



# Detection of Selection Signatures in Chinese Landrace and Yorkshire Pigs Based on Genotyping-by-Sequencing Data

Kai Wang<sup>\*†</sup>, Pingxian Wu<sup>†</sup>, Qiang Yang, Dejuan Chen, Jie Zhou, Anan Jiang, Jideng Ma, Qianzi Tang, Weihang Xiao, Yanzhi Jiang, Li Zhu, Xuewei Li and Guoqing Tang<sup>\*</sup>

Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu, China

## OPEN ACCESS

### Edited by:

Peng Xu,  
Xiamen University, China

### Reviewed by:

Fabyano Fonseca Silva,  
Universidade Federal de Viçosa, Brazil  
Shikai Liu,  
Ocean University of China, China

### \*Correspondence:

Kai Wang  
614351818@qq.com  
Guoqing Tang  
tyq003@163.com

<sup>†</sup> These authors have contributed equally to this work.

### Specialty section:

This article was submitted to  
Livestock Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 19 November 2017

**Accepted:** 26 March 2018

**Published:** 09 April 2018

### Citation:

Wang K, Wu P, Yang Q, Chen D, Zhou J, Jiang A, Ma J, Tang Q, Xiao W, Jiang Y, Zhu L, Li X and Tang G (2018) Detection of Selection Signatures in Chinese Landrace and Yorkshire Pigs Based on Genotyping-by-Sequencing Data. *Front. Genet.* 9:119. doi: 10.3389/fgene.2018.00119

The domestic pigs have been undergone intense selection pressures for these development of interested traits following domestication and modern breeding. This has altered many traits in most of pig breeds, such as growth rate, body weight, fertility, and immunity. Thus, the objectives of this study were to (1) detect these selection signatures and identify the candidate genes that show evidences of recent artificial selection at the level of whole genome, (2) be beneficial to understand the relationship between genomic structure and phenotypic diversity, and (3) highlight the key roles of these candidate genes in growth and development in the two breeds. The data consisted of total raw number of 345570 single nucleotide polymorphisms (SNPs) in 1200 individuals from the Chinese Landrace pigs (L,  $n = 600$ ) and Yorkshire pigs (Y,  $n = 600$ ). Based on these SNPs data, two complementary methods, population differentiation ( $F_{st}$ ) and composite likelihood ratio test (CLR), were carried out to detect the selection signatures in this study. A total of 540 potential selection regions (50 kb) which contained 111 candidate genes were detected for Landrace-Yorkshire pair (L-Y) by  $F_{st}$ . In addition, 73 and 125 candidate genes were found for Landrace pigs and Yorkshire pigs by CLR test based on 321 and 628 potential selection regions, respectively. Some candidate genes are associated with important traits and signaling pathways including the *ACACA*, *MECR*, *COL11A1*, *GHR*, *IGF1R*, *IGF2R*, *IFNG*, and *MTOR* gene. The *ACACA* and *MECR* gene are related to fatty acid biosynthesis. The *COL11A1* gene is essential for the development of the normal differentiation. The *GHR*, *IGF1R*, and *IGF2R* gene are significant candidate genes which play major roles in the growth and development in animals. The *IFNG* gene is associated with some aspects of immune response. The *MTOR* gene regulates many signaling pathways and signaling transduction pathway.

**Keywords:** pig, selection signature, artificial selection, domestication, genome

## INTRODUCTION

As a major protein source of humans, the swine is one of the most important livestock (Amaral et al., 2011; Wilkinson et al., 2013). The domestic pig originated from the wild boar (*Sus scrofa*) that were mainly distributed in Europe and Asia (Larson et al., 2005, 2007). Artificial selection altered the phenotypic traits of these pigs. Many pig breeds were objectively selected for desirable performance traits such as rapid growth, increased lean meat and enhanced prolificacy. A number of trait differentiations, therefore, occurred in a short time, resulting in the development of distinct pig breeds phenotypes. It is possible to investigate the effects of selective pressure on various livestock at the genome-wide level due to the development of high-throughput genotyping technologies. Studies have been concentrated on the specific traits relevant to the growth and development processes with the object of identifying and characterizing selection signatures, and then identifying the potential causal mutations in order to understand the genetic basis structure for phenotypic variation in the swine (Andersson, 2001; Andersson and Georges, 2004). Mutations conferring new favorable genotypes will be subject to a selective sweep, which is an increase rapidly in allele frequency under artificial selection. Moreover, the artificial selection led to the traits changes mainly related to growth and development. For example, a single nucleotide mutation in the *IGF2* gene lead to a major effect on muscle growth in some commercial pig breeds (Laere et al., 2003). The *MC4R* gene is related to growth and fatness both in pigs and humans (Kim et al., 2000; Ovilo et al., 2006).

When a beneficial mutation emerges and subsequently spreads in a population, this process, that is selective sweep (Smith and Haigh, 1974), will generate higher population differentiations, higher frequencies of segregating sites and linkage disequilibrium (LD) patterns (Grossman et al., 2010). In this study, the *Fst* and CLR test were carried out to assess population difference and detect artificial selection. Population differentiation was accessed by *Fst* based on the DNA polymorphism in populations (Wright, 1931; Hudson et al., 1992; Nielsen, 2005; Weir et al., 2005). While CLR test has a high power to detect selective sweeps using site frequency spectrum (SFS) patterns of SNP (Williamson et al., 2007).

Although many studies have been detected a lot of selection signatures in the pig, those findings are not entirely consistent. Andersson and Plastow (2011) and Rubin et al. (2012) detected selection signatures in European domestic pig and wild boar using whole-genome resequencing. Selective sweep analyses revealed strong selection signatures at three loci which were related to morphological changes in the domestic pig (Rubin et al., 2012). Fu et al. (2016) Identified 417 protein-coding genes which were mainly associated with developmental and metabolic processes by combining Enshi black pigs and Chinese wild boars. The reasons are not only the differences of the statistical methods, but also the varieties of SNP panels density and sample size. In addition, domestic pigs under different evolution conditions (different breeding objectives) show different selection signatures even in the same breed. Therefore, it is important to explore

selection signatures in more pig breeds and these findings will provide a foundation to investigate the artificial selection process and domestication history of the two commercial pig breeds.

