

1 **Genome-wide Association Study for Number of Vertebrae in an**
2 **F2 Large White × Minzhu Population of Pigs**

3 **Authors:** Longchao Zhang[§], Xin Liu[§], Jing Liang, Kebin Zhao, Hua Yan, Na Li, Lei
4 Pu, Yuebo Zhang, Huibi Shi, Ligang Wang¹, and Lixian Wang¹

5 **Institutional addresses:**

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

9 [§]These authors contributed equally to this work.

10 ¹ Corresponding author:

11 Lixian Wang, Institute of Animal Science, Chinese Academy of Agricultural Sciences,
12 Beijing 100193, China. Fax: +86-10-62818771. Email: iaswlx@263.net

13 Ligang Wang, Institute of Animal Science, Chinese Academy of Agricultural Sciences,
14 Beijing 100193, China. Fax: +86-10-62818771. Email: ligwang@126.com

15 **Emails of other authors:**

16 Longchao Zhang: zhlchias@163.com Xin Liu: firstliuxin@163.com

17 Jing Liang: jing_224@126.com Kebin Zhao: iaszkb@sina.com

18 Hua Yan: zcyyh@126.com Na Li: Lina_0507@126.com

19 Lei Pu: pulei87@126.com Yuebo Zhang: ybzhangfd@126.com

20 Huibi Shi: 785706092@qq.com

21 **ABSTRACT**

22 Porcine carcass that is approximately 800 mm long may be expected to have one
23 additional vertebra. Therefore, the number of vertebrae in pigs is an economically
24 important trait. To examine the genetic basis of this trait, we genotyped 593 F2 Large
25 White × Minzhu pigs using the Porcine SNP60K BeadChip. A genome-wide
26 association study identified 39 significant single-nucleotide polymorphisms (SNPs)
27 on the chromosomes SSC1 and SSC7. An 8.82-Mb region that contained all 21
28 significant SNPs on SSC1 harbored the gene *NR6A1*, previously reported to influence
29 the number of vertebrae in pigs. The remaining 18 significant SNPs on SSC7 were
30 concentrated in a 4.56-Mb region, which was within the quantitative trait loci interval
31 for number of vertebrae. A haplotype sharing analysis refined the locus to a segment
32 of ~637 Kb. The most significant SNP, SIRI0001067, was contained in this refined
33 region on SSC7 and located in one of the introns of *TGFB3*. As *TGFB3* influences the
34 development of vertebrae in mammalian embryos, the gene may be another strong
35 candidate for the number of vertebrae in pigs.

36 **KEYWORDS:** genome-wide association study; number of vertebrae; pig; SSC7;
37 *TGFB3*.

38 Vertebrae consist of the morphologically differentiated cervical, thoracic, lumbar,
39 sacral, and caudal vertebrae, and the number of these varies in pigs (King and Roberts
40 1960). Wild boar, the ancestor of domestic pigs, have a uniform 26 cervical–lumbar
41 vertebrae, whereas most Chinese indigenous pig breeds have 26 or 27 vertebrae
42 (Zhang *et al.* 1986). By comparison, Western commercial pig breeds, such as
43 Landrace and Large White, have 27–29 vertebrae (Mikawa *et al.* 2005). The number
44 of cervical–lumbar vertebrae is usually associated with carcass length and is an
45 economically important trait; a greater number of vertebrae in pigs is associated with
46 higher economic value. According to a study by King and Roberts, a carcass that is
47 approximately 800 mm long may be expected to have one additional vertebra (King
48 and Roberts 1960). Understanding the genetic basis for number of vertebrae can offer
49 insights into the mechanism for vertebral developmental in mammals and provide
50 genetic markers to aid molecular breeding of pigs.

51 Locating quantitative trait loci (QTLs) and gene-mining for number of vertebrae
52 have recently become research interests. Wada *et al.* (2000) first reported two QTLs on
53 wild boar (*Sus scrofa*) chromosomes SSC1 and SSC2 that are associated with the
54 number of vertebrae. Subsequently, using several different pig populations, two QTLs
55 for the number of vertebrae were mapped to SSC1 and SSC7 (Mikawa *et al.* 2005).
56 Using three F2 experimental families, a fine mapping of the QTL on SSC1 was
57 performed and showed that nuclear receptor subfamily 6, group A, member 1
58 (*NR6A1*), was a strong candidate gene that appeared to influence the number of
59 vertebrae in pig (Mikawa *et al.* 2007). Gene-mining to the other QTL on SSC7

60 suggested that the vertebrae development homologue (VRTN) could be a good
61 candidate (Mikawa *et al.* 2011). A genome-wide association study (GWAS), based on
62 the high density of genome-wide single-nucleotide polymorphisms (SNPs), has been a
63 more efficient method to not only identify QTLs but also mine major new genes.
64 Additional methods and different populations are required to identify loci that are
65 associated with the number of ribs and detect good candidate genes. The objectives of
66 this study were to identify SNPs associated with the number of vertebrae using
67 GWAS in an F2 Large White × Minzhu population of pigs.

