ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Genome-wide association studies for the number of teats and teat asymmetry patterns in Large White pigs

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Moscatelli, G., Dall'Olio, S., Bovo, S., Schiavo, G., Kazemi, H., Ribani, A., et al. (2020). Genome-wide association studies for the number of teats and teat asymmetry patterns in Large White pigs. ANIMAL GENETICS, 51(4), 595-600 [10.1111/age.12947].

Availability:
This version is available at: https://hdl.handle.net/11585/777142 since: 2020-11-02

Published:

DOI: http://doi.org/10.1111/age.12947

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

05 February 2025

SHORT COMMUNICATION

Genome-wide association studies for the number of teats and teat asymmetry patterns in Large White pigs

Giulia Moscatelli[*,1], Stefania Dall'Olio[*1], Samuele Bovo[*,1], Giuseppina Schiavo[*,1], Hamed Kazemi[*1], Anisa Ribani[*], Paolo Zambonelli[*], Silvia Tinarelli[*,†], Maurizio Gallo[†], Francesca Bertolini[‡] and Luca Fontanesi[*]

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

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

[‡]National Institute of Aquatic Resources, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

[1] These authors contributed equally to this work.

Corresponding author: Luca Fontanesi, Email: luca.fontanesi@unibo.it

Short title: GWAS for teat number parameters in heavy pigs.

Summary

The number of teats is a morphological trait that influences the mothering ability of the sows and thus their reproduction performances. In this study, we carried out genome-wide association analyses for the total number of teats and other 12 related parameters in 821 Italian Large White heavy pigs. All pigs were genotyped with the Illumina PorcineSNP60 BeadChip array. For four investigated parameters (total number of teats, TN; the number of teats of the left line, LTN; the number of teats of the right line, RTN; and the maximum number of teats comparing the two sides, MAX) significant markers were identified on SSC7, in the region of the *vertnin* (*VRTN*) gene. Significant markers for the numbers of posterior teats (NPT) and the absolute difference between anterior and posterior teat numbers (ADIFF-AP) were consistently identified on SSC6. The most significant SNP for these parameters was an intron variant in the *TOX high mobility group box family member 3* (*TOX3*) gene. For other four parameters (absolute difference between the two sides; anterior teats; the ratio between the posterior and the anterior number of teats; and the absence or the presence of extra teats) only suggestively significant markers were identified on several other chromosomes. This study further supported the role of the *VRTN* gene region in affecting the recorded variability of the number of teats in the Italian Large White pig population and identified a genomic region potentially affecting the biological mechanisms controlling the developmental programme of morphological features in pigs.

Keywords: GWAS; single nucleotide polymorphism; *Sus scrofa*; *TOX3*; *VRTN*.

43 **Text**

44 The number of teats is an important morphological trait that largely influences the mothering
45 ability of the sows and thus their reproduction performances (Pumfrey *et al.* 1980; Kim *et al.* 2005).
46 The number of teats in pigs is considered a quantitative trait with discrete and countable values, with
47 a medium/high level of heritability, showing a considerable variability among breeds as well as within
48 breeds and lines (e.g.: Willham & Whatley 1963; McKay & Rahnefeld 1990; Borchers *et al.* 2004;
49 Chalkias *et al.* 2013; Felleki & Lundeheim 2015; Rohrer & Nonneman 2017; Dall'Olio *et al.* 2018).

50 Other related parameters, that are not usually recorded, are i) the number of teats divided in the
51 two sides and their differences, which can identify directional and fluctuating asymmetry (bilateral
52 variation), ii) the number of anterior and posterior teats (having the navel as dividing line), which can
53 provide information on their longitudinal body distribution, and iii) the number of extra or additional
54 teats (i.e. teats not placed in parallel between the two sides). Only few studies have investigated these
55 phenotypes in pigs, which might provide information on developmental patterns, developmental
56 stability and instability (Palmer 1994; Fernández *et al.* 2004). Estimations of heritability for some of
57 these teat related parameters have reported lower values than the heritability for the total number of
58 teats (Willham & Whatley 1963; McKay & Rahnefeld 1990; Borchers *et al.* 2004; Fernández *et al.*
59 2004; Rohrer & Nonneman 2017).

60 Several quantitative traits locus (QTL) studies for the total number of teats carried out in
61 reference populations, obtained by crossing different pig breeds (including highly prolific Chinese
62 breeds), have shown that this parameter is affected by variants in several chromosome regions,
63 confirming that a complex polygenic influence contributes to determine its variability (Hu *et al.*
64 2019). Then, genome-wide association studies (GWAS) within breeds and lines added other regions
65 to the list of loci affecting this trait (e.g. Duijvesteijn *et al.* 2014; Verardo *et al.* 2015; Rohrer &
66 Nonneman 2017; Van Son *et al.* 2019). A major QTL located on porcine chromosome 7 (SSC7),
67 having pleiotropic effects including both the number of teats and the number of vertebrae, has been
68 shown to segregate in several pig populations and breeds (Mikawa *et al.* 2005; Duijvesteijn *et al.*

2014; Verardo *et al.* 2015; Yang *et al.* 2016; Rohrer & Nonneman 2017; Dall'Olio *et al.* 2018; Van Son *et al.* 2019). Mikawa *et al.* (2011) fine mapped this number of vertebrae QTL and identified a new gene (*vertnin*, also known as *vertebrae development associated* or *VRTN*), encoding for a DNA binding factor required for the development of thoracic vertebrae in mammals (Duan *et al.* 2018), as the gene affecting this trait.

In this study, we carried out genome-wide association analyses for the total number of teats (TN) and several related parameters (including several asymmetry patterns of teat number and disposition on the ventral side) in 843 Italian Large White heavy pigs (278 castrated males and 565 gilts, obtained from 86 boars and 377 litters). These animals were included in the sib-testing programme of the Italian Large White pig population (Fontanesi *et al.* 2014a). A total of 13 teat related parameters were obtained after slaughtering the pigs, as described below. The ventral side of the carcasses were photographed in vertical position, before evisceration and dissection. Three independent persons visually inspected all obtained photographs and counted: (i) the number of teats of the left line (LTN) and (ii) the number of teats of the right line (RTN), then (iii) the total number of teats was calculated as the sum of LTN and RTN; (iv) the number of anterior teats (NAT) and (v) the number of posterior teats (NPT), having the navel as dividing longitudinal line (McKay & Rahnefeld 1990). A few other parameters were also calculated based on these records: (vi) the maximum number of teats between the two sides (MAX, obtained comparing the LTN and the RTN and choosing the highest value between the two) which could be considered the best indicator of the genetic potential for proliferation of initial mammary buds (Rohrer & Nonneman 2017); (vii) the signed difference between the two sides (SDIFF = RTN – LTN), which describes the directional asymmetry; and (viii) the absolute difference between the two sides (ADIFF = |RTN – LTN|), which identifies the fluctuating asymmetry (Palmer 1994). The same measures were calculated between anterior and posterior teats: (ix) the maximum number of teats between the two parts (MAX-AP); (x) the signed difference between anterior and posterior teats (SDIFF-AP = NAT – NPT); and (xi) the absolute difference between anterior and posterior teats (ADIFF-AP = |NAT – NPT|). Additionally,

