

# Genomic and transcriptomic analyses reveal selection of genes for puberty in Bama Xiang pigs

Yang Yang<sup>1,2</sup>, Adeniyi C. Adeola<sup>1</sup>, Hai-Bing Xie<sup>1,\*</sup>, Ya-Ping Zhang<sup>1,2,\*</sup>

<sup>1</sup> State Key Laboratory of Genetic Resources and Evolution, Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming Yunnan 650204, China

<sup>2</sup> Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming Yunnan 650204, China

## ABSTRACT

The Bama Xiang pig (BMX) is a famous early-maturing Chinese indigenous breed with a two-end black coat. To uncover the genetic basis of the BMX phenotype, we conducted comparative genomic analyses between BMX and East Asian wild boars and Laiwu pigs, respectively. Genes under positive selection were enriched in pathways associated with gonadal hormone and melanin synthesis, consistent with the phenotypic changes observed during development in BMX pigs. We also performed differentially expressed gene analysis based on RNA-seq data from pituitary tissues of BMX and Large White pigs. The *CTTNBP2NL*, *FRS2*, *KANK4*, and *KATNAL1* genes were under selection and exhibited expressional changes in the pituitary tissue, which may affect BMX pig puberty. Our study demonstrated the positive selection of early maturity in the development of BMX pigs and advances our knowledge on the role of regulatory elements in puberty evolution in pigs.

**Keywords:** Puberty; Bama Xiang pig; Pituitary; Differentially expressed genes

## INTRODUCTION

Pigs are important protein food source and globally distributed animal domesticated from Eurasian wild boars 9 000 years ago (Larson et al., 2005). During their long domestication history, pigs have exhibited abundant phenotypic variation in coat color, reproduction, and growth. This phenotypic diversity provides valuable animal models for genetic studies (Ciobanu et al., 2001; Rubin et al., 2012).

The Bama Xiang pig (BMX) from Guangxi Zhuang Autonomous Region is a well-known variety noted for its early maturity and two-end black pigmentation. The female BMX pig matures at about three months, in sharp contrast to the

five months observed in most other East Asian domestic pigs (China National Commission of Animal Genetic Resources, 2011). Thus, the BMX pig provides an important model to study early maturity and pigmentation evolution during pig domestication.

In this study, we performed genomic screening of signatures of positive selection in the BMX pig in comparison with those of the East Asian wild boar (EAW) and Laiwu (LWU) pigs. The LWU pig has a normal maturity process (at six months) and black coat color (China National Commission of Animal Genetic Resources, 2011), but shares the same domestication origin as that of the BMX pig from East Asian wild boars (Wu et al., 2007). In combination with transcriptomic analysis, we identified genetic variations that were putatively associated with the maturity and pigment phenotype in the BMX pig. The primary goal of this study was to explore the (1) genomic evolutionary pattern shaped by development of the BMX pig, and (2) genetic variations associated with selection for early maturity and coat color in the BMX pig.

## MATERIALS AND METHODS

### Sample collection and resequencing of genomic DNA

We downloaded the genome resequencing data of six BMX, six LWU, and 13 EAW pigs from the NCBI database (Supplementary Table S1), as reported from earlier studies

Received: 30 March 2018; Accepted: 10 April 2018; Online: 19 June 2018

Foundation items: This work was supported by grants from the National Natural Science Foundation of China (31472000), National 973 Program of China (2013CB835203), and Animal Branch of the Germplasm Bank of Wild Species, Chinese Academy of Sciences (Large Research Infrastructure Funding)

\*Corresponding authors, E-mail: xiehb@mail.kiz.ac.cn; zhangyp@mail.kiz.ac.cn

DOI: 10.24272/j.issn.2095-8137.2018.068

(Ai et al., 2015; Groenonet al., 2012; Li et al., 2013; Rubin et al., 2012). Sequence reads were filtered by removing adaptors and low-quality bases using Trimmomatic (Bolger et al., 2014). Genomic reads were aligned to the *Sus scrofa* 10.2 genome (Groenonet al., 2012) with the BWA program (Li & Durbin, 2010). Alignment was performed with a maximum of six mismatches allowed for each 100-bp fragment.

### Single nucleotide polymorphism (SNP) calling and annotation

SNP calling was performed with the DNA sequencing data of the 25 pig genomes using HaplotypeCaller in the GATK toolkit (McKenna et al., 2010). Only polymorphic SNPs were included in this analysis. VariantFiltration and SelectVariants were used to filter SNPs by quality score (QUAL <20) and depth of coverage (DP <10). SNPs were annotated with the GTF file downloaded from the NCBI website (<https://www.ncbi.nlm.nih.gov/>). For annotation, SNPs were classified by their locations in protein-coding (synonymous or non-synonymous), intronic, intergenic, and untranslated region (UTR) sequences.

