Article

Identification of Selection Signatures and Loci Associated with Important Economic Traits in Yunan Black and Huainan Pigs

Yachun Han [1], Tao Tan [1], Zixin Li [2], Zheng Ma [1], Ganqiu Lan [2], Jing Liang [2], Kui Li [1] and Lijing Bai [1,*]

1   Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518124, China
2   College of Animal Science & Technology, Guangxi University, Nanning 530003, China
*   Correspondence: bailijing@caas.cn; Tel./Fax: +86-0755-2325-0160

Abstract: Henan Province is located in central China and rich in domestic pig populations; Huainan (HN) pigs are one of three Henan indigenous breeds with great performance, including early maturation, strong disease resistance and high meat quality. Yunan (YN) black pigs are a typical, newly cultivated breed, synthesized between HN pigs and American Duroc, and are subjected to selection for important traits, such as fast growth and excellent meat quality. However, the genomic differences, selection signatures and loci associated with important economic traits in YN black pigs and HN pigs are still not well understood. In this study, based on high-density SNP chip analysis of 159 samples covering commercial DLY (Duroc × Landrace × Large White) pigs, HN pigs and YN black pigs, we performed a comprehensive analysis of phylogenetic relationships and genetic diversity among the three breeds. Furthermore, we used composite likelihood ratio tests (CLR) and F-statistics (Fst) to identify specific signatures of selection associated with important economic traits and potential candidate genes. We found 147 selected regions (top 1%) harboring 90 genes based on genetic differentiation (Fst) in the YN-DLY group. In the HN-DLY group, 169 selected regions harbored 58 genes. In the YN-HN group, 179 selected regions harbored 77 genes. In addition, the QTLs database with the most overlapping regions was associated with triglyceride level, number of mummified pigs, hemoglobin and loin muscle depth for YN black pigs, litter size and intramuscular fat content for HN pigs, and humerus length, linolenic acid content and feed conversion ratio mainly in DLY pigs. Of note, overlapping 14 tissue-specific promoters' annotation with the top Fst 1% selective regions systematically demonstrated the muscle-specific and hypothalamus-specific regulatory elements in YN black pigs. Taken together, these results contribute to an accurate knowledge of crossbreeding, thus benefitting the evaluation of production performance and improving the genome-assisted breeding of other important indigenous pig in the future.

Keywords: signatures of selection; pig high-density SNP chip; QTL; annotation

1. Introduction

Pork is one of the most important forms of animal protein, with common use as a human dietary protein. Western commercial (Europe and America) pig breeds are well-known for superior performance, with fast growth, effective feed conversion efficiency and high carcass yield [1]. On the other hand, eastern (Asia) pigs, especially Chinese indigenous breeds, are characterized by early maturation, high reproduction rate, strong disease resistance, tolerance to harsh environments and high-quality meat. Therefore, the crossbreeding of western breeds and Chinese indigenous pigs has become the leading practice in China for developing new populations. Huainan (HN) pigs are one of northern China's oldest native breeds, with all-black color, large body size, high fertility, strong adaptability and good meat quality. Yunan (YN) black is a typical, newly hybrid pig breed,

synthesized between HN pig and American Duroc, selected for important traits such as growth and excellent meat quality [2]. In 2008, YN black pigs were registered as a new breed by China National Livestock and Poultry Genetic Resources Committee.

The process of domestication and modern breeding practice, with the subsequent breed formation and artificial directional selection, has left detectable selection signatures in numerous regions of the pig genome [3–7]. These regions can be analyzed by applying various genomic and population statistical methods. Li et al. characterized for under selection of distinct genomic regions involved in hypoxia, olfaction, energy metabolism and drug response in Tibetan wild boars based on whole-genome re-sequencing [8]. Ai et al. found a 14-Mb region with low recombination rate on the X chromosome playing a key role in regional adaptations to high and low latitude environments within China [9]. Moreover, Frantz et al. analyzed pig domestication using over 100 genome sequences and demonstrated that domestic pigs have strong signatures of selection at loci that have contributed to behavior and morphology [10]. Ma et al. identified signatures of artificial selection and loci related to intramuscular fat content and feed conversion ratio traits in Duroc Pigs [11]. In addition, a number of studies have been carried out to explore specific signatures of selection in Chinese indigenous pig breeds, including Anqing Six-End-White [12,13], Saba, Baoshan, Tunchang, Dingan [14], Meishan [15] and Lulai pigs [16]. Qiao et al. analyzed genetic structures and mapped genomic regions affecting backfat thickness of Yunan black pigs [17]. However, the genomic differences and integrated analysis of economical traits of HN and YN black pigs are still not well understood.

In this study, we analyzed 159 genotyped pigs, including one commercial breed (DLY), one Chinese indigenous breed (HN) and one Chinese cultivated breed (YN) with the 70 k Porcine Functional Variants Chip. SNPs were called and allele frequencies were calculated. We performed a comprehensive analysis of phylogenetic relationships and genetic diversity among the three breeds. Furthermore, two different statistics, the composite likelihood ratio (CLR) and fixation index (Fst) were used to identify specific signatures of selection associated with important economic traits and potential candidate genes in these three breeds. Our findings facilitate a better understanding of HN and YN black pigs' genome characteristics and provide novel insights for molecular-assisted breeding strategies and germplasm conservation in the near future.

2. Materials and Methods

2.1. Ethics Approval

All animals used in this study were sampled according to guidelines for the Care and Use of Experimental Animals established by the Ministry of Agriculture and Rural Affairs of China. The protocol was approved by the College of Animal Science & Technology, Guangxi University (ethics approval reference number: Gxu-2022-114).

2.2. DNA Sample Collection

A total of 159 subjects from three pig populations were collected from China. Specifically, 99 commercial Duroc × Landrace × Large White (DLY) pigs, collected from Guangxi University, Nanning City, Guangxi Province, China. Additionally, 30 Huainan pigs (HN) and 30 Yunan black pigs (YN) were collected from Gushi Country, Xinyang City, Henan Province, China.

2.3. Single-Nucleotide Polymorphism (SNP) Genotyping and Data Quality Control

Genomic DNA samples from the three pig populations were extracted from ear tissue using the traditional standard phenol/chloroform method, and the quality was assessed using the light absorption ratio (A260/280). The Tianpeng chip (Porcine SNP70 Functional Variants Genotyping Array), which contains 187,244 SNPs, was used for the genotyping of all three pig populations. The SNP data can be downloaded from figshare database (https://figshare.com/articles/dataset/Pigs_70K_functional_locus_gene_chip/21900417, accessed on 14 January 2023). The sequence variants were called using PLINK v1.90 software [17,18],

and quality control criteria were set as follows: (1) filtering out individuals with low detection rates according to the criterion of -mind 0.2; (2) removing SNP sites with a SNP deletion rate greater than 0.80 and unknown location or located on sex chromosome; and (3) removing SNP sites with minor allele frequency (MAF) less than 0.01. SNP genome coordinates were obtained from the Susscrofa 11.1 porcine genome reference assembly (https://www.ncbi.nlm.nih.gov, accessed on 12 September 2022).

