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

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

Francesca Bertolini[1,2,*], Giuseppina Schiavo[1,*], Silvia Tinarelli[1,3], Laura Santoro[4], Valerio Joe Utzeri[1], Stefania Dall'Olio[1], Leonardo Nanni Costa[1], Maurizio Gallo[3] and Luca Fontanesi[1]

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

[2]Department of Bio and Health Informatics, Technical University of Denmark, Kemitorvet, 2800 Kgs. Lyngby, Denmark

[3]Associazione Nazionale Allevatori Suini, Via L. Spallanzani 4, 00161 Roma, Italy

[4]ConSDABI - National Focal Point Italiano FAO, Contrada Piano Cappelle, 82100 Benevento, Italy

Corresponding author: Luca Fontanesi, e-mail: luca.fontanesi@unibo.it

* These authors contributed equally to this work.

Running title: SNPs and tail shape phenotype in pigs

Highlights

- Casertana is an autochthonous pig genetic resource reared in Central-South of Italy.
- Tail shape phenotype variability in the breed was investigated in a GWAS.
- A single nucleotide polymorphism on porcine chromosome 12 was associated with this trait.
- This marker is close to the *SRY-box 9* (*SOX9*) gene that is essential in skeletogenesis.

Abstract

Casertana is a local pig breed mainly raised in Central-South regions of Italy. Pigs of this breed are considered the descendants of the ancient Neapolitan population that largely influenced the constitution of the modern commercial pigs. The pigs of this breed are usually curly-tailed, like several other domestic pig populations. However, Casertana population shows some variability for this trait, including animals having straight tail as observed in wild boars. In this study, we run, for the first time, a genome wide association study (GWAS) comparing the curly tailed (no. = 53) and straight tailed (no. = 19) Casertana pigs to identify genomic regions associated with the tail shape phenotype in *Sus scrofa*. All animals were genotyped with the Illumina PorcineSNP60 BeadChip v.2. GEMMA software was used in the GWAS for which we were able to correct for stratification in the analysed cohort. A single nucleotide polymorphism (rs81439488), located on porcine chromosome 12, was significantly associated with the investigated trait. This marker is close to the *SRY-box 9* (*SOX9*) gene that encodes for a transcription factor that is required during sequential steps of the chondrocyte differentiation pathway, notochord maintenance and skeletogenesis. As the shape of the tail could be important in relation to the problem of tail biting in pigs, the obtained results might open new perspectives for defining selection programs answering indirectly animal welfare issues. This work demonstrated that autochthonous animal genetic resources might be used to disclose genetic factors affecting peculiar traits by exploiting segregating phenotypes and genetic variability.

Keywords: Animal genetic resource; autochthonous breed; GWAS; morphological trait; SNP; *Sus scrofa*.

1. Introduction

Conservation of animal genetic resources is mainly aimed to preserve genetic diversity and associated inheritable phenotypes characterizing different populations that might be interesting for current or future purposes, including potential use in breeding programs. These resources can be also useful to understand biological mechanisms determining unique phenotypes derived by diversity in selection pressures or as result of adaptation to environmental and production conditions (Leroy et al., 2016).

Casertana pigs constitute a local breed mainly raised in Central-South regions of Italy. Pigs of this breed are considered the descendants of the ancient Neapolitan pig population that largely influenced the constitution of the modern commercial pig breeds through introgression of blood into British pig populations during the 19th century (Porter, 1993). Neapolitan pigs were, in turn, influenced by Asian blood in the late 18th century (Porter, 1993). Casertana is enlisted among the endangered animal genetic resources as the herd book of this breed accounts for about 100 boars and sows currently registered (ANAS, 2016). Animals are mainly raised in extensive or semi-extensive production systems with possible contacts and crossbreeding with European wild boars that could have contributed, at least in part, to shape their morphological characteristics. Casertana pigs have a black or grey coat colour, wrinkled skin, forward ears, and usually a typical hairless phenotype. The pigs of this breed are usually curly-tailed, like several other domestic pig populations. However, Casertana population shows some variability for this trait, including animals having straight and wavy tail as in a few other pig breeds and in wild boars.

Domestication in mammals has been a complex and continuous process associated with a series of changes in the domesticated animals compared to the wild counterparts, derived by selective breeding of animals showing favourable production and reproduction performances, and increased docility that indirectly shaped the genome of domesticated populations (Wiener and Wilkinson, 2011; Larson and Burger, 2013; Carneiro et al., 2014; Wang et al., 2014; Wilkins et al., 2014). Several morphological features have been also directly or indirectly selected and, in most cases, fixed in

domesticated populations as result of the domestication process (Darwin, 1868). Coat colour is one of the most common phenotypic traits that has been modified as result of reduced selective pressure against colours with low fitness in the wild and of aesthetic preferences of the breeders, sometimes associated with higher production performances (Clutton-Brock, 1999). Among several other morphological characters, curliness of the tail and shape has been associated with domestication in mammals (Trut et al., 2009).

The tail is considered an extension of the spinal column usually composed of specifically shaped vertebrae. Spontaneous curly tail phenotypes in mice have been the matter of studies that investigated the role of embryonic development in this morphological anomaly (Copp et al., 1988; van Straaten and Copp, 2001; Ohnishi et al., 2017). Curly tail is also commonly observed in many dog breeds. Vaysse et al. (2011) compared the genome of dog breeds having curly tails with that of breeds with straight tails using single nucleotide polymorphisms (SNPs) chip data and identified a genomic region on chromosome 1 significantly associated with these alternative tail shapes.

