Genome-wide association analyses for coat colour patterns in the autochthonous Nero Siciliano pig breed

Published Version:

Schiavo, G., Bovo, S., Tinarelli, S., Gallo, M., Dall'Olio, S., Fontanesi, L. (2020). Genome-wide association analyses for coat colour patterns in the autochthonous Nero Siciliano pig breed. LIVESTOCK SCIENCE, 236, 1-6 [10.1016/j.livsci.2020.104015].

Published:

DOI: http://doi.org/10.1016/j.livsci.2020.104015

Genome-wide association analyses for coat colour patterns in the autochthonous Nero Siciliano pig breed

Giuseppina Schiavo[a], Samuele Bovo[a], Silvia Tinarelli[a,b], Maurizio Gallo[b], Stefania Dall’Olio[a] and Luca Fontanesi[a,*]

[a] Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

[c] Associazione Nazionale Allevatori Suini, Via Nizza 53, 00198 Roma, Italy

* Corresponding author

E-mail address: luca.fontanesi@unibo.it (L. Fontanesi)

Running title: Mining the genome of Nero Siciliano pigs

Highlights

- Nero Siciliano is an autochthonous pig genetic resource reared in Sicily (Italy).
- Animals of this breed are completely black, few pigs have white patterns.
- Genome-wide association and F_{ST} analyses were carried out in this breed for coat colour patterns.
- A significant marker (rs81329493), confirmed in the F_{ST} analysis, was identified on porcine chromosome 2.
- This marker is located on a quantitative trait locus for coat colour in pigs and close to a few candidate genes.

Abstract

Nero Siciliano (or Sicilian Black) is an Italian autochthonous pig breed reared in the Sicily island, mainly under extensive management systems. Nero Siciliano pigs are black (with black skin and black hair), but animals with white face or partially white face ("suino facciolo") can be registered to the breed herd book. Sometimes, other white patterns on extreme portions of legs could appear in this population. This study took advantage from the rare occurrence of pigs with white patterns in the Nero Siciliano population to carry out a genome-wide association study and comparative genome-wide Fixation index (F_{ST}) analysis to identify genomic regions that could affect coat colour variability (solid black *vs* white patterns over black) in this autochthonous pig breed. Analyses have been conducted on 66 Nero Siciliano pigs: 30 completely black and 36 black with white patterns. All samples have been genotyped for the *KIT* gene duplication and *MC1R* mutations, two genes well known to affect coat colours in pigs. Only pigs that did not carry any duplication of the *KIT* gene and were homozygous for the E^{D2} black dominant *MC1R* gene allele (n. = 26 completely black and n. 22 with white patterns) were genotyped with the Illumina PorcineSNP60 BeadChip. The genome-wide analyses identified on chromosome 2 a significant marker (rs81329493) associated with the coat colour white patterns in this breed. The homologous chromosome region in felids contains the gene responsible for the blotched tabby and striped coat colour patterns. Further studies, including a larger number of pigs, are needed to confirm this result and identify the causative mutation(s) affecting this coat colour diversity, which might be used to design a conservation programme in this breed aiming to maintain phenotypic homogeneity (i.e. solid black) that is typically associated with Nero Siciliano pigs. This study demonstrated how genetic diversity segregating in an autochthonous genetic resource can be explored to understand the genetic mechanisms affecting phenotypic traits in a livestock species.

Key words: Animal genetic resource; autochthonous breed; livestock; single nucleotide polymorphism; *Sus scrofa*.

1. Introduction

Most of the genetic diversity in livestock resides among many different breeds that have been shaped by the combined action of human driven and natural directional selection, including breeding programmes and adaptation to a variety of production systems and environments, that underwent several and complex genetic events (e.g. genetic drift, bottleneck, isolation, introgression, migration, crossbreeding). These genetic resources can be described by associated inheritable phenotypes, that usually describe the uniqueness of local animal breeds (Leroy et al., 2016). Some populations are not completely fixed for these traits as a result of mild or only recent selection pressure towards some features or due to recent introgression or crossbreeding that together increased within breed diversity. Exploring this variability could highlight genetic mechanisms and genes affecting production and morphological traits, as we recently demonstrated in the autochthonous Casertana pig breed using genome-wide association and selection signature analyses (Bertolini et al., 2018; Schiavo et al. 2018, 2019).

Nero Siciliano (or Sicilian Black) is another autochthonous pig breed reared in the Sicily island, mainly under extensive management systems. This breed, also known with other local names (“Nero dei Nebrodi”, “Nero delle Madonie” and “Nero dell’Etna”), is one of the six officially recognized Italian local pig breeds, under the national conservation programme managed by the Italian Pig Breeders Association (ANAS). The Herd Book of this breed currently accounts about 1000 registered sows and boars (ANAS, 2020).

Even if a pig population has been present in the Sicily island since the pre-Roman period, the genetic pool of the Nero Siciliano breed has been mainly shaped during the last centuries. Pig populations with Neapolitan blood were described in Sicily by Chicoli (1870). Then, Casertana pigs have been used for the constitution of some nuclei of districts in the province of Enna and

Messina (Faelli, 1928; Pino, 1947; Chiofalo, 1990). According to Montanaro (1939), Iberian lines were then introduced into this population and Porter (1993) indicated that the Sicilian pig population was subsequently improved by crossbreeding with Large Black and Large White animals. Subsequent introgression events, derived by other improved breeds (probably Pietrain and others) and wild boars, contributed to determine the large genetic variability of this breed, already recognized at the genome level by several previous studies (Russo et al., 2004; Muñoz et al., 2018, 2019; Schiavo et al., 2020).

Nero Siciliano pigs are usually completely black (with black skin and black hair) with a black dorsal stripe. Few animals can have a white face or a face with white portions (referred as "suino facciolo"). Sometimes, other white patterns could appear in the population (e.g. a white belt, white extremities of the legs or other white patterns) but these animals cannot be registered to the Herd Book of the breed.

