Exploiting single-marker and haplotype-based genome-wide association studies to identify QTL for the number of teats in Italian Duroc pigs

# Exploiting single-marker and haplotype-based genome-wide association studies to identify QTL for the number of teats in Italian Duroc pigs

Samuele Bovo[1], Mohamad Ballan[1], Giuseppina Schiavo[1], Anisa Ribani[1], Silvia Tinarelli[1,2], Stefania Dall'Olio[1], Maurizio Gallo[2] and Luca Fontanesi[1]

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

[2] Associazione Nazionale Allevatori Suini (ANAS), Via Nizza 53, 00198 Roma, Italy

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

Short title: **Genome scans for the number of teats in Italian Duroc pigs**

## Abstract

The number of teats is a morphological trait of high economic relevance for the pig industry. Here, to dissect the genomic architecture of this trait in the Italian Duroc pig population, we present the results of genome-wide association studies in this Italian heavy pig breed. A total of 1,162 pigs, for which the number of teats was recorded, was genotyped with two high-throughput single nucleotide polymorphism (SNP) genotyping platforms (60K and 70K). Genome-wide association analyses were based on a single-marker approach and on a haplotype-based approach. Two quantitative trait loci (QTL) affecting the number of teats were identified. The most significant QTL, identified by the single-marker analysis and confirmed by the haplotype-based method, was located on *Sus scrofa* chromosome (SSC) 7, in the region of the *vertnin* (*VRTN)* gene. Suggestively associated markers (SNPs and haplotypes) were located on SSC10, in the region of the *FERM domain containing 4A* (*FRMD4A*) gene, the second identified QTL. These findings confirm previous results obtained in a few other Duroc populations. Overall, this study further supported the important role of variability in the *VRTN* gene region in affecting the number of teats in pigs. Moreover, the results also indicated that this trait in the Italian Duroc breed, as in many other pig breeds, is affected by few QTL, with the contribution of many other genetic factors with small effects, following the classical theory of quantitative traits.

**Keywords:** Breed: GWAS; heavy pig; single nucleotide polymorphism; *Sus scrofa*

## 1. Introduction

The number of teats is one of the most relevant morphological traits related to the reproductive performance of the sows. This trait is indirectly associated with the mothering ability, that, in turn, determines the number of piglets weaned per sow and per year (Kim et al., 2005; Andersen et al., 2011). To maximize the number of weaned piglets, selection programs, that have as main objective an increased litter size (a trait with low heritability), need to select, in parallel, for increased number of functional teats. Selection to improve the number of teats is, to some extent, facilitated by the medium to high heritability of the trait itself, as demonstrated by most studies in several pig breeds or lines that estimated the fraction of the genetic variance over the phenotypic variance for this morphological feature (e.g., Willham and Whatley 1962; Toro et al., 1986; McKay and Rahnefeld 1990; Borchers et al., 2002; Chalkias et al., 2013; Felleki and Lundeheim 2015; Balzani et al., 2016).

A broad variability in the number of teats (which could range from about 8 to 21) has been observed both across and within pig breeds and lines (Borchers et al., 2002; Lopes et al., 2014; Verardo et al., 2016; Rohrer and Nonneman 2017; Dall'Olio et al., 2018; van Son et al., 2019). Despite being represented by discrete and countable values, the number of teats is considered a quantitative trait determined by the effect of a high number of genetic factors. Several studies have however identified a few quantitative trait loci (QTL) affecting this morphological trait in pigs. The first QTL investigations were based on reference populations (F2 and backcrosses) constructed using parental animals of different breeds or lines (some of which with divergent number of teats), including some hyper-prolific Chinese breeds (e.g., Wada et al., 2000; Hirooka et al., 2001; Rodríguez et al., 2005; Bidanel et al., 2008; Ding et al., 2009; Hernandez et al., 2014). Subsequent studies based on high density single nucleotide polymorphism (SNP)

genotyping data explored the within breed variability via genome-wide association studies (GWAS). These studies confirmed the previously reported QTL regions and/or identified other novel genome regions affecting this trait (e.g., Arakawa et al., 2015; Rohrer and Nonneman 2017; Tang et al., 2017; Lee et al., 2019; van Son et al., 2019). One of the most relevant QTL for the number of teats, confirmed by several studies, is located on *Sus scrofa* chromosome (SSC) 7 (Mikawa et al., 2007, 2011; Duijvesteijn et al., 2014; Rohrer and Nonneman 2017; Dall'Olio et al., 2018; van Son et al., 2019; Moscatelli et al., 2020). At this QTL, alleles segregating in many populations are determined by variability in the *vertnin* gene, also known as *vertebrae development associate*d gene (*VRTN*; Mikawa et al., 2011; Arakawa et al., 2015). Originally reported to have pleiotropic effects on the number of vertebrae (Mikawa *et al.*, 2011), *VRTN* encodes a novel DNA-binding transcription factor which regulates the transcription of a set of genes that harbor VRTN binding motifs and modulates somite segmentation via the Notch signaling pathway (Duan et al., 2018). Some studies in a few pig populations proposed that variability in other genes on SSC7 [*latent transforming growth factor binding protein 2* (*LTBP2*); *BRMS1 like transcriptional repressor* (*BRMS1L*); and *Fos proto-oncogene, AP-1 transcription factor subunit* (*FOS*)], close to the *VRTN* gene, are involved in affecting the number of teats and the number of vertebrae (Zhang et al., 2016; Park et al., 2018; Liu et al., 2020).

Genome-wide association studies are usually run with a model fitting one marker at a time. However, to further extract information at the genome-wide level, the use of haplotypes has been recommended to capture additional marker-phenotype associations that cannot be detected with the single-marker approach (Lorenz et al., 2010; Barendse 2011). We recently applied haplotype-based analyses in a few GWAS in pigs and demonstrated that it is possible to improve the identification of genomic regions affecting

a targeted phenotype, complementing and completing the results obtained via a single-marker approach (Bovo et al., 2019, 2020, 2021). The additional information that can be retrieved is population-dependent, as it is related to the level of linkage disequilibrium between QTL alleles and the genotyped markers, which may vary among breeds, lines and populations. An ascertainment bias derived by the SNP composition of the commercial genotyping panels used in the investigation is also another element that should be considered in this context.

We recently carried out GWAS for the number of teats in two heavy pig breeds, Italian Large White and Italian Landrace, and identified several genome regions affecting this trait, not overlapping between the two breeds (Moscatelli et al., 2020; Bovo et al., 2021). Only in Italian Large White, the *VRTN* gene region was associated with this trait whereas in Italian Landrace no significant effects of *VRTN* (or other markers of SSC7) were evidenced (Bovo et al., 2021). For the other heavy pig breed, Italian Duroc, included in the Italian selection program conducted by the Italian Pig Breeders Association (ANAS), only few information on the segregation of the *VRTN* gene alleles was reported thus far (Fontanesi et al., 2014a). This Duroc population is a close nucleus constituted in the 1990' that is under selection with the main aim to optimize the production of green legs transformed in Protected Designation of Origin (PDO) dry cured hams. Other purebred Duroc pig populations (e.g., from USA, Canada, China, Japan and The Netherlands) have been already investigated with GWAS to identify genome regions affecting the number of teats. The investigations involving these populations reported partially overlapping results in terms of QTL regions (Arakawa et al., 2015; Yang et al., 2016; Tan et al., 2017; van Son et al., 2019; Zhuang et al., 2020; Li et al., 2021).

In this study, we carried out two GWAS for the number of teats in Italian Duroc heavy pigs using single-marker and haplotype-based approaches. Obtained results were

compared with the QTL information for the same trait reported in other pure-bred Duroc populations and in the other two Italian heavy pig breeds.