In pigs, few studies have been reported on the genetic factors affecting tail shape. A putative recessive genetic defect known as kinky tail (or flexed or screw tail), derived by fused caudal vertebrae associated in some cases with other defects, has been described in the mid of the last century (Nordby, 1934; Donald, 1949; Brooksbank, 1958). It is not known if this defect could be, in some way, related or not to the normal curling of the tail that is common in domestic pigs. This signature of domestication, however, seems not fixed in all pig breeds (Porter, 1993) but no systematic study has been conducted so far, probably because the difficulties in retrieving phenotype information due to the usual practice of tail docking in most herds.

In this study, we took advantage from the variability of the shape of the tail that we recorded in the Casertana pig population and run a genome wide association study (GWAS) comparing the genome of curly-tailed and straight tailed animals to identify genomic regions associated with the tail shape phenotype in *Sus scrofa*.

2. Materials and methods

2.1. Animals

A total of 101 Casertana pigs (of about 7 to 20 months old) from six different farms were evaluated. Photographic records of each animal were obtained to capture information on the tail shape in standardized restraining conditions (including a direct evaluation of the personnel on this phenotype during this phase for the biological sampling) for all animals and after release (Figure 1). Pigs were classified as follows: 53 (25 males and 28 females) showed the curly tail phenotype; 19 (five males and 14 females) showed the straight tail phenotype; 29 were not classified and excluded from the study as tail docking, that was practised by the farmers as routine before weaning of the piglets, prevented the recording of any tail phenotype.

2.2. Genotyping

Hairs (with roots) were collected from the investigated pigs. DNA extraction was carried out using the Wizard® Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA) following the manufacturer's instructions. Genotyping of the extracted DNA was obtained with the Illumina PorcineSNP60 BeadChip v.2 (Illumina, Inc., San Diego, CA, USA) that interrogated 61,565 SNPs. Single nucleotide polymorphisms were assigned to the Sscrofa11.1 genome version, as previously described (Fontanesi et al., 2012). PLINK 1.9 software (Chang et al., 2015) was used to filter SNPs and genotyping data using the following criteria already used in a similar study (Schiavo et al., 2018): genotyping call rate >0.9, minor allele frequency >0.01 and Hardy-Weinberg equilibrium $P>0.001$.

2.3. Data analysis and genome wide association

To evaluate distance relationships among the animals of the investigated cohort, multidimensional scaling (MDS) was obtained with the PLINK 1.9 software (Chang et al., 2015). Genome wide association study was carried out by applying the univariate mixed model of GEMMA

(Zhou and Stephens, 2012) that can accommodate the centered relatedness matrix calculated from SNP genotypes to correct for population stratification in a case and control analysis. The model also included the farm and the sex as fixed effects. To be able to identify associated markers in this experiment that included a low number of animals (derived by the fact that the analysed pigs were almost a complete representation of the whole population of the Casertana breed) and that used a SNP chip that might originally have an ascertain bias (as local breeds were not used for the selection of the informative SNPs), the significant threshold was defined at the $P_{nominal\ value}$<5.00E-05 level, according to the Wellcome Trust Case Control Consortium (2007) and as also applied in several other GWAS in livestock (e.g. Fontanesi et al., 2012; Sanchez et al., 2014). Genomic inflation factor (λ) and quantile–quantile (Q–Q) plot were obtained with GenABEL (Aulchenko et al., 2007). Gene annotation information was retrieved from the Sscrofa11.1 genome version available at the Ensembl database (http://www.ensembl.org/Sus_scrofa/Info/Index), release 91.

3. Results and discussion

A recent phenotypic characterization of the endangered Casertana pig population that we carried out noted several morphological differences among distinct animals of this autochthonous breed (data not shown). For example, in addition to the hairless or hypotrichotic condition (that is the characteristic phenotype of the Casertana animals), we already described the presence of haired pigs in this population and this morphological variability was used for a GWAS that we have recently reported (Schiavo et al. 2018). Despite a limited number of animals was included in that study, we were able to identify genomic regions associated with the hairless phenotype, demonstrating that local animal genetic resources can be used to genetically describe phenotypic variability of simple traits (Schiavo et al., 2018).

Another morphological trait that is not fixed in this breed is the shape of the tail (Figure 1). Of the animals for which we could record this phenotype, 26% (19 out of 72) showed a straight tail without any curls, similarly to the usual shape of wild boars. This shape was clearly different from

158 the curly tails reported in the remaining investigated pigs (74%). There was no age effect and the two
159 groups included a comparable number of males and females so that sex did not explain the tail shape
160 phenotype. In addition, we could exclude the possible effect of the behavioral change of tail posture
161 on this phenotype (Zonderland et al., 2009). The recording system was based on standardized
162 conditions and subsequent photographic records of the animals confirmed, at least for the
163 photographic time point, their assignment to one or to the other group of tail shape phenotype. The
164 two groups were observed in animals from all six farms. Limited pedigree record prevented the
165 possibility to evaluate any potential founder effect.

