CrossMark

Animal (2019), **13:4**, pp 683–693 © The Animal Consortium 2018
doi:10.1017/S1751731118001714

animal

A genome scan for selection signatures in Taihu pig breeds using next-generation sequencing

Z. Wang[1,2a], H. Sun[1,2a], Q. Chen[1,2], X. Zhang[1,2], Q. Wang[1,2] and Y. Pan[1,2†]
[1]*Department of Animal Science, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, PR China;* [2]*Shanghai Key Laboratory of Veterinary Biotechnology, Shanghai 200240, PR China*

(Received 20 September 2017; Accepted 25 May 2018; First published online 10 July 2018)

Taihu pig breeds are the most prolific breeds of swine in the world, and they also have superior economic traits, including high resistance to disease, superior meat quality, high resistance to crude feed and a docile temperament. The formation of these phenotypic characteristics is largely a result of long-term artificial or natural selection. Therefore, exploring selection signatures in the genomes of the Taihu pigs will help us to identify porcine genes related to productivity traits, disease and behaviour. In this study, we used both intra-population (Relative Extend Haplotype Homozygosity Test (REHH)) and inter-population (the Cross-Population Extend Haplotype Homozygosity Test (XPEHH); F-STATISTICS, F_{ST}*) methods to detect genomic regions that might be under selection process in Taihu pig breeds. As a result, we found 282 (REHH) and 112 (XPEHH) selection signature candidate regions corresponding to 159.78 Mb (6.15%) and 62.29 Mb (2.40%) genomic regions, respectively. Further investigations of the selection candidate regions revealed that many genes under these genomic regions were related to reproductive traits (such as the* TLR9 *gene), coat colour (such as the* KIT *gene) and fat metabolism (such as the* CPT1A *and* MAML3 *genes). Furthermore, gene enrichment analyses showed that genes under the selection candidate regions were significantly over-represented in pathways related to diseases, such as autoimmune thyroid and asthma diseases. In conclusion, several candidate genes potentially under positive selection were involved in characteristics of Taihu pig. These results will further allow us to better understand the mechanisms of selection in pig breeding.*

Keywords: indigenous pig, commercial pig, single nucleotide polymorphism, haplotype homozygosity test, gene enrichment analyses

Implications

Taihu pig breeds are renowned for having high-fertility characteristics, superior meat quality and high resistance to disease. This study examined a genome-wide map of selection signatures within Taihu pigs and between Taihu and Western commercial pigs. We identified some porcine genes related to productivity, disease and behaviour traits. This study will provide an insight to better understand the underlying mechanisms of selection in pig breeding. In addition, the difference between Taihu and Western commercial pigs may be useful to help develop a good hybrid through a crossing production system.

Introduction

Taihu pig breeds (Meishan, Fengjing, Shawutou, Erhualian, Jiaxing Black and Mizhu) are the most prolific breeds of swine in the world. They are distributed mainly in a narrow region around Taihu Lake in the lower Yangtze River Valley of China. In addition to their renowned reproduction traits and high proliferation, they also have many other outstanding economic traits, such as high resistance to disease, superior meat quality, and high resistance to crude feed and a docile temperament (China National Commission of Animal Genetic Resources, 2011). Current phenotypes of pig breeds have dramatically changed compared to their wild counterparts as a result of domestication and the subsequent selection of traits by humans or nature. During this process, different selection pressures (e.g. a wide variety of environmental conditions and intentional breeding by humans) has remarkably subjected the pigs to display enormous amount of phenotypic variations, and their genomes also has experienced changes that are consistent with their current phenotypes. For example, characteristics related to coat colour and muscle growth are favoured by breeders and are strongly selected for in certain breeds. Therefore, the identification of regions that have been subjected to selective breeding would be beneficial for identifying genes related to traits of interest and biological processes.

To date, research studies involving selection signature analysis have largely been conducted in animals. Several studies

[a] These two authors contributed equally to this work.
[†] E-mail: panyuchun1963@aliyun.com

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

have revealed strong signatures of selection affecting genomic regions that harbour genes underlying the economic traits of pigs, such as body size (e.g. *NR6A1*, *PLAG1* and *LCORL*, or *ADAMTSL3* genes) (Rubin *et al*., 2012), adaptation to high altitude (e.g. *ADAMTS12*, *SIM1* and *NOS1* genes) (Ai *et al*., 2013), white belt coat colour (e.g. the *EDNRB* gene) (Ai *et al*., 2013) and muscle growth (e.g. *BARX2* gene) (Wang *et al*., 2014). Nevertheless, the number of genes identified under selection is still small and does not provide insight into changes within the genomes between domestic pigs and wild pigs. In addition, the selection patterns in pig breeds are also different because of their distinct evolutionary histories. As mentioned above, Taihu pigs display superior economic traits of high proliferation, high resistance to disease, high resistance to crude feed and tameness, which are probably as a result of long-term artificial or natural selection. Therefore, a scan of genome-wide selection signatures in the genomes of Taihu pigs will help us identify porcine genes related to productivity, disease and behaviour traits. This will also provide an insight to better understand the underlying mechanisms of selection in pig breeding.

In this study, we detected the selection signatures of Taihu pigs at the genome level using the sequencing data of 252 Taihu pigs and 120 Western pigs. Both intra-population (Relative Extend Haplotype Homozygosity Test (REHH) (Sabeti *et al*., 2002)) and inter-population (the Cross Population Extend Haplotype Homozygosity Test (XPEHH) (Sabeti *et al*., 2007); F-STATISTICS, F_{ST} (Weir and Cockerham, 1984)) statistical methods were used to identify the selection signatures.

Material and methods

Populations and sequencing data preparation

A total of 372 individuals from nine pig breeds, including six Chinese Taihu pig breeds (Meishan, Fengjing, Shawutou, Erhualian, Jiaxing Black and Mizhu; $n = 252$) and three Western pig breeds (Duroc, $n = 48$; Yorkshire, $n = 35$; Landrace, $n = 37$) were obtained for this study (Supplementary Table S1). The detailed information about all the individuals and the methods of acquiring the raw Illumina DNA sequence data have been reported in our previous studies (Wang *et al*., 2016; Zhang *et al*., 2016). Briefly, the genomic DNA was extracted from ear tissues. The DNA was digested by the restriction enzyme A*va*II. Then, the unique adapter-barcode sequence was ligated to each individual's digested fragments. Finally, the ligations of each individual were pooled to make up a DNA library. The library (digested fragments ranging from 200 to 300 bp) was sequenced. We aligned the reads to the pig reference genome (SGSC Sscrofa10.2; http://hgdownload.soe.ucsc.edu/goldenPath/susScr3/bigZips/) using Burrows-Wheeler Aligner (BWA ver 0.7.5) (Li and Durbin, 2009) with the default settings and using the steps of the genotyping by genome reducing and sequencing approach (Chen *et al*., 2013). The alignment results with mapping quality scores less than 20 were filtered before the single nucleotide polymorphism (SNP) was called.