2.4. Diversity, Phylogenetic, and Population Genetic Analyses

In order to gain a better understanding of the relationship among the three populations in this study, a phylogenetic tree (neighbor-joining tree, NJ-Tree) was constructed by using genetic distance matrix calculated by MEGA [19], then the tree was visualized online using the ITOL website (https://itol.embl.de accessed on 15 December 2022).

To investigate the pattern of genetic differentiation among different populations, principal component analysis (PCA) was performed on all autosomal SNPs using the smartpca program in the EIGENSOFT software package (https://mybiosoftware.com/eigensoft-population-structure-eigenanalysis-stratification.html accessed on 9 November 2022), significance of the eigenvalues was tested by Tracy–Widom distribution and the results were visualized using the R program ggplot2 package (https://www.R-project.org/ accessed on 28 November 2022).

For exploring the exactly genetic structure of the three populations, the population structure was deduced using the ADMIXTURE software v1.3.0 [20]. The K value from two to nine (K = 2–9) and the cross validation value (CV error) were calculated. To estimate and compare the linkage imbalance (LD) patterns of the three populations, PopLDdecay software (https://github.com/BGI-shenzhen/PopLDdecay accessed on 14 January 2023) was used to calculate the value of the squared correlation relationship coefficient (r^2) between any two SNPs in a 300 kb interval [21].

2.5. Genome-Wide Selection Sweeps Detection

The genome-wide selection signatures of the three populations were identified using two different statistical methods: composite likelihood ratio tests (CLR) and F-statistics (Fst). The CLR values of the three populations were calculated separately using sweepfinder2 (v1.24) software [22]. The quality control criteria were set as follows: all CLR values of autosomes were rank ordered separately with a 10 kb sliding window and 10 kb step size, the top 1% was considered as regions under selection, and then these candidate regions were genetically, functionally annotated using ANNOVAR (v2.1.1) software [23].

Estimations of Fst among the three populations were calculated using the VCFTOOLS (v0.1.13) package [24] with a 200 kb sliding window and 20 kb step size, accompanied with minor 20 variants per window. All Fst values were sorted, the top 1% level was considered as putative regions under selection. The Fst figure was drawn using R (v4.3.1) package [25].

2.6. Annotation of Promoters of Tissue-Specific Expression Genes in 14 Tissues

In this study, we overlapped the top 1% regions sorted according to the Fst values of YN-DLY, YN-HN and HN-DLY groups with the 14 tissue-specific promoters' annotation database [26].

2.7. Annotation of the Candidate Genes and Quantitative Trait Loci (QTL) Overlapping with Potential Selection Signatures

Genome annotation is based on Susscrofa 11.1 (https://www.animalgenome.org/blast/ accessed on 14 November 2022). Genes in these potentially selected regions were considered as candidate genes after RNAs and unidentified genes were filtered out. To explore the function of these identified genes, Gene Ontology (GO) [27] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [28] were used for enrichment analyses through the Database for Annotation, Visualization, and Integrated Discovery (DAVID) Version 6.8 [29,30]. Furthermore, the pig quantitative trait loci (QTL) database [31] (https://www.animal-genome.org/cgi-bin/QTLdb/SS/index, on SS11.1 in gff format, accessed

on 13 December 2022) was used to annotate economic traits associated with the selective sweep regions.

3. Results

3.1. Genotypes and Population Genetics Analysis

A total of 187,244 SNPs located on autosomes were genotyped in the SNP-chip. In the process of quality control, 65,779 SNPs and 2 individuals were discarded due to low SNP call rate and MAF. After the quality control, 121,465 SNPs from 157 individuals were retained for further study.

To describe the exact relationship among YN black, HN and DLY pigs, we first performed a neighbor-joining (NJ) tree analysis. The phylogenetic tree showed that the three populations were clearly divided into separate groups (Figure 1a). Secondly, PCA was conducted with PC1 and PC2, explaining 18.05% and 3.48% of the variations, respectively (Figure 1b). The PCA analysis clearly separated DLY from HN, with YN black dispersed in-between these two populations, as expected from the hybrid nature of YN black pigs. ADMIXTURE analyses were conducted with K = 2 to K = 9 (Figure S1) and cross validation analysis showed K = 8 was the optimal number of possible ancestral varieties (Figure 1c). When K = 2, all samples were divided into commercial pigs and Chinese indigenous pigs, both YN black and DLY pigs were clustered together. When K = 3, the YN black population could be well-distinguished from commercial pigs. When K = 8, the DLY pigs were separated into several sub-populations, and the YN population appeared a mixture of HN and DLY sub-populations (Figures 1d and S1). Estimation of LD (r2) demonstrated highest mean LD and slowest decay in the DLY pigs and lowest mean LD and fastest decay in the YN black pigs (Figure 1e).
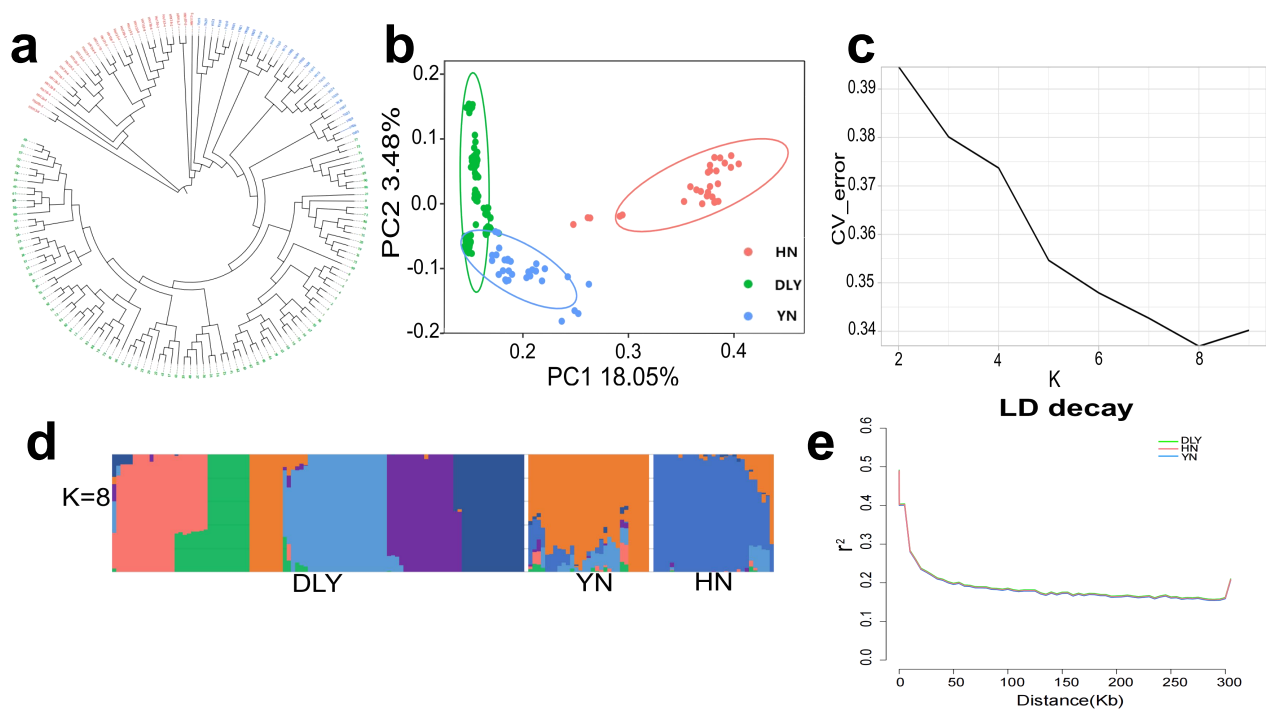
Figure 1. Genomic characteristics of three pig breeds. (**a**) Neighbor-joining (NJ) tree of the after-filtering 157 pigs covering 30 YN black pigs, 28 HN pigs and 99 DLY pigs. Green color: DLY pigs; blue color: YN black pigs; red color: 28 HN pigs. (**b**) Principal component analysis of 30 YN black pigs (blue), 28 HN pigs (red) and 99 DLY pigs (green). (**c**) Ancestral composition analysis K = 2–9. (**d**) Structure analysis with K = 8. (**e**) LD decay in different breeds with 300 kb.

