

RESEARCH ARTICLE

Genome-wide search for candidate genes determining vertebrae number in pigs

Longchao ZHANG, Jingwei YUE, Xin LIU, Jing LIANG, Kebin ZHAO, Hua YAN, Na LI, Lei PU, Yuebo ZHANG, Huibi SHI, Ligang WANG, Lixian WANG (✉)

Key Laboratory of Farm Animal Genetic Resources and Germplasm Innovation of Ministry of Agriculture of China/Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

Abstract Longer porcine carcasses may be expected to have more vertebrae. Therefore, vertebrae number in pigs is an economically important trait. To examine the genetic basis of this trait, we genotyped 578 F₂ Large White × Minzhu pigs using the Porcine SNP60K BeadChip. A genome-wide association study (GWAS) identified 36 significant single nucleotide polymorphisms (SNPs) on the chromosomes SSC1 (294.28–300.32 Mb) and SSC7 (102.22–109.39 Mb). A 6.04-Mb region that contained all 13 significant SNPs on SSC1 also contained the gene *NR6A1*, previously reported to influence the number of vertebrae in pigs. However, the reported putative casual mutation of *NR6A1* c.748C > T showed no genome-wide significant association with the trait, suggesting it was not a causal mutation in our population. The remaining 23 significant SNPs on SSC7 were concentrated in a 7.17-Mb region, which was within a quantitative trait locus interval for number of vertebrae. *TMED10* was the closest gene to the most significant SNP and might be a candidate. Haplotype sharing and block analysis refined the QTL to an interval of about 3 Mb containing 29 candidate genes. Of these 29 genes, the previously reported possible casual mutation of *VRTN* g.19034A > C was not found to be a causal mutation in our population. Exploration of these genes via additional genetic and functional studies in mammals revealed that *TGFβ3* could be a good candidate on SSC7. A mutation of *TGFβ3* c.1749G > A was detected by GWAS and could be proposed as a candidate causal mutation, or as closely linked to a causal mutation, for the number of vertebrae in pigs.

Keywords genome-wide association study, number of vertebrae, pig, *SSC7*, *TGFβ3*

1 Introduction

Backbones consist of the morphologically differentiated cervical, thoracic, lumbar, sacral and caudal vertebrae, and the number of these varies in pigs^[1]. Wild boars (*Sus scrofa*), the ancestor of domestic pigs, all have 19 thoracic-lumbar vertebrae, whereas most Chinese indigenous pig breeds have 19 or 20 vertebrae (thoracic-lumbar). By comparison, Western commercial pig breeds, such as Landrace and Large White, have 20–22 vertebrae^[1]. The number of vertebrae is usually associated with carcass length and is an economically important trait, because a greater number of vertebrae in pigs is associated with higher economic value. According to a study by King and Roberts, longer carcasses may be expected to have a higher number of vertebrae^[1]. Understanding the genetic basis for the number of vertebrae can offer insights into the mechanism for vertebral developmental in mammals and provide genetic markers to aid molecular breeding of pigs.

Locating quantitative trait loci (QTLs) and gene-mining for number of vertebrae have recently received research attention. Wada et al. first reported two QTLs on wild boar chromosomes SSC1 and SSC2 that are associated with the number of vertebrae^[2]. Subsequently, using several different pig populations, two QTLs for the number of vertebrae were mapped to SSC1 and SSC7^[3]. Using three F₂ experimental families, fine mapping of the QTL on SSC1 was performed and showed that nuclear receptor subfamily 6, group A, member 1 (*NR6A1*), was a strong candidate gene that appeared to influence the number of vertebrae in pigs^[4]. Gene-mining the other QTL on SSC7 suggested that the vertebrae development homolog (*VRTN*) could be a good candidate^[5]. A genome-wide association study (GWAS), based on the high density of genome-wide single nucleotide polymorphisms (SNPs), has proven to be a more efficient method to not only identify QTLs but also mine major new genes. Additional methods and different populations are required to identify

Received January 25, 2017; accepted May 9, 2017

Correspondence: iaswlx@263.net

loci associated with vertebrae number and detect good candidate genes. The objectives of this study were to identify SNPs associated with vertebrae number using GWAS in an F₂ Large White × Minzhu pig population.

2 Materials and methods

2.1 Population and phenotypic data

All animals used in this study were treated according to the guidelines for experimental animals established by the Council of China. Animal experiments were approved by the Science Research Department of the Institute of Animal Science, Chinese Academy of Agricultural Sciences (Beijing, China). The F₀ population was generated using four Large White boars and 16 Minzhu sows. In the F₁ generation, nine boars and 46 sows were selected to mate to produce 578 F₂ individuals. All animals in the F₂ generation were born in three parities and 94 litters. Each F₂ male was castrated 3 d after birth. All animals were maintained in uniform housing and were fed the same fodder. When F₂ animals were 240 ± 7 d-old, they were slaughtered in 52 batches. The numbers of thoracic-lumbar vertebrae together were counted from the carcasses and used as phenotypes.