## MATERIALS AND METHODS

### Genotype by Sequencing Data

In the study, we adopted a surfactant and the protease pyrolysis method to extract genomic DNA from ear tissue. A total of 600 Landrace and 600 Yorkshire pigs were genotyped using GBS technology (Elshire et al., 2011). The data were sequenced by using Illumina HiSeq PE150. In the raw reads, N contents with < 10% of sequence length or with high quality bases (>5) and a number < 50% of the sequence length were retained. The clean data were aligned against the *Scrofa* 10.2 using Burrows Wheeler Aligner (BWA) with the parameters 'mem -t 4 -k 32 -M' (Li and Durbin, 2009). Then we found a total of 10,445,924 SNPs by using the Genome Analysis Toolkit (GATK) with default parameters (Depristo et al., 2011). And a total of 345,570 SNP markers met the quality requirements using VCFtools with the parameters '-min-meanDP 3 -maf 0.01 -max-missing 0.2' (Danecek et al., 2011). The data that support the findings is publically available at figshare<sup>1</sup> under doi: 10.6084/m9.figshare.5960914.

### Populations and the Data Quality Control

A total of 1200 individuals from two commercial pig breeds, 600 Chinese Landrace pigs and 600 Chinese Yorkshire pigs, were genotyped by sequencing. Genotyping contains raw data with a total of 345570 SNPs. And after quality control we obtained 92114 SNPs. The quality control was determined using the PLINK program (Purcell et al., 2007). A quality control was adopted to access the high data quality by (1) removing SNPs loci with call rate less than 0.95 and unknown position, (2) removing SNPs loci with minor allele frequency (MAF) less than 0.05, (3) discarding the individuals with call rate less than 0.95, (4) removing SNPs loci with Hardy Weinberg balance test less than 10<sup>-6</sup> and (5) removing SNPs loci in sex chromosomes. Following the quality control, the missing genotypes was imputed using the BEAGLE (Browning and Browning, 2007). After imputation, the average estimated squared correlation (R<sup>2</sup>) between the allele dosage with highest posterior probability and the true allele dosage fore the marker is 0.9982. And then the principal component analysis (PCA) for population structure based on SNPs information was performed using the EIGENSOFT (Patterson et al., 2006).

### Methods for Detection of Selection Signatures

Two methods, *Fst* and CLR test, were implemented to detect the selection signatures. The two approaches are all directly handling the SNP genotype. Hudson's *Fst* (Hudson et al., 1992), a classical measure of population differentiation, was used to detect

<sup>1</sup><https://figshare.com/>

selection signatures between Landrace pigs and Yorkshire pigs at each SNP in this study. For each SNP in a pairwise comparison, the expected  $F_{st}$  was calculated with the PopGenome v2.2.4 in R (Pfeifer et al., 2015).

A recent selective sweep changes patterns of allele frequency at linked sites, excluding variation at closely linked loci and generating a relative excess of allele at very low and very high frequencies at more aloof loci (Williamson et al., 2007). A statistical method, CLR test (Nielsen et al., 2005), was implemented to search for the particular pattern of allele frequencies along a chromosome following a selective sweep. The CLR value was computed using the SweeD program (Pavlidis et al., 2013) at each SNP for the two breeds respectively.

## Identification of Candidate Genes Under Selection and Gene Annotation

An outlier-approach to obtain the candidate genes under selection was taken (Yang et al., 2014). For all loci, a SNP which corresponded to the upper 1% of the empirical genome-wide distribution of  $F_{st}$  was considered to be a high- $F_{st}$  outlier. Accordingly, candidate selected regions were deemed as the 99th percentile of the empirical genome-wide distribution of CLR. And then, the distance cutoff was limited to be 500 bp in order to obtain a high-quality gene for gene annotations. Thus, a SNP was judged to belong with a gene if it is located within the region defined by 500 bp upstream of gene start site and 500 bp downstream of the gene end site. Genes found within the intervals spanning the candidate regions were searched from the Ensembl genome browser<sup>2</sup> using the Sscrofa 11.1 reference genome and these were considered as candidate genes. Function annotation and enrichment of the candidate genes were displayed using the DAVID browser<sup>3</sup> and the KEGG<sup>4</sup>. Phenotype that are known to be affected by the identified candidate genes were inferred from the literature.

## RESULTS

### Information of Filtered Data and Population Structure

A genome-wide scan for selection signatures in two commercial pig breeds was carried out by estimating  $F_{st}$  and CLR at each marker. In the process of quality control, 171121 SNPs and 100 individuals were discarded due to missing genotype data, 45600 SNPs and 31825 SNPs were removed due to minor allele thresholds and Hardy Weinberg balance thresholds, and 4910 SNPs in sex chromosomes were discarded. After the quality control, 1100 individuals (546 Landrace pigs and 554 Yorkshire pigs) with 92114 SNPs were retained for this analysis (Tables 1, 2). Table 3 summaries the whole genome potential regions (50 kb) at 18 autonomies by  $F_{st}$  and CLR.

Population structure was investigated using PCA analysis. The PCA was carried out based on all available SNPs to examine the

population genetic structure in this study. As shown in Figure 1, most of Landrace pigs and Yorkshire pigs formed a separate cluster severally. Based on the population structure information by PCA, we founded that the raw determination data from the farm had some mistakes about genealogy so we adjusted genealogy for specific individuals in the two breeds.

### Selection Signatures Shared in Two Breeds – $F_{st}$ Approach

According to the empirical distribution of  $F_{st}$  (Figure 2), candidate regions under selection were defined as outliers falling with the upper 1% of the distribution of  $F_{st}$  ( $F_{st} > 0.243179$ ). Supporting information shows graphically the  $F_{st}$  of each SNP for all 18 autonomies (Supplementary Figure S1).

TABLE 1 | Information of filtered data and candidate genes detected by  $F_{st}$ .

Method	Items	L-Y
Fst	SNPs	92114
	Outliers	921
	Potential selection regions	540
	Candidate genes	111

TABLE 2 | Information of filtered data and candidate genes detected by CLR.

