Elucidating breed-specific variants of native pigs in Korea: insights into pig breeds’ genomic characteristics

Young-Sup Lee¹†, Seungwoo Son¹†, Hak-Kyo Lee¹,b, Ra Ham Lee¹ and Donghyun Shinb

¹Department of Animal Biotechnology, Jeonbuk National University, Jeonju, Republic of Korea; bDepartment of Agricultural Convergence Technology, Jeonbuk National University, Jeonju, Republic of Korea

ABSTRACT
Although conserving native pig breeds is important in Korea, research on the genomic aspects to identify breed-specific variations in native pig breeds is uncommon. Single nucleotide polymorphisms (SNPs) can be a powerful source for identifying breed-specific variants. We used whole genome sequencing data, including Jeju Native Pig (JNP), Korean Native Pig (KNP), Korean Wild Boar (KWB), and other western commercial pig breeds to determine native pig breed-specific SNPs. Furthermore, the goal was not only to determine the genomic specificity of native pig breeds but also to identify SNPs that carry breed-specific information (breed-informative SNPs) that can be related to breed characteristics. The representative characteristics of native pigs are their unique meat quality and disease resistance. We surveyed the gene ontology (GO) of native pigs with breed-specific SNPs. Examining the genes associated with GO terms were neuron projection development, cell surface receptor signaling pathway, ion homeostasis in JNP, cell adhesion and wound healing in KNP, and DNA repair and reproduction in KWB. We expect that this study of breed-specific SNPs will enable us to gain a deeper understanding of native pigs in Korea.

ARTICLE HISTORY
Received 1 August 2022
Revised 20 October 2022
Accepted 24 October 2022

KEYWORDS
Breed-specific SNPs; breed-informative; genomic characteristics; native pigs; whole genome sequencing

Introduction
Pigs were first domesticated in the Near East around 8500 BC and were brought into Europe by agriculturalists (Larson et al. 2007). Evidence supports the independent domestication of pigs in Asia and Europe from the wild boar subspecies (Giuffra et al. 2000). The answer to which pig domestication was independent or connected by diffusion from a single origin remains to be established, although recent research based on complete mitochondrial genomes of East Asian pigs suggests that wild boar in China may have been domesticated independently in the Mekong and in the middle to downstream Yangtze regions (Larson et al. 2010). Recently, western pig breeds have been imported into the Korean peninsula and crossed with native pigs in Korea to improve their growth and carcass-related traits, which has led to a dramatic drop in the native pig population. Although commercial pig breeds are superior in terms of growth and feed efficiency traits, native pigs in Korea are known for their better meat quality (high redness and intramuscular fat) and greater ability to thrive under low management conditions (Yeo et al. 2000, Kim et al. 2005).

Single nucleotide polymorphisms (SNPs) have been used to describe genetic variability and identify phenotype-associated candidate genes. Furthermore, SNPs can capture breed differences and be used to determine the breed of each pig. Informative SNPs have been used for comparative selection signature analyses, parentage assessment, and breed assignment of individual animals, as well as for several other applications (Kuehn et al. 2011, Asaf et al. 2014). The genomic diversity of specific breeds with specific phenotypic and thus genotypic characteristics can be uncovered in some ways (Czech et al. 2018). Determining breed-specific variants is one method for uncovering genetic characteristics. Breed-specific variants, including SNPs and nonsynonymous SNPs (nsSNPs), can not only differentiate the target breeds from other breeds but also reveal their characteristics. More specifically, breed-specific nsSNPs
alter the amino acid sequence and protein function and can play a role in delineating breed characteristics (Lee, Won et al. 2020).

The native pigs in Korea (JNP: Jeju native black pig, KNP: Korean native pig, KWB: wild boar in Korea) differ from the Western commercial breeds in appearance; their fur is black, the ears are folded forward, and are resistant to disease. They have a unique meat quality, which Koreans prefer. Specifically, JNP, a native pig breed in Korea, has survived for a long time without mixing with other pig breeds, in an independent environment away from inland. Since the 1960s, the number of JNPs has sharply declined, putting them on the verge of extinction. Since 1986, the pure line breeding project has been initiated, and ~260 black pigs have been preserved and managed. KWB has a maximum body length of 180 cm and weight ranging 50–300 kg. Their legs are short and slender, and they run fast. Body color varies widely from dark brown to light brown. To reveal the unique characteristics of Korea’s native pig breeds, we investigated the specific variants of the three breeds to determine the genomic cause of their characteristics. The breed-specific variants can be breed-informative. Kumar et al. attempted to determine cow breed-specific SNPs in India that are associated with the breeds’ characteristics, such as milk production, heat tolerance, disease resistance, and fertility (Kumar et al. 2021). Ramos et al. developed breed-specific SNPs in pig breeds, presenting breed-specific markers for breed assignment and traceability (Ramos et al. 2011).