Single nucleotide polymorphism calling and genotyping

Single nucleotide polymorphisms were identified and genotyped using SAMtools (ver 0.1.19) (Li *et al*., 2009) and the mpileup algorithm. These SNPs were retained for further analysis according to the following criteria: more than 111 (30%) genotyped samples and a sequencing depth greater than fivefold on average. The imputation analysis for the missing genotypes was performed with an iBLUP algorithm (Yang *et al*., 2014). Finally, a total of 114 596 SNPs (with an average depth for SNP calling of 10.9-fold; Supplementary Table S1) with minor allele frequencies of 0.05 were obtained from the selection signature analysis. In addition, the FASTPHASE program (Scheet and Stephens, 2006) was implemented to reconstruct haplotypes for every chromosome using the default parameters to acquire full-phase haplotype data.

Detection of selection signatures

Three methods, including the REHH, XPEHH and F_{ST} tests, were performed to detect the selection signatures. The REHH test can detect positive selection signatures that occurs within a population's genome, whereas the XPEHH and F_{ST} tests can detect selection signatures by comparing the genomes of two populations.

The program Sweep v.1.1 was implemented to compute the REHH statistic, and the core regions were defined as chromosome fragments that included at least four SNPs (Sabeti *et al*., 2002). We then placed the REHH values in order, based on the frequency of all the haplotypes, into 20 bins to calculate their significance and obtained the *P* values by log-transforming the REHH in the bin to reach normality and calculate the mean and standard deviation. The core haplotypes with extreme REHH values ($P < 0.01$) were considered to be statistically significant. We removed the core haplotypes with a frequency of less than 0.25 because regions harbouring the beneficial allele would have a high frequency due to a hitchhiking effect (Sabeti *et al*., 2002). Finally, we extended 250 kb towards both the upstream and downstream directions of the significant core haplotypes with lengths longer than 1 kb, and we regarded these extended regions as candidates for selection regions.

The program available at http://hgdp.uchicago.edu (coded by Joe Pickrell) was employed to perform the *XPEHH test* (Pickrell *et al*., 2009). In our study, we defined Taihu pigs as the observed population and Western pigs as the reference population. The positive XPEHH scores indicate that a selection occurred in the observed population, whereas negative scores suggest that a selection occurred in the reference population. As the calculation of the XPEHH requires genetic distance between adjacent SNPs, we considered a chromosome segment of 1 Mb to be 1 cM. After normalization, the XPEHH values approximately followed a standard norm distribution (Sabeti *et al*., 2007). We considered these outlying signals with a *P* value less than 0.01 to be statistically significant. We first merged the outlier signals with a physical distance less than 250 kb into one region and extended 250 kb towards the upstream and downstream directions of merged regions that were considered potential

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

candidates of the selection regions. The F_{ST} analysis described by Weir and Cockerham (1984) was also performed between Taihu pigs and Western pigs. The top 1% SNPs were considered as the candidate selection regions. Moreover, the difference of selection signatures among Taihu pigs and Western pigs were also performed, respectively.

Annotation and function analysis of identified genomic regions

The candidate selection regions were annotated to the closest genes based on the annotation for the *Susscrofa* assembly 10.2 available from the Ensembl databases (http://www.ensembl.org/index.html). The genes were defined as candidate genes when their positions were contained inside the candidate regions of selection. The functions and pathways in which these genes are involved were determined using the DAVID tool (ver 6.7) (Dennis *et al.*, 2003). Because the Gene Ontology (GO) annotation of pigs is incomplete, corresponding human orthologous Ensembl IDs were retrieved to determine enrichment function categories using a perl script. The enriched GO terms and pathways with P values < 0.05 after correction for multiple testing (with a false discovery rate of less than 25%) were considered to be statistically significant. In addition, we identified enriched networks in the Ingenuity Pathways Analysis (IPA, Ingenuity Systems and www.ingenuity.com) platform. Finally, the porcine quantitative trait locus (QTL) database (http://www.animalgenome.org) was searched to identify overlap with previously published porcine QTLs within the candidate regions. We removed QTLs with lengths more than 1 Mb to consider some QTLs with lengths that were too large for efficient post-processing, and we also merged the QTLs with overlapped QTLs. Ultimately, a total of 2031 non-overlapping QTL regions were acquired in the analysis (Supplementary Table S2).

Results

Sequencing and identification of single nucleotide polymorphisms

Approximately 1.47 billion reads of high quality (with approximately 3.95 million reads and a coverage of 2.7% on average for each individual) were used to call SNPs. The information of the average amount of coverage and the read number for each breed are shown in Supplementary Table S1. A density of 100 kb of a non-overlapping window for these 114 596 SNPs exhibited a relatively uniform distribution across each chromosome, except for the sex chromosomes (Supplementary Figure S1). The results of the population structure based on the principal component analysis, STRUCTURE and Neighbor-joining tree showed that Taihu pigs and Western pigs can be clearly separated into two different populations fostering comparison between the two populations (Supplementary Figure S2). Therefore, we conducted selection analysis by comparing the genomes of the two populations. However, there were also some differences among Taihu pig breeds and Western pig breeds. Therefore, the difference of selection signatures among Chinese native breeds and Western pig breeds were explored. Among the Taihu pig breeds, the sub-population number was set as 7 in the F_{ST} analysis. There were no obvious difference between Yorkshire and Landrace pigs even when the K is 10 in the STRUCTURE analysis. We then considered the Yorkshire and Landrace pigs as a whole to compare with the Duroc pigs.

Candidate selection regions

To detect signals of positive selection within the Taihu pig population as well as against the Western population, we used three statistical methods to comprehensively localize the source of selection. We first used the REHH test to detect the selected signals in the genomes of the Taihu and Western pigs. A total of 282 and 385 candidate selection regions with an overall length of 159.78 and 207.10 Mb, and average lengths of 0.57 and 0.58 Mb, accounted for approximately 6.15% and 7.96% of the genome, and they were identified in Taihu and Western pig populations, respectively (Table 1). Chromosome 14 in both Taihu and Western pigs showed the longest length of the selection regions compared to the rest of the chromosomes.

We then used the XPEHH and F_{ST} tests to compare Taihu and Western pig populations. A total of 666 and 896 outliers ($P < 0.01$) were detected in Taihu and Western pig populations in the XPEHH test. After merging the overlap outlier as described in the methods section, a total of 112 and 170 candidate selection regions with total lengths of 62.29 and 94.20 Mb, and average lengths of 0.56 and 0.55 Mb that accounted for approximately 2.40% and 3.63% of the genome, were identified in Taihu and Western pig populations, respectively (Table 1). In the genome of the Taihu pigs, chromosome 2 displayed the longest length of selection regions, and chromosome 18 was not identified in genomic regions under selection. Many neighbour loci showed consecutive lower P values (Figures 1 and 2), which indicated that more reliable evidence of those regions experienced selection. In the F_{ST} analysis, the top 1% SNPs were considered as the candidate selection regions. The average F_{ST} value between Taihu pigs and Western pigs was 0.198. Moreover, the average F_{ST} value among Taihu pigs and Western pigs were 0.123 and 0.101, respectively.