Method	Items	L	Y
CLR	SNPs	92114	92114
	Outliers	921	921
	Potential selection regions	321	628
	Candidate genes	73	125

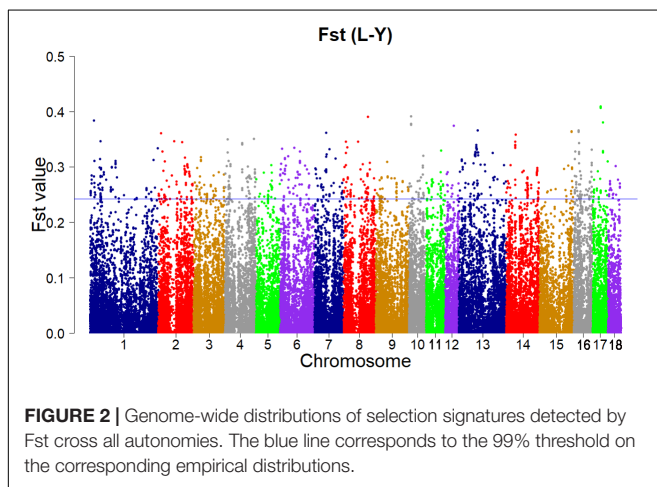
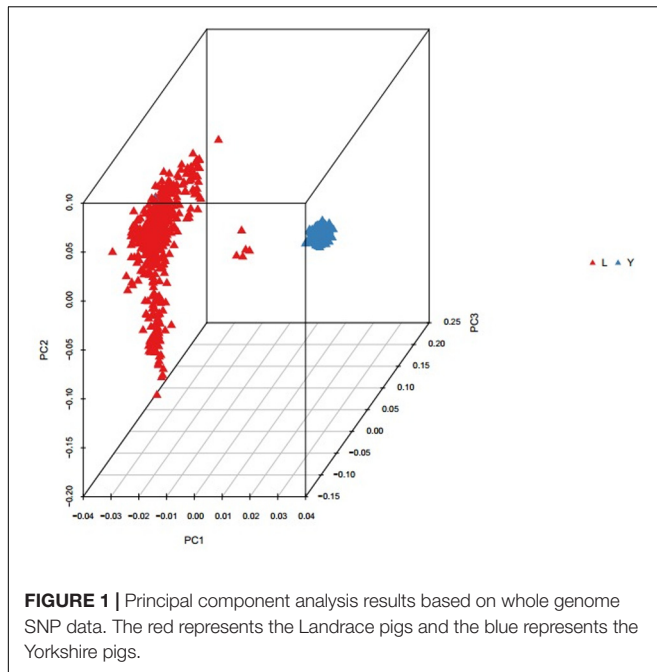
TABLE 3 | Summaries of the numbers of potential selection regions.

Chr	L-Y ( $F_{st}$ )	L (CLR)	Y (CLR)
1	50	49	66
2	36	25	35
3	30	15	42
4	37	16	41
5	23	17	39
6	37	17	39
7	29	21	43
8	43	24	39
9	27	8	51
10	19	16	26
11	19	22	35
12	11	10	10
13	52	19	42
14	42	22	30
15	16	17	27
16	41	9	30
17	16	6	19
18	12	8	14
Total	540	321	628

<sup>2</sup><http://www.ensembl.org/>

<sup>3</sup><http://david.ncicrf.gov/>

<sup>4</sup><http://www.genome.jp/kegg/pathway.html>



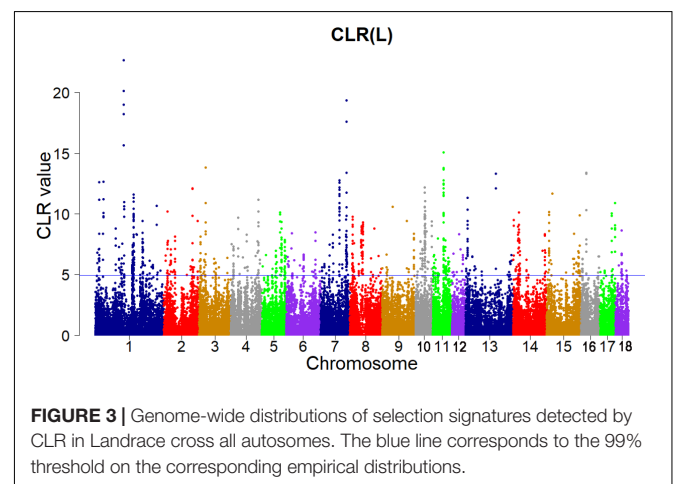
Using this strategy, the result of  $F_{st}$  revealed 111 candidate genes of high levels of differentiation (**Supplementary Table S1**). A gene on chromosome 12 (SSC 12), *ACACA* gene ( $F_{st} = 0.373979048$ ), plays a key role in fatty acid biosynthesis (Stachowiak et al., 2013). The acetyl-CoA carboxylase alpha that catalyzes the carboxylation of acetyl-CoA to malonyl-CoA is encoded by the *ACACA* gene. And malonyl-CoA regulates mitochondrial fatty acid  $\beta$ -oxidation and has effects on the expression of hypothalamic neuropeptides modulating energy homeostasis and lipid metabolism (Bionaz and Looor, 2008; Wakil and Abuelheiga, 2009). Some QTLs for average backfat thickness and birth body weight, the percentage of vaccenic, stearic, palmitic, and palmitoleic fatty acids and the total percentage of saturated fatty acids in backfat and/or skeletal muscles are corresponded with the precise localization of *ACACA* gene (Liu et al., 2007, 2008; Quintanilla et al., 2011).

On SSC 6, the *MTOR* gene ( $F_{st} = 0.33406672$ ) was identified as an important selection signature. *MTOR* gene is related to multiple signaling and signal transduction pathways in various of cellular processes (Hay and Sonenberg, 2004). The *MTOR* gene encodes mammalian target of rapamycin (mTOR), which regulates multiple biological processes such as cellular metabolism, growth and survival in response to hormones, growth factors, nutrients, energy and stress signals as a serine/threonine protein kinase. Moreover, the mTOR signaling pathway regulates many major biological processes and is related to many pathological conditions such as cancer, obesity, type 2 diabetes, and neurodegeneration (Guertin and Sabatini, 2007; Laplante and Sabatini, 2012). An important gene, *GHR* gene ( $F_{st} = 0.268147326$ ), which was implicated in promoting the growth and development of animals was deemed as a significant candidate gene on SSC 16 (Leung et al., 1987; Schnoebelen-Combes et al., 1996). The significance of growth hormone (GH) produced by the pituitary gland in supporting growth and development has been known for a long time (Dauncey et al., 1994). The major effect of GH is to promote postnatal longitudinal growth. And GH regulates the lipid, carbohydrate, nitrogen, and mineral metabolism within a cell through interaction with the GH receptor on target cells (Kopchick and Andry, 2000).