2.2 Genotyping and quality control

Genomic DNA was extracted from ear tissue using standard methods^[6]. The Illumina (San Diego, CA, USA) SNP60K chip for pigs was employed to genotype all individuals. In our previous study^[7], quality control parameters for single nucleotide polymorphisms included the following: call rate > 90%, minor allele frequency > 3%, and Hardy–Weinberg equilibrium with $P > 10^{-6}$.

2.3 Genome-wide association study

The genome-wide rapid association using the mixed model and regression-genomic control approach^[8,9] was used in the present study. Sex and litter were used as fixed and random effects, respectively. DMU and GenABEL software^[9] in the R environment were employed to estimate the residuals of traits for each individual and perform the GWAS.

In the first step, the mixed model was used, and data were analyzed using DMU software, using the formula:

$$y = 1\mu + Xb + Tc + Za + e$$

where y is the vector of the phenotypes of all F₂ individuals, b is the vector of fixed effects of the sex, c is the vector of litter effect as a random effect, $c \sim N(0, \sigma_c^2)$, a is the vector of random additive genetic effects with $a \sim N(0, A\sigma_a^2)$ (where A is the relationship matrix

calculated from the corrected pedigree and σ_c^2 is the additive genetic variance). X , T and Z are incidence matrices related to records of fixed and random effects in y , e is the vector of residual errors, $e \sim N(0, I\sigma_e^2)$ (where I is the identity matrix and σ_e^2 is the residual variance). The vector of residuals y^* is estimated as:

$$y^* = y - (1\hat{\mu} + X\hat{b} + T\hat{c} + Z\hat{a})$$

where \hat{b} , \hat{c} , and \hat{a} are estimates and predictors for b , c and a , respectively.

Second, the residuals were used as the dependent trait, via the formula:

$$y^* = 1\mu + kg + e^*$$

where y^* is the vector of the adjusted phenotypes in the first step, g is the vector of the genotypes which were coded as “0”, “1” and “2” corresponding to AA, AB, and BB, k is the regression coefficient, and e^* is the vector of random residuals. Based on single locus regression analysis; the analysis was performed in the R statistical environment using the GenABEL package.

Finally, the unadjusted test statistic factor of the i th SNP T_i^2 was calculated in the genomic control procedure as:

$$T_i^2 = \hat{k}_i^2 / \text{var}(\hat{k}_i)$$

where \hat{k}_i and $\text{var}(\hat{k}_i)$ are the estimate and sample variance of k , respectively. The deflation factor λ is estimated as $\lambda = \text{median}(T_1^2, T_2^2, \dots, T_i^2) / 0.456$, where 0.456 is simply the median of χ^2 distribution with one degree of freedom^[10]. The association of the i th SNP with the trait was examined by comparison of T_i^2 / λ with $\chi_{(1)}^2$.

When association studies are conducted with many SNPs, the tests performed on each SNP are usually not independent, depending on the correlation structure among the SNPs. This violation of the independence assumption limits the ability of the Bonferroni correction to effectively control type I error, and the point-wise error rate has to be adjusted in order to keep the experiment-wise error rate at a nominal level. Therefore, the number of independent tests, which is regarded as number of effective SNPs, can be correctly inferred and used in the standard Bonferroni correction to rapidly adjust for multiple testing. The number of effective SNPs, which was estimated using a simpleM method^[11], was 12039 (Table S1). A lower conventional Bonferroni P was calculated as 4.15×10^{-6} ($0.05/12039$)^[12] and was applied to avoid missing QTLs.

2.4 Haplotype sharing and linkage disequilibrium analysis

The genotypes of nine F₁ boars were determined using marker-assisted segregation analysis (MASS)^[13]. According to the two SNPs of INRA0027600 and H3GA0022821, the genotype of each boar was determined

from a Z-score corresponding to the \log_{10} likelihood ratio LH1/LH0. LH1 corresponds to the likelihood of the pedigree data assuming that the boar is of Qq genotype, and LH0 corresponds to the likelihood of the pedigree data assuming that the boar is of QQ or qq genotype. Boars were considered to be Qq when $Z > 2$; QQ or qq when $Z < -2$; and of undetermined genotype if $2 > Z > -2$. According to the MASS analysis, Q-bearing chromosomes in F₁ boars segregated and haplotype sharing analysis was performed using all 23 significant SNPs on SSC7. Haplotype block detection was performed on the chromosomal region which contained all the SNPs that were significantly associated with vertebrae number. The HAPLOVIEW V4.1 program^[14] was used to detect and visualize the haplotype blocks in this work.

2.4.1 NR6A1 and VRTN candidate causal variants genotyping

According to previous reports^[4,15], *NR6A1* c.748C > T, *VRTN* g.19034A > C and *VRTN* g.20311_20312ins291 were selected as candidate causal mutations to genotype for 578 F_{TGF} individuals. A set of primers (Table S2) was used to amplify genomic DNA. Amplification was performed in a routine way with 1.5 mmol·L⁻¹ of MgCl₂ and optimal annealing temperatures. All PCR products were bidirectionally sequenced with original PCR primers on the 3130XL Genetic Analyzer (Applied Biosystem, Carlsbad, CA, USA).