Purebred pigs are commercially important, and many breeders have requested them in their cross-breeding programs. Cross-breeding helps breeders discover new breeds with desirable traits, such as disease resistance and heat tolerance (Pasupa et al. 2020). When crossing native pigs in Korea with western commercial pig breeds, we aimed to determine the breed-specific variants of native pigs in Korea, which could help improve the conservation of the genetic resources of native breeds. We used whole genome sequencing data from native pig breeds, western commercial pig breeds, and variant information. The breed-specific SNPs and breed-specific nonsynonymous SNPs (nsSNPs) of native pig breeds were identified, and the characteristics of the SNPs were surveyed using gene ontology (GO) analysis.

Materials and methods

Whole genome sequencing

The pig fastq sequence data—Berkshire (BKS) 10, Duroc (DUR) 20, Jeju Native Pig (JNP) 20, Korean native pig (KNP) 6, Korean wild boar (KWB) 10, landrace (LDR) 13, Yorkshire (YKS) 15, and Yucatan Miniature Pig (YMP) 12—were checked using the FastQC software. Potential adapter sequences were removed using Trimomatic-0.32 (Bolger et al. 2014), followed by mapping paired-end sequence reads to the pig reference genome (Sscrofa 11.1) obtained from Ensembl. Picard tools (http://broadinstitute.github.io/picard/), SAMtools (Li and Durbin 2009), and the Genome Analysis Toolkit (GATK) (McKenna et al. 2010) were used for downstream processing and variant calling. Using ‘CreateSequence-Dictionary’ and ‘MarkDuplicates’ Picard command-line tools, we read reference FASTA sequences to generate bam files with only a sequence dictionary and filter potential PCR duplicates. The index files for the reference and BAM files were created using SAMtools. Using the GATK ‘Realigner-TargetCreator’ and ‘IndelRealigner’ arguments, local realignment of the sequence reads was performed to remove small insertions and deletions. Additionally, the base quality score was recalibrated to obtain accurate quality scores and correct for the variation in quality with the machine cycle and sequence context. To call variants, GATK ‘UnifiedGenotyper’ and ‘SelectVariants’ arguments were used with the following standards: 1) Phred-scaled quality score of less than 30, 2) read depth less than 5, 3) MQ0 (total count of mapping quality zero reads across all samples) > 4, and 4) a Phred-scaled P-value using Fisher’s exact test of more than 200 were filtered out to reduce false-positive calls due to strand bias.

The SNPs dataset (37,823,235) was imputed using Beagle 5.3 (Browning and Browning 2007), and QC (Quality Check) was performed with minor allele frequency (MAF < 0.05) and Hardy-Weinberg Equilibrium (HWE p-value < 1.0E-06), leaving 16,868,444 SNPs. Figures were drawn using the data after QC. Nevertheless, to determine the breed-specific SNPs, untreated data was used before QC. The SNPs were annotated using the SNP annotation tool SnpEff (version 4.1) (Cingolani et al. 2012).

Examining the population structure of native pig breeds in Korea

To verify the purity of each pig breed, we analyzed its genomic structures. The selection of JNP-, KNP-, and KWB-breed-specific SNPs should be based on the purity of each breed. Admixture analysis to check the purity of the pig breeds was performed using fastStructure, based on a variational Bayesian framework, and a genetic cluster of size 8 (Raj et al. 2014) was estimated. Breeds were classified as pure for the most part, but 5 JNP individuals were mixed with the DUR breed.
Therefore, we excluded five individuals from the subsequent analysis.

As a dimension reduction method, principal component analysis (PCA) combines multiple features into lower-dimensional features, that can explain most of the variance in a large dataset (Lee et al. 2021). PCA was conducted using genome-wide complex trait analysis (GCTA) to obtain eigenvalues and eigenvectors (Yang et al. 2011) and to assess the explanatory power of each breed-specific SNPs, which revealed that breed-specific SNPs can distinguish them from other breeds. The neighbor-joining tree (NJ tree) was constructed using a distance matrix calculated using VCF2Dis (https://github.com/BGI-shenzhen/VCF2Dis) and MEGA6 (https://megasoftware.net/) (Tamura et al. 2013). Treemix v1.13 was used to identify historical relationships among pig populations (Fitak 2021).