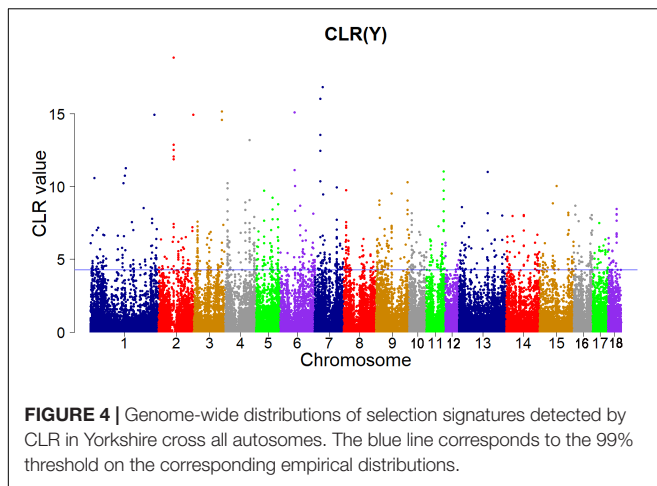
## Selection Signatures Unique to Individual Breeds– CLR Approach

Candidate regions were taken as outliers falling with the 99th percentile of the distribution of CLR (**Figures 3, 4**). Supporting information shows graphically the CLR of each SNP for all 18 autonomies of Landrace pigs (**Supplementary Figure S1**). Supporting information shows graphically the CLR of each SNP for all 18 autonomies of Yorkshire pigs (**Supplementary Figures S2, S3**).

Add up to 73 candidate genes were found for the Landrace pigs using this criterion (CLR > 4.928279, **Supplementary Table S2**). On SSC 4, the *COL11A1* gene (CLR = 11.18224) encodes one of the two alpha chains of type XI collagen which is a minor fibrillar collagen. The type XI collagen in







cartilage extracellular matrix is essential for the development of the normal differentiation and spatial organization structure of growth plate chondrocytes (Li et al., 1995). Furthermore, the *COL11A1* gene also shows high evidence of selection in Tibetan wild boar (Li et al., 2013). The *MECR* gene (CLR = 6.377755) on SSC 6 encodes an oxidoreductase that catalyzes the last step in mitochondrial fatty acid synthesis (mtFAS). The mtFAS is a highly conserved metabolic process and is necessary for mitochondrial respiration as a highly conserved metabolic process. It was recently reported that mitochondrial 2-enoyl-CoA/ACP reductase played an important role in placental development in mice (Nair et al., 2017) and the mutations of *MECR* resulted in childhood-onset dystonia and optic atrophy (Heimer et al., 2016).

A total of 125 candidate genes with high CLR score (CLR > 4.272431) were identified for the Yorkshire pigs based on the distribution of CLR (**Supplementary Table S3**). Some genes play key role in growth, immune and fertility of animals. Two genes on SSC 1, the *IGF1R* gene (CLR = 6.658113) and the *IGF2R* gene (CLR = 4.433161), were important candidate genes in growth and development of Yorkshire pigs. The insulin-like growth factor (IGF)-system includes insulin-like growth factors 1 and 2 (*IGF1* and *IGF2*) along with the type I (*IGF1R*) and type II (*IGF2R*) cell-surface receptors, the insulin receptors (*IR*) and circulating IGF-binding proteins (*IGFBPs*) (Wang et al., 2003). *IGF1R* and *IR* regulate the biological processes and lead to cell growth, differentiation and survival. In mammals, the activity of *IGF2* is mediated by *IGF2R*, which sequesters *IGF2* for internalization and degradation (Brown et al., 2008). The *IFNG* gene (CLR = 6.903143) encodes a soluble cytokine that belongs to the type II interferon (*IFN*) class. *IFNs* are cytokines that play a significant role in the resistance of mammalian hosts to pathogens (Boehm et al., 1997). *IFN- $\gamma$*  is secreted by thymus-derived (T) cells under specific activation conditions and by natural killer (NK) cells. The biological effects of *IFN- $\gamma$*  are regulation of some ways of the immune response, including stimulation of antigen presentation by class I and class II major histocompatibility complex (MHC) molecules, stimulation of bactericidal activity of phagocytes, effects on cell proliferation and

apoptosis, orchestration of leukocyte-endothelium interactions (Schroder et al., 2004).

## Functional Annotation of the Candidate Genes

The 111 candidate genes that were identified at high Fst values for Landrace-Yorkshire pair (L-Y), and the 73 and 125 candidate genes that were found by CLR for Landrace pigs and Yorkshire pigs respectively, were accessed for functional enrichment using the DAVID browser. Most of these genes were implicated in multiple signaling and signal transduction pathways in multifarious cellular and biochemical processes (**Supplementary Table S4**). In the cAMP signaling pathway seven candidate genes among them were identified (*CNGB1*, *GABBR1*, *GRIA2*, *HTR4*, *PPARA*, *RAC3*, *TIAMI*). Cyclic adenosine 3',5'-monophosphate is a member of the most common second messengers. The cAMP is a regulator of pivotal physiologic processes including metabolism, secretion, calcium homeostasis, muscle contraction, cell fate, and gene transcription (Fimia and Sassone-Corsi, 2001). The Jak-STAT signaling pathway contained five candidate genes under selection (*GHR*, *IFNG*, *IL20RB*, *IL23R*, *OSMR*). In mammals, the Jak-STAT pathway is the major signaling mechanism for a wide variety of cytokines and growth factors. Jak-STAT signaling is indispensable for lots of developmental and homeostatic processes, including hematopoiesis, immune cell development, stem cell maintenance, organismal growth, and mammary gland development (Imada and Leonard, 2000; Harrison, 2012; Kiu and Nicholson, 2012). In Wnt signaling pathway, three candidate genes were found (*BAMBI*, *FZD4*, *RAC3*). The Wnt signaling pathway is a regulator in the core biological processes of proliferation, differentiation, stem cell renewal and in the origin of cancer when deregulated (Huelsen and Behrens, 2002). And signaling molecules secreted by the Wnt family have been found to come into play in controlling embryonic development from hydra to human (Yang, 2010). A variety of candidate genes were also identified in pathways upstream and downstream of the mTOR signaling pathway including the insulin signaling pathway (*ACACA*, *FOXO1*, *MTOR*, *PPARA*), the AMPK signaling pathway (*FOXO1*, *IGF1R*), the MAPK signaling pathway (*PTPRR*, *RAC3*), the PI3K-Akt signaling pathway (*COL11A1*, *GHR*, *IGF1R*, *MDM2*, *MTOR*, *OSMR*,