The black coat colour in this breed may be determined by the dominant alleles at the melanocortin 1 receptor (*MC1R*) gene: E^{D1} or *MC1R*2* and E^{D2} or *MC1R*3* (the most frequent allele in this breed) of Asian and European origin, respectively (Kijas et al., 1998; Fang et al., 2009; Muñoz et al., 2018; Ribani et al., 2019). However, other *MC1R* alleles segregate in this breed (E^{+} or *MC1R*1*, the wild type allele; *e* or *MC1R*4*, the recessive red allele; and E^{P} or *MC1R*6*, the patched black allele. Kijas et al., 2001; Fang et al., 2009) supporting the genetic influence of several pig breeds/populations in shaping the genetic background of the Nero Siciliano breed (Fontanesi et al., 2010; Muñoz et al., 2018; Ribani et al., 2019).

White coat colours and white patterns in pigs are mainly influenced by the *Dominant white* locus which has a complex series of alleles determined by different regions of the *KIT* gene affected by copy number variations and other polymorphisms (Johansson Moller et al., 1996;

Marklund et al., 1998; Pielberg et al., 2002; Johansson et al., 2005; Fontanesi et al., 2010, 2016; Rubin et al., 2012). Epistatic interactions have been observed between the *Extension* and the *Dominat white* loci (Hirooka et al., 2002; Fontanesi et al., 2010).

White spotted patterns in Chinese breeds and in traditional European breeds have been associated with variability in the endothelin receptor type B (*EDNRB*) and KIT ligand (*KITLG*) gene regions (e.g. Wilkinson et al., 2013; Wang et al. 2015, 2018; Lü et al., 2016) and several QTLs for coat colour patterns have been reported in few other chromosome regions (Hirooka et al., 2002).

In this study, we took advantage from the rare occurrence of pigs with white patterns in the Nero Siciliano population to carry out a genome-wide association study and a comparative genome-wide Fixation index (F_{ST}) analysis with the final objective to identify genomic regions that could affect coat colour variability (solid black *vs* white patterns over black) in this autochthonous pig breed.

2. Material and methods

2.1. Animals and coat colour phenotypes

This study did not have any ethical implications as animals were not treated in any way. Operations on the animals were carried out under routine veterinary inspections. All animals were raised according to National and European legislation and no ethical permit was therefore needed.

A total of 66 Nero Siciliano pigs (about 6 to 24 months old) from four different farms were included in this study. Photographic and visual records of each animal were obtained to classify the animals as completely black (n. 30) or black with white patterns (n. 36). This latter group accounted for 20 pigs having both white patterns in the face and in the extreme portions of one or more legs and for 16 pigs with white portions only in the face. A few examples of Nero Siciliano pigs with different coat colours are reported in Figure 1.

2.2. Genotyping of *KIT* and *MC1R* mutations

Hair roots or blood samples collected from the studied pigs were used for DNA extraction that was carried out with the Wizard® Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA) following the manufacturer's instructions. DNA of all pigs was amplified with the diagnostic test for the duplication of the *KIT* gene using the primers described by Giuffra et al. (2002). The test detects a dominant genotype and cannot reveal the number of duplicated copies. The test was carried out in a duplex PCR with *ESR1* primers in order to evaluate if the lack of amplification for the *KIT* primers was due to failure of the PCR or by the presence of a single copy gene, as previously described (Fontanesi et al., 2010). PCR amplification, primers and genotyping conditions were as reported by Fontanesi et al. (2010).

Four mutations of the *MC1R* gene, that identify the five main alleles at this locus (E^{+} or *MC1R*1*, E^{D1} or *MC1R*2*, E^{D2} or *MC1R*3*, *e* or *MC1R*4* and E^{P} or *MC1R*6*; Kijas et al., 1998, 2001; Fang et al., 2009) have been genotyped in all pigs with the protocols described in Fontanesi et al. (2010, 2014).

2.3. Genotyping with the Illumina PorcineSNP60 BeadChip panel

All pigs that did not carry any duplication of the *KIT* gene (expected to have a wild type genotype at the *Dominant white* locus) and that were homozygous for the E^{D2} allele at the *MC1R* gene (26 pigs with completely black coat colour and 22 pigs with white patterns) were genotyped with the Illumina PorcineSNP60 BeadChip v.2 (Illumina, Inc., San Diego, CA, USA). The preselection of the pigs having a putative wild type genotype at the *Dominant white* locus and a dominant black genotype at the *Extension* locus was done to avoid interactions or effects of different alleles at these two loci that are well known to affect coat colour phenotypes (Fontanesi and Russo, 2013). Genotyping of the 61,565 single nucleotide polymorphisms (SNPs) of the commercial panel was carried out following the manufacturer's protocol. Genotyping data were filtered using PLINK 1.9 software (Chang et al., 2015) using criteria already applied in similar studies (Bertolini et al., 2018; Schiavo et al., 2018, 2019): genotyping call rate >0.9, minor allele frequency (MAF) >0.01 and Hardy-Weinberg equilibrium *P*>0.001. After filtering, the number of retained autosomal SNPs was 49266. The positions of the SNPs were assigned to the Sscrofa11.1 reference genome version by applying the mapping method described in Fontanesi et al. (2012).

2.4. Genome-wide analyses

PLINK 1.9 software (Chang et al., 2015) was used to obtain multi-dimensional scaling (MDS) plots and evaluate distance relationships among the animals of the investigated cohorts, included in the different comparative genome-wide analyses for the considered coat colour patterns. Possible clusters evidenced in MDS plots might only due to within breed stratification problems, as already reported in similar studies (Bertolini et al., 2018; Schiavo et al., 2018, 2019).

Genome-wide association studies were carried out using a single marker analysis by applying the univariate mixed model of GEMMA v. 0.98.1 (Zhou and Stephens, 2012). The centered relatedness matrix calculated from SNP genotypes was used to correct for population stratification in a case and control (completely black *vs* black with white patterns) analysis with a model that included the farm and the sex as fixed effects. The significant threshold was defined using a Bonferroni corrected *p*-value of 0.05. Suggestive significant threshold was defined at the

$P_{nominal\ value}$<5.0×10^{-05} level, as already applied in several other genome wide association studies in livestock (e.g. Fontanesi et al., 2012; Sanchez et al., 2014; Bovo et al., 2019), including similar case and control studies in pigs (Bertolini et al., 2018; Schiavo et al., 2018, 2019). Genomic inflation factor (λ) and quantile–quantile (QQ) plot in the genome wide association study were obtained using GenABEL (function estlamda and method median; Aulchenko et al., 2007).

