A genome-wide association study for the number of teats in European rabbits (Oryctolagus cuniculus) identifies several candidate genes affecting this trait

SHORT COMMUNICATION

A genome-wide association study for the number of teats in European rabbits (*Oryctolagus cuniculus*) identifies several candidate genes affecting this trait

Samuele Bovo[*], Giuseppina Schiavo[*], Valerio Joe Utzeri[*], Anisa Ribani[*], Michele Schiavitto[†], Luca Buttazzoni[‡], Riccardo Negrini[¶] 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 Coniglicoltori Italiani (ANCI), Contrada Giancola snc, 71030 Volturara Appula (FG), Italy

[‡] Research Centre for Animal Production and Aquaculture (CREA), Monterotondo (Roma), Italy

[¶] Associazione Italiana Allevatori, Via G. Tomassetti 9, 00161 Roma, Italy

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

Short title: Genome-wide scan for teat number in rabbits

Summary

In the European rabbit (*Oryctolagus cuniculus*), a polytocous livestock species, the number of teats indirectly impacts the doe reproduction efficiency and, in turn, the sustainable production of rabbit meat. In this study, we carried out a genome-wide association study (GWAS) for the total number of teats in 247 Italian White does included in the Italian White rabbit breed selection program, by applying a selective genotyping approach. Does had either 8 (N = 121) or 10 teats (N = 126). All rabbits were genotyped with the Affymetrix Axiom OrcunSNP Array. Genomic data from the two extreme groups of rabbits were also analysed with the single-marker Fixation Index (F_{ST}) statistic and combined with the GWAS results. The GWAS identified 50 significant SNPs and the F_{ST} analysis identified a total of 20 SNPs that trespassed the 99.98th percentile threshold, 19 of which confirmed the GWAS results. The most significant SNP ($P = 4.31\times10^{-11}$) was located on OCU1, close to the *NUDT2* gene, a breast carcinoma cells proliferation promoter. Another significant SNP identified as candidate gene *NR6A1*, which is well known to play an important role in affecting the correlated number of vertebrae in pigs. Other significant markers were close to candidate genes involved in determining body length in mice. Markers associated with increased number of teats could be included in selection programmes to speed up the improvement for this trait in rabbit lines that need to increase maternal performances.

Keywords: GWAS; single nucleotide polymorphism; *Oryctolagus cuniculus*; reproduction; *NR6A1*.

Text

The European rabbit (*Oryctolagus cuniculus*), thereafter indicated as rabbit, is a polytocous animal that is listed within the top ten most important meat species. Rabbit meat production is a growing industry, particularly in Southern Europe and in developing countries. With an estimated number of slaughtered animals of about 24.5 million heads, Italy is among the top three rabbit meat producing countries in the world (European Union 2017).

The rabbit national selection programme in Italy involves three meat breeds, i.e. Italian White (Bianca Italiana), Italian Spotted (Macchiata Italiana) and Italian Silver (Argentata Italiana), which are described in the rabbit Herd Book that is managed by the Italian Rabbit Breeders Association (ANCI). These breeds are used in a three-way crossbreeding scheme. Within breed selection programmes are mainly based on reproductive and growth traits in the two maternal lines (Italian White and Italian Spotted) and growth efficiency and carcass traits in the paternal line (Argentata Italiana; ANCI 2010). Among the considered reproduction related traits, the number of teats is selected to maintain at least 10 functional teats in the maternal breeds.

In rabbit, as in other mammals the number of teats indirectly impacts on the reproduction efficiency and, in turn, on growth rate of suckling rabbits and on the number of weaned rabbits per litter. This trait is considered a quantitative trait partially under genetic control (polygenic inheritance), with a high level of heritability that can facilitate selection to improve doe maternal performances (Szendrő *et al.* 1991, 1992).

The number of teats in the other main polytocous livestock species, the pig, has been the matter of several recent genome-wide analyses aimed to identify chromosome regions and candidate gene variants involved in its variability and in the variability of other pleiotropic traits, mainly related to the number of vertebrae (Rohrer 2000; Mikawa *et al.* 2005, 2011; Duijvesteijn *et al.* 2014; Verando *et al.* 2015; Rohrer & Nonneman 2017; Tan *et al.* 2017; Tang *et al.* 2017; Dall'Olio *et al.* 2018; Van Son *et al.* 2019; Moscatelli *et al.* 2020). However, nothing is currently known on the genetic architecture of the number of teats in rabbits. Genomic tools and genome-wide information, like an

assembled reference genome and SNP data, have been recently produced for this species (Bertolini *et al.* 2014; Carneiro *et al.* 2014) and used to develop a commercial SNP chip panel, already applied in a few genome-wide association studies (GWAS) for relevant production traits in selected rabbit lines (e.g. Sosa *et al.* 2020a, 2020b).

In this study, we carried out a GWAS for the total number of teats in 247 Italian White rabbit does (obtained from 86 bucks) included in the Italian White breed selection program (ANCI, 2010), by applying a selective genotyping approach. Does presented either 8 (N = 121) or 10 teats (N = 126), as recorded by the visual inspection of a trained operator. These two groups represent two extremes for the number of teats that could be identified in the population. Animal with less than 10 teats are discarded from selection. Rabbits presented standard breed characteristics and were chosen by avoiding highly related individuals (no full- or half-sibs). DNA was collected by means of buccal swaps and it was extracted by using the Wizard® Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA). Animals were genotyped with the Affymetrix Axiom OrcunSNP Array (Affymetrix Inc., Santa Clara, CA, USA), which analyses a total of 199692 DNA markers, following the manufacturer's procedures. A first quality control based on the Axiom™ Analysis Suite led to discard 65320 low quality markers. PLINK v. 1.9 (Chang *et al.* 2015) was used to discard DNA markers presenting a call rate < 0.90 and a minor allele frequency (MAF) < 0.01. After filtering, a total of 247 animals and 101503 DNA markers were used in the subsequent analyses. Genome-wide association was based on a case-control study (rabbits with 8 teats vs rabbits with 10 teats, coded as 0 and 1, respectively). An additive genetic model was used to specify the dependency of the trait on genotype categories: $\boldsymbol{y} = \boldsymbol{x}\beta + \boldsymbol{g} + \boldsymbol{e}$, where $\boldsymbol{y}$ ($n \times 1$) is a vector containing the phenotype for the n^{th} animal, $\boldsymbol{x}$ ($n \times 1$) is the vector containing genotypes for the i^{th} DNA marker, β is the additive 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 Gaussian polygenic effect, with covariance matrix proportional to the genomic centered 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. The assessment of the