Identifying breed-specific SNPs

Let $A'$ be the allele count in the object population, and $A$ be the allele count in the reference population. The ratio $\frac{A'}{A}$ is a measure of breed-specific SNPs. If the ratio is greater than 0.9, the SNPs are considered breed-specific.

$$\frac{A'}{A} > 0.9$$

where $A'$ is the allele count in the object population, and $A$ is the allele count in the reference population.

Thus, breed-specific SNPs that overlapped with nonsynonymous SNPs (nsSNPs) were used for the subsequent analysis. nsSNPs are an important type of SNPs that alter amino acid sequences and potentially affect protein structure and function (Krawczak et al. 2000, Wu and Jiang 2013). These breed-specific nsSNPs are likely to link pig phenotypes to genomic variations. To predict whether SNPs were nonsynonymous, we used the SnpEff program with the reference genome version Sscrofa 11.1. SnpEff is a variant annotation and effect prediction tool that is used to identify differences like amino acid changes (Cingolani, Platts, Wang, Coon, Nguyen, Wang, Land, Lu and Ruden 2012).

**GO analysis**

We analyzed the GO using breed-specific SNPs. For JNP, genes encompassing breed-specific SNPs were used for GO analysis. Genes encompassing breed-specific nsSNPs were selected for KNP and KWB. The gene catalogue was retrieved from the Ensembl DB (http://www.ensembl.org) and Database for Annotation, Visualization, and Integrated Discovery (DAVID 2021). A list of gene identifiers was uploaded to summarize the functional annotations associated with groups or each individual gene, and each biological process (BP)-related GO term was based on the number of genes and p-values (Huang et al. 2007).

**Results**

**SNP data description**

The number of SNPs per autosome (chromosome 1–18) ranged from 1,056,274–3,493,244. The number of nsSNPs ranged from 2,868–11,507, while the number of SNPs and nsSNPs per autosome tended to co-vary (Figure 1A). The average distance between adjacent SNPs per autosome (± standard deviation [SD]) ranged from 100 (± 181) to 186 (± 291). The number of rs-tagged nsSNPs ranged from 1,687–7,754, and the number of novel nsSNPs ranged from 551–3,753 (Figure 1B).

**Population stratification**

Structural analysis using total SNPs and setting the parameters K to 6, 7, and 8 revealed that some JNPs were mixed with DUR in five JNP individuals; thus, these five JNP pigs were eliminated from the analysis. After elimination, the purity of each breed was verified by structural analysis. Analyses with the parameter K = 4, using total SNPs after QC (upper panel) and breed-specific SNPs

| Breed-specific nonsynonymous SNPs (nsSNPs) |
|------------------------------------------|
| In JNP, breed-specific SNPs and encompassing genes were used for further analyses. Enormous SNPs were not detected as breed-specific SNPs in comparison with other native pigs. However, numerous breed-specific SNPs have been detected in the KNP and KWB. |
were performed to check the representativeness of breed-specific SNPs (Figure 1C). In the NJ tree, each native pig breed (JNP, KNP, and KWB) was distinct from other commercial pig breeds (Figure 1D). For Treemix, the results were congruent with those of the NJ tree (Figure 1E). The plot was drawn using Treemix program and could show the historical relationships among the populations. It showed that JNP, KNP and KWB were separate with western commercial pig breeds.

**Breed-specific SNPs**

To identify the genomic characteristics of the native pig breeds in Korea (JNP, KNP, and KWB), breed-specific SNPs and nsSNPs were surveyed. Most breed-specific SNPs in JNP, KNP, and KWB were from alternative alleles, indicating the divergence of native pig breeds from wild boar ancestors. The three SNPs encompassed in neuron navigator 1 (NAV1) were JNP-specific SNPs that showed the unity of JNP alleles (Figure 2A). The Venn diagram illustrating breed-specific SNPs and nsSNPs showed that ~0.3–2% of the total SNPs could be classified as breed-specific. The number of breed-specific SNPs in JNP, KNP, and KWB was 103,992, 196,180, and 687,201, respectively. The number of breed-specific nsSNPs was 227, 391, and 1,244, respectively. In particular, breed-specific SNPs and nsSNPs of JNP did not overlap with those of KNP and KWB (Figure 2B and C). It is likely that the JNP diverged well from the wild boars in Korea. PCA analysis showed that the breed-specific SNPs in JNP (Figure 3A) and the breed-specific nsSNPs in KNP and KWB (Figure 3B and C) were distinct from other pig breeds.