other two parameters were computed: (xii) the ratio between the number of posterior and the number of anterior teats (NPT/NAT); and (xiii) the absence or the presence of extra teats (EX-T, recorded as n. = 0, 1, 2), i.e. a different number of teats between the two sides (right and left) because of the extra teat between the regularly spaced teats. Figure S1 shows a schematic representation of all basic teat parameters considered in this study from which calculated parameters were then obtained. Table S1 summarizes statistics for all considered traits. Functional and non-functional teats could not be clearly distinguished by inspecting available photographic records, therefore this trait was not considered in this study.

All pigs were genotyped with the Illumina PorcineSNP60 BeadChip (v.1 or v.2) (Illumina Co., St. Diego, CA, USA). The assignment of the SNPs to the reference pig genome (Sscrofa11.1) was obtained as previously described (Fontanesi et *al.* 2012, 2014a). Genotyping data were filtered using PLINK software 1.9 (Chang *et al.* 2015) adopting the following criteria: SNPs were retained if located on autosomes and their call rates was > 0.9, minor allele frequency (MAF) > 0.02 and if they did not deviate from Hardy-Weinberg equilibrium (considering a *p*-value >0.0001); animals with call rate >0.90 were used for the analyses. After filtering, the dataset was composed by 821 animals and 50069 autosomal variants.

Genome-wide association analyses were carried out following a linear mixed effect model:

$$\boldsymbol{y} = \boldsymbol{W\alpha} + \boldsymbol{x}\beta + \boldsymbol{u} + \boldsymbol{e}$$

where $\mathbf{y}$ ($n \times 1$) is a vector containing parameter for the n^{th} animal, $\boldsymbol{W}$ ($n \times k$) is a matrix of a covariates (fixed effects) with $k = 2$ (a column of 1s, sex), $\boldsymbol{\alpha}$ is the k-dimensional vector of covariates effects, $\boldsymbol{x}$ ($n \times 1$) is the vector containing genotypes for the i^{th} SNP (coded as 0, 1, 2, according to the number of copies of the minor allele), β is the additive effect of the i^{th} SNP on the trait, $\mathbf{u} \sim N(\mathbf{0}, \sigma^2_u \mathbf{K})$ is a multivariate Gaussian polygenic effect, with covariance matrix proportional to the relatedness matrix $\mathbf{K}$ ($n \times n$) and $\mathbf{e} \sim N(\mathbf{0}, \sigma^2_e \mathbf{I})$ is a multivariate Gaussian vector of uncorrelated residuals. For each trait, the effect of the top associated SNP was evaluated by including in the model, as fixed effect, the

genotype of each tested individual. The genotype was coded with dummy variables that led to increase the size of covariate matrix *k* from 2 to 4 (three genotypes can be coded with N-1 dummy variables). Moreover, for each trait we evaluated the effect of the *VRTN* polymorphism by including it, as fixed effect, in the model. The assessment of the association between each SNP and trait was obtained by testing the null hypothesis H_0:β = 0. Significance was tested by using the Wald test. All analyses were performed using GEMMA v. 0.96 (Zhou & Stephens 2012) after computing the relatedness matrix **G** as a centered genomic matrix controlling the population structure. A Bonferroni corrected threshold equal to a nominal value of 0.05 was used to define significant markers (*p*-value = 0.05/50069 = 9.9×10^{-07}). To take into consideration also moderate associations and balance the risk of Type I and Type II errors, in our analyses we considered a suggestively significance threshold of *p*-value = 5.0×10^{-05}, as widely adopted in GWAS in farm animals (e.g. Fontanesi *et al.* 2012; Sanchez *et al.* 2014; Bovo *et al.* 2019). For each trait, GEMMA estimated from the whole set of available genotypes the chip heritability (or SNP heritability; h^2_{SNP}; Table S1). Quantile-quantile (QQ)-plots and Manhattan plots were generated in R v. 3.5.1 (R Core Team 2018) by using the "qqman" package whereas the genomic inflation factors (λ) were computed with the function "estlambda" (method: "regression") within the "GenABEL" package (Aulchenko *et al.* 2007). Associated peaks were annotated with Biomart (http://www.ensembl.org/biomart/martview/) and the pig QTL database (Hu *et al.* 2019; Table S2).

Among the 13 considered teat parameters, those that were based on total or sided counted teats (TN, RTN, MAX and LTN) had the highest h^2_{SNP} values (0.360, 0.258, 0.261 and 0.158, respectively). These estimates were in line to what was previously reported, even if these values were a little bit higher than what provided by Rohrer & Nonneman (2017) for the same traits. All other parameters had low or virtually null h^2_{SNP} values, suggesting that additive genetic factors might have low or negligible effects on most asymmetric patterns, as already reported by other authors (Fernández *et al.* 2004; Rohrer & Nonneman 2017).

Results of the GWAS are summarized in Table 1, which lists significant SNPs for six parameters (TN, LTN, RTN, MAX, NPT and ADIFF-AP). Figure 1 reports the Manhattan plots for these six parameters. QQ-plots for these GWAS are reported in Figure S2. For other four parameters (ADIFF, NAT, NPT/NAT and EX-T) only suggestively significant markers were identified (Figure S3). All suggestively significant markers are reported in Table S2. Figure S4 includes the Manhattan plots and QQ-plots for the remaining traits (SDIFF, MAX-AP, SDIFF-AP).

For TN, LTN, RTN and MAX a major significant peak was observed on SSC7, in the region of the *VRTN* gene (located at nucleotide positions from 97614707 to 97624273), with the most significant SNP (MARC0038565 or rs80894106, located at position 97652632) for all four parameters. To confirm that the identified QTL could be attributed to the *VRTN* gene, we genotyped in 821 pigs the indel determined by the insertion of a short interspersed nuclear element PRE1 of 291 bp (AB554652:g.20311_20312ins291), considered a causative mutation for the number of vertebrae/number of teats QTL (Mikawa *et al.* 2011), using the protocols already described (Fontanesi *et al.* 2014b). This marker was included in the genome-wide association analyses and resulted significant (TN and MAX) or suggestively significant (LTN and RTN) and in the same QTL peak already observed (Table 1 and Table S2). When this marker or the MARC0038565 SNP were conditionally included in the model of the GWAS all effects for these four parameters were erased, further supporting the presence of only one segregating QTL in this chromosome region.