3.2. Identification of Signatures of Selection within-Population by CLR

To study the positive selection within the three populations, whole genome scans were performed by CLR and the top 1% was selected as the candidate regions. YN black pigs yield 2501 regions (Table S1), HN pigs yield 2583 regions (Table S2) and 2114 regions were found for DLY pigs (Table S3). A total of 2829, 3242 and 3603 genes were harbored in these potential selection regions in YN, HN and DLY populations, respectively (Figure 2a–c).
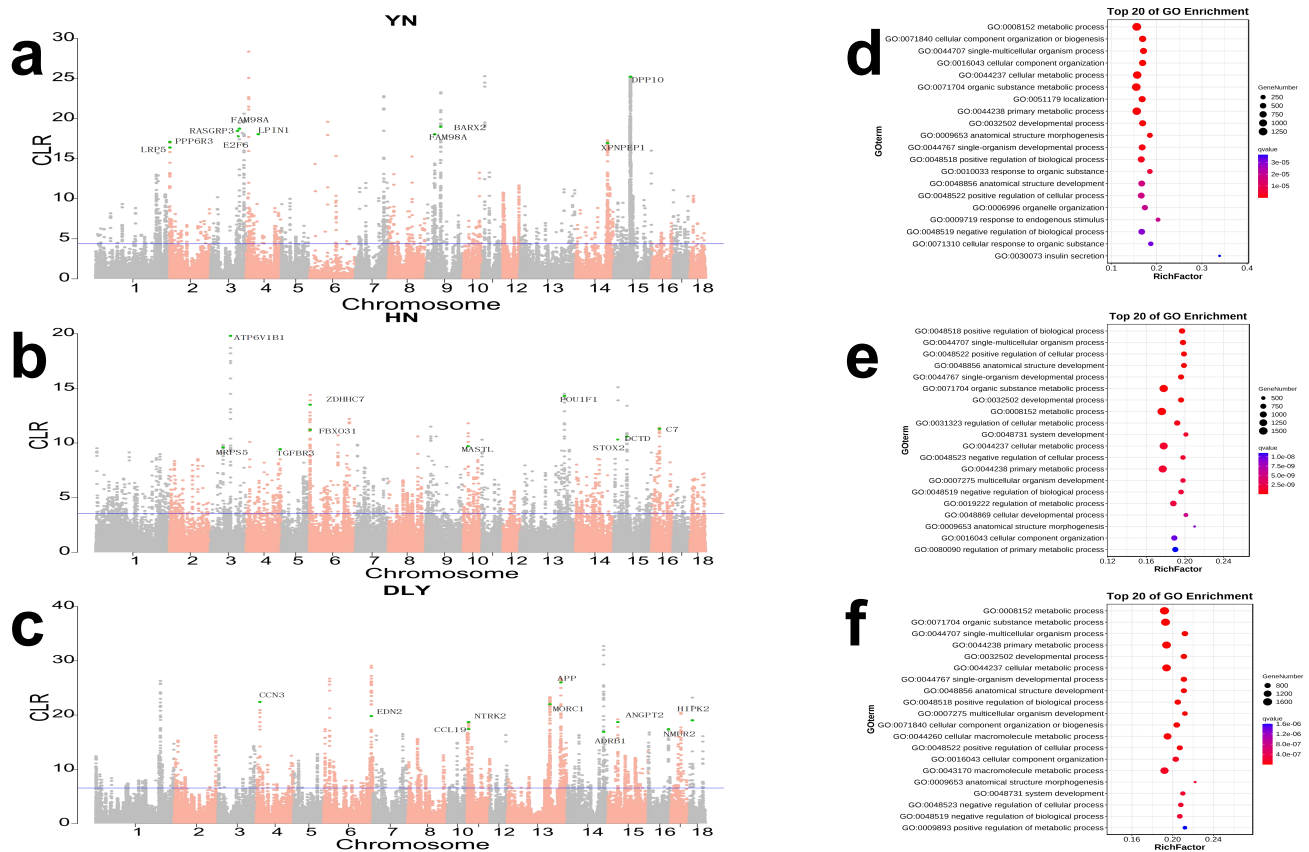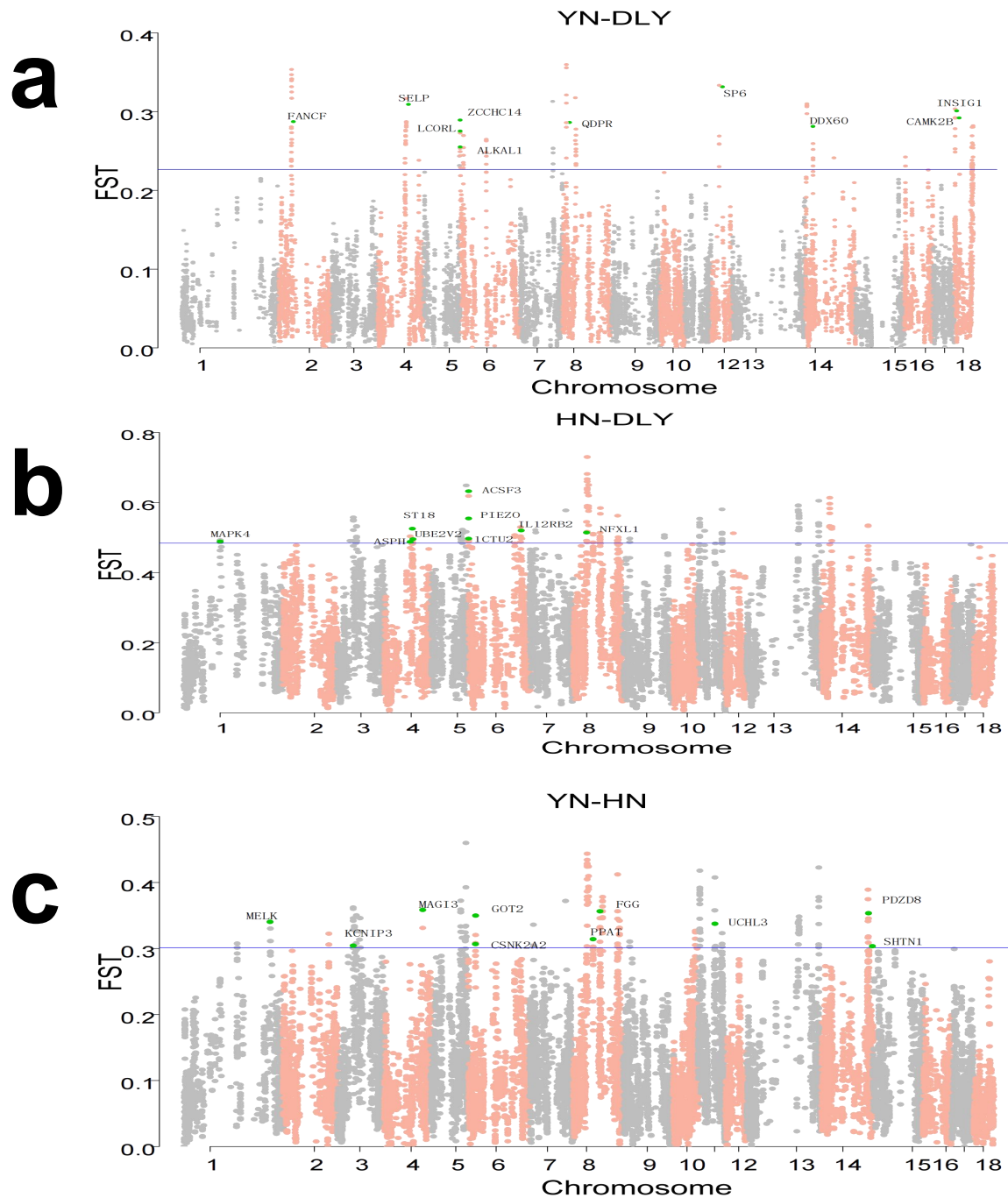

