

Evolutionary insights into porcine genomic structural variations based on a novel-constructed dataset from 24 worldwide diverse populations

Huanfa Gong^{1,2,3} | Weiwei Liu¹ | Zhongzi Wu¹ | Mingpeng Zhang¹ | Yingchun Sun¹ | Ziqi Ling¹ | Shijun Xiao¹ | Huashui Ai¹ | Yuyun Xing¹ | Bin Yang¹ | Lusheng Huang¹

¹State Key Laboratory of Pig Genetic Improvement and Production Technology, Jiangxi Agricultural University, Nanchang, China

²Key Laboratory of Molecular Animal Nutrition, Ministry of Education, College of Animal Sciences, Zhejiang University, Hangzhou, China

³Key Laboratory of Animal Nutrition and Feed Science in Eastern China, Ministry of Agriculture, College of Animal Sciences, Zhejiang University, Hangzhou, China

Correspondence

Bin Yang and Lusheng Huang, State Key Laboratory of Pig Genetic Improvement and Production Technology, Jiangxi Agricultural University, Nanchang 330045, China.
 Email: binyang@live.cn and lushenghuang@hotmail.com

Funding information

National Natural Science Foundation of China, Grant/Award Number: 31760657

Abstract

Structural variations (SVs) are important DNA polymorphisms that contribute to genetic diversity and evolution in humans, animals, and plants. In this study, we present a novel swine SV dataset of 79,919 deletions, 23,638 duplications, and 9333 inversions with average sequence depths of 24.1x from 24 varieties of worldwide pig populations, encompassing 305 individuals. Genotypes of SVs, particularly deletions, can accurately group individuals based on their population identity. We showed that exon-covering deletions were subject to negative selection. Fixation index and differential allele frequency analysis identified highly differentiated SVs between European and Asian indigenous pigs, including deletions in NR6A1 and PLAG1, which are significantly associated with vertebrate numbers and growth performances, respectively. The growth-enhancing allele at the deletion in PLAG1 was shared by European commercial pigs and Northern Chinese indigenous pigs including Laiwu and Min pigs, suggesting potential introgression from European commercial breeds into Chinese indigenous breeds. Moreover, we uncovered highly differentiated SVs in 139 genes between domesticated pigs and wild boars in Asia, temperature and altitude adaptation-associated SVs in 41 genes, and population-specific SVs in 718 genes. This study provides novel insights into the role of porcine SVs in domestication, environmental adaptation, and breed formation.

KEY WORDS

adaptation, breed formation, domestication, pig, population genetics, structural variation

Huanfa Gong and Weiwei Liu contributed equally to this study.

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Structural variations (SVs) characterized by large genomic alterations normally have lengths of >50 bp and can be categorized as balanced (inversion and translocation) and unbalanced (insertion, duplication, and deletion) variants. SVs account for the larger proportion of the genome and are demonstrated to have greater effects on gene expression than single-nucleotide polymorphisms (SNPs) (Alkan et al., 2011; Chaisson et al., 2019; Chiang et al., 2017). They also have a significant impact on human disease (Weischenfeldt et al., 2013) and on breed-defining phenotypes, such as coat color, in domesticated animal species (Georges et al., 2019). Moreover, accumulating evidence suggests that SVs are subject to selection in the evolutionary process of various species such as human (Almarri et al., 2020), dog (Wang et al., 2019), rice (Kou et al., 2020), and grapevine (Lucek et al., 2019; Wellenreuther et al., 2019; Yan et al., 2021; Zhou et al., 2019).

The pig (*Sus scrofa*) was domesticated independently in Europe (Near East) and Asia (China) approximately 10,000 years ago (Frantz et al., 2015). Modern domesticated pigs in Europe and Asia display divergent phenotypic characteristics owing to their independent domestication and artificial selection history. The Chinese indigenous pig breeds account for about one-third of the total number of breeds globally and have a higher overall genetic diversity than those of European pigs (Groenen et al., 2012). Many of the Chinese indigenous breeds display breed-defining characteristics such as high intramuscular fat content of Laiwu pigs and two-end black coat color in Bamaxiang pigs and have adapted to various environmental conditions such as extreme altitudes and temperatures. European pigs have undergone distinct evolutionary pressures than from Chinese indigenous pigs; for example, they are traditionally grazed outdoors, whereas Chinese indigenous pigs often stay indoors for most of their lifetime. After industrialization, European commercial pigs were intensively selected for their large body size, growth rate, low fatness content, and especially for their white coat colors (White, 2011; Yang et al., 2017). A number of pig breeds were also cultivated in America, such as Yucatan pigs, which were imported to the United States in 1960 from the arid Yucatan peninsula, and well-known for small body size, hairless, gentleness, and resistance to disease (Kim et al., 2015). Therefore, the different varieties of pig breeds found in China and Europe provide exceptional resources for investigating the genetic basis of their evolutionary process.

Currently, the majority of population genetic studies on pigs have focused on SNPs (Cho et al., 2019; Frantz et al., 2015). Recently, advances in genome sequencing and the use of detection algorithms in studies on humans have enabled the investigation of structural variations in diverse pig populations (Almarri et al., 2020; Chiang et al., 2017; Larson et al., 2019; Rubin et al., 2012). Progress has been made in identifying positive selected SVs that are associated with breed-specific traits (Du et al., 2021; Zhou et al., 2021). However, the genetic diversity and evolution of SVs, and their associations with environmental adaption and breed-specific phenotypes remain to be fully addressed.

In this study, we analyzed whole genome sequence data obtained from 305 individual pigs representing 24 worldwide populations and an outgroup, *Sus verrucosus* with a minimum sequence depth of 10 and average sequence depth of 24.1. By implementing a pipeline taking into account the heterogeneity in reading lengths and insert sizes in the data (Larson et al., 2019; Layer et al., 2014), we characterized the genetic diversity and evolutionary history of SVs. We identified highly stratified SVs between Asian (Chinese) domesticated breeds and European commercial breeds, and the sharing of derived alleles supporting the introgression/intercrossing history of Asian and European indigenous pigs. We also identified highly differentiated SVs between indigenous pigs and wild boars in Asia. Moreover, SVs that are potentially associated with local environmental adaptation in China and population-private SVs that are potentially associated with breed-defining characteristics were also identified.

2 | MATERIALS AND METHODS

2.1 | Samples collection and quality control

In this study, sequence data of 462 pigs were initially investigated. Adapter and low-quality reads (quality score < 20) were removed using fastp software with default parameters (Chen et al., 2018). Data obtained from 172 pigs representing breeds such as JH, BMX, LWH, GST, SCT, YNT, TT, LR, PI, DU, and LW were generated by our laboratory, and the remaining data were downloaded from public databases (Ai et al., 2015; Frantz et al., 2015; János et al., 2014; Kim et al., 2015; Ramírez et al., 2014) (Table 1; Table S2). Out of the 462 pigs, 305 were used for SV analysis after filtering out 157 individuals with sequence depth < 10 and genomic coverage < 0.96 (Table S2). The heterogeneous pig population was crossed with eight distinct phenotype breeds for the founder (F0) from European/American: Duroc, Landrace, Large white, Pietrain pig; Asian (China): Erhuanian, Bamaxiang, Laiwu, Tibetan pig. The first generation (F1) was generated by each Chinese breed (boars or sows) mated with one European breed (sows or boars) and finally mix the genetic ancestry through a rotating mating until the F6 generation in our study. The intercrossing strategy in detail was described in (Yang et al., 2022). The number of waist vertebra and vertebra were counted manually, the weight of carcass and head were measured by electronic scales, the carcass of straight length was measured using either rule or tape, and all of these five traits were recorded after slaughter (Gong et al., 2019).

2.2 | SV discovery and genotyping

Clean reads from each individual were mapped to *Sus scrofa* genome assembly 11.1 (*Sscrofa11.1*) using Burrows-Wheeler Aligner (BWA) with Mem parameter (Li & Durbin, 2009). SAMTools v1.11 was used to mark duplicate reads and sort the mapped reads (Li et al., 2009). The traditional pipeline of LUMPY v0.2.13 employing parameters

TABLE 1 Summary information of 305 samples in this study

Regions	Breeds	Abbreviation	Number	Source	Regions	Breeds	Abbreviation	Number	Source
SouthWest pig of China	Bamei	BM	5	PRJNA398176	Western Commerical	Landrance	LR	27	JXLAB, PRJNA260763
	Baoshan	BS	6	PRJNA398176		Pietrain	PI	6	JXLAB
	Neijiang	NJ	6	PRJNA398176	Duroc	DU	26	JXLAB, PRJNA260763	
Jianghai pig of China	Erhualian	EHL	29	PRJNA488327	French large white	LW	34	JXLAB	
	Jinhua	JH	6	JXLAB, PRJNA398176	LargeWhite	LW	26	JXLAB, PRJNA260763	
Southern pig of China	Bamaxiang	BMX	12	JXLAB, PRJNA213179	European Indigenous	Iberian pig	ED	1	PRJNA255085
	Luchuan	LC	6	PRJNA213179	American Indigenous	Red manganlica	ED	1	PRJNA239399
Northern pig of China	Wuzhishan	WZS	6	PRJNA213179	SwallowBelly	manganica	ED	1	PRJNA239399
	Hetao	HT	5	PRJNA213179	Yucatan miniature	YC	11	PRJNA260763	
Laiwu	LWH	15	JXLAB, PRJNA213179	Asian Wild boar	Southern wild boar of China (Zhejiang)	SWB	2	PRJNA213179	
	Min	Min	6	PRJNA213179	Southern wild boar of China (Shangyou)	SWB	2	PRJNA213179	
Tibetan	Gansu Tibetan	GST	18	JXLAB, PRJNA213179	Southern wild boar of China (Nanchang)	SWB	2	PRJNA213179	
Sichuan Tibetan	SCT	12	JXLAB, PRJNA213179	Korean wild boar	KWB	8	PRJNA260763		
	Yunan Tibetan	YNLT	12	JXLAB, PRJNA213179	Spanish wild boar	EWB	1	PRJNA255085	
Tibetan Tibetan	TT	12	JXLAB, PRJNA213179	Outgroup	<i>Sus verrucosus</i>	Outgroup	1	PRJEB1683	

Abbreviation: JXLAB, laboratory from Jiangxi agricultural university.

such as discordant_z = 2, back_distance = 20, weight = 1, min_mapping_threshold = 20 was used to infer the genotypes and locations of the SVs (Layer et al., 2014). The SVs generated for each individual were sorted and merged with the -p 20 option using svtools (Figure S1c) (Larson et al., 2019). We filtered out SVs supported by only one split read, located on unplaced contigs, with length <5 bp and >5 Mb, genotype quality score equal to 0 or call rate less than 50%. Platypus software was performed to call SNP based on the same bam file used in the SV detecting pipeline (Rimmer et al., 2014), Beagle was then implemented to impute missing genotypes (Browning & Browning, 2007). Minor Allele Frequency (MAF) low than 3% in each SNP was removed by PLINK v1.9 (Purcell et al., 2007). Finally, we obtained 58,733,085 SNPs for comparing the clustering tendency with SVs.

2.3 | Functional annotation of SV

We used the Variant Effect Predictor (VEP) program to annotate genomic features of the identified SVs (McLaren et al., 2016). From the annotation results, we grouped the SVs into those located in the coding region (CDS), untranslated region (UTR), intron (intron), and intergenic (intergenic) regions. Ensembl gene names were transformed into official gene symbols using the Biomart online tool (<http://asia.ensembl.org/biomart/martview/c96ad4072c30630bcfbf44459a1722bc>).

