

1 *This is the pre-peer reviewed version of the following article: “Taking advantage from phenotype*  
2 *variability in a local animal genetic resource: identification of genomic regions associated with the*  
3 *hairless phenotype in Casertana pigs” by Giuseppina Schiavo, Francesca Bertolini, Valerio Joe*  
4 *Utzeri, Anisa Ribani, Claudia Geraci, Laura Santoro, Cristina Óvilo, Ana I. Fernández, Maurizio*  
5 *Gallo and Luca Fontanesi, which has been published in final form at*  
6 <https://doi.org/10.1111/age.12665> *. This article may be used for non-commercial purposes in*  
7 *accordance with Wiley Terms and Conditions for Self-Archiving.”*

8

9 **SHORT COMMUNICATION**

10

11 **Taking advantage from phenotype variability in a local animal genetic resource: identification**  
12 **of genomic regions associated with the hairless phenotype in Casertana pigs**

13

14 Giuseppina Schiavo<sup>1,\*</sup>, Francesca Bertolini<sup>1,2,\*</sup>, Valerio Joe Utzeri<sup>1</sup>, Anisa Ribani<sup>1</sup>, Claudia Geraci<sup>1</sup>,  
15 Laura Santoro<sup>3</sup>, Cristina Óvilo<sup>4</sup>, Ana I. Fernández<sup>4</sup>, Maurizio Gallo<sup>5</sup> and Luca Fontanesi<sup>1</sup>

16

17 <sup>1</sup> Department of Agricultural and Food Sciences, Division of Animal Sciences, University of  
18 Bologna, Viale Fanin 46, 40127 Bologna, Italy

19 <sup>2</sup> Department of Animal Science, Iowa State University, 2255 Kildee Hall, 50011 Ames, Iowa,  
20 USA

21 <sup>3</sup> ConSDABI - National Focal Point Italiano FAO, Contrada Piano Cappelle, 82100 Benevento,  
22 Italy

23 <sup>4</sup> Departamento de Mejora Genética Animal, Instituto Nacional de Tecnología Agraria y  
24 Alimentaria (INIA), Ctra. de la Coruña km. 7.5, 28040 Madrid, Spain

25 <sup>5</sup> Associazione Nazionale Allevatori Suini, Via L. Spallanzani 4, 00161 Roma, Italy

26

27 \* These authors contributed equally to this work

28

29 Corresponding author: luca.fontanesi@unibo.it

30

31

32 **Running title:** Hairless in Casertana pigs

### 33 **Summary**

34 Casertana is an endangered autochthonous pig breed (raised in the Central-South of Italy) that  
35 is considered the descendant of the influential Neapolitan pig population that was used to improve  
36 British breeds in the 19<sup>th</sup> century. Casertana pigs are characterized by a typical, almost complete,  
37 hairless phenotype. Despite this phenotype is the characteristic trait of this breed, few Casertana pigs  
38 are normal-haired. In this work, using Illumina PorcineSNP60 BeadChip data, we carried out a  
39 genome wide association study (GWAS) and an  $F_{ST}$  analysis in this breed by comparing animals  
40 showing the classical hairless phenotype (n. 81) versus pigs classified as haired (n. 15). Combining  
41 results obtained with the two approaches, we identified two significant regions, one on porcine  
42 chromosome (SSC) 7 and one on SSC15. The SSC7 region contains the *forkhead box N3 (FOXN3)*  
43 gene, the most plausible candidate gene of this region, considering that mutations in another gene of  
44 the same family (forkhead box N1; *FOXN1*) are responsible for the *nude* locus in rodents and alopecia  
45 in humans. Another potential candidate gene, *Rho guanine nucleotide exchange factor 10*  
46 (*ARHGEF10*) is located on the SSC15 region. *FOXN3* and *ARHGEF10* have been detected as  
47 differentially expressed in androgenetic and senescent alopecia, respectively. This study in an  
48 autochthonous pig breed contributed to shed some lights on novel genes potentially involved in hair  
49 development and growth, demonstrating that local animal breeds can be valuable genetic resources  
50 to disclose genetic factors affecting unique traits, taking advantage from phenotype variability  
51 segregating in small populations.