**GO analysis for native pig breeds in Korea**

Gene ontology analysis of the genes encompassing JNP-specific SNPs revealed that neuron projection development (GO:0031175), cell surface receptor signaling pathway (GO:0007166), and ion homeostasis (GO:0050801) were the most frequent GO terms (Table 1). In the KNP-specific nsSNPs, cell adhesion
and wound healing (GO:0007155) were the most frequent GO terms (Table 2). The JNP and KNP unique pork quality can originate from genes that were affected by breed-specific variants. In KWB, the enriched GOs pertained to KWB-specific nsSNPs and the genes for DNA repair (GO:0006281) and reproduction (GO:0000003) (Table 3).

**Discussion**

Native pigs, including JNP, KNP, and KWB, are renowned for their good meat quality and juiciness. The fatty acid composition and higher redness value (CIE a*) explain the uniqueness of native pigs compared to other commercial pig breeds (Kim et al. 2009). In JNP, ion homeostasis, neuronal genes, and in KNP, cell adhesion-related genes can affect juicy and chewy pork (Sodhi et al. 2014, Kim et al. 2015).

Some native pigs are likely to be mixed with western commercial pig breeds, as mentioned above. In our study, some JNPs were mixed with DUR, which could interfere with the determination of breed-specific SNPs. Therefore, we eliminated some JNP individuals from further analysis. We did not find any compounded individuals in the NJ tree and treemix after elimination of some JNPs (Figure 1D and E).
Many SNPs were classified as breed-specific SNPs in the whole-genome sequencing data (Figure 2B and C). In the PCA using only breed-specific SNPs in JNP, KNP, and KWB, the results indicated obvious distinctness from other breeds, as seen in principal component 1 (Figure 3A, B, and C), implying that breed-

**Figure 3.** The principal component analysis (PCA) plot using JNP (A), KNP (B) and KWB (C) - specific SNPs. These PCA plot show that the breed-specific SNPs play a characteristic role in defining the breeds. The individuals in JNP, KNP and KWB were color-circled, respectively. The principal component 1 (PC1) was the determining factor in each PCA plot.