Figure 2. Genome-wide distribution of selection signatures detected by CLR. (**a**) CLR statistics across all autosomes in YN black pigs. (**b**) CLR statistics across all autosomes in HN pigs. (**c**) CLR statistics across all autosomes in DLY pigs. The blue line corresponds to the 99% threshold on the corresponding empirical distributions. (**d**–**f**) The top 20 GO enrichment analysis of selected genes in YN black pigs, HN pigs and DLY pigs.

GO terms enrichment were applied on the genes in the significant regions, and the top 20 *p*-value of each group are shown in Figure 2f. The most significant enriched terms of YN black pigs were insulin secretion (GO: 0030073, *p* = 0.00000426, 12 genes), cellular response to organic substance (GO: 0071310, *p* = 0.001716052, 273 genes) and negative regulation of peptide secretion (GO: 0002792, *p* = 0.000210512, 12 genes) (Figure 2d). The most significant enriched terms of HN pigs were regulation of primary metabolic process (GO: 0044238, *p* = 0.03215, 13 genes), cellular developmental process (GO: 0048869, *p* = 0.0002916, 608 genes) and regulation of metabolic process (GO: 0019222, *p* = 0.0000971, 1006 genes) (Figure 2e). For DLY pigs, the most significant enriched term was positive regulation of metabolic process (GO: metabolic process, *p* = 0.0000928, 10 genes) (Figure 2f).

### 3.3. Identification of Signatures of Selection between-Population by Fst

To further explore the specific signatures of selection between the three populations, Fst statistical analyses were conducted. We calculated Fst per site and averaged them in a non-overlapping 200 kb sliding window approach with a 20 kb step size across all autosomes. The windows where the scores of the statistics fell in the top 1% were considered

significant. For the YN-DLY comparison, with DLY pigs treated as reference population, 147 of 14,804 sliding windows were identified as potential selection signatures, including 90 candidate genes (Figure 3a, Table S4). For the HN-DLY comparison, with DLY pigs as the reference population, 169 out of 16,954 sliding windows were identified, including 58 genes (Figure 3b, Table S5). Finally, for the YN-HN comparison, with HN pigs as the reference population, 179 out of 17,961 sliding windows were identified, including 77 genes (Figure 3c, Table S6).

Figure 3. Genome-wide distribution of selection signatures detected by Fst. (**a**) Fst statistics across all autosomes in YN black pigs. (**b**) Fst statistics across all autosomes in HN pigs. (**c**) Fst statistics across all autosomes in DLY pigs. The blue line corresponds to the 99% threshold on the corresponding empirical distributions.

