AJAS

Asian-Australasian Journal of Animal Sciences

Whole-genome association and genome partitioning revealed variants and explained heritability for total number of teats in a Yorkshire pig population

Md. Rasel Uzzaman[1,a], Jong-Eun Park[1,a], Kyung-Tai Lee[1], Eun-Seok Cho[1], Bong-Hwan Choi[1,*], and Tae-Hun Kim[1,*]

* Corresponding Authors:
Bong-Hwan Choi
Tel: +82-63-238-7304, **Fax:** +82-63-238-7347,
E-mail: bhchoi@korea.kr
Tae-Hun Kim
Tel: +82-63-238-0790, **Fax:** +82-63-238-7347,
E-mail: thkim63@korea.kr

[1] Animal Genomics & Bioinformatics Division, National Institute of Animal Science, RDA, Wanju 55365, Korea

[a] Md. Rasel Uzzaman and Jong-Eun Park have contributed equally to this work.

ORCID
Bong-Hwan Choi
https://orcid.org/0000-0002-4795-3285
Tae-Hun Kim
https://orcid.org/0000-0003-1621-3281

Submitted Mar 7, 2017; Revised Jul 3, 2017;
Accepted Oct 9, 2017

Objective: The study was designed to perform a genome-wide association (GWA) and partitioning of genome using Illumina's PorcineSNP60 Beadchip in order to identify variants and determine the explained heritability for the total number of teats in Yorkshire pig.
Methods: After screening with the following criteria: minor allele frequency, MAF≤0.01; Hardy-Weinberg equilibrium, HWE≤0.000001, a pair-wise genomic relationship matrix was produced using 42,953 single nucleotide polymorphisms (SNPs). A genome-wide mixed linear model-based association analysis (MLMA) was conducted. And for estimating the explained heritability with genome- or chromosome-wide SNPs the genetic relatedness estimation through maximum likelihood approach was used in our study.
Results: The MLMA analysis and false discovery rate p-values identified three significant SNPs on two different chromosomes (rs81476910 and rs81405825 on SSC8; rs81332615 on SSC13) for total number of teats. Besides, we estimated that 30% of variance could be explained by all of the common SNPs on the autosomal chromosomes for the trait. The maximum amount of heritability obtained by partitioning the genome were 0.22±0.05, 0.16±0.05, 0.10±0.03 and 0.08±0.03 on SSC7, SSC13, SSC1, and SSC8, respectively. Of them, SSC7 explained the amount of estimated heritability along with a SNP (rs80805264) identified by genome-wide association studies at the empirical p value significance level of 2.35E-05 in our study. Interestingly, rs80805264 was found in a nearby quantitative trait loci (QTL) on SSC7 for the teat number trait as identified in a recent study. Moreover, all other significant SNPs were found within and/or close to some QTLs related to ovary weight, total number of born alive and age at puberty in pigs.
Conclusion: The SNPs we identified unquestionably represent some of the important QTL regions as well as genes of interest in the genome for various physiological functions responsible for reproduction in pigs.

Keywords: Genome-wide Association Study; Genome Partitioning; Pig; Teat Number

INTRODUCTION

As one of the most in-demand livestock species, pigs play a big role in the global meat market economy. The success of pig production is related to an important trait: total number of teats. This trait directly affects the mothering ability of sows [1]. An increased total number of teats can enhance the productivity of the sows functioning teats through some management factors/considerations [2]. Therefore, an increased number of weaned piglets from increased total number of teats. On the other hand, the trait is known to have low (<0.20) to medium (approximately 0.30) heritability [1,3,4], and this also indicates a reasonable expectation of response to genetic selection for total number of teats.