Genomic annotation

To further investigate the possible biological functions of these regions under selection, we revealed the QTLs and genes harboured in the selection regions. We found that selection regions of approximately 25% overlapped with QTLs (Table 2). We noted that the vast majority of these genes identified under selection in the two populations were different, and even the genes identified under selection in the same population with different methods only displayed a small proportion of overlap (Taihu, 102 genes; Western, 124 genes; Supplementary Figure S3). Furthermore, the results of the enrichment analysis showed that the two breeds presented a different over-representation of genes with GO

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

Table 1 *Summary of whole genome candidate selection regions identified in Taihu and Western pig populations (Mb)*

	Taihu				Western			
	Intra-population		Inter-population		Intra-population		Inter-population	
Chromosome	Number	Length (average)	Number	Length (average)	Number	Length (average)	Number	Length (average)
1	25	14.87 (0.59)	4	2.05 (0.51)	18	10.84 (0.60)	15	8.01 (0.53)
2	19	10.28 (0.54)	15	8.15 (0.54)	29	17.33 (0.60)	9	4.76 (0.53)
3	22	12.95 (0.59)	7	4.25 (0.61)	24	14.29 (0.60)	12	6.16 (0.51)
4	20	11.03 (0.55)	3	1.53 (0.51)	25	15.66 (0.63)	11	6.14 (0.56)
5	10	5.26 (0.53)	6	3.24 (0.54)	16	8.70 (0.54)	7	3.76 (0.54)
6	17	9.36 (0.55)	13	7.06 (0.54)	35	19.72 (0.56)	12	6.60 (0.55)
7	13	7.64 (0.59)	3	2.03 (0.68)	13	7.29 (0.56)	18	9.45 (0.52)
8	16	10.49 (0.66)	7	4.50 (0.64)	13	7.61 (0.59)	3	1.75 (0.58)
9	13	6.97 (0.54)	3	1.60 (0.53)	20	12.57 (0.63)	10	6.05 (0.60)
10	6	3.30 (0.55)	3	1.54 (0.51)	14	7.78 (0.56)	1	0.50 (0.50)
11	9	5.24 (0.58)	2	1.00 (0.50)	3	1.77 (0.59)	6	3.30 (0.55)
12	19	10.39 (0.55)	7	3.99 (0.57)	23	13.09 (0.57)	2	1.02 (0.51)
13	16	9.12 (0.57)	11	6.44 (0.59)	11	7.13 (0.65)	11	6.05 (0.55)
14	30	16.55 (0.55)	13	6.73 (0.52)	56	32.67 (0.58)	18	10.36 (0.58)
15	19	10.43 (0.55)	3	1.67 (0.56)	21	11.73 (0.56)	5	3.01 (0.60)
16	5	2.54 (0.51)	4	2.05 (0.51)	6	3.25 (0.54)	9	5.19 (0.58)
17	17	9.70 (0.57)	7	3.95 (0.56)	12	6.48 (0.54)	9	5.39 (0.60)
18	3	1.59 (0.53)	0	0.00 (0.00)	11	6.51 (0.59)	8	4.64 (0.58)
X	3	2.07 (0.69)	1	0.50 (0.50)	5	2.67 (0.53)	4	2.05 (0.51)
Total	282	159.78 (0.57)	112	62.29 (0.56)	355	207.10 (0.58)	170	94.20 (0.55)

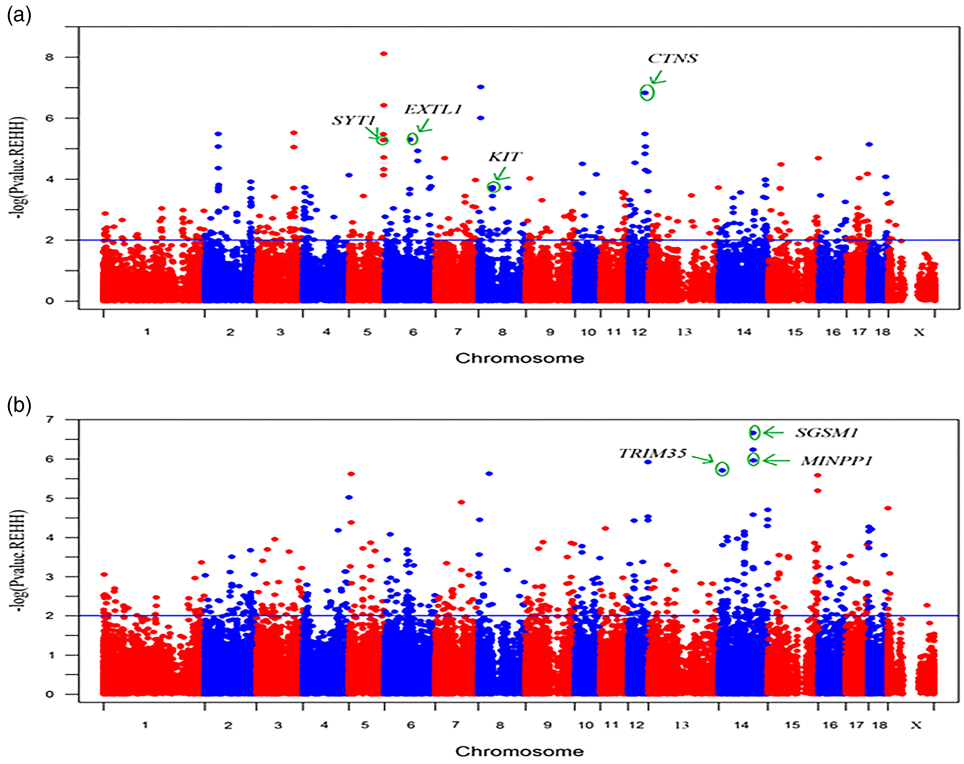


Figure 1 (colour online) The genome-wide distribution of selection signatures detected by the Relative Extend Haplotype Homozygosity Test (REHH) across each chromosome in Taihu and Western pigs. (a) Results of the REHH test for Taihu pigs. *SYT1* = Synaptotagmin 1; *EXTL1* = exostosin-like glycosyltransferase 1; *KIT* = KIT proto-oncogene receptor tyrosine kinase; *CTNS* = cystinosin, lysosomal cystine transporter. (b) Results of the REHH test for Western pigs. *TRIM35* = tripartite motif containing 35, *SGSM1* = small G protein signalling modulator 1; *MINPP1* = multiple inositol-polyphosphate phosphatase 1. The green circle represents the most significant genes under the selection.

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

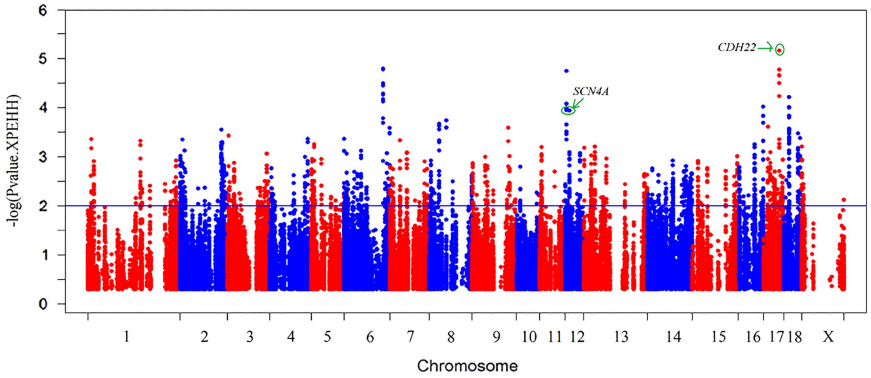


Figure 2 (colour online) The genome-wide distribution of selection signatures detected by the Cross-Population Extend Haplotype Homozygosity Test (XPEHH) across each chromosome in Taihu and Western pigs. *SCN4* = sodium voltage-gated channel alpha subunit 4; *CDH22* = cadherin 22. The Taihu pigs are defined as the observed population, and the Western pigs are the reference population. The green circle represents the most significant genes under selection.