2.4 | Population genetic analyses of SVs

We determined phylogenetic relationship and population structure based on pairwise genetic distances inferred from filtered SV genotype data using PLINK v1.9 (Purcell et al., 2007). The neighbor-joining phylogenetic tree (NJ tree) was constructed using FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). Principal component analysis (PCA) was performed using PLINK v1.9 (Purcell et al., 2007) and visualized employing the R program. Population differentiation analysis was performed using fixation index (F_{ST}) and absolute allele frequency difference (deltaAF) as described previously (Huang et al., 2020). The significance threshold was determined by the 1000 times permutation tests (Churchill & Doerge, 1994). Population-private SVs are those with alleles found only in one population and have frequency >0.5. The ancestor allele of an SV is determined as an allele that is fixed i.e., in homozygous status in the Javan Warty pig. When an SV is heterozygous in the Javan Warty pig, the major allele in the wild boars is determined as the ancestral allele, while the alternative allele is defined as a derived allele.

2.5 | Local assembly

All the aligned reads extracted from the BAM file were locally assembled using SOAPdenovo2 with parameters such as reverse_seq = 0,

asm_flags = 1, rank = 1, pair_num_cutoff = 3, and map_len = 32 (Luo et al., 2012). SV breakpoints were verified by mapping the contig to *Sus scrofa* 11.1 using Basic Local Alignment Search Tool (BLAST) from the Ensembl website.

3 | RESULTS

3.1 | Variant discovery and population structure

Sequence data of 462 pigs, with average read lengths of 100, 125, and 150 bp and insert sizes of 300 and 500 bp, were initially analyzed. As the accuracy of SV genotype inference is highly dependent on the read depth, we retained the sequences of 305 individuals (Table S2), including those of 304 *Sus scrofa* and a Javan Warty pig have sequence depths >10 and genomic coverage >0.96, for subsequent analysis (Materials and Methods; Figure S1a,b). The 304 individuals included 56 Chinese indigenous pigs, 6 Chinese wild pigs, 8 Korean wild pigs, 119 European commercial pigs, 3 European indigenous pigs, 1 European wild pig, and 11 Yucatan pigs originating from Mexico (Table 1). To avoid including too many Javan Warty-specific SV in the reference SV set, we used the 304 *Sus scrofa* individuals in SV discovery and quality control procedures. We identified an average of 19,100 SVs per sample, ranging from 6478 in a Large White pig to 26,776 in Yunnan Tibetan pig (Table S2). The number of identified SVs in the samples increased with the sequence depth and reached a plateau at a sequence depth of approximately 23x (Figure S2a). We merged all the identified SVs of the 304 individuals into a non-redundant set of 162,134 SVs using svtools (Larson et al., 2019) (Figure S1c).

After removing SVs classified as BND (breakends) (Danecek et al., 2011), mapped to unplaced contigs, with genotype call rates less than 90% and with sizes <50 bp or >5 Mb, we retained 112,890 SVs including 79,919 (70.8%) deletions, 23,638 (20.8%) duplications and 9333 (8.3%) inversions for subsequent analysis (Table S1; Figure 1a). Most of the deletions (DELS) had lengths ranging from 250 to 320 bp, which are comparable to the size of DELs reported in humans, where nearly 80% of the breakpoints of SVs overlapped with repetitive elements such as SINEs and LINEs, which have a length ~300 bp (Quan et al., 2021). The size of duplications (DUPs) and inversions (INVs) was generally greater (Figure 1b). These results agreed with those observed in humans (Chiang et al., 2017; Sudmant et al., 2015). The average call rate of the SVs was 98.5% (Figure S2b). The identified SVs hereby replicated 49.8%, 97.8%, and 46.3% of DELs, DUPs, and INVs, respectively, in an independent study (Figure S2c) (Zhao et al., 2016).

Next, we clustered the samples based on the 112,890 SVs and compared the neighbor-joining tree constructed from these SVs with that constructed using genome-wide SNPs (58,733,085) (Ai et al., 2021). By comparing SV- and SNP-based neighbor-joining (NJ) tree, we further excluded three Tibetan pigs that were misclassified by SV data. In the NJ trees constructed from both the SV and SNP

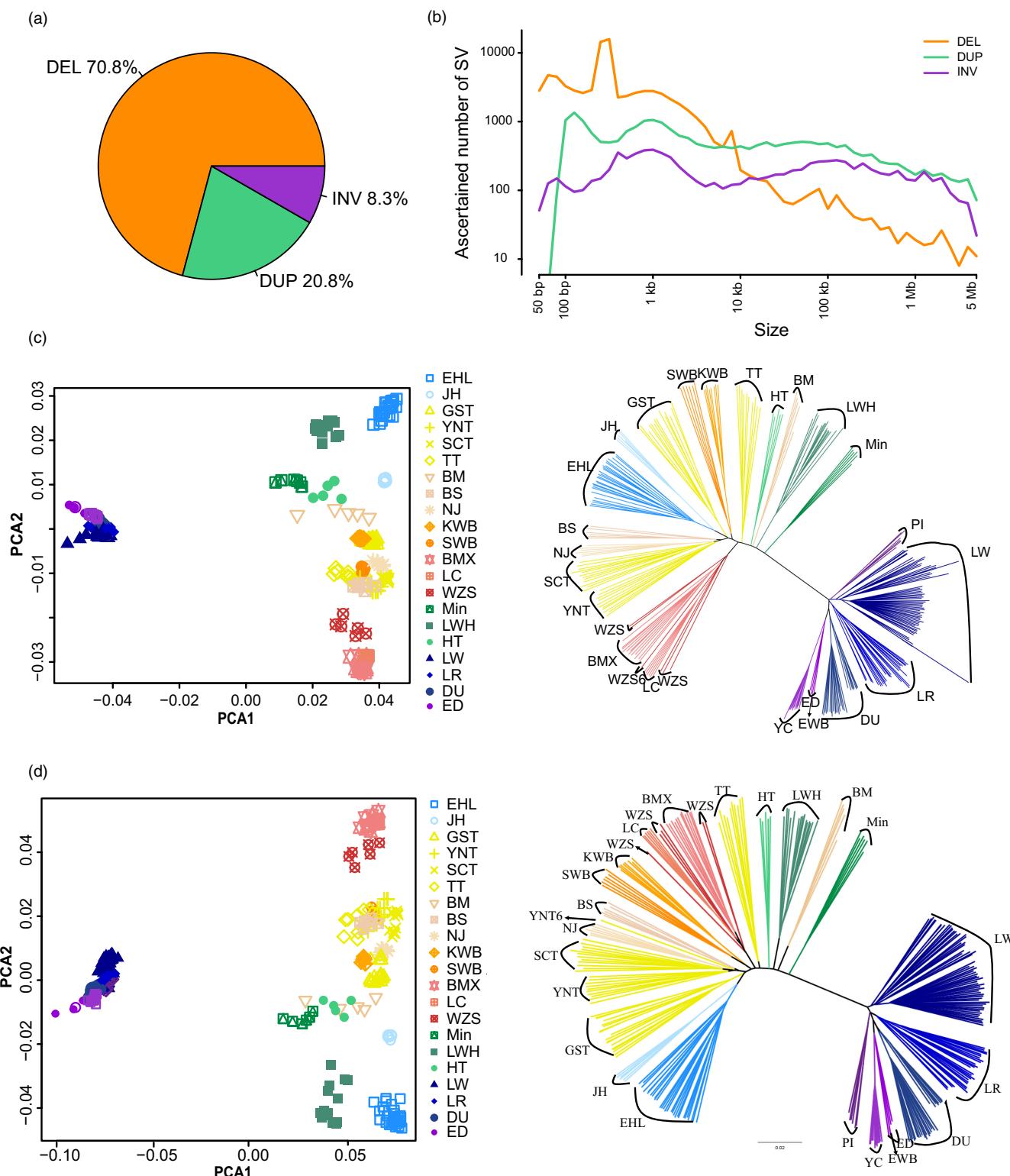


FIGURE 1 Properties of SVs across swine populations. (a) Pie plot showing the proportion of three categories SV. (b) Distribution of three categories SV size. Comparing with duplications (DUPs) and inversions (INVs), deletion (DELs) were smaller and approximately 40% DELs locate between 250 and 320 bp. (c) Principal component analysis (PCA) and neighbor-joining phylogenetic tree (NJ tree) inferred from genotype data of 58,733,085 SNPs in 301 pigs. (d) PCA and NJ tree inferred from genotype data of 112,890 SV in 301 pigs.

data of the rest 301 individuals, the pigs from Europe and Asia were clearly separated from each other, and individuals from the same breed were clustered together (Figure 1c,d). The DELs performed

better than DUPs and INVs in grouping individuals with known breed identity (Figure S3a,b), consistent with observations in humans (Almarri et al., 2020).

3.2 | Genetic diversity and evolutionary history of SVs

We further assessed the SV-based genetic diversity in different pig breeds based on the number of heterozygous SVs in each individual. Generally, Asian pigs have more heterozygous SVs per individual than European pigs (Figure S3c), consistent with previous observations based on SNP data that genetic diversity was greater in Asian pigs than in European pigs (Ai et al., 2015; Yang et al., 2017). To examine the evolutionary history of SV, we further investigated the

distribution of alleles in 1-24 pig populations (Figure 2a). The results showed that most of the SVs (40.4% DELs, 50.4% DUPs, and 51.8% INVs) shared alleles across all 24 populations (Figure 2b). The MAF of the DELs overlapping with CDS and UTR was significantly ($p\text{-value} = 2.2 \times 10^{-16}$, T-test) lower than those in intergenic and intronic regions, suggesting that the DELs in CDS and UTR might have been subjected to selection due to their deleterious effect on gene functions (Figure 2d). Additionally, we noted that DELs in CDS and UTR were more likely present in 1-3 populations, indicating that mutations in these SVs might have occurred recently (Figure S4a-d).