Significant markers for NPT and ADIFF-AP, and suggestive associated markers for NAT were consistently identified on SSC6 (Table 1 and Table S2; Figure 1 and Figure S5). The most significant SNP for NPT and ADIFF-AP was ASGA0100698 (rs81476132), an intron variant at position 32528964 within the *TOX high mobility group box family member 3* (*TOX3*) gene. This gene encodes a protein with a high-mobility-group motif that modifies chromatin structure by bending or unwinding DNA. It is involved in mediating calcium-dependent transcription and interacts with the cAMP response element-binding protein (Yuan *et al.* 2009). Mutations in this gene have been implicated in high breast cancer risks (Easton *et al.* 2007). Highly conserved genomic structures across mammals

might be involved in defining the function of this developmental regulator gene (Harmston *et al.* 2017) which has been also suggested to be involved in determining asymmetry patterns in embryonic development (Wilting & Hagedorn 2011). Even if the role and function of *TOX3* are still far to be completely understood, what is currently known might support its candidacy for a role on the teat asymmetry parameter (anterior-posterior numbers) measured in pigs. This parameter is derived by the relative position of the naval and could describe the effects of biological mechanisms controlling the developmental programme of morphological feature(s) in pigs (i.e. the navel position or anterior/poster shifts of the two teat lines).

Other suggestively significant markers in regions not related to the two previous QTL (Table S2) were identified on SSC1 (RTN and MAX), SSC3 (TN, RTN, MAX and EX-T), SSC4 (EX-T), SSC5 (ADIFF), SSC9 (NPT/NAT), SSC11 (ADIFF, NPT/NAT and EX-T), SSC14 (NPT/NAT), SSC16 (ADIFF and NPT/NAT) and SSC17 (NPT and ADIFF-AP). Some of these markers are located in regions where other studies have reported QTL for teat number related traits or other potentially related morphological traits (Table S2). No other QTL peak emerged in all GWAS that included MARC0038565 or *VRTN* markers as fixed effects in the models.

In conclusion, this study further supported the role of the *VRTN* gene region in affecting the recorded variability in the number of teats in the Italian Large White pig population (Dall'Olio *et al.* 2018). It seems that this chromosome region (SSC7) harbors one of the few major QTL for the number of teats for which alleles are still segregating in this breed. We already observed an increasing trend of the favorable allele frequency of this gene over the past few decades of directional selection towards a higher number of teats in this breed (Fontanesi *et al.* 2015) that might have also acted in fixating other QTL for this important trait. It will be interesting to understand if epigenetic mechanisms or other factors could explain the observed heterogeneity and variability for most of the other teat related measures that could provide information on developmental patterns and instability.

Acknowledgements

This study received funds from the Italian MiPAAF Innovagen project, the Italian MIUR PRIN2017 *PigPhenomics* project and the University of Bologna RFO 2018-2019 programmes. This study was also supported by the PSRN (Progetto di Sviluppo Rurale Nazionale) SUIS project (co-funded by the European Agricultural Fund for Rural Development of the European Union and by the Italian Ministry of Agriculture, Food, Forestry and Tourism - MiPAAFT).

Competing interests

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

References

Aulchenko Y.S., Ripke S., Isaacs A. & van Duijn C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294-6.

Borchers N., Reinsch N. & Kalm E. (2004) The number of ribs and vertebrae in a Pietrain cross: variation, heritability and effects on performance traits. *Journal of Animal Breeding and Genetics* **121**, 392-403.

Bovo S., Mazzoni G., Bertolini F., Schiavo G., Galimberti G., Gallo M., Dall'Olio S. & Fontanesi L. (2019) Genome-wide association studies for 30 hematological and blood clinical-biochemical traits in Large White pigs reveal genomic regions affecting intermediate phenotypes. *Scientific Reports* **9**, 7003.

Chalkias H., Rydhmer L. & Lundeheim N. (2013) Genetic analysis of functional and non-functional teats in a population of Yorkshire pigs. *Livestock Science* **152**, 127-34.

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.

221 Dall'Olio S., Ribani A., Moscatelli G., Zambonelli P., Gallo M., Nanni Costa L. & Fontanesi L. (2018)
222 Teat number parameters in Italian Large White pigs: Phenotypic analysis and association with
223 vertnin (*VRTN*) gene allele variants. *Livestock Science* **210**, 68-72.
224 Duan Y., Zhang H., Zhang Z., *et al.* (2018) VRTN is required for the development of thoracic
225 vertebrae in mammals. *International Journal of Biological Sciences* **14**, 667-81.
226 Duijvesteijn N., Veltmaat J.M., Knol E.F. & Harlizius B. (2014) High-resolution association mapping
227 of number of teats in pigs reveals regions controlling vertebral development. *BMC Genomics*
228 **15**, 542.
229 Easton D.F., Pooley K.A., Dunning A.M., *et al.* (2007) Genome-wide association study identifies
230 novel breast cancer susceptibility loci. *Nature* **447**, 1087-93.
231 Felleki M. & Lundeheim N. (2015) Genetic heteroscedasticity of teat count in pigs. *Journal of Animal*
232 *Breeding and Genetics* **132**, 392-8.
233 Fernández A., Toro M., Rodríguez C. & Silió L. (2004) Heterosis and epistasis for teat number and
234 fluctuating asymmetry in crosses between Jiaxing and Iberian pigs. *Heredity* **93**, 222-7.
235 Fontanesi L., Galimberti G., Calò D.G., *et al.* (2012) Identification and association analysis of several
236 hundred single nucleotide polymorphisms within candidate genes for back fat thickness in
237 Italian Large White pigs using a selective genotyping approach. *Journal of Animal Science*
238 **90**, 2450-464.
239 Fontanesi L., Schiavo G., Galimberti G., Calò D.G. & Russo V. (2014a) A genomewide association
240 study for average daily gain in Italian Large White pigs. *Journal of Animal Science* **92**, 1385-
241 94.
242 Fontanesi L., Schiavo G., Scotti E., *et al.* (2015) A retrospective analysis of allele frequency changes
243 of major genes during 20 years of selection in the Italian Large White pig breed. *Journal of*
244 *Animal Breeding and Genetics* **132**, 239-46.
245 Fontanesi L., Scotti E., Buttazzoni L., Dall'Olio S. & Russo V. (2014b) Investigation of a short
246 interspersed nuclear element polymorphic site in the porcine vertnin gene: allele frequencies