Generally, teats are grouped into functional (good and desirable) teats, inverted and supernumerary (extra) teats. A functional teat is one with a well-developed predominant sphincter whereas a teat that is turned inwards is called inverted. Supernumerary teats are usually small and shorter in size in relation to the normal (functional) teats. While visual inspection and classification of teats can be done during processing 2 to 3 days after birth, it is not possible to accurately differentiate between functional and non-functional teats. Lundheim et al [5] examined the genetic association between total number of teats and number of functional teats evaluated at the end of a performance test (100 kg body weight) in Yorkshire gilts. They found that the total number of teats and number of functional teats had a high and positive genetic correlation (0.82) indicating that selection for total number of teats would result in a significant increase in number of functional teats. Thus, from an objective measurement point of view the total number of teats is usually used as a proxy for the number of functional teats.

The genetic improvement of female reproductive traits is complex because of this low to medium heritability and their sex-limited expression, and because phenotyping is only possible late in a sow's life. These conditions constitute a challenge for traditional animal breeding programs. Exploration of the genetic architecture of reproductive traits is necessary because of the complex genetic and biological processes involved [6,7]. Thus, the use of genome-wide association studies (GWAS) can be useful in searching for chromosomal regions that can help to explain the genetic architecture of total number of teats in pigs.

Hundreds and thousands of single nucleotide polymorphisms (SNPs) triggered the genotyping technology that began the revolution leading to GWAS. Among available methods, an increasing interest has been noticed in using the mixed linear models (MLMs) for GWAS because of their demonstrated effectiveness in accounting for relatedness among samples and in controlling for population stratification and other confounding factors [8-13].

In our study, we conducted a genome-wide association analysis for the total number of teats in a Yorkshire population using MLM approach. We thereafter estimated the proportion of phenotypic variance explained by genome- or chromosome-wide SNPs using genetic relatedness estimation through maximum likelihood (GREML) approach.

MATERIALS AND METHODS

Animals and phenotypes

The studied animals were from a commercial population of Sunjin (http://www.sj.co.kr/eng/sunjin_eng/about.asp); a commercial pig breeding company in South Korea. The animals were of purebred Yorkshire pigs and born between 2007 and 2014. All animals were raised in similar dietary supply and housing management system. Phenotypic measurements of total number of teats were obtained from 1,061 pigs as a proxy for the number of functional teats. The sampled animals were half-sib and full-sib pigs produced from 171 sires and 477 dams of Yorkshire breed. The record for the total number of teats was performed at 3 to 5 days of age by visual inspection. And, all the animals were of the first parity origin of their parents.

Genotyping and SNP quality

Genotyping was performed using the Illumina Porcine SNP 60K Beadchip. Blood samples collected for DNA extraction were only used for the purposes of the breeding program, and procedures were strictly in line with Korean law on the protection of animals. DNA was extracted from whole blood and genotyped at Illumina, Inc. (San Diego, CA, USA). We removed SNPs with a GenCall score <0.7, call rate <95%, minor allele frequency <0.01, and the SNPs with no physical position on the pig genome (pig genome build 10.2). After quality control, a total of 42,953 SNPs remained for association analysis.

Statistical analysis

For studying the association between genotyped SNPs and the trait of interest the following MLM was used in our study:

$$y = a + bx + g- + e$$

Where y is the phenotype, a is the mean term, b is the additive effect (fixed effect) of the candidate SNP to be tested for association, x is the SNP genotype indicator variable coded as 0, 1, or 2, g- is the accumulated effect of all SNPs except those on the chromosome where the candidate SNP is located. The $var(g-)$ will be re-estimated each time when a chromosome is excluded from calculating the genomic relationship matrix (GRM) and e is the residual. For the ease of computation, the genetic variance, $var(g-)$, is estimated based on the null model i.e. $y = a + g- + e$ and then fixed while testing for the association between each SNP and the trait. We thereby used a designed option called 'mlma-loco' for the association analysis as implemented in the genome-wide complex trait analysis (GCTA) tool [14].