## 2. Materials and methods

### 2.1. Animals

All animals used in this study were kept according to Italian and European legislation for pig production and all procedures described were in compliance with national and European Union regulations for animal care and slaughtering. All animals were part of the routine Italian pig breeding program and were slaughtered in a commercial authorized abattoir following standard procedures. Animals were not raised or sampled for the purpose of this study. As no treatment was given to any animals, no ethical approval was needed according to the rules of the animal research ethics committee of the University of Bologna based on the Italian legislation, as reported in the "DECRETO LEGISLATIVO 4 marzo 2014, n. 26".

The animals included in this study were from the national selection program of heavy pig breeds that is run by ANAS. A total of 1,164 Italian Duroc pigs, born in the years 1996-2018, was investigated (727 males and 437 females). Animals were included in the sib-testing evaluation program run by ANAS, described in previous reports (Fontanesi et al., 2012, 2014b, 2015). The number of teats on these animals was routinely recorded by direct counting at the beginning of the testing period and animals having less than 12 teats were excluded from the herd book of this breed.

### 2.2. Genotyping and SNP quality

DNA was extracted by using the Wizard Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA) from blood samples routinely collected on all pigs

included in the national selection program. A total of 606 animals were genotyped with the GeneSeek 70K GGP Porcine BeadChip (which includes 68,516 SNPs) whereas the remaining 558 pigs were genotyped with the Illumina PorcineSNP60 BeadChip v.2 (which includes 61,565 SNPs). The two panels presented a total of 41,862 shared genomic positions. Genotyping followed standard procedures based on the supplier's recommendations. DNA markers shared between the two SNP platforms were used in the analysis. BLAST+ v.2.7.1 (Camacho et al., 2009) was used to map SNPs to the Sscrofa11.1 reference genome and markers assigned to more than one position or assigned to sex chromosomes were discarded. Genotypes were evaluated with PLINK v.1.09 (Chang et al., 2015). Individual pigs with a call rate >0.90 and SNPs with a call rate > 0.90, a minor allele frequency (MAF) > 0.01 and in Hardy-Weinberg equilibrium (P > 0.0001) were retained for further analyses. The final dataset counted 1,162 animals (727 males and 435 females) and 29,604 DNA markers.

### 2.3. Haplotype estimation

The software SHAPEIT v.2 (Delaneau et al., 2011) was used for genotype phasing considering: (i) a genomic window size of 2 Mb, (ii) an effective population size (Ne) estimated with SNeP v.1.1 (Barbato et al., 2015) and (iii) a chromosome specific recombination rate given by Tortereau et al., (2012). The R package GHap 1.2.2 (Utsunomiya et al., 2016) was further used to call haplotypes considering a genomic window of 400 kb with a sliding block of 100 kb (Veroneze et al., 2013). Each haplotype was treated as bi-allelic DNA variant (genotypes are: NN, NH and HH; H = haplotype allele and N = NULL = all other N alleles present in the haploblock. PLINK was used to filter out haplotypes having a MAF < 0.02. A total of 85590 haplotypes coming from 66,363 overlapping haploblocks were estimated.

164

165 **2.4. Linear mixed model analyses**

166 To assess the association between the DNA markers (SNP or haplotypes) and the
167 number of teats we used a linear mixed effect model. An additive genetic model assuming
168 a trend per copy of the minor allele that specify the dependency of the number of teats on
169 genotype categories was implemented as follows:

170 $$\boldsymbol{y} = \boldsymbol{W\alpha} + \boldsymbol{x}\beta + \boldsymbol{g} + \boldsymbol{e} \qquad (1)$$

171 where $\boldsymbol{y}$ ($n \times 1$) is a vector containing the phenotype (the number of teats) for the $n^{th}$
172 animal, $\boldsymbol{W}$ ($n \times k$) is a covariate matrix with $k = 2$ (a column of 1s and a column for the
173 sex of the animal) and $\boldsymbol{\alpha}$ is the $k$-dimensional vector of covariates effects, $\boldsymbol{x}$ ($n \times 1$) is the
174 vector containing genotypes for the $i^{th}$ DNA marker (SNP or haplotype), $\beta$ is the additive
175 fixed effect of the $i^{th}$ DNA marker on the phenotype, $\mathbf{g} \sim N(\mathbf{0}, \sigma^2_g \mathbf{K})$ is a multivariate
176 Gaussian polygenic effect, with covariance matrix proportional to the relatedness matrix
177 $\mathbf{K}$ ($n \times n$) and $\mathbf{e} \sim N(\mathbf{0}, \sigma^2_e \mathbf{I})$ is a multivariate Gaussian vector of uncorrelated residuals.
178 Polygenic effect and residuals represented the random effects. The assessment of the
179 association between each DNA marker and the total number of teats was obtained by
180 testing the null hypothesis $H_0: \beta = 0$. Significance was tested by using the Wald test. All
181 the models were fitted with GEMMA v.0.98 (Zhou and Stephens 2012) after computing
182 the relatedness matrices $\mathbf{K_1}$ and $\mathbf{K_2}$ as a centered genomic matrices, for SNPs and
183 haplotypes, respectively. Details about the genomic matrix construction are given in the
184 manual of GEMMA. Bonferroni correction was applied to account for multiple
185 comparisons, considering the total number of DNA markers used in each genome scan
186 and value of $\alpha = 0.05$. Markers presenting a $P < 5.5 \times 10^{-05}$ were considered suggestive for
187 associations (Wellcome Trust Case Control Consortium 2007). The proportion of

188 variance in phenotype explained (PVE) by each QTL was calculated as reported by Shim
189 et al., 2015.
190 SNPs and haplotypes that had the lowest *P* in chromosome regions separated by
191 at least 5 Mb were considered as tag DNA markers. GEMMA was also used to estimate
192 the genomic (chip) heritability ($h^2_G$). Genomic control inflation factor ($\lambda_{GC}$) was computed
193 in R v.3.6.0 (R Core Team 2018). Quantile-quantile plots (QQplots) and the Miami plot
194 were generated in R by using the *qqman* package (Turner 2018).
195
196 **2.5. Haploblock analysis and genome annotation**
197 HaploView v.4.2 (Barrett et al., 2011) was used to study the structure of the
198 haplotypes and linkage disequilibrium of the most interesting QTL regions. Each QTL
199 was annotated considering the protein coding genes spanning the region of ± 500 kb
200 around the evaluated DNA marker. The Sscrofa11.1 NCBI's GFF file and Bedtools
201 v.2.17.0 (Quinlan and Hall 2010) were used for this purpose.
202
203 **3. Results**
204 **3.1. Descriptive statistics**
205 The numbers of teats in this pig population ranged from 12 to 17. About 42% of the
206 pigs had 12 teats (Figure 1) as this represents the lower limit considered for including an
207 animal in the Herd Book of this Italian heavy pig breed (ANAS 2021). Detailed
208 information on the number of pigs with different numbers of teats are reported in Table
209 S1. On average, pigs had 12.95 ± 0.97 (mean ± standard deviation) teats.
210 Table 1 reports the number of pigs and DNA markers used in GWAS. Assessment
211 of results via QQplots (Figure S1) and the genomic control inflation factors (Table 1)
212 indicated a good control of the population stratification.

Genomic heritability estimated for the number of teats in Italian Duroc pigs was moderate, with the value that increased from the estimates based on single-marker to the value obtained with the haplotype information (Table 1). This suggested that the haplotype analysis might have captured an additional fraction of heritability, as already reported for the same trait in other pig populations (Bovo et al., 2020, 2021).

### 3.2. Single-marker and haplotype-based association studies

The top associated DNA markers (SNPs and haplotypes) for each QTL region are reported in Table 2, whereas Table S2 lists all the significant and suggestively associated markers. Miami plot including SNP and haplotype data is showed in Figure 2.