GO terms and KEGG pathway enrichment analyses on candidate genes were conducted. In the GO analysis, we found a total of 6 terms which were significantly enriched in YN black pigs (e.g., metabolic process (GO: 0008152, p = 0.093663, 17 genes), developmental process (GO: 0.032502, p = 0.36310, 11 genes) and growth (GO: 0040007, p = 0.039897, 2 genes)). The KEGG pathway analysis suggested that the candidate genes of YN black pigs were enriched in Tryptophan metabolism, cysteine methionine metabolism and inositol phosphate metabolism (Figure 4a). For HN pigs, there were 7 GO terms significantly enriched: metabolic process (GO: 0008152, p = 0.70383, 29 genes), developmental process (GO: 0032502, p = 0.41900, 15 genes), immune system process (GO: 0040007, p = 0.0245, 9 genes) and reproduction (GO: 0002376, p = 0.93621, 3 genes), among others. The KEGG pathway analysis showed that the HN pigs' candidate genes were enriched in the JAK-STAT signaling pathway, vitamin B6 metabolism and fatty acid biosynthesis (Figure 4b). Finally, 5 GO terms were significantly enriched in DLY pigs, relating to metabolic process (GO: 0008152, p = 0.03215, 12 genes), growth (GO: 0040007, 0.069601, 2 genes), developmental process (GO: 0032502, 0.015029, 21 genes) and reproductive process (GO: 0022414, p = 0.068548, 3 genes). The KEGG pathway analysis indicated that the candidate genes of DLY pigs were enriched in alanine, aspartate and glutamate metabolism, phosphonate and phosphonate metabolism and biosynthesis of amino acids (Figure 4c).
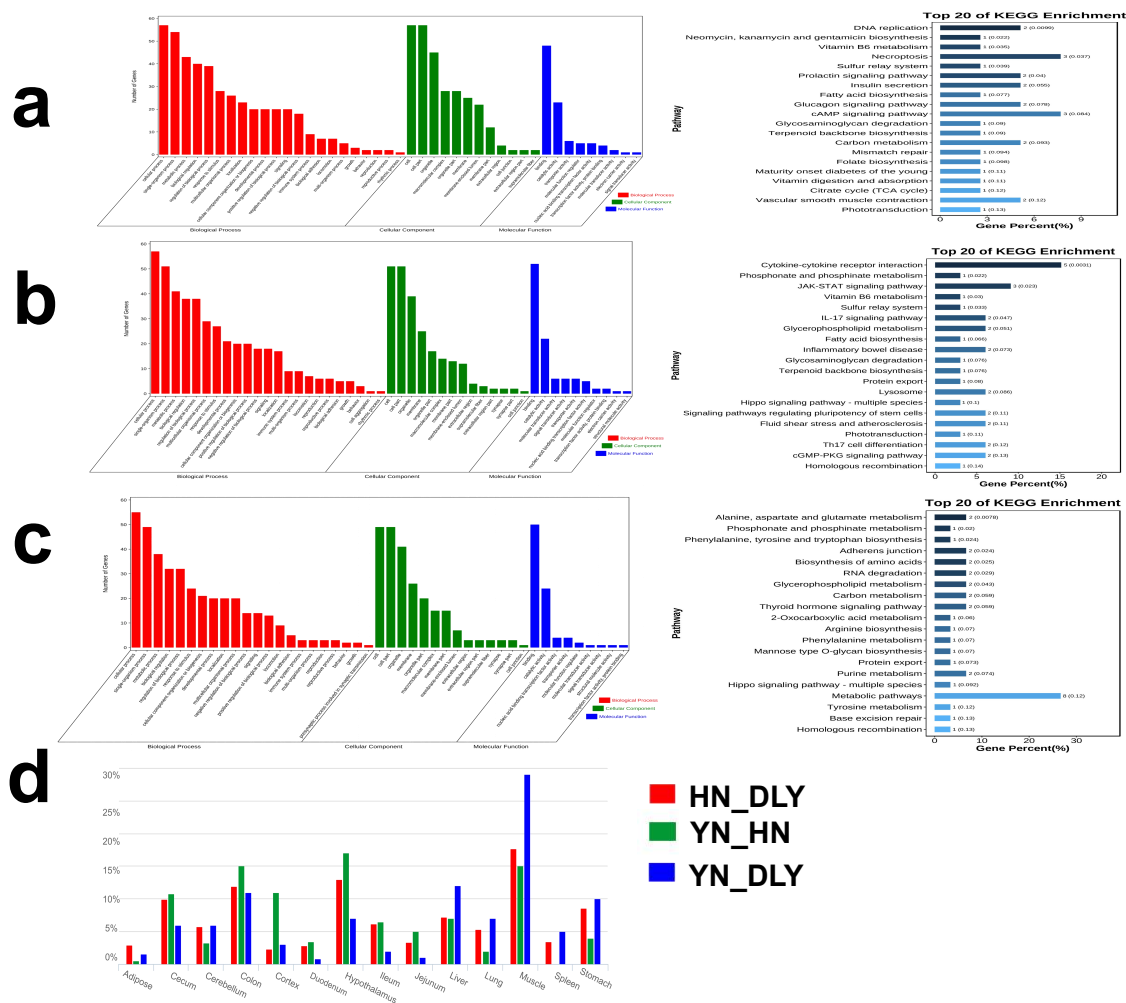
Figure 4. The function analysis of the selected genes. (**a**) KEGG pathway and top 20 GO enrichment analysis of differential expressed genes in YN-DLY group. (**b**) KEGG pathway and top 20 GO enrichment analysis of differential expressed genes in HN-DLY group. (**c**) KEGG pathway and top 20 GO enrichment analysis of differential expressed genes in YN-HN group. (**d**) Annotation of the non-coding selection regions with 14 tissues promoter database.

To gain a better understanding of the mechanisms by which non-coding elements in selective regions cause quantitative trait differentiation between two different populations, we identified 14 tissue-specific promoters in the top 1% regions, sorted according to Fst value. Interestingly, we found much more muscle-specific regulatory elements in active promoters of YN black pigs' selective regions compared to DLY pigs and HN pigs. In addition, the ratio of hypothalamus-specific regulatory elements in active promoters of the YN population's selective regions was also the highest when compared to the other two (Figure 4d).

3.4. QTL Overlapping with Potential Signatures of Selection

We employed the pig quantitative trait loci (QTL) database to annotate economic traits associated with the top 1% Fst candidate regions. QTL overlapping with the candidate selection regions of YN black pigs was associated with triglyceride level, number of mummified pigs, hemoglobin and loin muscle depth. QTL overlapping with the candidate selection regions of HN pigs was associated with litter size and intramuscular fat content. For DLY pigs, the associated QTL traits were ear weight, humerus length, linolenic acid content and feed conversion ratio (Table 1).

Table 1. Genomic regions identified as signatures of selection in three breeds overlapped with QTL database.