association was obtained by testing the null hypothesis H_0:$\beta = 0$ (Wald test). The model was fitted with GEMMA v. 0.98 (Zhou and Stephens 2012) after computation of the relatedness matrix **K** accounting for the population structure. DNA markers presenting a $P < 0.05$ (Bonferroni corrected) were considered associated. GEMMA was used to estimate the chip or SNP heritability (h^2_{SNP}). The genomic control inflation factors (λ_{GC}) was computed in R v. 3.6.0 (R Core Team, 2018). Quantile-quantile (QQ) plots and Manhattan plots were generated in R by using the *qqman* package (Turner 2018). Haploview software (Barret *et al*. 2005) was used to study the linkage disequilibrium (LD) patterns of the associated genomic regions. The proportion of variance in phenotype (PVE) explained by a given SNP was computed as described by Shim *et al*. 2015. To further evaluate the results and confirm the significant SNPs detected in GWAS, single-marker Fixation Index (F_{ST}) analysis was carried out in PLINK by comparing the two extreme groups of rabbits. We considered as outliers the SNPs presenting an F_{ST} value above the 99.98^{th} percentile of the related distribution. Then we combined the results of the two genome-wide analyses to identify significant SNPs that trespassed the thresholds in both analyses, as previously described (Schiavo *et al.* 2018, 2019). Biological annotation of DNA markers was carried out by retrieving from the OryCun2.0 NCBI's GFF file the annotated protein coding genes from a region spanning ± 0.5 Mb over the marker positions. Relevance of the genes was evaluated through Gene Cards information (https://www.genecards.org/), the NHGRI-EBI GWAS catalog (https://www.ebi.ac.uk/gwas/), the Mouse Genome Informatics database (http://www.informatics.jax.org/) and scientific literature.

Figure 1 shows the Manhattan plots obtained from the GWAS and F_{ST} analyses. The QQ-plot obtained from the GWAS is reported in Fig. S1. DNA markers that were detected simultaneously by both analyses are reported in Table 1 whereas Table S1 lists the full set of DNA markers detected within each genome scan (i.e. the GWAS and the F_{ST} analysis). The genomic control inflation factor was equal to 1.02, suggesting that population stratification in the analysed cohort did not affect the reliability of the results. Heritability of the number of teats, estimated from the genome-wide data (h^2_{SNP}) was equal to 0.64 (s.e. = 0.13). Szendrő *et al.* (1992) inferred high heritability values for teat

number in rabbits by following pedigree data with this trait recorded, even if a formal analysis was not provided. The genomic estimate here reported is higher than what was reported in a similar analysis in pigs ($h^2_{SNP} = 0.36$; Moscatelli *et al.* 2020) but it is in the range of other pedigree-based estimates obtained in several pig populations (Arakawa *et al.* 2015; Balzani *et al.* 2016; Chalkias *et al.* 2013; Rohrer & Nonneman 2017; Toro *et al.* 1986). However, we should carefully evaluate and compare these results as the two heritability estimation procedure are based on different data type (SNPs *vs* pedigree records).

The GWAS identified 50 significant SNPs located on 15 different rabbit chromosomes (OCU) and 14 genome scaffolds (Table S1). The F_{ST} analysis identified a total of 20 SNPs that trespassed the applied threshold (Table S1), 19 of which confirmed the GWAS results (Table 1). The combination of the two analyses reduced the probability to detect false positive results (associated markers) that are likely to emerge when GWAS is applied in a small population with small effective population size (Schiavo *et al.* 2018, 2020).

The most significantly associated DNA marker was AX-147080594 ($P = 4.31\times10^{-11}$), located on OCU1:19.4 Mb, within the *family with sequence similarity 219 member A* (*FAM219A*) gene. While *FAM219A* seems not linked to the phenotype (its function is not well known), this region harbours the *nudix hydrolase 2* (*NUDT2*) gene, a breast carcinoma cells proliferation promoter (Oka *et al.* 2011). Among the top associated markers, the SNP AX-147142871 ($P = 1.04\times10^{-10}$) was located on the scaffold GL018740: 0.97 Mb, in a region that contains the *ADP-ribosylation factor-like 4A* (*ARL4A*) gene, a member of the ADP-ribosylation factor family of GTP-binding proteins. Other members of this family have been linked to variation in body length (*ARL4D* in mice; Mouse Genome Informatics: 1933155) and teat number (*ARL4C* in pigs; van Son *et al.* 2019). Interestingly, another SNP (AX-147019338; GL018736: 1.8 Mb), located near the *ARL4C* gene, was significantly associated in this study with the number of teats ($P = 1.51\times10^{-08}$; Table S1) even if the F_{ST} analysis did not confirm this result as this marker was only included in the top 99.97th percentile list (F_{ST} = 0.152).

Other markers mapped close to genes that, according to the current information available on their functions, could be considered candidate genes affecting the investigated trait. For example, marker AX-147178840 (OCU6: 23.8 Mb) mapped close to the *activator of transcription and developmental regulator AUTS2* gene (*AUTS2*). Mutations in this gene have been linked to variability in body length (Mouse Genome Informatics: 1919847). Both the GWAS and F_{ST} analyses identified the DNA marker AX-147108764 (mapped in the scaffold GL018699: 2.9 Mb). This genome region harbours the *nuclear receptor subfamily 6 group A member 1* (*NR6A1*) gene, which affects the variability of the number of vertebrae in pigs and that has been also suggested to play an important role in determining the number of teats in this species (Mikawa *et al.* 2007; Ding *et al.* 2009; Duijvesteijn *et al.* 2014; Rohrer & Nonneman 2017). Mutational studies in mice linked this gene to the variability in body length (Mouse Genome Informatics: 1352459). Comparative genome analyses between the rabbit and human genomes using synteny information could place the rabbit *NR6A1* gene, at present not assembled in any chromosome in the OryCun2.0 genome version, on OCU1 (Fig. S2).

Other significant markers identified in this study were within or close to genes described to affect body length in mice (Table 1), a trait that is usually highly associated with the number of teats and the number of vertebrae in pigs (e.g. van Son *et al.* 2019). Among these genes we can mention: the *jumonji domain containing 1C* (*JMJD1C*) gene, close to the DNA marker AX-147027111 (OCU18: 22.56 Mb), the *zinc fingers and homeoboxes 3* (*ZHX3*) gene near the DNA marker AX-147118652 (GL018718: 1.02 Mb) and the *growth hormone releasing hormone* (*GHRH*) gene, close to the marker AX-147111821 (OCU4: 2.95 Mb). The related mutant phenotypes are described in the Mouse Genome Informatics database, entries MGI:1918614, MGI:2444772 and MGI:95709, respectively. Linkage disequilibrium analyses (Fig. S3) of the associated genomic regions further supported the involvement of the proposed genes except for *ZHX3* (marker AX-147027111) that seems part of a different LD block.