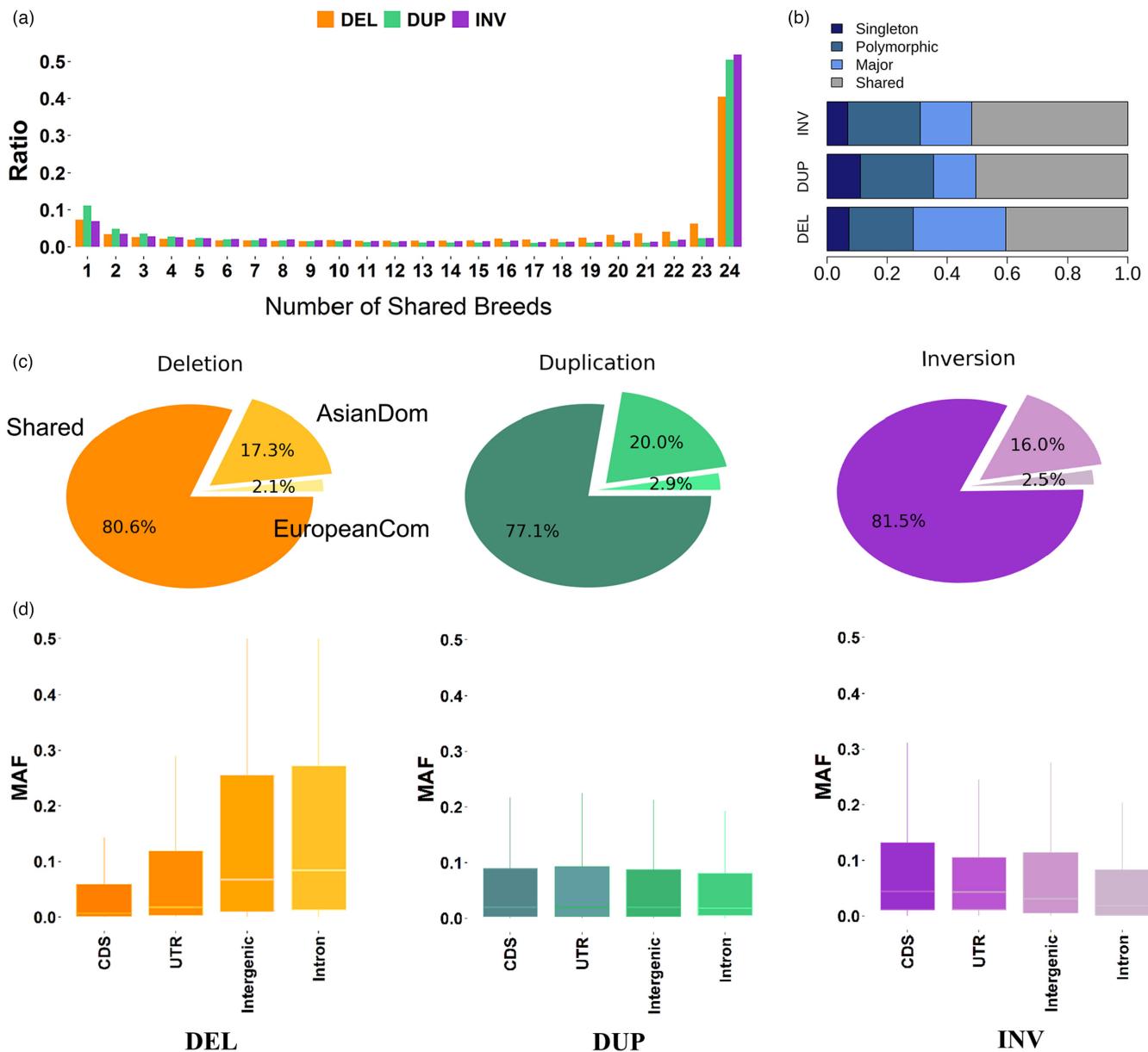


FIGURE 2 Evolutionary history of SV. (a) Distribution of number of breeds that share a SV allele by the three SV types: DEL, DUP, and INV. (b) The frequency for each SV type. (c) Proportion of shared and specific SV alleles in European commercial and Asian (Chinese) domesticated breeds. (d) Distribution of MAF of the three types of SVs in CDS (coding regions), UTR (5' UTR and 3' UTR), intergenic, and intron. DEL overlapped with CDS and UTR is significantly ($p\text{-value} = 2.2 \times 10^{-16}$, T-test) lower than its intergenic and intron.

3.3 | Highly differentiated SVs between Asian and European domestic pigs

Domestic pig breeds in Europe and Asia have different evolution, domestication, and selection history, and display dramatic phenotypic divergences (Frantz et al., 2013; Yang et al., 2017). The majority of the SV alleles were shared by the Asian and European commercial pigs, while private SV in Asian domesticated breeds was much greater than those in European commercial breeds for all three SV types (Figure 2c), agreeing with the observations that the genetic diversity is higher in Asian than in European domesticated pigs (Groenen et al., 2012; Wu et al., 2021). We assumed that at least a part of the SVs displaying extreme allele frequency differences between these two groups of pigs may have been subjected to selection and may have caused phenotypic differences between the two groups of pigs. We combined F_{ST} and deltaAF (Materials and Methods) to identify stratified SVs exceeding the top 1% between Asian and European pigs—a total of 1007 SVs ($F_{ST} > 0.870$ and deltaAF > 0.864), including 988 DELs, 6 DUPs, and 13 INVs (Table S3; Figure 3a), which reflected the divergent evolution and selection history of European and Asian pigs. We then performed GO enrichment analysis of these genes using the ClueGO program in Cytoscape v3.7.1, and uncovered pathways related to the regulation of the release of sequestered calcium ions into the cytosol (*JPH1*, *LYN*, *MTTP*, *MYO5A*, *SLC8A1*, *TFRC*; adjusted *p*-value = 0.02), positive regulation of cell-matrix adhesion (*DAPK2*, *MTTP*, *ROCK1*; adjusted *p*-value = 0.05), T cell migration (*CXCR3*, *DOCK8*, *LRCH1*, *MSN*, *NOX1*; adjusted *p*-value = 0.03), regulation of establishment of cell polarity (*DOCK8*, *KANK1*, *ROCK1*; adjusted *p*-value = 0.04) and neuromuscular junction development (*COL4A5*, *DAPK2*, *TFRC*; adjusted *p*-value = 0.03) (Table S5), which may reflect differences in the body shape, immunity and motor behavior of the pig breeds from the two domestication centers.

Among these SVs, we identified a total of 270 SVs (accounting for 64.4% of all differentiated SVs) located in the 48-Mb low recombination region on chromosome X (44.0–91.5 Mb) that was identified previously (Table S3) (Ai et al., 2015). These SVs overlapped with 63 genes, including genes related to mitochondria (*ABCB7*, *ATP7A*, *APOOL*, *TRMT2B*, *COX8B*, and *COX7B*) (Bhattacharjee et al., 2016; Laptev et al., 2020; Pondarré et al., 2006; Weber et al., 2013), regulation of transcription (*ZNF449*, *TAF9B*, *TBX22*, and *DACH2*) (Braybrook et al., 2001; Zheng Chen & Manley, 2003; Cohen et al., 2007; Luo et al., 2006), autism spectrum disorder (*PTCHD1*, *PCDH19*) (Jamal et al., 2010; Torrico et al., 2015), and adipogenesis (*TNMD*) (Senol-Cosar et al., 2016), indicating the potential roles of these genes on phenotypic differences, such as behavior, between European commercial and Asian domesticated breeds.

Among the highly differentiated SVs on the autosome, we identified 135 SVs covering 130 genes (Table S3). Among the top 10 SV-associated genes (SV genes) (*ANKRD11*, *PDGFRA*, *KLHL2*, *PLAG1*, *PLPPR1*, *NR6A1*, *CORIN*, *LEO1*, and *CPE*) ranked by their population differentiation (Figure 3a; Table 2), we identified *ANKRD11* (DEL

309 bp) associated with the KBG syndrome, a disease characterized by short stature, skeletal anomalies, delay in global developmental and regulation of bone homeostasis (Lim et al., 2014). *PDGFRA* (DEL 54 bp) is a paralog of the *KIT* gene, which is related to the dominant white coat color in western pigs (Johansson et al., 1992). Notably, the top 10 SV genes included two DELs in *NR6A1* and 1 DEL in *PLAG1* (Figure 3c,d and Table S6). *NR6A1* has been reported to be a candidate gene related to the number of vertebrae in pigs (Rubin et al., 2012) and *PLAG1* (zinc finger protein) is a candidate causal gene related to bovine stature (Karim et al., 2011). Based on whole genome sequence data, we determined genotypes of the three DELs in 836 individuals from eight breeds that crossed the 6th generation heterogeneous population, and examined their associations with five growth traits (Table S6). We observed and showed that the *PLAG1*-SV is associated with weight of head (*p*-value = 1.3×10^{-10}), carcass weight (*p*-value = 5.3×10^{-9}) and straight length of carcass (*p*-value = 6.7×10^{-10}) (Figure 3e,f; Figure S5a,b), *NR6A1*-SVs (chr1:265,568,219–265,568,536; Table S6) have significantly associated with straight length of carcass (*p*-value = 7.5×10^{-4}) (Figure S5c), the number of vertebrae (*p*-value = 1.0×10^{-39}) (Figure 3g,h) and waist (*p*-value = 3.0×10^{-15}) (Figure S5d). The alleles that increase the growth performance at these two loci were inferred to be derived alleles by referring to the genotypes of the Javan Warty pig (Table S1). Notably, the allele of *PLAG1*-SV that increased the body weight remained fixed in European commercial breeds but was segregated in Asian indigenous breeds such as Laiwu (LWH) and Min pigs, indicating a plausible introgression of the genome sequence at two gene regions from European breeds to two Chinese indigenous breeds (Figure 3b,e). These findings agreed with historical records that stated that European commercial pigs were introduced to be crossbred with northern local breeds in China (Wang et al., 2011; Wang et al., 2021).

3.4 | Highly differentiated SVs between indigenous pigs and wild boars in Asia

We next investigated the SVs that were highly differentiated between indigenous pigs and wild boars in Asia, and identified a total of 789 SVs including 770 DELs, 4 DUPs, and 15 INVs covering 137, 1, and 1 genes, respectively (Table S4; Figure 4a). The genes were enriched for GO terms (Table S5) such as endothelial cell development (*AFDN*, *DMD*, *TNMD*; adjusted *p*-value = 0.01), DNA interstrand cross-link repair (*ERCC6L*, *FAAP100*, *FANCA*; adjusted *p*-value = 0.01), and cellular response to amino acid stimulus (*ASS1*, *ATP7A*, *COL5A2*; adjusted *p*-value = 0.003), suggesting a potential role pertaining to movement ability, exposure to UV effect and nutrient absorption, respectively (Tang & Conti, 2004; Tretyakova et al., 2015). Among the top 10 SV genes (Table 2; *PMS1*, *SFT2D2*, *COL4A5*, *ASS1*, *PRIM1*, *HSD17B6*, *TBX19*, *NUP42*, and *CPKOW*), we identified genes including *PMS1* (DEL 1148 bp) involving that is involved in the repair of DNA mismatches and has been reported as