Fst Top 1% Overlaps Candidate Regional Genes and Traits with Porcine QTL						
Group	CHR	START	END	FST	QTL	Genes
YN_DLY	Chr.1	26,823,251	26,823,255	0.93663199	Triglyceride level QTL (56,029)	DOLK, CRAT, PTPA, DOLPP1
	Chr.12	28,689,860	28,689,864	0.73554397	Hemoglobin QTL (22,161)	PNPO
	Chr.12	29,033,566	29,033,570	0.61874989	Loin muscle depth QTL (255,351)	NPEPPS, OSBPL7, SRCN2
	Chr.14	136,511,430	136,511,434	0.93663199	Number of mummified pigs QTL (178,885)	MYOF, CEP55, ADAM12,
YN_HN	Chr.8	38,946,696	38,946,700	0.93663199	Litter size QTL (130,289)	DCUN1D4, SPATA18, PAICS, PPAT
	Chr.11	47,893,515	47,893,519	0.61874989	Intramuscular fat content QTL (147,429)	UCHL3, KATNAL1
HN_DLY	Chr.1	100,049,958	141,109,887	0.93663199	Ear weight QTL (8853)	MAPK4
	Chr.4	77,799,657	77,799,661	0.93663199	Humerus length QTL (57,328)	UBE2V2
	Chr.4	79,493,441	79,493,445	0.61874989	Feed conversion ratio QTL (139,907)	ST18
	Chr.6	145,288,072	145,288,076	0.73554397	Linolenic acid content QTL (193,502)	PIEZO1, ACSF3, CTU2

4. Discussion

Henan is one of China's central provinces, and includes one of the earliest pig domestication areas, with a rich diversity of indigenous pig breeds. As an excellent, newly cultivated black pig breed, Yunan (YN) black pigs are characterized by desirable body shape, tender local pig flavor and robust disease resistance. In the present study, we analyzed the population genetic diversity of YN black, HN and commercial DLY pigs. The results revealed that genetic stratification has occurred in these three populations. As expected from their hybrid nature, the YN black pig population displayed a genetic variation in-between DLY and HN pigs. The genetic diversity of a population is always recognized as a key factor in ensuring the survival and/or evolution of a species. Our results indicated Yunan black pigs have a richer diversity and are thus a valuable contribution to China's germplasm bank.

In this study, we employed CLR and Fst to explore selection footprints within and between populations. The CLR method uses the combined likelihood of multiple markers to detect the selected sweeps in genomic regions [3]. CLR is especially robust in scenarios where changes in allele frequency occur fast [32]. The Fst method is classical and effective for detecting selection footprints in a single locus, based on population differentiation [33]. Voight et al. advised empirical cutoffs using the top 1% or 5% genome-wide, on all autosomes, to determine the significance of test statistic [34]. We adopted the same principle of data filtering to ensure high accuracy of the two methods in the present study, our results showed this strategy was more reasonable.

Selection signals can reflect loci and harboring genes that explain potential genetic mechanisms of phenotype polymorphism and provide scientific interpretation of the consequence of long-term domestication/breeding/evolution. In the results of CLR, we found that more genes related to fatty acid biosynthesis were selected in YN black pigs. We performed an extensive and accurate search of the human genome literature. One of the most interesting genes, LPIN1, acts as a magnesium-dependent phosphatidate phosphatase enzyme, which catalyzes the conversion of phosphatidic acid to diacylglycerol during triglyceride, phosphatidylcholine and phosphatidylethanolamine biosynthesis, and therefore controls the metabolism of fatty acids at different levels [35,36]. Additionally, the low-density lipoprotein receptor-related protein 5 (LRP5) gene may activate the canonical Wnt signaling pathway that controls cell fate determination and self-renewal during embryonic development and adult tissue regeneration [37]. For HN pigs, FOU1F1 is a pituitary-specific transcription factor which plays an essential role in the specification of the somatotroph, lactotroph and thyrotroph lineages, and specifically activates GH1, PRL and TSHβ transcription [38]. FBOX31 has a role in regulating the cell cycle as well as dendrite growth and neuronal migration [39]. Interestingly, genes related to growth were selected by a more intense artificial selection in the DLY pig breed; CCN3,essential for human hematopoietic stem or progenitor cells functional integrity [40], plays an important role in cardiovascular, skeletal development and fibrosis [41,42]. MORC is testis-specific in mouse but transcribed in multiple somatic tissues in human, and serves important biological functions in both meiotic and mitotic cells [43]. Meanwhile, the result of Fst has supported our findings mentioned above. As a newly cultivated pig breed of economic importance, leanness and flavor has consistently been considered as an objective trait of YN black pigs. When comparing YN black with DLY pigs, several genes were found related to growth rate: SP6 gene caused severe hypoplastic amelogenesis imperfecta [44] and LCORL was reported to be associated with adult height in individuals of African ancestry [45]. When comparing YN black with HN pigs, several selected genes were found to be related to brain and neuron development. PDZD8 has been shown to be required for calcium ion (Ca^{2+}) uptake in mammalian neurons [46]. Maternal embryonic leucine zipper kinase (MELK) regulates multipotent neural progenitor proliferation [47]. When comparing HN with DLY pigs, more selected genes were found to be related to metabolic procession. ACSF3 deficiency was the first human disorder identified and caused human metabolic disorders [48]. Furthermore, ACSF3 protein is a malonyl-CoA that could be used for de

novo fatty acid synthesis in mammalian mitochondria [49]. PIEZO1 plays a critical role via intracellular accumulation of glucose metabolites and regulating insulin secretion [50].

The functional enrichment analysis of genes usually only focuses on biological information, hence the intrinsic relationship between genes and important quantitative traits is usually less conclusive. QTL mapping can serve as a reference or potential clue to elucidate the identified signatures of selection. In this study, the pig QTL database and enrichment analyses were used to annotate economic traits associated with overlapped selective regions. Through the overlapped results, we confirmed the detected selection regions were not only related to fat deposition, but also to growth, and meat quality and fertility in YN black pigs and HN pigs, respectively.

5. Conclusions

In summary, we used the 70 k Porcine Functional Variants Chip to detect the heritable variation information among Yunan (YN) black pigs, Huainan (HN) pigs and DLY pigs. The characterization of genomic diversity and population structure demonstrated the direction for genetic assessment and the development of reasonable breeding strategies for YN black pigs. Moreover, CLR analysis was implemented for selective sweep detection within-populations and Fst estimation was conducted for selective signatures detection between-populations. Both methods identified several selective regions in accompanied representative groups associated with important quality traits. In addition, overlapping 14 tissue-specific promoters' annotation with the Fst top 1% selective regions systematically described the muscle-specific and hypothalamus-specific regulatory elements in YN black pigs specifically. Taken together, these results provide a valuable resource for further genetic research to improve the genome-assisted breeding of other important indigenous pigs in the future.