To further evaluate the results of the genome-wide association studies, F_{ST} analyses were performed using the same SNP genotyping data and by comparing the two groups of pigs (completely black *vs* black with white patterns) in a single marker analysis using PLINK 1.9 software (Chang et al., 2015). F_{ST} relevant differences were considered for the SNPs having values over the 99.9th percentile distribution.

Annotation of the genomic regions that had SNPs that trespassed the defined thresholds in both genome-wide association and genome-wide F_{ST} analyses was obtained from the information available on the Sscrofa11.1 reference genome, retrieved at the Ensembl database (http://www.ensembl.org/Sus_scrofa/Info/Index), release 98 (September 2019).

3. Results and discussion

A total of 26 out of 30 completely black pigs were homozygous E^{D2}/E^{D2} at the *MC1R* gene. Three pigs had genotype E^{D2}/E^{p} and one had genotype E^{D2}/e. These results confirmed that Nero Siciliano pigs are not fixed for one allele at this gene, as already reported in previous studies (Russo et al., 2004; Fontanesi et al., 2010; Muñoz et al., 2018; Ribani et al., 2019). None of the completely black pigs were positive at the test that detects the duplication of the *KIT* gene, thus a total of 26 completely black pigs were considered for further studies.

About 39% of the pigs with white patterns (14 out of 36) were positive at this *KIT* duplication DNA test, therefore they were carriers of alleles constituted by multiple copies of the *KIT* gene (Giuffra et al., 2002; Fontanesi et al., 2010). The high frequency in this breed of pigs carrying a duplicate *KIT* allele confirms the results we already reported in a previous investigation (Fontanesi et al., 2010). The observed high frequency, however, can be considered a biased estimation as, in this study, we purposely sampled pigs with white patterns. It would be useful to further evaluate this matter in the Nero Siciliano pigs and obtain an unbiased estimated frequency of the animal carrying duplicate *KIT* alleles in this breed, probably introgressed into this population by crossbreeding with other white and more productive breeds. As duplicated *KIT* alleles were only observed in pigs with white patterns, it could be possible that copy number variation at this locus might contribute to produce these coat colour phenotypes. Not all pigs with white patterns, however, were positive at the duplication test, indicating that variability in other genomic regions or other *KIT* gene variants may be involved in determining the presence of white patterns in pigs, as already suggested by other studies (Hirooka et al., 2002). In addition, all remaining 22 pigs (negative at the duplication *KIT* gene DNA test) were homozygous E^{D2}/E^{D2} at the *MC1R* gene. Therefore, we could also exclude any other effects on the white pattern phenotypes derived by the *Extension* locus.

A total of 48 pigs were then genotyped with the Illumina PorcineSNP60 BeadChip panel. Potential stratification in the analysed pig population, probably derived by the limited number of farms from which animals were sampled, could be evidenced from the MDS plot obtained using this SNP dataset (Figure 2).

Results of the genome-wide association study comparing the two groups of pigs with different coat colour phenotypes (26 completely black *vs* 22 black with white patterns) are reported in Table S1. Figure 3A includes the Manhattan plot obtained in this analysis. Despite the potential stratification of the analysed population, the centered relatedness matrix calculated from SNP genotypes was able to correct this sampling bias, as demonstrated by the genomic inflation factor value (equal to 1.047) and the obtained QQ plot (Figure S1).

The two groups of pigs were also compared using a single marker F_{ST} analysis. Figure 3B shows the Manhattan plot obtained using the F_{ST} values for the analysed SNPs. Table 1 reports all SNPs that trespassed the 99.9th percentile distribution with their corresponding *p*-value in the genome-wide association study. Only the top SNP in both analyses was included in this list. The most significant marker in the genome-wide association study had also the highest F_{ST} value. This SNP (ALGA0102645 or rs81329493) is an intergenic variant located on porcine chromosome 2 (SSC2) at position 120728602. This SSC2 position overlaps with the QTL region for black coat colour reported by Hirooka et al. (2002) in a F2 reference family constructed by crossing Meishan pigs with Dutch commercial animals.

The closest annotated gene on Sscrofa11.1 is semaphoring 6A (*SEMA6A*), which is located from nucleotide position 120560095 to 120687106. SEM6A is a member of the semaphoring family of secreted and transmembrane proteins that, coupled with plexin receptors, regulate critical cellular developmental processes involving actin cytoskeleton, cell proliferation and differentiation, which also affect pigmentation (Prislei et al., 2008; Weiner et al., 2014). SEMA6A plays a role in growth and survival of human melanoma cells (Loria et al., 2015). Disruption or inhibition of other semaphorins also cause skin pigmentation defects (Yu et al., 2004; Scott et al., 2008; Maier et al., 2011).

Interestingly, the SSC2 region spanning the most significant SNPs is overlapping with a corresponding chromosome region in cats and other felids harboring a large haplotype containing the gene responsible for the blotched tabby coat colour pattern (Kaelin et al., 2012). Mutations in the transmembrane aminopeptidase Q (*Taqpep*) gene, that is homologous to the laeverin (*LVRN*) gene, also known as aminopeptidase Q (APQ; annotated from nucleotide positions 120117652 to 120193715 in the pig chromosome), have been shown to determine the change from a spotted phenotype to blotches and stripe patterns (Kaelin et al., 2012). Even if this across species comparative analysis could be speculative, it could be an interesting hint for further investigations.

4. Conclusion

This study was designed to take advantage from a phenotypic diversity on a breed-specific trait (i.e. coat colour) segregating in a local pig population. The design excluded the potential effect on coat colour patterns of main alleles at two major coat colour loci (*KIT* and *MC1R*), as our aim was to identify other loci affecting this phenotype. However, considering the low number of pigs showing white marks that we were able to sample (animals with this phenotype are rare in this population), this study could not include a large number of animals and relied only on the possibility to identify QTL with large effects, not explained by the *Dominant white* and *Extension* loci. Other studies are therefore needed to validate the obtained results.

