BMC Genomics

RESEARCH

Open Access

The effects of runs-of-homozygosity on pig domestication and breeding

Lin Tao[1,3†], Hang Liu[1,3†], Adeniyi C. Adeola[1], Hai-Bing Xie[1*], Shu-Tang Feng[2,5*] and Ya-Ping Zhang[1,4*]

Abstract

Background Since their domestication, recent inbreeding together with intensive artificial selection and population bottlenecks have allowed the prevalence of deleterious mutations and the increase of runs-of-homozygosity (ROH) in domestic pigs. This makes pigs a good model to understand the genetic underpinnings of inbreeding depression.

Results Here we integrated a comprehensive dataset comprising 7239 domesticated pigs and wild boars genotyped by single nucleotide polymorphism (SNP) chips, along with phenotypic data encompassing growth, reproduction and disease-associated traits. Our study revealed differential ROH landscapes during domestication and artificial selection of Eurasian pigs. We observed associations between ROH burden and phenotypic traits such as body conformation and susceptibility to diseases like scrotal hernia. By examining associations of whole-genome and regional ROH burden with gene expression, we identified specific genes and pathways affected by inbreeding depression. Associations of regional ROH burden with gene expression also enabled the discovery of novel regulatory elements. Lastly, we inferred recessive lethal mutations by examining depletion of ROH in an inbred population with relatively small sample size, following by fine mapping with sequencing data.

Conclusions These findings suggested that both phenotypic and genetic variations have been reshaped by inbreeding, and provided insights to the genetic mechanisms underlying inbreeding depression.

Keywords Inbreeding depression, Deleterious mutations, Inbred pigs, RNA-seq

Background

Intensive artificial selection, inbreeding and population bottlenecks during domestication and breeding allowed the increase in runs-of-homozygosity (ROH) and the prevalence of deleterious mutations. Domestic pigs play dual roles as pivotal agricultural animals and valuable biomedical models [1]. Pig breeding has evolved along two primary trajectories: genetic improvement within breeds for growth, reproduction, and disease resistance [2], and the development of miniature/inbred lines for laboratory research [3]. These processes present opportunities to assess how ROH impact economic traits in pigs. The relevance of ROH in understanding inbreeding depression has been highlighted by its enrichment with deleterious variations [4]. Thus, detecting and mitigating deleterious mutations are critical for future breeding efforts [5]. Despite previous efforts [6–8], the impact of

†Lin Tao and Hang Liu contributed equally to this work.

*Correspondence:
Hai-Bing Xie
xiehb@mail.kiz.ac.cn
Shu-Tang Feng
fst@grand-livescience.com
Ya-Ping Zhang
zhangyp@mail.kiz.ac.cn
[1] Key Laboratory of Genetic Evolution & Animal Models and Yunnan Key Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[2] Beijing Grand-Life Science and Technology Company, Beijing 102206, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China
[4] State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, School of Life Science, Yunnan University, Kunming 650091, China
[5] Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

ROH on economic traits has not been fully understood in pigs. Furthermore, several fundamental questions regarding ROH in pig domestication and breeding remain unsolved. For instance, what is the ROH landscape during domestication, and which genes and pathways contribute to inbreeding depression at the gene expression level? Utilizing ROH as a proxy for homozygosity analysis may provide insights into these unsolved questions.

Identification of lethal genotypes causing pregnancy loss and perinatal mortality poses challenges due to their absence in living individuals. Deficit of homozygosity and haplotype has been applied to infer genetic causes of recessive lethality in humans and farm animals, which however relies heavily on ultra-large sample size [9–11]. Therefore, the existing research for pigs was limited in highly managed commercial populations [12–14], leading to our insufficient understanding of recessive lethality. In contrast, the availability of inbred population offers a feasible alternative for inferring recessive deleterious mutations with smaller sample size, given that inbred individuals are expected to carry enrich recessive deleterious alleles that trigger the risk of inbreeding depression, and that only on average 1% heterozygosity remains in genome of individuals that have been mated brother×sister or parent×offspring for 20 consecutive generations [15]. Inbred pig lines, such as the ZhongXu Wuzhishan minipig inbred line (ZXWIPIG) and Banna minipig both of which have been inbred for over 20 generations [16, 17], provide unique models due to relaxed ethical constraints compared to humans, facilitating the study of inbreeding depression in large mammals.