166 A total of 36,533 autosomal SNPs, mapped to a unique position in the Sscrofa11.1 genome
167 version, was used for MDS. The obtained MDS plot showed some structures not well defined in the
168 analysed pigs that however did not clearly separate the curly and straight tailed Casertana pigs (Figure
169 2). A stratified sample could be a critical point in GWAS in a very small population where, to some
170 extent, all animals might be related. Figure S1 reports the genomic inflation factor (λ) and Q–Q plot
171 that did not show any biased test statistic distribution, suggesting that the investigated cohort was
172 corrected for a possible stratification effect.

173 Figure 3 reports the Manhattan plot obtained in this GWAS. One significant SNP (P=2.3E-05)
174 was identified on porcine chromosome 12 (SSC12). This marker indicated as ALGA0064877
175 (rs81439488 C/A) is located at position 10,301,075 of this chromosome. Allele A of this SNP was
176 more frequent in the straight tail group (0.17) than in the curly tailed group (0.03).

177 The complete list of genes annotated around the ALGA0064877 marker (with known functions)
178 is reported in Table S1. Figure S2 report a screenshot of this SSC12 region retrieved from ENSEMBL
179 database with the annotated genes. One of the closest gene in this desert chromosome region is the
180 *SRY-box 9* (*SOX9*) gene (positions 8,641,629-8,647,764, encoded by the -1 strand; actually the closest
181 upstream gene to the significant SNP), that, according to its function, might be the most plausible
182 candidate gene, explaining the recorded phenotypic variability. It is well established that the
183 expression of this gene at the embryonal level marks the onset of cartilage differentiation (Wright et

al., 1995; Healy et al., 1996). *SOX9* encodes for a transcription factor that is required during sequential steps of the chondrocyte differentiation pathway, notochord maintenance and skeletogenesis (Akiyama et al., 2002; Barrionuevo et al., 2006; Montero et al., 2017). Continued expression of *Sox9* in differentiated chondrocytes is essential for subsequent hypertrophy and sustains chondrocyte-specific survival mechanisms (Ikegami et al., 2011). Although *SOX9* seems to be a master gene for chondrocytes differentiation, the whole region surrounding the marker ALGA0064877 is downstream closer to a few genes [i.e. potassium voltage-gated channel subfamily J member 2 (*KCNJ2*), 61169 bp of distance from the SNP; potassium voltage-gated channel subfamily J member 16 (*KCNJ16*), 147461 bp of distance; mitogen-activated protein kinase kinase 6 (*MAP2K6*), 570407 bp of distance]. Alterations in the genes of this region have been also observed to affect the Wnt pathway (Kurth et al., 2009). Heterozygous mutations within and around human *SOX9* cause campomelic dysplasia that is a malformation syndrome characterized by cartilage derived skeletal structure defects (Foster et al., 1994; Wagner et al., 1994). These mutations, most of which reduce the level of expression of this gene, are located upstream spanning a large region (from 50 kb to more than 1 Mb) in which regulatory elements are present (Wunderle et al., 1998; Bagheri-Fam et al., 2006). Close upstream mutations produce more severe defects whereas far upstream mutations cause mild defects (Pfeifer et al., 1999; Velagaleti et al., 2005; Leipoldt et al., 2007). About 40 k SNPs, located between *SOX9* and *MAP2K6*, are reported in the Sscrofa11.1 genome version and could be considered for future studies to identify the causative mutation(s). Among them, 20 are also present in the PorcineSNP60 BeadChip but only ALGA0064877 reached the threshold of significance, probably due to the biased chip design that might not be able to capture particular haplotype structures of this breed.

Based on these studies in other species it is tempting to suggest a possible regulatory mechanism affecting *SOX9* expression in porcine developing chondrocytes that would, in turn, produce a mild cartilage/skeletal effect determining the shape of the tail. This hypothesis might be worth of further investigation starting from a precise characterization of the structure and morphology of the pig tail with different shapes for which, at present, there is no detailed investigation. Our phenotype records

were based only on an external morphological evaluation of the shape of the tail. Furthermore, analysis of gene expression of *SOX9* at different developmental stages should be also carried out to evaluate the role of this gene in the phenotype observed in pigs.

The results we obtained might have broader impacts than those that would be limited to a simple morphological characterization. The shape of the tail could be important in relation to the problem of tail biting in pigs. Tail biting is a widespread behavioral vice with significant animal welfare implications and economic losses in commercial pig farms (Bracke et al., 2004). A few studies have established correlations between tail posture and tail biting incidence suggesting limited damages and related welfare complications with behaviors of the pigs that tended to have a tail posture up that those with tail posture down (Zonderland et al., 2009; Lahrmann et al., 2018). It would be interesting to evaluate if pigs with genetically determined curly tails (as a possible adaptation derived by the domestication process) are less affected by tail biting damages than pigs with straight tails.

4. Conclusions

This work demonstrated that autochthonous animal genetic resources, even constituted by very small populations, might be used to disclose genetic factors affecting peculiar traits by exploiting segregating phenotypes and genetic variability. To our knowledge, this is the first study that reported a frequency distribution of the tail shape phenotype in a pig population. Our results indicated that this morphological trait is associated with a marker close to an important gene involved in embryonic development, opening other hypothesis, worth of further investigations. It will be important to validate the results we obtained in this GWAS in other breeds and populations, including a more precise anatomical characterization of this trait, to further extend the impact of the results reported in Casertana pigs. It would be however first needed to know if diversity for this morphological characteristic is common in commercial pig populations as at present, there is not information on this aspect, mainly due to the usual practice of tail docking that prevents the recording of this phenotype. Considering the potential relationship between tail shape and tail biting damages (that, however,

remains to be formally demonstrated), it could be possible to envisage practical applications of the identified marker in selection programs aimed to respond to animal welfare issues. From this study it emerges that conservation strategies of autochthonous pig genetic resources should take also into account the preservation of phenotypic variability within populations. Our study represents one of the few examples of exploitation of animal genetic resources to recover information that might have potential impacts in commercial populations.

