

Genome-wide association study for reproductive traits in a Large White pig population

Y. Wang^{*1}, X. Ding^{*1}, Z. Tan^{*}, K. Xing^{*}, T. Yang^{*}, Y. Pan[†], Y. Wang[†], S. Mi[‡], D. Sun^{*} and C. Wang^{*}

^{*}Key Laboratory of Animal Genetics and Breeding of Ministry of Agriculture, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100094, China. [†]Beijing Shunxin Agriculture Co., Ltd., Beijing 101300, China. [‡]Beijing Liuma Pig Breeding Technology Co., Ltd., Beijing 101308, China.

Summary

Using the PorcineSNP80 BeadChip, we performed a genome-wide association study for seven reproductive traits, including total number born, number born alive, litter birth weight, average birth weight, gestation length, age at first service and age at first farrowing, in a population of 1207 Large White pigs. In total, we detected 12 genome-wide significant and 41 suggestive significant SNPs associated with six reproductive traits. The proportion of phenotypic variance explained by all significant SNPs for each trait ranged from 4.46% (number born alive) to 11.49% (gestation length). Among them, 29 significant SNPs were located within known QTL regions for swine reproductive traits, such as corpus luteum number, stillborn number and litter size, of which one QTL region associated with litter size contained the *ALGA0098819* SNP for total number born. Subsequently, we found that 376 functional genes contained or were near these significant SNPs. Of these, 14 genes—*BHLHA15*, *OCM2*, *IL1B2*, *GCK*, *SMAD2*, *HABP2*, *PAQR5*, *GRB10*, *PRELID2*, *DMKN*, *GPI*, *GPIHBP1*, *ADCY2* and *ACVR2B*—were considered important candidates for swine reproductive traits based on their critical roles in embryonic development, energy metabolism and growth development. Our findings contribute to the understanding of the genetic mechanisms for reproductive traits and could have a positive effect on pig breeding programs.

Keywords candidates, genetic mechanism, GWAS, single nucleotide polymorphism, phenotypic variance

In pig production systems, reproductive traits, such as total number born (TNB), number born alive (NBA), litter birth weight (LBW), average birth weight (ABW), gestation length (GL), age at first service (AFS) and age at first farrowing (AFF), play key roles in production efficiency and economic profits. Traditional breeding technologies based on best linear unbiased prediction (Holm *et al.* 2004) is limited for significant genetic improvement of these reproductive traits due to their low heritability (Chen *et al.* 2003). In addition, piglet records can be collected for sows only later in life. Thus, for improved breeding programs, it is essential to better understand the genetic determinants of

these traits. Candidate genes and quantitative trait loci (QTL) for reproductive traits were identified in previous studies (Onteru *et al.* 2009). Until now, a total of 405 QTL had been found on different swine chromosomes for TNB, NBA, LBW, ABW and GL (<http://www.animalgenome.org/cgi-bin/QTLdb/SS/index>, Apr 27, 2017).

Compared with previous methods, genome-wide association studies (GWASs) provide a more powerful strategy for genetic dissection and have been performed with domestic animals for various economic traits (Zhang *et al.* 2013). To date, several GWASs for pig reproductive traits, such as number of teats (Verardo *et al.* 2016), litter size (Sell-Kubiak *et al.* 2015) and ovulation rate (He *et al.* 2017), have been conducted, but knowledge about the complex genetic mechanism for reproductive performance still remains insufficient. Therefore, the aim of this study was to detect novel genetic variants and identify candidate genes for reproductive traits by performing a GWAS in a Large White pig population.