2.4.2 Resequencing of the *TGFβ3* and polymorphism detecting and genotyping

To resequence the porcine transforming growth factor, beta 3 (*TGFβ3*), a set of primers (Table S2) were used to amplify genomic DNA of the 20 F₀ individuals. The resulting polymerase chain reaction (PCR) products covered all of the untranslated regions (UTR) and exons of genes. Similar to mutations of *NR6A1* and *VRTN* genotyping, a total of 578 F₂ individuals were genotyped according to the detected mutations by PCR.

3 Results and discussion

3.1 Detection 36 genome-wide significant SNPs on SSC1 and SSC7

The pigs had 19 ($n = 243$), 20 ($n = 292$), 21 ($n = 42$), and 22 ($n = 1$) vertebrae with a mean of 19.7. The final data set that passed quality control screening and was used in the analysis contained 43837 SNPs and came from 578 F₂ individuals. The distribution of SNPs after quality control and the average distance between adjacent SNPs on each chromosome are shown in Table 1. The results of the

Table 1 Distribution of SNPs on each chromosome (after quality control) in a population of Large White × Minzhu pigs and the average distances between SNPs on each chromosome¹

Chromosome	No. of SNPs	Average distance/kb
1	5340	58.98
2	2775	58.48
3	2262	63.72
4	2941	48.75
5	1972	56.40
6	2627	59.86
7	2785	48.34
8	2319	63.62
9	2684	57.18
10	1504	52.34
11	1581	55.42
12	1316	48.19
13	3384	64.57
14	3154	48.71
15	2378	66.20
16	1565	55.50
17	1387	49.88
18	1135	53.64
X	728	187.25
Total	43837	

Note: ¹Data from *Sus scrofa* Build 10.2.

GWAS for vertebrae number showed that a total of 36 genome-wide significant SNPs were detected on SSC1 and SSC7. The Manhattan plot obtained from GWAS is shown in Fig. 1. These results are consistent with previously reported QTLs^[3]. When thoracic-lumbar vertebrae were used as the phenotypes instead of total vertebrae, the QTLs for number of thoracic vertebrae were focused on SSC7^[15]. Also in their research, a putative QTL for number of lumbar vertebrae was detected on SSC1 although it did not reach a significant level. Therefore, the total number of thoracic-lumbar vertebrae used as phenotype may be the possible reason for two QTLs being detected in the current study and that of Mikawa et al.^[3].

3.2 6.04-Mb region includes all 13 genome-wide significant SNPs on SSC1

On SSC1, there were 13 genome-wide significant SNPs associated with vertebrae number, within a 6.04-Mb (294.28–300.32 Mb) region. The most significant SNP could explain 25.6% of phenotypic variance. Several previous studies have mapped the QTL for vertebrae number to a similar interval on SSC1^[2–5,16,17]. Using MASS, five F₁ boars proved to be heterozygous Qq

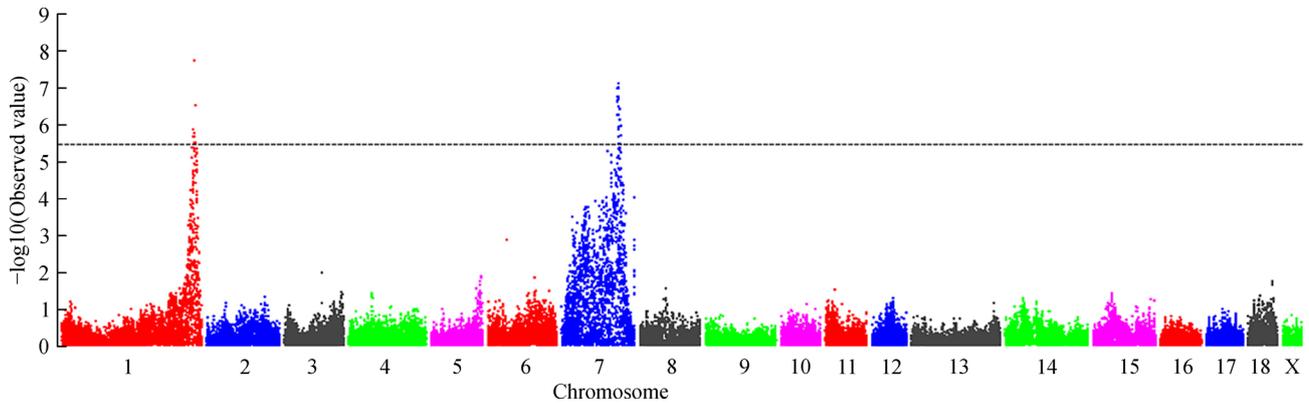


Fig. 1 Manhattan plot of genome-wide association study of vertebrae number in a population of Large White × Minzhu pigs. Chromosomes 1–18, and X are shown in different colors. The horizontal line indicates the genome-wide significance level ($-\log_{10}(4.15 \times 10^{-6})$).