| Term                                      | Count | P-Value  | Fold Enrichment | Genes                                                                 |
|-------------------------------------------|-------|----------|-----------------|----------------------------------------------------------------------|
| GO:0007166~cell surface receptor signaling pathway | 127   | 2.26E-04 | 1.4             | HHIP, CTNND1, HHAT, RASSF2, KDR, ENP1, PD2K, MED1, SEMA6A, FBXW7, IL1R2, CD180, MULT1, PRKCA, HNF1B, RUNX2, DDK3, DEPDC1B, SFRP2, CCNY, SMO, ADGRB2, ROR1, IL6ST, IFNAR1, BLK, GPIF, NPF2, CTR9, PRDM14, TGFA, NREP, DEPD2, CD39, CD79A, CALCR, KLK1, ADGRG6, FGFR2, EIF4EBP1, ANKR0D1, LPXN, PLCG1, PLXNA4, WT13, FAM83B, PRNP, TGFBR2, SMAD3, FZD4, SMAD6, PTK2, BMP6, BMP5, IL22RA1, ZCHC3, GH1, CXCL12, FGFR1, ITGA1, AK1, GRB2, OSTN, GLP1R, ITG85, DDX47, PTPRO, CELF4, SPP1L, SHRIF1, ROBO1, GLI2, GHR, LRTM4, LRTM1, BOC, IQUB, DACT2, JAK3, PHACTR4, ANX4A, PEDF4, GLP2R, GAB1, RGM1, LAT2, F7, ADGRF1, ADGRF4, KIT, ADGRF5, GFRAL, ITGA8, IGFB, ITG8, GSC, ZMYND11, TRL2, LEP, PRO, NAL, PC4M4, AEP, LEP, GLUT1, KCP, F2RL1, SPRY1, GHR, ADGR4L |
| GO:0031175~neuron projection development   | 49    | 6.38E-04 | 1.6             | ITSN2, TOP2B, PTPRO, STN2M, KNDCl1, ROBO1, GLI2, SYNGAP1, ZFYVE27, CAP2ZB, KISCF5, DYSPL2, CTCLN1, BOC, EFHD1, EPHB2, PRKGl, CAMK1D, SEMA6A, FBXW8, MAP15, MULT1, KEL, AMIGO1, PLPPR4, PLPPR5, LPAR3, ADCY1, BLOC1S5, CUX2, PTPT2, ZWIM6, WNT3, PLXNA4, PACS1N1, NTNG1, ITGA2, RYK, CCK, LAMBl, BFAPl, SEMA4F, BAIAP2, BMP5, LRFN2, CXCL12, KIF26B, GHR |
| GO:0008081~ion homeostasis                | 45    | 7.88E-04 | 1.7             | GLP1R, CAMK2D, ITPR3, SPP3L, ADRA1A, CLN6, SLC4A5, CACNA1G, SCGN, CALCR, RAB20, ATP6VOA2, CDH23, KDR, TAC4, ENP1, CAPN3, PLCG1, CNM1M2, PD2K, ATP6VOA1, SWAP70, TRPC1, PD4E4, ATP2BI, SYPL2, GPCR1, CAP, ERC1, BMP6, HCRTR2, SLC9A2, SLC9A4, KEL, TRPV4, SLC9A9, AGTR1, F2RL1, GHR, FMAD20, GPD1L, SLC2A6, ISCU, CFTR, CCLN1, CLSTN2, AMIGOl, PLPPR5, WASL, FBNP11, SCN, CUXL2, LRTM1, SYNGDI1, KDR, SNX9, ARFPI1, EPHB2, BCAA3, SEMA2D, GSN, CNIT6L, SWAP70, BFAFl, HNF1B, VPS37B, TERFl, BAIAP2, MPP7, CLIP1, PSM3, CDC424EP4, F2RL1, GRB2, GHR, EPHAI |
| GO:0044089~positive regulation of cellular component biogenesis | 32    | 8.78E-04 | 1.9             | GLP1R, CAMK2D, ITPR3, SPP3L, ADRA1A, CLN6, SLC4A5, CACNA1G, SCGN, CALCR, RAB20, ATP6VOA2, CDH23, KDR, TAC4, ENP1, CAPN3, PLCG1, CNM1M2, PD2K, ATP6VOA1, SWAP70, TRPC1, PD4E4, ATP2BI, SYPL2, GPCR1, CAP, ERC1, BMP6, HCRTR2, SLC9A2, SLC9A4, KEL, TRPV4, SLC9A9, AGTR1, F2RL1, GHR, FMAD20, GPD1L, SLC2A6, ISCU, CFTR, CCLN1, CLSTN2, AMIGOl, PLPPR5, WASL, FBNP11, SCN, CUXL2, LRTM1, SYNGDI1, KDR, SNX9, ARFPI1, EPHB2, BCAA3, SEMA2D, GSN, CNIT6L, SWAP70, BFAFl, HNF1B, VPS37B, TERFl, BAIAP2, MPP7, CLIP1, PSM3, CDC424EP4, F2RL1, GRB2, GHR, EPHAI |
| GO:0006885~regulation of pH                | 12    | 0.0012   | 3.1             | SLC4A7, SLC9A4, RAB20, SLC9A9, ATP6VOA2, SLC26A4, CLN6, CFTR, PD2K, SLC4A5, ATP6VOA |
| GO:0051222~positive regulation of protein transport | 30    | 0.0012   | 1.9             | BLK, SPP3L, WDR4E, IL1R1, RUNBPL3, HHA, MED1, BCAA3, TGBF2, SMAD3, FBXW7, ACTL6A, ACSL3, SREBF2, PT2K, GCK, BMP6, GLUD1, SFRP2, SMO, TRPV4, LEP, F2RL1, PPARG, GHR, SAE1, GAS6, CFTR, TRL2 |
| GO:0003030~cell projection organization   | 69    | 0.0019   | 1.4             | ITSN2, CEPI26, TOP2B, SFP2, PTPRO, STN2M, KNDCl1, FBNP11, ROBO1, GLI2, SYNGAP1, ZFYVE27, CAPZB, KISCF5, DYSPL2, CTCLN1, BOC, EFHD1, EPHB2, PRKGl, BCAA3, CAMK1D, SEMA6A, FBXW8, MAP15, MULT1, KEL, ENP1, CAPN3, PLCG1, CNM1M2, PD2K, ATP6VOA1, SWAP70, TRPC1, PD4E4, ATP2BI, SYPL2, GPCR1, CAP, ERC1, BMP6, HCRTR2, SLC9A2, SLC9A4, KEL, TRPV4, SLC9A9, AGTR1, F2RL1, GHR, FMAD20, GPD1L, SLC2A6, ISCU, CFTR, CCLN1, CLSTN2, AMIGOl, PLPPR5, WASL, FBNP11, SCN, CUXL2, LRTM1, SYNGDI1, KDR, SNX9, ARFPI1, EPHB2, BCAA3, SEMA2D, GSN, CNIT6L, SWAP70, BFAFl, HNF1B, VPS37B, TERFl, BAIAP2, MPP7, CLIP1, PSM3, CDC424EP4, F2RL1, GRB2, GHR, EPHAI |
specific SNPs can be useful genetic markers for genomics. Breed-specific SNPs can elucidate the genomic landscape related to the breed’s unique characteristics in comparison to other breeds, which implies breed informative SNPs. Therefore, breed-specific SNPs may be useful genetic markers. Furthermore, because breed-specific nsSNPs alter amino acid sequence and protein function, they directly affect breed characteristics. The deleterious nsSNPs in the breed can actually be deleterious, but we identified that nsSNPs with tolerated effects can be the determining factors of the unique characteristics of native pig breeds.