Acknowledgements

The Authors thank all farmers and personnel that contributed to collect phenotypes and samples.

Funding

This work has received funding from the Italian Ministry of Agriculture, Food and Forestry (MiPAAF) under the project INNOVAGEN and 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 authors' view and the European Union Agency is not responsible for any use that may be made of the information it contains. This study was associated with the PSRN program (2017-2019).

Conflict of interest statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Akiyama, H., Chaboissier, M.C., Martin, J.F., Schedl, A., de Crombrugghe, B. 2002. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. Genes Dev. 16, 2813–2828.

ANAS, 2016. *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.

Bagheri-Fam, S., Barrionuevo, F., Dohrmann, U., Günther, T., Schüle, R., Kemler, R., Mallo, M., Kanzler, B., Scherer, G. 2006. Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. Dev. Biol. 291, 382–397.

Barrionuevo, F., Taketo, M. M., Scherer, G., Kispert, A. 2006. Sox9 is required for notochord maintenance in mice. Dev. Biol. 295, 128–140.

Bracke, M.B.M., Hulsegge, B., Keeling, L., Blokhuis, H.J. 2004. Decision support system with semantic model to assess the risk of tail biting in pigs. 1. Modelling. Appl. Anim. Behav. Sci. 87, 31–44.

Brooksbank, N.H. 1958. Congenital deformity of the tail in pigs. Br. Vet. J. 114, 50–55.

Carneiro, M., Rubin, C.J., Di Palma, F., Albert, F.W., Alföldi, J., Martinez Barrio, A., Pielberg, G., Rafati, N., Sayyab, S., Turner-Maier, J., Younis, S., Afonso, S., Aken, B., Alves, J.M., Barrell, D., Bolet, G., Boucher, S., Burbano, H.A., Campos, R., Chang, J.L., Duranthon, V., Fontanesi, L., Garreau, H., Heiman, D., Johnson, J., Mage, R.G., Peng, Z., Queney, G., Rogel-Gaillard, C., Ruffier, M., Searle, S., Villafuerte, R., Xiong, A., Young, S., Forsberg-Nilsson, K., Good, J.M., Lander, E.S., Ferrand, N., Lindblad-Toh, K., Andersson, L. 2014. Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. Science 345, 1074–1079.

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.

Clutton-Brock, J. 1999. *A Natural History of Domesticated Mammals.* 2nd Edition, Cambridge, UK: Cambridge University Press.

286 Copp, A. J., Brook, F. A., & Roberts, H. J. 1988. A cell-type-specific abnormality of cell proliferation
287 in mutant (curly tail) mouse embryos developing spinal neural tube defects. Development 104,
288 285–295.
289 Darwin, C. 1868. *The Variation of Animals and Plants under Domestication.* London, UK: John
290 Murray.
291 Donald, H. P., 1949. The inheritance of a tail abnormality associated with urogenital disorders in pigs.
292 J. Agric. Sci. 39, 164–173.
293 Fontanesi, L., Schiavo, G., Galimberti, G., Calò, D.G., Scotti, E., Martelli, P.L., Buttazzoni, L.,
294 Casadio, R., Russo, V. 2012. A genome wide association study for backfat thickness in Italian
295 Large White pigs highlights new regions affecting fat deposition including neuronal genes.
296 BMC Genomics 13, 583.
297 Foster, J.W., Dominguez-Steglich, M.A., Guioli, S., Kwok, C., Weller, P.A., Stevanović, M.,
298 Weissenbach, J., Mansour, S., Young, I.D., Goodfellow, P.N., et al. 1994. Campomelic
299 dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature 372,
300 525–530.
301 Healy, C., Uwanogho, D., Sharpe, P.T. 1996. Expression of the chicken Sox9 gene marks the onset
302 of cartilage differentiation. Ann. N. Y. Acad. Sci. 785**,** 261–262.
303 Ikegami, D., Akiyama, H., Suzuki, A., Nakamura, T., Nakano, T., Yoshikawa, H., Tsumaki, N. 2011.
304 Sox9 sustains chondrocyte survival and hypertrophy in part through Pik3ca-Akt pathways.
305 Development 138, 1507–1519.
306 Kurth, I., Klopocki, E., Stricker, S., van Oosterwijk, J., Vanek, S., Altmann, J., Santos, H.G., van
307 Harssel, J.J., de Ravel, T., Wilkie, A.O., Gal, A., Mundlos, S. 2009 Duplications of
308 noncoding elements 5' of SOX9 are associated with brachydactyly-anonychia. Nat. Genet.
309 41, 862–863.
310 Lahrmann, H. P., Hansen, C. F., D'Eath, R., Busch, M. E., & Forkman, B. (2018). Tail posture
311 predicts tail biting outbreaks at pen level in weaner pigs. Appl. Anim. Behav. Sci. 200, 29–35.