A total of 1207 Large White pigs from the Beijing Shunxin Agriculture Co., Ltd. and Beijing Liuma Pig Breeding Technology Co., Ltd. pig breeding farms (Beijing,

Address for correspondence

C. Wang and D. Sun, Key Laboratory of Animal Genetics and Breeding of Ministry of Agriculture, National Engineering Laboratory of Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100094, China.
E-mails: cdwang@cau.edu.cn (CW); sundx@cau.edu.cn (DS)

¹Yuan Wang and Xiangdong Ding contributed equally to this work

Accepted for publication 27 November 2017

China) were genotyped using the GeneSeek PorcineSNP80 BeadChip. Using PLINK software (Purcell *et al.* 2007), DNA samples with genotyping call rates less than 90% were removed, and we also excluded SNPs with call rates less than 90%, minor allele frequencies less than 0.03 or Hardy–Weinberg equilibrium P -value $< 1.00E-06$ in SNPs with no position information or located on sex chromosomes were also excluded from the dataset. Missing genotypes were imputed using BEAGLE v4.0 (Browning & Browning 2009) based on information from remaining SNP genotypes using, and SNPs with $R^2 > 0.3$ (Browning & Browning 2007) were retained. After quality control, 1198 individuals and 51 443 SNPs qualified for the study.

The phenotypic data for first parity were collected from 2010 to 2015, and statistics for the phenotypes of seven reproductive traits were calculated (Table S1). TNB, NBA, LBW, ABW and GL were approximately normally distributed, but for AFS and AFF, phenotype values were converted using the `rntransform` function in the GENABEL R package (Aulchenko *et al.* 2007).

To better control population structure, we first conducted a principal components analysis to reduce spurious associations derived from the presence of individual relatedness. All autosomal SNPs were pruned using the `indep-pairwise` option in PLINK software, using a window size of 25 SNPs, a step of five SNPs and an r^2 threshold of 0.2 (Gu *et al.* 2011), which resulted in 6993 independent SNPs.

The genome-wide association study was implemented with a mixed model approach using GEMMA software (Zhou & Stephens 2012) for each trait in the first parity. The centered relatedness matrix was calculated using all autosomal markers, and a Wald test was performed for each SNP against the null hypothesis of $g = 0$. The statistical model used is as follows:

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

where \mathbf{y} is an $n \times 1$ vector of phenotype values for all individuals; \mathbf{W} is an $n \times c$ matrix of covariates (fixed effects that contain the first PC, herd, farrowing season and a column of 1s); $\boldsymbol{\alpha}$ is a $c \times 1$ vector of the corresponding coefficients including the intercept; \mathbf{x} is an $n \times 1$ vector of genotypes of a marker at the locus tested; β is the effect size of the marker; \mathbf{u} is an $n \times 1$ vector of random polygenic effects with a covariance structure as $\boldsymbol{\varepsilon} \sim N(0, \mathbf{K}V_g)$, where \mathbf{K} is an $n \times n$ genetic relatedness matrix and V_g is the polygenic additive variance; and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of residual errors with $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}V_e)$, where \mathbf{I} is $n \times n$ identity matrix and V_e is the residual variance component.

To properly decide the thresholds for genome-wide significant/suggestive associations, we calculated the effectively independent tests based on the estimated number of independent markers and linkage disequilibrium blocks for autosome markers (Nicodemus *et al.* 2005). A linkage disequilibrium block was defined as a set of adjacent SNPs with pairwise r^2 values greater than 0.40 (Gu *et al.* 2011). A total of 11 315 effectively independent tests was suggested, following Lander & Kruglyak (1995), and the threshold P -value for genome-wide significance association was $4.42E-6$ ($0.05/11\ 315$) and for suggestive association was $8.84E-5$ ($1/11\ 315$). The genomic inflation factor λ was calculated using GENABEL packages. In addition, GCTA software (Yang *et al.* 2011) was used to calculate phenotypic variances explained by significant SNPs.

The functional genes containing or near the identified significant SNPs, less than 1 Mb away from significant SNPs, were selected based on the *Sus scrofa* 10.2 genome assembly. Further functional annotation was carried out based on the NCBI database (<https://www.ncbi.nlm.nih.gov/>), and Gene Ontology analysis was conducted using DAVID Bioinformatics Resources (<http://david.abcc.ncifcrf.gov/>).

Table 1 Genome-wide significant SNPs for five reproductive traits.