68

69 **MATERIALS AND METHODS**

70 **Ethics statement**

71 All animal procedures were performed according to the guidelines developed by
72 the China Council on Animal Care, and all protocols were approved by the Animal
73 Care and Use Committee at Institute of Animal Science, Chinese Academy of
74 Agricultural Sciences. All efforts were made to minimize the suffering of animals.

75

76 **Population and phenotypic data**

77 The F0 population was generated using four Large White boars and 16 Minzhu
78 sows. In the F1 generation, nine boars and 46 sows were selected to mate to produce
79 598 F2 individuals. All animals in the F2 generation were born in three parties and 94
80 litters. Each F2 male was castrated 3 days after birth. All animals were maintained
81 under uniform housing conditions and were fed the same fodder. When F2 animals

82 arrived at 240 ± 7 days, they were slaughtered in 52 batches. The numbers of thoracic
83 and lumbar vertebrae were counted from the carcasses.

84

85 **Genotyping and quality control**

86 Genomic DNA was extracted using standard methods (Miller *et al.* 1988). The
87 Illumina SNP60K chip for pigs was employed to genotype all individuals. As our
88 previous study (Zhang *et al.* 2014), quality control parameters for single nucleotide
89 polymorphisms included the following: call rate $> 90\%$, minor allele frequency
90 (MAF) $> 3\%$, and Hardy–Weinberg equilibrium (HWE) with a P -value $> 10^{-6}$.

91

92 **GWAS**

93 The genome-wide rapid association using the mixed model and regression-genomic
94 control approach (Aulchenko *et al.* 2007; Amin *et al.* 2007) was used in the present
95 study. Three fixed effects (i.e., sex, parity, and batch), a random effect (litter), and a
96 covariate (body weight) were used as inputs. DMU (Madsen *et al.* 2006) and
97 GenABEL software (Amin *et al.* 2007) in the R environment were employed to
98 estimate the residuals of traits for each individual and perform the GWAS. After
99 Bonferroni correction, genome-wide significance was accepted at a p -value of
100 $0.05/\text{number of SNPs passing quality control}$ (Yang *et al.* 2005).

101

102 **Haplotype sharing and linkage disequilibrium analysis**

103 The genotypes of F1 sires were confirmed using marker-assisted segregation
104 analysis (MASS) (Nezer *et al.* 2003). According to the SNP, H3GA0004881, the
105 genotype of each boar was checked from a z-score as the \log_{10} likelihood ratio
106 (L_{H1}/L_{H0}), where L_{H1} supposes that the sire is heterozygous (Qq), and L_{H0} supposes
107 that the sire is homozygous (QQ or qq). Sires were regarded as Qq, if $z > 2$; QQ or qq
108 if $z < -2$; and of uncertain genotype when $2 > z > -2$. The Q-bearing chromosomes of
109 F1 boars and top 10% of F2 population were segregated by MASS. Haplotype sharing
110 analysis was done according to all of significant SNPs on SSC7.

111

112 **RESULTS AND DISCUSSION**

113 **Phenotype descriptions**

114 The number of vertebrae in the Large White \times Minzhu intercross population may
115 be found in Table 1. The pigs had 24 (n = 1), 25 (n = 9), 26 (n = 244), 27 (n = 296),
116 28 (n = 42), or 29 (n = 1) vertebrae and the mean number of ribs was 26.63.

117

118 **SNPs on SSC1 containing NR6A1**

119 The final data set that passed quality control screening and was used in the analysis
120 contained 47,615 SNPs and came from 585 F2 individuals. The distribution of SNPs
121 after quality control and the average distance between adjacent SNPs on each
122 chromosome are shown in Table 2. The results of the GWAS for number of vertebrae
123 are shown in Table 3. The Manhattan plot obtained from GWAS is shown in Figure 1.
124 On SSC1, there were 21 genome-wide significant SNPs associated with number of

125 vertebrae, within an 8.82-Mb (293.93–302.75 Mb) region. Several previous studies
126 have mapped the QTL for number of vertebrae to a similar interval on SSC1 (Mikawa
127 *et al.* 2005, 2007, 2011; Wada *et al.* 2000; Sato *et al.* 2003; Ren *et al.* 2012). Fine
128 mapping to the QTL showed that *NR6A1* is a strong candidate for the QTL and that
129 the most likely causative mutation is a base substitution, *NR6A1* c.748 C > T, which
130 results in a proline to leucine substitution at codon 192 (Mikawa *et al.* 2007). The
131 strongest signature of selection was observed for a locus on chromosome 1, which
132 includes *NR6A1* (Rubin *et al.* 2012). The distribution of *NR6A1* c.748 C > T in
133 different pig breeds showed that the C and T alleles were almost fixed in wild boars
134 and western commercial breeds, respectively (Burgos *et al.* 2014; Fontanesi *et al.*
135 2014). All three genotypes are represented in European and Chinese indigenous pig
136 breeds (Burgos *et al.* 2014; Yang *et al.* 2009). In the present study, the most
137 significant SNP, H3GA0004881, was near *NR6A1*, indicating that this gene is a strong
138 candidate for the number of vertebrae in pigs.