Two types of p value plots had been demonstrated as part of the standard presentation of GWAS results: $-\log_{10} (P)$ genome-wide association plots (Manhattan plots) and quantile-quantile (QQ) plots. Manhattan plot represented the p values on the y-axis against genomic order by chromosome and position on the x-axis. Because of local correlation of the genetic variants, arising from infrequent genetic recombination, groups of significant p values tend to rise up high on the Manhattan plot. The QQ plot was used to assess the number and magnitude of observed associations between genotyped SNPs

and trait under study, compared to the association statistics expected under the null hypothesis of no association [15]. Observed association statistic ($-\log_{10}$ p-values) were ranked in order from smallest to largest on the y-axis and plotted against the expected values from a theoretical x^2-distribution under the null hypothesis of no association on the x-axis of the QQ plot. Deviations from the identity line suggested either that the assumed distribution was incorrect or that the sample contained values arising in some other manner, such as by true association. Since the underlying assumption in GWA studies is that the clear majority of assayed SNPs are not associated with the trait, strong deviations from the null suggest either a very highly associated and excessively genotyped locus (i.e. an associated gene with many genotyped SNPs) or spurious associations. On the other hand, very little deviation from the expected values could be attributed to cryptic relatedness in the study population, in the genotyping, or either true association.

For the estimation of the explained variances by genome-wide SNPs the basic concept behind our method was to fit the effects of all the SNPs simultaneously as random effects in a MLM without testing for associations of any individual SNP with the trait. The linear equation could be written using the following model:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{W}\mathbf{u} + \varepsilon \text{ with var}(\mathbf{y}) = V = \mathbf{W}\mathbf{W}'\sigma^2_\mu + \mathbf{I}\sigma^2_\varepsilon \quad (I)$$

Where $\mathbf{y}$ is an n×1 vector of phenotypes, n is the sample size, β is a vector of fixed covariate herd-year season (HYS) and eigenvectors from principal component analyses (PCAs). The HYS had 4 levels of season (winter, summer, spring, and fall) since 2007 to 2014. Individual GRMs were used to calculate the first 20 eigenvalues and eigenvectors and the eigenvector of the covariance matrix was subsequently used to calculate the PCA (PC1 and PC2), where U is a vector of SNP effects with $\mathbf{u} \sim N(0, \mathbf{I}\sigma^2_u)$, $\mathbf{I}$ is an nxn identity matrix, and ε is a vector of residual effects with $\varepsilon \sim N(0, \mathbf{I}\sigma^2_\varepsilon)$. $\mathbf{W}$ is a standardized genotype matrix with the ijth element $w_{ij} = (x_{ij} - 2p_i)/\sqrt{(2p_i(1-p_i))}$, where x_{ij} (coded as 0, 1, or 2) is the number of copies of the reference allele for the ith SNP of the jth individual, and p_i is the frequency of the reference allele of SNP i. If we define $\mathbf{A} = \mathbf{W}\mathbf{W}'/N$, and define σ^2_g as the variance explained by all the SNPs, i.e., $\sigma^2_g = N\sigma^2_u$, with N being the number of SNPs then the equation I will be equivalent to

$$\mathbf{y} = \mathbf{X}\beta + g + \varepsilon \text{ with } \mathbf{V} = \mathbf{A}\sigma^2_g + \mathbf{I}\sigma^2_\varepsilon \quad (II)$$

Where g is an n×1 vector of the total genetic effects of the individuals with $g \sim N(0, \mathbf{A}\sigma^2_g)$, and A is interpreted as the GRM between individuals with its kjth element being $A_{jk} = \frac{1}{N}\sum_{i=1}^{N}(x_{ij} - 2p_i)(x_{ik} - 2p_i)/2p_i(1-p_i)$, and the other notations are defined as in equation I. Therefore, we could estimate σ^2_g by the

restricted maximum likelihood (REML) approach, relying on the GRM estimated from all SNPs, and the model in general is represented by the following equation:

$$\mathbf{y} = \mathbf{X}\beta + \sum_{i=1}^{r} g_i + \varepsilon \quad (III)$$

Where g_i is a vector of random genetic effects, which could be the total genetic effect for the whole genome or for a single chromosome. In this model, the phenotypic variance (σ^2_p) is partitioned into the variance explained by each of the genetic factors and the residual variance,

$$\mathbf{V} = \sum_{i=1}^{r} \mathbf{A}_i\sigma^2_i + \mathbf{I}\sigma^2_\varepsilon, \quad (IV)$$