Trait	SNP	Chr	Position (bp)	P -value	MAF	β	CPV% (SE)	Nearest gene/ candidate gene ¹	Location (bp) ²
TNB	WU_10.2_2_162527469	2	162 527 469	9.72E-07	0.27 (C/A)	-0.59	2.18 (0.03)	IFITM2	174 050
TNB	WU_10.2_3_44631648	3	44 631 648	1.19E-06	0.31 (G/A)	-0.55	2.09 (0.03)	BARX1/IL1B2	23 864/53 8271
TNB	WU_10.2_3_44862084	3	44 862 084	2.83E-06	0.27 (G/A)	-0.57	2.01 (0.03)	BARX1/IL1B2	254 300/307 835
TNB	ALGA0098819	18	56 535 534	3.46E-06	0.17 (A/G)	-0.62	1.72 (0.03)	LOC102165380/GCK	within/617 340
NBA	WU_10.2_3_44631648	3	44 631 648	9.07E-07	0.31 (G/A)	-0.53	2.08 (0.03)	BARX1/IL1B2	23 864/538 271
NBA	ALGA0098819	18	56 535 534	3.10E-06	0.17 (A/G)	-0.61	1.79 (0.03)	LOC102165380/GCK	within/617 340
NBA	WU_10.2_2_162527469	2	162 527 469	4.42E-06	0.27 (C/A)	-0.53	1.93 (0.03)	IFITM2	174 050
GL	ALGA0061535	11	25 305 148	1.42E-06	0.05 (A/G)	-1.04	2.26 (0.03)	AKAP11	99 469
GL	WU_10.2_4_1247716	4	1 247 716	3.32E-06	0.16 (A/G)	-0.66	1.93 (0.03)	ZC3H3/GPIHBP1	14 516/96 877
GL	ASGA0023643	4	1 166 037	3.87E-06	0.16 (A/G)	-0.66	1.88 (0.03)	MAFA	7303
AFS	ALGA0111336	16	80 445 920	2.23E-06	0.38 (A/C)	-0.22	2.26 (0.03)	ADCY2	within
AFF	ALGA0111336	16	80 445 920	3.53E-06	0.38 (A/C)	-0.22	2.19 (0.03)	ADCY2	within

Chr, swine Chromosome; MAF, allele frequency of first listed marker; β , allele substitution effect; CPV% (SE), contribution to phenotypic variance (standard error); TNB, total number born; NBA, number born alive; GL, gestation length; AFS, age at first service; AFF, age at first farrowing.

¹Gene names in bold type represent candidate genes with less than 1.0 Mb of the SNPs.

²Locations in bold type represent the distance between a significant SNP and the candidate gene.

As a result, 53 SNPs, including 12 genome-wide significant (Table 1) and 41 suggestive significant (Table S2) SNPs, were detected for TNB, NBA, LBW, GL, AFS and AFF on SSC1, 2, 3, 4, 5, 6, 9, 10, 11, 13, 14, 15, 16 and 18 (there were no significant SNPs detected for LBW). The lambda values were 1.01, 1.02, 1.02, 1.01, 1.01 and 1.01 for each trait respectively, which meant lower population stratification (Price *et al.* 2010). Manhattan plot and quantile–quantile plots for TNB, NBA and ABW are shown in Fig. 1, and the plots for GL, AFS and AFF are shown in Fig. S1.

The phenotypic variance explained by each significant SNP ranged from 1.24% to 2.26% (Tables 1 & S2). The phenotypic variance explained by all significant SNPs was 6.77% (SE = 0.03) and 4.46% (SE = 0.03) for TNB and NBA respectively. Of note, among them, individuals with

genotype AA at SNP *WU_10.2_2_162527469* on SSC2 had higher TNB than did those with genotypes AC or CC, whereas individuals with genotype GG at SNP *WU_10.2_3_44631648* on SSC3 showed lower NBA than did those with genotypes AG and AA (Table S3). For AFS and AFF, the phenotypic variance explained by all significant SNPs was 7.98% (SE = 0.048) and 9.38% (SE = 0.05) respectively. The peaking SNP *ALGA0111336* on SSC16 revealed that individuals with genotype CC had higher AFS and AFF than did those with other genotypes (Table S3). In addition, for two other traits, ABW and GL, the phenotypic variance explained by all significant SNPs was equal to 7.11% (SE = 0.05) and 11.49% (SE = 0.05) respectively.