139

140 **Genome-wide association of SNPs on SSC7 with number of vertebrae**

141 The genome-wide significant SNPs on SSC7 are shown in Table 2. A total of 18
142 SNPs were significantly associated within a 4.56-Mb region (102.22–106.78 Mb) on
143 SSC7. The most significant SNP was H3GA0004881. Several previous studies have
144 reported similar findings when they mapped the major QTL associated with the
145 number of vertebrae on SSC7 using different populations. This QTL was first reported
146 to be on SSC7 in an F2 Meishan × Duroc resource population (Sato *et al.* 2003).

147 Subsequently, research in several populations derived from crosses between Western
148 pig breeds (Large White, Duroc, Berkshire, and Landrace) and Chinese indigenous
149 pig breeds (Meishan and Jinhua), revealed that the identical QTL was located on
150 SSC7 (Mikawa *et al.* 2005). Moreover, this QTL on SSC7 was repeatedly identified in
151 an F2 population crossed from the commercial breeds Duroc and Pietrain (Edwards *et*
152 *al.* 2008). Gene-mining in the QTL region revealed that *VRTN* could be a strong
153 candidate gene and the Q/Q homozygotes could increase ~1 vertebra over the wt/wt
154 (Mikawa *et al.* 2011; Fan *et al.* 2013). In addition, the *VRTN* Q/Q genotype was
155 reported to significantly increase body length by ~1 cm (Hirose *et al.* 2013). *VRTN*
156 was also located in this GWAS region and could be a good candidate for number of
157 vertebrae in pigs.

158

159 **Haplotype sharing analysis to refine the QTL on SSC7**

160 Using the most significant SNP, SIRI0001067, MASS identified three out of nine
161 F1 boars as heterozygous (Qq genotype), while the rest were of undetermined
162 genotype (Figure 2). F2 homozygous (QQ on SIRI0001067) individuals, which were
163 in the top 10% of the population, were used to determine the shared haplotype. Visual
164 examination of all the previously mentioned populations revealed a ~637-Kb shared
165 haplotype, SIRI0001067, which contained the most significant SNPs (Figure 3). The
166 sharing region contained a total of 16 genes in GenBank. Finkel-Biskis-Jinkins
167 murine osteosarcoma viral oncogene homologue (*FOS*), one of 16 genes, plays
168 essential roles in the osteoclastic differentiation of precursor cells and the

169 upregulation of Receptor Activator of Nuclear Factor κ B (RANK) expression in
170 osteoclast precursors within the bone environment (Arai *et al.* 2012). This gene has
171 been reported to be a candidate for number of vertebrae (Ren *et al.* 2012).

172 However, the most interesting gene in this region is transforming growth factor,
173 beta 3 (*TGFB3*). This GWAS located the most significant SNP, SIRI0001067, in the
174 intron of *TGFB3*. The development of the vertebral column is a consequence of a
175 segment-specific balance between proliferation, apoptosis, and differentiation of
176 mesoderm cells in embryos (Christ *et al.* 2000). Gene expression analysis shows that
177 *TGFB3* is abundant in the growing undifferentiated mesoderm (Lorda-Diez *et al.*
178 2010). In zebrafish, *TGFB3* is moderately expressed in 14- to 24-somite embryos
179 (Cheah *et al.* 2005). In mammals, *TGFB3* promotes chondrogenesis in posterfrontal
180 suture-derived mesenchymal cells, influencing different stages of chondrogenic
181 differentiation and proliferation (James *et al.* 2009). *TGFB3* plays a critical role in
182 vertebral column development by increasing the proliferation of mesoderm cells.
183 *TGFB3* protein, in combination with its downstream factor, TGF beta receptor type I
184 (*ALK5*), regulates the differentiation and proliferation of the spinal column (Zhao *et al.*
185 2014) and has been associated with the biological functions of *TGFB3* proteins.
186 Compared with the wild type, *Alk5^{+/+}* and *Alk5^{+/-}* mice had an additional 3 and 5
187 vertebrae, respectively (Andersson *et al.* 2006). Although *TGFB3* knockout mice
188 (*TGFB3^{-/-}*) have been created, they died soon after birth and the effect of gene
189 knockout on number of vertebrae was not recorded (Kartinen *et al.* 1995). We
190 believe that these GWAS findings support the idea that *TGFB3* might be a strong

191 candidate of the QTL for the number of vertebrae.