Genome partitioning

The principle of this method based on partitioning phenotypic variance across multiple additive genetic variance components to test whether these components explain significant amounts of phenotypic variance. Each genetic component represents an arbitrary group of genetic markers, usually chromosomes or smaller linkage groups. Additive genetic variance components are estimated from identity-by-state GRMs rather than from pedigree data [14,16]. The heritability of genome- or chromosome-wide SNPs was estimated as the ratio [V(G)/Vp] of the genetic variance to the phenotypic variance estimate of the animals using the GRM-restricted maximum likelihood (GREML) approach. Chromosome-specific heritability (h^2) was estimated by generating one GRM for each chromosome using only SNPs on that chromosome and fitting this GRM alongside a second GRM based on all genome-wide SNPs apart from those on that chromosome. A likelihood-ratio test (LRT) was then carried out between the full model and the model excluding the chromosome-specific component. The threshold of an LRT would depend both on the log-likelihood values as well as parameter numbers of the two models. In case of positive LRT (>10) values the generated probability is often large and conclusive.

RESULTS AND DISCUSSION

Genome-wide association

We estimated the GRM between pairs of individuals based on 42,953 autosomal SNPs, followed by a mixed linear model-based association analysis (MLMA). The summary of the MLMA for total number of teats with 42,953 genome-wide SNPs was produced. A false discovery rate (FDR level: 0.05) correction of p-values derived from MLMA identified three significant (q≤0.01) SNPs (Table 1). The Manhattan plot of the GWA analysis for the trait is given in Figure 1A. The QQ plot to illustrate the level of potential p-value inflation is shown in Figure 1B. Several effective statistical methods are available to correct for population structure and are a standard component

Table 1. A false discovery rate (FDR level: 0.05) correction of p-values derived from mixed linear model-based association analysis identified three significant (q≤0.01) SNPs for total number of teats in a Yorkshire pig population

Chr	SNP	refSNP	bp	A1	A2	Freq	b	se	p-value	q-value
8	ASGA0102835	rs81476910	145,098,780	A	G	0.01	−1.54	0.25	1.66E−09	7.12E−05
8	ALGA0050175	rs81405825	145,451,027	A	G	0.07	−0.49	0.09	7.03E−07	0.010
13	H3GA0052641	rs81332615	12,004,442	A	C	0.03	−0.83	0.15	1.34E−07	0.003

Chr: chromosome; SNP, single nucleotide polymorphism; refSNP, reference SNP identification from the SNP database; bp, physical position; A1, reference allele; A2, other allele; Freq, frequency of the reference allele; b, SNP effect; se, standard error.