The combined results obtained by the genome-wide association study and the genome-wide F_{ST} analysis indicated that a region on SSC2 affects, at least in part, the observed coat colour patterns in the Nero Siciliano pig population. Then, adding a candidate gene analysis based on the functions of the genes located in the identified genomic region, we further evaluated the potential role of this chromosome position in the association with the presence of small white patterns on a black background. Further studies, including a larger number of pigs, are needed to confirm this result and identify the causative mutation(s) affecting this coat colour diversity, which might be used to design conservation programme in this breed aiming to maintain phenotypic homogeneity (i.e. solid black) that is typically associated with Nero Siciliano pigs. Our study demonstrated how genetic diversity segregating in an autochthonous genetic resource can be explored to understand the genetic mechanisms affecting phenotypic traits in a livestock species.

Declaration of competing interests

The authors declare no conflict of interest.

Acknowledgements

The Authors thank all farmers and personnel that contributed to collect phenotypes and samples. This study was supported by the PSRN (Progetto di Sviluppo Rurale Nazionale) SUIS project (co-funded by the European Agricultural Fund for Rural Development of the European Union and by the Italian Ministry of Agriculture, Food, Forestry and Tourism - MiPAAFT) and received funding from the University of Bologna RFO 2018 and 2019 programmes.

References

ANAS, 2020. *Registro Anagrafico*. http://www.anas.it/.

Aulchenko, Y.S., Ripke, S., Isaacs, A., van Duijn C.M., 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23, 1294–1296. https://doi.org/10.1093/bioinformatics/btm108.

Bertolini, F., Schiavo, G., Tinarelli, S., Santoro, L., Utzeri, V.J., Dall'Olio, S., Nanni Costa, L., Gallo, M., Fontanesi, L., 2018. Exploiting phenotype diversity in a local animal genetic resource: Identification of a single nucleotide polymorphism associated with the tail shape phenotype in the autochthonous Casertana pig breed. Livest. Sci. 216, 148–152. https://doi.org/10.1016/j.livsci.2018.08.007.

Bovo, S., Mazzoni, G., Bertolini, F., Schiavo, G., Galimberti, G., Gallo, M., Dall'Olio, S., Fontanesi, L., 2019. Genome-wide association studies for 30 haematological and blood clinical-biochemical traits in Large White pigs reveal genomic regions affecting intermediate phenotypes. Sci. Rep. 9, 7003. https://doi.org/10.1038/s41598-019-43297-1.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7. https://doi.org/10.1186/s13742-015-0047-8.

Chicoli, N., 1870. Riproduzione, allevamento e miglioramento degli animali domestici in Sicilia. Stamperia di G. 4. Lorsnaider, Palermo.

Chiofalo, L., 1990. Il suino nero dei Nebrodi: condizioni di allevamento, realtà e prospettive. Convegno "Suino Nero", Longi, Messina, 9 dicembre 1990.

Faelli, F., 1928. Razze bovine, equine, suine, ovine, caprine. Hoepli, Milano, Italy.

Fang, M., Larson, G., Ribeiro, H.S., Li, N., Andersson, L., 2009. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. PLoS Genet. 5, e1000341. https://doi.org/10.1371/journal.pgen.1000341.

Fontanesi, L., D'Alessandro, E., Scotti, E., Liotta, L., Crovetti, A., Chiofalo, V., Russo, V., 2010. Genetic heterogeneity and selection signature at the *KIT* gene in pigs showing different coat colours and patterns. Anim. Genet. 41, 478–492. https://doi.org/10.1111/j.1365-2052.2010.02054.x.

Fontanesi, L., Ribani, A., Scotti, E., Utzeri, V.J., Veličković, N., Dall'Olio, S., 2014. Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the MC1R and NR6A1 genes. Meat Sci. 98, 781–784. https://doi.org/10.1016/j.meatsci.2014.07.026.

Fontanesi, L., Russo, V., 2013. Molecular genetics of coat colour in pigs. Acta Agric. Slov. 4, 16. http://aas.bf.uni-lj.si/zootehnika/supl/4-2013/PDF/4-2013-15-20.pdf.

Fontanesi, L., Schiavo, G., Galimberti, G., Calò, D.G., Scotti, E., Martelli, P.L., Buttazzoni, L., Casadio, R., Russo, V., 2012. A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. BMC Genomics 13, 583. https://doi.org/10.1186/1471-2164-13-583.

Fontanesi, L., Scotti, E., Gallo, M., Costa, L.N., Dall'Olio, S., 2016. Authentication of "mono-breed" pork products: Identification of a coat colour gene marker in Cinta Senese pigs useful to this purpose. Livest. Sci. 184, 71–77. https://doi.org/10.1016/j.livsci.2015.12.007.

Giuffra, E., Törnsten, A., Marklund, S., Bongcam-Rudloff, E., Chardon, P., Kijas, J.M., Anderson, S.I., Archibald, A.L., Andersson, L., 2002. A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking *KIT*. Mamm. Genome 13, 569-577. https://doi.org/10.1007/s00335-002-2184-5.

Hirooka H., de Koning D.J., van Arendonk J.A.M., Harlizius B., de Groot P.N., Bovenhuis H. 2002. Genome scan reveals new coat color loci in exotic pig cross. J. Hered. 93, 1–8. https://doi.org/10.1093/jhered/93.1.1.

Johansson Moller, M., Chaudhary, R., Hellmén, E., Höyheim, B., Chowdhary, B., Andersson, L., 1996. Pigs with the dominant white coat color phenotype carry a duplication of the KIT gene encoding the mast/stem cell growth factor receptor. Mamm. Genome 7, 822–830. https://doi.org/10.1007/s003359900244.

Johansson, A., Pielberg, G., Andersson, L., Edfors-Lilja, I., 2005. Polymorphism at the porcine *Dominant white/KIT* locus influence coat colour and peripheral blood cell measures. Anim. Genet. 36, 288–296. https://doi.org/10.1111/j.1365-2052.2005.01320.x.