Table 2 *The number of candidate selection regions under quantitative trait loci in Taihu and Western pig populations*

Population (method)	Counts[1] (%)	Meat and carcass	Health	Production	Reproduction	Exterior
Taihu (intra-population)	78 (27.7)	41	7	22	27	8
Western (intra-population)	69 (19.4)	41	11	16	21	9
Taihu (inter-population)	28 (25.0)	12	3	9	11	9
Western (inter-population)	48 (28.2)	32	3	14	12	9

[1]Some selection regions may be related to more than one trait, but they were only counted once.

annotations and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The Taihu pigs showed over-representation of genes with pathways related to diseases, whereas the Western pigs displayed an over-representation of genes with pathways related to lipids and olfaction.

Quantitative trait loci overlap with selection regions

The Pig QTLdb database includes almost all of the QTLs reported in the past few decades, and these data can provide information about genome regions that have been experimentally identified to be correlated with a specific trait. Therefore, we can explore the potential function of genomic regions under selection with traits by comparing their physical locations with QTLs.

The QTLs located in or overlapping with selection regions (REHH test) in each chromosome with the lowest *P* value are shown in Supplementary Table S3. The results show that a series of these QTLs is related to reproduction traits in Taihu pigs. Take the 268 556 705 to 268 561 835 bp selection signal of chromosome 1 for instance: one QTL (ID = 18061) influencing the total number of piglets born alive was mapped in this region. Many of the QTLs in Western pigs are related to meat and carcass traits, for example, the 25 248 186 to 25 252 699 bp selection signal of chromosome 16 overlapping with the QTL (ID = 22 311 to 22 314), which influenced the loin muscle area. In addition, we also found that some selection regions overlapped with QTLs that influence multiple traits, for example, the 7 116 095 to 7 124 680 bp selection signal of chromosome 12 in Taihu pigs was related to reproduction, production, health and exterior traits. Similarly, the 34 286 910 to 34 301 822 bp selection signals in Western pigs overlapped with QTLs related to meat and quality as well as reproduction and production traits.

Some of the QTLs located in or overlapping with the selection regions identified by the XPEHH test are summarized in Supplementary Table S4. In Taihu pigs, we found the 2 728 371 to 2 728 382 bp selection signal of chromosome 2 overlapped with many QTLs related to multiple traits (e.g. the total number of piglets born alive and backfat thickness), and these results also overlapped with the carnitine palmitoyltransferase 1A (*CPT1A*) gene. The *CPT1A* gene, which plays an important role in the regulation of fatty acid metabolism in newborn pigs (Lin *et al.*, 2010), has also been previously reported to be under selection (Ma *et al.*, 2015). Compared to the Taihu pigs, Western pigs experienced more selection in meat and carcass traits (e.g. average backfat thickness, ID = 22 292; drip loss, ID = 8053).

Genes under potential selection regions

A total of 1219 (REHH test) and 742 genes (XPEHH test) in Taihu pigs and 1609 (REHH test) and 1061 (XPEHH test) genes in Western pigs were found under the selection regions (Supplementary Figure S3). For the F_{ST} analysis, a total of 202 genes were detected between Taihu pigs and Western pigs (Supplementary Table S5). In addition, a total of 199 and 188 genes were detected among the Taihu pigs and Western pigs, respectively.

For the REHH test, we found that three genes in the Taihu population (*CTNS*, *EXTL1* and *SYT1* genes; Figure 1a) and

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

three different genes in the Western population (*SGMS1*, *MINPP1* and *TRIM35* genes; Figure 1b) are under selection regions showing the smallest *P* values (the top 5) (Supplementary Table S6). For the XPEHH test, we found three out of five outliers with the smallest *P* values located in two genes (*SCN4A* and *CDH22* genes; Figure 2). The *SCN4A* gene was identified under selection in the Taihu population, and the *CDH22* gene was identified under selection in the Western population. For the F_{ST} test between Taihu and Western pigs, the most significant gene in the chromosome 7 was *LTBP2* (F_{ST} value 0.967). Among the Taihu pigs analysis, the most significant gene was *SLC27A1* (F_{ST} value 0.866). Among the Western pigs analysis, the most significant genes were *GNAL* (F_{ST} value 0.966) and *CHST10* (F_{ST} value 0.948).

Involved biological processes under selection

To see the functions of the putative genes under selection in the entire gene network, we performed an enrichment analysis for these genes using DAVID. According to the criterion mentioned above (see methods), we found seven and six significant pathways in total for the REHH test as well as 5 and 16 significant GOs in the Taihu and Western populations, respectively. For the XPEHH test, the Taihu pigs had only one significant pathway and seven significant GOs, and the Western pigs had two significant pathways and nine significant GOs. The significant GO and KEGG pathways of over-represented genes are shown in Supplementary Table S6.

The results suggested that the two populations present a different pattern of over-representation of genes in the GO annotations and KEGG pathways. For Taihu pigs (REHH test), the majority of the pathways were associated with disease, for example, asthma, autoimmune thyroid disease, the intestinal immune network for IgA production and graft-*v.*-host disease. The GO terms were associated with embryonic development, ossification and enzyme activity. For the Western pigs, the significant pathways were associated with reproduction (e.g. the GnRH signalling pathway) and lipid metabolism (e.g. ether lipid metabolism and alpha-Linolenic acid metabolism), and the significant GO terms were associated with development (e.g. embryonic development and regulation of developmental processes).

By comparing the genome of Western pigs, the genes under selection regions in Taihu pigs were over-represented in the GO terms related to reproduction (e.g. female pregnancy, endomembrane system) and health (e.g. stress response). In contrast, the genes in selection regions in Western pigs were over-represented in the pathway related to olfaction (e.g. the Olfactory transduction pathway), and the GO terms are related to oxygen binding and temperature homeostasis when compared to the genome of Taihu pigs. In addition, for the F_{ST} test between Taihu and Western pigs, the significant term (GO:0019838 ~ growth factor binding, *P* value was 3.17E-05) was associated with growth was detected.

Molecular subnetwork

To assess the key roles of genes under selection regions in networks, we performed Fisher's exact test based on the IPA molecular library, and we found a total of 49 (Taihu: 25 REHH; 6 XPEHH; Western: 10 REHH, 8 XPEHH) significant molecular networks (Supplementary Table S7). We also found that genes under selection regions of the Taihu pig genome were enriched in functions related to productivity and immunity, for example, one of the networks (ID = 5, REHH; Figure 3) consists of 26 selected genes, and its most essential functions comprise cellular development, cellular growth and proliferation and embryonic development. Another network (ID = 2, XPEHH) includes 17 selected genes, and the functions of cellular movement, haematological system development and function as well as immune cell trafficking were enriched in this structure. The results of this network show that the genes under selection regions of the Western pig genome were enriched in functions related to development and the nervous system, for example, one of the networks (ID = 5, REHH) comprised 30 selected genes principally related to embryonic development, organismal development, skeletal and muscular system development and function, and the main functions of another network (ID = 3, XPEHH; Figure 4), with 16 selection genes participating in the system, involve behaviour, nervous system development and function, and skeletal and muscular system development and function.