Here, we leverage ROH from 6724 domestic pigs, 461 wild boars, and 54 feral pigs to estimate the genetic diversity of global pigs, to describe the ROH landscape during pig domestication, and to link recent inbreeding to deleterious variations associated with both economic traits and inbreeding depression (Fig. 1). This dataset consists of genotypes from 106 populations of domesticated pigs, feral pigs from Europe, America and Oceania, and wild boars from Asia, Europe and Africa (Table S1). We test for the connection between ROH burden and a series of growth, reproduction and disease traits in commercial and synthetic breeds. Additionally, we investigate the transcriptomic effects of ROH in commercial pig herds using sequencing data. Furthermore, we develop an alternative approach to infer recessive lethal loci associated with inbreeding depression by the depletion of ROH in an inbred population with subsequent validation using sequencing variations.

Methods

Genotypes and phenotypes

We merged autosomal genotype data from a total of 7239 pigs, including 6724 domestic pigs, 461 wild boars, and 54 feral pigs (Table S1). These animals distributed in 41 countries were genotyped by Illumina PorcineSNP60 BeadChip or GeneSeek GGP Porcine 50 K/80 K BeadChip. All genomic coordinates were updated to the sus11.1 reference genome. Missing genotypes were imputed by SHAPEIT5 [18] with a reference panel of 937 genomes [19], resulting in a set of 61,114 autosomal SNPs.

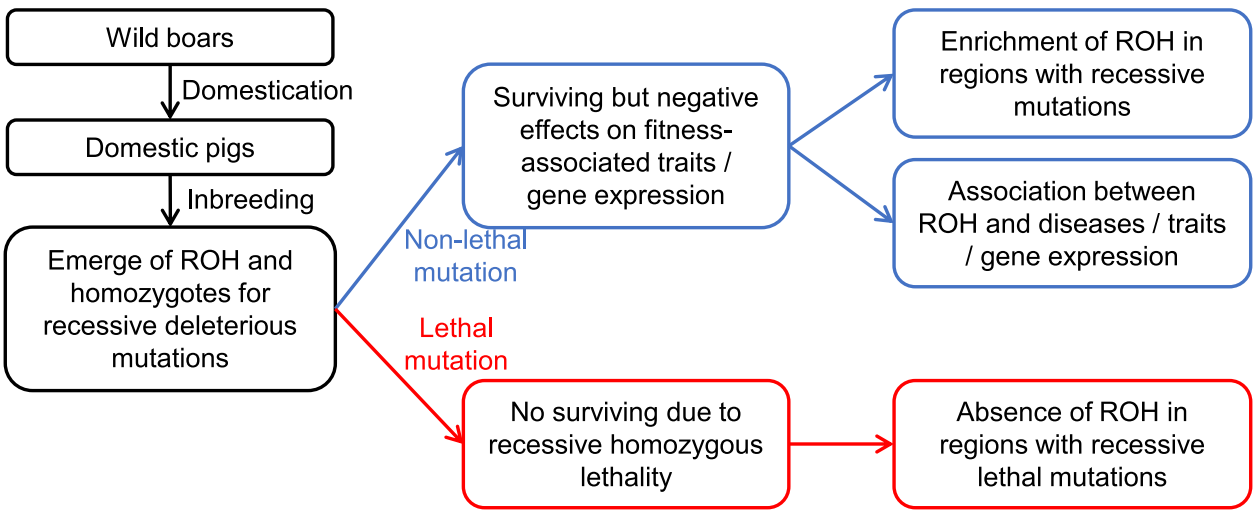

Fig. 1 Potential mechanisms for runs-of-homozygosity (ROH) in pig domestication and breeding. Population bottleneck, strong artificial selection and recent inbreeding during domestication may change ROH landscape. In pig populations, recessive non-lethal mutations (blue) enriched in ROH may be linked to phenotypes, diseases and gene expression, whereas lethal mutations (red) may lead to the depletion of ROH