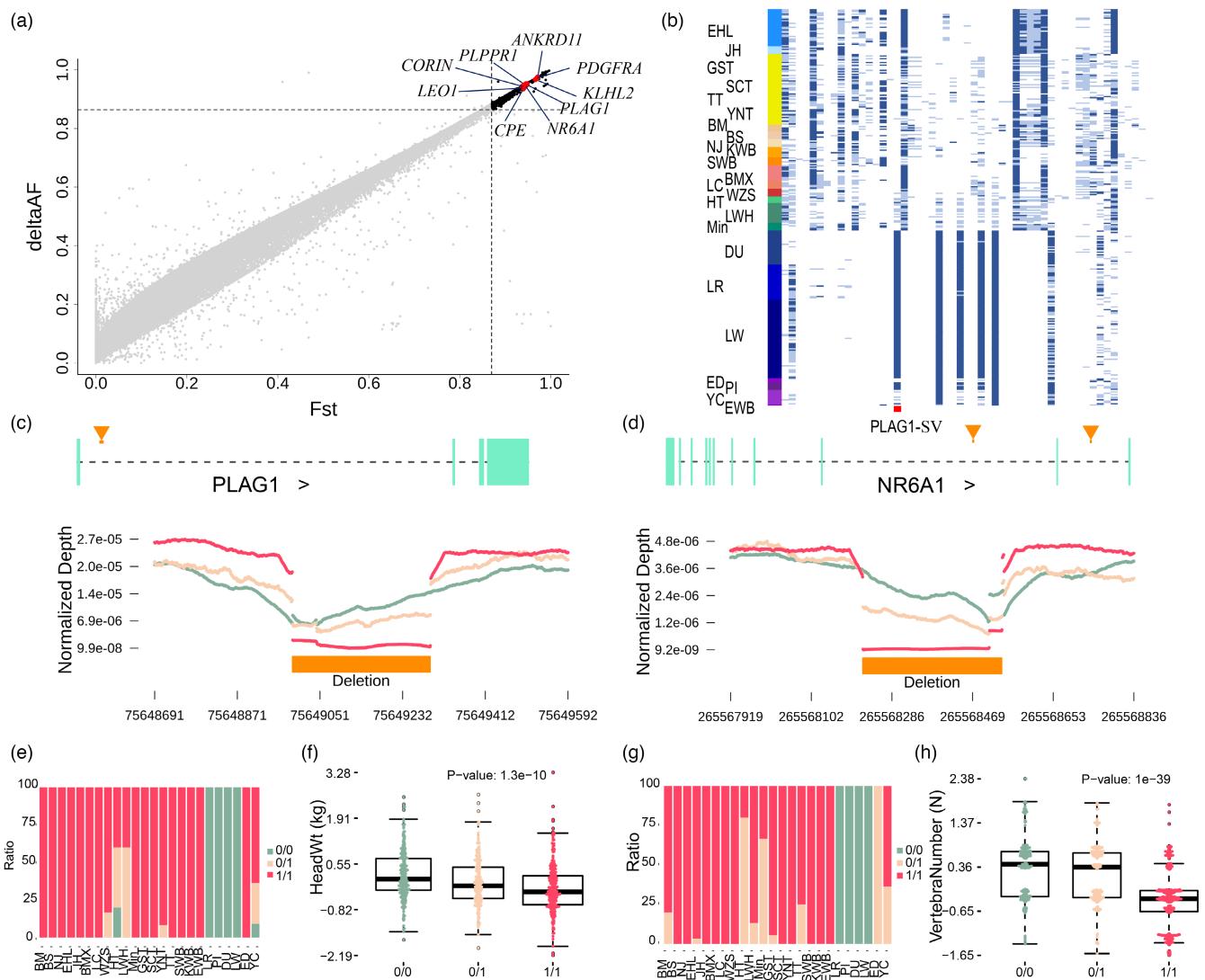


FIGURE 3 Highly differentiated and introgression SVs between Asian and European domestic pigs. (a) Dark gray points represent the SVs exceeding top 5% (p -value <0.05) of both absolute ΔAF and F_{ST} ; red points (top10 gene on autosome) were discussed and listed in Table 2. (b) The distribution of SVs in chr4:75,276,459–76,476,146 region surrounding the *PLAG1* gene, showing the distribution of SV allele across the 24 populations; red rectangle indicates the SVs within *PLAG1* gene; dark blue, light blue, and white rectangles represent the 0/0, 0/1 and 1/1, respectively. (c, d) Normalized depth for deletion within *PLAG1* (301 bp chr4:75,648,991–75,649,292) and *NR6A1* (317 bp chr1:265,568,219–265,568,536). Normalized read count was employed for comparing the regional depth between groups using bedtools v2.28.0 with coverage parameter. Normalized read count = SRN/RTR, SRN is site of reads number, while RTR is region of total reads, a region is always large than the target gene. (e, g) Distribution of allele frequency among 24 populations within *PLAG1* and *NR6A1*. (f, h) Genotyping the DEL in 6th generation individuals from a pig heterogeneous population ($N = 836$) using the whole genome sequence data showed significant association of the *PLAG1*-SV with head weight (p -value = 1.33×10^{-10} , ANOVA test) and *NR6A1*-SV with the number of vertebra (p -value = 1.04×10^{-39} , ANOVA test), respectively. The value of Y axis is the residual of the corresponding phenotype corrected for sex, transport batch, and slaughter using *lm()* in R. Listed in Table S6.

a candidate gene for serum ferritin in a Chinese population (Liao et al., 2014). *SFT2D2* (DEL 273 bp) was detected as a candidate gene for adipose deposition, consistent with fat content diversity between wild and domesticated pigs in Asia (Zhang et al., 2019). Two DELs were detected within the *NUP42* gene (nucleoporin 42; DEL 292 and 288 bp) and domestic to wild boars in Asia. Interestingly, we found a DUP covering the 5' UTR and the first exon of *TBX19* (DUP 9466 bp; Figure 4c). *TBX19* gene is a member of the phylogenetically conserved T-box gene family and encodes transcription

factors involved in the regulation of developmental processes, which have been reported to regulate the behavior of humans and other animals (Zhu et al., 2017). Mutations in the *TBX19* gene cause isolated ACTH deficiency (Peng et al., 2020). A 9466 bp DUP was located in a 70 kb haplotype shared in Chinese breeds, and was highly differentiated from the haplotypes of European pigs and all wild boars (Figure 4b) (Zhu et al., 2017). A similar study revealed that a large copy number variation (~12 kb) encompassing the 14th exon is associated with increased GC expression and low resistance

TABLE 2 Highly differentiated SVs (top 10) among groups

Chr_SV_Start	Group	SV length	SV type	Annotation For SV	Ensembl ID	F _{ST}	deltaAF	Average value	Gene name	Gene description
6_533365	EuropeanComVASianDom	309	DEL	Intron	ENSSSCG000000039541	0.970797	0.974359	0.972577789	ANKRD11	ankyrin repeat domain 11 [Source:VGNC Symbol;Acc:VGNC:85323]
8_40982396	EuropeanComVASianDom	54	DEL	Intron	ENSSSCG000000008841	0.966487	0.967168	0.966827276	PDGFRA	platelet-derived growth factor receptor alpha [Source:VGNC Symbol;Acc:VGNC:98179]
8_43772551	EuropeanComVASianDom	369	DEL	Intron	ENSSSCG000000008858	0.947984	0.952139	0.95006144	KLHL2	kelch-like family member 2 [Source:VGNC Symbol;Acc:VGNC:89518]
4_75648991	EuropeanComVASianDom	301	DEL	Intron	ENSSSCG00000006247	0.945408	0.951923	0.948665376	PLAG1	PLAG1 zinc finger [Source:VGNC Symbol;Acc:VGNC:91509]
1_242797376	EuropeanComVASianDom	275	DEL	Intron	ENSSSCG00000005393	0.944978	0.950927	0.94795226	PLPPR1	phospholipid phosphatase related 1 [Source:VGNC Symbol;Acc:VGNC:103153]
1_265486649	EuropeanComVASianDom	321	DEL	Intron	ENSSSCG00000005589	0.941793	0.948718	0.945255523	NR6A1	nuclear receptor subfamily 6 group A member 1 [Source:VGNC Symbol;Acc:VGNC:90887]
1_265568219	EuropeanComVASianDom	317	DEL	Intron	ENSSSCG00000005589	0.941793	0.948718	0.945255523	NR6A1	nuclear receptor subfamily 6 group A member 1 [Source:VGNC Symbol;Acc:VGNC:90887]
8_37721904	EuropeanComVASianDom	260	DEL	Intron	ENSSSCG00000008813	0.94345	0.943951	0.943700463	CORIN	corin, serine peptidase [Source:VGNC Symbol;Acc:VGNC:86911]
1_11948286	EuropeanComVASianDom	271	DEL	Intron	ENSSSCG00000004625	0.937254	0.94352	0.940387064	LEO1	LEO1 homolog, Paf1/RNA polymerase II complex component [Source:VGNC Symbol;Acc:VGNC:89683]
8_43627960	EuropeanComVASianDom	292	DEL	Intron	ENSSSCG00000008854	0.939965	0.935763	0.937863664	CPE	carboxypeptidase E [Source:VGNC Symbol;Acc:VGNC:86936]
15_94385189	DomVSWildnAsia	1148	DEL	Intron	ENSSSCG00000016045	0.980377	0.892857	0.936617163	PMS1	PMS1 homolog 1, mismatch repair system component [Source:VGNC Symbol;Acc:VGNC:98205]
4_82722178	DomVSWildnAsia	273	DEL	Intron	ENSSSCG00000031617	0.919863	0.925824	0.922843653	SFT2D2	SFT2 domain containing 2 [Source:VGNC Symbol;Acc:VGNC:98856]

TABLE 2 (Continued)

Chr_SVStart	Group	SV length	SV type For SV	Annotation Ensembl ID	F_{ST}	deltaAF	Average value	Gene name	Gene description
X_88980615	DomVSWildnAsia	202	DEL	Intron	ENSSSCG000000012572	0.859995	0.919872	0.889933564	COL4A5 collagen type IV alpha 5 chain [Source:VGNC Symbol;Acc:VGNC:86876]
1_270587309	DomVSWildnAsia	369	DEL	Intergenic	ENSSSCG000000005701	0.852344	0.861264	0.856803893	ASS1 argininosuccinate synthase 1 [Source:VGNC Symbol;Acc:VGNC:85590]
5_22087052	DomVSWildnAsia	301	DEL	Intron	ENSSSCG00000026055	0.911439	0.769689	0.840563948	PRIM1 DNA primase subunit 1 [Source:VGNC Symbol;Acc:VGNC:103313]
5_22087052	DomVSWildnAsia	301	DEL	Intron	ENSSSCG00000027854	0.911439	0.769689	0.840563948	HSD17B6 hydroxysteroid 17-beta dehydrogenase 6 [Source:VGNC Symbol;Acc:VGNC:103289]
4_82658000	DomVSWildnAsia	9466	DUP	CDS	ENSSSCG00000006299	0.80097	0.858516	0.829743416	TBX19 T-box transcription factor 19 [Source:VGNC Symbol;Acc:VGNC:93797]
9_9190818	DomVSWildnAsia	292	DEL	Intron	ENSSSCG00000023793	0.937348	0.714286	0.825816657	NUP42 nucleoporin 42 [Source:VGNC Symbol;Acc:VGNC:90981]
X_43194315	DomVSWildnAsia	267	DEL	Intergenic	ENSSSCG00000012291	0.721637	0.830128	0.77588256	GPKOW G-patch domain and KOW motifs [Source:VGNC Symbol;Acc:VGNC:88589]

Note: EuropeanComVSAsianDom, group between European commercial breeds and Asian domesticated pigs, we selected top 10 SV genes from the autosome for this group; DomVSWildnAsia, group between domesticated and wild pig in Asia.

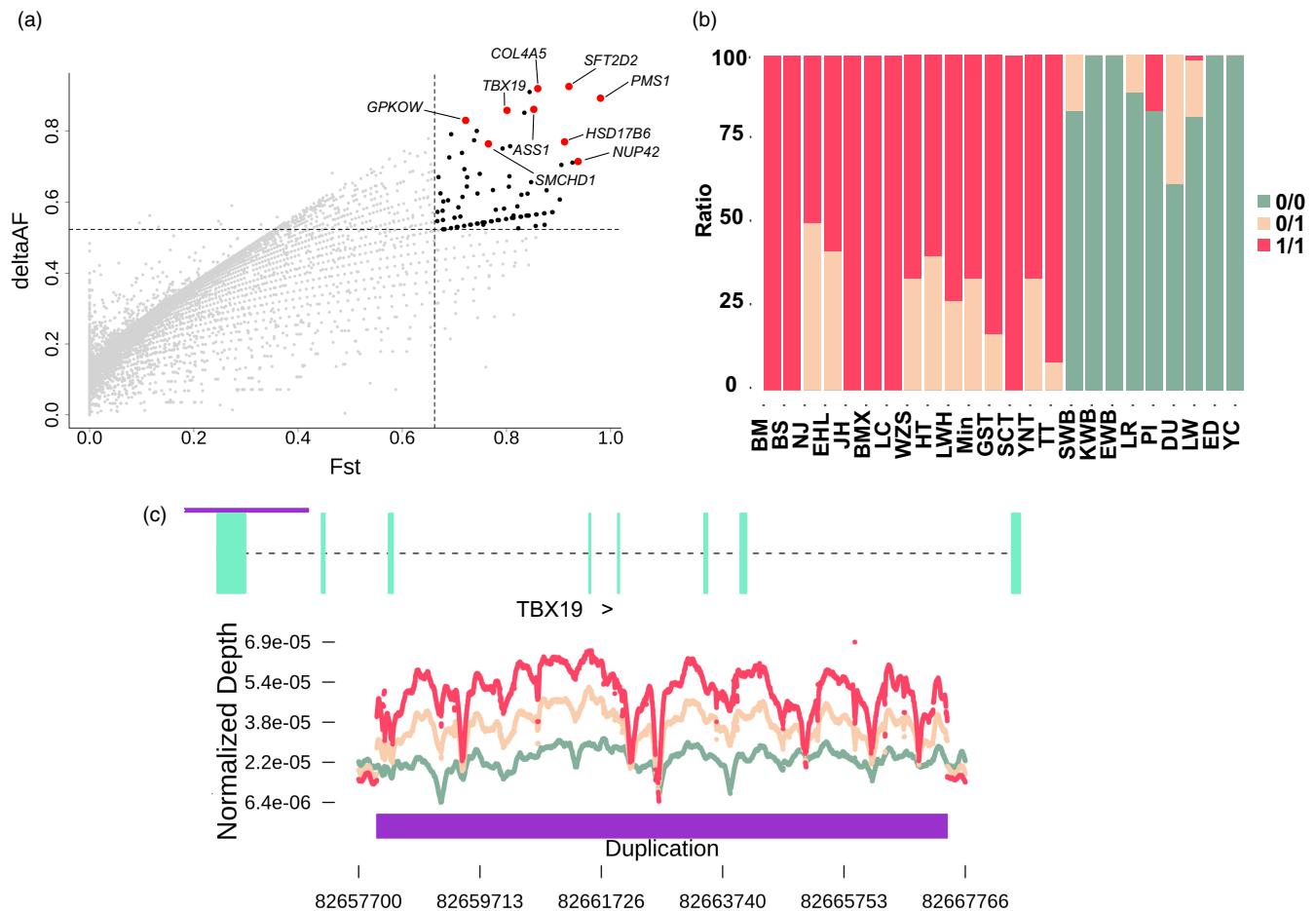


FIGURE 4 Highly differentiated SVs between indigenous pigs and wild boars in Asia. (a) Dark gray points represent the SV exceeding top 5% (p -value <0.05) of both absolute deltaAF and F_{ST} ; red points (top 10 gene) were selected for discussion and listed in Table 2. (b) Distribution of DUP (9466 bp chr4:82,658,000–82,667,466) allele frequency among 24 populations within TBX19. (c) Normalized depth for duplication in TBX19, which includes the 3' UTR and exon 8.

to clinical mastitis in dairy cattle (Lee et al., 2021). Thus, we speculate that this duplication would increase the expression of TBX19 gene and play an important evolutionary role in Asian (Chinese) domestic breeds.