Kaelin, C.B., Xu, X., Hong, L.Z., David, V.A., McGowan, K.A., Schmidt-Küntzel, A., Roelke, M.E., Pino, J., Pontius, J., Cooper, G.M., Manuel, H., Swanson, W.F., Marker, L., Harper, C.K., van Dyk, A., Yue, B., Mullikin, J.C., Warren, W.C., Eizirik, E., Kos, L., O'Brien, S.J., Barsh, G.S., Menotti-Raymond, M. 2012. Specifying and sustaining pigmentation patterns in domestic and wild cats. Science 337, 1536–1541. https://doi.org/10.1126/science.1220893.

Kijas, J. M. H., Moller, M., Plastow, G., Andersson, L., 2001. A frameshift mutation in *MC1R* and high frequency of somatic reversions cause black spotting in pigs. Genetics 158, 779–785.

Kijas, J.M., Wales, R., Törnsten, A., Chardon, P., Moller, M., Andersson, L., 1998. Melanocortin receptor 1 (*MC1R*) mutations and coat color in pigs. Genetics 150, 1177-1185.

Leroy, G., Besbes, B., Boettcher, P., Hoffmann, I., Capitan, A., Baumung, R., 2016. Rare phenotypes in domestic animals: unique resources for multiple applications. Anim. Genet. 47, 141–153. https://doi.org/10.1111/age.12393.

Loria, R., Bon, G., Perotti, V., Gallo, E., Bersani, I., Baldassari, P., Porru, M., Leonetti, C., Di Carlo, S., Visca, P., Brizzi, M.F., Anichini, A., Mortarini, R., Falcioni, R., 2015. Sema6A and Mical1 control cell growth and survival of BRAFV600E human melanoma cells. Oncotarget 6, 2779–2793. https://doi.org/10.18632/oncotarget.2995.

Lü, M.D., Han, X.M., Ma, Y.F., Irwin, D.M., Gao, Y., Deng, J.K., Adeola, A.C., Xie, H.B., Zhang, Y.P., 2016. Genetic variations associated with six-white-point coat pigmentation in Diannan small-ear pigs. Sci. Rep. 6, 27534. https://doi.org/10.1038/srep27534.

Maier, V., Jolicoeur, C., Rayburn, H., Takegahara, N., Kumanogoh, A., Kikutani, H., Tessier-Lavigne, M., Wurst, W., Friedel, R.H., 2011. Semaphorin 4C and 4G are ligands of Plexin-B2 required in cerebellar development. Mol. Cell. Neurosci. 46, 419–431. https://doi.org/10.1016/j.mcn.2010.11.005.

Marklund, S., Kijas, J., Rodriguez-Martinez, H., Rönnstrand, L., Funa, K., Moller, M., Lange, D., Edfors-Lilja, I., Andersson, L., 1998. Molecular basis for the dominant white phenotype in the domestic pig. Genome Res. 8, 826–833. https://doi.org/10.1101/gr.8.8.826.

Montanaro, G., 1939. Per il miglioramento della suinicoltura siciliana. Critica zootecnica no. 8.

Muñoz, M., Bozzi, R., García, F., Núñez, Y., Geraci, C., Crovetti, A., García-Casco, J., Alves, E., Škrlep, M., Charneca, R., Martins, J.M., Quintanilla, R., Tibau, J., Kušec, G., Djurkin-Kušec, I., Mercat, M.J., Riquet, J., Estellé, J., Zimmer, C., Razmaite, V., Araujo, J.P., Radović, Č., Savić, R., Karolyi, D., Gallo, M., Čandek-Potokar, M., Fontanesi, L., Fernández, A.I., Óvilo, C., 2018. Diversity across major and candidate genes in European local pig breeds. PLoS One 13, e0207475. https://doi.org/10.1371/journal.pone.0207475.

Muñoz, M., Bozzi, R., García-Casco, J., Núñez, Y., Ribani, A., Franci, O., García, F., Škrlep, M., Schiavo, G., Bovo, S., Utzeri, V.J., Charneca, R., Martins, J.M., Quintanilla, R., Tibau, J., Margeta, V., Djurkin-Kušec, I., Mercat, M.J., Riquet, J., Estellé, J., Zimmer, C., Razmaite, V., Araujo, J.P., Radović, Č., Savić, R., Karolyi, D., Gallo, M., Čandek-Potokar, M., Fernández, A.I., Fontanesi, L., Óvilo, C., 2019. Genomic diversity, linkage disequilibrium and selection signatures in European local pig breeds assessed with a high density SNP chip. Sci. Rep. 9, 13546. https://doi.org/10.1038/s41598-019-49830-6.

Pielberg, G., Olsson, C., Syvänen, A.C., Andersson, L., 2002. Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. Genetics 160, 305–311.

Pino, N., 1947. Il patrimonio suino della Sicilia e la sua etnologia alla luce di ricerche biometriche su alcuni caratteri razziali. Zootecnia e Veterinaria, La fecondazione artificiale, II, 1, 1-15.

Porter, V. 1993. *Pigs: A Handbooks to the Breeds of the World*. Cornell University Press.

Prislei, S., Mozzetti, S., Filippetti, F., De Donato, M., Raspaglio, G., Cicchillitti, L., Scambia, G., Ferlini, C. 2008. From plasma membrane to cytoskeleton: a novel function for semaphorin 6A. Mol. Cancer Ther. 7, 233–241. https://doi.org/10.1158/1535-7163.MCT-07-0390.

Ribani, A., Utzeri, V.J., Geraci, C., Tinarelli, S., Djan, M., Veličković, N., Doneva, R., Dall'Olio, S., Nanni Costa, L., Schiavo, G., Bovo, S., Usai, G., Gallo, M., Radović, Č., Savić, R., Karolyi, D., Salajpal, K., Gvozdanović, K., Djurkin-Kušec, I., Škrlep, M., Čandek-Potokar, M., Ovilo, C., Fontanesi, L., 2019. Signatures of de-domestication in autochthonous pig breeds and of domestication in wild boar populations from *MC1R* and *NR6A1* allele distribution. Anim. Genet. 50, 166–171. https://doi.org/10.1111/age.12771.