of rigorous GWA analyses [17,18]. Determining the population structure by PCA indicated that the sampled animals were from a little bit disperse but still a single population (Figure 2). Genome-wide significant hits on SSC8 were detected, namely
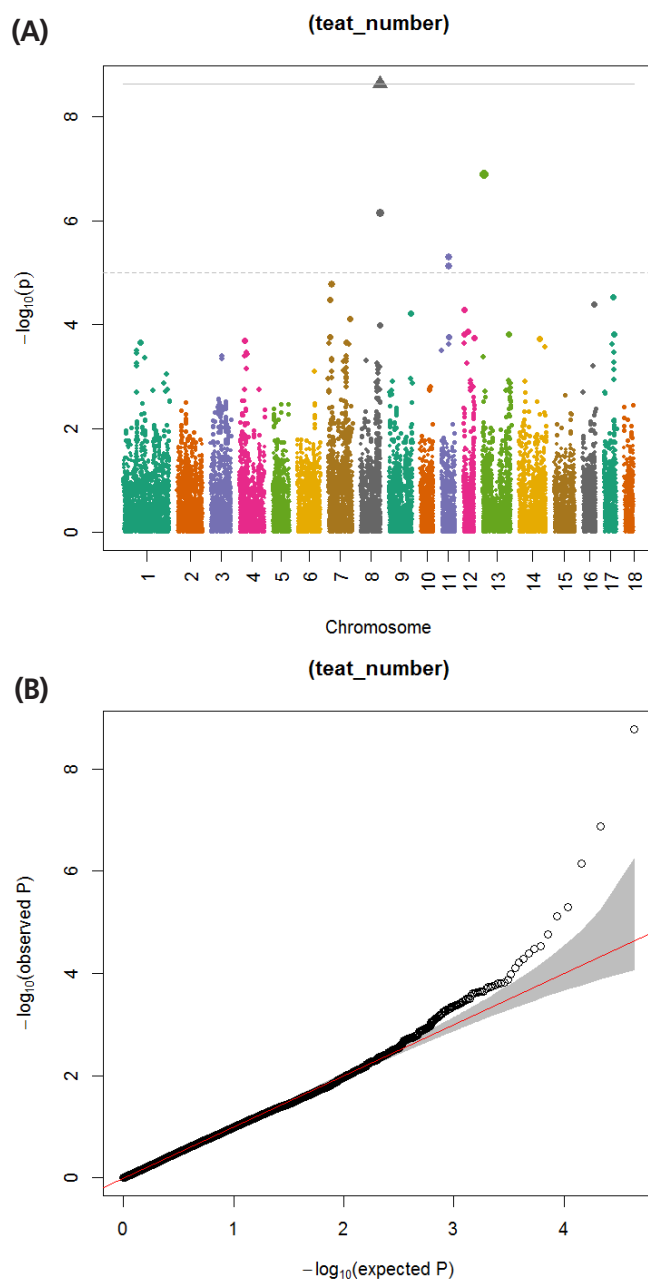

Figure 1. (A) Manhattan plot for the total number of teats examined in the genome-wide association study using Korean Yorkshire pigs. (B) The QQ plot in the upper-right corner compares expected versus observed −log$_{10}$ (p-values) for all 42,953 single nucleotide polymorphisms included in the genome-wide association (GWA) analysis, with the red line corresponding to the null hypothesis of no association.
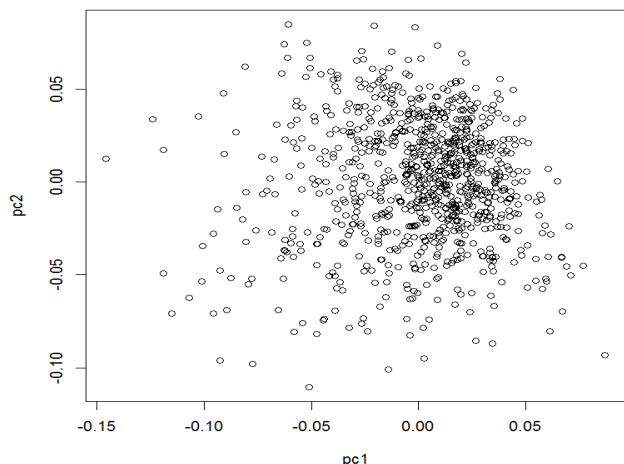
Figure 2. Population structure revealed by principal component analysis (PCA). Individuals genomic relationship matrix (GRM) was used to calculate first 20 eigenvalues and eigenvectors and subsequently the eigenvector of covariance matrix was used to calculate the PCA (PC1 and PC2).