This study provided, for the first time, information on genetic markers associated with the number of teats in *Oryctolagus cuniculus*, a recently domesticated species (Carneiro *et al.* 2014). It will be interesting to evaluate if the domestication process and then the directional selection pressure towards more productive does could have contributed to shape the rabbit genome in regions affecting the number of teats, similarly to what may have happened for the pig genome (Rubin *et al.* 2012; Ribani *et al.* 2019; Bovo *et al.* 2020). Other morphological changes, including number of vertebrae and body length, could be also derived considering the high general correlation that might exists between these traits. These hypotheses are worthy of further investigations and other studies are needed to clarify these matters in the rabbit. Markers associated with increased number of teats could be included in selection programmes to speed up the improvements for this trait in rabbit lines that need to increase maternal performances.

Acknowledgements

The study was funded by the PSRN (Progetto di Sviluppo Rurale Nazionale) Cun-Fu project (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) and by the University of Bologna RFO 2019 programme.

Competing interests

The authors declare that they do not have any competing interests.

Availability of data

Data reported in this work can be shared after signature of an agreement on their use with University of Bologna.

References

196 ANCI (2010) Disciplinare del Libro genealogico e del Registro anagrafico della specie cunicola
197 https://www.anci-aia.it/wp-content/uploads/2019/07/Disciplinare-del-Libro-genealogico-e-
198 del-Registro-anagrafico-della-specie-cunicola-ANCI-2010.pdf
199 Arakawa A., Okumura N., Taniguchi M., *et al.* (2015) Genome-wide association QTL mapping for
200 teat number in a purebred population of Duroc pigs. *Animal Genetics* **46**, 571-5.
201 Balzani A., Cordell H.J., Sutcliffe E., Edwards S.A. (2016) Heritability of udder morphology and
202 colostrum quality traits in swine. *Journal of Animal Science* **94**, 3636-44.
203 Barrett J.C., Fry B., Maller J., Daly M.J. (2005) Haploview: analysis and visualization of LD and
204 haplotype maps. *Bioinformatics* **21**, 263-5.
205 Bertolini F., Schiavo G., Scotti E., Ribani A., Martelli P.L., Casadio R. & Fontanesi L. (2014) High
206 throughput SNP discovery in the rabbit (*Oryctolagus cuniculus*) genome by next generation
207 semiconductor based-sequencing. *Animal Genetics* **45**, 304-7.
208 Bovo S., Ribani A., Muñoz M., *et al.* (2020). Whole-genome sequencing of European autochthonous
209 and commercial pig breeds allows the detection of signatures of selection for adaptation of
210 genetic resources to different breeding and production systems. *Genetics Selection Evolution*
211 **52**, 1-19.
212 Carneiro M., Rubin C.J., Di Palma F., *et al.* (2014) Rabbit genome analysis reveals a polygenic basis
213 for phenotypic change during domestication. *Science* **345**, 1074-9.
214 Chalkias H., Rydhmer L. & Lundeheim N. (2013) Genetic analysis of functional and nonfunctional
215 teats in a population of Yorkshire pigs. *Livestock Science* **152**, 127-34.
216 Chang C.C., Chow C.C., Tellier L.C. Vattikuti S., Purcell S.M. & Lee J.J. (2015) Second-generation
217 PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7.
218 Dall'Olio S., Ribani A., Moscatelli G., Zambonelli P., Gallo M., Nanni Costa L. & Fontanesi L. (2018)
219 Teat number parameters in Italian Large White pigs: Phenotypic analysis and association with
220 vertnin (*VRTN*) gene allele variants. *Livestock Science* **210**, 68-72.

Ding N., Guo Y., Knorr C., et al. (2009) Genome-wide QTL mapping for three traits related to teat number in a White Duroc × Erhualian pig resource population. *BMC Genetics* **10**, 6.

Duijvesteijn N., Veltmaat J.M., Knol E.F. & Harlizius B. (2014) High-resolution association mapping of number of teats in pigs reveals regions controlling vertebral development. *BMC Genomics* **15**, 542.

European Union (2017) Overview Report: Commercial Rabbit Farming in the European Union. Luxembourg: Publications Office of the European Union. 16 s. DOI: 10.2772/62174. https://op.europa.eu/en/publication-detail/-/publication/5029d977-387c-11e8-b5fe-01aa75ed71a1/language-en

Mikawa S., Hayashi T., Nii M., Shimanuki S., Morozumi T. & Awata T. (2005) Two quantitative trait loci on *Sus scrofa* chromosomes 1 and 7 affecting the number of vertebrae. J*ournal of Animal Science* **83**, 2247-54.

Mikawa S., Morozumi T., Shimanuki S.-I., *et al.* (2007) Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1). *Genome Research* **17**, 586-93.

Mikawa S., Sato S., Nii M., *et al.* (2011) Identification of a second gene associated with variation in vertebral number in domestic pigs. *BMC Genetics* **12**, 5.

Moscatelli G., Dall'Olio S., Bovo S., Schiavo G., Kazemi H., Ribani A., Zambonelli P., Tinarelli S., Gallo M., Bertolini F. & Fontanesi L. (2020) Genome-wide association studies for the number of teats and teat asymmetry patterns in Large White pigs. *Animal Genetics* **51**, 595-600.

Oka K., Suzuki, T., Onodera, Y., *et al.* (2011) Nudix-type motif 2 in human breast carcinoma: a potent prognostic factor associated with cell proliferation. *International Journal of Cancer* **128**, 1770-82.