### Selective sweep detection

Population fixation index ( $F_{ST}$ ) (Weir & Cockerham, 1984), cross population extended haplotype homozygosity (XPEHH) (Sabeti et al., 2007), and nucleotide diversity ( $\pi$ ) (Tajima, 1983) were used to search for signals of selection in the BMX pig genome. We analyzed selective sweeps in 10-kb non-overlapping sliding windows. Windows containing <5 SNPs were removed from the analysis.

The  $F_{ST}$  values between BMX and other pigs (LWU and EAW) were calculated using VCFtools (Danecek et al., 2011).

Haplotypes were inferred using Beagle software (Browning & Browning, 2007). Comparing the BMX pigs with LWU and EAW pigs, XPEHH was calculated to examine haplotypes that showed low levels of linkage decay in the BMX populations (Coop et al., 2009; Pickrell et al., 2009; Pritchard et al., 2010).

The nucleotide diversities of BMX ( $\pi_{BMX}$ ), EAW ( $\pi_{EAW}$ ), and LWU ( $\pi_{LWU}$ ) pigs were then calculated separately using VCFtools (Danecek et al., 2011). To identify regions that showed low nucleotide diversity in the BMX pigs compared to the LWU and EAW pigs, we calculated  $\pi_{EAW}$  divided by  $\pi_{BMX}$  and  $\pi_{LWU}$  divided by  $\pi_{BMX}$ , respectively.

Genomic windows showing high  $F_{ST}$ , low nucleotide diversity, and high XPEHH values were identified as having signatures of selection in BMX pigs. Candidate selective sweeps were chosen from genomic windows identified simultaneously by all three statistics as having values above the top 10% threshold in their empirical distributions.

### RNA sample collection, sequencing, and identification of differentially expressed genes

Pituitary tissues were collected from three BMX pigs and three Large White pigs for RNA sequencing (RNA-seq). These pigs were females at 85 days after birth. The RNA-seq libraries with an inserted size of 250 bp were prepared using the Illumina standard RNA-seq library preparation pipeline and sequenced on the Illumina HiSeq 2000 platform, with 150-bp paired-end reads generated. In total, the transcriptomes of

six pituitaries (GSA number: CRA000876) were used for comparative analysis. The RNA-seq clean reads were mapped onto the *Sus scrofa* 10.2 genome using Tophat2 (Kim et al., 2013). A new merged GTF annotation file was generated using Cufflinks and Cuffmerge (Trapnell et al., 2010).

Differentially expressed genes (DEGs) were identified by DESeq2 (Love et al., 2014). As DESeq2 required a raw read count table as input, we obtained the raw read count using HTseq (Anders et al., 2014). An FDR of <0.05 was used as the cutoff for DEGs. The ClueGO plugin in Cytoscape (Mlecnik et al., 2018) was used for gene ontology (GO) enrichment analysis. GO terms with a  $P$ -value of <0.05 were defined as enriched.