Discussion

Taihu pig breeds are an invaluable component of the world's porcine genetic resources. However, the genetic architecture of their economic important traits is still unclear and this might impede the full utilization of their genetic potential. Exploring the selection signatures in the genome of Taihu

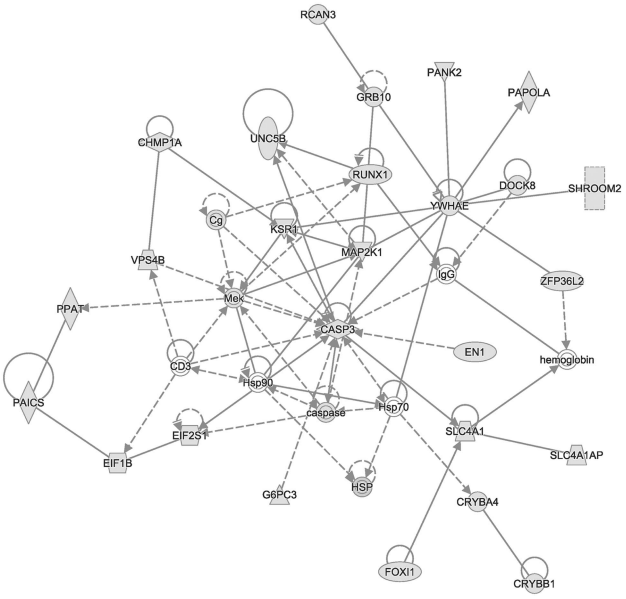


Figure 3 Significant ingenuity pathway analysis molecular networks under selection regions in Taihu pigs. Networks involving cellular development, cellular growth and proliferation as well as embryonic development in Taihu pigs.

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

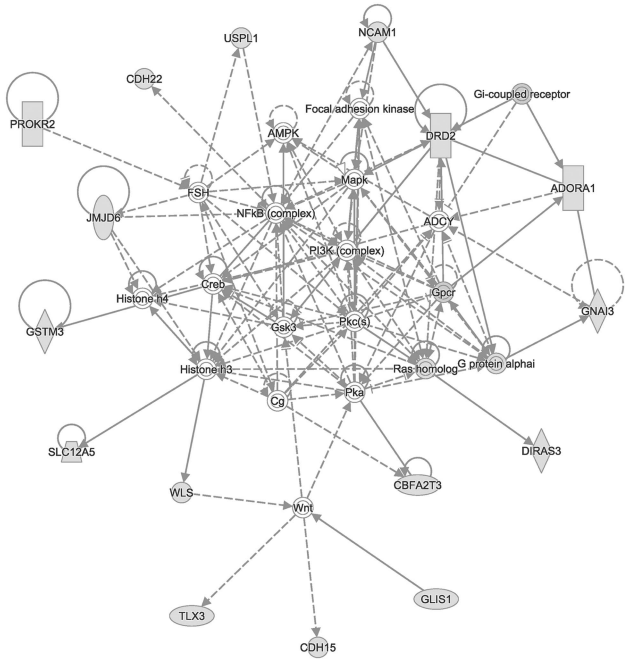


Figure 4 Significant ingenuity pathway analysis molecular networks under selection regions in Western pigs. Networks involving behaviour, nervous system development and function, and skeletal and muscular system development and function in Western pigs.

pigs will help us to identify porcine genes related to productivity, disease and behaviour traits as well as allow us to better understand the underlying mechanisms of selection in pig breeding. In this study, we used three complementary methods (REHH, XPEHH, F_{ST} tests) to detect the selection signatures. The three approaches are necessary and helpful to construct a comprehensive selection map for the pig genome and to reveal genetic mechanisms of phenotypic diversity.

We noted that Taihu pigs exhibited less overall length of selection regions compared to the Western pigs, implying that Taihu pigs experienced less intensive selection unlike the Western pigs. This observation agrees with the history of the breeding programmes of the Western pig breed, which experienced a more intensive selection in the past few decades than indigenous Chinese pigs. Furthermore, the majority of genes under selection regions in Taihu pigs were different from selection genes in Western pigs, indicating that the Taihu and Western pigs experienced selection in different regions of the genome. Approximately 10 000 years ago, pigs from Asia and Europe were independently domesticated (Larson *et al*., 2007). During this period, artificial selection and breeding has led to highly phenotypic changes among breeds on these continents (Megens *et al*., 2008). The different selection pressures on their phenotypes eventually led to different selection in their genomes.

In addition, for each population, the selection regions identified by the two approaches (REHH and XPEHH) also showed a low fraction of overlap (Supplementary Figure S3). This outcome is presumably due to the methods being sensitive to different categories of selection signatures (Oleksyk *et al*., 2010; Suzuki, 2010), for example, the REHH test is sensitive for detecting ongoing selection signatures with low-to-moderate frequency, whereas the XPEHH test is effective for detecting nearly complete selection signatures (Sabeti *et al*., 2007).

The functions associated with the putative selection regions were investigated at three levels (i.e. trait level, gene level and gene set level) based on several genomic annotation databases, for example, the QTL database and KEGG. For the traits level, we first explored the pig QTL database to test whether the detected selection signatures were as a result of human selection during pig breeding. Interestingly, we found that about 25% of the selection regions were located in the newly defined QTLs, which is consistent with previous studies showing that the most important economic traits of pigs have been subjected to strong selection pressure (Chen *et al*., 2007). Since Taihu pigs are especially famous for their reproduction traits, we expect to be able to identify selection signatures related to reproduction traits. Our findings supported this expectation with a series of selection regions located in QTLs of each chromosome (with the smallest *P* value) that were related to reproduction traits (Supplementary Table S3). Similarly, many of the QTLs overlapping with the selection regions in Western pigs were related to meat and carcass traits (Supplementary Table S3), and this observation correlates with the history of meat quality selection (such as lean meat percentage) in Western pig breeding programmes (Vidal *et al*., 2005; Amaral *et al*., 2011). Furthermore, based on the selection regions identified by the XPEHH test, we found that Taihu pigs had a greater percentage (Taihu: 39.2%, 11/28; Western: 25%, 12/48; Table 2) of selection regions overlapping with the reproduction QTLs compared to Western pigs, whereas Western pigs had a greater percentage (Taihu: 42.8%, 12/28; Western: 66.7%, 32/48; Table 2) of selection regions overlapping with the meat quality and carcass QTLs compared to Taihu pigs. To a certain extent, the results indicate that the Taihu pigs experienced more selection in production traits compared to the Western pigs, whereas the Western pigs experienced more selection in meat quality and carcass traits compared to the Taihu pigs.

For the gene level, the genes under the most significant selection regions in Taihu pigs were relevant to disease and the nervous system (Figure 1a), for example, the *CTNS* gene has been reported to be related to argyrophilic grain disease, which is a progressive neurodegenerative disease of the human brain (Villela *et al*., 2013). The *EXTL1* gene is a member of the *EXT* tumour suppressor gene family (Wuyts *et al*., 1999), and the *SCN4A* gene has been described as being associated with neuromuscular disorders such as paramyotonia congenita, sodium channel myotonias and congenital myasthenic syndrome in human studies (Vicart *et al*., 2005). The *SYT1* gene plays an important role in neurons and is associated with feeding behaviour traits (Do *et al*., 2013). Furthermore, the *SGMS1* gene, which was under the most significant selection regions in Western pigs was relevant to production traits, and it was reported to be

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

associated with the residual feed intake (RFI) of pigs. This observation is consistent with the history of Western pig breeding in which the RFI is used as a trait for selection to improve feed efficiency (Onteru *et al*., 2013). The outcome indicated that genes related to feed intake may have experienced certain selection content.