To determine the genomic factors that explain the characteristics of native pig breeds in Korea, we surveyed the GO terms of each native pig breed (JNP, KNP, and KW). In JNP, the notable genes containing nsSNPs were Transforming growth factor beta 2 protein (TGFB2), glutathione S-transferase kappa 1 (GSTK1) and ATP-binding cassette subfamily A member 8 (ABCA8). In previous study, TGFB2 gene were reported to be related to the developmental growth in JNP in comparison to the Berkshire pigs (Lee, Shin, et al. 2020). Members of the glutathione S-transferase (GST) family constitute a large class of cytoplasmic, membrane-bound, multigene, and multifunctional enzymes (Hayes et al. 2005). Sperm GSTs are not only important for sperm function but are also used as biomarkers for estimating sperm quality. The protein encoding GST kappa 1 (GSTK1, also named Dsb-L; nsSNP rs326967994: T→C, Ile→Val) is a highly conserved mitochondrial enzyme involved in lipid metabolism and energy production. Recent studies have suggested that GSTK1 plays an essential role in sperm function (Gao et al. 2009, Petit et al. 2013). ATP binding cassette subfamily A member 8 (ABCA8; nsSNP rs328587109: G→T, Ala→Ser) is a member of the ATP-binding cassette (ABC) transporter superfamily. ABCA8 is a transmembrane transporter responsible for the transport of organics such as drug efflux and cholesterol. This transcript is mediated by microRNA-374b-5p (miR-374b-5p). According to a luciferase reporter assay, miR-374b-5p binds to the ABCA8 3′-untranslated region (3′UTR) (Cui et al. 2020). The notable genes in breed-specific SNPs in JNP were calcium-dependent protein kinase II delta (CAMK2D: GO:0050801, ion homeostasis) and estrogen receptor 1 (ESR1: GO:0050801, ion homeostasis). CAMK2D regulates multiple ion channels by phosphorylation, including Ca2 + channels, sarcolemmal cardiac Na + channels, and K+ channels. By activating these channels, CAMK2D causes an extracellular influx of ions (Grootjans et al. 2017). In a Chinese-European pig line, ESR1 was found to be related to litter size (Muñoz et al. 2007). ESR1 encodes an estrogen-and a ligand-activated transcription factor. The encoded protein regulates the transcription of many estrogen-inducible genes that play roles in growth, metabolism, gestation, and sexual development.

In KNP, the notable gene encompassing nsSNPs was integrin beta 3 (ITGB3, GO:0007155→cell adhesion). Researchers have found that integrin beta 3 (ITGB3; nsSNP: rs341036511 A→G, Ile→Thr, rs336055648 T→C, Met→Val, rs339763435 T→C, Tyr→Cys) affects immune cell adhesion and leukocyte movement (Jana et al. 2021). Integrins are ubiquitously expressed adhesion molecules that are highly glycosylated and contain Ca2 + or Mg2 + ions that are essential for ligand binding. Immunoglobulin superfamily containing leucine-rich repeat (ISLR) mRNA levels correlate negatively with LDL-cholesterol (Jacobsen et al. 2019).
KWB, the notable genes were DNA cross-link repair 1C (DCLRE1C, GO:0006974~cellular response to DNA damage stimulus) and Deltex 3-like protein (DTX3L, GO:0006974~cellular response to DNA damage stimulus). Severe combined immunodeficiency (SCID) is a group of inherited disorders characterized by T lymphocyte differentiation related to abnormal lymphocyte development. SCID caused by mutations in DCLRE1C (nsSNP: rs323944749 T→C, Tyr→His) results in the absence of both B and T cells (Iqbal et al. 2019). Previous studies have shown that DTX3L (nsSNP rs331172640 A→C, Leu→Arg) is more highly expressed in male pigs than in female pigs. DTX3L is a crucial member of the Notch signaling pathway that controls myogenesis (Zhang et al. 2013).