Rubin, C.J., Megens, H.J., Martinez Barrio, A., Maqbool, K., Sayyab, S., Schwochow, D., Wang, C., Carlborg, Ö., Jern, P., Jørgensen, C.B., Archibald, A.L., Fredholm, M., Groenen, M.A., Andersson, L., 2012. Strong signatures of selection in the domestic pig genome. Proc. Natl. Acad. Sci. USA 109, 19529–19536. https://doi.org/10.1073/pnas.1217149109.

Russo, V., Fontanesi, L., Davoli, R., Chiofalo, L., Liotta, L., Zumbo, A., 2004. Analysis of single nucleotide polymorphisms in major and candidate genes for production traits in Nero Siciliano pig breed. Ital. J. Anim. Sci. 3, 19–29. https://doi.org/10.4081/ijas.2004.19.

Sanchez M.P., Tribout T., Iannuccelli N., Bouffaud, M., Servin, B., Tenghe, A., Dehais, P., Muller, N., Del Schneider, M.P., Mercat, M.J., Rogel-Gaillard, C., Milan, D., Bidanel, J.P., Gilbert, H., 2014. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. Genet. Sel. Evol. 46, 12. https://doi.org/10.1186/1297-9686-46-12.

Schiavo, G., Bertolini, F., Galimberti, G., Bovo, S., Dall’Olio, S., Nanni Costa, L., Gallo, M., Fontanesi, L., 2020. A machine learning approach for the identification of population-informative markers from high-throughput genotyping data: application to several pig breeds. Animal 14, 223–232. https://doi.org/10.1017/S1751731119002167.

Schiavo, G., Bertolini, F., Utzeri, V.J., Ribani, R., Geraci, C., Santoro, L., Óvilo, C., Fernández, A.I., Gallo, M., Fontanesi, L., 2018. Taking advantage from phenotype variability in a local animal genetic resource: identification of genomic regions associated with the hairless phenotype in Casertana pigs. Anim. Genet. 49, 321–325. https://doi.org/10.1111/age.12665.

Schiavo, G., Bovo, S., Tinarelli, S., Bertolini, F., Dall'Olio, S., Gallo, M., & Fontanesi, L., 2019. Genome-wide association analyses for several exterior traits in the autochthonous Casertana pig breed. Livest. Sci. 230, 103842. https://doi.org/10.1016/j.livsci.2019.103842.

Scott, G.A., McClelland, L.A., Fricke, A.F., 2008. Semaphorin 7a promotes spreading and dendricity in human melanocytes through beta1-integrins. J. Invest. Dermatol. 128, 151–161. https://doi.org/10.1038/sj.jid.5700974.

Wang, C., Wang, H., Zhang, Y., Tang, Z., Li, K., Liu, B., 2015. Genome-wide analysis reveals artificial selection on coat colour and reproductive traits in Chinese domestic pigs. Mol. Ecol. Resour. 15, 414–424. https://doi.org/10.1111/1755-0998.12311.

Wang, C., Wang, X., Tang, J., Chen, H., Zhang, J., Li, Y., Lei, S., Ji, H., Yang, B., Ren, J., Ding, N., 2018. Genome-wide association studies for two exterior traits in Chinese Dongxiang spotted pigs. Anim. Sci. J. 89, 868–875. https://doi.org/10.1111/asj.13003.

Weiner, L., Fu, W., Chirico, W.J., Brissette, J.L., 2014. Skin as a living coloring book: how epithelial cells create patterns of pigmentation. Pigment Cell Melanoma Res. 27, 1014–1031. https://doi.org/10.1111/pcmr.12301.

Wilkinson, S., Lu, ZH., Megens, HJ., Archibald, AL., Haley, C., Jackson, IJ., Groenen, M.A., Crooijmans, R.P., Ogden, R., Wiener, P., 2013. Signatures of diversifying selection in European pig breeds. PLoS Genet. 9, e1003453. https://doi.org/10.1371/journal.pgen.1003453.

Yu, Q., Shen, Y., Chatterjee, B., Siegfried, B.H., Leatherbury, L., Rosenthal, J., Lucas, J.F., Wessels, A., Spurney, C.F., Wu, Y.J., Kirby, M.L., Svenson, K., Lo, C.W., 2004. ENU induced mutations causing congenital cardiovascular anomalies. Development 131, 6211–6223. https://doi.org/10.1242/dev.01543.

Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44, 821–824. https://doi.org/10.1038/ng.2310.

Table 1. List of all single nucleotide polymorphisms (SNPs) that trespassed the 99.9th percentile distribution in the genome-wide F_{ST} analysis, with corresponding *p*-value in the genome-wide association study (GWAS).

SSC[1]	**SNP position**[2]	**SNP**[3]	**F_{ST} value**[4]	**p-value in GWAS**[5]	**Minor allele**[6]	**Major allele**[7]	**MAF**[8]
1	233174803	INRA0006475	0.427	2.07E-04	A	C	0.312
2	120728602	ALGA0102645	0.511	6.01E-08	C	A	0.333
9	43369155	ASGA0042830	0.418	3.86E-04	A	G	0.365
10	54113880	ASGA0048383	0.414	1.18E-03	G	A	0.490
15	131582702	ALGA0088040	0.416	1.43E-04	C	A	0.490

[1] Porcine chromosome.

[2] Position of the SNP on the chromosome.

[3] SNP identifier.

[4] F_{ST} value for the SNP.

[5] *p-value* in the association analysis.

[6] Minor allele for the SNP.

[7] Major allele for the SNP.

[8] Minor Allele Frequency.

Figure 1. Nero Siciliano pigs with completely black coat colour (A) and with white patterns (B).

