F, Silló L, et al. Haplotypic diversity of porcine LEP and LEPR genes involved in growth and fatness regulation. *J Appl Genet*. 2015; 56: 525–533. <https://doi.org/10.1007/s13353-015-0284-7> PMID: 25910874
84. Muñoz G, Ovilo C, Silló L, Tomás A, Noguera JL, Rodríguez MC. Single- And joint-population analyses of two experimental pig crosses to confirm quantitative trait loci on *Sus scrofa* chromosome 6 and leptin receptor effects on fatness and growth traits. *J Anim Sci*. 2009; 87: 459–468. <https://doi.org/10.2527/jas.2008-1127> PMID: 18952727
85. Uemoto Y, Kikuchi T, Nakano H, Sato S, Shibata T, Kadowaki H, et al. Effects of porcine leptin receptor gene polymorphisms on backfat thickness, fat area ratios by image analysis, and serum leptin concentrations in a Duroc purebred population. *Anim Sci J*. 2012; 83: 375–385. <https://doi.org/10.1111/j.1740-0929.2011.00963.x> PMID: 22574789
86. Ros-Freixedes R, Gol S, Pena RN, Tor M, Ibáñez-Escriche N, Dekkers JCM, et al. Genome-wide association study singles out SCD and LEPR as the two main loci influencing intramuscular fat content and fatty acid composition in duroc pigs. *PLoS One*. 2016; 11. <https://doi.org/10.1371/journal.pone.0152496> PMID: 27023885
87. Fan B, Onteru SK, Plastow GS, Rothschild MF. Detailed characterization of the porcine MC4R gene in relation to fatness and growth. *Anim Genet*. 2009; 40: 401–409. <https://doi.org/10.1111/j.1365-2052.2009.01853.x> PMID: 19397528
88. Fontanesi L, Buttazzoni L, Galimberti G, Calò DG, Scotti E, Russo V. Association between melanocortin 4 receptor (MC4R) gene haplotypes and carcass and production traits in Italian Large White pigs evaluated with a selective genotyping approach. *Livest Sci*. 2013; 157: 48–56. <https://doi.org/10.1016/j.livsci.2013.07.006>
89. Fontanesi L, Scotti E, Buttazzoni L, Davoli R, Russo V. The porcine fat mass and obesity associated (FTO) gene is associated with fat deposition in Italian Duroc pigs. *Anim Genet*. 2009; 40: 90–93. <https://doi.org/10.1111/j.1365-2052.2008.01777.x> PMID: 18786155
90. Stinckens A, Luyten T, Bijttebier J, Van Den Maagdenberg K, Dieltiens D, Janssens S, et al. Characterization of the complete porcine MSTN gene and expression levels in pig breeds differing in muscularity. *Anim Genet*. 2008; 39: 586–596. <https://doi.org/10.1111/j.1365-2052.2008.01774.x> PMID: 18822098
91. Guimaraes SEF, Stahl CH, Lonergan SM, Geiger B, Rothschild MF. Myostatin promoter analysis and expression pattern in pigs. *Livest Sci*. 2007; 112: 143–150. <https://doi.org/10.1016/j.livsci.2007.02.004>
92. Liu D, Xu Q, Zang L, Liang S, Wu Y, Wei S, et al. Identification and genetic effect of haplotypes in the promoter region of porcine myostatin gene. *Anim Genet*. 2011; 42: 6–14. <https://doi.org/10.1111/j.1365-2052.2010.02081.x> PMID: 20497155

93. Wood JD, Enser M, Fisher A V, Nute GR, Sheard PR, Richardson RI, et al. Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci.* 2008/04/01. 2008; 78: 343–358. <https://doi.org/10.1016/j.meatsci.2007.07.019> [pii] PMID: 22062452
94. Fernández AI, Óvilo C, Barragán C, Carmen Rodríguez M, Silió L, Folch JM, et al. Validating porcine SCD haplotype effects on fatty acid desaturation and fat deposition in different genetic backgrounds. *Livest Sci.* 2017; 205: 98–105. <https://doi.org/10.1016/j.livsci.2017.09.021>
95. Tibau J, Torrentó N, Jaume J, González J, Čandek-Potokar M, Batorek-Lukač N, et al. Majorcan Black Pig: a sustainable production system for high quality meat products. *EAAP– 69th Annual Meeting.* 2018.