## RESULTS

### Selective signatures in BMX pig genomes

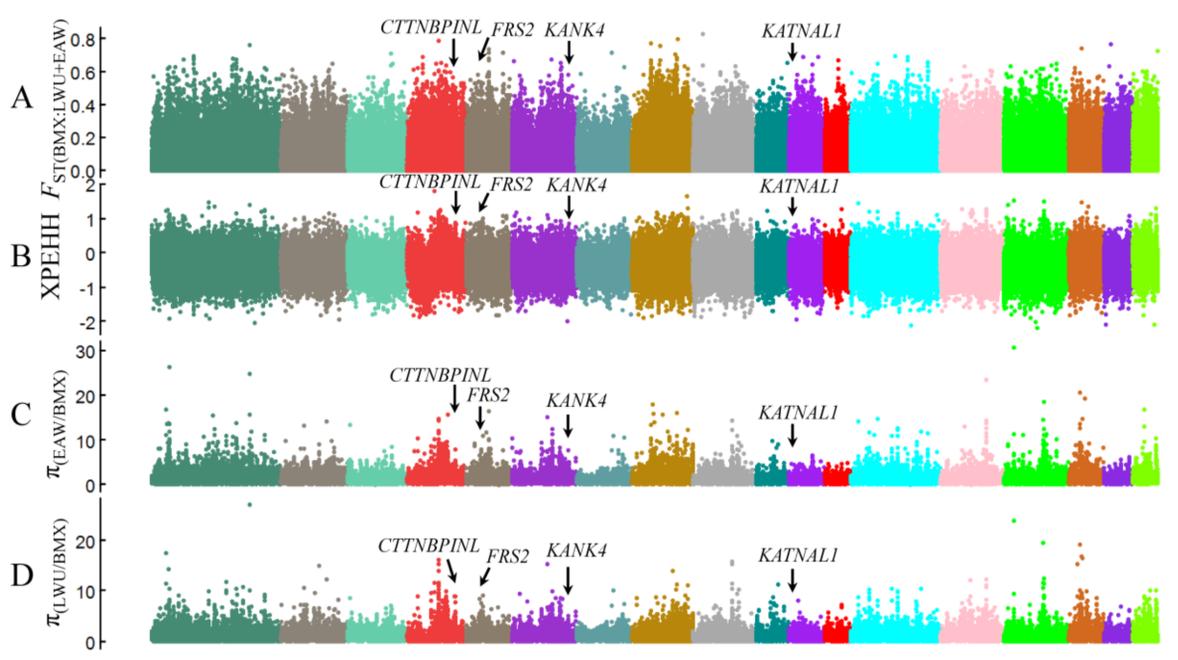
The whole genomes of six BMX, 13 EAW, and six LWU pigs were used to identify genomic variants underlying the development of BMX pigs. Genomic reads were quality filtered and mapped to the pig reference genome (*Sus scrofa* 10.2) (see Materials and Methods). A total of 43.5 million SNPs were identified in the 25 individuals. To identify signatures of selection in BMX pig genomes, we analyzed the  $F_{ST}$  (Weir & Cockerham, 1984),  $\pi$  (Tajima, 1983), and XPEHH (Sabeti et al., 2007) metrics along chromosomes in comparison to those of EAW and LWU pigs (Figure 1). Selective signatures were screened in 10-kb non-overlapping sliding windows and adjacent selective windows were combined. As recent artificial selection produces long stretches of haplotypes in a population (Amaral et al., 2008), a clustering of selective windows with intervals of less than nine windows was chosen as an indicator of a putatively selective sweep. Sweep regions greater than 30 kb in length were included in the following analyses. This procedure resulted in the identification of 257 genes within 339 putative selective sweeps in the BMX pig genomes (Supplementary Table S2).

Functional enrichment analysis of the gene set revealed significant overrepresentation of 59 terms (Table 1). Interestingly, several biological processes related to sexual maturity were involved, including “pituitary gland development” (*BMP4*, *FGF10*, and *TBX19*), “endocrine system development” (*BMP4*, *CRH*, *FGF10*, *NEUROD1*, and *TBX19*), and “negative regulation of reproductive process” (*BMP4*, *CSN2*, and *ZP4*). The pituitary is the most important endocrine gland involved in growth and reproduction regulation. In addition, five genes (*FGF10*, *FGF2*, *MITF*, *PDGFC*, and *PDGFD*) were involved in “Melanoma”, which might be associated with the two-end black pigmentation observed in BMX pigs, which differs from most Chinese domestic pigs with black coats (China National Commission of Animal Genetic Resources, 2011).

Although many missense mutations are associated with the domestication of pigs (Fang et al., 2009; Li et al., 2010; Ren et al., 2011; Rubin et al., 2012), whether regulatory variants play a role and their effects on the domestication processes compared to coding variants remain unknown. To address this issue, we compared the evolutionary patterns for SNPs identified from the selective regions in the BMX pig genomes. A

total of 830 949 SNPs were identified from 339 sweeps. Among these, 3 668 SNPs were located in coding sequences and 827 542 SNPs were located in noncoding sequences. We first focused on protein coding region variations located in the 339 sweeps. We examined highly differentiated missense mutations and analyzed their presence in DNA sequences encoding conserved protein sequences. We concentrated on missense SNPs with  $F_{ST}$  values up to the chromosomal top 0.1% threshold and compared BMX pigs to LWU and EAW

pigs, respectively. We identified a dataset consisting of 11 missense SNPs from 10 genes (Supplementary Table S3). Alignment of the 10 protein sequences showed that the flanking sequences at the residues encoded by the 11 SNPs displayed low level cross-species conservation, implying weak functional constraints on these protein-coding variants. This indicated that regulatory elements rather than missense variants may have greater effects on BMX pig development.