312 Larson, G., Burger, J. 2013. A population genetics view of animal domestication. Trends Genet. 29,
313 197–205.
314 Leipoldt, M., Erdel, M., Bien-Willner, G.A., Smyk, M., Theurl, M., Yatsenko, S.A., Lupski, J.R.,
315 Lane, A.H., Shanske, A.L., Stankiewicz, P., Scherer, G. 2007. Two novel translocation
316 breakpoints upstream of SOX9 define borders of the proximal and distal breakpoint cluster
317 region in campomelic dysplasia. Clin. Genet. 71, 67–75.
318 Leroy, G., Besbes, B., Boettcher, P., Hoffmann, I., Capitan, A., Baumung, R. 2016. Rare phenotypes
319 in domestic animals: unique resources for multiple applications. Anim. Genet. 47, 141–153.
320 Montero, J.A., Lorda-Diez, C.I., Francisco-Morcillo, J., Chimal-Monroy, J., Garcia-Porrero, J.A.,
321 Hurle, J.M. 2017. Sox9 expression in Amniotes: Species-specific differences in the formation
322 of digits. Front. Cell Dev. Biol. 5, 23.
323 Nordby J.E. 1934. Kinky tail in swine. J. Hered. 25, 171–174.
324 Ohnishi, T., Miura, I., Ohba, H., Shimamoto, C., Iwayama, Y., Wakana, S., Yoshikawa, T. 2017. A
325 spontaneous and novel Pax3 mutant mouse that models Waardenburg syndrome and neural tube
326 defects. Gene 607, 16–22.
327 Pfeifer, D., Kist, R., Dewar, K., Devon, K., Lander, E.S., Birren, B., Korniszewski, L., Back, E.,
328 Scherer, G. 1999. Campomelic dysplasia translocation breakpoints are scattered over 1 Mb
329 proximal to SOX9: evidence for an extended control region. Am. J. Hum. Genet. 65, 111–124.
330 Porter, V. 1993. *Pigs: A Handbooks to the Breeds of the World*. Cornell University Press,
331 Sanchez M.P., Tribout T., Iannuccelli N., Bouffaud, M., Servin, B., Tenghe, A., Dehais, P., Muller,
332 N., Del Schneider, M.P., Mercat, M.J., Rogel-Gaillard, C., Milan, D., Bidanel, J.P., Gilbert, H.
333 2014. A genome-wide association study of production traits in a commercial population of
334 Large White pigs: evidence of haplotypes affecting meat quality. Genet. Sel. Evol. 46, 12.
335 Schiavo, G., Bertolini, F., Utzeri, V.J., Ribani, R., Geraci, C., Santoro, L., Óvilo, C., Fernández, A.I.,
336 Gallo, M., Fontanesi, L. 2018. Taking advantage from phenotype variability in a local animal

337 genetic resource: identification of genomic regions associated with the hairless phenotype in
338 Casertana pigs. Anim. Genet. 49, 321–325.
339 Trut, L., Oskina, I., Kharlamova, A. 2009. Animal evolution during domestication: the domesticated
340 fox as a model. Bioessays 31, 349–360.
341 van Straaten, H. W., Copp, A. J. 2001. Curly tail: a 50-year history of the mouse spina bifida model.
342 Anat. Embryol. 203, 225–238.
343 Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., Fall,
344 T., Seppälä, E.H., Hansen, M.S., Lawley, C.T., Karlsson, E.K.; LUPA Consortium, Bannasch,
345 D., Vilà, C., Lohi, H., Galibert, F., Fredholm, M., Häggström, J., Hedhammar, A., André, C.,
346 Lindblad-Toh, K., Hitte, C., Webster, M.T. 2011. Identification of genomic regions associated
347 with phenotypic variation between dog breeds using selection mapping. PLoS Genet. 7,
348 e1002316.
349 Velagaleti, G.V., Bien-Willner, G.A., Northup, J.K., Lockhart, L.H., Hawkins, J.C., Jalal, S.M.,
350 Withers, M., Lupski, J.R., Stankiewicz, P. 2005. Position effects due to chromosome
351 breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream
352 of SOX9 in two patients with campomelic dysplasia. Am. J. Hum. Genet. 76, 652–662.
353 Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F.D., Keutel,
354 J., Hustert, E., Wolf, U., Tommerup, N., Schempp, W., Scherer, G. 1994. Autosomal sex
355 reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene
356 SOX9. Cell 79, 1111–1120.
357 Wang, G.D., Xie, H.B., Peng, M.S., Irwin, D., Zhang, Y.P. 2014. Domestication genomics: evidence
358 from animals. Annu. Rev. Anim. Biosci. 2, 65–84.
359 Wiener, P., Wilkinson, S. 2011. Deciphering the genetic basis of animal domestication. Proc. Biol.
360 Sci. 278, 3161–3170.
361 Wilkins, A. S., Wrangham, R. W., Fitch, W. T. 2014. The “domestication syndrome” in mammals: a
362 unified explanation based on neural crest cell behavior and genetics. Genetics 197, 795–808.

Wright, E., Hargrave, M.R., Christiansen, J., Cooper, L., Kun, J., Evans, T., Gangadharan, U., Greenfield, A., Koopman, P. 1995. The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. Nat. Genet. 9, 15–20.