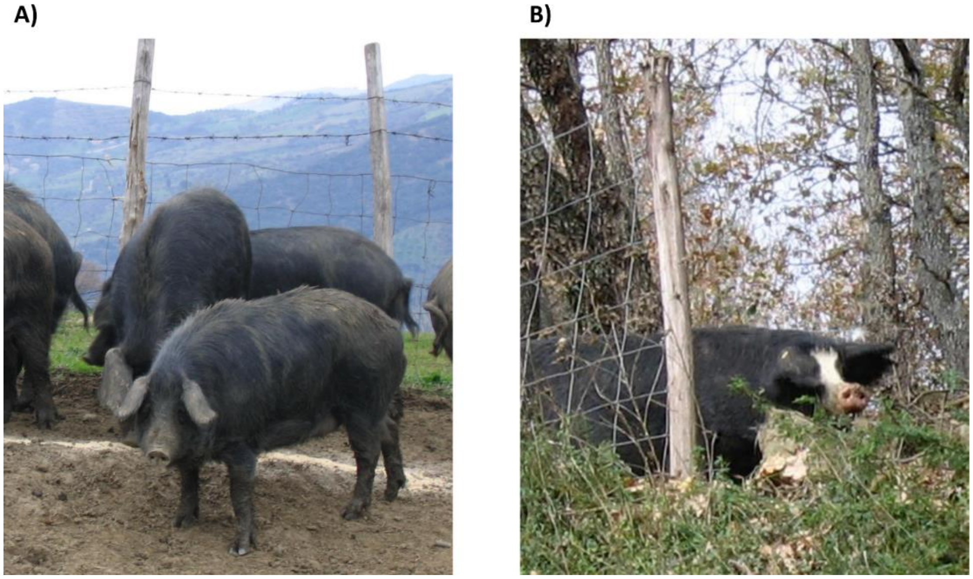

Figure 2. Multidimensional scaling (MDS) plot obtained for the pigs classified according to the alternative coat colour phenotypes (black dots: pigs with completely black coat colour; red dots: pigs having white patterns).

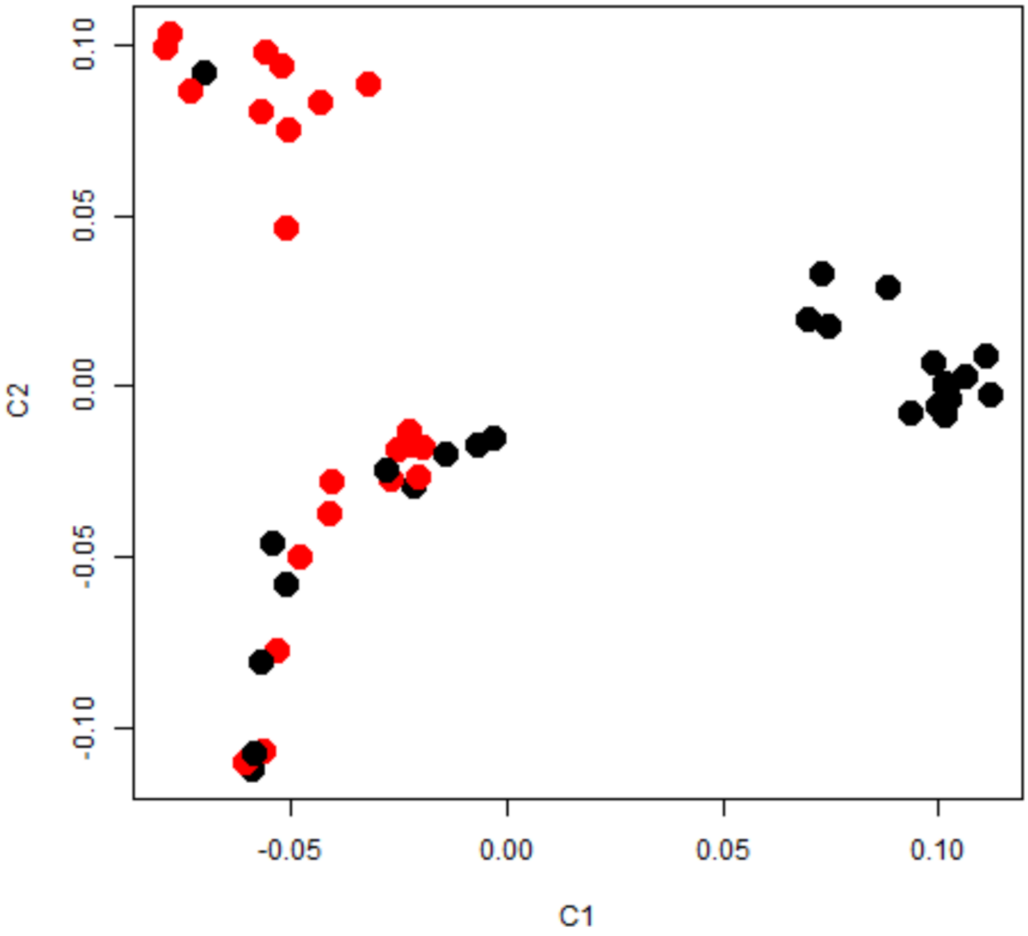

Figure 3. Manhattan plots of the genome-wide association study (A) and of the genome-wide F_{ST} analysis (B) obtained by comparing the two groups of pigs with different coat colours. Thresholds in the genome-wide association study are for *p*-value = 0.05 Bonferroni corrected (red line), *p*-value = 5.0E-05 (blue line). The threshold in the F_{ST} analysis identifies the SNPs having values over the 99.9th percentile distribution.