247 and association study with meat quality, carcass and production traits in Italian Large White
248 pigs. *Italian Journal of Animal Science* **13**, 3090.
249 Harmston N., Ing-Simmons E., Tan G., Perry M., Merkenschlager M. & Lenhard B. (2017)
250 Topologically associating domains are ancient features that coincide with Metazoan clusters of
251 extreme noncoding conservation. *Nature Communications* **8**, 441.
252 Hu Z., Park C.A. & Reecy J.M. (2019) Building a livestock genetic and genomic information
253 knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic*
254 *Acids Research* **47**, 701-10.
255 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
256 on litter size in gilts. *Animal Reproduction Science* **90**, 111-6.
257 McKay R.M. & Rahnefeld G.W. (1990) Heritability of teat number in swine. *Canadian Journal of*
258 *Animal Science* **70**, 425-30.
259 Mikawa S., Hayashi T., Nii M., Shimanuki S., Morozumi T. & Awata T. (2005) Two quantitative
260 trait loci on *Sus scrofa* chromosomes 1 and 7 affecting the number of vertebrae. *Journal of*
261 *Animal Science* **83**, 2247-54.
262 Mikawa S., Sato S., Nii M., *et al.* (2011) Identification of a second gene associated with variation in
263 vertebral number in domestic pigs. *BMC Genetics* **12**, 5
264 Palmer A.R. (1994) Fluctuating asymmetry analyses: a primer, in Markow T.A. (Eds.)
265 Developmental instability: its origins and evolutionary implications. *Kluwer Academic*
266 *Publishers: Dordrecht, The Netherlands*, pp. 335-64.
267 Pumfrey R.A., Johnson R.K., Cunningham P.J. & Zimmerman D.R. (1980) Inheritance of teat number
268 and its relationship to maternal traits in swine. *Journal of Animal Science* **50**, 1057-60.
269 R Development Core Team (2018) R: A language and environment for statistical computing. R
270 Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Rohrer G.A. (2000) Identification of quantitative trait loci affecting birth characters and accumulation of backfat and weight in a Meishan-White Composite resource population. *Journal of Animal Science* **78**, 2547-53.

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. *Genetics Selection Evolution* **49**, 4.

Sanchez M., Tribout T., Iannuccelli N., *et al.* (2014) A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. *Genetics Selection Evolution* **46,** 12.

Van Son M., Lopes M.S., Martell H.J., *et al.* (2019) A QTL for number of teats shows breed specific effects on number of vertebrae in pigs: bridging the gap between molecular and quantitative genetics. *Frontiers in Genetics* **10**, 272.

Verardo L.L., Silva F.F., Varona L., Resende M.D.V., Bastiaansen J.W.M., Lopes P.S. & Guimarães S.E.F. (2015) Bayesian GWAS and network analysis revealed new candidate genes for number of teats in pigs. *Journal of Applied Genetics* **56**, 123-32.

Willham V.R.L & Whatley J. (1963) Genetic variation in nipple number in swine. *Journal of Animal Breeding and Genetics* **78**, 350-63.

Wilting J. & Hagedorn M. (2011) Left-right asymmetry in embryonic development and breast cancer: common molecular determinants? *Current Medicinal Chemistry* **18**, 5519-27.

Yang J., Huang L., Yang M., *et al.* (2016) Possible introgression of the *VRTN* mutation increasing vertebral number, carcass length and teat number from Chinese pigs into European pigs. *Scientific Reports* **6,** 19240.

Yuan S.H., Qiu Z. & Ghosh A. (2009) TOX3 regulates calcium-dependent transcription in neurons. *Proceedings of the National Academy of Sciences of the USA* **106**, 2909-14.

Zhou X. & Stephens M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* **44**, 821-4.

297 **Table 1.** Significant markers identified in the genome-wide association studies.

Parameter[1]	**SSC**[2]	**Marker position**[3]	**Marker**	**Alleles**[4]	**MAF**[5]	***p*-value**[6]	**Candidate gene**[7]
TN	7	97652632	MARC0038565	G/A	0.494	1.62×10^{-10}	*VRTN* (28359)
	7	97619490	VRTN*	A/G	0.393	3.32×10^{-07}	*VRTN* (0)
LTN	7	97652632	MARC0038565	G/A	0.494	4.04×10^{-07}	*VRTN* (28359)
RTN	7	97652632	MARC0038565	G/A	0.494	1.11×10^{-09}	*VRTN* (28359)
	7	97881096	H3GA0022659	A/G	0.464	9.91×10^{-07}	*VRTN* (256823)
MAX	7	97652632	MARC0038565	G/A	0.494	1.16×10^{-11}	*VRTN* (28359)
	7	97881096	H3GA0022659	A/G	0.464	1.05×10^{-07}	*VRTN* (256823)
	7	97619490	VRTN*	A/G	0.393	2.08×10^{-07}	*VRTN* (0)
	7	97795647	M1GA0010653	A/G	0.482	3.70×10^{-07}	*VRTN* (171374)
NPT	6	32528964	ASGA0100698	A/G	0.209	5.55×10^{-08}	*TOX3* (0)
	6	32678775	ALGA0035080	G/A	0.275	5.66×10^{-07}	*TOX3* (68847)
ADIFF-AP	6	32528964	ASGA0100698	A/G	0.209	4.18×10^{-07}	*TOX3* (0)

298

299 [1] TN, total number of teats; LTN, number of left line teats; RTN, number of right line teats; NP, number of posterior teats; MAX, maximum number

300 of teats between the two sides; ADIFF-AP, absolute difference between anterior and posterior teats (|NAT – NPT|).

301 [2] Porcine chromosome.

302 [3] Position of the marker on the corresponding chromosome in the Sscrofa11.1 genome version.

303 [4] Minor and major alleles.

304 [5] Minor allele frequency.