We also found several other genes (*MINPP1*, *TRIM35* and *CDH22* genes) under the most significant selection regions in Western pigs. Among them, the *MINPP1* gene was reported to affect milk fatty acid traits in dairy cattle (Li *et al*., 2014), and the *CDH22* gene was reported to likely play a significant role during the development and the adult functions of the mouse brain (Mayer *et al*., 2010). In addition, we performed a χ^2 test of the genotypes and allele frequency distribution of these loci under the most significant selection regions for the two populations. We speculated that the two populations would exhibit a different pattern of distribution in genotypes or alleles due to different selection pressures (i.e. the selection pressures of Taihu pigs are mainly from natural selection and conventional breeding, whereas for Western pigs, the pressures largely arose from modern artificial selection). The results showed that all these loci displayed different distribution patterns that are extremely significant in the two populations (Supplementary Table S8), implying that the different selection pressures are likely to result in different frequency distributions of genotypes or alleles.

In addition, we screened the list of genes under the selection regions and found that many genes were related to economic traits (Table 3). For the Taihu pigs, we found that some genes under selection were related to lipogenesis and reproduction traits. For example, the *CPT1A* and *MAML3* genes, which were reported to be relevant to lipogenesis in pigs (Lin *et al*., 2010; Revilla *et al*., 2014), overlapped with the selected region 2 728 371 to 2 728 382 bp on chromosome 2 and 92 696 500 to 92 790 169 bp on chromosome 8, respectively. Another example is the *TLR9* gene, which overlapped with the selected region 37 851 476 to 38 200 552 bp on chromosome 13 and was reported to be related to porcine reproductive traits involved in antiviral responses. Furthermore, the *KIT* gene (a well-known coat colour gene that is an important factor regulating coat colour) was found to be under the selection region of 43 339 560 to 43 408 258 bp on chromosome 8 (Figure 1, Table 3), and it has also been reported as a selection gene in several previous studies (Rubin *et al*., 2012; Bosse *et al*., 2015; Ma *et al*., 2015). For the Western pigs, we found some genes related to growth and meat quality traits, for example, the *IRX4* gene overlapping with selection region 85 397 504 to 85 537 619 bp on chromosome 16 was reported to be related to BW at 71 days of age (Jung *et al*., 2014), and the *GNAS* gene overlapping with the selection regions identified by both tests (REHH, 66 507 415 to 66 555 775 bp; XPEHH, 66 335 420 to 66 335 758 bp; on chromosome 17) was reported to affect growth in pigs (Oczkowicz *et al*., 2013). Furthermore, we also found that some genes were related to immunity, for example, the *BCL10* gene was identified under selection regions in both Taihu (143 418 209 to

Table 3 *Representative candidate genes located in potential selection regions in Taihu and Western pig populations*

Chromosomes	Location[1]	*P* value (method)[2]	Population	Candidate gene	Gene function
2	2 728 371 to 2 728 382	-(XPEHH)	T	*CPT1A* (carnitine palmitoyltransferase 1A)	Related to fatty acid metabolism in new born pig
4	143 418 209 to 143 454 176	0.0000736 (REHH)	T	*BCL10* (B-cell CLL/lymphoma 10)	Immunity
	143 195 558 to 143 454 176	0.00000954 (REHH)	W	*BCL10*	
5	106 969 622 to 106 972 821	0.00000529 (REHH)	T	*SYT1* (Synaptotagmin 1)	Related with neuronal functions
6	77 090 305 to 77 127 881	0.00000508 (REHH)	T	*STMN1* (stathmin 1)	Related to muscle development
	77 061 736 to 77 090 307	-(XPEHH)	T	*STMN1*	
8	92 696 500 to 92 790 169	0.000191 (REHH)	T	*MAML3* (mastermind-like transcriptional coactivator 3)	Related to lipogenesis
8	43 339 560 to 43 408 258	0.00021 (REHH)	T	*KIT* (KIT proto-oncogene receptor tyrosine kinase)	Related to coat colour
13	37 851 476 to 38 200 552	-(XPEHH)	T	*TLR9* (toll like receptor 9)	Related to antiviral responses and reproductive
16	85 397 504 to 85 537 619	-(XPEHH)	W	*IRX4* (Iroquois homeobox 4)	Related to growth
17	66 507 415 to 66 555 775	0.000153 (REHH)	W	*GNAS* (GNAS complex locus)	Related to growth
	66 335 420 to 66 335 758	-(XPEHH)	W	*GNAS*	

[1]The position of the start and end of a core region (Relative Extend Haplotype Homozygosity Test (REHH)) or outliers (the Cross-Population Extend Haplotype Homozygosity Test (XPEHH)), and extending 250 kb towards the upstream and downstream directions of one outlier would be a potential selection region.
[2]'-' represents a selected region including more than one outlier, and the *P* value is not shown.
T = Taihu population; W = Western population.

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

143 454 176 bp, chromosome 4, REHH test) and Western (143 195 558 to 143 454 176 bp, chromosome 4, REHH test) pig populations.

The gene set level functional enrichment analysis shows that there are strong selection candidates in the genomes of Taihu pigs that are associated with diseases, for example, asthma ($P = 0.00949$), autoimmune thyroid disease ($P = 0.00204$), the intestinal immune network for IgA production ($P = 0.00657$) and graft- *v.*-host disease ($P = 0.00811$). This result is likely to be consistent with the observation of the outstanding economic traits of Taihu pigs concerning high resistance to disease. Since the phenotypes of many traits (e.g. immune processes and behaviour) cannot be easily measured or scored for QTL/association studies, it is quite possible to cause the QTL mapping to lose its power for detecting genes subject to artificial selection during domestication. Nevertheless, the enrichment analysis for selected gene sets might be an effective way to identify some genes related to some specific phenotypes that are not easily measured (e.g. immune and disease traits). Overall, despite the Taihu pigs having a higher resistance to disease than Western pigs, we observed that asthma has been more common in recent years compared to the past for Taihu pigs. Interestingly, we found the genes under the selection regions to be over-represented for asthma (*IL9*, *SLA-DQB1*, *SLA-DRB1*, *SLA-DQA*, *SLA-DRA* and *ENSSSCG00000001458*) in Taihu pigs. This result indicates that the common occurrence of asthma may become one kind of selection pressure that leads to selection in the genomes of Taihu pigs. Currently, the asthma pathway contains 30 genes, and we speculated that other porcine homologous genes (except for the six genes identified under the selection regions in our study) could also experience selection pressure from asthma disease. In this case, asthma may result in a different frequency distribution of genotypes or alleles of these asthma-related genes between Taihu and Western pigs. To test our hypothesis, we implemented a χ^2 test for the genotypes and allele frequency distributions of these asthma-related genes in the two populations. We found that 11 out of 16 SNPs corresponding to eight genes (located in nine genes related to the asthma pathway in this study) exhibited a significantly different ($P < 0.01$) frequency distribution of alleles, and all of these SNPs displayed a significantly different ($P < 0.01$) frequency distribution of genotypes (Supplementary Table S9). In particular, the *SLA-DQB1* and *SLA-DQA* genes, which were under the selection regions, showed a different frequency distribution of both genotypes and alleles that was extremely significant. Although the different frequency distribution of genotypes and alleles in genes related to asthma disease might be a result from some other factors (e.g. population differentiation), we speculated that the genes in the asthma pathway are more likely to experience selection in Taihu pigs. Furthermore, for Western pigs, we found that genes under the selection regions were over-represented in pathways related to lipids and reproduction. During the past 100 years, breeding programmes for Western pigs based on pedigree information have achieved a remarkable genetic progress in reducing backfat (−75%) for carcass quality and increasing the growth rate (+100%) for production efficiency (Merks *et al.*, 2012). Although very small or negligible gains were made in reproduction traits, we still observed genetic progress in litter size (Merks *et al.*, 2012). Interestingly, we also found the genes under selection to be over-represented in pathways related to lipids (e.g. ether lipid metabolism, $P = 0.0169$; alpha-Linolenic acid metabolism, $P = 0.0203$) and reproduction (e.g. GnRH signalling pathway, $P = 0.00562$). These results indicate that genes related to lipids and reproduction traits (especially for litter size) have experienced intensive selection, and this correlates with the history of Western pig breeding programmes in terms of selection for carcass quality and reproduction traits. In addition, the genes related to reproduction and health traits may have experienced more selection in Taihu pigs compared to Western pigs. Based on the results of the genome comparison analysis (XPEHH test), we found that genes under the selection regions in Taihu pigs were over-represented in the GO terms related to reproduction (e.g. the female pregnancy and endomembrane system) and health (e.g. stress response), whereas the genes under selection regions in western pigs over-represented pathways related to olfaction (olfactory transduction pathway), which was also reported in the another study (Amaral *et al.*, 2011).