52

53 **Key words:** alopecia, animal genetic resource, animal model, baldness, hairless,  $F_{ST}$ , GWAS, SNP.

54 **Text**

55 Local animal genetic resources might be characterized by specific and inheritable phenotypes  
56 with relevant importance for current or potential future use in breeding programs or for many other  
57 purposes, including the definition of new biological models or to understand mechanisms of  
58 biological adaptations to different environments (Leroy *et al.* 2016).

59 Casertana is an endangered autochthonous pig breed mainly raised in the Central-South of Italy,  
60 accounting for about 100 boars and sows currently registered to its herd book (ANAS 2016).  
61 Casertana pigs are usually raised in extensive or semi-extensive systems to produce niche pork  
62 products. This local breed is considered the descendant of the influential Neapolitan breed of the late  
63 18th and 19th centuries that was used to improve British pig populations from which several modern  
64 commercial breeds were derived (Porter 1993). Casertana pigs are characterized by a black or grey  
65 coat colour, wrinkled skin, forward ears, two goatlike wattles (not always present) and a typical,  
66 almost complete, hairless phenotype. This later characteristic is also reported in one of its local names,  
67 i.e. *Pelatella* (that means plucked or bald). Despite the hairless phenotype is the characteristic  
68 phenotype of this breed, Casertana population shows some variability for this trait, including animals  
69 having from almost complete absence of hairs (hairless; the most common pigs) to few animals  
70 having abundant hairs (normal-haired pigs; Figure 1a). The hairless phenotype is also present in other  
71 pig breeds like the Creole hairless Mexican breed (also known as Pelón Mexicano) and the black  
72 hairless Iberian strains, including the Guadyerbas strain maintained as isolated population (Toro *et*  
73 *al.* 2000; Lemus-Flores *et al.* 2001). Casertana and all these other hairless pigs seem historically  
74 connected through exchange of pig genetic material determined by commercial activities in the 18<sup>th</sup>  
75 and 19<sup>th</sup> centuries (Porter 1993), suggesting a potential common origin of the hairless phenotype.

76 Hairless or hairlessness in pigs can be better described as hypotrichosis or congenital deficiency  
77 of hairs, as animals classified as “hairless” usually show a small number rather than a complete  
78 absence of hairs. Roberts & Carroll (1931) were the first authors that reported a possible inheritance  
79 model for this hypotrichotic condition in Mexican pigs, suggesting the presence of a monogenic factor

80 with a recessive mutated *h* allele that could give the hairless phenotype when homozygous.  
81 Homozygous pigs for the wild type allele *H* might be normal-haired whereas heterozygous *Hh* pigs  
82 might show an intermediate phenotype. This early study was not followed by any other genetic  
83 investigations on the hairless condition in pigs. More recently, variability in the porcine *hairless* gene  
84 (known as *HR*, *lysine demethylase and nuclear receptor corepressor*), located on porcine  
85 chromosome (SSC) 14, was evaluated in a candidate gene approach to study the hairless phenotype  
86 in Iberian pigs but no association with this trait was reported (Fernández *et al.* 2003, 2006). Mutations  
87 in the *HR* gene have been shown to impair hair growth in different mammals (i.e. Stoye *et al.* 1988;  
88 Ahmad *et al.* 1998; Finocchiaro *et al.* 2003). A high number of other genes in humans and rodents  
89 have been implicated in abnormal hair development and hypotrichosis (Shimomura & Christiano  
90 2010; Ramot & Zlotogorski 2015), making impractical a candidate gene approach to successfully  
91 identify polymorphisms associated with the hairless phenotype in pig populations.