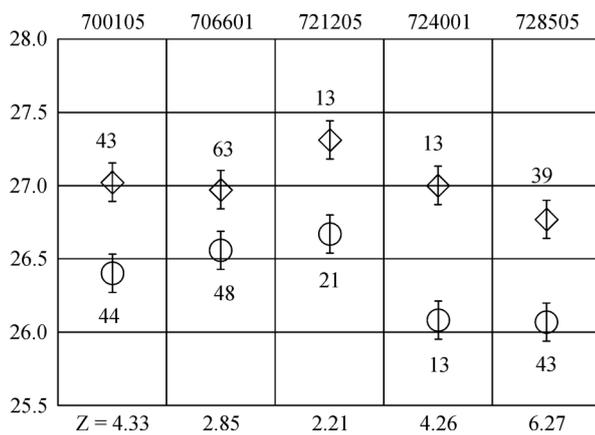


Fig. 2 The marker-assisted segregation analysis on SSC1 for F₁ boars in a population of Large White × Minzhu pigs. The marker-assisted segregation analysis on SSC1 for F₁ boars. The graphs show for half-sib pedigrees for five F₁ boars (700105, 706601, 721205, 724001 and 728505) the phenotypic mean ± standard errors of the offspring sorted in two groups according to the homolog inherited from the sire. The number of offspring in each group is given above the respective error bars. The graph corresponds to the boars that were shown to be heterozygous Qq and gives a Z-score for each pedigree. Q alleles associated with a positive allele substitution effect on vertebrae number are marked by a diamond, q alleles by a circle.

genotypes (Fig. 2). However, haplotype sharing analysis showed that all five chromosomes in the Q pool share an identical descent haplotype spanning all 13 SNPs and the analysis did not refine this 6.04-Mb region using the 13 significant SNPs (data not shown). Previous fine mapping to the QTL showed that *NR6A1* was a strong candidate for the QTL and that the most likely causative mutation is a base substitution, *NR6A1* c.748C > T, which results in a proline to leucine substitution at codon 192^[4]. The strongest selection signature was observed for a locus on chromosome 1, which includes *NR6A1*^[18]. The distribu-

tion of *NR6A1* c.748C > T in different pig breeds showed that the C and T alleles were almost fixed in wild boars and western commercial breeds, respectively^[19,20]. All three genotypes are represented in European and Chinese indigenous pig breeds^[19,21]. Although the T allele was fixed in all F₀ Large White in the current work, the allele frequency was over 50% in Minzhu (Table S3). In the current study, the association of *NR6A1* c.748C > T with trait did not reach significant genome-wide association (Table 2). From a previous report, the *NR6A1* c.748C > T was predicted as causal mutation although its further functional identification is lacking^[4]. Hence more work is required to identify the function of this gene in the future.

3.3 7.17-Mb region includes all 23 genome-wide significant SNPs on SSC7

The significant genome-wide SNPs on SSC7 are shown in Table 2. A total of 23 SNPs were significantly associated within a 7.17-Mb region (102.22–109.39 Mb) on SSC7. The most significant SNP was ASGA0035537 and could explain 20.6% of the phenotypic variance. Similar to this result, a previous study reported that the QTL on SSC7 for vertebrae number could explain 4.2%–37.3% of the phenotypic variance in several distinct pig populations^[3]. Several previous studies have reported similar findings when they mapped the major QTL associated with the vertebrae number on SSC7 using different populations. This QTL was first reported to be on SSC7 in an F₂ Meishan × Duroc resource population^[16]. Subsequently, research in several populations derived from crosses between Western pig breeds (Large White, Duroc, Berkshire and Landrace) and Chinese indigenous pig breeds (Meishan and Jinhua), revealed that the identical QTL was located on SSC7^[3]. Moreover, this QTL on SSC7 was repeatedly identified in an F₂ population crossed from the commercial breeds Duroc and Pietrain^[22].

Table 2 Genome-wide significant SNPs associated with vertebrae number in a population of Large White × Minzhu pigs