192 In summary, this work is a GWAS focusing on the number of vertebrae in pigs. A
193 genome-wide scan identified 21 significant SNPs within an 8.82-Mb region
194 containing the reported causal gene *NR6A1* on SSC1. An additional 18 SNPs were
195 identified within a 4.56-Mb region on SSC7 that showed genome-wide association
196 with the number of vertebrae. Finally, haplotype sharing analysis refined the 4.56-Mb
197 region to a region about 637 Kb in size, encompassing the gene *TGFB3* on SSC1.
198 Exploration of the gene via additional genetic and functional studies in mammals
199 revealed that *TGFB3* could be a strong candidate for the number of vertebrae in pigs.

200

201 **Conflict of Interests**

202 The authors have declared that no conflict of interest exists.

203

204 **ACKNOWLEDGMENTS**

205 This research was supported by the Agricultural Science and Technology Innovation
206 Program (ASTIP-IAS02), National Key Technology R&D Program of China
207 (No.2011BAD28B01), earmarked fund for Modern Agro-industry Technology
208 Research System, and Chinese Academy of Agricultural Sciences Foundation
209 (No.2014ZL006).

210 **LITERATURE CITED**

211 Amin, N., C. M. van Duijn, and Y. S. Aulchenko, 2007 A genomic background based
212 method for association analysis in related individuals. *PLoS One* 2: e1274.

213 Andersson, O., E. Reissmann, and C. F. Ibáñez, 2006 Growth differentiation factor 11
214 signals through the transforming growth factor-beta receptor ALK5 to regionalize the
215 anterior-posterior axis. *EMBO Rep.* 7: 831-837.

216 Arai, A., T. Mizoguchi, S. Harada, Y. Kobayashi, Y. Nakamichi *et al.*, 2012 Fos plays
217 an essential role in the upregulation of RANK expression in osteoclast precursors
218 within the bone microenvironment. *J. Cell Sci.* 125: 2910-2917.

219 Aulchenko, Y. S., D. J. de Koning, and C. Haley, 2007 Genome-wide rapid
220 association using mixed model and regression: a fast and simple method for
221 genome-wide pedigree-based quantitative trait loci association analysis. *Genetics* 177:
222 577-585.

223 Burgosa, C., P. Latorrea, J. Altarribab, J. A. Carrodeguasc, L. Varona *et al.*, 2014
224 Allelic frequencies of NR6A1 and VRTN, two genes that affect vertebrae number in
225 diverse pig breeds: A study of the effects of the VRTN insertion on phenotypic traits
226 of a DurocxLandrace-Large White cross. *Meat Sci.* 100C: 150-155.

227 Cheah, F. S., E. W. Jabs, and S. S. Chong, 2005 Genomic, cDNA, and embryonic
228 xpression analysis of zebrafish transforming growth factor beta 3 (*tgfbeta3*). *Dev.*
229 *Dyn.* 232: 1021-1030.

230 Christ, B., R. Huang, and J. Wilting, 2000 The development of the avian vertebral
231 column. *Anat. Embryol. (Berl)* 202: 179-194.

- 232 Edwards, D. B., C. W. Ernst, N. E. Raney, M.E. Doumit, M.D. Hoge *et al.*, 2008
233 Quantitative trait locus mapping in an F2 Duroc6Pietrain resource population: II.
234 carcass and meat quality trait. *J. Anim. Sci.* 86: 254–266.
- 235 Fan, Y., Y. Xing, Z. Zhang, H. Ai, Z. Ouyang *et al.* 2013 A further look at porcine
236 Chromosome 7 reveals *VRTN* variants associated with vertebral number in Chinese
237 and Western Pigs. *PLoS One* 8: e62534.
- 238 Fontanesi, L., A. Ribani, E. Scotti, V. J. Utzeri, N. Veličković *et al.* 2014
239 Differentiation of meat from European wild boars and domestic pigs using
240 polymorphisms in the *MC1R* and *NR6A1* genes. *Meat Sci.* 98: 781-784.
- 241 Hirose, K., S. Mikawa, N. Okumura, G. Noguchi, K. Fukawa *et al.* 2013 Association
242 of swine *vertnin* (*VRTN*) gene with production traits in Duroc pigs improved using a
243 closed nucleus breeding system. *Anim. Sci. J.* 84: 213-221.
- 244 James, A. W., Y. Xu, J. K. Lee, R. Wang, and M. T. Longaker, 2009 Differential
245 effects of TGF-beta1 and TGF-beta3 on chondrogenesis in posterofrontal cranial
246 suture-derived mesenchymal cells in vitro. *Plast. Reconstr. Surg.* 123: 31-43.
- 247 Kaartinen, V., J. W. Voncken, C. Shuler, D. Warburton, D. Bu *et al.* 1995 Abnormal
248 lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of
249 epithelialmesenchymal interaction. *Nat. Genet.* 11: 415-421.
- 250 King, J. W. B., and R. C. Roberts, 1960 Carcass length in the bacon pig: its
251 association with vertebrae numbers and prediction from radiographs of the young pig.
252 *Anim. Prod.* 2: 59-65.
- 253 Lorda-Diez, C. I., J. A. Montero, J. A. Garcia-Porrero, and J. M. Hurle, 2010 Tgfbeta2

254 and 3 are coexpressed with their extracellular regulator *Ltbpl* in the early limb bud
255 and modulate mesodermal outgrowth and BMP signaling in chicken embryos. *BMC*
256 *Dev. Biol.* 10: 69.