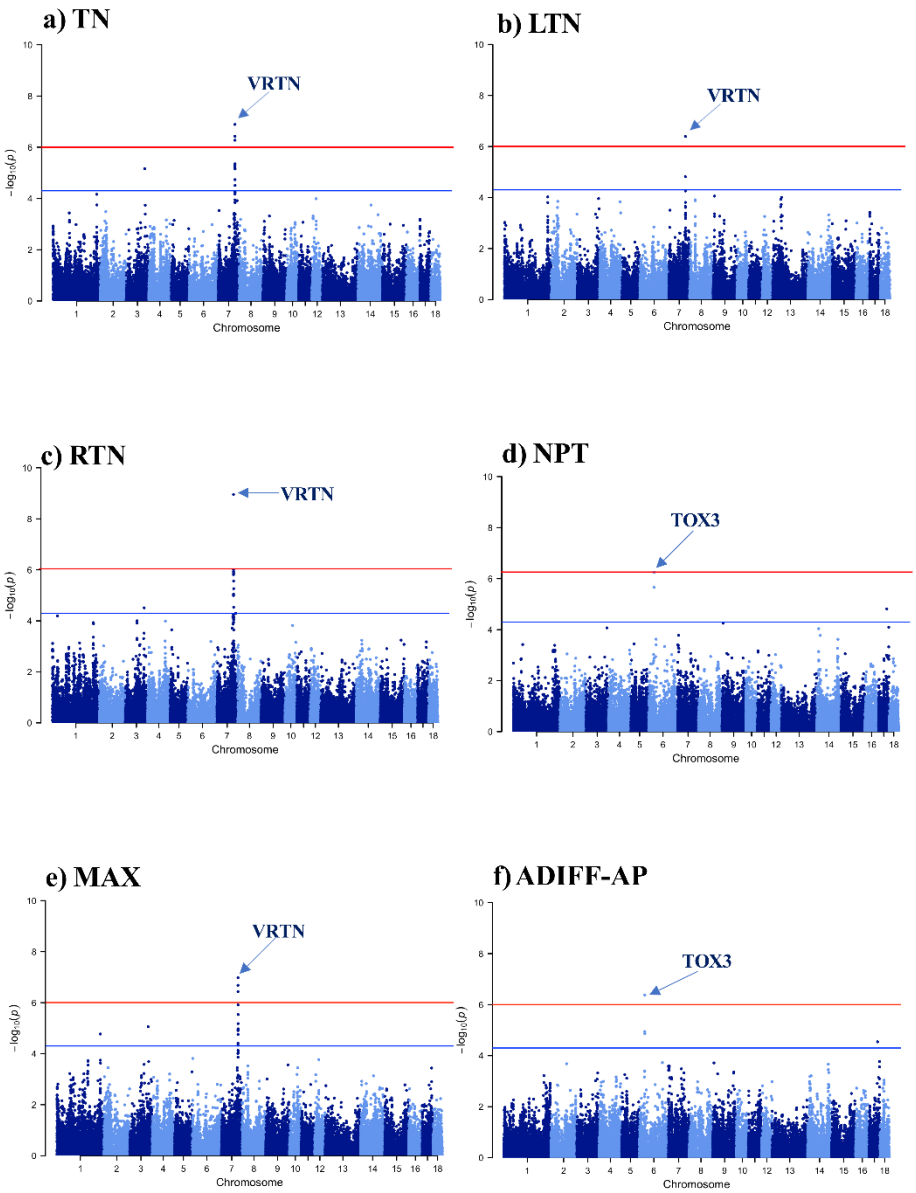
305 [6] *p*-value from GEMMA (Wald test). Only single nucleotide polymorphisms (SNPs) significantly associated after Bonferroni correction are reported.

306 [7] Candidate genes identified in the significant SNP region (± 500 kbp) according to their function and potential role in the analysed phenotypes, in

307 brackets is reported the distance of the marker to the gene in bp. Zero means that the marker was within the gene.

308 *In house genotyped AB554652:g.20311_20312ins291 *VRTN* related polymorphism (Fontanesi *et al.* 2014b).

Figure 1. Manhattan plots obtained in the GWAS with teat number related parameters. The plots are reported only for the six parameters which had significant markers. The redline identifies the threshold for statistical significance, the blue line indicates the suggestively significance threshold. a) TN, total number of teats; b) LTN, number of left line teats; c) RTN, number of right line teats; d) NPT, number of posterior teats; e) MAX, the maximum number of teats between the two sides; f) ADIFF-AP, absolute difference between anterior and posterior teats (|NAT – NPT|).



Supplementary Materials

for

Genome-wide association studies for the number of teats and teat asymmetry patterns in Large White pigs

Giulia Moscatelli, Stefania Dall'Olio, Samuele Bovo, Giuseppina Schiavo, Hamed Kazemi, Anisa Ribani, Paolo Zambonelli, Silvia Tinarelli, Maurizio Gallo, Francesca Bertolini and Luca Fontanesi

Table S1. Summary of basic statistics on the number of teats and related parameters included in this study observed in Italian Large White population. Single nucleotide polymorphism heritability (h^2_{SNP}) is reported together with the genomic inflation factor value (λ) obtained in the corresponding genome-wide association studies.

Parameters	**Acronym**	**No. of pigs[1]**	**Mean[2]**	**s.d.[3]**	**Median[4]**	h^2_{SNP}	**s.e.[5] of** h^2_{SNP}	**λ**
Total number of teats	TN	821	14.79	0.96	15	0.360	0.072	1.019
Total number of left line teats	LTN	821	7.34	0.58	7	0.158	0.055	1.021
Total number of right line teats	RTN	821	7.45	0.60	7	0.258	0.070	1.025
Total number of anterior teats	NAT	667	6.66	0.85	7	<0.001	0.036	0.966
Total number of posterior teats	NPT	667	8.13	0.62	8	0.034	0.036	1.012
Max number of teats comparing the two sides	MAX	821	7.61	0.59	8	0.261	0.068	1.026
Signed difference between RTN and LTN (RTN - LTN)	SDIFF	821	0.10	0.69	0	<0.001	0.037	0.963
Absolute difference between RTN and LTN (|RTN – LTN|)	ADIFF	821	0.44	0.55	0	0.051	0.041	1.023
Max number of teats between anterior and posterior parts	MAX-AP	667	8.18	0.58	8	<0.001	0.036	0.966
Signed difference between anterior and posterior teats (NAT - NPT)	SDIFF-AP	667	-1.47	1.18	-2	<0.001	0.045	0.916
Absolute difference between anterior and posterior teats (|NAT – NPT|)	ADIFF-AP	667	1.56	1.06	2	0.030	0.033	1.007
Ratio between posterior and anterior teats	NPT/NAT	667	1.24	0.22	1.28	0.033	0.039	1.064
Extra teats	EX-T	821	0.43	0.54	0	<0.001	0.187	0.982

[1]Number of pigs with the phenotype information and that were included in the genome-wide association studies: for some parameters it was not possible to obtain the information for all animals. Particularly, the position of the navel was not always possible to clearly distinguish from the available pictures.
[2]Mean: average of the number of teats.
[3]s.d.: standard deviation of the number of teats.
[4]Median: median of the number of teats.
[5]s.e. of h^2_{SNP}: standard error of the chip heritability.

Table S2. Suggestively associated markers identified in the genome-wide association studies.