R Development Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

246 Ribani A., Utzeri V.J., Geraci C., *et al.* (2019) Signatures of de-domestication in autochthonous pig
247 breeds and of domestication in wild boar populations from *MC1R* and *NR6A1* allele
248 distribution. *Animal Genetics* **50**, 166-71.
249 Rohrer G.A. & Nonneman D.J. (2017) Genetic analysis of teat number in pigs reveals some
250 developmental pathways independent of vertebra number and several loci which only affect a
251 specific side. *Genetics Selection Evolution* **49**, 4.
252 Rohrer G.A. (2000) Identification of quantitative trait loci affecting birth characters and accumulation
253 of backfat and weight in a Meishan-White Composite resource population. *Journal of Animal*
254 *Science* **78**, 2547-53.
255 Rubin C.J., Megens H.J., Barrio A.M., *et al.* (2012). Strong signatures of selection in the domestic
256 pig genome. *Proceedings of the National Academy of Sciences of the USA* **109**, 19529-36.
257 Schiavo G., Bertolini F., Utzeri V.J., Ribani A., Geraci C., Santoro L., Óvilo C., Fernández A.I., Gallo
258 M. & Fontanesi L. (2018) Taking advantage from phenotype variability in a local animal
259 genetic resource: identification of genomic regions associated with the hairless phenotype in
260 Casertana pigs. *Animal Genetics* **49**, 321-325.
261 Schiavo G., Bovo S., Tinarelli S., Bertolini F., Dall'Olio S., Gallo M. & Fontanesi L. (2019) Genome-
262 wide association analyses for several exterior traits in the autochthonous Casertana pig breed.
263 *Livestock Science* **230**, 103842.
264 Schiavo G., Bovo S., Tinarelli S., Gallo M., Dall'Olio S. & Fontanesi L. (2020) Genome-wide
265 association analyses for coat colour patterns in the autochthonous Nero Siciliano pig breed.
266 *Livestock Science* **236**, 104015.
267 Shim H., Chasman D.I., Smith J.D., Mora S., Ridker P.M., Nickerson D.A., Krauss R.M. & Stephens
268 M. (2015) A multivariate genome-wide association analysis of 10 LDL subfractions, and their
269 response to statin treatment, in 1868 Caucasians. *PLoS One* **10**, e0120758.

Sosa-Madrid B.S., Hernández P., Blasco A., Haley C.S., Fontanesi L., Santacreu M.A., Pena R.N., Navarro P. & Ibañez-Escriche N. (2020a) Genomic regions influencing intramuscular fat in divergently selected rabbit lines. *Animal Genetics* **51**, 58-69.

Sosa-Madrid B.S., Santacreu M.A., Blasco A., Fontanesi L., Pena R.N. & Ibáñez-Escriche N. (2020b) A genome-wide association study in divergently selected lines in rabbits reveals novel genomic regions associated with litter size traits. *Journal of Animal Breeding and Genetics* **137**, 123-38.

Szendrő Zs., Mohamed M.M.A. & Biró-Németh E. (1991) Teat number of new-born rabbits depending on the teat number of their parents. *Journal of Applied Rabbit Research* **14**, 133-5.

Szendrő Zs., Mohamed M.M.A., Biró-Németh E. & Radnai, I. (1992) Heritability of teat number of rabbits. *Journal of Applied Rabbit Research* **15**, 174-80.

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. *Genetics Selection Evolution* **49**, 35.

Tang J., Zhang Z., Yang B., *et al.* (2017) Identification of loci affecting teat number by genome-wide association studies on three pig populations. *Asian-Australasian Journal of Animal Science* **30**, 1-7.

Toro M.A., Dobao M.T., Rodriganez J. & Silio L. (1986) Heritability of a canalized trait: teat number in Iberian pigs. *Genetics Selection Evolution* **18**, 173-84.

Turner S.D. (2014) qqman: an R package for visualizing GWAS results using QQ and Manhattan plots. Biorxiv 005165.

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.

296 Zhou X. & Stephens M. (2012) Genome-wide efficient mixed-model analysis for association studies.
297 *Nature Genetics* **44**, 821-4.
298

Table 1. List of single nucleotide polymorphisms (SNPs) associated with the number of teats in rabbits and detected in both GWAS ($P < 0.05$, Bonferroni corrected) and F_{ST} analysis (99.98th percentile). Data are ordered by *P* of association. The full lists of significant markers identified in the GWAS and all SNPs of the 99.98th percentile in the F_{ST} analysis are given in Table S1.

Marker[1]	**OCU**[2]	**Pos**[3]	**Min**[4]	**Maj**[5]	**MAF**[6]	***β***[7]	***P***[8]	**F_{ST}**[9]	**PVE**[10]	**Genes**[11]
AX-147080594	1	19404556	C	A	0.109	0.512	4.31×10^{-11}	0.186	16.2	*DCAF12; DCTN3;* ***FAM219A****; DNAI1; UBAP2; UBAP1; LOC100347619; CCL21; CCL27; GALT; LOC108176349; NUDT2; RPP25L; FAM205C; LOC100347365; CNTFR; PHF24; KIAA1161; ARID3C; CCL19; ENHO; C1H9orf24; KIF24; IL11RA; SIGMAR1*
AX-147173940	1	20900106	T	C	0.099	0.538	5.67×10^{-11}	0.177	16.0	*DDX58; NDUFB6; TMEM215; ACO1; APTX; TOPORS*
AX-147162558	7	68775751	T	G	0.099	0.532	5.81×10^{-11}	0.176	16.0	*LOC100347899; ACTR3*
AX-147142871	GL018740	965628	T	C	0.109	0.497	1.04×10^{-10}	0.177	15.6	*LOC103352516; LOC100356405; LOC100356924; LOC108175498; PPP2R3B; LOC100356149; LOC100357437; LOC100355902; CRLF2;* ***ARL4A***
AX-147036777	19	7022754	T	C	0.105	0.513	1.26×10^{-10}	0.176	15.4	*ARHGAP44; MYOCD; ELAC2*
AX-147178840	6	23830178	G	A	0.114	0.451	1.49×10^{-10}	0.192	15.4	***AUTS2****; WBSCR17*
AX-147037069	3	40049384	G	T	0.095	0.532	1.61×10^{-10}	0.169	15.3	*EBF1*
AX-147108764	GL018699	2909467	G	A	0.092	0.532	2.33×10^{-10}	0.168	15.0	*NEK6; OLFML2A; ARPC5L;* ***NR6A1****; RPL35; ADGRD2; LOC103352017; DENND1A; GOLGA1; LOC100349446; NR5A1; LHX2; WDR38*
AX-147033079	GL018755	1507787	T	C	0.104	0.492	4.02×10^{-10}	0.174	14.7	*PPP1R3D; SYCP2; PHACTR3; FAM217B*
AX-147027111	18	22564826	G	T	0.100	0.505	4.41×10^{-10}	0.163	14.6	*REEP3; NRBF2;* ***JMJD1C***
AX-147154664	13	59533425	T	G	0.112	0.461	5.32×10^{-10}	0.174	14.5	*LOC103350105*
AX-147118652	GL018718	1022162	A	G	0.092	0.524	5.54×10^{-10}	0.168	14.4	*EMILIN3; LPIN3; CHD6;* ***ZHX3****; TOP1; PLCG1*
AX-147114988	3	25019807	G	C	0.103	0.483	8.60×10^{-10}	0.171	14.2	*NR3C1; ARHGAP26; FGF1*
AX-146993766	1	6087385	C	T	0.103	0.464	1.00×10^{-09}	0.169	14.1	*LOC103347974; LOC100357282; KLF4; RAD23B; LOC108178114*
AX-147144556	GL018864	275213	G	A	0.100	0.489	1.14×10^{-09}	0.163	14.0	*BRI3BP; LOC100343523; LOC100342501; SCARB1; AACS; TMEM132B; DHX37*
AX-147126930	9	35546563	T	C	0.087	0.479	1.45×10^{-09}	0.171	13.8	*FAM19A1*
AX-147161194	GL018803	762803	A	G	0.097	0.494	1.48×10^{-09}	0.160	13.8	*IAH1; CPSF3; YWHAQ; KIDINS220; MBOAT2; ADAM17; ITGB1BP1; ASAP2*
AX-147111821	4	2949855	T	G	0.083	0.509	3.45×10^{-09}	0.159	13.2	*SRC; RPN2;* ***GHRH****; SAMHD1; RBL1; LOC103348200; DSN1; TLDC2; BLCAP; SLA2; LOC100342874;*