Table 2. Estimation of variance component for 1,048 animals using REML

Source	Variance	SE
V(G)	0.32	0.06
V(e)	0.76	0.05
Vp	1.09	0.05
V(G)/Vp	0.30	0.05
L(H$_0$)[1]	−558.524	-
L(H$_1$)[2]	−527.841	-
LRT[3]	61.366	-
df	1	-
Pval	2.387e-15	-
n	1,048	-

REML, restricted maximum likelihood; SE, standard error; V(G), genetic variance estimate; V(e), residual variance estimate; Vp, phenotypic variance estimate; V(G)/Vp, ratio of genetic variance to phenotypic variance estimate.
[1] log-likelihood under the null hypothesis that $\sigma^2_g = 0$.
[2] log-likelihood under the null hypothesis that $\sigma^2_g \neq 0$.
[3] log likelihood ratio test statistic, $LRT = 2[L(H_1)-L(H_0)]$, L(H$_1$); log-likelihood for alternative model, L(H$_0$); log-likelihood for null model.

rs81476910 (p = 1.66E−09) and rs81405825 (p = 7.03E−07) located in the 145,098,780- and 145,451,027-bp genomic regions, respectively. The other significant SNP was rs81332615 (p = 1.34E−07), located in the 12,004,442-bp genomic region on SSC13 (Table 1). The three highly significant loci did not fall within the previously reported quantitative trait loci (QTL) for the teat number trait in pigs. Apart from it, a SNP (rs8080 5264) identified on SSC7 by GWAS at the empirical p value significance level of 2.35E−05 was found likely to be concordant being in a nearby QTL (103,789,642 to 105,224,235) region on SSC7 for the teat number trait as identified in a recent study [19]. The most promising candidate gene nearby the QTL on this chromosome is *VRTN* (vertebrae development associated). This gene had been suggested to be related to the total number of vertebrae, a trait correlated with the total number of teats [20]. Additionally, rs81476910 and rs81405825 were found within the reproduction QTL (140,282,531 to 148,491,826 bp) on SSC8 for pig ovary weight. On the other hand, rs81332615 were seen nearby the QTLs for total number born alive (13,012,892 to 13,038,269 bp) and age at puberty (9,843,648 to 45,496,370 bp) on SSC13 for pig reproduction traits (http://www.animalgenome.org/cgi-bin/QTLdb/SS/download?file=gbpSS_10.2). Therefore, significant SNPs along with the genomic regions derived from our genome-wide association analysis should undergo for further studies to explore the genomic underpinnings of total number of teats in Yorkshire pigs.

Genome partitioning and explained variances

The heritability of teat number varies over a wide range, from 0.07 to 0.79, but most estimates are at the low to medium level, from 0.20 to 0.50 [21]. The estimated explained variance i.e. heritability estimating with the variance component V(G)/Vp for the total number of teats was 0.30±0.05 (Table 2) in our study. The partitioning of the total genetic variance to chromosome-specific by fitting the GRMs with the SNPs on each of the chromosomes is shown in Table 3.

An estimated heritability of 0.34±0.05 for teat number was reported using a GRM constructed from all SNP markers, whereas using the additive relationship matrix with pedigree information the heritability estimate was found to be 0.43±0.04 in a purebred Duroc population [22]. Again, using pedigree information, an estimated heritability of 0.44±0.04 for teat number was obtained in a population of purebred Berkshire

Table 3. Genome partitioning of genetic variation associated with total number of teats using autosomal SNPs

Chromosome number	Length (Mb)	Explained heritability
1	315.32	0.10
2	162.57	0.04
3	144.79	0.07
4	143.47	0.04
5	111.51	0.00
6	157.77	0.02
7	134.76	0.22
8	148.49	0.08
9	153.67	0.06
10	79.10	0.07
11	87.69	0.03
12	63.59	0.07
13	218.64	0.16
14	153.85	0.04
15	157.68	0.02
16	86.90	0.05
17	69.70	0.04
18	61.22	0.02

SNPs, single nucleotide polymorphisms.

pigs [23]. Therefore, a difference was observed between the heritability estimates obtained using pedigree information and that based on genomic information. Studies suggest that one reason for this difference in heritability estimates is the imperfect LD between SNPs and QTL. Accordingly, it would be impossible to detect all the QTL that contributed to genetic variation in teat number using the SNP information available thus far in the Porcine SNP60 BeadChip.