References

1. Briggs, H.M. *Modern Breeds of Livestock Hardcover*; Collier Macmillan Ltd.: New York, NY, USA, 1969; pp. 293–379. ISBN 978-002-314-720-3.
2. China National Commission of Animal Genetic Resources. In *Animal Genetic Resources in China: Pigs*; China Agriculture Press: Beijing, China, 2011; pp. 60–63, 441–444. ISBN 978-710-915-882-5.
3. Wang, K.; Wu, P.; Yang, Q.; Chen, D.; Zhou, J.; Jiang, A.; Ma, J.; Tang, Q.; Xiao, W.; Jiang, Y.; et al. Detection of Selection Signatures in Chinese Landrace and Yorkshire Pigs Based on Genotyping-by-Sequencing Data. *Front. Genet.* **2018**, *9*, 119. [CrossRef] [PubMed]
4. Ma, Y.; Zhang, H.; Zhang, Q.; Ding, X. Identification of selection footprints on the X chromosome in pig. *PLoS ONE* **2014**, *9*, e94911. [CrossRef] [PubMed]
5. Xu, J.; Ruan, Y.; Sun, J.; Shi, P.; Huang, J.; Dai, L.; Xiao, M.; Xu, H. Association analysis of PRKAA2 and MSMB polymorphisms and growth traits of Xiangsu hybrid pigs. *Genes* **2022**, *14*, 113. [CrossRef] [PubMed]
6. Li, H.; Xu, C.; Meng, F.; Yao, Z.; Fan, Z.; Yang, Y.; Meng, X.; Zhan, Y.; Sun, Y.; Ma, F.; et al. Genome-wide association studies for flesh color and intramuscular fat in (Duroc x landrace x large white) crossbred commercial pigs. *Genes* **2022**, *13*, 2131. [CrossRef] [PubMed]
7. Wang, H.; Wang, X.; Li, M.; Sun, H.; Chen, Q.; Yan, D.; Dong, X.; Pan, Y.; Lu, S. Genome-wide association study of growth traits in a four-way crossbred pig population. *Genes* **2022**, *13*, 1990. [CrossRef]
8. Li, M.; Tian, S.; Jin, L.; Zhou, G.; Li, Y.; Zhang, Y.; Wang, T.; Yeung, C.K.; Chen, L.; Ma, J.; et al. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nat. Genet.* **2013**, *45*, 1431–1438. [CrossRef]
9. Ai, H.; Fang, X.; Yang, B.; Huang, Z.; Chen, H.; Mao, L.; Zhang, F.; Zhang, L.; Cui, L.; He, W.; et al. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat. Genet.* **2015**, *47*, 217–225. [CrossRef]
10. 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. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat. Genet.* **2015**, *47*, 1141–1148. [CrossRef]
11. Ma, Y.; Zhang, S.; Zhang, K.; Fang, C.; Xie, S.; Du, X.; Li, X.; Ni, D.; Zhao, S. Genomic analysis to identify signatures of artificial selection and loci associated with important economic traits in duroc pigs. *G3 (Bethesda)* **2018**, *8*, 3617–3625. [CrossRef]
12. Chen, Y.; Wu, X.; Wang, J.; Hou, Y.; Liu, Y.; Wang, B.; Hu, X.; Zheng, X.; Zhang, X.; Ding, Y.; et al. Detection of selection signatures in Anqing six-end-white pigs based on resequencing data. *Genes* **2022**, *13*, 2310. [CrossRef]
13. Guo, L.; Sun, H.; Zhao, Q.; Xu, Z.; Zhang, Z.; Liu, D.; Qadri, Q.R.; Ma, P.; Wang, Q.; Pan, Y. Positive selection signatures in Anqing six-end-white pig population based on reduced-representation genome sequencing data. *Anim. Genet.* **2021**, *52*, 143–154. [CrossRef] [PubMed]
14. Diao, S.; Huang, S.; Chen, Z.; Teng, J.; Ma, Y.; Yuan, X.; Chen, Z.; Zhang, H.; Li, J.; Zhang, Z. Genome-wide signatures of selection detection in three south china indigenous pigs. *Genes* **2019**, *10*, 346. [CrossRef] [PubMed]
15. Zhao, P.; Yu, Y.; Feng, W.; Du, H.; Yu, J.; Kang, H.; Zheng, X.; Wang, Z.; Liu, G.E.; Ernst, C.W.; et al. Evidence of evolutionary history and selective sweeps in the genome of Meishan pig reveals its genetic and phenotypic characterization. *Gigascience* **2018**, *7*, giy058. [CrossRef] [PubMed]
16. Cao, R.; Feng, J.; Xu, Y.; Fang, Y.; Zhao, W.; Zhang, Z.; Zhang, Z.; Li, M.; Wang, Q.; Pan, Y. Genomic signatures reveal breeding effects of Lulai pigs. *Genes* **2022**, *13*, 1969. [CrossRef]
17. 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.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
18. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [CrossRef]
19. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [CrossRef]
20. Zhang, C.; Dong, S.S.; Xu, J.Y.; He, W.M.; Yang, T.L. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **2019**, *35*, 1786–1788. [CrossRef]
21. Nielsen, R.; Williamson, S.; Kim, Y.; Hubisz, M.J.; Clark, A.G.; Bustamante, C. Genomic scans for selective sweeps using SNP data. *Genome Res.* **2005**, *15*, 1566–1575. [CrossRef]
22. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [CrossRef]
23. 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.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef]
24. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **2010**, *11*, R14. [CrossRef]

25. Pan, Z.; Yao, Y.; Yin, H.; Cai, Z.; Wang, Y.; Bai, L.; Kern, C.; Halstead, M.; Chanthavixay, G.; Trakooljul, N.; et al. Pig genome functional annotation enhances the biological interpretation of complex traits and human disease. *Nat. Commun.* **2021**, *12*, 5848. [CrossRef] [PubMed]

26. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef] [PubMed]

27. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]

28. Huang Da, W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13. [CrossRef] [PubMed]

29. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [CrossRef] [PubMed]

30. Hu, Z.L.; Park, C.A.; Reecy, J.M. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* **2019**, *47*, D701–D710. [CrossRef]

31. Hu, Z.L.; Park, C.A.; Reecy, J.M. Bringing the Animal QTLdb and CorrDB into the future: Meeting new challenges and providing updated services. *Nucleic Acids Res.* **2022**, *50*, D956–D961. [CrossRef]