										CTNNBL1; MANBAL; MROH8; NNAT; NDRG3; TGIF2; SOGA1; MYL9
AX-147105176	GL018704	479221	A	G	0.121	0.417	8.29×10^{-09}	0.170	12.6	*GNL2; POU3F1; LOC100356330; RSPO1; LOC103352105; FHL3; MTF1; LOC100353815; YRDC; INPP5B; DNALI1; UTP11; SNIP1; SF3A3; LOC100345534; CDCA8; LOC103352122; MANEAL; MEAF6; LOC100349353*

[1] SNP identifier used by the Affymetrix Axiom OrcunSNP Array.

[2] *Oryctolagus cuniculus* chromosome.

[3] Position in base pairs on the *O. cuniculus* reference genome (OryCun2.0).

[4] Minor allele.

[5] Major allele.

[6] Minor allele frequency.

[7] Regression coefficient.

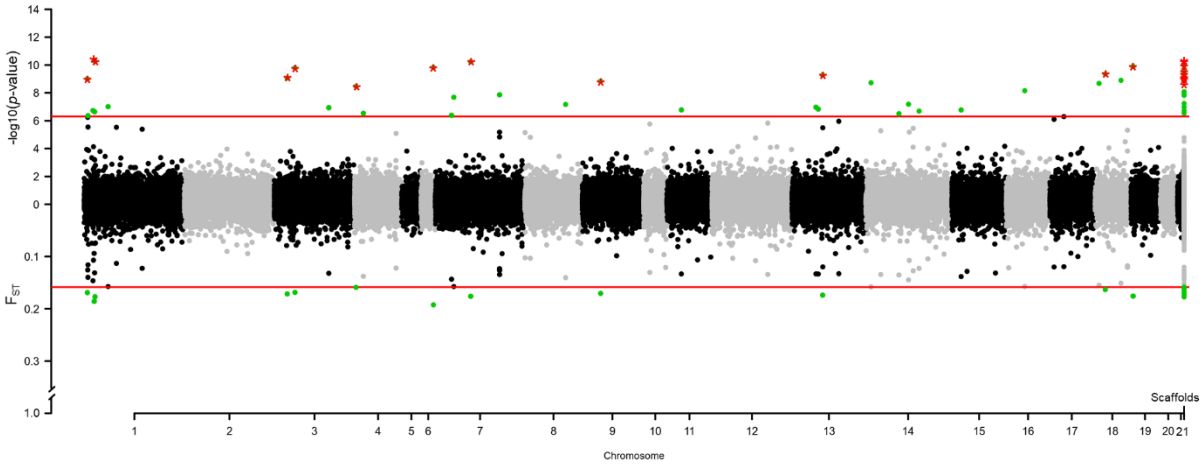
[8] *P* at the Wald test (GEMMA) of the GWAS.

[9] Fixation index value.

[10] Percentage phenotypic variation explained.

[11] Genes overlapping the SNP position ± 0.5 Mb as annotated in the OryCun2.0 genome version. In bold are reported the candidate genes discussed in the text.

Figure 1. Manhattan plots of the GWAS (top) and the F_{ST} analysis (bottom) for the number of teats in rabbits. Each dot represents a single nucleotide polymorphism (SNP). Scaffolds are grouped tighter at the end of the plot. Reds lines represent the thresholds in GWAS ($P < 0.05$, Bonferroni corrected) and F_{ST} analysis (99.98th percentile). Green dots represent significant SNPs. Red stars mark SNPs detected in GWAS and confirmed by the F_{ST} analysis.



Legends to Supplementary Material

Figure S1. Quantile-Quantile-plot of the genome-wide association study.

Figure S2. Schematic representation of the human chromosome 9 region containing the *NR6A1* gene and the position of the syntenic regions in the rabbit genome (OCU1 and scaffold GL0118699, where the significant SNP close to the rabbit *NR6A1* gene is located).

Figure S3. Pairwise linkage disequilibrium (LD) analysis of genomic regions carrying candidate genes for the number of teats in rabbits. LD was measured as r^2 and it is presented in each box coloured considering the magnitude of linkage. The associated SNP is highlighted with a red star symbol. The proposed candidate gene is also reported and its position is marked with a blue line. A) DNA marker AX-147080594. B) DNA marker AX-147142871 C) DNA marker AX-147178840. D) DNA marker AX-147108764. E) DNA marker AX-147027111. F) DNA marker AX-147118652. G) DNA marker AX-147111821.

Table S1. List of single nucleotide polymorphisms (SNPs) detected in the GWAS ($P < 0.05$, Bonferroni corrected) and F_{ST} analysis (99.98th percentile) for the number of teats in rabbits. Data are ordered by *P* of association.