Partitioning of genome revealed that the four chromosomes (SSC1, SSC7, SSC8, and SSC13), would exhibit higher explained heritability as compared to all the other chromosomes (Figure 3). SSC13 and SSC8 explained an estimated heritability of 0.16±0.05 and 0.08±0.03, respectively, reflecting the effects of significant SNPs detected on these chromosomes. Conversely, SSC1 explained an estimated heritability of 0.10± 0.03 having no significant SNPs detected on it. On the other hand, SSC7 explained an estimated heritability of 0.22±0.05 along with a SNP (rs80805264) identified on it by GWAS at the empirical p value significance level of 2.35E-05 in our study.

However, for the chromosome 1, 7, where none of the three highly significant SNPs was detected, the explained proportion of genetic variance was greater (10%, 22%) than the chromosome 8 (8%). In GCTA, the model hereby assumes that expected heritability does not vary with either linkage disequilibrium (LD) or genotype certainty. On the contrary, by analyzing imputed data for a large number of human traits, researchers empirically derived a model that more accurately describes how heritability varies with minor allele frequency (MAF), LD, and genotype certainty [24]. Their improved model across 19 traits leaded to estimates of common SNP heritability on average 43% (standard deviation 3) higher than those obtained from the widely used software GCTA and 25% (standard deviation 2) higher than those from the recently proposed extension GCTA-LDMS (LD+MAF Stratification). So, the chromosome(s) with the highly significant SNPs ex-

plaining lower proportion of genetic variance might be under this pitfall. And, in particular rs81476910 has been noticed with lower MAF (0.01) but highly significant SNP on SSC8 (Table 1).

On the other hand, the summed values of explained heritability were 1.13, which was quite different from 1.00. This would be because of the LD structure of the genome which is commonly seen while partitioning the heritability into the contributions of genomic regions [25].

While estimating the proportion of (narrow sense) heritability of a trait a substantial portion of missing heritability may arise from over estimation of the denominator. And, in such case the numerator is the proportion of the phenotypic variance explained by the additive effects of known variants and the denominator is the proportion of the phenotypic variance attributable to the additive effects of all variants, including those not yet discovered. This over estimation of denominator thus under estimate explained heritability by a set of known genetic variants and therefore may fall far short of 100%. This gap is known as phantom heritability [26]. Possibly due to this reason our chromosome-specific heritability remains as shortfall to 100% of those explained in genome-wide scale for the trait (Table 3).

In conclusion, the SNPs we identified unquestionably represent some of the important QTL regions as well as genes of interest in the genome for various physiological functions responsible for reproductive traits. In addition, chromosome-specific partitioning was concordant, while phenotypic variances were explained by highly significant SNPs together with other SNPs on certain chromosomes. Therefore, these SNP markers along with the important genes in those QTL regions as observed on the four chromosomes (SSC1, SSC7, SSC8, and SSC13) should be considered for further validating research on total number of teats in Yorkshire pigs.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.
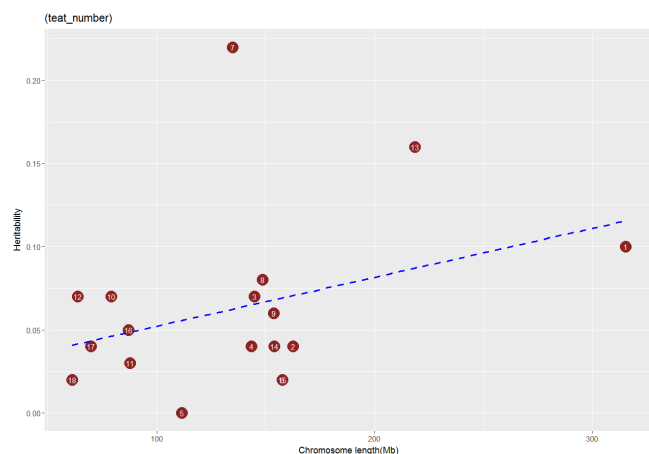
ACKNOWLEDGMENTS

REFERENCES

Figure 3. Estimates of the variance explained by each chromosome plotted against chromosome length.