**Figure 1 Whole genome selective signatures of BMX pigs**

X-axis shows the positions of the windows along the 1–18 autosome chromosomes. Different chromosomes are displayed in different colors. Y-axis shows the  $F_{ST}$  (A), XPEHH (B),  $\pi_{EAW/BMX}$  (C), and  $\pi_{LWU/BMX}$  (D) values.  $F_{ST}$  and XPEHH values were calculated between BMX and (LWU+EAW) pigs.  $\pi_{EAW/BMX}$  and  $\pi_{LWU/BMX}$  were obtained from  $\pi_{EAW}$  and  $\pi_{LWU}$  divided by  $\pi_{BMX}$ , respectively.

### Comparative transcriptomic analysis of pituitary tissues

We conducted transcriptomic profiling to examine gene expression changes due to selection in the BMX genomes as regulatory variants exhibited a profound role during BMX pig development. We chose pituitary tissue as the pituitary gland secretes reproduction-related hormones involved in maturity regulation. In our experimental design, we collected pituitary tissues from three BMX and three Large White pigs at 85 days after birth. At this stage, BMX pigs are close to maturity, whereas Large White pigs do not reach maturity until eight months of age (China National Commission of Animal Genetic Resources, 2011). The transcriptomes from the pituitary tissues were sequenced, from which we obtained 35 million mapped reads, on average, for each individual. We compared the gene expression levels in the pituitary and found 1 600 DEGs between BMX and Large White pigs. Although Large White pigs were domesticated from European wild boars (Larson et

al., 2005;Rubin et al., 2012), the DEGs between BMX and Large White pigs provided valuable information on expressional changes in BMX pigs if they also exhibited selective signatures in the BMX genome. The DEGs were enriched in the following biological processes: “Estrogen signaling pathway” (*ADCY6*, *ADCY8*, *CREB3L1*, *CREB5*, *FOS*, *GABBR2*, *GNAS*, *GRM1*, *KCNJ9*, *MMP9*, *PIK3CG*, *PIK3R5*, and *PRKACA*), “Ovarian steroidogenesis” (*ADCY6*, *ADCY8*, *ALOX5*, *GNAS*, *LDLR*, *LHCGR*, and *PRKACA*), “Oocyte meiosis” (*ADCY6*, *ADCY8*, *CAMK2B*, *CAMK2D*, *CCNE1*, *CDC20*, *CPEB2*, *ESPL1*, *MAD2L2*, *MOS*, *PRKACA*, and *YWHAZ*), “Developmental maturation” (*C12H17orf47*, *CDC20*, *EDNRB*, *FAM101B*, *LTF*, *PLP1*, *PRKACA*, *RBPJ*, *TAL1*, *WASH1*, and *WNT10B*), and “Progesterone-mediated oocyte maturation” (*ADCY6*, *ADCY8*, *CPEB2*, *MAD2L2*, *MOS*, *PIK3CG*, *PIK3R5*, and *PRKACA*) (Supplementary Table S4). These pathways all participate in sexual maturity.

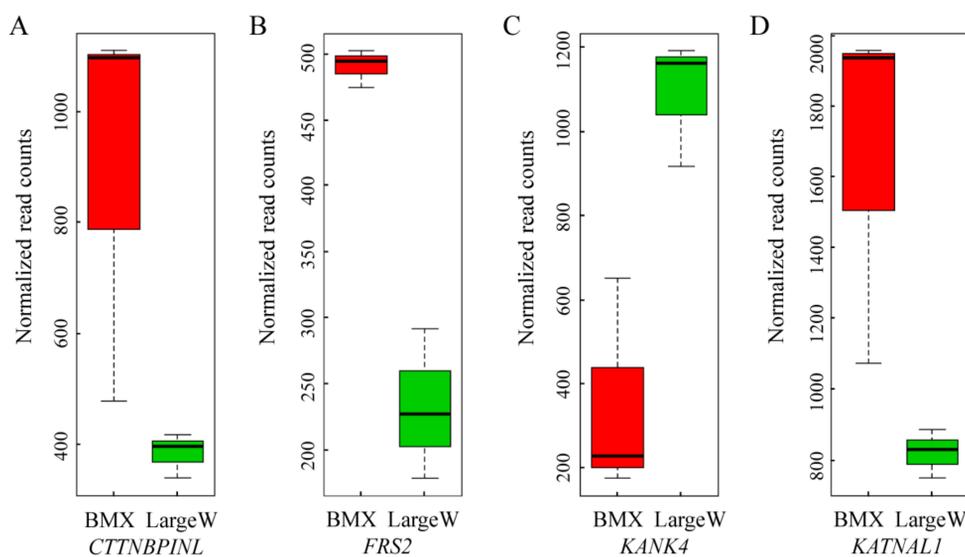
**Table 1 Functional enrichment terms of positively selected genes in BMX pigs.**