SNP	Chromosome	Position ¹	Number	P-value	Nearest gene ²	Distance/bp	Var ³ /%
ALGA0010073	1	294277597	516	1.42×10 ⁻⁶	<i>LOC100152823</i>	within	
INRA0007489	1	294523343	523	2.20×10 ⁻⁶	<i>NDUFA8</i>	21483	
ASGA0007558	1	297461876	523	2.26×10 ⁻⁶	<i>RABGAP1</i>	within	
DRGA0002521	1	297585240	523	2.26×10 ⁻⁶	<i>STRBP</i>	24284	
ASGA0007561	1	297725378	523	3.17×10 ⁻⁶	<i>LOC100627889</i>	10890	
H3GA0004851	1	298042326	523	2.16×10 ⁻⁶	<i>LOC102165702</i>	within	
ALGA0010388	1	298117818	523	3.64×10 ⁻⁶	<i>DENND1A</i>	within	
ASGA0007592	1	298890319	523	1.85×10 ⁻⁶	<i>NEK6</i>	within	
H3GA0004878	1	298951152	523	3.29×10 ⁻⁶	<i>PSMB7</i>	within	
H3GA0004881	1	298972575	523	1.91×10 ⁻⁶	<i>PSMB7</i>	within	25.62
NR6A1748C > T ⁴	1	299084752	500	8.5×10 ⁻⁶	<i>NR6A1</i>	within	
SIRI0001487	1	299578776	523	3.19×10 ⁻⁷	<i>SCAI</i>	within	
INRA0007601	1	299599913	523	3.19×10 ⁻⁷	<i>SCAI</i>	within	
ALGA0010455	1	300316974	523	3.3×10 ⁻⁶	<i>PBX3</i>	within	
INRA0027600	7	102222702	523	2.46×10 ⁻⁷	<i>RBM25</i>	within	
INRA0027605	7	102739058	523	5.85×10 ⁻⁷	<i>ACOT6</i>	3748	
ALGA0043941	7	103001352	523	2.26×10 ⁻⁷	<i>FAM161B</i>	within	
ALGA0043942	7	103020388	523	1.89×10 ⁻⁷	<i>COQ6</i>	within	
INRA0027623	7	103366820	523	5.81×10 ⁻⁷	<i>VSX2</i>	8605	
VRTN19034A > C ⁵	7	103457401	498	6.85×10 ⁻⁷	<i>VRTN</i>	within	
M1GA0010654	7	103796933	523	1.12×10 ⁻⁷	<i>FCF1</i>	within	
ALGA0043962	7	103816521	523	1.12×10 ⁻⁷	<i>YLPM1</i>	2238	
H3GA0022664	7	103910821	523	3.62×10 ⁻⁶	<i>PROX2</i>	within	
ASGA0035536	7	104108293	514	1.05×10 ⁻⁷	<i>ACYPI</i>	within	
ASGA0035537	7	104219054	523	8.38×10 ⁻⁸	<i>TMED10</i>	12301	20.56
ALGA0108658	7	104546250	522	1.20×10 ⁻⁶	<i>BATF</i>	21512	
DBNP0000926	7	104807152	523	2.11×10 ⁻⁷	<i>LOC100525758</i>	within	
ASGA0035551	7	105022346	523	5.70×10 ⁻⁷	<i>LOC100739559</i>	within	
TGFβ3c.1749G > A ⁶	7	105179474	521	2.98×10 ⁻⁷	<i>TGFβ3</i>	within	
SIRI0001067	7	105182819	523	3.37×10 ⁻⁷	<i>TGFβ3</i>	within	
MARC0027367	7	105314744	523	1.47×10 ⁻⁶	<i>C7H14orf118</i>	56124	
ALGA0044022	7	105383136	523	1.89×10 ⁻⁷	<i>C7H14orf118</i>	within	
H3GA0022720	7	105752757	523	1.36×10 ⁻⁶	<i>ESRRB</i>	9247	
ALGA0044071	7	106028315	523	2.25×10 ⁻⁶	<i>C7H14orf166B</i>	within	
MARC0034477	7	106779379	523	4.00×10 ⁻⁷	<i>AHSA1</i>	14272	
ALGA0044124	7	106793499	521	3.47×10 ⁻⁶	<i>LOC102162795</i>	within	
ALGA0044211	7	107725310	523	8.00×10 ⁻⁷	<i>NRXN3</i>	within	
ASGA0035786	7	109200786	522	2.05×10 ⁻⁶	<i>DIO2</i>	98537	
H3GA0022821	7	109387763	523	1.16×10 ⁻⁶	<i>LOC100519376</i>	32903	

Note: ¹Data from *Sus scrofa* Build 10.2; ²Gene symbols are as in GenBank; ³Phenotypic variation explained by the most significant SNPs on SSC 1 and 7; ⁴The reported putative casual mutation of *NR6A1* c.748C > T; ⁵The reported putative casual mutation of *VRTN* 19034A > C; ⁶The detected mutation in *TGFβ3* in this study.

3.4 Haplotype sharing and candidate gene analysis of SSC7

Using MASS, five of the nine F₁ boars proved to be heterozygous Qq genotypes (Fig. 3). Further, we judged the QTL status of the five F₁ boars with Qq genotypes by multiple comparisons of targeted chromosomes with the reference Q (number-increase) or q (wild-type). Visual examination of the Q pools immediately reveals that all five chromosomes in the Q pool share an identical by descent haplotype spanning the ASGA0035536-ALGA0044124 interval (Fig. 4). A total of 29 annotated genes in GenBank were contained within the 2.69-Mb region containing 13 significant SNPs on SSC7. Of these genes, transmembrane p24 trafficking protein 10 (*TMED10*) was the closest gene to the most significant SNP on SSC7. *TMED10* belongs to the highly conserved p24 family of type I transmembrane proteins and its main function is to regulate anterograde and retrograde transport in the early secretory pathway between the endoplasmic reticulum and Golgi apparatus^[23–25]. According to a previous study in mouse embryos, *TMED10* was found to be expressed at embryonic day 10.5^[26], which is the critical period for vertebral differentiation^[27]. Although little is known concerning the influence of *TMED10* on vertebral differentiation, this gene may be a candidate as it is the most significant SNP in the current study. The Finkel-Biskis-Jenkins murine osteosarcoma viral oncogene homolog (*FOS*), one of 29 genes, has an essential role in the osteoclastic differentiation of precursor cells and the