Supplementary material

A genome-wide association study for the number of teats in European rabbits (*Oryctolagus cuniculus*) identifies several candidate genes affecting this trait

Samuele Bovo[*], Giuseppina Schiavo[*], Valerio Joe Utzeri[*], Anisa Ribani[*], Michele Schiavitto[†], Luca Buttazzoni[‡], Riccardo Negrini[¶] 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 Coniglicoltori Italiani (ANCI), Contrada Giancola snc, 71030 Volturara Appula (FG), Italy

[‡] Research Centre for Animal Production and Aquaculture (CREA), Monterotondo (Roma), Italy

[¶] Associazione Italiana Allevatori, Via G. Tomassetti 9, 00161 Roma, Italy

Figure S1. Quantile-Quantile-plot of the genome-wide association study.

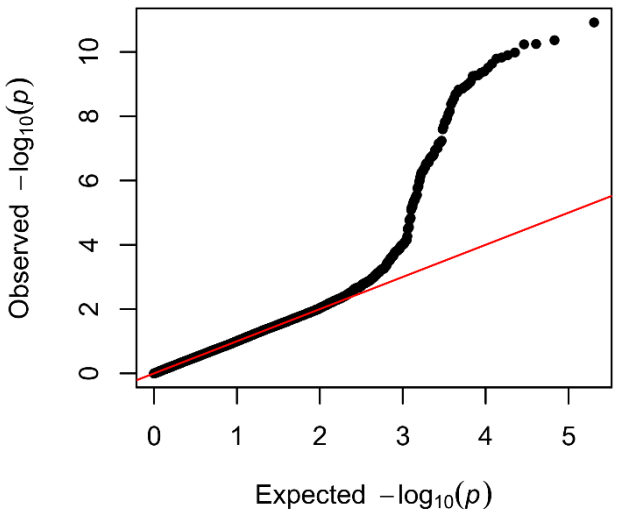


Figure S2. Schematic representation of the human chromosome 9 region containing the *NR6A1* gene and the position of the syntenic regions in the rabbit genome (OCU1 and scaffold GL0118699, where the significant SNP close to the rabbit *NR6A1* gene is located).

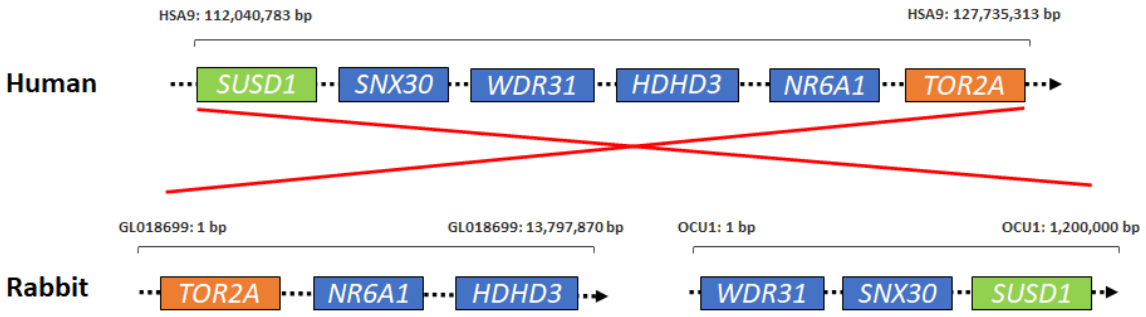


Figure S3. Pairwise linkage disequilibrium (LD) analysis of genomic regions carrying candidate genes for the number of teats in rabbits. LD was measured as r^2 and it is presented in each box coloured considering the magnitude of linkage. The associated SNP is highlighted with a red star symbol. The proposed candidate gene is also reported and its position is marked with a blue line. A) DNA marker AX-147080594. B) DNA marker AX-147142871 C) DNA marker AX-147178840. D) DNA marker AX-147108764. E) DNA marker AX-147027111. F) DNA marker AX-147118652. G) DNA marker AX-147111821.