Three SNPs and one haplotype on SSC7 were significantly associated with the number of teats. The most significant SNP was MARC0038565 (rs80894106; $P = 8.58\times10^{-10}$), located on SSC7 at position 97,652,632 bp, in the region of the *VRTN* gene (located on SSC7:97.61-97.62). This marker had a MAF = 0.44 (allele A) and a negative regression coefficient ($\beta$ = -0.29), indicating a decrease in number of teats while increasing the number of copies of the minor allele. Hence, the Italian Duroc population had a higher frequency of the favorable allele (allele G) at this SNP position. This marker had also the highest proportion of variance in phenotype explained (PVE% = 3.22). The other two significant markers were closely located to the mentioned SNP and had a medium-high linkage disequilibrium ($r^2$ = 0.64, Figure 3) with the most significant marker. In general, this region did not show a very high level of linkage disequilibrium spanning the genotyped markers (Figure 3).

The peak of association on SSC7 was also highlighted by the haplotype analysis (SSC7:97.6-97.8 Mb; haplotype GAG; $P = 2.51\times10^{-8}$), confirming the involvement of the *VRTN* gene region in affecting the number of teats in the Italian Duroc pig breed. The

DNA marker MARC0038565 was included in this haplotype [first position in the haplotype sequence AAG; Figure 3]. Hence, the haplotype GAG [higher frequent in the population with f(GAG) = 0.52; favorable allele] and marker MARC0038565 presented similar MAF and $\beta$ values.

A suggestively associated chromosome region identified both using the single-marker and the haplotype-based analyses was detected in the region of the *FERM domain containing 4A* (*FRMD4A*) gene located on SSC10:47.3-48.1 Mb. The results for the most significant SNP in this region, located at position 47,947,214 bp (ASGA0090802; rs81309209; $P = 4.80\times10^{-06}$), were confirmed by the haplotype analysis ($P = 5.08\times10^{-06}$). Both regions (SSC7 and SSC10) overlapped with the two QTL regions already reported in a previous study to affect the number of teats in Duroc pigs (van Son et al., 2019).

## 4. Discussion

Duroc pig populations are well known to have, in general, lower reproduction performances than other cosmopolitan pig breeds, as they are usually selected to maximize production performances and meat quality traits for the use of Duroc boars in terminal crosses (e.g., Cameron 1990; Gaugler et al., 1984; Skorupski et al., 1996; Hoque et al., 2007: Alam et al., 2021). However, to maintain a sustainable breeding in the selection nuclei of this breed, it is also needed to improve reproduction efficiency starting from traits with medium-high heritability.

A lower number of teats has been reported in Duroc populations than in the other two main cosmopolitan breeds, Large White and Landrace, which usually exceed the threshold of 14 teats. As such, 14 is the lower limit that is applied by ANAS to register animals to the Italian Large White and Italian Landrace herd books. In a recent survey that we carried out, the mean number of teats in these two breeds was 14.88 ± 0.92

(standard deviation) and 14.77 ± 0.87, respectively (Bovo et al., 2021), that are larger than what we observed in Italian Duroc pigs (12.95 ± 0.97). This number is however larger than what was reported in other Duroc populations. For example, Zhuang et al., (2020) reported the mean number of 10.90 ± 1.16 and 10.92 ± 1.14 in Duroc populations derived from US and Canadian nuclei and Tan et al., (2017) reported an average value of 10.72 ± 1.72 in a Chinese Duroc population. It is worth to mention that our statistics are biased as, in Italy, pigs with less than 12 teats are not registered in the Italian Duroc herd book and are therefore not considered for the calculation of the population mean in our study. Similar mean values to what we reported in the Italian Duroc breed were reported in a Dutch Duroc line (12.93 ± 1.05; van Son et al., 2019) and in another Chinese Duroc herd (13.17 ± 1.12; Li et al., 2021). These results suggest that some heterogeneity exists in Duroc populations, probably due to different selection pressures that have been applied on this trait in nuclei of this cosmopolitan breed around the world. Genomic heritability that we estimated in the Italian Duroc breed was close to what we already reported in Italian Landrace pigs using both single-marker and haplotype-based approaches ($h^2_G$ ± standard error: 0.25 ± 0.02 and 0.30 ± 0.03, respectively) and a little lower than what we estimated in Italian Large White pigs (0.30 ± 0.03 and 0.43 ± 0.04, respectively; Bovo et al., 2021). Genomic heritability estimates for the number of teats were also reported in other Duroc populations. The single-marker heritability in an US nucleus was 0.19 ± 0.02, in a Canadian population was 0.34 ± 0.03 (Zhuang et al., 2020) and in a Chinese nucleus was 0.29 ± 0.05 (Li et al., 2021). These other studies, however, did not report any estimation based on haplotypes. Genomic heritability estimated using haplotype information indicated that this approach can potentially capture additional fractions of the so called “missed heritability” that would not be disclosed using single-marker approaches (Bovo et al., 2021). Therefore, the use of haplotypes might be useful to further

exploit genomic information for this purpose (Lorenz et al., 2010; Barendse 2011).

Only two QTL for the number of teats, one highlighted by significant markers and one identified by suggestively significant markers, were identified in the Italian Duroc population. Results obtained by the single-marker analysis were confirmed by the haplotype-based analysis. To continue the comparison with the results obtained in other Duroc populations, it is worth to mention that a limited number of segregating QTL for this trait has been also reported in other Duroc populations (Tan et al., 2017; van Son et al., 2019; Zhuang et al., 2020; Li et al., 2021). For example, by investigating a Dutch Duroc population, van Son et al., (2019) reported only the same two QTL regions that we identified in the Italian Duroc pigs. The larger number of pigs that was investigated in the Dutch study made it possible to reach significant results even for markers in the region of the *FRMD4A* gene that, however, remained less significantly associated than what was reported for the SNPs on SSC7, located in the region of *VRTN*. The *VRTN* region also included the most significant SNPs identified in GWAS carried out in Canadian Duroc pigs and in two Chinese Duroc populations (Tan et al., 2017; Zhuang et al., 2020; Li et al., 2021). No significant QTL for the number of teats were however identified in an American derived nucleus of Duroc pigs (Zhuang et al., 2020). Comparison with the results obtained in a Japanese Duroc population is not possible due to the different statistical approaches that were used to detect QTL regions for the number of teats (BayesC methods; Arakawa et al., 2015). The *VRTN* region, however, emerged as one of the most significant regions also in this population (Arakawa et al., 2015).

The proportion of variance in phenotype explained (PVE) by the most significant marker in the *VRTN* region that we identified in the Italian Duroc population was about half the PVE of the most significant marker of this region identified in a Dutch Duroc line (van Son et al., 2019): 3.2% *vs* 6.0%. This difference might be due to different

313 levels of linkage disequilibrium between the markers identified in these two Duroc
314 populations (the Italian and the Dutch populations) and the causative mutation(s) and to
315 the different experimental power determined by the number of animals analysed in our
316 study and by van Son et al. (2019). Markers in this region have a much higher PVE for
317 vertebral numbers are directly related phenotypes (van Son et al., 2019), which are
318 probably the primary traits affected by *VRTN* variability (Mikawa et al., 2011), with
319 pleiotropic effects on teat number.

320 A QTL pattern similar to what was identified in Italian Duroc pigs was already
321 reported in Italian Large White pigs where, in addition to the *VRTN* and *FRMD4A* gene
322 regions, only another significant haplotype region was identified on SSC12 (Bovo et al.,
323 2021). The Italian Landrace population was however completely different, in terms of
324 QTL regions for the number of teats, from both the Italian Duroc and Italian Large White
325 breeds (Bovo et al., 2021).