upregulation of Receptor Activator of Nuclear Factor κ B (*RANK*) expression in osteoclast precursors within the bone environment^[28]. This gene has been reported to be a candidate for vertebrae number^[17]. However, a candidate gene of interest is *TGF β 3*, in which a significant SNP, SIRI0001067 was located in this GWAS. The development of the vertebral column is a consequence of a segment-specific balance between proliferation, apoptosis and differentiation of mesoderm cells in embryos^[29]. In mammals, *TGF β 3* promotes chondrogenesis in posterofrontal suture-derived mesenchymal cells, influencing different stages of chondrogenic differentiation and proliferation^[30]. *TGF β 3* protein, in combination with its downstream factor, TGF beta receptor type I (*ALK5*), regulates the differentiation and proliferation of the spinal column^[31]. We resequenced all of UTRs and exons of *TGF β 3* and found only one SNP, which was located in the 3'-UTR (1749 bp of NM_214198 in GenBank), i.e., c.1749G > A. Also, a fixed G allele (increasing number) was detected in each F₀ Large White pig (Table S3). Rerunning GWAS suggested that *TGF β 3* c.1749G > A had significant ($P = 2.98 \times 10^{-7}$) genome-wide association with vertebrae number in the F₂ population. These findings support the proposal that *TGF β 3* c.1749G > A might be a candidate causal mutation, or is closely linked to a causal mutation, on SSC7 for vertebrae number in our population. However, *VRTN*, which has been reported to be a strong candidate gene^[5,15,32], was not contained in this shared region. Two putative causal mutations, i.e., g.19034A > C and g.20311_20312ins291 of *VRTN*, have been reported previously^[15]. Genotyping the two SNPs in our population revealed that only *VRTN* g.19034A > C has polymorphisms. Rerunning GWAS suggested that *VRTN* g.19034A > C showed significant genome-wide association (Table 2). Genotype detection in the F₀ pigs showed that none of alleles of *VRTN* g.19034A > C were fixed in either breed (Table S3). When the *VRTN* g.19034A > C was treated as a fixed effect to rerun GWAS, there was also an obvious peak (Fig. S1) for association, even if the top SNP (M1GA0010629, $P = 8.85 \times 10^{-5}$) did not reach significant chromosome-wide association. These results suggest *VRTN* g.19034A > C could be proposed as a casual mutation influencing the vertebral number.

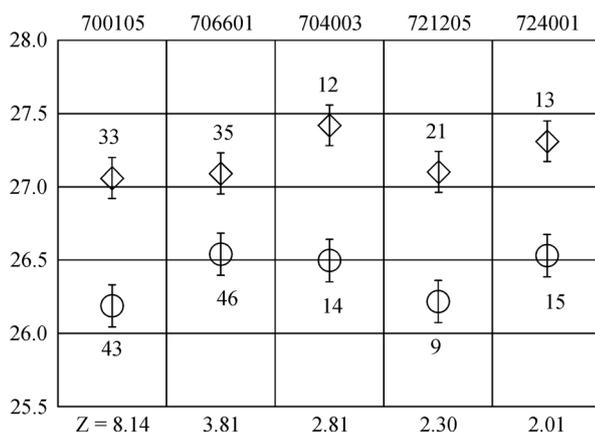


Fig. 3 The marker-assisted segregation analysis on SSC7 for F₁ boars in a population of Large White \times Minzhu pigs. The marker-assisted segregation analysis on SSC7 for F₁ boars. The graphs show for half-sib pedigrees of five F₁ boars (700105, 706601, 704003, 721205 and 724001) the phenotypic mean \pm standard errors of the offspring sorted in two groups according to the homolog inherited from the sire. The number of offspring in each group is given above the respective error bars. The graph corresponds to the boars that were shown to be heterozygous Qq and reports a Z-score for each pedigree. Q alleles associated with a positive allele substitution effect on vertebrae number are marked by a diamond, q alleles by a circle.

4 Conclusions

In summary, this study focused on vertebrae numbers in pigs. A genome-wide search identified 13 significant SNPs within a 6.04-Mb region containing the reported causal gene *NR6A1* on SSC1. An additional 23 SNPs were identified within a 7.17-Mb region on SSC7 that showed genome-wide association with vertebrae number in our population. By identifying a haplotype sharing and haplotype block analysis, the QTL was refined to a chromosome segment of about 3 Mb on SSC7. In addition

SNP	H3GA0022821	ASGA0038786	ALGA0044211	ALGA004124	MARCO034477	ALGA0044071	H3GA0022720	ALGA0044022	MARCO027367	SIRI0001067	ASGA0035551	DBNP0000926	ALGA0108658	ASGA0035537	ASGA0035536	H3GA0022664	ALGA0043962	MIGA0010654	VRTN19034A>C	INRA0027623	ALGA0043942	ALGA0043941	INRA0027605	INRA0027600		
724001	2	1	1	2	2	1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	Q
700105	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	Q
706601	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	Q
721205	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	Q
704003	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	Q