92 In this work, with the aim to restrict the number of potential causative genes involved in the  
93 hypotrichotic phenotype in pigs, we carried out a genome wide association study (GWAS) and a  
94 genome wide *F<sub>ST</sub>* analysis in the Casertana breed by comparing animals showing the classical hairless  
95 phenotype (n. 81) versus pigs classified as haired (n. 15; a quite rare phenotype in this breed), without  
96 any distinction between possible different hair levels that could not be precisely recorded in outdoor  
97 animals. Casertana breed offers a unique possibility to investigate this phenotype that is segregating  
98 within the same population. This is one of the first population based genome wide study in a local pig  
99 breed that is not only useful to characterize a breed specific trait but also to obtain basic biology  
100 information that could be important to better define an interesting animal model for alopecia or related  
101 phenotypes in humans (Shimomura 2012).

102 Blood or hair roots were collected from all these Casertana pigs raised in six different farms  
103 (having from 5 to 49 pigs each, with unknown relationships) and extracted DNA was used for  
104 genotyping with the Illumina PorcineSNP60 BeadChip v.2 (Illumina, Inc., San Diego, CA, USA)  
105 interrogating 61,565 single nucleotide polymorphisms (SNPs). Genotyping data were processed with

106 PLINK 1.9 software (Chang *et al.* 2015) using the following criteria to filter SNPs: call rate >0.9,  
107 minor allele frequency >0.01 and Hardy-Weinberg equilibrium  $P > 0.001$ . A total of 36,533 autosomal  
108 SNPs assigned to a unique position in the Sscrofa11.1 genome version was then used in the GWAS  
109 that was carried out by applying the univariate mixed model of GEMMA (Zhou & Stephens 2012).  
110 The centered relatedness matrix calculated from SNP genotypes was included in the model to correct  
111 for population stratification. Figure S1 reports the genomic inflation factor ( $\lambda$ ) and quantile–quantile  
112 (Q–Q) plot, obtained with GenABEL (Aulchenko *et al.* 2007). Figure 1b reports the Manhattan plot  
113 obtained in this GWAS. At the  $P < 0.05$  Bonferroni corrected level ( $P$  nominal value  $< 1.37E-06$ ),  
114 three SNPs were significant whereas at the  $P < 0.1$  Bonferroni corrected threshold ( $P$  nominal value =  
115  $2.74E-06$ ) other three SNPs were suggestively significant (Table 1). Two of these SNPs were located  
116 on SSC7 (170.17 kb apart) and four on SSC15, in two distinct regions of approximately 1.14 Mb and  
117 338.61 kb.

118  $F_{ST}$  analysis was performed on the same dataset using PLINK 1.9 software. Missing SNPs were  
119 imputed using the Beagle 3.3.2 software (Browning and Browning, 2009). Figure 1c reported the  
120 Manhattan plot of the  $F_{ST}$  analysis. The top 0.9998 SNPs of the percentile distribution ( $F_{ST}=0.345$ )  
121 were considered as the most divergent across the comparison and therefore retained for subsequent  
122 evaluation (Table 1). A total of 8 SNPs was above the selected threshold: one on SSC4, one on SSC2,  
123 two on SSC7 (170.17 kb apart), two on SSC15 (1.14 Mb apart) and two on SSC17 (32.00 kb apart).

124 The comparison among GEMMA and  $F_{ST}$  genome-wide analyses identified two overlapping  
125 regions encompassing two SNPs on SSC7 and two SNPs on SSC15 that constituted the 1.14 Mb  
126 region previously mentioned (Table 1). A total of eight and nine genes were annotated in the SSC7  
127 and SSC15 regions, respectively (in a window  $\pm 500$  kb from the first and the last SNPs; Table 1).  
128 The most plausible candidate gene in the SSC7 region was the *forkhead box N3* (*FOXP3*) gene  
129 (position: 111036492-111454106 bp), that is 66.56 kb far from INRA0028322 (one of the two most  
130 significant SNPs in the GWAS; Table 1). This gene has a role in the regulation of hepatic glucose  
131 utilization (Karanth *et al.* 2016), craniofacial development (Samaan *et al.* 2010) and growth and