**TABLE 4 |** Overlapped candidate genes between Fst and CLR.

Genes	SSC	Methods
CNIH3	10	Fst, CLR(Y)
COL14A1	4	Fst, CLR(Y)
DOCK9	11	CLR(L), CLR(Y)
GRM7	13	Fst, CLR(Y)
IL23R	5	Fst, CLR(L)
LRGUK	18	Fst, CLR(Y)
OAT	14	CLR(L), CLR(Y)
SPAG17	4	CLR(L), CLR(Y)
TBC1D5	13	Fst, CLR(Y)

*PKN2*, *RAC3*), the glycerolipid metabolism (*AGK*, *DGKH*, *LCLAT1*), the VEGF signaling pathway (*RAC3*), the regulation of autophagy (*ATG7*, *IFNG*) and the regulation of actin cytoskeleton (*PAK5*, *RAC3*, *TIAMI*) (Saltiel and Kahn, 2001; Bevan, 2001; Chen et al., 2001; Cross et al., 2003; Meijer and Codogno, 2004; Engelman et al., 2006; Prentki and Madiraju, 2008; Steinberg and Kemp, 2009; Lee and Dominguez, 2010). Besides, the pathway neuroactive ligand-receptor interaction contained 10 candidate genes (*THRB*, *GRIA2*, *GRIK1*, *GRM7*, *GLRA3*, *HTR4*, *GABBR1*, *NPPFR1*, *PLG*, *GHR*); the metabolic pathways contained 27 candidate genes (*ME1*, *NDST3*, *KYNU*, *HSD17B12*, *ASL*, *STT3B*, *TPK1*, *GALC*, *GALNT16*, *PAFAH1B1*, *RPIA*, *AGK*, *CYP4A24*, *MOCS2*, *MANIA2*, *BST1*, *ACACA*, *DGKH*, *AK7*, *POLR3C*, *CD38*, *LCLAT1*, *INPP4B*, *GULO*, *OAT*, *MECR*, *PCCB*).

## DISCUSSION

Over the past 300 years, intense artificial selection for production traits has led to the progress of many pig breeds with specialized phenotypic traits. In this study, two different methods, Fst and CLR, were carried out for detecting selection signatures. And then various genes could be deemed as candidate genes based on function or previous study with interesting traits in pig breeds. The same principle of data filtering was adopted to ensure the accuracy of the two methods.

### The Effect of the Methods on the Results

The Fst method is a classical approach for detecting selection signatures based on population differentiation. Under the condition of neutral evolution, the degree of differentiation between populations depends on the genetic drift intensity, as well as the difference in mutation rate and recombination rate in genome. However, when a locus was positively selected on only one population and remained neutral on another, or if the two populations are selected in different directions at same site, the two populations will increase genetic differentiation at that site. Thus, significant differences in allele frequencies between populations can be used as a selective effect of the left blot. On the other hand, if there is a very low degree of differentiation between the two populations in

the locus, it may also be caused by a balance selection, a purification selection, or a positive selection in the same direction.

The CLR method uses the combined likelihood of multiple markers to detect the genomic region selected. Besides, the CLR method is an available method for detecting sweeps that is not highly sensitive to assumptions about the underlying recombination rate or recombination hotspots (Williamson et al., 2007; Frantz et al., 2015). Thus, we used CLR test that compared allele frequencies in genome regions to the background pattern of variation so that we can obtain a relatively accurate result. Many other studies indicate that the CLR test has power to detect old selective sweep which occurred in wide time scale. And as we know, domestic pigs originated from wild boar about 9000 years ago so that some old selective sweeps may be detected by CLR test. However, we can not rule out a probability that the CLR test has not detected some old selective sweeps if these genome regions had selected frequently. Even so, we believe that these identified results are very useful to future studies.

### Pig Production Traits

Selection signatures were found that may be involved in some desirable pig production traits. **Table 5** lists the comparison of data on growth traits between two populations.

Traditional pig breeding program have concentrated on growth rate and leanness (Hammond, 1998). However, in order to satisfy the demands of consumers, meat quality has become a an important objective in genetic selection during the last few years (Vidal et al., 2005; Uimari et al., 2013). Fat deposition is a significant trait that directly affects the meat quality. According to the result of Fst, the *ACACA* gene was regarded as a candidate gene that plays an important role in fatty acid biosynthesis. Therefore, genes related to fatty acid biosynthesis were selected by a more intense artificial selection in two breeds. The result of CLR has supported this point: the *MECR* gene detected in Landrace breed was associated with mtFAS. However, no more genes associated with fatty acid biosynthesis were detected by Fst and CLR. The possible explanation is that the two breeds had similar breeding objectives about growth and development for a long time, however, meat quality has not been selected systematically in the last decades.

**TABLE 5** | Comparison of data on growth traits between two populations.

Breeds	Terms	Weight(kg)	BF(mm)	LMT(%)	B100(mm)	D100(d)
Landrace	Mean	107.41	10.10	60.75	9.45	162.64
	Max	136.00	20.30	81.00	20.48	204.20
	Min	75.00	5.00	44.00	4.87	135.90
	SD	14.95	2.51	5.94	2.06	11.49
Yorkshire	Mean	107.30	10.18	60.70	9.58	161.08
	Max	159.00	20.00	86.00	15.51	274.45
	Min	75.00	5.00	40.00	4.42	120.59
	SD	13.28	2.51	5.56	2.07	11.85

BF, backfat thickness; LMT, lean muscle thickness; B100, backfat thickness to 100 kg; D100, days to 100 kg; Max, maximum; Min, minimum; SD, standard deviation.

Therefore, most candidate genes were associated with growth and development, and yet a few candidate genes were involved in meat quality.

According to the functional annotation of the candidate genes, several pathways were directly associated with growth and development of animals. For example, the insulin signaling pathway that regulate individual development begin with the binding of insulin to insulin receptors, and thus trigger a series of intracellular signal transduction, ultimately playing a role after reaching the organ (White and Kahn, 1994); the PI3K-Akt signaling pathway is activated by many kinds of cellular stimuli or toxic insults activate the PI3k-Akt signaling pathway which is a regulator for fundamental cellular function including transcription, translation, proliferation, growth and survival (Song et al., 2005). The genes associated with growth and development were identified by Fst and CLR. This suggested that growth and development (growth rate and feed ratio) had been the major breeding goal in the last few decades.