257 Madsen, P., P. Sørensen, G. Su, L. H. Damgaard, H. Thomsen *et al.* 2006 DMU- a
258 package for analyzing multivariate mixed models. In 8th World Congress on Genetics
259 Applied to Livestock Production, Brasil.

260 Mikawa, S., T. Hayashi, M. Nii, S. Shimanuki, T. Morozumi *et al.* 2005 Two
261 quantitative trait loci on *Sus scrofa* chromosomes 1 and 7 affecting the number of
262 vertebrae. *J. Anim. Sci.* 83: 2247-2254.

263 Mikawa, S., T. Morozumi, S. Shimanuki, T. Hayashi, H. Uenishi *et al.* 2007 Fine
264 mapping of a swine quantitative trait locus for number of vertebrae and analysis of an
265 orphan nuclear receptor, germ cell nuclear factor (NR6A1). *Genome Res.* 17:
266 586-593.

267 Mikawa, S., S. Sato, M. Nii, T. Morozumi, G. Yoshioka *et al.* 2011 Identification of a
268 second gene associated with variation in vertebral number in domestic pigs. *BMC*
269 *Genet.* 12: 5.

270 Miller, S. A., D. D. Dykes, and H. F. Polesky, 1988 A simple salting out procedure for
271 extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16: 1215.

272 Nezer, C., C. Collette, L. Moreau, B. Brouwers, J. J. Kim *et al.* 2003 Haplotype
273 sharing refines the location of an imprinted quantitative trait locus with major effect
274 on muscle mass to a 250-kb chromosome segment containing the porcine *IGF2* gene.
275 *Genetics* 165: 277-285.

- 276 Ren, D. R., J. Ren, G. F. Ruan, Y. M. Guo, L. H. Wu *et al.* 2012 Mapping and fine
277 mapping of quantitative trait loci for the number of vertebrae in a White Duroc ×
278 Chinese Erhualian intercross resource population. *Anim. Genet.* 43: 545–551.
- 279 Rubin, C. J., H. J. Megens, A. Martinez Barrio, K. Maqbool, S. Sayyab *et al.* 2012
280 Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. U. S.*
281 *A.* 109: 19529-19536.
- 282 Sato, S., Y. Oyamada, K. Atsuji, K. Atsuji, T. Nade, S. Sato *et al.* 2003 Quantitative
283 trait loci analysis for growth and carcass traits in a Meishan × Duroc F2 resource
284 population. *J. Anim. Sci.* 81: 2938-2949.
- 285 Wada, Y., T. Akita, T. Awata, T. Furukawa, N. Sugai *et al.* 2000 Quantitative trait loci
286 (QTL) analysis in a Meishan × Göttingen cross population. *Anim. Genet.* 31:
287 376-384.
- 288 Yang, G., J. Ren, Z. Zhang, and L. Huang, 2009 Genetic evidence for the
289 introgression of Western NR6A1 haplotype into Chinese Licha breed associated with
290 increased vertebral number. *Anim. Genet.* 40: 247-250.
- 291 Yang, Q., J. Cui, I. Chazaro, L. A. Cupples, and S. Demissie, 2005 Power and type I
292 error rate of false discovery rate approaches in genome-wide association studies.
293 *BMC Genet.* 6: S134.
- 294 Zhang, L. C., N Li., X. Liu, J. Liang, H. Yan *et al.* 2014 A genome-wide association
295 study of limb bone length using a Large White × Minzhu intercross population.
296 *Genet. Sel. Evol.* 46: 56.
- 297 Zhang, Z. G., B. D. Li, and X. H. Chen, 1986 Pig breeds in China. Shanghai Scientific

298 and Technical Publisher, Shang, China.

299 Zhao, S. F., M. Z. Chai, M. Wu, Y. H. He, T. Meng *et al.* 2014 Effect of vitamin B12

300 on cleft palate induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin and dexamethasone in

301 mice. *J. Zhejiang Univ. Sci. B* 15: 289-294.