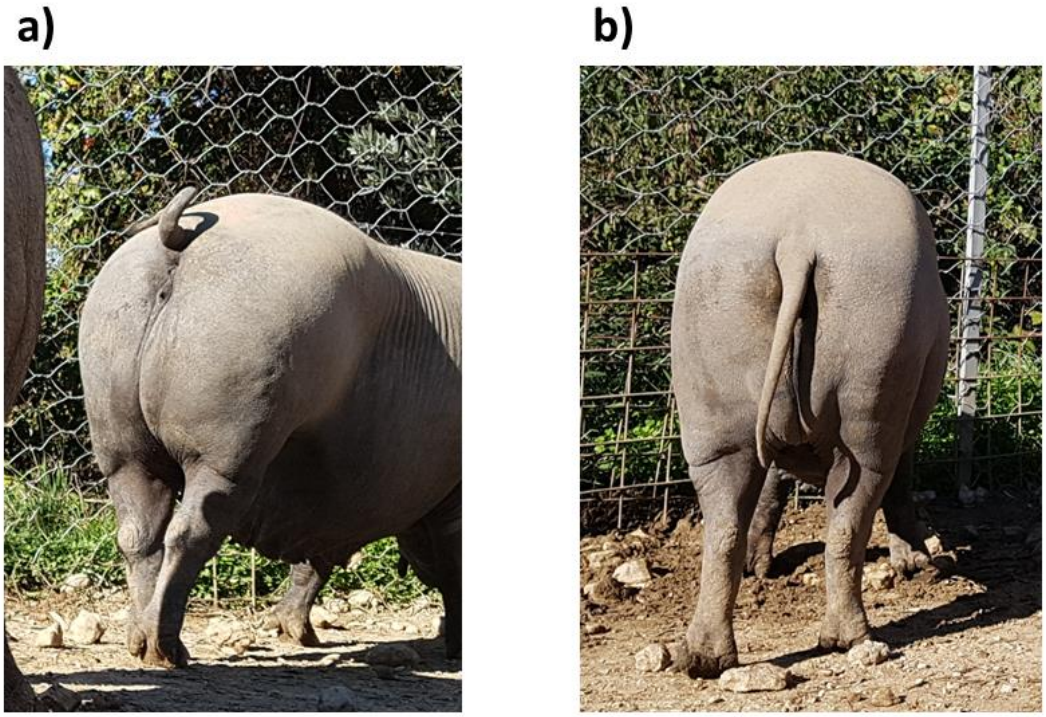
Wunderle, V.M., Critcher, R., Hastie, N., Goodfellow, P.N., Schedl, A. 1988. Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. Proc. Natl. Acad. Sci USA 95, 10649–10654.

Zhou, X., Stephens, M. 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44, 821–824.

Zonderland, J.J., van Riel, J.W., Bracke, M.B.M., Kemp, B., den Hartog, L.A., Spoolder, H.A.M., 2009. Tail posture predicts tail damage among weaned piglets. Appl. Anim. Behav. Sci. 121, 165–170.

374 **Figure 1.** Tail shape of Casertana pigs: a) curly tail; b) straight tail.