Several genes were identified as candidate genes by both Fst and CLR approaches (Table 4). The *IL23R* gene encodes interleukin-23 (IL23) receptor. IL23 is a key factor in innate and adaptive immunity and may participate in acute response to infection in peripheral tissues. IL23 is responsible for autoimmune inflammatory diseases and is important for inflammatory bowel diseases in human (Duerr et al., 2006). Three genes (*DOCK9*, *OAT* and *SPAG17*) were identified by CLR in both Landrace pigs and Yorkshire pigs. The *DOCK9* gene encodes dedicator of cytokinesis protein 9. The *OAT* gene encodes ornithine aminotransferase (OAT) which plays crucial physiological roles in amino acid metabolism (Levillain et al., 2007). The *SPAG17* gene encodes a central pair protein present in the axonemes of cells with a “9 + 2” organization of microtubules. The *SPAG17* gene plays a vital role in the function and structure of motile cilia (Teves et al., 2013).

## Comparison With Previous Studies

Selection signatures detected in this study were compared with previous studies. The *KIT* gene and the *MC1R* gene were important candidate genes relevant to coat color (Andersson and Plastow, 2011; Rubin et al., 2012). The two breeds are all white color, so the two genes were not detected under selection in this study. In addition, the *IGFBP7* gene and the *UNC13C* gene were overlapped with Yang's report (Yang et al., 2014). Furthermore, a total of 7 genes (*COL11A1*, *COL14A1*, *IFNG*, *IGF2R*, *MTOR*, *MYO10*, *PTPRR*) were overlapped with Li's study (Li et al., 2013).

## REFERENCES

- Amaral, A. J., Ferretti, L., Megens, H. J., Crooijmans, R. P., Nie, H., Ramos-Onsins, S. E., et al. (2011). Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLoS One* 6:e14782. doi: 10.1371/journal.pone.0014782
- Andersson, L. (2001). Genetic dissection of phenotypic diversity in farm animals. *Nat. Rev. Genet.* 2, 130–138. doi: 10.1038/35052563

## ETHICS STATEMENT

All experimental procedures were performed in accordance with the Institutional Review Board (IRB14044) and the Institutional Animal Care and Use Committee of the Sichuan Agricultural University under permit number DKY-B20140302.

## AUTHOR CONTRIBUTIONS

GT, XL, LZ, YJ, and KW designed the experiments. KW, QY, PW, JZ, and DC performed the data collection. GT, KW, and PW performed the experiments. KW, PW, GT, JM, QT, YJ, AJ, and WX analyzed the data. GT, KW, and PW developed some of the analysis software. KW and GT wrote the manuscript. All authors read and approved the final manuscript.

## FUNDING

The study was supported by grants from the National Natural Science Foundation of China (31530073), the National Natural Science Foundation of China #C170102, the Chinese National Science and Tech Support Program (Nos. 2015BAD03B01 and 2015GA810001), and the earmarked fund for the China Agriculture Research System (No. CARS-36-05B).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00119/full#supplementary-material>

**FIGURE S1** | Fst value for L-Y breed pair on all autosomes. The red dash line corresponds to the 99% threshold on the corresponding empirical distributions.

**FIGURE S2** | CLR value for Landrace pigs on all autosomes. The red dash line corresponds to the 99% threshold on the corresponding empirical distributions.

**FIGURE S3** | CLR value for Yorkshire pigs on all autosomes. The red dash line corresponds to the 99% threshold on the corresponding empirical distributions.

**TABLE S1** | Summary of candidate genes under selection in high Fst for L-Y breed pair.

**TABLE S2** | Summary of candidate genes under selection in high CLR in Landrace.

**TABLE S3** | Summary of candidate genes under selection in high CLR in Yorkshire.

**TABLE S4** | Enrichment of signaling pathways of the candidate genes by KEGG.

- Andersson, L., and Georges, M. (2004). Domestic-animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5, 202–212. doi: 10.1038/nrg1294
- Andersson, L., and Plastow, G. (2011). “Molecular genetics of coat colour variation,” in *The Genetics of the Pig*, eds M. F. Rothschild and A. Ruvinsky (Wallingford: CAB International).
- Bevan, P. (2001). Insulin signalling. *J. Cell Sci.* 114(Pt 8), 1429–1430.
- Bionaz, M., and Looor, J. J. (2008). Gene networks driving bovine milk fat synthesis during the lactation cycle. *BMC Genomics* 9:366. doi: 10.1186/1471-2164-9-366