132 migration of colon cancer cells (Dai *et al.* 2017). The *FOXN3* gene was also found differentially  
133 expressed in a case-control study for androgenetic alopecia in humans (Mirmirani & Karnik 2010).  
134 Forkhead box proteins constitute a family of transcription factors involved in embryo and fetal  
135 development and function of adult organisms (Hannenhalli & Kaestner 2009). This group of proteins  
136 list about 50 members in mammals, divided in 19 subfamilies indicated with the letters from A to S  
137 (Jackson *et al.* 2010; Benayoun *et al.* 2011). Among the N subfamily, forkhead box N1 (FOXN1)  
138 regulates keratin gene expression and the gene is responsible for the *nude* locus in rodents (Flanagan  
139 1966; Meier *et al.* 1999). Mutations in this gene determine hairlessness, alopecia and other pleiotropic  
140 effects in mice and rats (Nehls *et al.* 1994) and congenital alopecia, nail dystrophy, and primary T-  
141 cell immunodeficiency in humans (Frank *et al.* 1999). Therefore, considering the phylogenetic  
142 relationships and the partially conserved domains between the *FOXN1* and *FOXN3* genes (Benayoun  
143 *et al.* 2011), it seems plausible that FOXN3 might have conserved similar regulatory functions of  
144 FOXN1 that could explain the effect of this SSC7 chromosome region on the hairless phenotype in  
145 Casertana pigs. This indication might contribute to understand the involvement of forkhead box  
146 proteins in hair development and, if confirmed by functional studies, adds another candidate gene to  
147 the list of those potentially involved in alopecia and baldness.

148 No strong candidate gene could be identified in the SSC15 region. A possible candidate could  
149 be *Rho Guanine Nucleotide Exchange Factor 10 (ARHGEF10)* gene. ARHGEF10 is involved in  
150 neural morphogenesis and connectivity and in the regulation of small RhoGTPases (Verhoeven *et al.*  
151 2003). The *ARHGEF10* has been reported to be differentially expressed in a case-control study of  
152 senescent alopecia in human (Mirmirani & Karnik 2010), supporting, to some extent, its possible role  
153 in the hairless phenotype in the Casertana breed. According to the available functional information,  
154 no other gene in the two identified regions might be involved in hair or follicle development or  
155 phenotypes similar to the hairless condition we investigated.

156 The combination of the GWAS and  $F_{ST}$  results with the annotated gene functions was useful to  
157 draft a possible biological explanation of the hairless phenotype in Casertana pigs and to identify

158 significant regions, excluding other regions that reached or were close to the defined thresholds in  
159 one or the other genome wide investigation methods derived by several confounding factors that  
160 could not be better managed in our study (i.e. genetic drift, population structure, ascertain bias of the  
161 SNP chip tool). However, the results obtained in this breed, even if based on a small group of pigs  
162 with normal-haired phenotype (that is a quite rare in this breed) in contrast with the hairless group,  
163 seems to support the presence of more than one locus affecting this trait. A few of the associated  
164 genomic regions contain candidate genes that, based on their function or inferred function may be  
165 involved in the hypotricotic condition of the Casertana pigs, with the hypothesis that this trait might  
166 be more complex than previously suggested.

167 This work demonstrated that endangered animal genetic resources could be investigated to  
168 disclose genetic factors affecting unique traits taking advantage from phenotype variability  
169 segregating within a small population. Other investigations are needed to refine these results obtained  
170 in Casertana and to evaluate if the hairless condition in other pig breeds is derived by the same genetic  
171 factors identified in this study.

172

### 173 **Competing interests**

174 The authors declare that they do not have competing interests. Data reported in this work can be  
175 shared after signature of an agreement on their use with University of Bologna.