96. Uremović M, Uremović Z, Luković Z, Konjačić Z, Neđeral Nakić S. Meat quality of the Black Slavonian breed and the F1 generation with the Large White breed. *42nd Croatian & 2nd International Symposium on Agriculture.* 2007. pp. 577–580.
97. Parunovic N, Petrovic M, Djordjevic V, Petronijevic R, Lakicevic B, Petrovic Z, et al. Cholesterol content and fatty acids composition of Mangalitsa pork meat. *Procedia Food Sci. Elsevier;* 2015; 5: 215–218.
98. Petrović M, Wähner M, Radović Č, Radojković D, Parunović N, Savić R, et al. Fatty acid profile of m. longissimus dorsi of Mangalitsa and Moravka pig breeds. *Arch Anim Breed.* 2014; 57: 17.
99. Franci O, Pugliese C. Italian autochthonous pigs: progress report and research perspectives. *Ital J Anim Sci.* Taylor & Francis; 2007; 6: 663–671.
100. Sans P, Andrade MJ, Ventanas S, Ruiz J. Quality characteristics of fresh meat from pigs of the Gascon breed. *Food Sci Technol Int.* Sage Publications; 2004; 10: 29–34.
101. Teixeira A, Rodrigues S. Pork meat quality of Preto Alentejano and commercial Largewhite Landrace cross. *J Integr Agric. Elsevier;* 2013; 12: 1961–1971.
102. Rowe SJ, Karacaören B, de Koning DJ, Lukic B, Hastings-Clark N, Velander I, et al. Analysis of the genetics of boar taint reveals both single SNPs and regional effects. *BMC Genomics.* 2014; 15. <https://doi.org/10.1186/1471-2164-15-424> PMID: 24894739
103. Bernau M, Schwanitz S, Kremer-Rücker PV, Kreuzer LS, Scholz AM. Size matters: Boar taint in relationship with body composition and testis volume measured by magnetic resonance imaging. *Livest Sci.* 2018; <https://doi.org/10.1016/j.livsci.2018.06.015>
104. Aluwé M, Millet S, Bekaert KM, Tuytens FAM, Vanhaecke L, De Smet S, et al. Influence of breed and slaughter weight on boar taint prevalence in entire male pigs. *Animal.* 2011; <https://doi.org/10.1017/S1751731111000164> PMID: 22440181
105. Nonneman D, Lindholm-Perry AK, Shackelford SD, King DA, Wheeler TL, Rohrer GA, et al. Predictive markers in calpastatin for tenderness in commercial pig populations. *J Anim Sci.* 2011; 89: 2663–2672. <https://doi.org/10.2527/jas.2010-3556> PMID: 21454860
106. Dai LH, Xiong YZ, Deng CY, Jiang SW, Zuo B, Zheng R, et al. Association of the A-G polymorphism in porcine adiponectin gene with fat deposition and carcass traits. *Asian-Australasian J Anim Sci.* 2006; 19: 779–783. <https://doi.org/10.5713/ajas.2006.779>
107. D'Alessandro E, Dall'Olio S, Escotti E. Analysis of PRKAG3 gene polymorphisms in Italian local pig breeds. *Ital J Anim Sci.* 2013;
108. Galve A, Burgos C, Varona L, Carrodegua JA, Cánovas A, López-Buesa P. Allelic frequencies of PRKAG3 in several pig breeds and its technological consequences on a Duroc x Landrace-Large White cross. *J Anim Breed Genet.* 2013; 130: 382–393. <https://doi.org/10.1111/jbg.12042> PMID: 24074175
109. Hess AS, Islam Z, Hess MK, Rowland RRR, Lunney JK, Doeschl-Wilson A, et al. Comparison of host genetic factors influencing pig response to infection with two North American isolates of porcine reproductive and respiratory syndrome virus. *Genet Sel Evol.* 2016; 48: 43. <https://doi.org/10.1186/s12711-016-0222-0> PMID: 27324857