- Boehm, U., Klamp, T., Groot, M., and Howard, J. C. (1997). Cellular responses to interferon-gamma. *Annu. Rev. Immunol.* 15, 749–795. doi: 10.1146/annurev.immunol.15.1.749
- Brown, J., Delaine, C., Zaccheo, O. J., Siebold, C., Gilbert, R. J., Boxel, G. V., et al. (2008). Structure and functional analysis of the IGF-II/IGF2R interaction. *EMBO J.* 27, 265–276. doi: 10.1038/sj.emboj.7601938
- Browning, S. R., and Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097. doi: 10.1086/521987
- Chen, Z., Gibson, T. B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., et al. (2001). Map kinases. *Chem. Rev.* 101, 2449–2476. doi: 10.1021/cr000241p
- Cross, M. J., Dixelius, J., Matsumoto, T., and Claessonwelsh, L. (2003). Vegf-receptor signal transduction. *Trends Biochem. Sci.* 28, 488–494. doi: 10.1016/S0968-0004(03)00193-2
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., Depristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Dauncey, M. J., Burton, K. A., White, P., Harrison, A. P., Gilmour, R. S., Duchamp, C., et al. (1994). Nutritional regulation of growth hormone receptor gene expression. *FASEB J.* 8, 81–88. doi: 10.1096/fasebj.8.1.7507871
- Depristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498. doi: 10.1038/ng.806
- Duerr, R. H., Taylor, K. D., Brant, S. R., Rioux, J. D., Silverberg, M. S., Daly, M. J., et al. (2006). A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314, 1461–1463. doi: 10.1126/science.1135245
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379
- Engelman, J. A., Luo, J., and Cantley, L. C. (2006). The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nat. Rev. Genet.* 7, 606–619. doi: 10.1038/nrg1879
- Fimia, G. M., and Sassone-Corsi, P. (2001). Cyclic AMP signalling. *J. Cell Sci.* 114, 1971–1972.
- Frantz, L. A., Schraiber, J. G., Madsen, O., Megens, H. J., Cagan, A., Bosse, M., et al. (2015). Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat. Genet.* 47, 1141–1148. doi: 10.1038/ng.3394
- Fu, Y., Li, C., Tang, Q., Tian, S., Jin, L., Chen, J., et al. (2016). Genomic analysis reveals selection in Chinese native black pig. *Sci. Rep.* 6:36354. doi: 10.1038/srep36354
- Grossman, S. R., Shlyakhter, I., Karlsson, E. K., Byrne, E. H., Morales, S., Frieden, G., et al. (2010). A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* 327, 883–886. doi: 10.1126/science.1183863
- Guertin, D. A., and Sabatini, D. M. (2007). Defining the role of mTOR in cancer. *Cancer Cell* 12, 9–22. doi: 10.1016/j.ccr.2007.05.008
- Hammond, K. (1998). “Genetic resources and the global programme for their management,” in *Genetics of the Pig*, eds M. F. Rothschild and A. Ruvinsky (Wallingford: CAB International).
- Harrison, D. A. (2012). The Jak/STAT pathway. *Cold Spring Harb. Perspect. Biol.* 4, 829–841. doi: 10.1101/cshperspect.a011205
- Hay, N., and Sonenberg, N. (2004). Upstream and downstream of mTOR. *Genes Dev.* 18, 1926–1945. doi: 10.1101/gad.1212704
- Heimer, G., Kerätär, J. M., Riley, L. G., Balasubramaniam, S., Eyal, E., Pietikäinen, L. P., et al. (2016). MECP mutations cause childhood-onset dystonia and optic atrophy, a mitochondrial fatty acid synthesis disorder. *Am. J. Hum. Genet.* 99, 1229–1244. doi: 10.1016/j.ajhg.2016.09.021
- Hudson, R. R., Slatkin, M., and Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* 132, 583–589.
- Huelsken, J., and Behrens, J. (2002). The Wnt signalling pathway. *J. Cell Sci.* 115(Pt 21), 3977–3978. doi: 10.1242/jcs.00089
- Imada, K., and Leonard, W. J. (2000). The Jak-STAT pathway. *Mol. Immunol.* 37, 1–11. doi: 10.1016/S0161-5890(00)00018-3
- Kim, K. S., Larsen, N., Short, T., Plastow, G., and Rothschild, M. F. (2000). A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm. Genome* 11, 131–135. doi: 10.1007/s003350010025
- Kiu, H., and Nicholson, S. E. (2012). Biology and significance of the JAK/STAT signalling pathways. *Growth Factors* 30, 88–106. doi: 10.3109/08977194.2012.660936
- Kopchick, J. J., and Andry, J. M. (2000). Growth hormone (GH), GH receptor, and signal transduction. *Mol. Genet. Metab.* 71, 293–314. doi: 10.1006/mgme.2000.3068
- Laere, A. S. V., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., et al. (2003). A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 425, 832–836. doi: 10.1038/nature02064
- Laplante, M., and Sabatini, D. M. (2012). mTOR signaling in growth control and disease. *Cell* 149, 274–293. doi: 10.1016/j.cell.2012.03.017
- Larson, G., Albarella, U., Dobney, K., Rowleyconwy, P., Schibler, J., Tresset, A., et al. (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15276–15281. doi: 10.1073/pnas.0703411104
- Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoosmith, E., Robins, J., et al. (2005). Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* 307, 1618–1621. doi: 10.1073/pnas.0703411104
- Lee, S. H., and Dominguez, R. (2010). Regulation of actin cytoskeleton dynamics in cells. *Mol. Cells* 29, 311–325. doi: 10.1007/s10059-010-0053-8
- Leung, D. W., Spencer, S. A., Cachianes, G., Hammonds, R. G., Collins, C., Henzel, W. J., et al. (1987). Growth hormone receptor and serum binding protein: purification, cloning and expression. *Nature* 330, 537–543. doi: 10.1038/330537a0
- Levillain, O., Ventura, G., Déchaud, H., Hobeika, M., Meseguer, A., Moinard, C., et al. (2007). Sex-differential expression of ornithine aminotransferase in the mouse kidney. *Am. J. Physiol. Renal Physiol.* 292, F1016–F1027. doi: 10.1152/ajprenal.00408.2006
- Li, H., and Durbin, R. (2009). *Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform*. Oxford: Oxford University Press.
- Li, M., Tian, S., Jin, L., Zhou, G., Li, Y., Zhang, Y., et al. (2013). Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nat. Genet.* 45, 1431–1438. doi: 10.1038/ng.2811
- Li, Y., Lacerda, D. A., Warman, M. L., Beier, D. R., Yoshioka, H., Ninomiya, Y., et al. (1995). A fibrillar collagen gene, Col11a1, is essential for skeletal morphogenesis. *Cell* 80, 423–430. doi: 10.1016/0092-8674(95)90492-1
- Liu, G., Jennen, D. G. J., Tholen, E., Juengst, H., Kleinwächter, T., Hölker, M., et al. (2007). A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc-Pietrain resource population. *Anim. Genet.* 38, 241–252. doi: 10.1111/j.1365-2052.2007.01592.x
- Liu, G., Kim, J. J., Jonas, E., Wimmers, K., Ponsuksili, S., Murani, E., et al. (2008). Combined line-cross and half-sib QTL analysis in Duroc-Pietrain population. *Mamm. Genome* 19, 429–438. doi: 10.1007/s00335-008-9132-y
- Meijer, A. J., and Codogno, P. (2004). Regulation and role of autophagy in mammalian cells. *Int. J. Biochem. Cell Biol.* 36, 2445–2462. doi: 10.1016/j.biocel.2004.02.002
- Nair, R. R., Kerätär, J. M., Autio, K. J., Masud, A. J., Finnilä, M. A., Autio-Harmanen, H. I., et al. (2017). Genetic modifications of Mecr reveal a role for mitochondrial 2-enoyl-CoA/ACP reductase in placental development in mice. *Hum. Mol. Genet.* 26, 2104–2117. doi: 10.1093/hmg/ddx105
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218. doi: 10.1146/annurev.genet.39.073003.112420
- Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., and Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. *Genome Res.* 15, 1566–1575. doi: 10.1101/gr.4252305
- Ovilo, C., Fernández, A., Rodríguez, M. C., Nieto, M., and Silió, L. (2006). Association of MC4R gene variants with growth, fatness, carcass composition and meat and fat quality traits in heavy pigs. *Meat Sci.* 73, 42–47. doi: 10.1016/j.meatsci.2005.10.016
- Patterson, N., Price, A. L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet.* 2:e190. doi: 10.1371/journal.pgen.0020190
- Pavlidis, P., Zivkovic, D., Stamatakis, A., and Alachiotis, N. (2013). SweeD: likelihood-based detection of selective sweeps in thousands of genomes. *Mol. Biol. Evol.* 30, 2224–2234. doi: 10.1093/molbev/mst112