GO ID	GO term	P-value*
GO:0004080	Neuroactive ligand-receptor interaction	24.0E-15
GO:0060449	Bud elongation involved in lung branching	100.0E-6
GO:0032740	Positive regulation of interleukin-17 production	140.0E-6
GO:0060393	Regulation of pathway-restricted SMAD protein phosphorylation	230.0E-6
GO:0060389	Pathway-restricted SMAD protein phosphorylation	250.0E-6
GO:0032946	Positive regulation of mononuclear cell proliferation	310.0E-6
GO:0050671	Positive regulation of lymphocyte proliferation	310.0E-6
GO:0070665	Positive regulation of leukocyte proliferation	370.0E-6
GO:0048636	Positive regulation of muscle organ development	400.0E-6
GO:0045844	Positive regulation of striated muscle tissue development	400.0E-6
GO:1901863	positive regulation of muscle tissue development	450.0E-6
GO:0005412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	490.0E-6
GO:0005218	Melanoma	560.0E-6
GO:0045740	Positive regulation of DNA replication	600.0E-6
GO:0009798	Axis specification	600.0E-6
GO:0090090	Negative regulation of canonical Wnt signaling pathway	600.0E-6
GO:0032660	Regulation of interleukin-17 production	660.0E-6
GO:0051155	Positive regulation of striated muscle cell differentiation	660.0E-6
GO:0032620	Interleukin-17 production	780.0E-6
GO:0042506	Tyrosine phosphorylation of Stat5 protein	780.0E-6
GO:0051145	Smooth muscle cell differentiation	800.0E-6
GO:0061037	Negative regulation of cartilage development	920.0E-6
GO:0048546	Digestive tract morphogenesis	950.0E-6
GO:0005410	Hypertrophic cardiomyopathy (HCM)	990.0E-6
GO:0051149	Positive regulation of muscle cell differentiation	1.2E-3
GO:0035270	Endocrine system development	1.3E-3
GO:0055025	Positive regulation of cardiac muscle tissue development	1.3E-3
GO:0042310	Vasoconstriction	1.5E-3
GO:0060602	Branch elongation of an epithelium	1.5E-3
GO:0001709	Cell fate determination	2.0E-3
GO:0021536	Diencephalon development	1.8E-3
GO:0021983	Pituitary gland development	2.2E-3
GO:0042307	Positive regulation of protein import into nucleus	2.3E-3
GO:0000561	Glycerolipid metabolism	2.4E-3
GO:0004730	Long-term depression	2.6E-3
GO:0003338	Metanephros morphogenesis	2.4E-3
GO:0030199	Collagen fibril organization	2.4E-3
GO:0060441	Epithelial tube branching involved in lung morphogenesis	3.0E-3
GO:0030219	Megakaryocyte differentiation	3.3E-3
GO:0042383	Sarcolemma	3.3E-3
GO:0009799	Specification of symmetry	3.3E-3
GO:0051153	Regulation of striated muscle cell differentiation	3.3E-3
GO:0010092	Specification of organ identity	3.3E-3
GO:1900182	Positive regulation of protein localization to nucleus	3.7E-3
GO:0005033	Nicotine addiction	4.0E-3
GO:0004140	Regulation of autophagy	5.9E-3
GO:2000242	Negative regulation of reproductive process	5.9E-3
GO:0070664	Negative regulation of leukocyte proliferation	5.1E-3
GO:0003401	Axis elongation	5.1E-3
GO:0009948	Anterior/posterior axis specification	5.5E-3
GO:0055024	Regulation of cardiac muscle tissue development	5.1E-3
GO:0004940	Type I diabetes mellitus	7.4E-3
GO:0048259	Regulation of receptor-mediated endocytosis	6.9E-3
GO:0007193	Adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway	11.0E-3
GO:0009880	Embryonic pattern specification	11.0E-3
GO:0005160	Transforming growth factor beta receptor binding	9.5E-3
GO:0031623	Receptor internalization	11.0E-3
GO:0048286	Lung alveolus development	9.5E-3
GO:0010862	Positive regulation of pathway-restricted SMAD protein phosphorylation	10.0E-3

\*: P-value obtained from ClueGO plugin in Cytoscape using default parameters. Only terms with  $P < 0.05$  are listed.