176

### 177 **Acknowledgements**

178 This work has received funding from the Italian Ministry of Agriculture, Food and Forestry  
179 (MiPAAF) under the project INNOVAGEN and from the European Union's Horizon 2020 research  
180 and innovation programme under grant agreement No 634476 (Project acronym: TREASURE). The  
181 content of this paper reflects only the author's view and the European Union Agency is not  
182 responsible for any use that may be made of the information it contains.

183



184 **References**

- 185 Ahmad W., Haque M.F., Brancolini V., *et al.* (1998) Alopecia universalis associated with a  
186 mutation in the human hairless gene. *Science* **219**, 720-4.
- 187 ANAS (2016) Registro Anagrafico. <http://www.anas.it/>.
- 188 Aulchenko Y.S., Ripke S., Isaacs A. & van Duijn C.M. (2007) GenABEL: an R library for genome-  
189 wide association analysis. *Bioinformatics* **23**, 1294-6.
- 190 Benayoun B.A., Caburet, S. & Veitia R.A. (2011) Forkhead transcription factors: key players in  
191 health and disease. *Trends in Genetics* **27**, 224-32.
- 192 Browning B.L. & Browning S.R. (2009) A unified approach to genotype imputation and haplotype  
193 phase inference for large data sets of trios and unrelated individuals. *American Journal of*  
194 *Human Genetics* **84**, 210-23.
- 195 Chang C.C., Chow C.C., Tellier L.C., Vattikuti S., Purcell S.M. & Lee J.J. (2015) Second-generation  
196 PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**:7.
- 197 Dai Y., Wang M., Wu H., Xiao M., Liu H. & Zhang D. (2017) Loss of FOXN3 in colon cancer  
198 activates beta-catenin/TCF signaling and promotes the growth and migration of cancer cells.  
199 *Oncotarget* **8**, 9783-93.
- 200 Fernández A., Silió L., Noguera JL., Sánchez A. & Óvilo C. (2003) Linkage mapping of the porcine  
201 hairless gene (*HR*) to chromosome 14. *Animal Genetics* **34**, 317-8.
- 202 Fernández A.I., Silió L. & Óvilo C. (2006) Caracterization del gen hairless, candidato para el  
203 fenotipo lampino característico de una variedad de cerdo Ibérico. *Proceedings of the XIII*  
204 *Reunión Nacional de Mejora Genética Animal*, 28-30 June 2006, Gijón. Spain.
- 205 Finocchiaro R., Portolano B., Damiani G., *et al.* (2003) The hairless (*hr*) gene is involved in the  
206 congenital hypotrichosis of Valle del Belice sheep. *Genetics Selection and Evolution* **35**,  
207 S147-56.
- 208 Flanagan S.P. (1966) 'Nude' a new hairless gene with pleiotropic effects in the mouse. *Genetic*  
209 *Research* **8**, 295-309.

210 Frank J., Pignata C., Panteleyev A.A., *et al.* (1999) Exposing the human nude phenotype. *Nature*  
211 **398**:473-4.

212 Hannenhalli S. & Kaestner K.H. (2009) The evolution of Fox genes and their role in development  
213 and disease. *Nature Reviews Genetics* **10**, 233-40.

214 Jackson B.C., Carpenter C., Nebert D.W. & Vasiliou V. (2010) Update of human and mouse  
215 forkhead box (FOX) gene families. *Human Genomics* **4**, 345-52.

216 Karanth S., Zinkhan E.K., Hill J.T., Yost H.J. & Schlegel A. (2016) FOXN3 regulates hepatic  
217 glucose utilization. *Cell Reports* **15**, 2745-55.

218 Lemus-Flores C., Ulloa-Arvizu R., Ramos-Kuri M., Estrada F.J. & Alonso R.A. (2001) Genetic  
219 analysis of Mexican hairless pig populations. *Journal of Animal Science* **79**, 3021-6.

220 Leroy G., Besbes B., Boettcher P., Hoffmann I., Capitan A. & Baumung R. (2016) Rare phenotypes  
221 in domestic animals: unique resources for multiple applications. *Animal Genetics* **47**, 141-53.