326 These comparative analyses, in general, showed that one major QTL for the
327 number of teats segregates in the most important cosmopolitan pig breeds and derived
328 lines and populations (like the Italian Duroc breed). This QTL is located on SSC7 for
329 which the most plausible causative mutation has been reported in the *VRTN* gene
330 (Mikawa et al., 2011). In Italian Duroc pigs, the frequency of the insertion allele at the
331 *VRTN* gene suggested to be the putative favorable allele (allele Q; Fontanesi et al., 2014a),
332 which might increase the number of vertebrae and the number of teats (Mikawa et al.,
333 2011), is almost equal to the frequency of the positive allele of the most significant SNP
334 that we identified on SSC7 (allele G of MARC0038565). This information could
335 indirectly suggest a very high linkage disequilibrium between the two markers.
336 Additional studies are needed to complete the linkage disequilibrium pattern of this SSC7
337 genome region in Italian Duroc, including genotyping data of the *VRTN* gene which might

confirm the significant results obtained for the MARC0038565 SNP. Additional studies are also needed to characterize the variability in the *FRMD4A* gene and identify the causative mutation(s) affecting this trait. This gene encodes a FERM domain-containing protein that regulates epithelial polarity by connecting ADP ribosylation factor 6 (ARF6), which is considered an important factor involved in actin cytoskeleton dynamics and membrane trafficking (Ikenouchi and Umeda 2010). Therefore, the role of this gene in affecting the number of teats is not obvious. Thus, it will be also important to clarify the molecular mechanisms that could lead to this effect and eventually evaluate the effect of other close genes in determining the QTL located on SSC10.

## 4. Conclusion

In terms of genotyped animals, this is the largest GWAS carried out in the Italian Duroc pig breed thus far. The estimated genomic heritability that we obtained confirmed that genetic factors affect this trait in a similar way as it was evidenced in other Duroc populations. The limited number of major QTL influencing this trait and segregating in the Italian Duroc breed indicates that a marker assisted selection based on *VRTN* alleles and eventually on *FRMD4A* gene markers, might be effective only in part as it would not completely exploit the genetic variance related to this trait for which many other genes are involved. Their effect could be captured including this trait in a genomic selection program that can take into account genome-wide variability.

## Acknowledgements

The authors acknowledge the help of personnel of ANAS for the collaboration in this study.

## Funding

This study received funds from the MIUR PRIN2017 PigPhenomics project and by the University of Bologna RFO 2019-2021 programmes. This study was also supported by the PSRN (Progetto di Sviluppo Rurale Nazionale) SUIS and SUIS2.0 projects (co-funded by the European Agricultural Fund for Rural Development of the European Union and by the Italian Ministry of Agriculture, Food and Forestry - MiPAAF).

## Conflict of interest statement

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

## References

Alam, M., Chang, H-K., Lee, S-S., Choi, T-J., 2021. Genetic analysis of major production and reproduction traits of Korean Duroc, Landrace and Yorkshire pigs. Animals 11, 1321. https://doi.org/10.3390/ani11051321

ANAS, 2019. Libro Genealogico. Retrieved on the 2nd September 2019. from http://www.anas.it/.

Andersen, I.L., Nævdal, E., Bøe, K.E., 2011. Maternal investment, sibling competition, and offspring survival with increasing litter size and parity in pigs (*Sus scrofa*). Behav. Ecol. Sociobiol. 65, 1159-1167. https://doi.org/10.1007/s00265-010-1128-4

Arakawa, A., Okumura, N., Taniguchi, M., Hayashi, T., Hirose, K., Fukawa, K., Ito, T., Matsumoto, T., Uenishi, H., Mikawa, S., 2015. Genome-wide association QTL mapping for teat number in a purebred population of Duroc pigs. Anim. Genet. 46, 571-575. https://doi.org/10.1111/age.12331

388 Balzani, A., Cordell, H.J., Sutcliffe, E., Edwards, S.A., 2016. Heritability of udder
389 morphology and colostrum quality traits in swine. J. Anim. Sci. 94, 3636-3644.
390 https://doi.org/10.2527/jas.2016-0458
391 Barbato, M., Orozco-terWengel, P., Tapio, M., Bruford, M.W., 2015. SNeP: a tool to
392 estimate trends in recent effective population size trajectories using genome-wide
393 SNP data. Front. Genet. 6, 109. https://doi.org/10.3389/fgene.2015.00109
394 Barendse, W., 2011. Haplotype analysis improved evidence for candidate genes for
395 intramuscular fat percentage from a genome wide association study of cattle. PLoS
396 One 6, e29601. https://doi.org/10.1371/journal.pone.0029601
397 Barrett, J.C., Fry, B., Maller, J., Daly, M.J., 2005. Haploview: analysis and visualization
398 of LD and haplotype maps. Bioinformatics 21, 263-265.
399 https://doi.org/10.1093/bioinformatics/bth457
400 Bidanel, J.P., Rosendo, A., Iannuccelli, N., Riquet, J., Gilbert, H., Caritez, J.C., Billon,
401 Y., Amigues, Y., Prunier, A., Milan, D., 2008. Detection of quantitative trait loci
402 for teat number and female reproductive traits in Meishan × Large White F2 pigs.
403 Animal 2, 813-820. https://doi.org/10.1017/S1751731108002097
404 Borchers, N., Reinsch, N., Kalm, E., 2002. Teat number, hairiness and set of ears in a
405 Piétrain cross: variation and effects on performance traits. Arch. Anim. Breed. 45,
406 465-480. https://doi.org/10.5194/AAB-45-465-2002
407 Bovo, S., Ballan, M., Schiavo, G., Gallo, M., Dall'Olio, S., Fontanesi, L., 2020.
408 Haplotype-based genome-wide association studies reveal new loci for
409 haematological and clinical-biochemical parameters in Large White pigs. Anim.
410 Genet. 51, 601-606. https://doi.org/10.1111/age.12959
411 Bovo, S., Ballan, M., Schiavo, G., Ribani, A., Tinarelli, S., Utzeri, V.J., Dall'Olio, S.,
412 Gallo, M., Fontanesi, L., 2021. Single-marker and haplotype-based genome-wide

association studies for the number of teats in two heavy pig breeds. Anim. Genet. 52, 440-450. https://doi.org/10.1111/age.13095

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

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Cameron, N.D., 1990. Comparison of Duroc and British Landrace pigs and the estimation of genetic and phenotypic parameters for growth and carcass traits. Anim. Sci. 50, 141-153. doi:10.1017/S0003356100004542

Chalkias, H., Rydhmer, L., Lundeheim, N., 2013. Genetic analysis of functional and non-functional teats in a population of Yorkshire pigs. Livest. Sci. 152, 127-134. https://doi.org/10.1016/j.livsci.2013.01.003

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

Dall'Olio, S., Ribani, A., Moscatelli, G., Zambonelli, P., Gallo, M., Nanni Costa, L., Fontanesi, L., 2018. Teat number parameters in Italian Large White pigs: phenotypic analysis and association with *vertnin* (*VRTN*) gene allele variants. Livest. Sci. 210, 68-72. https://doi.org/10.1016/j.livsci.2018.01.020

436 Delaneau, O., Marchini, J., Zagury, J-F., 2011. A linear complexity phasing method for
437 thousands of genomes. Nat. Methods 9, 179-181.
438 https://doi.org/10.1038/nmeth.1785
439 Ding, N., Guo, Y., Knorr, C., Ma, J., Mao, H., Lan, L., Xiao, S., Ai, H., Haley, C.S.,
440 Brenig, B., Huang, L., 2009. Genome-wide QTL mapping for three traits related to
441 teat number in a White Duroc x Erhualian pig resource population. BMC Genet.
442 10, 6. https://doi.org/10.1186/1471-2156-10-6
443 Duan, Y., Zhang, H., Zhang, Z., Gao, J., Yang, J., Wu, Z., Fan, Y., Xing, Y., Li, L., Xiao,
444 S., Hou, T., Ren, J., Huang, L., 2018. *VRTN* is required for the development of
445 thoracic vertebrae in mammals. Int. J. Biol. Sci. 14, 667-681.
446 https://doi.org/10.7150/ijbs.23815
447 Duijvesteijn, N., Veltmaat, J.M., Knol, E.F., Harlizius, B., 2014. High-resolution
448 association mapping of number of teats in pigs reveals regions controlling vertebral
449 development. BMC Genomics 15, 542. https://doi.org/10.1186/1471-2164-15-542
450 Felleki, M., Lundeheim, N., 2015. Genetic heteroscedasticity of teat count in pigs. J.
451 Anim. Breed. Genet. 132, 392-398. https://doi.org/10.1111/jbg.12134
452 Fontanesi, L., Schiavo, G., Galimberti, G., Calò, D.G., Scotti, E., Martelli, P.L.,
453 Buttazzoni, L., Casadio, R., Russo, V., 2012. A genome wide association study for
454 backfat thickness in Italian Large White pigs highlights new regions affecting fat
455 deposition including neuronal genes. BMC Genomics. 13, 583.
456 https://doi.org/10.1186/1471-2164-13-583
457 Fontanesi, L., Scotti, E., Buttazzoni, L., Dall'Olio, S., Russo, V., 2014a. Investigation of
458 a Short Interspersed Nuclear Element polymorphic site in the porcine *vertnin* gene:
459 allele frequencies and association study with meat quality, carcass and production

traits in Italian Large White pigs. Ital. J Anim. Sci. 13, 3090. https://doi.org/10.4081/ijas.2014.3090