Parameter[1]	SSC[2]	Marker position[3]	Marker	Alleles[4]	MAF[5]	*p*-value[6]	Closest gene[7]	QTL[8]
TN	7	97795647	M1GA0010653	A/G	0.482	1.10×10^{-06}	*VRTN* (171374)	TVN
	7	97881096	H3GA0022659	A/G	0.464	1.66×10^{-06}	*VRTN* (256823)	TVN
	3	108957096	ALGA0020825	A/C	0.032	8.47×10^{-06}	*LCLAT1* (75074)	-
	7	97147161	H3GA0022644	G/A	0.297	1.49×10^{-05}	*ELMSAN1* (24195)	TVN
	7	98155866	M1GA0010658	C/A	0.393	1.53×10^{-05}	*ENSSSCG00000050616* (0)	VN, TVN
	7	97752476	DIAS0000795	G/A	0.439	1.59×10^{-05}	*LTBP2* (0)	TVN
	7	97732109	ASGA0035500	A/G	0.438	1.84×10^{-05}	*NPC2*(0)	TVN
	7	98595714	ALGA0122954	G/A	0.402	2.24×10^{-05}	*JDP2* (0)	VN, TVN, RTN
LTN	7	97619490	VRTN*	A/G	0.393	1.52×10^{-05}	*VRTN* (0)	NNF, TVN
RTN	7	97619490	VRTN*	A/G	0.393	1.16×10^{-06}	*VRTN* (0)	NNF, TVN
	7	98155866	M1GA0010658	C/A	0.393	1.32×10^{-06}	*ENSSSCG00000050616* (0)	VN, TVN
	7	97795647	M1GA0010653	A/G	0.482	1.53×10^{-06}	*VRTN* (171374)	TVN
	7	98264173	ASGA0035536	C/A	0.477	2.75×10^{-06}	*ACYP1* (0)	VN, TVN
	7	98595714	ALGA0122954	G/A	0.402	5.45×10^{-06}	*JDP2* (0)	VN, TVN, RTN
	7	97752476	DIAS0000795	G/A	0.439	8.87×10^{-06}	*LTBP2* (0)	TVN
	7	97732109	ASGA0035500	A/G	0.438	1.01×10^{-05}	*NPC2*(0)	TVN
	7	98566049	ALGA0112470	G/A	0.485	2.93×10^{-05}	*JDP2* (8149)	VN, TVN, RTN
	3	112227667	ALGA0020993	A/G	0.468	3.08×10^{-05}	*DPYSL5* (0)	TVN
	1	159785087	ASGA0101718	A/G	0.151	4.65×10^{-05}	*CDH20* (32470)	TN
	7	110378434	DRGA0008163	A/G	0.363	5.03×10^{-05}	*PTPN21* (0)	TN
NAT	6	32502689	ASGA0027999	A/G	0.303	5.25×10^{-05}	*TOX3* (0)	TN
NPT	6	32604025	ALGA0035073	G/A	0.239	2.19×10^{-06}	*TOX3* (0)	TN
	6	32529342	ASGA0100674	A/G	0.170	1.42×10^{-05}	*TOX3* (0)	TN
	17	49368877	ASGA0077390	A/G	0.070	1.53×10^{-05}	*ZMYND8* (0)	LVN, THVN
MAX	7	98155866	M1GA0010658	C/A	0.393	1.23×10^{-06}	*ENSSSCG00000050616* (0)	VN, TVN
	7	98595714	ALGA0122954	G/A	0.402	2.92×10^{-06}	*JDP2* (0)	VN, TVN, RTN
	7	98264173	ASGA0035536	C/A	0.477	6.63×10^{-06}	*ACYP1* (0)	VN, TVN
	7	95984726	ALGA0109584	A/G	0.222	6.91×10^{-06}	*DPF3* (0)	TN
	3	108957096	ALGA0020825	A/C	0.032	8.79×10^{-06}	*LCLAT1* (75074)	TN
	7	98440639	MARC0050386	A/G	0.476	1.06×10^{-05}	-	VN, TVN
	7	98066911	H3GA0022664	A/G	0.463	1.11×10^{-05}	*PROX2* (0)	TVN

	7	97347282	INRA0027622	A/G	0.358	1.20×10^{-05}	*BBOF1* (0)	TVN
	7	98089286	ASGA0035527	G/A	0.466	1.24×10^{-05}	*DLST* (1837)	TVN
	1	254612941	ALGA0009455	G/A	0.386	1.69×10^{-05}	*ZNF618* (0)	TN
	7	98116120	DIAS0001088	G/A	0.464	1.76×10^{-05}	*RPS6KL1* (0)	TVN
	7	97752476	DIAS0000795	G/A	0.439	3.86×10^{-05}	*LTBP2* (0)	TVN
	7	97732109	ASGA0035500	A/G	0.438	4.67×10^{-05}	*NPC2*(0)	TVN
ADIFF	11	11961785	ALGA0060867	A/G	0.277	2.17×10^{-06}	*DCLK1* (0)	TN
	5	102805263	M1GA0008203	C/A	0.086	4.38×10^{-05}	*NAV3* (208535)	TVN
	16	70736511	MARC0039548	A/G	0.224	5.26×10^{-05}	*NMUR2* (60195)	TN
ADIFF-AP	6	32604025	ALGA0035073	G/A	0.239	1.13×10^{-05}	*TOX3* (0)	TN
	6	32678775	ALGA0035080	G/A	0.275	1.37×10^{-05}	*TOX3* (68847)	TN
	17	49368877	ASGA0077390	A/G	0.070	2.87×10^{-05}	Z*MYND8* (0)	LVN, THVN
NPT/NAT	16	76153481	ASGA0099992	G/A	0.295	1.28×10^{-05}	-	TN
	11	9046973	H3GA0031250	G/A	0.264	1.30×10^{-05}	*PDS5E* (0)	TN
	9	5519984	ASGA0041143	A/G	0.242	3.12×10^{-05}	-	-
	14	132709906	DBWU0000067	G/A	0.070	3.81×10^{-05}	*ATP5PB* (0)	-
	9	3184033	MARC0002885	A/G	0.298	4.01×10^{-05}	*RRP8* (21421)	-
EX-T	4	23524316	ALGA0024067	G/A	0.399	8.06×10^{-06}	*TRPS1* (365547)	TN
	11	72970630	CASI0009242	G/A	0.074	1.20×10^{-05}	-	-
	3	124103822	H3GA0010881	G/A	0.167	1.33×10^{-05}	*TRIB2* (66855)	TN
	4	23121547	INRA0013304	A/G	0.279	3.32×10^{-05}	*TRPS1* (0)	TN

[1] TN, total number of teats; LTN, number of left line teats; RTN, number of right line teats; NAT, number of anterior teats; NPT, number of posterior teats; MAX, maximum number of teats between the two sides; ADIFF, absolute difference between RTN and LTN (|RTN – LTN|); ADIFF-AP, absolute difference between anterior and posterior teats (|NAT – NPT|); NPT/NAT, ratio between the number of poster and the number of anterior teats; EX-T, extra teats. Suggestively significant markers were also identified for all traits for which significant markers were reported in Table 1.

[2] Porcine chromosome.

[3] Position of the marker on the corresponding chromosome in the Sscrofa11.1 genome version.