222 Meier N., Dear T.N. & Boehm T. (1999) Wnt and mHa3 are components of the genetic hierarchy  
223 controlling hair follicle differentiation. *Mechanisms of Development* **89**, 215-21.

224 Mirmirani P. & Karnik P. (2010) Comparative gene expression profiling of senescent and  
225 androgenetic alopecia using microarray analysis. In: *Aging Hair*. (Trueb R.M. & Tobin D.J.,  
226 eds), New York: Springer, pp. 67–76.

227 Nehls M., Pfeifer D., Schorpp M., Hedrich H. & Boehm T. (1994) New member of the winged-  
228 helix protein family disrupted in mouse and rat nude mutations. *Nature* **372**, 103-7.

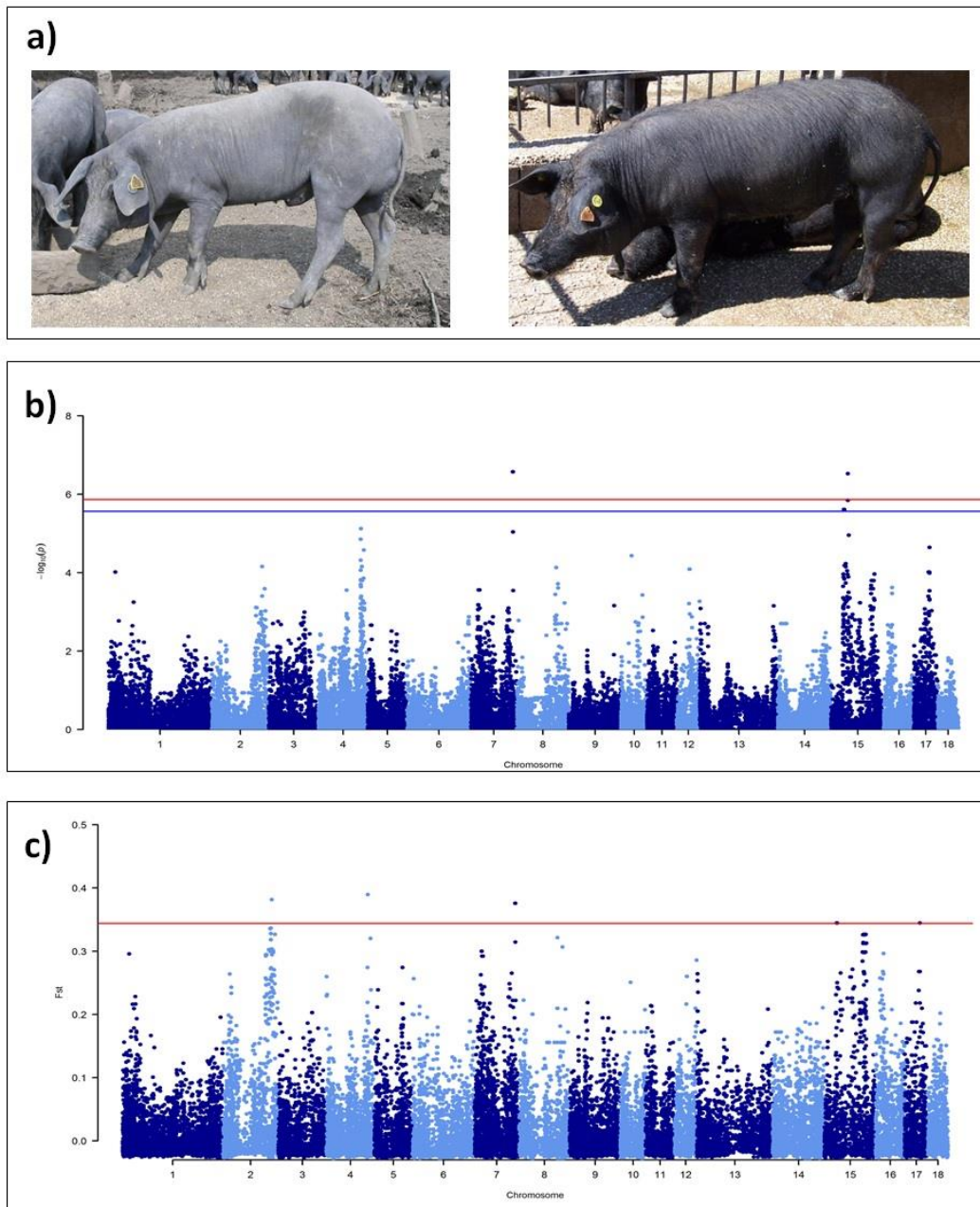
229 Porter V. (1993) *Pigs: A Handbook to the Breeds of the World*. Cornell University Press,