**Conclusion**

Native pig breeds in Korea (Jeju Native Pig: JNP, Korean Native Pig: KNP, Korean Wild Boar: KWB) have unique characteristics, such as juicy and very tasty meat quality and disease resistance. We identified genes that represent the genomic characteristics of pigs native to Korea. In JNP, the notable GO term was neuron projection development, cell surface receptor signaling pathway, and ion homeostasis; in KNP, cell adhesion and wound healing; and in KWB, DNA repair and reproduction. Genes containing breed-specific SNPs may help explain the unique characteristics of native pig breeds.

**Availability of data and materials**

The datasets analyzed during the current study are not publicly available because of intellectual property considerations, but are available from the corresponding author upon reasonable request.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by the Ministry of Education of the Republic of Korea, the National Research Foundation of Korea (NRF-2022R1A2C4002510), and the Cooperative Research Program for Agriculture Science & Technology Development, Rural Development Administration, Republic of Korea (PJ01620403).

**References**

Asaf VM, Kumar A, Rahim A, Sebastian R, Mohan V, Dewangan P, Panigrahi M. 2014. An overview on single nucleotide polymorphism studies in mastitis research. Vet World. 7:416–421.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 30:2114–2120.

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. Am J Human Genet. 81:1084–1097.
Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms. SnpedEff: SNPs in the Genome of Drosophila Melanogaster Strain w1118; iso-2; iso-3. Fly. 6:80–92.

Cui Y, Liang S, Zhang S, Zhang C, Zhao Y, Wu D, Wang J, Song R, Wang J, Yin D. 2020. Circular RNA circRHB83 acts as a sponge for miR-654-3p inhibiting gastric cancer growth. J Exp Clin Cancer Res. 39:1–14.

Czech B, Frączczak M, Mięczarek M, Szyja J. 2018. Identification and annotation of breed-specific single nucleotide polymorphisms in Bos Taurus genomes. PLoS One. 13:e0198419.

Fikacek RR. 2021. `OptM`: estimating the optimal number of migration edges on population trees using `Treemix`. Biol Methods Protoc. 6:bpab017.

Gao F, Fang Q, Zhang R, Lu J, Lu H, Wang C, Ma X, Xu J, JiA W, XiANG K. 2009. Polymorphism of DsBA-L gene associates with insulin secretion and body fat distribution in Chinese population. Endocr J. 56(3):487.

Giuffra E, Kijas J, Amarger V, Carlborg Ö, Jeon J, Andersson L. 2005. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics. 154:1785–1791.

Grootjans S, Vanden Berghe T, Vandenabeele P. 2017. Initiation and execution mechanisms of necroptosis: an overview. Cell Death Differ. 24:1184–1195.

Hayes JD, Flanagan JU, Jowsey IR. 2005. Glutathione transferases. Annu Rev Pharmacol Toxicol. 45:51–88.

Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC. 2007. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 35:W169–W175.

Iqbal MA, Hong K, Kim JH, Choi Y. 2019. Severe combined immunodeficiency pig as an emerging animal model for human diseases and regenerative medicines. BMB Rep. 52:625.

Jacobsen MJ, Havagaard JH, Anthon C, Mentzel CMJ, Cirera S, Krogh PM, Pundhir S, Karlskov-Mortensen P, Bruun CS, Lesnik P. 2019. Epigenetic and transcriptomic characterization of pure adipocyte fractions from obese pigs identifies candidate pathways controlling metabolism. Front Genet. 10:1268.

Jana B, Kaczmarek MM, Romaniewicz M, Brzozowska M. 2021. Profile for mRNA transcript abundances in the pig endometrium where inflammation was induced by Escherichia coli. Anim Reprod Sci. 232:106824.

Kim D, Seong P, Cho S, Kim J, Lee J, Jo C, Lim D. 2009. Fatty acid composition and meat quality traits of organically reared Korean native black pigs. Livest Sci. 120:96–102.