A total of 376 different functional genes (Table S4) that contained or were near the significant SNPs were selected. Combined with Gene Ontology analysis (Table S5), 14 genes

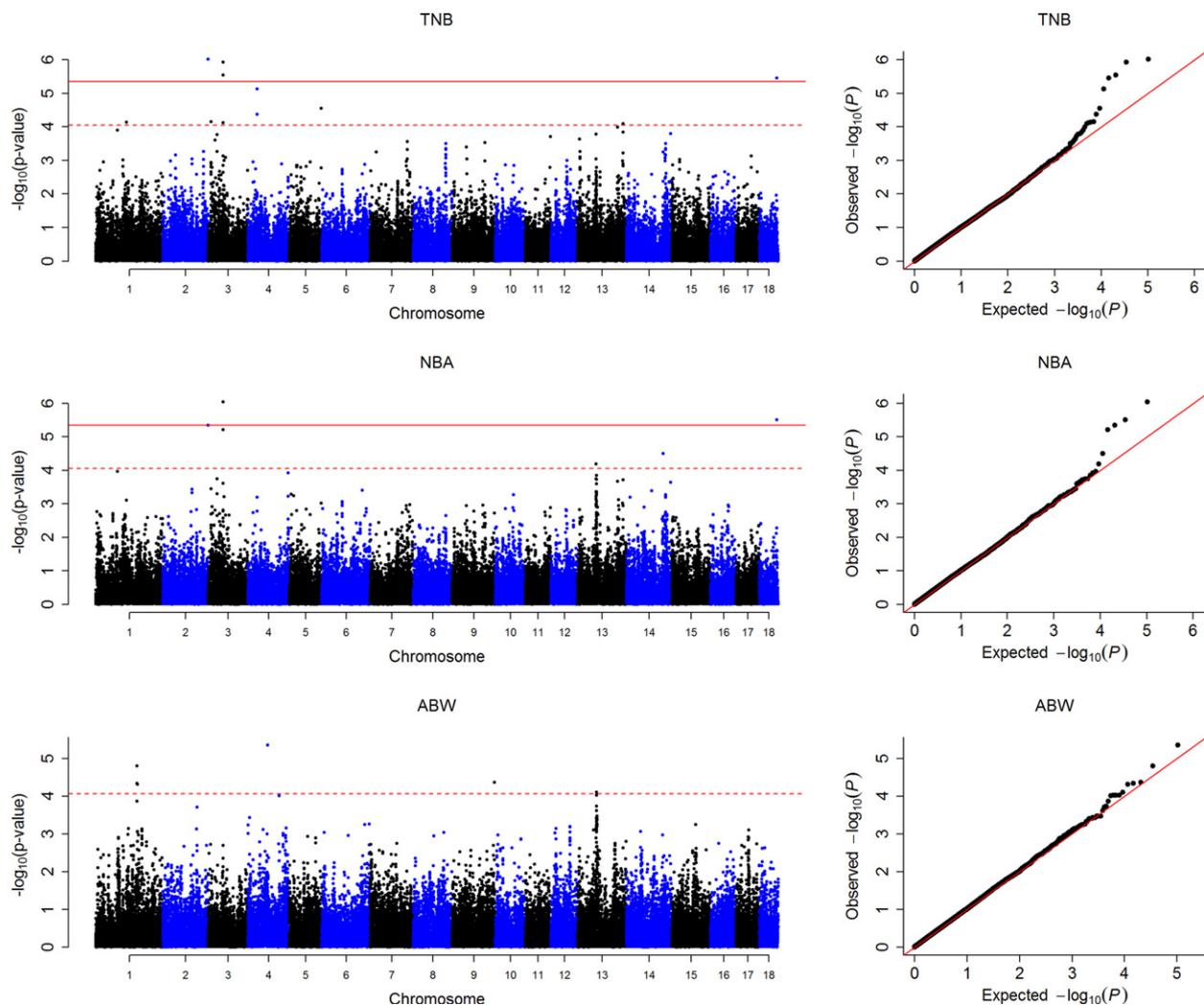


Figure 1 Manhattan plots and quantile–quantile (Q–Q) plots of the observed P -values for total number born (TNB), number born alive (NBA) and average birth weight (ABW). The horizontal red and red dashed lines in the Manhattan plots indicate the genome-wide (4.42×10^{-6}) and suggestive significance (8.84×10^{-5}) thresholds respectively. The Q–Q plots show the observed $-\log_{10}$ -transformed P -values (y-axis) and the expected $-\log_{10}$ -transformed P -values (x-axis).