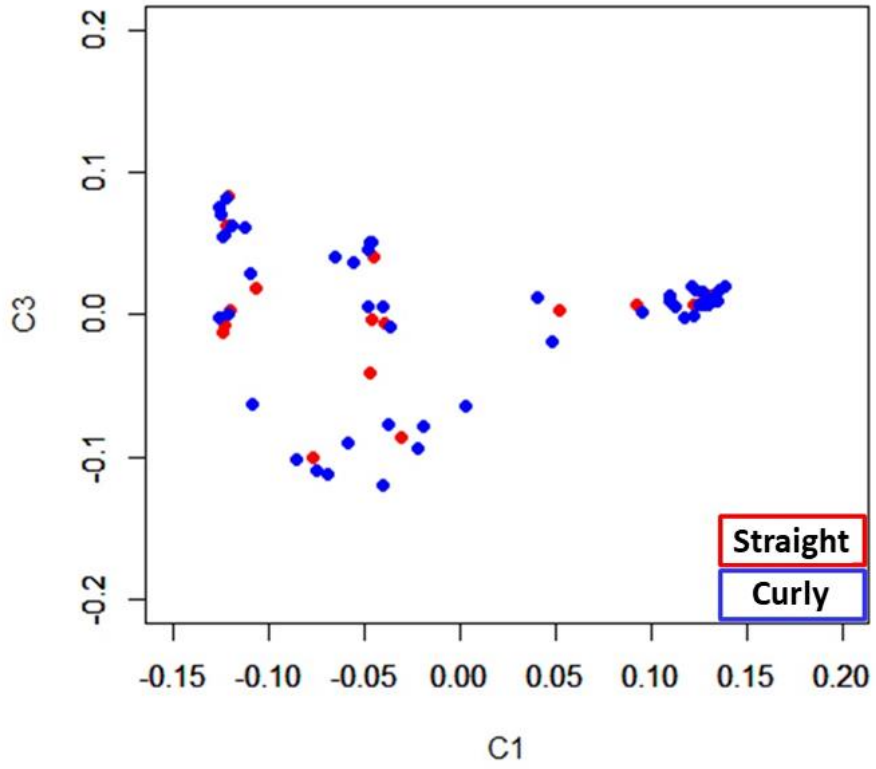
375



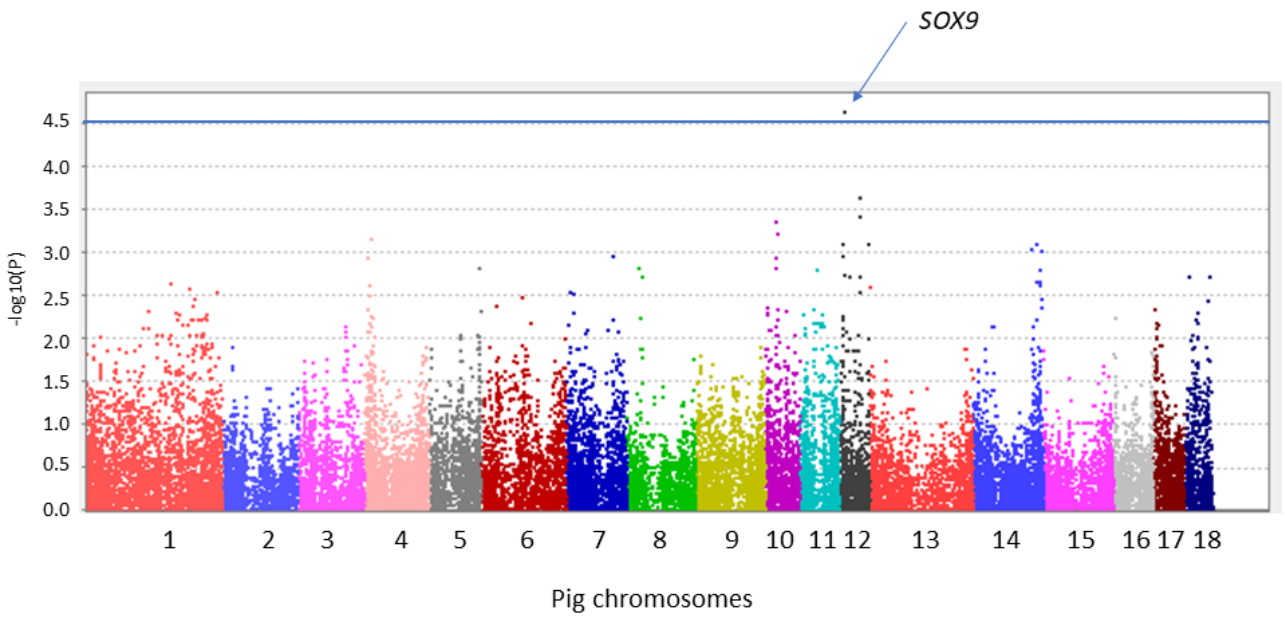
376

377

Figure 2. Multidimensional scaling (MDS) with represented the pigs (dots) included in this study divided in the two groups of tail shape.



381 **Figure 3.** Manhattan plot obtained for the genome wide association study.



382

Supplementary material

Figure S1. Quantile–quantile (Q–Q) plot obtained from the genome wide association analysis.

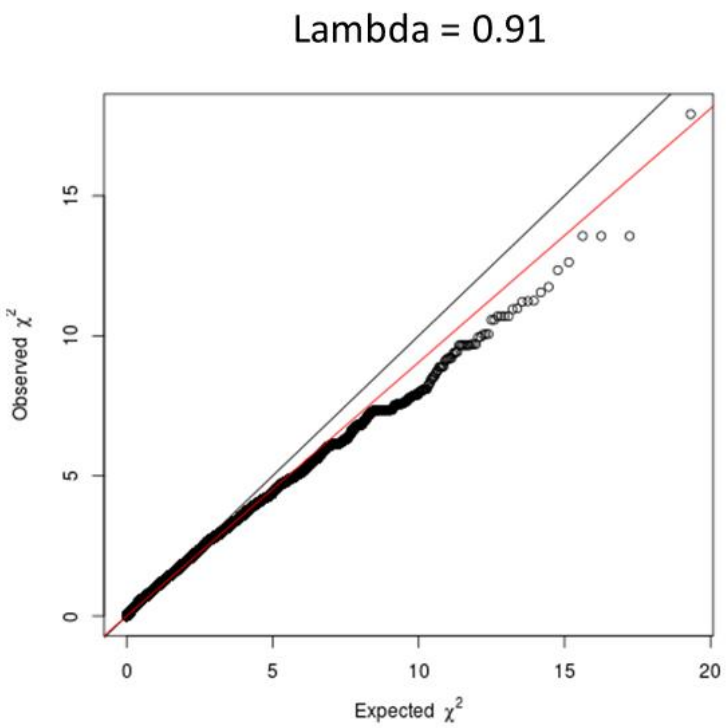
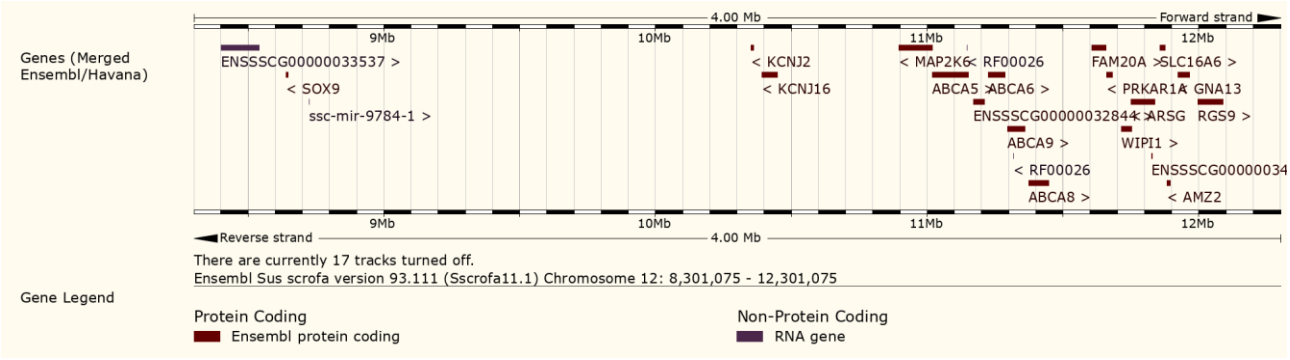