302 **Table 1 Summary of number of vertebrae in Large White × Minzhu intercross**

303 **population**

| Total Number of Vertebrae | | | | | | Total | Average |
|---------------------------|----|-----|-----|----|----|-------|---------|
| 24 | 25 | 26 | 27 | 28 | 29 | | |
| Number of Individuals | | | | | | | |
| 1 | 9 | 244 | 296 | 42 | 1 | 593 | 26.63 |

304

305

306 **Table 2 Distribution of SNPs after quality control and the average distances**

307 **between SNPs on each chromosome¹**

| Chromosome | No. SNPs | Average distance(kb) |
|-------------------|-----------------|-----------------------------|
| 1 | 5323 | 59.16 |
| 2 | 2759 | 58.82 |
| 3 | 2228 | 64.69 |
| 4 | 2937 | 48.81 |
| 5 | 1971 | 56.43 |
| 6 | 2620 | 60.02 |
| 7 | 2779 | 48.45 |
| 8 | 2308 | 63.93 |
| 9 | 2668 | 57.52 |
| 10 | 1496 | 52.62 |
| 11 | 1574 | 55.67 |
| 12 | 1308 | 48.49 |
| 13 | 3384 | 64.57 |
| 14 | 3136 | 48.99 |
| 15 | 2361 | 66.68 |
| 16 | 1565 | 55.50 |
| 17 | 1381 | 50.09 |
| 18 | 1133 | 53.74 |
| X | 733 | 195.70 |
| 0 ² | 3951 | |
| Total | 47615 | |

308 ¹Data from *Sus scrofa* Build 10.2

309 ²These SNPs are not assigned to any chromosomes.

310

311 **Table 3 Genome-wide significant SNPs associated with number of vertebrae**

| SNP | Chromosome | Position ¹ | P-value | Nearest gene |
|-------------|------------|-----------------------|----------|-----------------|
| H3GA0004881 | 1 | 298972575 | 2.68E-10 | <i>PSMB7</i> |
| SIRI0001487 | 1 | 299578776 | 1.16E-08 | <i>SCAI</i> |
| INRA0007601 | 1 | 299599913 | 1.16E-08 | <i>SCAI</i> |
| ASGA0007558 | 1 | 297461876 | 6.93E-08 | <i>GPR21</i> |
| DRGA0002521 | 1 | 297585240 | 6.93E-08 | <i>STRPB</i> |
| H3GA0004851 | 1 | 298042326 | 7.19E-08 | <i>DENND1A</i> |
| M1GA0001554 | 1 | 297217239 | 1.24E-07 | <i>OR5C1</i> |
| ALGA0010388 | 1 | 298117818 | 1.46E-07 | <i>DENND1A</i> |
| ALGA0010455 | 1 | 300316974 | 1.47E-07 | <i>PBX3</i> |
| ASGA0007597 | 1 | 299033446 | 2.03E-07 | <i>GPR144</i> |
| ALGA0010073 | 1 | 294277597 | 2.27E-07 | <i>DAB2IP</i> |
| INRA0007489 | 1 | 294523343 | 2.59E-07 | <i>NDUFA8</i> |
| ALGA0010213 | 1 | 296003591 | 3.32E-07 | <i>OR1J1</i> |
| ASGA0007561 | 1 | 297725378 | 4.05E-07 | <i>STRBP</i> |
| INRA0007578 | 1 | 297186811 | 4.24E-07 | <i>OR5C1</i> |
| ASGA0007592 | 1 | 298890319 | 4.96E-07 | <i>NEK6</i> |
| H3GA0004837 | 1 | 297148442 | 6.31E-07 | <i>OR5C1</i> |
| DBNP0000338 | 1 | 302746429 | 6.49E-07 | <i>FATP4</i> |
| DRGA0002465 | 1 | 293928612 | 6.87E-07 | <i>DAB2IP</i> |
| ALGA0010606 | 1 | 302716687 | 7.44E-07 | <i>SWI5</i> |
| ALGA0010178 | 1 | 295559918 | 1.02E-06 | <i>OR1J1</i> |
| INRA0027600 | 7 | 102222702 | 2.28E-07 | <i>RBM25</i> |
| INRA0027605 | 7 | 102739058 | 4.8E-07 | <i>ACOT6</i> |
| ALGA0043941 | 7 | 103001352 | 4.38E-07 | <i>FAM161B</i> |
| ALGA0043942 | 7 | 103020388 | 8.56E-07 | <i>COQ6</i> |
| INRA0027623 | 7 | 103366820 | 2.11E-07 | <i>VSX2</i> |
| M1GA0010653 | 7 | 103637930 | 8.6E-07 | <i>LTBP2</i> |
| M1GA0010654 | 7 | 103796933 | 1.47E-07 | <i>FCF1</i> |
| ALGA0043962 | 7 | 103816521 | 1.47E-07 | <i>FCF1</i> |
| ASGA0035536 | 7 | 104108293 | 1.95E-07 | <i>ACYPI</i> |
| ASGA0035537 | 7 | 104219054 | 3.59E-07 | <i>TMED10</i> |
| ALGA0108658 | 7 | 104546250 | 5.6E-07 | <i>BATF</i> |
| DBNP0000926 | 7 | 104807152 | 1.66E-07 | <i>FLVCR2</i> |
| ASGA0035551 | 7 | 105022346 | 4.13E-07 | <i>TTLL5</i> |
| SIRI0001067 | 7 | 105182819 | 1.42E-07 | <i>TGFB3</i> |
| ALGA0044022 | 7 | 105383136 | 4.05E-07 | <i>GPATCH2L</i> |
| H3GA0022720 | 7 | 105752757 | 8.57E-07 | <i>ESRRB</i> |
| ALGA0044071 | 7 | 106028315 | 7.33E-07 | <i>LRRC74A</i> |
| MARC0034477 | 7 | 106779379 | 7.89E-07 | <i>ISM2</i> |