with biological functions, such as carbohydrate metabolism, lipid metabolism and embryonic development, were selected as promising candidates for swine reproductive traits (Tables 1 and S2).

For TNB and NBA, we selected six functional genes. Both the *IL1B2* and *GCK* genes simultaneously associated with two traits. The *IL1B2* gene promotes follicular growth, corpus luteum formation and embryo development (Ross *et al.* 2003). The *GCK* gene encodes an enzyme that regulates glucose level (Muller *et al.* 2014) and affects the supplement of fetal energy substrate. Three other genes—*BHLHA15*, *OCM2* and *SMAD2*—associated with TNB. The *BHLHA15* gene has a critical role in mouse embryonic development, especially in gastrulae and plantule (Pin *et al.* 2000). The *OCM2* gene contributes to the transport of calcium and affects fetal skeletal mineralization (Belkacemi *et al.* 2002). We also selected the *SMAD2* gene, which promotes actions of follistatin on blastocyst development in early embryogenesis (Zhang *et al.* 2015). Furthermore, the *HABP2* gene is another candidate gene for NBA and plays a role in the integrity of mouse cumulus extracellular matrix (Chen *et al.* 1993).

The *PAQR5* and *GRB10* genes were selected as important candidates for ABW. The *PAQR5* gene promotes fast regulation of progesterone through a combination of specific receptors on the cell membrane (Thomas 2008), and the expression of *GRB10* in the placenta directly regulates placental size and efficiency in mouse (Charalambous *et al.* 2010).

Four genes—namely *PRELID2*, *DMNK*, *GPI* and *GPIHBP1*—associated with GL. *PRELID2* is involved in mouse embryogenesis during mid-later gestation (Gao *et al.* 2009), whereas *DMNK* plays a role in the process of embryonic implantation (Paria *et al.* 2002). The function of the *GPI* gene is similar to that of *GCK*: regulating glucose homeostasis. Like glucose, lipids play critical roles in fetal growth, and the *GPIHBP1* gene participates in the transportation of lipids, including cholesterol, triglycerides and other lipids (Herrera 2002).

A common candidate gene associated with AFS and AFF is the *ADCY2* gene, which contained the peak SNP *ALGA0111336* and is involved in catalyzing the synthesis of the secondary messenger cyclic adenosine monophosphate (cAMP), which promotes the functions of follicle stimulating hormone and luteinizing hormone on ovaries (Richards *et al.* 1995). Another candidate for AFF, *ACVR2B*, plays an important role in the production of estrogen and progesterone, oocyte maturation and follicle stimulating hormone receptor expression (Findlay 1993).

The sample size in this study, consisting of 1207 individuals, was larger than that used by Schneider *et al.* (2012), who used a composite population with 1152 individuals, and Onteru *et al.* (2011, 2012), who used 683 and 818 commercial sows respectively. Our study enriches the understanding of genetic mechanisms for reproductive traits, especially for AFS and AFF in pigs.

Compared with previous studies, there were 29 significant SNPs located on known QTL regions for reproductive traits (Table S6), such as number of teats (Verardo *et al.* 2016), number of stillborn and litter size (Onteru *et al.* 2012). One QTL region associated with litter size (SSC18, 52.3–77.6 MB) contained the *ALGA0098819* SNP for TNB; the *GCK* gene is located in this QTL region.