Figure S2. Graphical representation of the annotated genes in the porcine chromosome 12 region around the ALGA0064877 (rs81439488 C/A; nucleotide position 10,301,075. The screenshot has been retrieved from the ENSEMBL database (http://www.ensembl.org/index.html; release 93, July 2018).



393 **Table S1.** List of genes annotated in the Sscrofa11.1 genome version around the ALGA0064877

394 (rs81439488 C/A; nucleotide position 10,301,075) marker (2 Mbp on both directions) on porcine

395 chromosome 12.

Position	Gene symbol	Gene name	Role/function[1]
8,641,629-8,647,764	*SOX9*	SRY-box 9	The encode protein is involved in chondrocyte differentiation and, with steroidogenic factor 1, regulates transcription of the anti-Muellerian hormone (AMH) gene. Deficiencies lead to the skeletal malformation syndrome campomelic dysplasia
10,351,715-10,362,244	*KCNJ2*	potassium voltage-gated channel subfamily J member 2	The protein is an integral membrane protein and inward-rectifier type potassium channel. It probably participates in establishing action potential waveform and excitability of neuronal and muscle tissues. Mutations in this gene have been associated with Andersen syndrome in humans, which is characterized by periodic paralysis, cardiac arrhythmias, and dysmorphic features.
10,393,179-10,448,536	*KCNJ16*	potassium voltage-gated channel subfamily J member 16	Similar to the function of the previous gene product. It may act in fluid and pH balance regulation.
10,897,104-11,018,943	*MAP2K6*	mitogen-activated protein kinase kinase 6	It encodes for a member of the dual specificity protein kinase family, which functions as a mitogen-activated protein (MAP) kinase kinase. MAP kinases, also known as extracellular signal-regulated kinases (ERKs), act as an integration point for multiple biochemical signals. This gene is involved in many cellular processes such as stress induced cell cycle arrest, transcription activation and apoptosis.
11,018,687-11,152,913	*ABCA5*	ATP binding cassette subfamily A member 5	The encode protein is is a member of the superfamily of ATP-

			binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intracellular membranes. Its function is not well known.
11,169,809-11,211,801	*ABCA10* (ENSSSCG00000032844)	ATP-binding cassette sub-family A member 10	Similar to the previous gene.
11,225,408-11,286,018	*ABCA6*	ATP binding cassette subfamily A member 6	Similar to the previous gene.
11,295,697-11,359,520	*ABCA9*	ATP-binding cassette sub-family A member 9	Similar to the previous gene.
11,374,028-11,447,292	*ABCA8*	ATP-binding cassette sub-family A member 8	Similar to the previous gene.
11,607,233-11,657,737	*FAM20A*	FAM20A, golgi associated secretory pathway pseudokinase	This gene encodes a protein that is likely secreted and may function in hematopoiesis. A mutation at this locus has been associated with amelogenesis imperfecta and gingival hyperplasia syndrome in humans.
11,661,872-11,681,435	*PRKAR1A*	protein kinase cAMP-dependent type I regulatory subunit alpha	This gene encodes for one of the regulatory subunits of the kinase holoenzyme. This protein was found to be a tissue-specific extinguisher that down-regulates the expression of several liver genes. Mutations in this gene cause Carney complex (CNC) in humans. Other functions have been inferred based on the protein structure.
11,715,274-11,752,377	*WIPI1*	WD repeat domain, phosphoinositide interacting 1	It encodes a WD40 repeat protein, which is a key components of many essential biologic functions, by regulating the assembly of multiprotein complexes.
11,826,578-11,827,867	ENSSSCG00000034083	Novel gene	Not known
11,857,792-11,875,313	*SLC16A6*	solute carrier family 16 member 6	Not well defined.
11,882,800-11,895,486	*AMZ2*	archaelysin family metallopeptidase 2	The encoded protein is a zinc metalloprotease that displays some activity against angiotensin-3. The encoded protein is inhibited by the aminopeptidase inhibitor amastatin, as well as by the general inhibitors o-phenanthroline and batimastat. Defects in this

			gene may be associated with lung tumorigenesis.
11,922,572-11,966,747	*GNA13*	G protein subunit alpha 13	Not well defined.
11,996,142-12,089,466	*RGS9*	regulator of G protein signaling 9	This gene encodes a member of the RGS family of GTPase activating proteins that function in various signaling pathways by accelerating the deactivation of G proteins. This protein is anchored to photoreceptor membranes in retinal cells and deactivates G proteins in the rod and cone phototransduction cascades. Mutations in this gene result in bradyopsia in humans.

[1] Information has been adapted from the NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene/; August 2018).