1. Hirooka H, De Koning D, Harlizius B, et al. A whole-genome scan for quantitative trait loci affecting teat number in pigs. J Anim Sci 2001;79:2320-6.

2. Kim S, Easter R, Hurley W. The regression of unsuckled mammary glands during lactation in sows: the influence of lactation stage, dietary nutrients, and litter size. J Anim Sci 2001;79: 2659-68.

3. Clayton G, Powell J, Hiley P. Inheritance of teat number and teat inversion in pigs. Anim Prod 1981;33:299-304.

4. McKay R, Rahnefeld G. Heritability of teat number in swine. Can J Anim Sci 1990;70:425-30.

5. Lundheim N, Chalkias H, Rydhmer L. Genetic analysis of teat number and litter traits in pigs. Acta Agric Scand A Anim Sci 2013;63:121-5.

6. Lande R, Thompson R. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 1990;124: 743-56.

7. Rothschild M, Jacobson C, Vaske D, et al. The estrogen receptor locus is associated with a major gene influencing litter size in pigs. Proc Natl Acad Sci USA 1996;93:201-5.

8. Kang HM, Sul JH, Service SK, et al. Variance component model to account for sample structure in genome-wide association studies. Nat Genet 2010;42:348-54.

9. Kang HM, Zaitlen NA, Wade CM, et al. Efficient control of population structure in model organism association mapping. Genetics 2008;178:1709-23.

10. Listgarten J, Kadie C, Schadt EE, Heckerman D. Correction for hidden confounders in the genetic analysis of gene expression. Proc Natl Acad Sci USA 2010;107:16465-70.

11. Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. Nat Rev Genet 2010;11:459-63.

12. Yu J, Pressoir G, Briggs WH, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 2006;38:203-8.

13. Zhang Z, Ersoz E, Lai C-Q, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet 2010; 42:355-60.

14. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011; 88:76-82.

15. Burton PR, Clayton DG, Cardon LR, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.

16. William GH. Quantitative genetics in the genomics era. Curr Genomics 2012;13:196-206.

17. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet 2003;361:598-604.

18. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype–phenotype associations. Nature 2007;447:655-60.

19. Verardo LL, Silva FF, Lopes MS, et al. Revealing new candidate genes for reproductive traits in pigs: combining Bayesian GWAS and functional pathways. Genet Sel Evol 2016;48:9.

20. Lopes MS, Bastiaansen JW, Harlizius B, Knol EF, Bovenhuis H. A genome-wide association study reveals dominance effects on number of teats in pigs. PloS One 2014;9:e105867.

21. Borchers N, Reinsch N, Kalm E. Teat number, hairiness and set of ears in a Pietrain cross: variation and effects on performance traits. Archiv Tierzucht 2002;45:465-80.

22. Arakawa A, Okumura N, Taniguchi M, et al. Genome-wide association QTL mapping for teat number in a purebred population of Duroc pigs. Anim Genet 2015;46:571-5.

23. Tomiyama M, Oikawa T, Hoque M, Kanetani T, Mori H. Influence of early postweaning traits on genetic improvement of meat productivity in purebred Berkshire pigs. J Anim Sci 2009;87:1613-9.

24. Speed D, Cai N, Johnson MR, et al. Re-evaluation of SNP heritability in complex human traits. Nat Genet 2017;49:986-92.

25. Kostem E, Eskin E. Improving the accuracy and efficiency of partitioning heritability into the contributions of genomic regions. Am J Hum Genet 2013;92:558-64.

26. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci USA 2012;109:1193-8.