Fontanesi, L., Schiavo, G., Galimberti, G., Calò, D.G., Russo, V., 2014b. A genomewide association study for average daily gain in Italian Large White pigs. J. Anim. Sci. 92, 1385-1394. https://doi.org/10.2527/jas.2013-7059

Fontanesi, L., Schiavo, G., Scotti, E., Galimberti, G., Calò, D.G., Samorè, A.B., Gallo, M., Russo, V., Buttazzoni, L., 2015. A retrospective analysis of allele frequency changes of major genes during 20 years of selection in the Italian Large White pig breed. J. Anim. Breed. Genet. 132, 239-246. https://doi.org/10.1111/jbg.12127

Gaugler, H.R., Buchanan, D.S., Hintz, R.L., Johnson, R.K. 1984., Sow productivity comparisons for four breeds of swine: purebred and crossbred litters. J. Anim. Sci. 59, 941-947. https://doi.org/10.2527/jas1984.594941x

Hernandez, S.C., Finlayson, H.A., Ashworth, C.J., Haley, C.S., Archibald, A.L., 2014. A genome-wide linkage analysis for reproductive traits in F2 Large White × Meishan cross gilts. Anim. Genet. 45, 191–197. https://doi.org/10.1111/age.12123

Hirooka, H., de Koning, DJ., Harlizius, B., van Arendonk, JA., Rattink, A.P., Groenen, M.A., Brascamp, E.W., Bovenhuis, H., 2001. A whole-genome scan for quantitative trait loci affecting teat number in pigs. J. Anim. Sci. 79, 2320-2326. https://doi.org/10.2527/2001.7992320x

Hoque, M.A., Suzuki, K., Kadowaki, H., Shibata, T., Oikawa, T., 2007. Genetic parameters for feed efficiency traits and their relationships with growth and carcass traits in Duroc pigs. J. Anim. Breed. Genet. 124, 108-116. https://doi.org/10.1111/j.1439-0388.2007.00650.x

Ikenouchi, J., Umeda, M., 2010. *FRMD4A* regulates epithelial polarity by connecting Arf6 activation with the PAR complex. Proc. Natl. Acad. Sci. USA 107, 748-753. https://doi.org/10.1073/pnas.0908423107

Kim, J.S., Jin, D.I., Lee, J.H., Son, D.S., Lee, S.H., Yi, Y.J., Park, C.S., 2005. Effects of teat number on litter size in gilts. Anim. Reprod. Sci. 90, 111-116. https://doi.org/10.1016/j.anireprosci.2005.01.013

Lee, J., Lee, S., Park, J-E., Moon, S-H., Choi, S-W., Go, G-W., Lim, D., Kim, J-M., 2019. Genome-wide association study and genomic predictions for exterior traits in Yorkshire pigs. J. Anim. Sci. 97, 2793-2802. https://doi.org/ 10.1093/jas/skz158

Li, Y., Pu, L., Shi, L., Gao, H., Zhang, P., Wang, L., Zhao, F., 2021. Revealing new candidate genes for teat number relevant traits in Duroc pigs using genome-wide association studies. Animals 11, 806. https://doi.org/ 10.3390/ani11030806

Liu, Q., Yue, J., Niu, N., Liu, X., Yan, H., Zhao, F., Hou, X., Gao, H., Shi, L., Wang, L., Wang, L., Zhang, L., 2020. Genome-wide association analysis identified *BMPR1A* as a novel candidate gene affecting the number of thoracic vertebrae in a Large White × Minzhu intercross pig population. Animals 10, E2186. https://doi.org/10.3390/ani10112186

Lopes, M.S., Bastiaansen, J.W.M., Harlizius, B., Knol, E.F., Bovenhuis, H., 2014. A Genome-wide association study reveals dominance effects on number of teats in pigs. PLoS One 9, e105867. https://doi.org/10.1371/journal.pone.0105867

Lorenz, A.J., Hamblin, M.T., Jannink, J-L., 2010. Performance of single nucleotide polymorphisms versus haplotypes for genome-wide association analysis in barley. PLoS One 5, e14079. https://doi.org/10.1371/journal.pone.0014079

McKay, R.M., Rahnefeld, G.W., 1990. Heritability of teat number in swine. Can. J. Anim. Sci. 70, 425-430. https://doi.org/10.4141/cjas90-054

508 Mikawa, S., Morozumi, T., Shimanuki, S-I., Hayashi, T., Uenishi, H., Domukai, M.,
509 Okumura, N., Awata, T., 2007. Fine mapping of a swine quantitative trait locus for
510 number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear
511 factor (*NR6A1*). Genome Res. 17, 586-593. https://doi.org/10.1101/gr.6085507
512 Mikawa, S., Sato, S., Nii, M., Morozumi, T., Yoshioka, G., Imaeda, N., Yamaguchi, T.,
513 Hayashi, T., Awata, T., 2011. Identification of a second gene associated with
514 variation in vertebral number in domestic pigs. BMC Genet. 12, 5.
515 https://doi.org/10.1186/1471-2156-12-5
516 Moscatelli, G., Dall'Olio, S., Bovo, S., Schiavo, G., Kazemi, H., Ribani, A., Zambonelli,
517 P., Tinarelli, S., Gallo, M., Bertolini, F., Fontanesi, L., 2020. Genome-wide
518 association studies for the number of teats and teat asymmetry patterns in Large
519 White pigs. Anim. Genet. 51, 595-600. https://doi.org/ 10.1111/age.12947
520 Park, H-B., Lee, J-B., Han, S-H., Kim, S-G., Shin, H-S., Shin, M-C., Kang, Y-J., Cho, I-
521 C., 2018. Joint linkage and linkage disequilibrium mapping reveals association of
522 *BRMS1L* with total teat number in a large intercross between Landrace and Korean
523 native pigs. Livest. Sci. 218, 2–96. https://doi.org/ 10.1016/j.livsci.2018.09.014
524 Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing
525 genomic features. Bioinformatics 26, 841-842.
526 https://doi.org/10.1093/bioinformatics/btq033
527 R Core Team. 2018. R: A language and environment for statistical computing. R
528 Foundation for Statistical Computing, Vienna, Austria.
529 Rodríguez, C., Tomás, A., Alves, E., Ramirez, O., Arqué, M., Muñoz, G., Barragán, C.,
530 Varona, L., Silió, L., Amills, M., Noguera, J.L., 2005. QTL mapping for teat
531 number in an Iberian-by-Meishan pig intercross. Anim. Genet. 36, 490-496.
532 https://doi.org/10.1111/j.1365-2052.2005.01358.x

Rohrer, G.A., Nonneman, D.J., 2017. Genetic analysis of teat number in pigs reveals some developmental pathways independent of vertebra number and several loci which only affect a specific side. Genet. Sel. Evol. 49, 4. https://doi.org/10.1186/s12711-016-0282-1