[4] Minor and major alleles.

[5] Minor allele frequency.

[6] *p*-value from GEMMA (Wald test).

[7] The closest gene to the indicated marker in a region of ± 500 kbp. The distance of the marker to the gene is indicated in brackets (bp). Zero means that the marker was within the gene.

[8] QTL identified in a region of ± 200 kbp with what reported in PigQTL db, related to morphological traits that might be related/correlated to the investigated teat parameters. Short names for QTL are: TVN, Thoracic vertebra number; NNF, Number of non-viable fetuses; VN, Vertebra number; RTN, Right teat number; THVN, Thoracolumbar vertebra number; LVN, lumbar vertebra number; TN, teat number.
*In house genotyped AB554652:g.20311_20312ins291 *VRTN* related polymorphism (Fontanesi *et al.* 2014b).

Figure S1. Schematic representation of the eight basic teat number related parameters included in the study: a) TN, total number of teats; b) LTN, number of left line teats; c) RTN, number of right line teats; d) NAT, number of anterior teats; e) NPT, number of posterior teats; f) MAX, maximum number of teats comparing the two sides; g) MAX-AP, maximum number of teats between anterior and posterior parts; h) EX-T, extra teats. The other parameters are derived from these basic parameters: SDIFF, signed difference between RTN and LTN (RTN - LTN); ADIFF: absolute difference between RTN and LTN (|RTL – LTN|); SDIFF-AP, signed difference between NAT and NPT (NAT - NPT); ADIFF-AP, absolute difference between NAT and NPT (|NAT - NPT|); ratio between posterior and anterior teats (NPT/NAT).

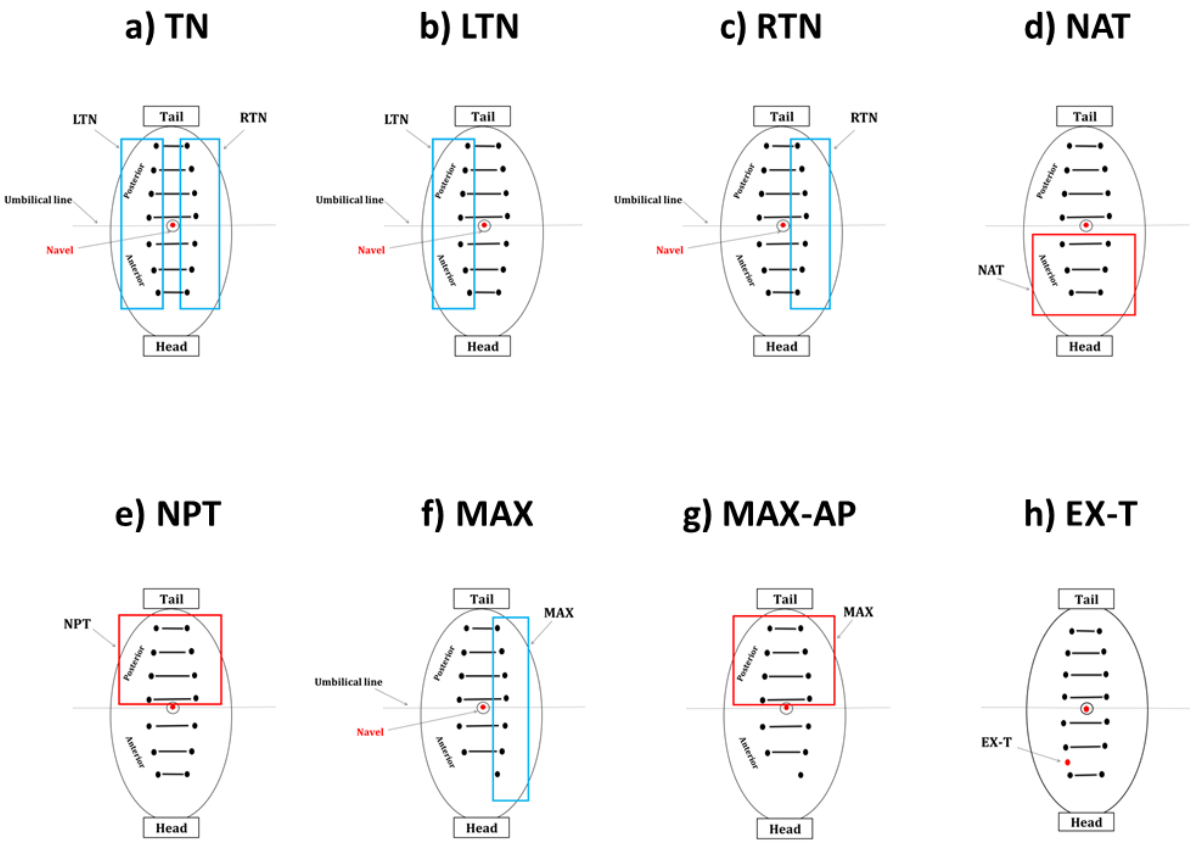


Figure S2. Quantile-quantile plots of the six genome-wide association studies described in Table 1: a) TN, total number of teats; b) LTN, number of left line teats; c) RTN, number of right line teats; d) NPT, number of posterior teats; e) MAX, the maximum number of teats between the two sides; f) ADIFF-AP, absolute difference between anterior and posterior teats (|NAT – NPT|).