3.5 | Identifying SVs that are associated with environmental adaptation in Chinese indigenous pigs

To assess the potential role of SVs that play a role related to environmental adaptation in indigenous Chinese pigs, we focused on two environmental variables: altitude and local temperature. We performed a genome-wide F_{ST} analysis (Figure 5a) between Tibetan pigs (GST, SCT, TT, and YNT) and non-Tibetan Chinese domestic pigs (BM, BMX, BS, EHL, HT, JH, LC, LWH, Min, NJ, and WZS). At an empirical p -value threshold of 0.05 ($F_{ST} = 0.49$; 1000 times permutation test), we identified a total of 12 differentiated SVs. Among these, we identified a 61 bp DEL in the *MYBPC1* gene, which had allele frequency ranging from 0.25 to 0.56 in Tibetan pig population, low allele frequency in EHL, LC, and WZS pigs, and was absent in the other breeds. *MYBPC1* is related to the development of skeletal

and cardiac muscle (Chen et al., 2011; Dhoot & Perry, 2005). Hence, it can be assumed that SVs could be involved in the adaptation of high-plateau pigs to the low-oxygen environment by affecting muscle and heart functions.

A similar method was performed between breeds from hot (WZS, BMX, and LC) and cold regions (Tibetan pigs, Min, LWH, and HT) in China. A total of 248 SVs covering 30 genes exceeding the threshold ($F_{ST} = 0.75$; $p = 0.05$, 1000 permutation test) were identified (Figure 5b). Notably, we observed an SV hotspot in a previously reported low recombination region on chromosome X (44.0–57.8 Mb) (Ai et al., 2015), covering genes associated with the nervous system or mental disorders (PHF8, WNK3, FAM120C, LIN7C, HUWE1, KLF8) (Table S7) (Abidi et al., 2007; Lossi et al., 2002; Moortgat et al., 2018; Qiao et al., 2008; Shinawi et al., 2011), congenital diaphragmatic hernia (*STRAD8*, *YIPF6*, and *OPHN1*) (Petit et al., 2011). We also found that disruption of the *EDA* gene could result in sparse hair and defects in cutaneous glands (Escouflaire et al., 2019), and androgen receptor gene (*AR*), a gene associated with reproductive abilities (Eisermann et al., 2013).

We found 11 SVs in 11 genes (*LIN7C*, *CYP3A29*, *STRN4*, *KIF16B*, *TLK1*, *FMO1*, *MYO5B*, *GRIK2*, *TCAIM*, *ATP5H*, and *SFRP1*) related to

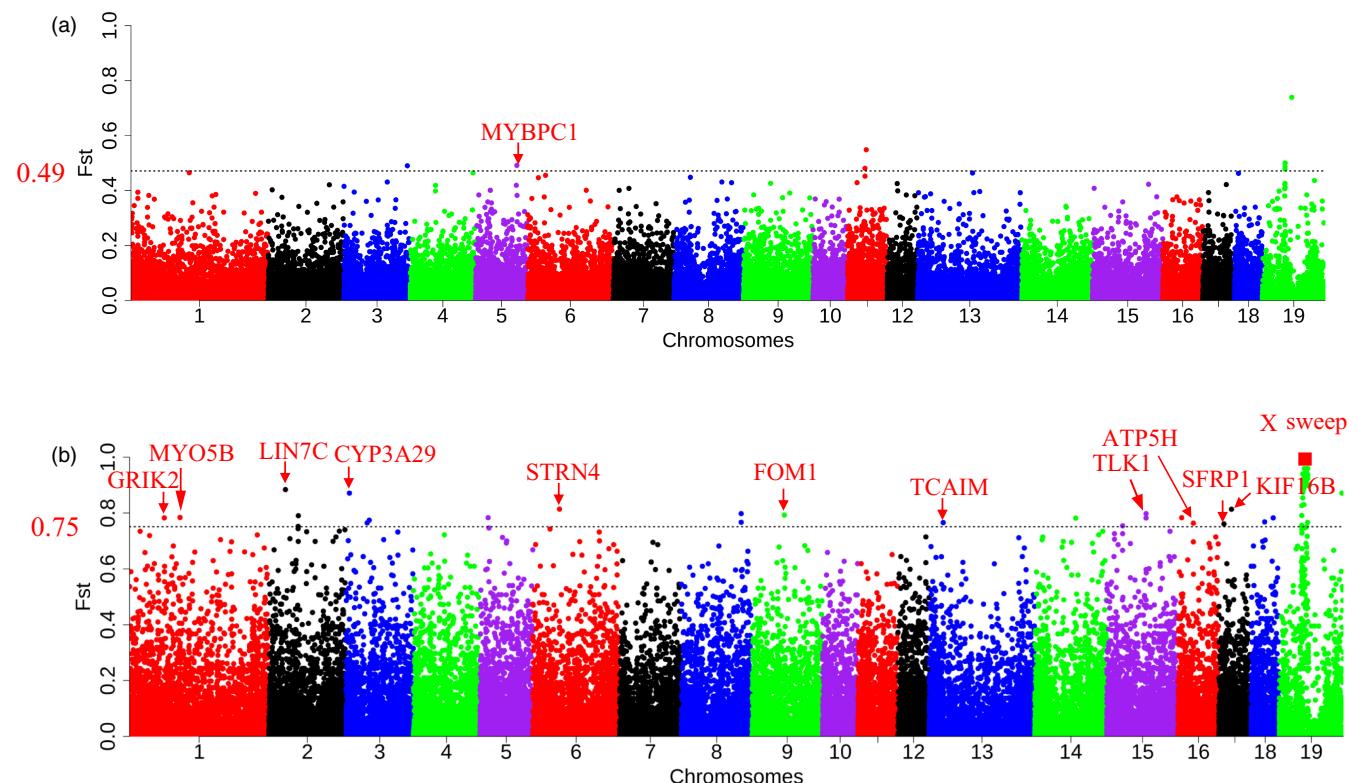


FIGURE 5 Population differentiation of SV in Chinese domesticated pigs. Manhattan plot showing the population differentiation analysis (F_{ST}) on pig populations from the (a) hot and cold, and (b) high and low altitude areas in China. Significance thresholds were determined empirically by permutation testing (1000 times). The red rectangle in (a) marks the SV hotspot on chromosome X.

autosome; these SVs exceeded the significance threshold (Table S7). Of these, *TCAIM* (DEL 808 bp) has functions related to mitochondria, which might be involved in adaptation to cold temperatures (Vogel et al., 2015). The *SFRP1* gene (DEL 276 bp) was included in the QTL region that influences intramuscular fat (IMF) content and marbling; correlation analysis revealed that the expression of *SFRP1* gene was significantly ($r = 0.45$, p -value = 1.69×10^{-2}) correlated with IMF content in the Duroc breed (Zhao et al., 2019).

3.6 | Identifying population-private SVs

To investigate the SVs that are potentially associated with the breed-specific characteristics, we focused on SVs with alleles that were specific to one population and with allele frequency >0.5 for each of the 24 populations (population-private) (Materials and Methods). Based on these criteria, we identified 718 SV genes including 379 DEL genes (Figure 6a), 155 DUP genes, and 93 INV genes (Table S8). Of these genes, we identified LWH, which is renowned for exceptionally high IMF, private SV within *FMO5* gene (DEL 65 bp) related to meat flavors (Glenn et al., 2007), and *OXR1* gene (DEL 115 bp) related to carcass meat yield (Y. Wu et al., 2016). *LCORL* (DEL 193 bp) has been reported to increase porcine body length, consistent with human-mediated introgression of pigs having European haplotype, to increase growth performance (Wang et al., 2021). Interestingly, we also found that a 5239 bp private DEL removed 2nd exon within

the *EP400* gene in LWH (Figure 6b). The expression of this gene has been reported to be increased during fat cell differentiation, thus promoting adipogenesis by incorporating a histone variant in the PPAR γ target gene (Couture et al., 2012).

Two BM-private DELs (DEL 444 bp; DEL 2815 bp) were in the *MYL1* gene (myosin light chain 1), which plays a role in porcine skeletal muscle formation (Fontanesi et al., 2000). Two Min-private SV (DEL 131 bp; DUP 632 bp) within *HDAC1* have been reported to be involved in the regulation of gene expression during carcinogenesis and spermatogenesis (Omisano et al., 2007). Jinhua (JH) pigs account for the largest proportion of private SV genes including 51 DEL genes, 20 DUP genes, and 3 INV genes. In addition, we also observed that western commercial breeds, such as Duroc (DU) harbor two private DELs (537 and 9211 bp) within the *NSUN7* gene, which is associated with sperm motility in mice (Khosronezhad et al., 2015) and is a potential genetic factor for its high reproductive performance.

The Yucatan pig carries 15 population-private SVs (9 DELs, 4 DUPs, and 2 INV) fixed in all individuals ($n = 11$; Figure 6c; Table S8) covering *DDR2*, *EFHC2*, and *FRAS1* genes. *DDR2*-deficient mice have a smaller body size and exhibit short limbs and irregular growth of flat bones, and a short snout mutation; in humans, it is found to be associated with dwarfism (Figure S6a) (Bargal et al., 2009; Labrador et al., 2001). Two DELs were observed within the *EFHC2* gene, which has been reported associated with social cognitive abilities in men (Figure S6b) (Startin et al., 2015). We focused on the *FRAS1* (Fraser syndrome 1) gene due to the findings of a previous study indicating