In order to check the signatures provided by method based on XPEHH, the F_{ST} analysis was also performed. Using the F_{ST} analysis, we detected the genes which were associated with growth. The most significant gene in the chromosome 7 was *LTBP2*. It was reported that this gene was associated with the number of ribs (Zhang *et al.*, 2016). For the genes annotation analysis, the term 'GO:0019838 ~ growth factor binding' associated with growth was also detected. Within this term, the *IGF1R* gene was detected in our results. It was reported that *IGF1R* was associated with growth and BW in pigs (Pierzchala *et al.*, 2012). These results may reflect the difference in growth between Taihu pig breeds and Western pig breeds. Interestingly, we also found that some selection signatures detected based on the XPEHH method were also detected based on the F_{ST} method. Within the area Chr16: 85 397 504 to 85 537 619 (Table 3), the SNP rs346134186 was detected with the F_{ST} value of 0.956 and the XPEHH *P* value of 0.003. This SNP was near the *IRX4* gene (21 kb apart) and has been reported to be related with BW at 71 days of age (Jung *et al.*, 2014). Within the area Chr13: 37 851 476 to 38 200 552 (Table 3), we detected a synonymous SNP rs330592274 (Chr13: 37 865 536) located in the *STAB1* gene. The XPEHH *P* value of this SNP was 0.009 and its ref allele frequency were 0.188 and 0.956 in Western pig breeds and Taihu pig breeds, respectively. The *STAB1* gene has been reported to be involved in immune response as a result of strong artificial in the intensive production systems in pig (Herrero-Medrano *et al.*, 2014). Therefore, the rs330592274 may be the candidate marker of selection signature.

Due to some important differences observed within the Taihu and Western groups, we explored the difference of

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

selection signatures among Chinese native breeds and Western pig breeds. The most significant gene detected among Taihu area pig breeds was *SLC27A1*. It was reported that this gene was associated with the lipid metabolism in pig (Melo *et al*., 2013). Among the Western pig breeds, the significant genes *CHST10* and *OCA2* were detected. It was reported that *CHST10* has a potential role for reproduction (Suzuki-Anekoji *et al*., 2013), and the *OCA2* gene was associated with pigmentation (Donnelly *et al*., 2012). These results were in accordance with the real situation among the Western pig breeds.

We also reported a genome-wide map of selection signatures of Taihu pigs. Further investigation of these selection signatures showed that many genes or QTLs were related to reproductive traits, coat colour and fat metabolism. In addition, gene enrichment analyses revealed that genes under selection candidate regions were significantly over-represented in pathways related to some complex diseases, such as autoimmune thyroid and asthma disease. This study provides insight into the architecture of selection in pig breeding, especially for production and health traits. Breeding has entered the genome era. It will be more effective to use genomic information in breeding plans most especially markers that has been associated with important economic traits. Both Chinese native pigs and Western pig breeds have their specific merits. It may be useful in pig breeding by combining their merits through a crossing production system.

Supplementary materials

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731118001714

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. U1402266, 31472069, 31370043). The authors thank Babatunde Shittu Olasege for improving the English of this paper.

Declaration of interest

The authors declare that they have no conflicts of interest.

Ethics statement

Animals were treated according to institutional guidelines, and the study protocol was approved by the Research Committee of Shanghai Jiao Tong University, Shanghai, China.

Software and data repository resources

The sequencing data information were obtained from the previous studies (Wang *et al*., 2016; Zhang *et al*., 2016). The Bioproject number is PRJNA281578 in the NCBI Sequence Read Archive.

References

Ai H, Huang L and Ren J 2013. Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide markers. PLoS One 8, e56001.

Amaral AJ, Ferretti L, Megens HJ, Crooijmans RP, Nie H, Ramos-Onsins SE, Perez-Enciso M, Schook LB and Groenen MA 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. PLoS One 6, e14782.

Bosse M, Lopes MS, Madsen O, Megens HJ, Crooijmans RP, Frantz LA, Harlizius B, Bastiaansen JW and Groenen MA 2015. Artificial selection on introduced Asian haplotypes shaped the genetic architecture in European commercial pigs. Proceedings of the Royal Society B 282, 20152019.

Chen K, Baxter T, Muir WM, Groenen MA and Schook LB 2007. Genetic resources, genome mapping and evolutionary genomics of the pig (Sus scrofa). International Journal of Biological Sciences 3, 153–165.

Chen Q, Ma Y, Yang Y, Chen Z, Liao R, Xie X, Wang Z, He P, Tu Y, Zhang X, Yang C, Yang H, Yu F, Zheng Y, Zhang Z, Wang Q and Pan Y 2013. Genotyping by genome reducing and sequencing for outbred animals. PLoS One 8, e67500.

China National Commission of Animal Genetic Resources 2011. Animal genetic resources in China pigs. China Agriculture Press, Beijing.

Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA 2003. DAVID: database for annotation, visualization, and integrated discovery. Genome Biology 4, P3.

Do DN, Strathe AB, Ostersen T, Jensen J, Mark T and Kadarmideen HN 2013. Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping. PLoS One 8, e71509.

Donnelly MP, Paschou P, Grigorenko E, Gurwitz D, Barta C, Lu RB, Zhukova OV, Kim JJ, Siniscalco M, New M, Li H, Kajuna SLB, Manolopoulos VG, Speed WC, Pakstis AJ, Kidd JR and Kidd KK 2012. A global view of the OCA2-HERC2 region and pigmentation. Human Genetics 131, 683–696.

Herrero-Medrano JM, Megens HJ, Groenen MAM, Bosse M, Perez-Enciso M and Crooijmans RPMA 2014. Whole-genome sequence analysis reveals differences in population management and selection of European low-input pig breeds. BMC Genomics 15, 601.