In summary, 53 significant SNPs were detected to be associated with six reproductive traits. Further, 14 functional genes were identified to be important candidates for TNB, NBA, ABW, GL, AFS and AFF, namely, *BHLHA15*, *OCM2*, *IL1B2*, *GCK*, *SMAD2*, *HABP2*, *PAQR5*, *GRB10*, *PRELID2*, *DMKN*, *GPI*, *GPIHBP1*, *ADCY2* and *ACVR2B*. Our findings provide important knowledge on the understanding of genetic architecture for swine reproductive traits.

Acknowledgements

We gratefully acknowledge our colleagues on the molecular quantitative genetics team for their technical assistance and helpful comments on the manuscript. This work was supported by the Beijing Innovation Consortium of Agriculture Research System (BAIC02-2016), the China Agriculture Research System (CARS-35), the National Natural Science Foundation of China (31671327), the Beijing City Committee of Science and Technology Key Project (D151100004615004), and the Program for Changjiang Scholar and Innovation Research Team in University (grant number IRT_15R621). The authors have no conflict of interest to declare.

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.
- Belkacemi L., Simoneau L. & Lafond J. (2002) Calcium-binding proteins: distribution and implication in mammalian placenta. *Endocrine* **19**, 57–64.
- Browning S.R. & Browning B.L. (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics* **81**, 1084–97.
- Browning B.L. & Browning S.R. (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *American Journal of Human Genetics* **84**, 210–23.
- Charalambous M., Cowley M., Geoghegan F., Smith F.M., Radford E.J., Marlow B.P., Graham C.F., Hurst L.D. & Ward A. (2010) Maternally-inherited *Grb10* reduces placental size and efficiency. *Developmental Biology* **337**, 1–8.
- Chen L., Russell P.T. & Larsen W.J. (1993) Functional significance of cumulus expansion in the mouse: roles for the preovulatory synthesis of hyaluronic acid within the cumulus mass. *Molecular Reproduction and Development* **34**, 87–93.