110. Fontanesi L, Bertolini F, Dall'Olio S, Buttazzoni L, Gallo M, Russo V. Analysis of association between the MUC4 g. 8227C> G polymorphism and production traits in Italian heavy pigs using a selective genotyping approach. *Anim Biotechnol.* Taylor & Francis; 2012; 23: 147–155. <https://doi.org/10.1080/10495398.2011.653462> PMID: 22870870
111. Jørgensen CB, Cirera S, Anderson SI, Archibald AL, Raudsepp T, Chowdhary B, et al. Linkage and comparative mapping of the locus controlling susceptibility towards *E. coli* F4ab/ac diarrhoea in pigs. *Cytogenet Genome Res.* Karger Publishers; 2003; 102: 157–162. <https://doi.org/10.1159/000075742> PMID: 14970696
112. He XP, Peng YB, Fan B, Li K, Zhu MJ, Peng ZZ. Study on the detection of swine  $\alpha$ -1 fucosyltransferase polymorphism and its correlation with traits. *Swine Hlth Prod.* 2006; 3: 26–28.

113. Bao W Bin, Ye L, Pan ZY, Zhu J, Zhu GQ, Huang XG, et al. Beneficial genotype of swine FUT1 gene governing resistance to *E. coli* F18 is associated with important economic traits. *J Genet.* 2011; 90: 315–318. <https://doi.org/10.1007/s12041-011-0059-9> PMID: 21869481
114. Short TH, Rothschild MF, Southwood OI, McLaren DG, De Vries A, Van Der Steen H, et al. Effect of the estrogen receptor locus on reproduction and production traits in four commercial pig lines. *J Anim Sci.* 1997; 75: 3138–3142. <https://doi.org/10.2527/1997.75123138x> PMID: 9419986
115. Baba T, Mimura J, Nakamura N, Harada N, Yamamoto M, Morohashi K-i., et al. Intrinsic Function of the Aryl Hydrocarbon (Dioxin) Receptor as a Key Factor in Female Reproduction. *Mol Cell Biol.* 2005; 25: 10040–10051. <https://doi.org/10.1128/MCB.25.22.10040-10051.2005> PMID: 16260617
116. Hernández-Ochoa I, Karman BN, Flaws JA. The role of the aryl hydrocarbon receptor in the female reproductive system. *Biochemical Pharmacology.* 2009. pp. 547–559. <https://doi.org/10.1016/j.bcp.2008.09.037> PMID: 18977336
117. Laval G, Iannuccelli N, Legault C, Milan D, Groenen MA, Giuffra E, et al. Genetic diversity of eleven European pig breeds. *Genet Sel Evol.* 2000; 32: 187–203. <https://doi.org/10.1186/1297-9686-32-2-187> PMID: 14736401
118. Peretti V, Ciotola F, Ianuzzi L. Characterization, conservation and sustainability of endangered animal breeds in Campania (Southern Italy). *Nat Sci.* 2013; 5: 1–9.
119. Rischkowsky B, Pilling D. List of breeds documented in the Global Databank for Animal Genetic Resources. The State of the World's Animal Genetic Resources for Food and Agriculture Rome: Food and Agriculture Organization of the United Nations. 2007. p. Annex.
120. Herrero-Medrano JM, Megens HJ, Groenen MAM, Ramis G, Bosse M, Pérez-Enciso M, et al. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC Genet.* 2013; 14. <https://doi.org/10.1186/1471-2156-14-106> PMID: 24172017
121. Megens HJ, Crooijmans R, Cristobal MS, Hui X, Li N, Groenen MAM. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genet Sel Evol.* 2008; 40: 103–128. <https://doi.org/10.1186/1297-9686-40-1-103> PMID: 18096118
122. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 2010; 11: 94. <https://doi.org/10.1186/1471-2156-11-94> PMID: 20950446