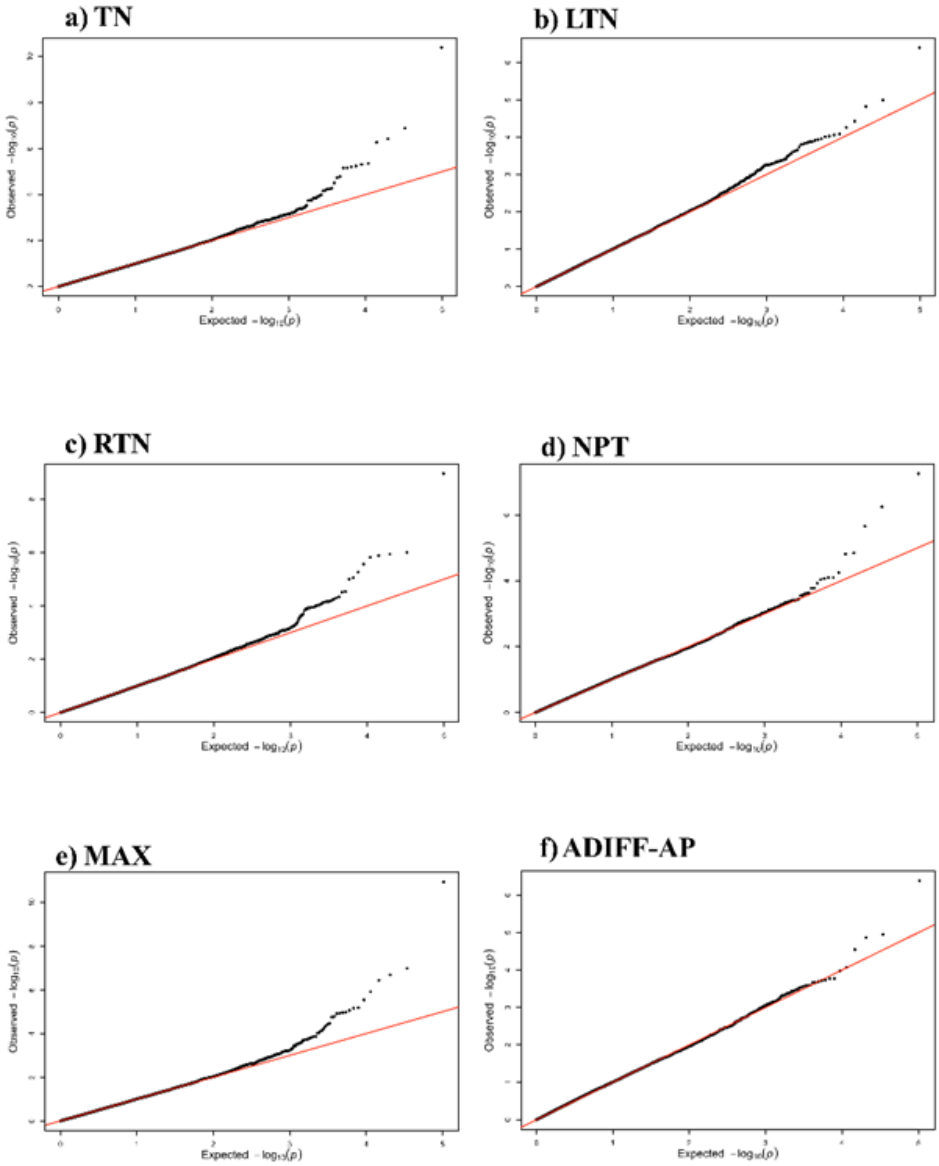


Figure S3. Manhattan plots (left) and quantile-quantile plots (right) of the four genome-wide association studies for which suggestively significant markers were identified: a) ADIFF, absolute difference between RTN and LTN (|RTL – LTN|); b) NAT, number of anterior teats; c) NPT/NAT, ratio between posterior and anterior teats; d) EX-T, extra teats.

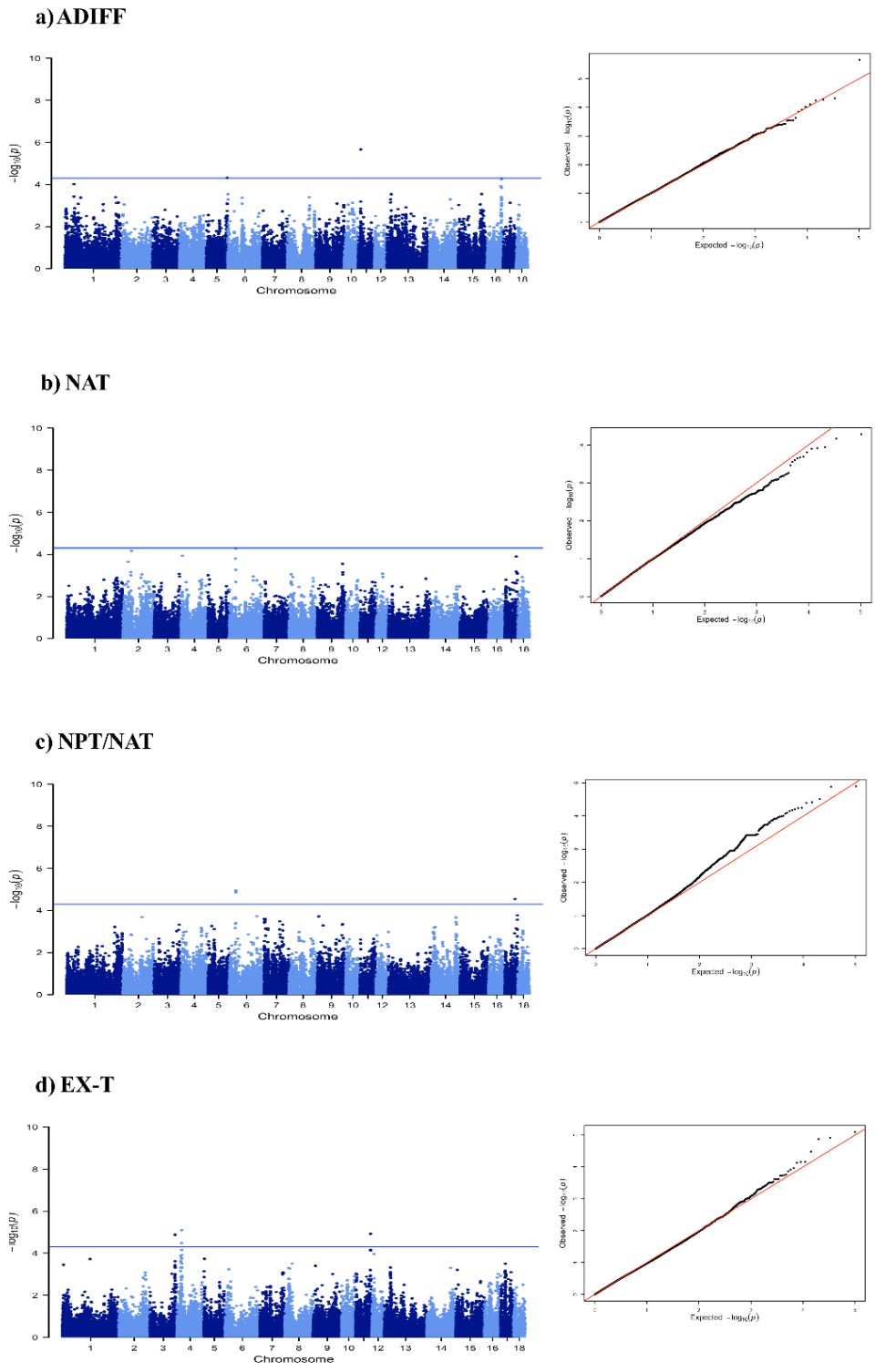


Figure S4. Manhattan plots (left) and quantile-quantile plots (right) of the three genome-wide association studies for which no significant or suggestively significant markers were identified: a) SDIFF, signed difference between RTN and LTN (RTN - LTN); b) MAX-AP, maximum number of teats between anterior and posterior parts; c) SDIFF-AP, signed difference between anterior and posterior teats (NAT - NPT).