Shim, H., Chasman, D.I., Smith, J.D., Mora, S., Ridker, P.M., Nickerson, D.A., Krauss, R.M., Stephens, M., 2015. A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. PLoS One 10, e0120758. https://doi.org/10.1371/journal.pone.0120758

Skorupski, M.T., Garrick, D.J., Blair, H.T., 1996. Estimates of genetic parameters for production and reproduction traits in three breeds of pigs. New Zealand J. Agric. Res. 39, 387-395. https://doi.org/10.1080/00288233.1996.9513198

Tan, C., Wu, Z., Ren, J., Huang, Z., Liu, D., He, X., Prakapenka, D., Zhang, R., Li, N., Da, Y., Hu, X., 2017. Genome-wide association study and accuracy of genomic prediction for teat number in Duroc pigs using genotyping-by-sequencing. Genet. Sel. Evol. 49, 35. https://doi.org/10.1186/s12711-017-0311-8

Tang, J., Zhang, Z., Yang, B., Guo, Y., Ai, H., Long, Y., Su, Y., Cui, L., Zhou, L., Wang, X., Zhang, H., Wang, C., Ren, J., Huang, K., Ding, N., 2017. Identification of loci affecting teat number by genome-wide association studies on three pig populations. Asian-Australas J. Anim. Sci. 30, 1-7. https://doi.org/10.5713/ajas.15.0980

Toro, M., Dobao, M.T., Rodrigáñez, J., Silio, L., 1986. Heritability of a canalized trait: teat number in Iberian pigs. Genet. Sel. Evol. 18, 173-184. https://doi.org/10.1186/1297-9686-18-2-173

Tortereau, F., Servin, B., Frantz, L., Megens, H-J., Milan, D., Rohrer, G., Wiedmann, R., Beever, J., Archibald, AL., Schook, L.B., Groenen, M.A., 2012. A high density recombination map of the pig reveals a correlation between sex-specific

558 recombination and GC content. BMC Genomics 13, 586.
559 https://doi.org/10.1186/1471-2164-13-586
560 Turner, S., 2018. qqman: an R package for visualizing GWAS results using Q-Q and
561 manhattan plots. J. Open Source Softw. 3, 731. https://doi.org/ 10.21105/joss.00731
562 Utsunomiya, Y.T., Milanesi, M., Utsunomiya, A.T.H., Ajmone-Marsan, P., Garcia, J.F.,
563 2016. GHap: an R package for genome-wide haplotyping. Bioinformatics 32, 2861-
564 2862. https://doi.org/10.21105/joss.00731
565 van Son, M., Lopes, M.S., Martell, H.J., Derks, M.F.L., Gangsei, L.E., Kongsro, J., Wass,
566 M.N., Grindflek, E.H., Harlizius, B., 2019. A QTL for number of teats shows breed
567 specific effects on number of vertebrae in pigs: bridging the gap between molecular
568 and quantitative genetics. Front. Genet. 10, 272.
569 https://doi.org/10.3389/fgene.2019.00272
570 Verardo, L.L., Silva, F.F., Lopes, M.S., Madsen, O., Bastiaansen, J.W.M., Knol, E.F.,
571 Kelly, M., Varona, L., Lopes, P.S., Guimarães, S.E.F., 2016. Revealing new
572 candidate genes for reproductive traits in pigs: combining Bayesian GWAS and
573 functional pathways. Genet. Sel. Evol. 48, 9. https://doi.org/10.1186/s12711-016-
574 0189-x
575 Veroneze, R., Lopes, P.S., Guimarães, S.E.F., Silva, F.F., Lopes, M.S., Harlizius., B.,
576 Knol, E.F., 2013. Linkage disequilibrium and haplotype block structure in six
577 commercial pig lines. J. Anim. Sci. 91, :3493-3501.
578 https://doi.org/10.2527/jas.2012-6052
579 Wada, Y., Akita, T., Awata, T., Furukawa, T., Sugai, N., Inage, Y., Ishii, K., Ito, Y.,
580 Kobayashi, E., Kusumoto, H., Matsumoto, T., Mikawa, S., Miyake, M., Murase,
581 A., Shimanuki, S., Sugiyama, T., Uchida, Y., Yanai, S., Yasue, H., 2000.

582 Quantitative trait loci (QTL) analysis in a Meishan x Göttingen cross population.
583 Anim. Genet. 31, 376-384. https://doi.org/10.1046/j.1365-2052.2000.00696.x
584 Wellcome Trust Case Control Consortium. 2007. Genome-wide association study of
585 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447,
586 661-678. https://doi.org/10.1038/nature05911
587 Willham, R.L., Whatley, J.A., 1962. Genetic variation in nipple number in swine.
588 Zeitschrift für Tierzüchtung und Züchtungsbiologie 78, 350-363.
589 https://doi.org/10.1111/j.1439-0388.1962.tb01289.x
590 Yang, J., Huang, L., Yang, M., Fan, Y., Li, L., Fang, S., Deng, W., Cui, L., Zhang, Z.,
591 Ai, H., Wu, Z., Gao, J., Ren, J., 2016. Possible introgression of the *VRTN* mutation
592 increasing vertebral number, carcass length and teat number from Chinese pigs into
593 European pigs. Sci. Rep. 6, 19240. https://doi.org/10.1038/srep19240
594 Zhang, L.-C., Yue, J.-W., Pu, L., Wang, L.-G., Liu, X., Liang, J., Yan, H., Zhao, K.-B.,
595 Li, N., Shi, H.-B., Zhang, Y.-B., Wang, L.-X., 2016. Genome-wide study refines
596 the quantitative trait locus for number of ribs in a Large White × Minzhu intercross
597 pig population and reveals a new candidate gene. Mol. Genet. Genomics 291, 1885-
598 1890. https://doi.org/10.1007/s00438-016-1220-1
599 Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for
600 association studies. Nat. Genet. 44, 821-824. https://doi.org/10.1038/ng.2310
601 Zhuang, Z., Ding, R., Peng, L., Wu, J., Ye, Y., Zhou, S., Wang, X., Quan, J., Zheng, E.,
602 Cai, G., Huang, W., Yang, J., Wu, Z., 2020. Genome-wide association analyses
603 identify known and novel loci for teat number in Duroc pigs using single-locus and
604 multi-locus models. BMC Genomics 21, 344. https://doi.org/10.1186/s12864-020-
605 6742-6

606 **Table 1**. Datasets used in the single-marker and haplotype-based genome-wide
607 association studies carried out in Italian Duroc pigs.
608

| **Parameter** | **Single-marker** | **Haplotype-based** |
|---|---|---|
| Animals used in GWAS (*n*) | 1162 | 1162 |
| Genotyped DNA markers used in GWAS (*n*) | 29604 | 85590 |
| Genomic heritability $h^2_G$ (standard error) | 0.25 (0.05) | 0.30 (0.06) |
| Inflation factor ($\lambda_{GC}$) | 1.015 | 1.010 |

609

**Table 2**. Genomic regions associated and significantly associated with the number of teats in Italian Duroc pigs. Results are stratified by marker position.

| Genome scan[1] | SSC[2] | Pos[3] | Marker[4] | Min/Maj[5] | MAF[6] | *B*[7] | *s.e.*[8] | *P*[9] | PVE(%)[10] | Candidate gene |
|---|---|---|---|---|---|---|---|---|---|---|
| Haplotype-based | 7 | 97600001 | CHR7_B920_GAG | N/H | 0.480 | -0.256 | 0.046 | $2.51\times10^{-8}$ | 2.60 | *VRTN* |
| Single-marker | 7 | 97652632 | MARC0038565 | A/G | 0.442 | -0.286 | 0.046 | $8.58\times10^{-10}$ | 3.22 | *VRTN* |
| | 7 | 98066911 | H3GA0022664 | A/G | 0.442 | -0.233 | 0.047 | $8.92\times10^{-7}$ | 2.07 | *VRTN* |
| | 7 | 98089286 | ASGA0035527 | G/A | 0.442 | -0.233 | 0.047 | $8.96\times10^{-7}$ | 2.07 | *VRTN* |
| Haplotype-based | 10 | 47000001 | CHR10_B437_AGA | N/H | 0.460 | 0.211 | 0.046 | $5.08\times10^{-6}$ | 1.78 | *FRMD4A* |
| Single-marker | 10 | 47947214 | ASGA0090802 | A/G | 0.450 | -0.209 | 0.045 | $4.80\times10^{-6}$ | 1.82 | *FRMD4A* |