As both artificially selected genes and DEGs showed functional enrichment in maturity regulation, it is possible that differential gene expression was favored by the artificial selection of variants from some sweep loci. We therefore identified artificially selected genes with altered gene expression. In our analysis, a total of 22 artificially selected genes exhibited differential gene expression in the pituitary ( $P < 0.05$ ;  $FDR < 0.05$ ; Table 2). Interestingly, four genes (*CTTNBP2NL*, *FRS2*, *KANK4*, and *KATNAL1*) were related to reproduction (Figure 2). Compared to that in the Large White pigs, the expression of the *CTTNBP2NL* gene was increased in BMX pigs. *CTTNBP2NL* has been reported to affect conception rates in bovines (Sugimoto et al., 2013). *FRS2* encodes a

docking protein bound to FGF receptors, with a knockout study in mice indicating its involvement in epididymal development and sperm maturation (Xu et al., 2013). In the pituitary tissue, we found that the expression level of *FRS2* in BMX pigs was about twice that found in Large White pigs. In addition, *KANK4* was found to be differentially expressed in the pituitary tissues of BMX and Large White pigs. Although few studies have reported on the function of *KANK4*, it is located in a significant pig quantitative trait locus (QTL) for age at puberty (Bidanel et al., 2008). *KATNAL1*, which encodes Katanin Catalytic subunit A1 like 1, has been implicated in both meiosis and spermiogenesis (Smith et al., 2012).



**Figure 2** Normalized read counts of significantly differentially expressed genes *CTTNBP2NL* (A), *FRS2* (B), *KANK4* (C), and *KATNAL1* (D) in the pituitary

For each subgraph, red and green boxes represent gene expression in the pituitaries of BMX and Large White pigs, respectively. Upper whisker is maximum value, black horizontal line in box is median, and lower whisker is minimum value.

## DISCUSSION

We conducted a comparative genomic analysis of BMX pigs with domestic pigs and wild boars from East Asia, and revealed selective signatures potentially associated with selection for early maturity and the two-end black phenotype. During domestication, selection is based on those functions and traits that are favored in pig breeds. As described earlier, selective signatures were detected for the specific phenotype of the BMX pig. This study sheds light on the evolution of puberty in pigs and provides important information and candidate genes for pig breeding.

Puberty is an important trait related to economic output, with early maturity resulting in early slow growth. For instance, Western commercial breeds are characterized with late maturity for enlarged body size (Rubin et al., 2012). The

present study revealed four loci (*CTTNBP2NL*, *FRS2*, *KANK4*, and *KATNAL1*) showing strong signatures of selection and functional association with puberty. In addition, the four genes exhibited differential expression between early and late maturity in pig breeds in the pituitary. This result indicates that regulatory elements may have considerable effect on BMX pig development.

Coat color is a trait that is usually under strong selection in different breeds. In previous studies, several genes have been shown to be associated with coat color variants, including *MC1R*, *KIT*, and *EDNRB* (Kijas et al., 1998; Wiener & Wilkinson, 2011; Wilkinson et al., 2013). However, we found no selective signature of *MC1R*. The reason for this may be that selection of *MC1R* occurred during early domestication, and most Chinese indigenous pig breeds share the same haplotype (Li et al., 2010). Earlier research has indicated that mutations in *MITF*,

*PAX3*, *EDNRB*, and *KIT* together may explain a large proportion of white spotting phenotypes in horses (Hauswirth et al., 2012). In BMX pigs, the two-end black phenotype is a complex trait resulting in the diversification of white areas on the body. We hypothesize that a correlation exists between coat color variants and different combinations of candidate genes.

**Table 2** Overlap between positively selected genes and differentially expressed genes from pituitary tissue in BMX pigs.

Gene	Normalized count fold change (BMX/LW)	P-value*
<i>MGAT5B</i>	0.08	7.57E-05
<i>BHLHE22</i>	2.57	0.000 538
<i>FAM174B</i>	0.43	0.000 589
<i>SFT2D2</i>	0.28	0.001 381
<i>TRHDE</i>	2.13	0.005 172
<i>COL23A1</i>	0.12	0.006 991
<i>FRMD4B</i>	2.45	0.007 354
<i>CDKN2B</i>	0.14	0.007 441
<i>FRS2</i>	2.09	0.008 626
<i>ATG4C</i>	2.16	0.010 881
<i>PELI2</i>	1.90	0.012 313
<i>FEZ2</i>	2.12	0.018 321
<i>LGALS12</i>	0.13	0.020 835
<i>TYRP1</i>	0.13	0.025 025
<i>KCND3</i>	1.81	0.030 48
<i>FNDC1</i>	0.22	0.030 808
<i>ATL3</i>	0.46	0.033 021
<i>KANK4</i>	0.34	0.036 521
<i>KATNAL1</i>	1.99	0.037 825
<i>CTTNBP2NL</i>	2.27	0.039 048
<i>ANO3</i>	2.50	0.044 986
<i>DYNC112</i>	1.67	0.047 03

\*: P-values are corrected by Benjamini-Hochberg FDR<0.05.