Kim J, Cho S, Caetano-Anolles K, Kim H, Ryu Y-C. 2015. Genome-wide detection and characterization of positive selection in Korean native black pig from Jeju Island. BMC Genet. 16:1–9.

Kim T, Kim K, Choi B, Yoon D, Jang G, Lee K, Chung H, Lee H, Park H, Lee J. 2005. Genetic structure of pig breeds from Korea and China using microsatellite loci analysis. J Anim Sci. 83:2255–2263.

Krawczak M, Ball EV, Fenton I, Stenson PD, Abeyasinghe S, Thomas N, Cooper DN. 2000. Human gene mutation database? A biomedical information and research resource. Hum Mutat. 15:45–51.

Kuehn L, Keele J, Bennett G, McDaneld T, Smith T, Snelling W, Sonstegard T, Thallman R. 2011. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project1,2. J Anim Sci. 89:1742–1750.

Kumar H, Panigrahi M, Chhotaray S, Parida S, Chauhan A, Bhushan B, Gaur G, Mishra B, Singh R. 2021. Comparative analysis of five different methods to design a breed-specific SNP panel for cattle. Anim Biotechnol. 32:130–136.

Lee Y-S, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A, Vigne J-D, Edwards CJ, Schlumbaum A, Dinu A. 2007. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. Proc Natl Acad Sci USA. 104:15276–15281.

Lee Y-S, Liu R, Zhao X, Yuan J, Fuller D, Barton L, Dobney K, Fan Q, Gu Z, Liu X-H. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. Proc Natl Acad Sci USA. 107:7686–7691.

Lee Y-S, Shin D, Won K-H, Kim DC, Lee SC, Song K-D. 2020. Genome-wide scans for detecting the selection signature of the Jeju-island native pig in Korea. Asian-Australasian Journal of Animal Sciences. 33(4):539–546.

Lee Y-S, Son S, Heo J, Shin D. 2021. Detecting the differential genomic variants using cross-population phenotype-associated variant (XP-PAV) of the Landrace and Yorkshire pigs in Korea. Animal Cells and Systems. 25(6):416–423.

Lee Y-S, Won K, Shin D, Oh J-D. 2020. Risk prediction and marker selection in nonsynonymous single nucleotide polymorphisms using whole genome sequencing data. Animal Cells and Systems. 24(6):321–328.

Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. bioinformatics. 25:1754–1760.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M. 2010. The genome analysis toolkit: a map reduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297–1303.

Muñoz G, Ovillo C, Estellé J, Síllo L, Fernández A, Rodríguez C. 2007. Association with litter size of new polymorphisms on ESR1 and ESR2 genes in a Chinese-European pig line. Genet Sel Evol. 39:1–12.

Pasupa K, Rathasamuth W, Tongsim S. 2020. Discovery of significant porcine SNPs for swine breed identification by a hybrid of information gain, genetic algorithm, and frequency feature selection technique. BMC Bioinformatics. 21:1–28.

Petit F, Serres C, Bourgeon F, Pineau C, Auer J. 2013. Identification of sperm head proteins involved in zona pellucida binding. Hum Reprod. 28:852–865.

Raj A, Stephens M, Pritchard JK. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics. 197:573–589.

Ramos A, Megens H, Crooijmans R, Schook L, Groenen M. 2011. Identification of high utility SNPs for population assignment and traceability purposes in the pig using high-throughput sequencing. Anim Genet. 42:613–620.
Sodhi SS, Park WC, Ghosh M, Kim JN, Sharma N, Shin KY, Cho IC, Ryu YC, Oh SJ, Kim SH. 2014. Comparative transcriptomic analysis to identify differentially expressed genes in fat tissue of adult Berkshire and Jeju Native Pig using RNA-seq. Mol Biol Rep. 41:6305–6315.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30:2725–2729.

Wu J, Jiang R. 2013. Prediction of deleterious nonsynonymous single-nucleotide polymorphism for human diseases. Scientific World J.

Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. The American J Human Genet. 88:76–82.

Yeo J-S, Kim J-W, Chang T-K, Park Y, Nam D-H. 2000. Utilization of DNA marker-assisted selection in Korean native animals. Biotechnol Bioprocess Eng. 5:71–78.

Zhang J, Zhou C, Ma J, Chen L, Jiang A, Zhu L, Shuai S, Wang J, Li M, Li X. 2013. Interdisciplinary dialogue for education, collaboration, and innovation: intelligent biology and medicine in and beyond 2013. BMC Genomics. 14:S1–S9.