32. Chen, H.; Patterson, N.; Reich, D. Population differentiation as a test for selective sweeps. *Genome Res.* **2010**, *20*, 393–402. [CrossRef]

33. Gianola, D.; Simianer, H.; Qanbari, S. A two-step method for detecting selection signatures using genetic markers. *Genet. Res. (Camb)* **2010**, *92*, 141–155. [CrossRef] [PubMed]

34. Voight, B.F.; Kudaravalli, S.; Wen, X.; Pritchard, J.K. A map of recent positive selection in the human genome. *PLoS Biol.* **2006**, *4*, e72. [CrossRef]

35. Han, G.S.; Carman, G.M. Characterization of the human LPIN1-encoded phosphatidate phosphatase isoforms. *J. Biol. Chem.* **2010**, *285*, 14628–14638. [CrossRef] [PubMed]

36. Li, T.Y.; Song, L.; Sun, Y.; Li, J.; Yi, C.; Lam, S.M.; Xu, D.; Zhou, L.; Li, X.; Yang, Y.; et al. Tip60-mediated lipin 1 acetylation and ER translocation determine triacylglycerol synthesis rate. *Nat. Commun.* **2018**, *9*, 1916. [CrossRef]

37. Mao, J.; Wang, J.; Liu, B.; Pan, W.; Farr, G.H., 3rd; Flynn, C.; Yuan, H.; Takada, S.; Kimelman, D.; Li, L.; et al. Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Mol. Cell* **2001**, *7*, 801–809. [CrossRef]

38. Sobrier, M.L.; Tsai, Y.C.; Pérez, C.; Leheup, B.; Bouceba, T.; Duquesnoy, P.; Copin, B.; Sizova, D.; Penzo, A.; Stanger, B.Z.; et al. Functional characterization of a human POU1F1 mutation associated with isolated growth hormone deficiency: A novel etiology for IGHD. *Hum. Mol. Genet.* **2016**, *25*, 472–483. [CrossRef]

39. Fishilevich, S.; Nudel, R.; Rappaport, N.; Hadar, R.; Plaschkes, I.; Iny Stein, T.; Rosen, N.; Kohn, A.; Twik, M.; Safran, M.; et al. GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. *Database* **2017**, *2017*, bax028. [CrossRef]

40. Gupta, R.; Hong, D.; Iborra, F.; Sarno, S.; Enver, T. NOV (CCN3) functions as a regulator of human hematopoietic stem or progenitor cells. *Science* **2007**, *316*, 590–593. [CrossRef]

41. Shimoyama, T.; Hiraoka, S.; Takemoto, M.; Koshizaka, M.; Tokuyama, H.; Tokuyama, T.; Watanabe, A.; Fujimoto, M.; Kawamura, H.; Sato, S.; et al. CCN3 inhibits neointimal hyperplasia through modulation of smooth muscle cell growth and migration. *Arterioscler. Thromb. Vasc. Biol.* **2010**, *30*, 675–682. [CrossRef]

42. Sakamoto, K.; Yamaguchi, S.; Ando, R.; Miyawaki, A.; Kabasawa, Y.; Takagi, M.; Li, C.L.; Perbal, B.; Katsube, K. The nephroblastoma overexpressed gene (NOV/ccn3) protein associates with Notch1 extracellular domain and inhibits myoblast differentiation via Notch signaling pathway. *J. Biol. Chem.* **2002**, *277*, 29399–29405. [CrossRef]

43. Inoue, N.; Hess, K.D.; Moreadith, R.W.; Richardson, L.L.; Handel, M.A.; Watson, M.L.; Zinn, A.R. New gene family defined by MORC, a nuclear protein required for mouse spermatogenesis. *Hum. Mol. Genet.* **1999**, *8*, 1201–1207. [CrossRef] [PubMed]

44. Kim, Y.J.; Lee, Y.; Zhang, H.; Song, J.S.; Hu, J.C.; Simmer, J.P.; Kim, J.W. A Novel De Novo SP6 mutation causes severe hypoplastic amelogenesis imperfecta. *Genes* **2021**, *12*, 346. [CrossRef]

45. N'Diaye, A.; Chen, G.K.; Palmer, C.D.; Ge, B.; Tayo, B.; Mathias, R.A.; Ding, J.; Nalls, M.A.; Adeyemo, A.; Adoue, V.; et al. Identification, replication, and fine-mapping of Loci associated with adult height in individuals of african ancestry. *PLoS Genet.* **2011**, *7*, e1002298. [CrossRef]

46. Hirabayashi, Y.; Kwon, S.K.; Paek, H.; Pernice, W.M.; Paul, M.A.; Lee, J.; Erfani, P.; Raczkowski, A.; Petrey, D.S.; Pon, L.A.; et al. ER-mitochondria tethering by PDZD8 regulates Ca^{2+} dynamics in mammalian neurons. *Science* **2017**, *358*, 623–630. [CrossRef] [PubMed]

47. Nakano, I.; Paucar, A.A.; Bajpai, R.; Dougherty, J.D.; Zewail, A.; Kelly, T.K.; Kim, K.J.; Ou, J.; Groszer, M.; Imura, T.; et al. Maternal embryonic leucine zipper kinase (MELK) regulates multipotent neural progenitor proliferation. *J. Cell Biol.* **2005**, *170*, 413–427. [CrossRef] [PubMed]

48. Sloan, J.L.; Johnston, J.J.; Manoli, I.; Chandler, R.J.; Krause, C.; Carrillo-Carrasco, N.; Chandrasekaran, S.D.; Sysol, J.R.; O'Brien, K.; Hauser, N.S.; et al. Exome sequencing identifies ACSF3 as a cause of combined malonic and methylmalonic aciduria. *Nat. Genet.* **2011**, *43*, 883–886. [CrossRef]

49. Witkowski, A.; Thweatt, J.; Smith, S. Mammalian ACSF3 protein is a malonyl-CoA synthetase that supplies the chain extender units for mitochondrial fatty acid synthesis. *J. Biol. Chem.* **2011**, *286*, 33729–33736. [CrossRef]

50. Ye, Y.; Barghouth, M.; Dou, H.; Luan, C.; Wang, Y.; Karagiannopoulos, A.; Jiang, X.; Krus, U.; Fex, M.; Zhang, Q.; et al. A critical role of the mechanosensor PIEZO1 in glucose-induced insulin secretion in pancreatic β-cells. *Nat. Commun.* **2022**, *13*, 4237. [CrossRef]