230 Ramot Y. & Zlotogorski A. (2015) Molecular genetics of alopecias. *Current Problems in*  
231 *Dermatology* **47**, 87-96.

232 Roberts E. & Carroll W.E. (1931) The inheritance of hairlessness in swine. *Journal of Heredity* **22**,  
233 125-32.

- 234 Samaan G., Yugo D., Rajagopalan S., *et al.* (2010) Foxn3 is essential for craniofacial development  
235 in mice and a putative candidate involved in human congenital craniofacial defects.  
236 *Biochemistry Biophysics Research Communications* **400**, 60-5.
- 237 Shimomura Y. (2012) Congenital hair loss disorders: rare, but not too rare. *Journal of Dermatology*  
238 **39**, 3-10.
- 239 Shimomura Y. & Christiano A.M. (2010) Biology and genetics of hair. *Annual Review of Genomics*  
240 *and Human Genetics* **11**, 109-32.
- 241 Stoye J.P., Fenner S., Greenoak G.E., Moran C. & Coffin J.M. (1988) Role of endogenous  
242 retroviruses as mutagens: the hairless mutation of mice. *Cell* **54**, 383-91.
- 243 Toro M.A., Rodriganez J., Silio L. & Rodriguez C. (2000) Genealogical analysis of a closed herd of  
244 black hairless Iberian pigs. *Conservation Biology* **14**, 1843-51.
- 245 Verhoeven K., De Jonghe P., Van de Putte T., *et al.* (2003) Slowed conduction and thin myelination  
246 of peripheral nerves associated with mutant rho Guanine-nucleotide exchange factor 10.  
247 *American Journal of Human Genetics* **3**, 926-32.
- 248 Zhou X. & Stephens M. (2012) Genome-wide efficient mixed-model analysis for association studies.  
249 *Nature Genetics* **44**, 821-4.

250 **Figure 1.** Casertana pigs and results of the genome wide association study (GWAS). a) Casertana  
251 pigs with the hairless (left) and haired (right) phenotypes. b) Manhattan plot of the GWAS results  
252 showing Bonferroni significant (red line:  $P < 0.05$ ) and suggestively significant (blue line:  $P < 0.10$ )  
253 single nucleotide polymorphisms (SNPs; thresholds are Bonferroni corrected  $P$  values). c)  $F_{ST}$  plot.  
254 Single nucleotide polymorphisms above the red line ( $F_{ST} = 0.345$ ) are the top 0.9998 SNPs.  
255



256

257

258 **Table 1.** List of significant ( $P < 0.05$ ) and suggestively significant ( $0.05 < P < 0.10$ ; Bonferroni corrected) single nucleotide polymorphisms (SNPs)  
259 obtained in the genome wide association study (GWAS) in the Casertana pigs (GEMMA) and the top 0.998 detected in the  $F_{ST}$  analysis. For the  
260 overlapping regions among the two approaches, annotated genes nearby the SNPs ( $\pm 500$  kb from the first to the last SNP of the region) were reported  
261 (Sscrofa11.1 genome version). The candidate genes that could be involved in the hair phenotype are indicated with the “\*” symbol.  $P$ ,  $F_{ST}$  and  
262 annotated genes are reported only for the SNPs and regions for which both  $P$  and  $F_{ST}$  values trespassed the indicated thresholds.