724001	2	1	1	1	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	q
700105	2	2	1	1	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	q
706601	2	2	1	1	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	q
721205	2	2	1	1	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	q
704003	2	2	1	1	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	q

Fig. 4 Haplotype sharing analysis in the 7.17-Mb region on SSC7 in a population of Large White × Minzhu pigs. Shared haplotypes of five F₁ boars with Q chromosome (Large White × Minzhu intercross population) were analyzed. Polymorphisms are displayed at the respective SNP markers. For these markers the alleles associated with high or lower vertebrae number QQ genotypes are denoted 1 or 2, respectively. The shared haplotype blocks are highlighted.

to the reported candidates, *VRTN* and *TMED10*, *TGFβ3* on SSC7 could also be good candidate for determining vertebrae number in pigs.

Supplementary materials The online version of this article at <http://dx.doi.org/10.15302/J-FASE-2017163> contains supplementary materials (Tables S1–S3; Fig. S1).

Acknowledgements This research was supported by the Agricultural Science and Technology Innovation Program (ASTIP-IAS02), National Key Technology R&D Program of China (2015BAD03B02), Earmarked Fund for Modern Agro-industry Technology Research System.

Compliance with ethics guidelines Longchao Zhang, Jingwei Yue, Xin Liu, Jing Liang, Kebin Zhao, Hua Yan, Na Li, Lei Pu, Yuebo Zhang, Huibi Shi, Ligang Wang, and Lixian Wang declare that they have no conflict of interest.

All applicable institutional and national guidelines for the care and use of animals were followed.

References

- King J, Roberts R. Carcass length in the bacon pig; its association with vertebrae numbers and prediction from radiographs of the young pig. *Animal Science*, 1960, **2**(1): 59–65
- Wada Y, Akita T, Awata T, Furukawa T, Sugai N, Inage Y, Ishii K, Ito Y, Kobayashi E, Kusumoto H, Matsumoto T, Mikawa S, Miyake M, Murase A, Shimanuki S, Sugiyama T, Uchida Y, Yanai S, Yasue H. Quantitative trait loci (QTL) analysis in a Meishan × Göttingen cross population. *Animal Genetics*, 2000, **31**(6): 376–384
- Mikawa S, Hayashi T, Nii M, Shimanuki S, Morozumi T, Awata T. Two quantitative trait loci on *Sus scrofa* chromosomes 1 and 7 affecting the number of vertebrae. *Journal of Animal Science*, 2005, **83**(10): 2247–2254
- Mikawa S, Morozumi T, Shimanuki S, Hayashi T, Uenishi H, Domukai M, Okumura N, Awata T. 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*, 2007, **17**(5): 586–593
- Mikawa S, Sato S, Nii M, Morozumi T, Yoshioka G, Imaeda N, Yamaguchi T, Hayashi T, Awata T. Identification of a second gene associated with variation in vertebral number in domestic pigs. *BMC Genetics*, 2011, **12**(1): 5
- Miller S A, Dykes D D, Polesky H F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 1988, **16**(3): 1215
- Zhang L C, Li N, Liu X, Liang J, Yan H, Zhao K B, Pu L, Shi H B, Zhang Y B, Wang L G, Wang L X. A genome-wide association study of limb bone length using a Large White × Minzhu intercross population. *Genetics, Selection, Evolution*, 2014, **46**(1): 56
- Aulchenko Y S, de Koning D J, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics*, 2007, **177**(1): 577–585
- Amin N, van Duijn C M, Aulchenko Y S. A genomic background based method for association analysis in related individuals. *PLoS One*, 2007, **2**(12): e1274
- Hao K, Li C, Rosenow C, Wong W H. Detect and adjust for population stratification in population-based association study using genomic control markers: an application of Affymetrix Genechip Human Mapping 10K array. *European Journal of Human Genetics*, 2004, **12**(12): 1001–1006
- Gao X, Starmer J, Martin E R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology*, 2008, **32**(4): 361–369
- Yang Q, Cui J, Chazaro I, Cupples L A, Demissie S. Power and type I error rate of false discovery rate approaches in genome-wide