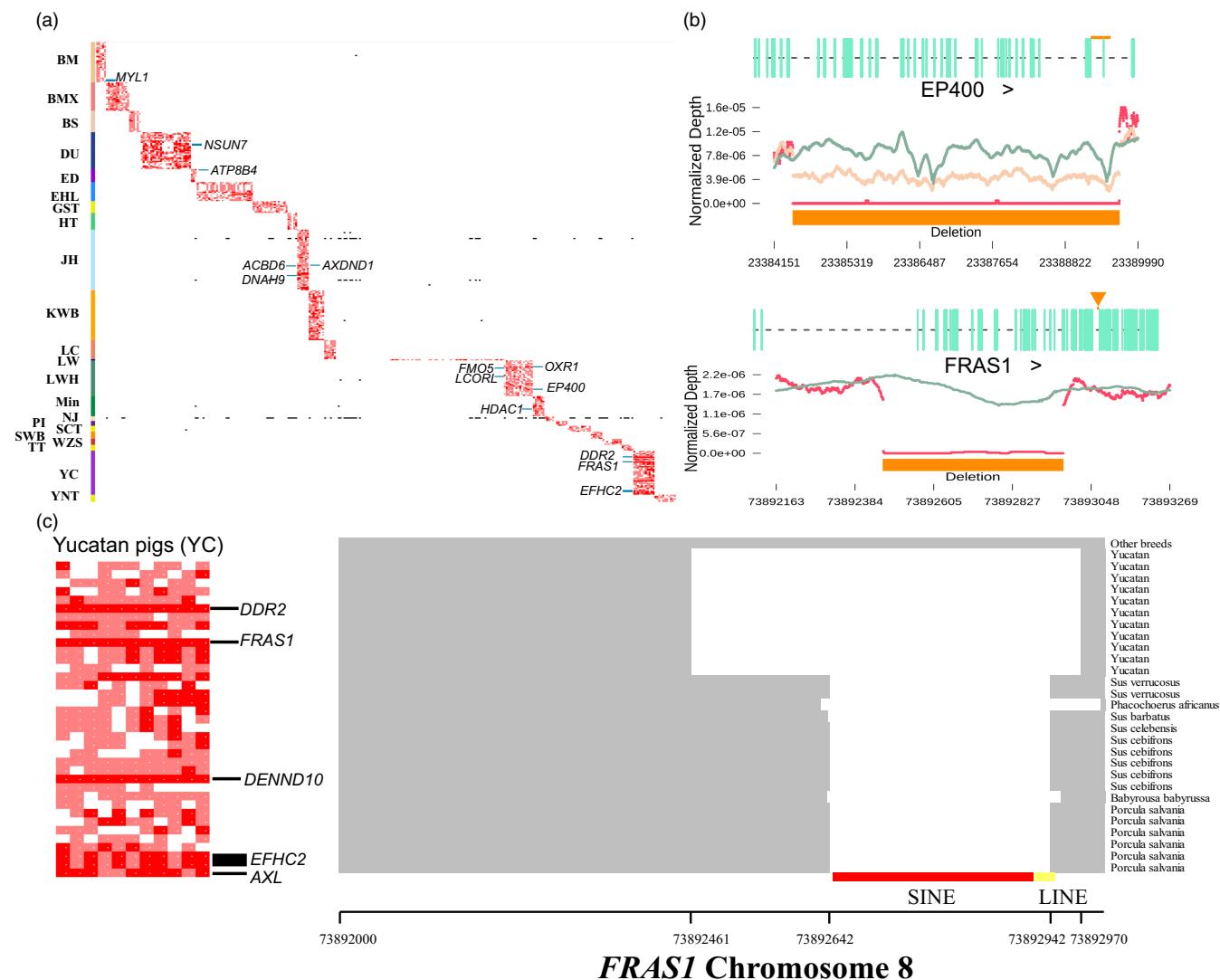


FIGURE 6 Population-private SVs in 24 diverse populations. (a) Heatmap of private DELs in 22 populations (EWB was excluded as it contains only one sample, while LR harbors none population-private SV in this study). White, light red, and dark red points represent for 0/0, 0/1, and 1/1, respectively. Dark points represent missing genotypes. (b) Normalized depth for DEL within EP400 (top) and FRAS1 gene (bottom). (c) Private SVs of Yucatan pigs (left). Comparing with *Sscrofa*11.1, the white region showed that contig (chr8:73,892,000–73,893,000) from local assembling of 11 Yucatan pigs and 7 Suidae including 17 individuals within FRAS1 gene are missing. Other breeds in this study are all nondeletion type. Red and yellow rectangle represent the 261 bp SINE and 25 bp LINE, respectively.

that this gene might influence the development of mouse skin consistent with the Yucatan hairless phenotype living in the hot region (Pitera et al., 2012). A blebbled phenotype emerges in mice and humans because of the absence of the *FRAS1* gene in early stages; direct knockout of this gene in mice causes early death (Pitera et al., 2012). A 506 bp deletion (chr8:73,892,463–73,892,969) locate on an intron that is situated 65 bp apart from the downstream 48th exon of *FRAS1* covers a short interspersed transposable element (SINE) (chr8:73,892,642–73,892,942), which is possibly associated with the occurrence of this DEL. We also found a 284 bp deletion in *Sus verrucosus* and assembled another 7 outgroups (*Suidae*) (Table S9; Materials and Methods). The majority *Suidae* contains a

DEL about 299 bp (chr8:73,892,642–73,892,942), comparing 508 bp DEL (chr8:73,892,461–73,892,970) within *Sscrofa*11.1 (Figure 6c).

4 | DISCUSSION

In this study, we have presented a comprehensive porcine SV dataset obtained from diverse populations. A total of 112,890 SVs including 79,919 DELs, 23,638 DUPs, and 9333 INVs were identified, most of which (84.5%) had not been documented till date; a relatively large proportion of SVs was included in our study compared to those reported in a previous study (Zhao et al., 2016), demonstrating the

power of SV discovery in large and population-level samples, thus providing a new resource for future porcine SV studies.

In comparison with DUPS and INVs, DELs overlapping with CDS and UTR regions were under purifying selection, consistent with the observation in grapevines (Zhou et al., 2019). Remarkably, we observed that the SVs present in NR6A1 and PLAG1, which are reported to influence the height and body length in humans and livestock, have introgressed from the dominant allele of European commercial breeds of pigs into a part of Chinese domesticated breeds, potentially increasing the number of vertebrae and body length (Rubin et al., 2012; Takasuga, 2016). To the best of our knowledge, this is the firstly report related to the fact that PLAG1-SVs have introgressed from western pigs into indigenous Chinese pigs to improve their growth performance.

To date, the fate of SVs during porcine domestication has not yet been investigated. Here, we uncovered a positive selection of SV genes during domestication; we were particularly interested in ~9 kb duplication within the TBX19 gene, suggesting a potential human-mediated selection for timidity traits in captivity. We scanned genomes of pigs across different altitudes and temperatures to identify SVs under natural selection. As expected, the SV hotspot in the SSCX genes that were identified related to ancient introgression regarding adaptation to cold and hot environments (Ai et al., 2015). Simultaneously, we also found that multiple SV genes including MYBPC1, TCAIM, and SFRP1 are potentially involved in environmental adaptation.

With relatively high-genome coverage data obtained from 24 diverse populations of pigs, we explored the swine population-private SVs. Breeds from the European domestication center possess fewer private SV; moreover, breeds such as Landrace (LR) possess no private SVs under these criteria. Chinese indigenous breeds harbor the largest proportion private SVs, potentially contributing to unique characteristics such as good meat quality and strong resistance to diseases. We noted that although Yucatan pigs are genetically closer to Yorkshires and Landraces pigs to the SNP-based phylogenetic tree (Kim et al., 2015), this breed harbors large private genomic differences, which might be highly related to its characteristics. Among the Chinese indigenous pig breeds, the greatest number of private SVs was observed in Jinhua (JH) pigs, which have a typical two-end black coat color similar to that of Bamaxiang, Luchuan, and Shaziling pigs. Notably, except for Jinhua pigs, the two-end-black phenotype of the other Chinese indigenous breeds was linked to a causal variant in EDNRB gene (Huang et al., 2020), reflecting the uniqueness of Jinhua pigs.

Altogether, our population-level SV study highlights SVs as an important set of genomic variations to be investigated for their associations with evolution, selection, and variations in the complex traits of pigs. SVs, especially the highly differentiated SVs between indigenous pigs and wild boars are potentially associated with local environmental adaptation, or display allele frequency only in a particular population, providing valuable information for the ongoing endeavors on breed conservation in China. Meanwhile, we recognize the limitation and potential effect of high error rates of short

reads sequencing, especially on detecting the large and complex SVs, which could be better captured by new technologies such as third sequencing and optical mapping (Ho et al., 2019).

ACKNOWLEDGEMENTS

This work was supported by funds from the National Natural Science Foundation of China (31760657). We are grateful to colleagues from the lab for the supportive work.

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY STATEMENT

Publicly database in this study was downloaded from the NCBI Sequence Read Archive under accession SRA: PRJNA398176, PRJNA488327, PRJNA213179, PRJNA260763, PRJNA239399, PRJNA255085, and PRJEB1683.

ORCID

Bin Yang  <https://orcid.org/0000-0001-6773-4971>

Lusheng Huang  <https://orcid.org/0000-0002-6940-667X>

REFERENCES

- Abidi, F., Miano, M., Murray, J., & Schwartz, C. (2007). A novel mutation in the PHF8 gene is associated with X-linked mental retardation with cleft lip/cleft palate. *Clinical Genetics*, 72(1), 19–22.
- 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., ... Huang, L. (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nature Genetics*, 47(3), 217–225.
- Ai, H., Zhang, M., Yang, B., Goldberg, A., Li, W., Ma, J., Brandt, D., Zhang, Z., Nielsen, R., & Huang, L. (2021). Human-mediated admixture and selection shape the diversity on the modern swine (*Sus scrofa*) Y chromosomes. *Molecular Biology and Evolution*, 38(11), 5051–5065.
- Alkan, C., Coe, B. P., & Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, 12(5), 363–376. <https://doi.org/10.1038/nrg2958>
- Almarri, M. A., Bergström, A., Prado-Martinez, J., Yang, F., Fu, B., Dunham, A. S., Chen, Y., Hurles, M. E., Tyler-Smith, C., & Xue, Y. (2020). Population structure, stratification, and introgression of human structural variation. *Cell*, 182(1), 189–199.e15.
- Bargal, R., Cormier-Daire, V., Ben-Neriah, Z., Le Merrer, M., Sosna, J., Melki, J., Chen, Y., Hurles, M. E., Tyler-Smith, C., & Belostotsky, R. (2009). Mutations in DDR2 gene cause SMED with short limbs and abnormal calcifications. *The American Journal of Human Genetics*, 84(1), 80–84.
- Bhattacharjee, A., Yang, H., Duffy, M., Robinson, E., Conrad-Antoville, A., Lu, Y.-W., Capps, T., Braiterman, L., Wolfgang, M., Murphy, M. P., Yi, L., Kaler, S. G., Lutsenko, S., & Ralle, M. (2016). The activity of Menkes disease protein ATP7A is essential for redox balance in mitochondria. *Journal of Biological Chemistry*, 291(32), 16644–16658.
- Braybrook, C., Doudney, K., Marçano, A. C. B., Arnason, A., Bjornsson, A., Patton, M. A., Goodfellow, P. J., Moore, G. E., & Stanier, P. (2001). The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. *Nature Genetics*, 29(2), 179–183.
- Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association

- studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, 81(5), 1084–1097.
- Chaisson, M. J. P., Sanders, A. D., Zhao, X., Malhotra, A., Porubsky, D., Rausch, T., Gardner, E. J., Rodriguez, O. L., Guo, L., Collins, R. L., Fan, X., Wen, J., Handsaker, R. E., Fairley, S., Kronenberg, Z. N., Kong, X., Hormozdiari, F., Lee, D., Wenger, A. M., & Lee, C. (2019). Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nature Communications*, 10(1), 1784. <https://doi.org/10.1038/s41467-018-08148-z>
- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen, Z., & Manley, J. L. (2003). In vivo functional analysis of the histone 3-like TAF9 and a TAF9-related factor, TAF9L. *Journal of Biological Chemistry*, 278(37), 35172–35183.
- Chen, Z., Zhao, T.-J., Li, J., Gao, Y.-S., Meng, F.-G., Yan, Y.-B., & Zhou, H.-M. (2011). Slow skeletal muscle myosin-binding protein-C (MyBPC1) mediates recruitment of muscle-type creatine kinase (CK) to myosin. *Biochemical Journal*, 436(2), 437–445.
- Chiang, C., Scott, A. J., Davis, J. R., Tsang, E. K., Li, X., Kim, Y., Hadzic, T., Damani, F. N., Ganel, L., GTEx Consortium, Montgomery, S. B., Battle, A., Conrad, D. F., & Hall, I. M. (2017). The impact of structural variation on human gene expression. *Nature Genetics*, 49(5), 692–699.
- Cho, I. C., Park, H. B., Ahn, J. S., Han, S. H., Lee, J. B., Lim, H. T., Yoo, C. K., Jung, E. J., Kim, D. H., Sun, W. S., Ramayo-Caldas, Y., Kim, S. G., Kang, Y. J., Kim, Y. K., Shin, H. S., Seong, P. N., Hwang, I. S., Park, B. Y., Hwang, S., ... Lee, J. W. (2019). A functional regulatory variant of MYH3 influences muscle fiber-type composition and intramuscular fat content in pigs. *PLoS Genetics*, 15(10), e1008279. <https://doi.org/10.1371/journal.pgen.1008279>
- Churchill, G. A., & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics*, 138(3), 963–971.
- Cohen, T. J., Waddell, D. S., Barrientos, T., Lu, Z., Feng, G., Cox, G. A., Bodine, S. C., & Yao, T.-P. (2007). The histone deacetylase HDAC4 connects neural activity to muscle transcriptional reprogramming. *Journal of Biological Chemistry*, 282(46), 33752–33759.
- Couture, J.-P., Nolet, G., Beaulieu, E., Blouin, R., & Gévr, N. (2012). The p400/Brd8 chromatin remodeling complex promotes adipogenesis by incorporating histone variant H2A Z at PPAR γ target genes. *Endocrinology*, 153(12), 5796–5808.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.
- Dhoot, G. K., & Perry, S. V. (2005). Expression of slow skeletal myosin binding C-protein in normal adult mammalian heart. *Journal of Muscle Research and Cell Motility*, 26(2–3), 143–148.
- Du, H., Zheng, X., Zhao, Q., Hu, Z., Wang, H., Zhou, L., & Liu, J.-F. (2021). Analysis of structural variants reveal novel selective regions in the genome of Meishan pigs by whole genome sequencing. *Frontiers in Genetics*, 12, 99.
- Eisermann, K., Wang, D., Jing, Y., Pascal, L. E., & Wang, Z. (2013). Androgen receptor gene mutation, rearrangement, polymorphism. *Translational Andrology and Urology*, 2(3), 137.
- Escofflaire, C., Rebours, E., Charles, M., Orellana, S., Cano, M., Rivière, J., Grohs, C., Hayes, H., & Capitan, A. (2019). A de novo 3.8-Mb inversion affecting the EDA and XIST genes in a heterozygous female calf with generalized hypohidrotic ectodermal dysplasia. *BMC Genomics*, 20(1), 715.
- Fontanesi, L., Davoli, R., Dall'Olio, S., & Russo, V. (2000). Linkage assignment of the fast skeletal alkali myosin light polypeptide 1 (MYL1) gene to porcine chromosome 15. *Animal Genetics*, 31(6), 415–416.
- Frantz, L. A., Schraiber, J. G., Madsen, O., Megens, H.-J., Bosse, M., Paudel, Y., Semidi, G., Meijaard, E., Li, N., Crooijmans, R. P., Archibald, A. L., Slatkin, M., Schook, L. B., Larson, G., & Groenen, M. A. (2013). Genome sequencing reveals fine scale diversification and reticulation history during speciation in Sus. *Genome Biology*, 14(9), R107.
- Frantz, L. A., Schraiber, J. G., Madsen, O., Megens, H. J., Cagan, A., Bosse, M., Paudel, Y., Crooijmans, R. P., Larson, G., & Groenen, M. A. (2015). Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nature Genetics*, 47(10), 1141–1148. <https://doi.org/10.1038/ng.3394>
- Georges, M., Charlier, C., & Hayes, B. (2019). Harnessing genomic information for livestock improvement. *Nature Reviews. Genetics*, 20(3), 135–156. <https://doi.org/10.1038/s41576-018-0082-2>
- Glenn, K., Ramos, A., & Rothschild, M. (2007). Analysis of FMO genes and off flavour in pork. *Journal of Animal Breeding and Genetics*, 124(1), 35–38.
- Gong, H., Xiao, S., Li, W., Huang, T., Huang, X., Yan, G., Huang, Y., Qiu, H., Jiang, K., Wang, X., Zhang, H., Tang, J., Li, L., Li, Y., Wang, C., Qiao, C., Ren, J., Huang, L., & Yang, B. (2019). Unravelling the genetic loci for growth and carcass traits in Chinese Baamaxiang pigs based on a 1.4 million SNP array. *Journal of Animal Breeding and Genetics*, 136(1), 3–14.
- Groenen, M. A., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H. J., Li, S., Larkin, D. M., Kim, H., Frantz, L. A., Caccamo, M., Ahn, H., Aken, B. L., Anselmo, A., Anthon, C., & Schook, L. B. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, 491(7424), 393–398. <https://doi.org/10.1038/nature11622>
- Ho, S. S., Urban, A. E., & Mills, R. E. (2019). Structural variation in the sequencing era. *Nature Reviews. Genetics*, 21(3), 171–189. <https://doi.org/10.1038/s41576-019-0180-9>
- Huang, M., Yang, B., Chen, H., Zhang, H., Wu, Z., Ai, H., Ren, J., & Huang, L. (2020). The fine-scale genetic structure and selection signals of Chinese indigenous pigs. *Evolutionary Applications*, 13(2), 458–475.
- Jamal, S. M., Basran, R. K., Newton, S., Wang, Z., & Milunsky, J. M. (2010). Novel de novo PCDH19 mutations in three unrelated females with epilepsy female restricted mental retardation syndrome. *American Journal of Medical Genetics Part A*, 152(10), 2475–2481.
- János, M., Nagy, T., Stéger, V., Tóth, G., Marincs, F., & Barta, E. (2014). Genome sequencing and analysis of Mangalica, a fatty local pig of Hungary. *BMC Genomics*, 15, 761.
- Johansson, M., Ellegren, H., Marklund, L., Gustavsson, U., Ringmar-Cederberg, E., Andersson, K., Edfors-Lilja, I., & Andersson, L. (1992). The gene for dominant white color in the pig is closely linked to ALB and PDGFRA on chromosome 8. *Genomics*, 14(4), 965–969.
- Karim, L., Takeda, H., Lin, L., Druet, T., Arias, J. A., Baurain, D., Cambisano, N., Davis, S. R., Farnir, F., Grisart, B., Harris, B. L., Keehan, M. D., Littlejohn, M. D., Spelman, R. J., Georges, M., & Coppelters, W. (2011). Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. *Nature Genetics*, 43(5), 405–413.
- Khosronezhad, N., Colagar, A. H., & Jorsarayi, S. G. A. (2015). T26248G-transversion mutation in exon7 of the putative methyltransferase Nsun7 gene causes a change in protein folding associated with reduced sperm motility in asthenospermic men. *Reproduction, Fertility and Development*, 27(3), 471–480.
- Kim, H., Song, K. D., Kim, H. J., Park, W., Kim, J., Lee, T., Shin, D. H., Kwak, W., Kwon, Y. J., Sung, S., Moon, S., Lee, K. T., Kim, N., Hong, J. K., Eo, K. Y., Seo, K. S., Kim, G., Park, S., Yun, C. H., ... Kim, H. (2015). Exploring the genetic signature of body size in Yucatan miniature pig. *PLoS One*, 10(4), e0121732.
- Kou, Y., Liao, Y., Toivainen, T., Lv, Y., Tian, X., Emerson, J., Gaut, B. S., & Zhou, Y. (2020). Evolutionary genomics of structural variation in Asian rice (*Oryza sativa*) domestication. *Molecular Biology and Evolution*, 37(12), 3507–3524.

- Labrador, J. P., Azcoitia, V., Tuckermann, J., Lin, C., Olaso, E., Mañes, S., Brückner, K., Goergen, J. L., Lemke, G., Yancopoulos, G., Angel, P., Martínez, C., & Klein, R. (2001). The collagen receptor DDR2 regulates proliferation and its elimination leads to dwarfism. *EMBO Reports*, 2(5), 446–452.
- Laptev, I., Shvetsova, E., Levitskii, S., Serebryakova, M., Rubtsova, M., Bogdanov, A., Kamenski, P., Sergiev, P., & Dontsova, O. (2020). Mouse Trmt2B protein is a dual specific mitochondrial methyltransferase responsible for m5U formation in both tRNA and rRNA. *RNA Biology*, 17(4), 441–450.
- Larson, D. E., Abel, H. J., Chiang, C., Badve, A., Das, I., Eldred, J. M., Layer, R. M., & Hall, I. M. (2019). svtools: Population-scale analysis of structural variation. *Bioinformatics*, 35(22), 4782–4787. <https://doi.org/10.1093/bioinformatics/btz492>
- Layer, R. M., Chiang, C., Quinlan, A. R., & Hall, I. M. (2014). LUMPY: A probabilistic framework for structural variant discovery. *Genome Biology*, 15, R84.
- Lee, Y. L., Takeda, H., Costa Monteiro Moreira, G., Karim, L., Mullaart, E., Coppelters, W., GplusE Consortium, Appeltant, R., Veerkamp, R. F., MAM, G., Georges, M., Bosse, M., Druet, T., Bouwman, A. C., & Charlier, C. (2021). A 12 kb multi-allelic copy number variation encompassing a GC gene enhancer is associated with mastitis resistance in dairy cattle. *PLoS Genetics*, 17(7), e1009331.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
- Liao, M., Shi, J., Huang, L., Gao, Y., Tan, A., Wu, C., Lu, Z., Yang, X., Zhang, S., Hu, Y., Qin, X., Li, J., Chen, G., Xu, J., Mo, Z., & Zhang, H. (2014). Genome-wide association study identifies variants in PMS1 associated with serum ferritin in a Chinese population. *PLoS One*, 9(8), e105844.
- Lim, J.-H., Seo, E.-J., Kim, Y.-M., Cho, H.-J., Lee, J.-O., Cheon, C. K., & Yoo, H.-W. (2014). A de novo microdeletion of ANKRD11 gene in a Korean patient with KBG syndrome. *Annals of Laboratory Medicine*, 34(5), 390–394.
- Lossi, A., Laugier-Anfossi, F., Depetris, D., Gecz, J., Gedeon, A., Kooy, F., Schwartz, C., Mattei, M. G., Croquette, M. F., & Villard, L. (2002). Abnormal expression of the KLF8 (ZNF741) gene in a female patient with an X; autosome translocation t(X; 21)(p11. 2; q22. 3) and non-syndromic mental retardation. *Journal of Medical Genetics*, 39(2), 113–117.
- Lucek, K., Gompert, Z., & Nosil, P. (2019). Genomics studies on the nature of species: Adaptation and speciation in Mimulus. *Molecular Ecology*, 28(6), 2601–2609.
- Luo, K., Li, J., Cui, Y., Xu, M., Yuan, J., Tang, W., Wan, B., & Yu, L. (2006). Identification and characterization of the human SCAN domain zinc-finger gene ZNF449. *Molecular Biology Reports*, 33(1), 51–57.
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., ... Wang, J. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience*, 1(1), 18.
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., & Cunningham, F. (2016). The ensembl variant effect predictor. *Genome Biology*, 17(1), 122.
- Moortgat, S., Berland, S., Aukrust, I., Maystadt, I., Baker, L., Benoit, V., Caro-Llopis, A., Cooper, N. S., Debray, F. G., Faivre, L., Gardeitchik, T., Haukanes, B. I., Houge, G., Kivuva, E., Martinez, F., Mehta, S. G., Nassogne, M. C., Powell-Hamilton, N., Pfundt, R., ... Newbury-Ecob, R. A. (2018). HUWE1 variants cause dominant X-linked intellectual disability: a clinical study of 21 patients. *European Journal of Human Genetics*, 26(1), 64–74.
- Omisano, O. A., Biermann, K., Hartmann, S., Heukamp, L. C., Sonnack, V., Hild, A., Brehm, R., Bergmann, M., Weidner, W., & Steger, K. (2007). DNMT1 and HDAC1 gene expression in impaired spermatogenesis and testicular cancer. *Histochemistry and Cell Biology*, 127(2), 175–181.
- Peng, C., Sun, G., Tang, Z., & Hou, X. (2020). Congenital isolated ACTH deficiency caused by TBX19 gene mutation: A family report. *Frontiers in Pediatrics*, 7, 546.
- Petit, F., Andrieux, J., Holder-Espinasse, M., Bouquillon, S., Pennaforte, T., Storme, L., & Manouvrier-Hanu, S. (2011). Xq12q13. 1 microduplication encompassing the EFNB1 gene in a boy with congenital diaphragmatic hernia. *European Journal of Medical Genetics*, 54(5), e525–e527.
- Pitera, J. E., Turmaine, M., Woolf, A. S., & Scambler, P. J. (2012). Generation of mice with a conditional null Fraser syndrome 1 (Fras1) allele. *Genesis*, 50(12), 892–898. <https://doi.org/10.1002/dvg.22045>
- Pondarré, C., Antiochos, B. B., Campagna, D. R., Clarke, S. L., Greer, E. L., Deck, K. M., McDonald, A., Han, A. P., Medlock, A., Kutok, J. L., Anderson, S. A., Eisenstein, R. S., & Fleming, M. D. (2006). The mitochondrial ATP-binding cassette transporter Abcb7 is essential in mice and participates in cytosolic iron–sulfur cluster biogenesis. *Human Molecular Genetics*, 15(6), 953–964.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Qiao, Y., Liu, X., Harvard, C., Hildebrand, M., Rajcan-Separovic, E., Holden, J., & Lewis, M. (2008). Autism-associated familial microdeletion of Xp11. 22. *Clinical Genetics*, 74(2), 134–144.
- Quan, C., Li, Y., Liu, X., Wang, Y., Ping, J., Lu, Y., & Zhou, G. (2021). Characterization of structural variation in Tibetans reveals new evidence of high-altitude adaptation and introgression. *Genome Biology*, 22(1), 159.
- Ramírez, O., Burgos-Paz, W., Casas, E., Ballester, M., Bianco, E., Olalde, I., Santpere, G., Novella, V., Gut, M., Lalueza-Fox, C., Saña, M., & Pérez-Enciso, M. (2014). Genome data from a sixteenth century pig illuminate modern breed relationships. *Heredity*, 114(2), 175–184. <https://doi.org/10.1038/hdy.2014.81>
- Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S. R., WGS500 Consortium, Wilkie, A. O. M., McVean, G., & Lunter, G. (2014). Integrating mapping-, assembly-and haplotype-based approaches for calling variants in clinical sequencing applications. *Nature Genetics*, 46(8), 912–918.
- Rubin, C.-J., Megens, H.-J., Barrio, A. M., Maqbool, K., Sayyab, S., Schwochow, D., Wang, C., Carlborg, Ö., Jern, P., Jørgensen, C. B., Archibald, A. L., Fredholm, M., Groenen, M. A., & 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.
- Senol-Cosar, O., Flach, R. J. R., DiStefano, M., Chawla, A., Nicoloro, S., Straubhaar, J., Hardy, O. T., Noh, H. L., Kim, J. K., Wabitsch, M., Scherer, P. E., & Czech, M. P. (2016). Tenomodulin promotes human adipocyte differentiation and beneficial visceral adipose tissue expansion. *Nature Communications*, 7(1), 10686.
- Shinawi, M., Sahoo, T., Maranda, B., Skinner, S., Skinner, C., Chinault, C., Zascavage, R., Peters, S. U., Patel, A., Stevenson, R. E., & Beaudet, A. L. (2011). 11p14. 1 microdeletions associated with ADHD, autism, developmental delay, and obesity. *American Journal of Medical Genetics Part A*, 155(6), 1272–1280.
- Startin, C. M., Fiorentini, C., de Haan, M., & Skuse, D. H. (2015). Variation in the X-linked EFHC2 gene is associated with social cognitive abilities in males. *PLoS One*, 10(6), e0131604.
- Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Fritz, M. H., Konkel, M. K., Malhotra, A., Stütz, A. M., Shi, X., Casale, F. P., Chen, J., Hormozdiari, F., & Durbin, R. (2015). Global variation in human X-linked haplotypes. *Nature*, 526(7572), 240–244.