One limitation of this study was the identification of the DEG direction in RNA-seq analysis. As both BMX and Large White pigs are domesticated, it was difficult to identify the direction of gene expression without the use of an outgroup. We analyzed the overlap between DEGs (between BMX and Large White pigs) and the artificially selected genes in BMX pigs (compared to LWU and EAW pigs). The shared genes were more likely to be changed in BMX pigs. We did not focus on DEGs that did not exhibit selective signatures in the BMX pig genome. In addition, future analysis of Asian and European pig genome data will help clarify breed development in pigs.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTIONS

Y.P.Z., H.B.X., and Y.Y. designed the study. H.B.X. supervised the analyses. Y.Y. performed genome assembly and annotation. Y.Y. extracted RNA and wrote the manuscript with the other authors' input. Y.Y. submitted RNA-seq data to GSA. A.C.A., H.B.X. and Y.P.Z. revised the manuscript. All authors read and approved the final manuscript.

## REFERENCES

- Ai H, Fang X, Yang B, Huang Z, Chen H, Mao L, Zhang F, Zhang L, Cui L, He W, Yang J, Yao X, Zhou L, Han L, Li J, Sun S, Xie X, Lai B, Su Y, Lu Y, Yang H, Huang T, Deng W, Nielsen R, Ren J, Huang L. 2015. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nature Genetics*, **47**(3): 217–225.
- Amaral AJ, Megens HJ, Crooijmans RPMA, Heuven HCM, Groenen MAM. 2008. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics*, **179**(1): 569–579.
- Anders S, Pyl PT, Huber W. 2014. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**(2): 166–169.
- Bidanel JP, Rosendo A, Iannuccelli N, Riquet J, Gilbert H, Caritez JC, Billon Y, Amigues Y, Prunier A, Milan D. 2008. Detection of quantitative trait loci for teat number and female reproductive traits in Meishan x Large White F2 pigs. *Animal*, **2**(6): 813–820.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15): 2114–2120.
- Browning SR, Browning BL. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics*, **81**(5): 1084–1097.
- China National Commission of Animal Genetic Resources. 2011. Animal Genetic Resources in China: Pigs. Beijing: China Agriculture Press. (in Chinese)
- Ciobanu DC, Day AE, Nagy A, Wales R, Rothschild MF, Plasto GS. 2001. Genetic variation in two conserved local Romanian pig breeds using type 1 DNA markers. *Genetics Selection Evolution*, **33**(4): 417–432.
- Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, Myers RM, Cavalli-Sforza LL, Feldman MW, Pritchard JK. 2009. The role of geography in human adaptation. *PLoS Genetics*, **5**(6): e1000500.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156–2158.
- Fang M, Larson G, Ribeiro HS, Li N, Andersson L. 2009. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genetic*, **5**(1): e1000341.
- Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, Li S, Larkin DM, Kim H, Frantz LA, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi SH, Chow W, Clark RC, Clee C, Crooijmans RP, Dawson HD, Dehais P, De Sapio F, Dibbits B, Drou N, Du ZQ, Eversole K, Fadista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JG, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu ZL, Humphray SJ, Hunt T, Hornshøj H, Jeon JT, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim JH, Kim KW, Kim TH, Larson G, Lee K, Lee KT, Leggett R, Lewin HA, Li Y, Liu W, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L, McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni P, Oh SJ, Onteru S, Panitz F, Park