312 ¹SNP location adjusted on chromosomes in the *Sus scrofa* Build 10.2 assembly.

313 **Figure legends:**

314 **Figure 1 Manhattan plots of genome-wide association study with porcine**

315 **number of vertebrae**

316 Chromosomes 1-18, and X are shown in different colors. The horizontal line indicates
317 the genome-wide significance level ($-\log_{10}(1.05E-06)$).

318

319 **Figure 2 The marker-assisted segregation analysis for F1 boars.**

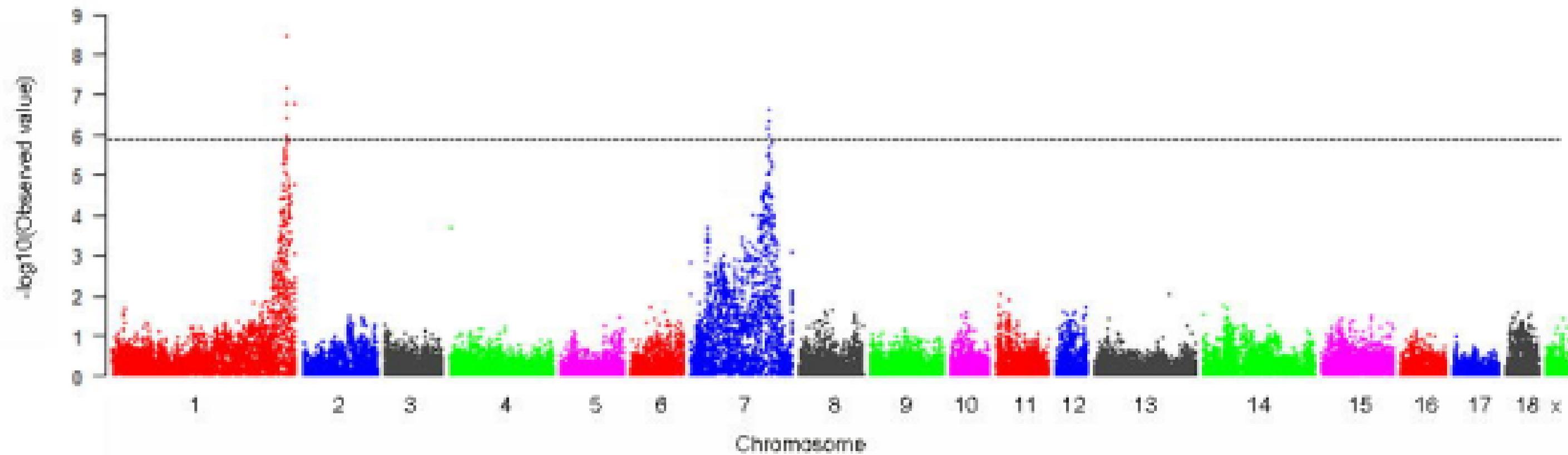
320 The marker-assisted segregation analysis for F1 boars. The graphs show, for 3 F1
321 boars' half-sib pedigrees (105, 4003, 6601), the phenotypic mean \pm standard errors of
322 the offspring sorted in two groups according to the homolog inherited from the sire.
323 The number of offspring in each group is given above the error bars, respectively. The
324 graph corresponds to the boars that were shown to be heterozygous Qq and reports a
325 Z-score for each pedigree. Q alleles associated with a positive allele substitution
326 effect on number of vertebrae are marked by a diamond, q alleles by a circle. The
327 number within the symbols differentiates the Q and q alleles according to the
328 associated marker genotype.