[1] Genome scans performed in the Italian Duroc population.
[2] *Sus scrofa* chromosome.
[3] Position, in base pairs, on the *Sus scrofa* reference genome (version Sscrofa11.1).
[4] DNA marker identifier reported in the chip panel. For haplotypes, the haploblock identifier (chromosome specific) and the allele is reported.
[5] Minor/Major alleles. Haplotypes have been treated as bi-allelic variants (H = haplotype allele and N = other *N* alleles).
[6] Minor allele frequency.
[7] Regression coefficient. A positive value indicates that the no. of teats increases with the increasing of the number of copies of the minor allele. A negative value indicates that the no. of teats decreases with the increasing of the number of copies of the minor allele.
[8] Standard error of the regression coefficient.
[9] *P* at the Wald test of GEMMA.
[10] Proportion of variance in phenotype explained (%) by the DNA marker.

**Figure 1.** Distribution plot of the number of teats in the Italian Duroc population. The number of animals is reported at the top of each bar.

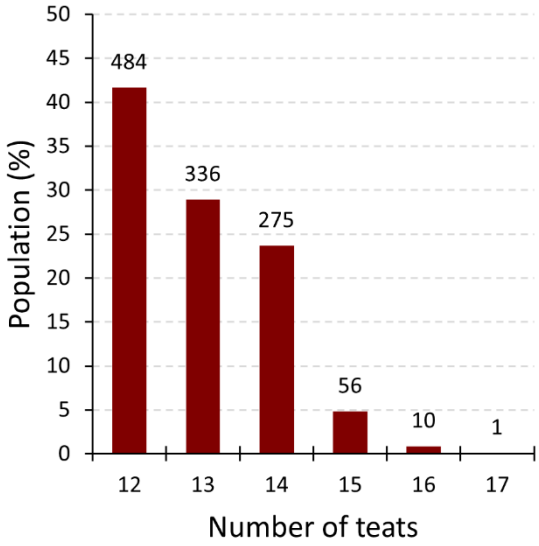



**Figure 2**. Miami plot of the number of teats in the Italian Duroc population. Results of the single-marker GWAS are on the top part of the plot whereas results of the haplotype-based analysis are on the bottom. Dots represent the single nucleotide polymorphism (SNP) or the haplotype markers. The red lines identify the significance thresholds (Bonferroni correction; $\alpha = 0.05$). Statistically associated SNPs are highlighted in green.

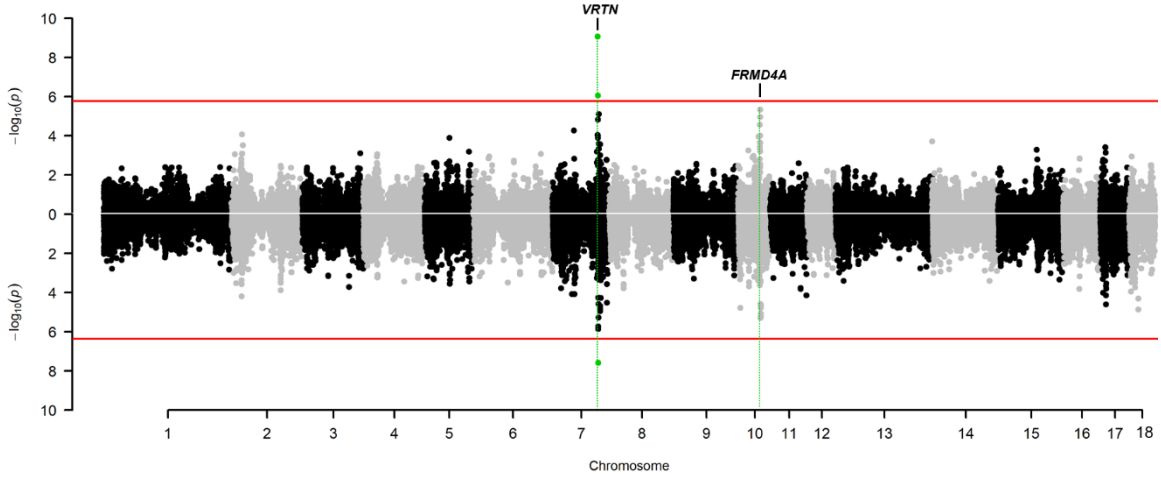
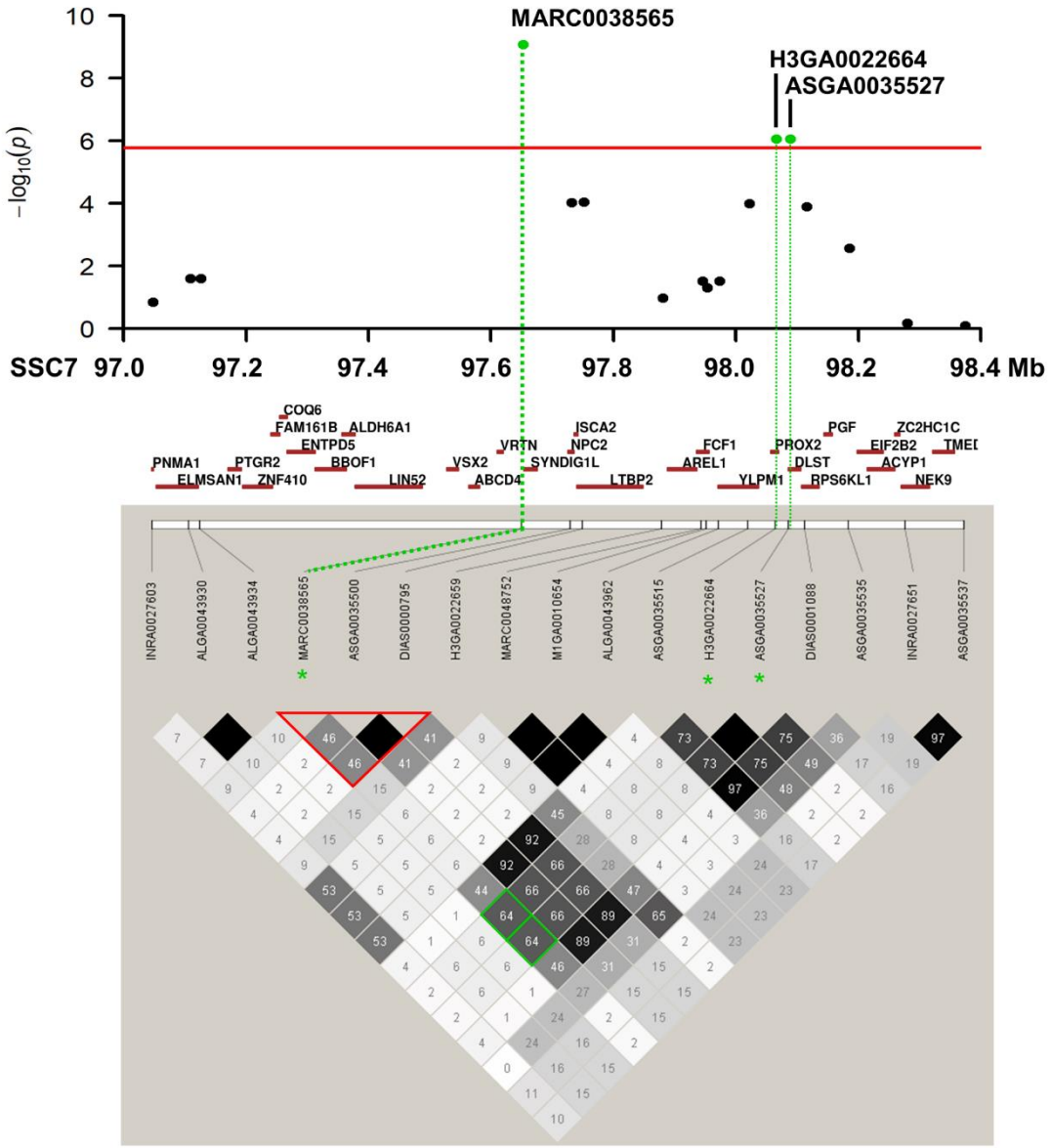


**Figure 3**. Regional association plot (single-marker analysis) for the SSC7:97.0-98.4 Mb genomic region. Linkage disequilibrium (LD) was measured between SNP pairs as $r^2$ and it is reported/showed in each box coloured in relation to its magnitude. The associated SNPs are marked with a green star symbol whereas DNA markers within the top associated haplotype (CHR7_B920_97400001_97800001_GAG) are marked with a red triangle.