- Chen P., Baas T.J., Mabry J.W., Koehler K.J. & Dekkers J.C. (2003) Genetic parameters and trends for litter traits in U.S. Yorkshire, Duroc, Hampshire, and Landrace pigs. *Journal of Animal Science* **81**, 46–53.
- Findlay J.K. (1993) An update on the roles of inhibin, activin, and follistatin as local regulators of folliculogenesis. *Biology of Reproduction* **48**, 15–23.
- Gao M., Liu Q., Zhang F., Han Z., Gu T., Tian W., Chen Y. & Wu Q. (2009) Conserved expression of the *PRELI* domain containing 2 gene (*prelid2*) during mid-later-gestation mouse embryogenesis. *Journal of Molecular Histology* **40**, 227–33.
- Gu X., Feng C., Ma L. *et al.* (2011) Genome-wide association study of body weight in chicken F2 resource population. *PLoS One* **6**, e21872.
- He L.C., Li P.H., Ma X., Sui S.P., Gao S., Kim S.W., Gu Y.Q., Huang Y., Ding N.S. & Huang R.H. (2017) Identification of new single nucleotide polymorphisms affecting total number born and candidate genes related to ovulation rate in Chinese Erhualian pigs. *Animal Genetics* **48**, 48–54.
- Herrera E. (2002) Lipid metabolism in pregnancy and its consequences in the fetus and newborn. *Endocrine* **19**, 43–55.
- Holm B., Bakken M., Klemetsdal G. & Vangen O. (2004) Genetic correlations between reproduction and production traits in swine. *Journal of Animal Science* **82**, 3458–64.
- Lander E. & Kruglyak L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics* **11**, 241–7.
- Muller Y.L., Piaggi P., Hoffman D. *et al.* (2014) Common genetic variation in the *glucokinase* gene (*GCK*) is associated with type 2 diabetes and rates of carbohydrate oxidation and energy expenditure. *Diabetologia* **57**, 1382–90.
- Nicodemus K.K., Liu W., Chase G.A., Tsai Y.Y. & Fallin M.D. (2005) Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. *BMC Genetics* **6** (Suppl 1), S78.
- Onteru S.K., Ross J.W. & Rothschild M.F. (2009) The role of gene discovery, QTL analyses and gene expression in reproductive traits in the pig. *Society of Reproduction & Fertility Supplement* **66**, 87.
- Onteru S.K., Fan B., Nikkila M.T., Garrick D.J., Stalder K.J. & Rothschild M.F. (2011) Whole-genome association analyses for lifetime reproductive traits in the pig. *Journal of Animal Science* **89**, 988–95.
- Onteru S.K., Fan B., Du Z.Q., Garrick D.J., Stalder K.J. & Rothschild M.F. (2012) A whole-genome association study for pig reproductive traits. *Animal Genetics* **43**, 18–26.
- Paria B.C., Reese J., Das S.K. & Dey S.K. (2002) Deciphering the cross-talk of implantation: advances and challenges. *Science* **296**, 2185–8.
- Pin C.L., Bonvissuto A.C. & Konieczny S.F. (2000) *Mist1* expression is a common link among serous exocrine cells exhibiting regulated exocytosis. *Anatomical Record* **259**, 157–67.
- Price A.L., Zaitlen N.A., Reich D. & Patterson N. (2010) New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics* **11**, 459–63.
- Purcell S., Neale B., Todd-Brown K. *et al.* (2007) *PLINK*: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.
- Richards J.S., Fitzpatrick S.L., Clemens J.W., Morris J.K., Alliston T. & Sirois J. (1995) Ovarian cell differentiation: a cascade of multiple hormones, cellular signals, and regulated genes. *Recent Progress in Hormone Research* **50**, 223–54.
- Ross J.W., Malayer J.R., Ritchey J.W. & Geisert R.D. (2003) Characterization of the interleukin-1beta system during porcine trophoblastic elongation and early placental attachment. *Biology of Reproduction* **69**, 1251–9.
- Schneider J.F., Rempel L.A., Snelling W.M., Wiedmann R.T., Nonneman D.J. & Rohrer G.A. (2012) Genome-wide association study of swine farrowing traits. Part II: Bayesian analysis of marker data. *Journal of Animal Science* **90**, 3360–7.
- Sell-Kubiak E., Duijvesteijn N., Lopes M.S., Janss L.L., Knol E.F., Bijma P. & Mulder H.A. (2015) Genome-wide association study reveals novel loci for litter size and its variability in a Large White pig population. *BMC Genomics* **16**, 1049.
- Thomas P. (2008) Characteristics of membrane progesterin receptor alpha (mPRalpha) and progesterone membrane receptor component 1 (PGMRC1) and their roles in mediating rapid progesterin actions. *Frontiers in Neuroendocrinology* **29**, 292–312.
- Verardo L.L., Silva F.F., Lopes M.S., Madsen O., Bastiaansen J.W.M., Knol E.F., Kelly M., Varona L., Lopes P.S. & Guimaraes S.E.F. (2016) Revealing new candidate genes for reproductive traits in pigs: combining Bayesian GWAS and functional pathways. *Genetics Selection Evolution* **48**, 9.
- Yang J., Lee S.H., Goddard M.E. & Visscher P.M. (2011) *GCTA*: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics* **88**, 76–82.
- Zhang H., Wang Z., Wang S. & Li H. (2013) Retraction note: progress of genome wide association study in domestic animals. *Journal of Animal Science & Biotechnology* **4**, 1.
- Zhang K., Rajput S.K., Lee K.B., Wang D., Huang J., Folger J.K., Knott J.G., Zhang J. & Smith G.W. (2015) Evidence supporting a role for SMAD2/3 in bovine early embryonic development: potential implications for embryotropic actions of follistatin. *Biology of Reproduction* **93**, 86.
- Zhou X. & Stephens M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* **44**, 821–4.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1 Manhattan plots and quantile–quantile plots of the observed *P*-values for gestation length (GL), age at first service (AFS) and age at first farrowing (AFF).

Table S1 Descriptive statistics of reproductive traits in the Large White population.

Table S2 Suggestive significant SNPs for six reproductive traits.

Table S3 Genotype–phenotype correlations of the most significant SNPs for four reproductive traits

Table S4 Annotated genes with less than 1 Mb of significant SNPs.

Table S5 Significant Gene Ontology terms for five reproductive traits.

Table S6 Significant SNPs located in known QTL regions for reproductive traits.