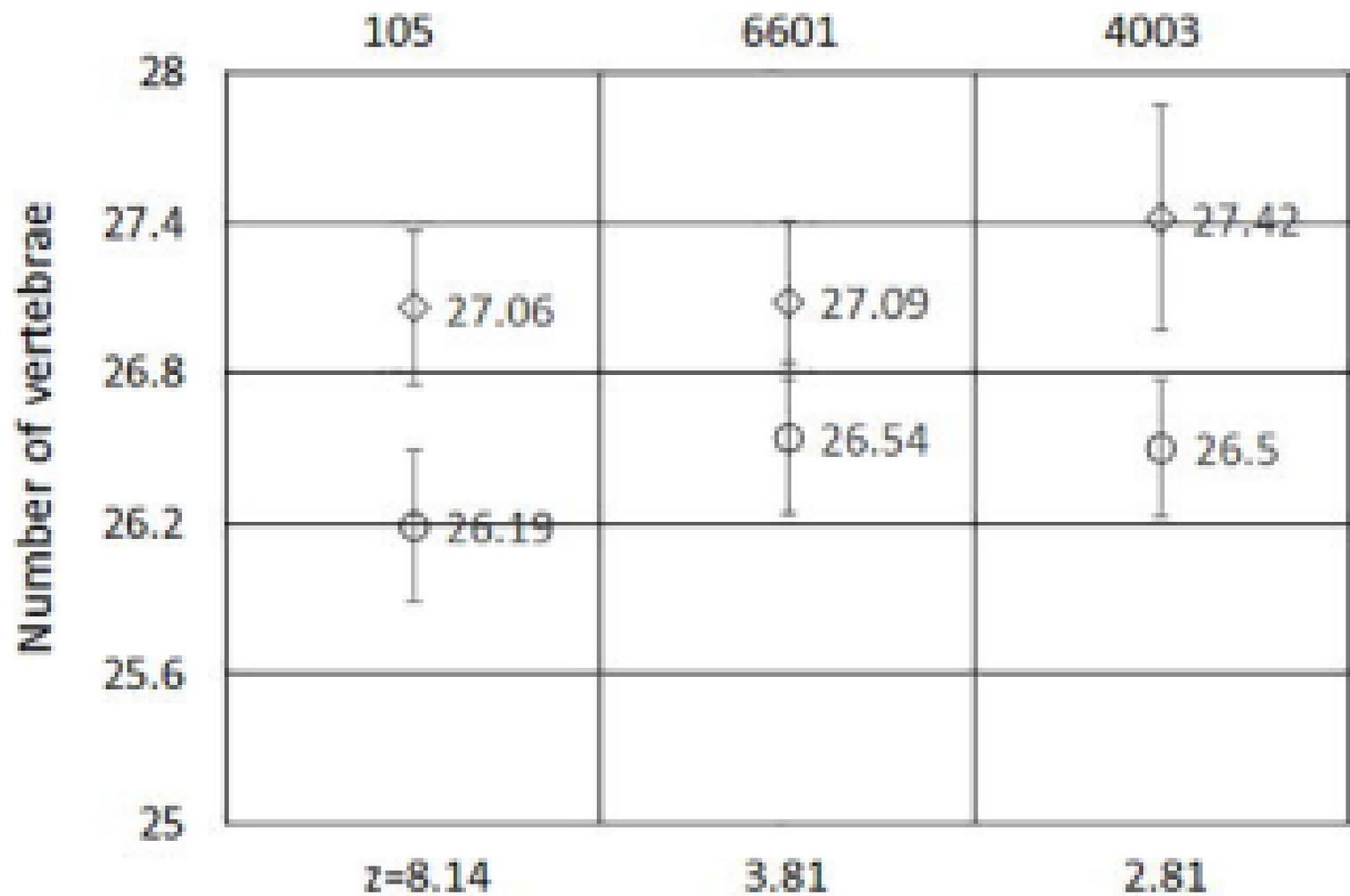
329

330 **Figure 3 Haplotype sharing analysis in the 4.56-Mb region on SSC7.**

331 Shared haplotypes of Qq F1 boars (Large White \times Minzhu intercross population) and
332 top 10% F2 individuals with Q chromosomes were analyzed. Polymorphisms are
333 displayed at the respective SNP markers. SNP alleles are shown by 1 and 2 for the

334 major and minor alleles, respectively. Identities of animals carrying the Q
335 chromosome are given in the left axis.





| ID | INRA0027600 | INRA0027605 | ALGA0043941 | ALGA0043942 | INRA0027623 | MI GA0010653 | MI GA0010654 | ALGA0043962 | ASGA0035536 | ASGA0035537 | ALGA0108658 | DBNP0000926 | ASGA0035551 | SIRI0001067 | ALGA0044022 | H3GA0022720 | ALGA0044071 | MARCO034477 | |
|-----------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|---|
| F1 105 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 6601 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 4003 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 105 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | q |
| 6601 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | q |
| 4003 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | q |
| F2 915904 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 915909 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 916106 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 917101 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 917104 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 918609 | 2 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 918707 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 930101 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 932702 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 932708 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | Q |
| 933104 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 933108 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1004403 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1008909 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1008911 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1012207 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1013411 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1013713 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1015902 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1015913 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1019602 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1019604 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1027204 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1030609 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1032006 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1035309 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1037311 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | Q |
| 1038407 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | Q |