- F., Dayama, G., Chen, K., ... Korbel, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*, 526(7571), 75–81.
- Takasuga, A. (2016). PLAG1 and NCAPG-LCRTL in livestock. *Animal Science Journal*, 87(2), 159–167.
- Tang, D. G., & Conti, C. J. (2004). Endothelial cell development, vasculogenesis, angiogenesis, and tumor neovascularization: An update. Paper presented at the Seminars in thrombosis and hemostasis.
- Torrico, B., Fernández-Castillo, N., Hervás, A., Milà, M., Salgado, M., Rueda, I., Buitelaar, J. K., Rommelse, N., Oerlemans, A. M., Bralten, J., Freitag, C. M., Reif, A., Battaglia, A., Mazzone, L., Maestrini, E., Cormand, B., & Toma, C. (2015). Contribution of common and rare variants of the PTCHD1 gene to autism spectrum disorders and intellectual disability. *European Journal of Human Genetics*, 23(12), 1694–1701.
- Tretyakova, N. Y., Groehler, A., IV, & Ji, S. (2015). DNA–protein cross-links: Formation, structural identities, and biological outcomes. *Accounts of Chemical Research*, 48(6), 1631–1644.
- Vogel, S. Z., Schlickeiser, S., Jürchott, K., Akyuez, L., Schumann, J., Appelt, C., Vogt, K., Schröder, M., Vaeth, M., Berberich-Siebelt, F., Lutz, M. B., Grütz, G., & Sawitzki, B. (2015). TCAIM decreases T cell priming capacity of dendritic cells by inhibiting TLR-induced Ca²⁺ influx and IL-2 production. *The Journal of Immunology*, 194(7), 3136–3146.
- Wang, G.-D., Shao, X.-J., Bai, B., Wang, J., Wang, X., Cao, X., Liu, Y. H., Wang, X., Yin, T. T., Zhang, S. J., Lu, Y., Wang, Z., Wang, L., Zhao, W., Zhang, B., Ruan, J., & Zhang, Y. P. (2019). Structural variation during dog domestication: Insights from gray wolf and dhole genomes. *National Science Review*, 6(1), 110–122.
- Wang, L., Wang, A., Wang, L., Li, K., Yang, G., He, R., & Pan, Y. (2011). China National Commission of animal genetic resources. *Animal genetic resources in China*. Beijing: Chinese Agriculture Press. [Google Scholar].
- Wang, X., Zhang, H., Huang, M., Tang, J., Yang, L., Yu, Z., Li, D., Li, G., Jiang, Y., Sun, Y., Wei, S., Xu, P., & Ren, J. (2021). Whole-genome SNP markers reveal conservation status, signatures of selection, and introgression in Chinese Laiwu pigs. *Evolutionary Applications*, 14(2), 383–398.
- Weber, T. A., Koob, S., Heide, H., Wittig, I., Head, B., van der Bliek, A., Brandt, U., Mittelbronn, M., & Reichert, A. S. (2013). APOOL is a cardiolipin-binding constituent of the Mitoflin/MINOS protein complex determining cristae morphology in mammalian mitochondria. *PLoS One*, 8(5), e63683.
- Weischenfeldt, J., Symmons, O., Spitz, F., & Korbel, J. O. (2013). Phenotypic impact of genomic structural variation: Insights from and for human disease. *Nature Reviews. Genetics*, 14(2), 125–138. <https://doi.org/10.1038/nrg3373>
- Wellenreuther, M., Merot, C., Berdan, E., & Bernatchez, L. (2019). Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Molecular Ecology*, 28(6), 1203–1209. <https://doi.org/10.1111/mec.15066>
- White, S. (2011). From globalized pig breeds to capitalist pigs: A study in animal cultures and evolutionary history. *Environmental History*, 16(1), 94–120.
- Wu, Y., Davies, K. E., & Oliver, P. L. (2016). The antioxidant protein Oxr1 influences aspects of mitochondrial morphology. *Free Radical Biology and Medicine*, 95, 255–267.
- Wu, Z., Gong, H., Zhang, M., Tong, X., Ai, H., Xiao, S., Perez-Enciso, M., Yang, B., & Huang, L. (2021). A worldwide map of swine short tandem repeats and their associations with evolutionary and environmental adaptations. *Genetics Selection Evolution*, 53(1), 39.
- Yan, S. M., Sherman, R. M., Taylor, D. J., Nair, D. R., Bortvin, A. N., Schatz, M. C., & McCoy, R. C. (2021). Local adaptation and archaic introgression shape global diversity at human structural variant loci. *eLife*, 10, e67615.
- Yang, B., Cui, L., Perez-Enciso, M., Traspov, A., Crooijmans, R. P., Zinovieva, N., Schook, L. B., Archibald, A., Gatphayak, K., Knorr, C., Triantafyllidis, A., Alexandri, P., Semiadi, G., Hanotte, O., Dias, D., Dovč, P., Uimari, P., Iacolina, L., Scandura, M., ... Megens, H. J. (2017). Genome-wide SNP data unveils the globalization of domesticated pigs. *Genetics Selection Evolution*, 49(1), 71.
- Yang, H., Wu, J., Huang, X., Zhou, Y., Zhang, Y., Liu, M., Liu, Q., Ke, S., He, M., Fu, H., Fang, S., Xiong, X., Jiang, H., Chen, Z., Wu, Z., Gong, H., Tong, X., Huang, Y., Ma, J., ... Huang, L. (2022). ABO genotype alters the gut microbiota by regulating GalNAc levels in pigs. *Nature*, 606, 358–367.
- Zhang, T., Gao, H., Sahana, G., Zan, Y., Fan, H., Liu, J., Shi, L., Wang, H., Du, L., Wang, L., & Zhao, F. (2019). Genome-wide association studies revealed candidate genes for tail fat deposition and body size in the Hulun Buir sheep. *Journal of Animal Breeding and Genetics*, 136(5), 362–370.
- Zhao, P., Li, J., Kang, H., Wang, H., Fan, Z., Yin, Z., Wang, J., Zhang, Q., Wang, Z., & Liu, J.-F. (2016). Structural variant detection by large-scale sequencing reveals new evolutionary evidence on breed divergence between Chinese and European pigs. *Scientific Reports*, 6, 18501.
- Zhao, X., Hu, H., Wang, C., Wang, Y., Lin, H., & Wang, J. (2019). Transcriptome analysis of muscle reveals potential candidate genes and pathways affecting intramuscular fat content in Duroc pigs.
- Zhou, R., Li, S. T., Yao, W. Y., Xie, C. D., Chen, Z., Zeng, Z. J., Wang, D., Xu, K., Shen, Z. J., Mu, Y., Bao, W., Jiang, W., Li, R., Liang, Q., & Li, K. (2021). The Meishan pig genome reveals structural variation-mediated gene expression and phenotypic divergence underlying Asian pig domestication. *Molecular Ecology Resources*, 21(6), 2077–2092.
- Zhou, Y., Minio, A., Massonet, M., Solares, E., Lv, Y., Beridze, T., Cantu, D., & Gaut, B. S. (2019). The population genetics of structural variants in grapevine domestication. *Nature Plants*, 5(9), 965–979.
- Zhu, Y., Li, W., Yang, B., Zhang, Z., Ai, H., Ren, J., & Huang, L. (2017). Signatures of selection and interspecies introgression in the genome of Chinese domestic pigs. *Genome Biology*, 9(10), 2592–2603.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gong, H., Liu, W., Wu, Z., Zhang, M., Sun, Y., Ling, Z., Xiao, S., Ai, H., Xing, Y., Yang, B., & Huang, L. (2022). Evolutionary insights into porcine genomic structural variations based on a novel-constructed dataset from 24 worldwide diverse populations. *Evolutionary Applications*, 15, 1264–1280. <https://doi.org/10.1111/eva.13455>