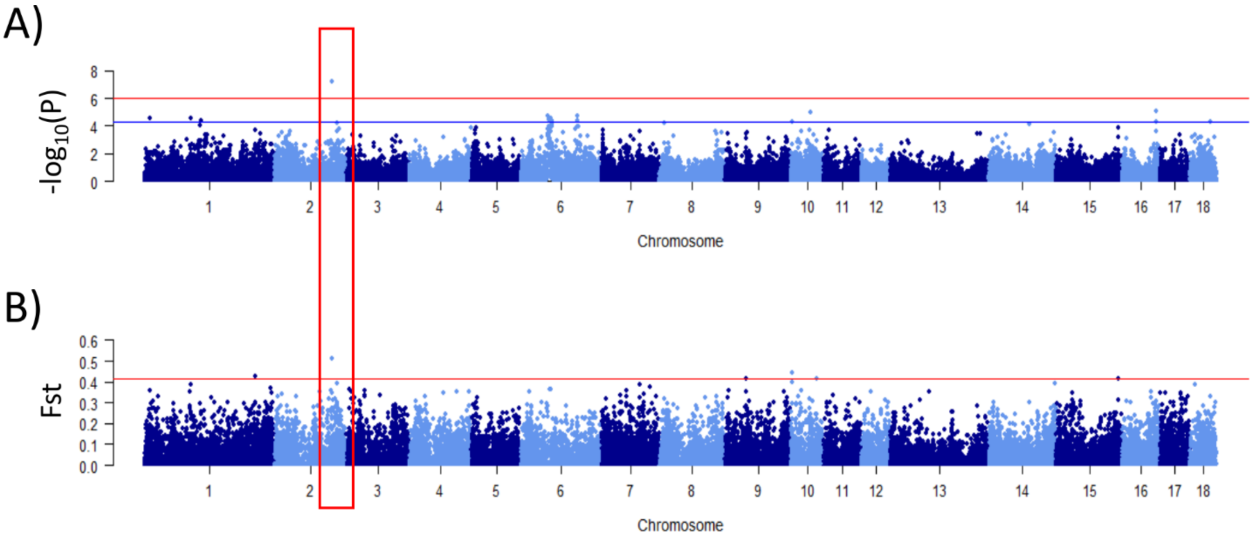

Figure S1. Quantile–quantile (QQ) plot for the genome-wide association study.

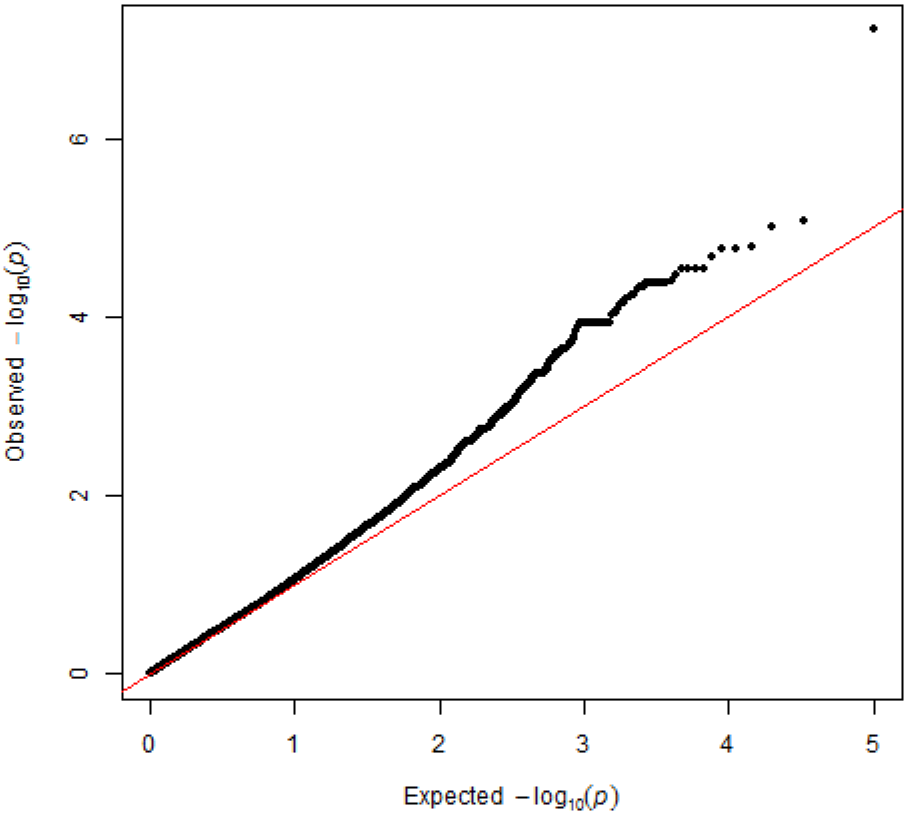

Table S1. List of all significant and suggestively significant single nucleotide polymorphisms (SNPs) in the genome-wide association study.

SSC[1]	**SNP**[2]	**SNP position**[3]	**Minor Allele**[4]	**Major Allele**[5]	**MAF**[6]	***p*-value**[7]
2	ALGA0102645	120728602	C	A	0.333	6.01E-08
16	MARC0113179	71074387	C	A	0.188	8.19E-06
10	H3GA0030074	41868864	G	A	0.052	9.94E-06
6	ALGA0035355	56360299	A	G	0.188	1.73E-05
6	INRA0021527	56427227	A	G	0.188	1.73E-05
6	ALGA0036490	119127413	G	A	0.309	2.08E-05
6	M1GA0009140	62190493	A	G	0.490	2.85E-05
6	MARC0020138	62350290	A	G	0.490	2.85E-05
1	MARC0029100	96759885	A	G	0.354	2.88E-05
1	ALGA0000898	10330993	A	G	0.208	2.89E-05
6	ALGA0112704	57922951	A	G	0.125	3.45E-05
6	ASGA0094008	119146958	A	G	0.312	3.96E-05
6	ASGA0083587	65466319	A	G	0.219	4.14E-05
6	H3GA0054032	65189989	G	A	0.219	4.14E-05
6	H3GA0055528	65284657	A	G	0.219	4.14E-05
6	H3GA0056560	63336711	A	G	0.219	4.14E-05
6	MARC0054213	65230892	C	A	0.219	4.14E-05
6	MARC0094194	65267505	A	G	0.219	4.14E-05
1	ALGA0005754	119505578	G	A	0.490	4.53E-05
10	ALGA0056456	3293112	A	G	0.479	4.65E-05
16	H3GA0047094	71120522	G	A	0.198	4.90E-05

[1] Porcine chromosome.

[2] SNP identifier.

[3] Position of the SNP on the porcine chromosome (Sscrofa11.1 reference genome).

[4] Minor allele for the SNP.

[5] Major allele for the SNP.

[6] Minor Allele Frequency.

[7] *p*-value in the association analysis.