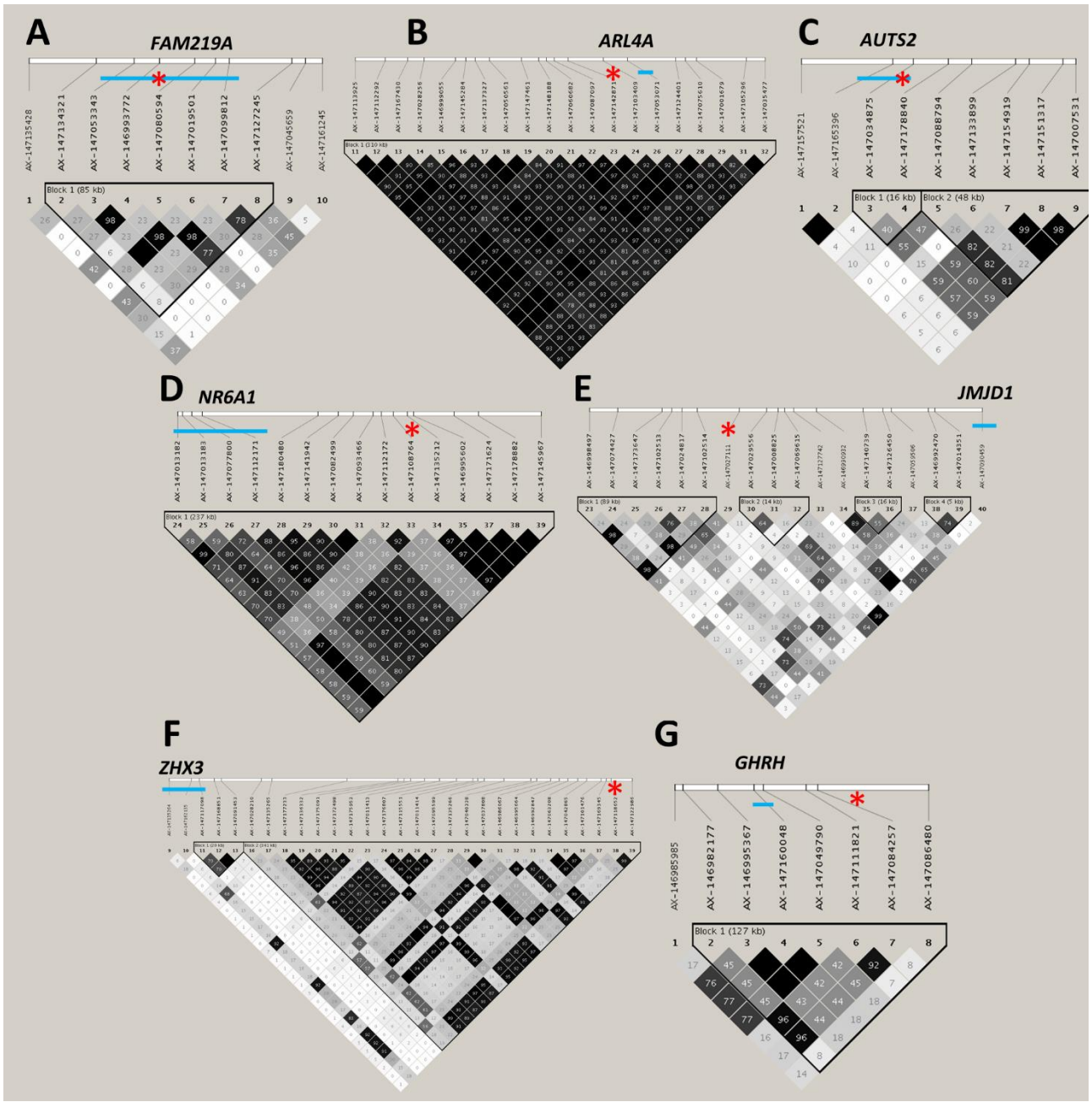


Table S1. List of single nucleotide polymorphisms (SNPs) detected in the GWAS ($P < 0.05$, Bonferroni corrected) and F_{ST} (99.98th percentile) analyses for the number of teats in rabbits. Data are ordered by *P* of association.

Marker[1]	OCU[2]	Pos[3]	Min/Maj[4]	MAF[5]	Beta[6]	*P*[7]	F_{ST}[8]	PVE[9]	Analysis[10]	Genes[11]
AX-147180998	GL019669	6241	C/T	0.096	-0.526	1.21E-11	0.126	17.0	GWAS	*DCAF12, DCTN3, FAM219A, DNAI1, UBAP2, UBAP1, LOC100347619, CCL21, CCL27, GALT, LOC108176349, NUDT2, RPP25L, FAM205C, LOC100347365, CNTFR, PHF24, KIAA1161, ARID3C, CCL19, ENHO, C1H9orf24, KIF24, IL11RA, SIGMAR1*
AX-147080594	1	19404556	C/A	0.109	0.512	4.31E-11	0.186	16.2	GWAS- F_{ST}	*DTD1, SCP2D1, SLC24A3*
AX-147173940	1	20900106	T/C	0.099	0.538	5.67E-11	0.177	16.0	GWAS- F_{ST}	*UBAP2, DNAJA1, AQP3, AQP7, LOC103345142, NFX1, DCAF12, APTX, BAG1, SMU1, TMEM215, B4GALT1, NOL6, SPINK4, UBE2R2, CHMP5*
AX-147162558	7	68775751	T/G	0.099	0.532	5.81E-11	0.176	16.0	GWAS- F_{ST}	*LSAMP, GAP43, LOC100339409*
AX-147142871	GL018740	965628	T/C	0.109	0.497	1.04E-10	0.177	15.6	GWAS- F_{ST}	*APC, SRP19, NREP, REEP5, EPB41L4A*
AX-147036777	19	7022754	T/C	0.105	0.513	1.26E-10	0.176	15.4	GWAS- F_{ST}	*BACE2, LOC103346181*
AX-147178840	6	23830178	G/A	0.114	0.451	1.49E-10	0.192	15.4	GWAS- F_{ST}	*MECOM, LRRIQ4, LRRC34, MYNN, ACTRT3*
AX-147037069	3	40049384	G/T	0.095	0.532	1.61E-10	0.169	15.3	GWAS- F_{ST}	*TSHZ2, ZNF217, BCAS1*
AX-147108764	GL018699	2909467	G/A	0.092	0.532	2.33E-10	0.168	15.0	GWAS- F_{ST}	*SCEL, SLAIN1, EDNRB, MYCBP2*
AX-147087265	23	28420299	G/T	0.082	0.539	3.10E-10	NA	14.9	GWAS	*FOXP2, LOC103348720, PPP1R3A, LOC100346274*
AX-147033079	GL018755	1507787	T/C	0.104	0.492	4.02E-10	0.174	14.7	GWAS- F_{ST}	*GRHPR, ZBTB5, FBXO10, PAX5, MELK, POLR1E, ZCCHC7*
AX-147027111	18	22564826	G/T	0.1	0.505	4.41E-10	0.163	14.6	GWAS- F_{ST}	*B3GNT5, LOC100353443, ST6GAL1, SST, MCF2L2, MASP1, LOC108177987, RTP4, RTP2, ADIPOQ*
AX-147154664	13	59533425	T/G	0.112	0.461	5.32E-10	0.174	14.5	GWAS- F_{ST}	*FREM3, USP38, SMARCA5, GAB1*
AX-147118652	GL018718	1022162	A/G	0.092	0.524	5.54E-10	0.168	14.4	GWAS- F_{ST}	*TMEM261*
AX-147070994	23	92444500	T/C	0.095	0.514	5.62E-10	NA	14.5	GWAS	*TMEM38B, SLC44A1, FKTN, FSD1L, TAL2*
AX-147114988	3	25019807	G/C	0.103	0.483	8.60E-10	0.171	14.2	GWAS- F_{ST}	*LOC108177271, FER, FBXL17*
AX-146993766	1	6087385	C/T	0.103	0.464	1.00E-09	0.169	14.1	GWAS- F_{ST}	*SPAG17, WDR3, GDAP2, MAN1A2, FAM46C*
AX-147144556	GL018864	275213	G/A	0.100	0.489	1.14E-09	0.163	14.0	GWAS- F_{ST}	*LOC100351959, RPS5, CHMP2A, LOC103345283, LOC100351453, MZF1, RNF225, A1BG, ZNF324, TRIM28, LOC100353729, ZNF329, LOC100352209, LOC100352972, LOC108175676, ZNF584, SLC27A5, LOC108175679, LOC108175678,*