- Pfeifer, B., Wittelsbürger, U., Ramosonsins, S. E., and Lercher, M. J. (2015). PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* 31, 1929–1936. doi: 10.1093/molbev/msu136
- Prentki, M., and Madiraju, S. R. (2008). Glycerolipid metabolism and signaling in health and disease. *Endocr. Rev.* 29, 647–676. doi: 10.1210/er.2008-0007
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Quintanilla, R., Pena, R. N., Gallardo, D., Cánovas, A., Ramírez, O., Díaz, I., et al. (2011). Porcine intramuscular fat content and composition are regulated by quantitative trait loci with muscle-specific effects. *J. Anim. Sci.* 89, 2963–2971. doi: 10.2527/jas.2011-3974
- Rubin, C. J., Megens, H. J., Barrio, A. M., Maqbool, K., Sayyab, S., and Schwochow, D. (2012). Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19529–19536. doi: 10.1073/pnas.1217149109
- Saltiel, A. R., and Kahn, C. R. (2001). Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 414, 799–806. doi: 10.1038/414799a
- Schoebelen-Combes, S., Louveau, I., Postel-Vinay, M. C., and Bonneau, M. (1996). Ontogeny of GH receptor and GH-binding protein in the pig. *J. Endocrinol.* 148, 249–255. doi: 10.1677/joe.0.1480249
- Schroder, K., Hertzog, P. J., Ravasi, T., and Hume, D. A. (2004). Interferon-gamma: an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* 75, 163–189. doi: 10.1189/jlb.0603252
- Smith, J. M., and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet. Res.* 23, 23–35. doi: 10.1017/S0016672300014634
- Song, G., Ouyang, G., and Bao, S. (2005). The activation of AKT/PKB signaling pathway and cell survival. *J. Cell Mol. Med.* 9, 59–71. doi: 10.1111/j.1582-4934.2005.tb00337.x
- Stachowiak, M., Nowacka-Wozuk, J., Szydlowski, M., and Switonski, M. (2013). The ACACA and SREBF1 genes are promising markers for pig carcass and performance traits, but not for fatty acid content in the longissimus dorsi muscle and adipose tissue. *Meat Sci.* 95, 64–71. doi: 10.1016/j.meatsci.2013.04.021
- Steinberg, G. R., and Kemp, B. E. (2009). AMPK in health and disease. *Physiol. Rev.* 89, 1025–1078. doi: 10.1152/physrev.00011.2008
- Teves, M. E., Zhang, Z., Costanzo, R. M., Henderson, S. C., Corwin, F. D., Zweit, J., et al. (2013). Spag17 is essential for motile cilia function and neonatal survival. *Am. J. Respir. Cell Mol. Biol.* 48, 765–772. doi: 10.1165/rcmb.2012-0362OC
- Uimari, P., Sironen, A., and Sevónaimonen, M. L. (2013). Evidence for three highly significant QTL for meat quality traits in the Finnish Yorkshire pig breed. *J. Anim. Sci.* 91, 2001–2011. doi: 10.2527/jas.2012-5811
- Vidal, O., Noguera, J. L., Amills, M., Varona, L., Gil, M., Jiménez, N., et al. (2005). Identification of carcass and meat quality quantitative trait loci in a Landrace pig population selected for growth and leanness. *J. Anim. Sci.* 83, 293–300. doi: 10.2527/2005.832293x
- Wakil, S. J., and Abuelheiga, L. A. (2009). Fatty acid metabolism: target for metabolic syndrome. *J. Lipid Res.* 50(Suppl.), S138–S143. doi: 10.1194/jlr.R800079-JLR200
- Wang, Y., Price, S. E., and Jiang, H. (2003). Cloning and characterization of the bovine class 1 and class 2 insulin-like growth factor-I mRNAs. *Domest. Anim. Endocrinol.* 25, 315–328. doi: 10.1016/j.domaniend.2003.06.001
- Weir, B. S., Cardon, L. R., Anderson, A. D., Nielsen, D. M., and Hill, W. G. (2005). Measures of human population structure show heterogeneity among genomic regions. *Genome Res.* 15, 1468–1476. doi: 10.1101/gr.4398405
- White, M. F., and Kahn, C. R. (1994). The insulin signaling system. *J. Biol. Chem.* 269, 1–4.
- Wilkinson, S., Lu, Z. H., Megens, H. J., Archibald, A. L., Haley, C., Jackson, I. J., et al. (2013). Signatures of diversifying selection in European pig breeds. *PLoS Genet.* 9:e1003453. doi: 10.1371/journal.pgen.1003453
- Williamson, S. H., Hubisz, M. J., Clark, A. G., Payseur, B. A., Bustamante, C. D., and Nielsen, R. (2007). Localizing recent adaptive evolution in the human genome. *PLoS Genet.* 3:e90. doi: 10.1371/journal.pgen.0030090
- Wright, S. (1931). Evolution in mendelian populations. *Genetics* 16, 97–159.
- Yang, S., Li, X., Li, K., Fan, B., and Tang, Z. (2014). A genome-wide scan for signatures of selection in Chinese indigenous and commercial pig breeds. *BMC Genet.* 15:7. doi: 10.1186/1471-2156-15-7
- Yang, Y. (2010). Wnts and wing: Wnt signaling in vertebrate limb development and musculoskeletal morphogenesis. *Birth Defects Res. C Embryo Today Rev.* 69, 305–317. doi: 10.1002/bdrc.10026

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Wang, Wu, Yang, Chen, Zhou, Jiang, Ma, Tang, Xiao, Jiang, Zhu, Li and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.