Jung EJ, Park HB, Lee JB, Yoo CK, Kim BM, Kim HI, Kim BW and Lim HT 2014. Genome-wide association analysis identifies quantitative trait loci for growth in a Landrace purebred population. Animal Genetics 45, 442–444.

Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A, Vigne JD, Edwards CJ, Schlumbaum A, Dinu A, Balacsescu A, Dolman G, Tagliacozzo A, Manaseryan N, Miracle P, Van Wijngaarden-Bakker L, Masseti M, Bradley DG and Cooper A 2007. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. Proceedings of the National Academy of Sciences of the United States of America 104, 15276–15281.

Li C, Sun D, Zhang S, Wang S, Wu X, Zhang Q, Liu L, Li Y and Qiao L 2014. Genome wide association study identifies 20 novel promising genes associated with milk fatty acid traits in Chinese Holstein. PLoS One 9, e96186.

Li H and Durbin R 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R, and Genome Project Data Processing S 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Lin X, Shim K and Odle J 2010. Carnitine palmitoyltransferase I control of acetogenesis, the major pathway of fatty acid {beta}-oxidation in liver of neonatal swine. American Journal of Physiology-Regulatory Integrative and Comparative Physiology 298, R1435–1443.

Ma Y, Wei J, Zhang Q, Chen L, Wang J, Liu J and Ding X 2015. A genome scan for selection signatures in pigs. PLoS One 10, e0116850.

Mayer M, Bercsenyi K, Geczi K, Szabo G and Lele Z 2010. Expression of two type II cadherins, Cdh12 and Cdh22 in the developing and adult mouse brain. Gene Expression Patterns 10, 351–360.

Megens HJ, Crooijmans RP, San Cristobal M, Hui X, Li N and Groenen MA 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. Genetics Selection Evolution 40, 103–128.

Melo C, Gallardo D, Quintanilla R, Zidi A, Castello A, Diaz I, Amills M and Pena RN 2013. An association analysis between polymorphisms of the pig solute

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press

carrier family 27A (SLC27A), member 1 and 4 genes and serum and muscle lipid traits. Livestock Science 152, 143–146.

Merks JW, Mathur PK and Knol EF 2012. New phenotypes for new breeding goals in pigs. Animal 6, 535–543.

Oczkowicz M, Ropka-Molik K and Tyra M 2013. Analysis of the associations between polymorphisms in GNAS complex locus and growth, carcass and meat quality traits in pigs. Molecular Biology Reports 40, 6419–6427.

Oleksyk TK, Smith MW and O'Brien SJ 2010. Genome-wide scans for footprints of natural selection. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 365, 185–205.

Onteru SK, Gorbach DM, Young JM, Garrick DJ, Dekkers JC and Rothschild MF 2013. Whole genome association studies of residual feed intake and related traits in the pig. Plos One 8, e61756.

Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW and Pritchard JK 2009. Signals of recent positive selection in a worldwide sample of human populations. Genome Research 19, 826–837.

Pierzchala M, Pareek CS, Urbanski P, Goluch D, Kamyczek M, Rozycki M, Smoczynski R, Horbanczuk JO and Kuryl J 2012. Study of the differential transcription in liver of growth hormone receptor (GHR), insulin-like growth factors (IGF1, IGF2) and insulin-like growth factor receptor (IGF1R) genes at different postnatal developmental ages in pig breeds. Molecular Biology Reports 39, 3055–3066.

Revilla M, Ramayo-Caldas Y, Castello A, Corominas J, Puig-Oliveras A, Ibanez-Escriche N, Munoz M, Ballester M and Folch JM 2014. New insight into the SSC8 genetic determination of fatty acid composition in pigs. Genetics Selection Evolution 46, 28.

Rubin CJ, Megens HJ, Martinez Barrio A, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg O, Jern P, Jorgensen CB, Archibald AL, Fredholm M, Groenen MA and Andersson L 2012. Strong signatures of selection in the domestic pig genome. Proceedings of the National Academy of Sciences of the United States of America 109, 19529–19536.

Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R and Lander ES 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature 419, 832–837.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF and Lander ES, International HapMap C, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Johnson TA, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R and Stewart J 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449, 913–918.

Scheet P and Stephens M 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. American Journal of Human Genetics 78, 629–644.

Suzuki Y 2010. Statistical methods for detecting natural selection from genomic data. Genes & Genetic Systems 85, 359–376.

Suzuki-Anekoji M, Suzuki A, Wu SW, Angata K, Murai KK, Sugihara K, Akama TO, Khoo KH, Nakayama J, Fukuda MN and Fukuda M 2013. In vivo regulation of steroid hormones by the Chst10 sulfotransferase in mouse. Journal of Biological Chemistry 288, 5007–5016.

Vicart S, Sternberg D, Fontaine B and Meola G 2005. Human skeletal muscle sodium channelopathies. Neurological Sciences 26, 194–202.

Vidal O, Noguera JL, Amills M, Varona L, Gil M, Jimenez N, Davalos G, Folch JM and Sanchez A 2005. Identification of carcass and meat quality quantitative trait loci in a Landrace pig population selected for growth and leanness. Journal of Animal Science 83, 293–300.

Villela D, Kimura L, Schlesinger D, Goncalves A, Pearson PL, Suemoto CK, Pasqualucci C, Krepischi AC, Grinberg LT and Rosenberg C 2013. Germline DNA copy number variation in individuals with Argyrophilic grain disease reveals CTNS as a plausible candidate gene. Genetics and Molecular Biology 36, 498–501.

Wang Z, Chen Q, Liao R, Zhang Z, Zhang X, Liu X, Zhu M, Zhang W, Xue M, Yang H, Zheng Y, Wang Q and Pan Y 2016. Genome-wide genetic variation discovery in Chinese Taihu pig breeds using next generation sequencing. Animal Genetics 48, 38–47.

Wang Z, Chen Q, Yang Y, Yang H, He P, Zhang Z, Chen Z, Liao R, Tu Y, Zhang X, Wang Q and Pan Y 2014. A genome-wide scan for selection signatures in Yorkshire and Landrace pigs based on sequencing data. Animal Genetics 45, 808–816.

Weir BS and Cockerham CC 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Wuyts W, Spieker N, Van Roy N, De Boulle K, De Paepe A, Willems PJ, Van Hul W, Versteeg R and Speleman F 1999. Refined physical mapping and genomic structure of the EXTL1 gene. Cytogenetics and Cell Genetics 86, 267–270.

Yang Y, Wang Q, Chen Q, Liao R, Zhang X, Yang H, Zheng Y, Zhang Z and Pan Y 2014. A new genotype imputation method with tolerance to high missing rate and rare variants. Plos One 9, e101025.

Zhang LC, Yue JW, Pu L, Wang LG, Liu X, Liang J, Yan H, Zhao KB, Li N, Shi HB, Zhang YB and Wang LX 2016. Genome-wide study refines the quantitative trait locus for number of ribs in a Large White x Minzhu intercross pig population and reveals a new candidate gene. Molecular Genetics and Genomics 291, 1885–1890.

Zhang Z, Wang Z, Yang Y, Zhao J, Chen Q, Liao R, Chen Z, Zhang X, Xue M, Yang H, Zheng Y, Wang Q and Pan Y 2016. Identification of pleiotropic genes and gene sets underlying growth and immunity traits: a case study on Meishan pigs. Animal 10, 550–557.

https://doi.org/10.1017/S1751731118001714 Published online by Cambridge University Press