| SSC | SNP         | position  | GWAS, $P$ nominal value | $F_{ST}$ value | Annotated genes   |
|-----|-------------|-----------|-------------------------|----------------|---|
| 2   | ALGA0016212 | 134598604 | -                       | 0.381          | -   |
| 4   | INRA0016870 | 113277535 | -                       | 0.390          | -   |
| 7   | INRA0028322 | 111520662 | 2.68E-07                | 0.376          | <i>LOC106504536, PSMC1, EFCAB11, NRDE2, CALM1, TDPI, KCNK13, FOXN3*</i>                       |
| 7   | ALGA0044817 | 111690832 | 2.68E-07                | 0.376          |   |
| 15  | MARC0009352 | 33679138  | 2.45E-06                | 0.345          | <i>C110257074, CLN8, KBTBD11, DLGAP2, LOC106509653, ARHGEF10*, LOC106506202, CSMD1, MYOM2</i> |
| 15  | ALGA0084906 | 34793592  | 2.45E-06                | 0.345          |   |
| 15  | H3GA0044265 | 44006149  | 3.00E-07                | -              | -   |
| 15  | INRA0049225 | 44344760  | 1.43E-06                | -              | -   |
| 17  | DRGA0016747 | 41675886  | -                       | 0.345          | -   |
| 17  | H3GA0049027 | 41643251  | -                       | 0.345          | -   |

263  
264

265 **Supplementary Material**

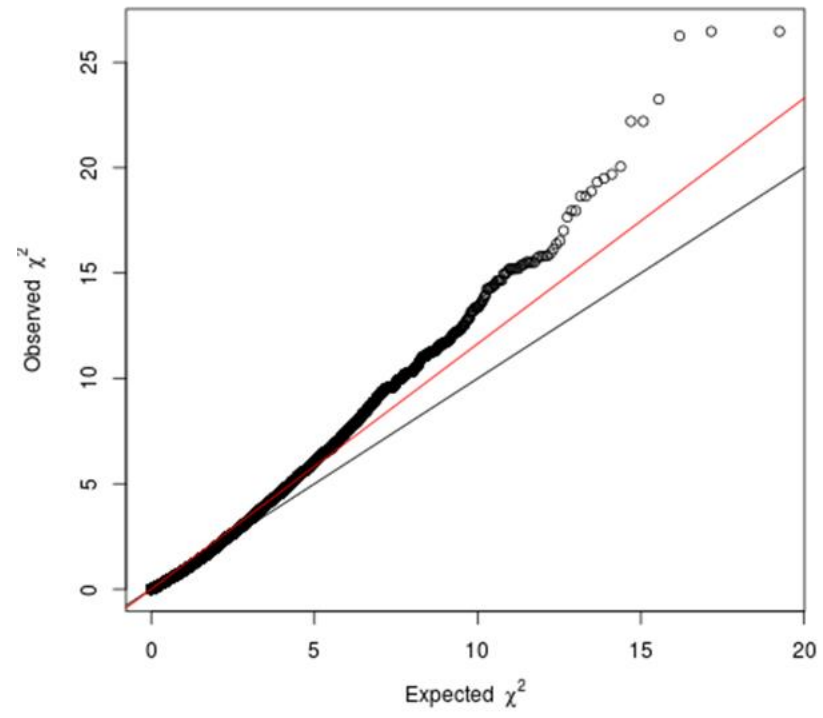
266

267 **Taking advantage from phenotype variability in a local animal genetic resource: identification of genomic regions associated with the**  
268 **hairless phenotype in Casertana pigs**

269

270 **Figure S1. Q-Q plot.**

$\lambda = 1.16$



271

272

273