- association studies. *BMC Genetics*, 2005, **6**(Suppl 1): S134
13. Nezer C, Collette C, Moreau L, Brouwers B, Kim J J, Giuffra E, Buys N, Andersson L, Georges M. Haplotype sharing refines the location of an imprinted quantitative trait locus with major effect on muscle mass to a 250-kb chromosome segment containing the porcine IGF2 gene. *Genetics*, 2003, **165**(1): 277–285
 14. Barrett J C, Fry B, Maller J, Daly M J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 2005, **21**(2): 263–265
 15. Fan Y, Xing Y, Zhang Z, Ai H, Ouyang Z, Ouyang J, Yang M, Li P, Chen Y, Gao J, Li L, Huang L, Ren J. A further look at porcine chromosome 7 reveals VRTN variants associated with vertebral number in Chinese and Western pigs. *PLoS One*, 2013, **8**(4): e62534
 16. Sato S, Oyamada Y, Atsugi K, Nade T, Sato S, Kobayashi E, Mitsuhashi T, Nirasawa K, Komatsuda A, Saito Y, Terai S, Hayashi T, Sugimoto Y. Quantitative trait loci analysis for growth and carcass traits in a Meishan x Duroc F2 resource population. *Journal of Animal Science*, 2003, **81**(12): 2938–2949
 17. Ren D R, Ren J, Ruan G F, Guo Y M, Wu L H, Yang G C, Zhou L H, Li L, Zhang Z Y, Huang L S. Mapping and fine mapping of quantitative trait loci for the number of vertebrae in a White Duroc × Chinese Erhualian intercross resource population. *Animal Genetics*, 2012, **43**(5): 545–551
 18. Rubin C J, Megens H J, Barrio A M, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg O, Jern P, Jorgensen C B, Archibald A L, Fredholm M, Groenen M A M, Andersson L. Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, **109**(48): 19529–19536
 19. Burgos C, Latorre P, Altarriba J, Carrodegua J A, Varona L, López-Buesa P. Allelic frequencies of NR6A1 and VRTN, two genes that affect vertebrae number in diverse pig breeds: a study of the effects of the VRTN insertion on phenotypic traits of a Duroc × Landrace-Large White cross. *Meat Science*, 2015, **100**: 150–155
 20. Fontanesi L, Ribani A, Scotti E, Utzeri V J, Veličković N, Dall'Olio S. Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the *MC1R* and *NR6A1* genes. *Meat Science*, 2014, **98**(4): 781–784
 21. Yang G, Ren J, Zhang Z, Huang L. Genetic evidence for the introgression of Western NR6A1 haplotype into Chinese Licha breed associated with increased vertebral number. *Animal Genetics*, 2009, **40**(2): 247–250
 22. Edwards D B, Ernst C W, Raney N E, Doumit M E, Hoge M D, Bates R O. Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. Carcass and meat quality traits. *Journal of Animal Science*, 2008, **86**(2): 254–266
 23. Rojo M, Pepperkok R, Emery G, Kellner R, Stang E, Parton R G, Gruenberg J. Involvement of the transmembrane protein p23 in biosynthetic protein transport. *Journal of Cell Biology*, 1997, **139**(5): 1119–1135
 24. Rojo M, Emery G, Marjomäki V, McDowall A W, Parton R G, Gruenberg J. The transmembrane protein p23 contributes to the organization of the Golgi apparatus. *Journal of Cell Science*, 2000, **113**(6): 1043–1057
 25. Bonnon C, Wendeler M W, Paccaud J P, Hauri H P. Selective export of human GPI-anchored proteins from the endoplasmic reticulum. *Journal of Cell Science*, 2010, **123**(10): 1705–1715
 26. Jerome-Majewska L A, Achkar T, Luo L, Lupu F, Lacy E. The trafficking protein Tmed2/p24β(1) is required for morphogenesis of the mouse embryo and placenta. *Developmental Biology*, 2010, **341**(1): 154–166
 27. Shipley J M, Mecham R P, Maus E, Bonadio J, Rosenbloom J, McCarthy R T, Baumann M L, Frankfater C, Segade F, Shapiro S D. Developmental expression of latent transforming growth factor β binding protein 2 and its requirement early in mouse development. *Molecular and Cellular Biology*, 2000, **20**(13): 4879–4887
 28. Arai A, Mizoguchi T, Harada S, Kobayashi Y, Nakamichi Y, Yasuda H, Penninger J M, Yamada K, Udagawa N, Takahashi N. Fos plays an essential role in the upregulation of RANK expression in osteoclast precursors within the bone microenvironment. *Journal of Cell Science*, 2012, **125**(12): 2910–2917
 29. Christ B, Huang R, Wilting J. The development of the avian vertebral column. *Anatomy and Embryology*, 2000, **202**(3): 179–194
 30. James A W, Xu Y, Lee J K, Wang R, Longaker M T. Differential Effects of transforming growth factor-beta1 and -beta3 on chondrogenesis in postero-frontal cranial suture-derived mesenchymal cells *in vitro*. *Plastic and Reconstructive Surgery*, 2009, **123**(1): 31
 31. Andersson O, Reissmann E, Ibáñez C F. Growth differentiation factor 11 signals through the transforming growth factor-β receptor ALK5 to regionalize the anterior-posterior axis. *EMBO Reports*, 2006, **7**(8): 831–837
 32. Hirose K, Mikawa S, Okumura N, Noguchi G, Fukawa K, Kanaya N, Mikawa A, Arakawa A, Ito T, Hayashi Y, Tachibana F, Awata T. Association of swine vertnin (*VRTN*) gene with production traits in Duroc pigs improved using a closed nucleus breeding system. *Animal Science Journal*, 2013, **84**(3): 213–221