**Supplementary Material**

**Table S1.** Distribution of the number of teats in the Italian Duroc pigs.

| Teat number class | Total no. of pigs (%)[§] | No. of females (%) | No. of males (%) |
|---|---|---|---|
| 12 | 484 (41.65) | 177 (40.69) | 307 (42.23) |
| 13 | 336 (28.92) | 130 (29.89) | 206 (28.34) |
| 14 | 275 (23.67) | 106 (24.37) | 169 (23.25) |
| 15 | 56 (4.82) | 20 (4.60) | 36 (4.95) |
| 16 | 10 (0.86) | 2 (0.46) | 8 (1.10) |
| 17 | 1 (0.09) | 0 (0) | 1 (0.14) |
| TOTAL | 1162 | 435 | 727 |

[§] Percentages are given within class (population, males, females).

**Table S2.** Genomic regions affecting the number of teats ($P < 5.5 \times 10^{-05}$) in Italian Duroc pigs. Results are sorted by chromosome and *P*.

| Genome scan[1] | SSC[2] | Pos[3] | Marker[4] | Min/Maj[5] | MAF[6] | *β*[7] | *s.e.*[8] | *P*[9] |
|---|---|---|---|---|---|---|---|---|
| Single-marker | 7 | 97652632 | MARC0038565 | A/G | 0.44 | -0.286 | 0.046 | 8.58E-10 |
| Haplotype-based | 7 | 97600001 | CHR7_B920_97400001_97800001_GAG | N/H | 0.48 | -0.256 | 0.046 | 2.51E-08 |
| Single-marker | 7 | 98066911 | H3GA0022664 | A/G | 0.43 | -0.233 | 0.047 | 8.92E-07 |
| Single-marker | 7 | 98089286 | ASGA0035527 | G/A | 0.43 | -0.233 | 0.047 | 8.96E-07 |
| Haplotype-based | 7 | 97600001 | CHR7_B920_97400001_97800001_AGA | H/N | 0.33 | -0.237 | 0.049 | 1.36E-06 |
| Haplotype-based | 7 | 97500001 | CHR7_B919_97300001_97700001_AA | H/N | 0.35 | -0.231 | 0.048 | 1.84E-06 |
| Haplotype-based | 7 | 98600001 | CHR7_B930_98400001_98800001_ACG | N/H | 0.41 | -0.210 | 0.046 | 5.23E-06 |
| Single-marker | 7 | 100073102 | ALGA0044076 | A/G | 0.41 | -0.220 | 0.049 | 7.96E-06 |
| Haplotype-based | 7 | 101900001 | CHR7_B963_101700001_102100001_GAAAAGC | H/N | 0.39 | -0.211 | 0.048 | 1.13E-05 |
| Haplotype-based | 7 | 101800001 | CHR7_B962_101600001_102000001_GGGAAAAG | H/N | 0.39 | -0.209 | 0.048 | 1.47E-05 |
| Single-marker | 7 | 98469335 | ASGA0035543 | G/A | 0.41 | -0.205 | 0.047 | 1.54E-05 |
| Haplotype-based | 7 | 101700001 | CHR7_B961_101500001_101900001_AGAGGGGAAAA | H/N | 0.39 | -0.205 | 0.048 | 2.02E-05 |
| Haplotype-based | 7 | 98700001 | CHR7_B931_98500001_98900001_GG | N/H | 0.38 | -0.211 | 0.050 | 2.53E-05 |
| Haplotype-based | 7 | 116400001 | CHR7_B1106_116200001_116600001_AGAAGGA | H/N | 0.04 | 0.424 | 0.101 | 2.93E-05 |
| Single-marker | 10 | 47947214 | ASGA0090802 | A/G | 0.45 | -0.209 | 0.045 | 4.80E-06 |
| Haplotype-based | 10 | 47000001 | CHR10_B437_46800001_47200001_AGA | N/H | 0.46 | 0.211 | 0.046 | 5.08E-06 |
| Haplotype-based | 10 | 47200001 | CHR10_B439_47000001_47400001_AGAAAGGA | N/H | 0.50 | 0.209 | 0.046 | 6.14E-06 |
| Haplotype-based | 10 | 47100001 | CHR10_B438_46900001_47300001_AGAAAG | N/H | 0.46 | 0.208 | 0.046 | 7.83E-06 |
| Haplotype-based | 10 | 47400001 | CHR10_B441_47200001_47600001_AAGGAGGCA | N/H | 0.49 | 0.212 | 0.047 | 8.37E-06 |
| Single-marker | 10 | 47963773 | ASGA0094083 | A/G | 0.46 | -0.201 | 0.046 | 1.12E-05 |
| Haplotype-based | 10 | 3200001 | CHR10_B31_3000001_3400001_GAGGA | H/N | 0.04 | 0.475 | 0.110 | 1.64E-05 |
| Haplotype-based | 10 | 47500001 | CHR10_B442_47300001_47700001_GAGGCAGC | H/N | 0.48 | -0.201 | 0.046 | 1.65E-05 |
| Haplotype-based | 10 | 47600001 | CHR10_B443_47400001_47800001_GGCAGCACG | H/N | 0.49 | -0.192 | 0.045 | 1.77E-05 |
| Haplotype-based | 10 | 47300001 | CHR10_B440_47100001_47500001_AAAGGAG | N/H | 0.49 | 0.201 | 0.047 | 1.79E-05 |
| Haplotype-based | 10 | 47700001 | CHR10_B444_47500001_47900001_GCAGCACGG | H/N | 0.46 | -0.194 | 0.046 | 2.30E-05 |
| Haplotype-based | 10 | 46900001 | CHR10_B436_46700001_47100001_AG | N/H | 0.31 | 0.211 | 0.050 | 2.59E-05 |
| Single-marker | 10 | 47026618 | ALGA0059028 | G/A | 0.31 | 0.209 | 0.050 | 2.80E-05 |

[1] Genome scans. SNPs and Haplotypes indicates which DNA markers have been used to carry out the genome scans.

[2] *Sus scrofa* chromosome.
[3] Position, in base pairs, on the *Sus scrofa* reference genome (Sscrofa11.1).
[4] DNA marker identifier reported in the chip panels. For haplotypes, we report; chromosome, haploblock, start, end, haplotype.
[5] Minor/Major alleles. Haplotypes have been treated as bi-allelic variants (H = haplotype allele and N = other *N* alleles).
[6] Minor allele frequency.
[7] Regression coefficient. A positive value indicates that the no. of teats increases with the increasing of the number of copies of the minor allele. A negative value indicates that the no. of teats decreases with the increasing of the number of copies of the minor allele.
[8] Standard error of the regression coefficient.
[9] *P* at the Wald test of GEMMA.

**Figure S1.** Quantile-quantile plots of the genome-wide association studies carried out in the Italian Duroc population. Inflation factor ($\lambda_{GC}$) is reported.