										ZNF446, UBE2M, ZSCAN18, ORYCUNV1R1556, ZBTB45, LOC100352717, ZNF274, ZNF132, ZNF135
AX-147134772	18	52989526	A/G	0.085	0.525	1.26E-09	0.151	13.9	GWAS	*IAH1, CPSF3, YWHAQ, KIDINS220, MBOAT2, ADAM17, ITGB1BP1, ASAP2*
AX-146985826	23	29778013	G/A	0.084	0.522	1.40E-09	NA	13.8	GWAS	*ISCU, TMEM119, DAO, LOC103345380, SELPLG, CORO1C, SVOP, FICD, SSH1, SART3, CMKLR1*
AX-147126930	9	35546563	T/C	0.087	0.479	1.45E-09	0.171	13.8	GWAS- F_{ST}	*PHTF1, LOC103350094, LRIG2, SLC16A1, PPM1J, RSBN1, MAGI3, LOC108177607, FAM19A3*
AX-147161194	GL018803	762803	A/G	0.097	0.494	1.48E-09	0.160	13.8	GWAS- F_{ST}	*GNAT3, GNAI1*
AX-147161353	14	10088641	A/G	0.092	0.502	1.88E-09	0.158	13.6	GWAS	*HPCAL1, PDIA6, KCNF1, NOL10, RRM2, PQLC3, ATP6V1C2, ROCK2, CYS1, ODC1*
AX-147081568	18	10464751	C/T	0.086	0.521	2.08E-09	0.155	13.5	GWAS	*KCNMA1*
AX-147111821	4	2949855	T/G	0.083	0.509	3.45E-09	0.159	13.2	GWAS- F_{ST}	*HJURP, MROH2A, SH3BP4, TRPM8, SPP2, ARL4C*
AX-147059338	16	35713998	A/C	0.086	0.497	6.96E-09	0.157	12.7	GWAS	*SRC, RPN2, GHRH, SAMHD1, RBL1, LOC103348200, DSN1, TLDC2, BLCAP, SLA2, LOC100342874, CTNNBL1, MANBAL, MROH8, NNAT, NDRG3, TGIF2, SOGA1, MYL9*
AX-147105176	GL018704	479221	A/G	0.121	0.417	8.29E-09	0.170	12.6	GWAS- F_{ST}	*TMEM64, OSGIN2, CALB1, RIPK2, NBN, DECR1*
AX-147121683	GL018795	362725	T/C	0.115	0.455	1.11E-08	0.144	12.4	GWAS	*SORCS3*
AX-147007632	7	1.25E+08	A/G	0.082	0.5	1.37E-08	0.134	12.3	GWAS	-
AX-147019338	GL018736	1822716	C/T	0.084	0.48	1.51E-08	0.152	12.2	GWAS	*NKIRAS1, LOC103350360, THRB, RPL15, NR1D2, LOC100354960*
AX-147047442	7	35911714	G/A	0.091	0.469	2.07E-08	0.157	12.0	GWAS	*GNL2, POU3F1, LOC100356330, RSPO1, LOC103352105, FHL3, MTF1, LOC100353815, YRDC, INPP5B, DNALI1, UTP11, SNIP1, SF3A3, LOC100345534, CDCA8, LOC103352122, MANEAL, MEAF6, LOC100349353*
AX-147124436	GL018763	1423677	T/C	0.097	0.446	5.79E-08	0.132	11.3	GWAS	*LOC103352516, LOC100356405, LOC100356924, LOC108175498, PPP2R3B, LOC100356149, LOC100357437, LOC100355902, CRLF2, ARL4A*
AX-147048907	14	82851725	C/T	0.093	0.424	6.49E-08	0.145	11.2	GWAS	-
AX-147052744	8	79279002	A/G	0.072	0.496	6.73E-08	0.14	11.1	GWAS	*LOC103347974, LOC100357282, KLF4, RAD23B, LOC108178114*
AX-147028523	1	46414727	T/C	0.108	0.343	9.80E-08	0.157	10.9	GWAS	*MAP10, DISC1, SIPA1L2*
AX-146982385	GL018712	2963020	G/A	0.076	0.483	1.02E-07	0.139	10.9	GWAS	-
AX-147087748	13	46528135	T/C	0.110	0.396	1.07E-07	0.133	10.8	GWAS	*AUTS2, WBSCR17*
AX-147037138	3	1.06E+08	T/C	0.063	0.514	1.15E-07	0.132	10.8	GWAS	*ARHGAP44, MYOCD, ELAC2*
AX-147171391	13	50949460	G/C	0.083	0.465	1.46E-07	0.133	10.6	GWAS	*NEK6, OLFML2A, ARPC5L, NR6A1, RPL35, ADGRD2, LOC103352017, DENND1A, GOLGA1, LOC100349446, NR5A1, LHX2, WDR38*

AX-147028777	11	28090766	T/C	0.079	0.465	1.67E-07	0.133	10.5	GWAS	*EBF1*
AX-147005356	15	21130687	A/G	0.078	0.462	1.68E-07	0.138	10.5	GWAS	*EMILIN3, LPIN3, CHD6, ZHX3, TOP1, PLCG1*
AX-147156218	1	16996176	T/C	0.098	0.405	1.86E-07	0.146	10.4	GWAS	*FAM19A1*
AX-147080455	GL018832	683056	G/A	0.087	0.441	1.87E-07	0.138	10.5	GWAS	*LOC103350105*
AX-147036367	14	1.03E+08	A/G	0.060	0.502	2.03E-07	0.123	10.4	GWAS	*ADAMTS13, CACFD1*
AX-147150946	1	20183530	A/G	0.073	0.468	2.28E-07	0.131	10.3	GWAS	*LOC100347899, ACTR3*
AX-147080391	GL018744	1600109	A/G	0.085	0.445	2.83E-07	0.127	10.1	GWAS	*DDX58, NDUFB6, TMEM215, ACO1, APTX, TOPORS*
AX-147059974	4	17266266	C/A	0.073	0.475	2.92E-07	0.138	10.1	GWAS	*REEP3, NRBF2, JMJD1C*
AX-147078864	14	64269915	G/A	0.114	0.396	3.12E-07	0.135	10.1	GWAS	*PPP1R3D, SYCP2, PHACTR3, FAM217B*
AX-147032537	7	31518328	C/T	0.100	0.370	4.10E-07	0.143	9.9	GWAS	*BRI3BP, LOC100343523, LOC100342501, SCARB1, AACS, TMEM132B, DHX37*
AX-147167042	1	7235487	T/C	0.087	0.400	4.35E-07	0.140	9.8	GWAS	*NR3C1, ARHGAP26, FGF1*
AX-146999069	GL018777	142249	T/C	0.172	0.238	2.98E-05	0.159	6.8	F_{ST}	*ZNF804A*

[1] SNP identifier used by the Affymetrix Axiom OrcunSNP Array;
[2] *Oryctolagus cuniculus* chromosome;
[3] Position, in basepairs, on the *O. cuniculus* reference genome (OryCun2.0);
[4] Minor/Major alleles;
[5] Minor allele frequency;
[6] Regression effect;
[7] *P* at the Wald test (GEMMA) of the GWAS.
[8] Fixation Index value;
[9] Percentage phenotypic variation explained;
[10] Analysis;
[11] Genes overlapping the SNP position ± 0.5 Mb