- EW, Park HS, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang Y, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang G, Zhang J, Zhang J, Zhao S, Rogers J, Churcher C, Schook LB. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, **491**(7424): 393–398.
- Hauswirth R, Haase B, Blatter M, Brooks SA, Burger D, Drögemüller C, Gerber V, Henke D, Janda J, Jude R, Magdesian KG, Matthews JM, Poncet PA; Svansson V, Tozaki T, Wilkinson-White L, Penedo MCT, Rieder S, Leeb T. 2012. Mutations in MITF and PAX3 cause “splashed white” and other white spotting phenotypes in horses. *PLoS Genetics*, **8**(4): e1002653.
- Kijas JMH, Wales R, Törnsten A, Chardon P, Moller M, Andersson L. 1998. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics*, **150**(3): 1177–1185.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, **14**(4): R36.
- Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, Robins J, Lowden S, Finlayson H, Brand T, Willerslev E, Rowley-Conwy P, Andersson L, Cooper A. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, **307**(5715): 1618–1621.
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**(5): 589–595.
- Li J, Yang H, Li JR, Li HP, Ning T, Pan XR, Shi P, Zhang YP. 2010. Artificial selection of the melanocortin receptor 1 gene in Chinese domestic pigs during domestication. *Heredity*, **105**(3): 274–281.
- Li M, Tian S, Jin L, Zhou G, Li Y, Zhang Y, Wang T, Yeung CKL, Chen L, Ma J, Zhang J, Jiang A, Li J, Zhou C, Zhang J, Liu Y, Sun X, Zhao H, Niu Z, Lou P, Xian L, Shen X, Liu S, Zhang S, Zhang M, Zhu L, Shuai S, Bai L, Tang G, Liu H, Jiang Y, Mai M, Xiao J, Wang X, Zhou Q, Wang Z, Stothard P, Xue M, Gao X, Luo Z, Gu Y, Zhu H, Hu X, Zhao Y, Plastow GS, Wang J, Jiang Z, Li K, Li N, Li X, Li R. 2013. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nature Genetics*, **45**(12): 1431–1438.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**(12): 550.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297–1303.
- Mlecnik B, Galon J, Bindea G. 2018. Comprehensive functional analysis of large lists of genes and proteins. *Journal of Proteomics*, **171**: 2–10. DOI: doi.org/10.1016/j.jprot.2017.03.016.
- Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, Srinivasan BS, Barsh GS, Myers RM, Feldman MW, Pritchard JK. 2009. Signals of recent positive selection in a worldwide sample of human populations. *Genome Research*, **19**(5): 826–837.
- Pritchard JK, Pickrell JK, Coop G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**(4): R208–R215.
- Ren J, Duan Y, Qiao R, Yao F, Zhang Z, Yang B, Guo Y, Xiao S, Wei R, Ouyang Z, Ding N, Ai H, Huang L. 2011. A missense mutation in PPARD causes a major QTL effect on ear size in pigs. *PLoS Genetics*, **7**(5): e1002043.
- Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, Archibald AL, Fredholm M, Groenen MAM, 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**(48): 19529–19536.
- Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, The International HapMap Consortium. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature*, **449**(7164): 913–918.
- Smith LB, Milne L, Nelson N, Eddie S, Brown P, Atanassova N, O'Bryan MK, O'Donnell L, Rhodes D, Wells S, Napper D, Nolan P, Lalanze Z, Cheeseman M, Peters J. 2012. KATNAL1 regulation of sertoli cell microtubule dynamics is essential for spermiogenesis and male fertility. *PLoS Genetics*, **8**(5): e1002697.
- Sugimoto M, Sasaki S, Gotoh Y, Nakamura Y, Aoyagi Y, Kawahara T, Sugimoto Y. 2013. Genetic variants related to gap junctions and hormone secretion influence conception rates in cows. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(48): 19495–19500.
- Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**(2): 437–460.
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, **28**(5): 511–515.
- Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**(6): 1358–1370.
- Wiener P, Wilkinson S. 2011. Deciphering the genetic basis of animal domestication. *Proceedings of the Royal Society B: Biological Sciences*, **278**(1722): 3161–3170.
- Wilkinson S, Lu ZH, Megens HJ, Archibald AL, Haley C, Jackson IJ, Groenen MAM, Crooijmans RPMA, Ogden R, Wiener P. 2013. Signatures of diversifying selection in European pig breeds. *PLoS Genetics*, **9**(4): e1003453.
- Wu GS, Yao YG, Qu KX, Ding ZL, Li H, Palanichamy MG, Duan ZY, Li N, Chen YS, Zhang YP. 2007. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. *Genome Biology*, **8**(11): R245.
- Xu B, Yang L, Hinton BT. 2013. The role of fibroblast growth factor receptor substrate 2 (FRS2) in the regulation of two activity levels of the components of the extracellular signal-regulated kinase (ERK) pathway in the mouse epididymis. *Biology of Reproduction*, **89**(2): 48.