To identify and manage duplicate samples, relationship inference was conducted using KING v2.3.0 with the "–duplicate" parameter [20]. Samples were treated as duplicates if they met two criteria: (1) a kinship coefficient of at least 0.5, and (2) assignment to the same population or breed across different studies, often with shared authorship. Duplicate individuals were randomly excluded from this analysis. However, individuals with differential phenotype values from the same study, indicative of potential monozygotic twins, were kept.

The phenotypes we focused on were related to growth, reproduction, and health. Growth performance included body/carcass weight, body conformation (length, height, chest circumference, chest width, hip width, cannon bone circumference, leg buttock circumference, rib number), metabolic traits (acid and neutral detergent fiber), and meat quality traits (backfat thickness (BF), loin muscle depth (LMD), intramuscular fat content (IMF) and marbling scores (MS)). Reproduction performance included teat number, litter size, the average birth weight (ABW) and birth interval of piglets (ABI), and piglet mortality at birth. The health traits were the susceptibility to post-weaning diarrhea and scrotal hernia.

Principal component analysis (PCA) and recent demographic history

PCA was performed by Plink v1.9 [21] to assess genetic relationships and population structure among samples. Effective population size (Ne) of commercial and minipigs was estimated by GONE based on linkage disequilibrium with default parameters to infer recent demographic history [22].

Genetic differentiation during domestication of Eurasian pigs

Weir and Cockerham's F_{ST} was calculated SNP-by-SNP in VCFtools v0.1.13 by comparing Eurasian domestic pigs to their local wild relatives [23]. Sites with negative F_{ST} values were removed for analyses.

Phenotypic association test

ROH were called by Plink v1.9 with a sliding window scan of 50 SNPs in 5 Mb length regions [21]. One heterozygote was allowed to account for possible genotyping errors to avoid underestimation. A minimal length of 1 Mb was used to exclude the relatively short stretches due to linkage disequilibrium blocks. To avoid the effect of homozygous segments in regions with sparse SNPs, we set the minimum marker density to one SNP per 50 kb and the maximum distance of two adjacent SNPs within one homozygous segment to 100 kb.

Individual inbreeding coefficient F_{ROH} was calculated to quantify cumulative whole-genome ROH burden by using an autosomal length of 2266 Mb. To test for the effects of whole-genome ROH burden on economic traits, we first employed standard linear and logistic regression models for quantitative and binary traits, respectively. To account for population structure, we then run general linear mixed models (linear and logistic mixed models for quantitative and binary traits, respectively), which were implemented in GMMAT v1.4.2 [24]. A genetic relatedness matrix calculated by GEMMA v0.98.5 was included in general linear mixed models as a random effect to correct for the nonindependence of observations [25]. Fixed effects were included in these models if it was available in original research. Wald statistic was used to test the null hypothesis of no association. A significant association was declared when P value < 0.05.

To test the roles of significant common GWAS variations, we reanalyzed the association of F_{ROH} with phenotypes by filtering out samples that possessed ROH covered GWAS variations reported by the original research [26]. If the association is no longer significant, inbreeding depression is primarily caused by the ROH encompassing these significant GWAS variations. A more significant association means that these significant GWAS variations do not play primary roles in inbreeding depression.

Gene expression association test

To explore the effects of ROH on gene expression, we analyzed the RNA sequencing (RNA-seq) data from duodenum, liver, and muscle tissues in three commercial breeds (Large White, Duroc and Landrace), which belongs to the GENE-SWitCH project [27]. We initially had 100 samples for each breed and each tissue. However, 16 samples were removed due to low quality, including four from liver, four from duodenum, and eight from muscle. The details of sample information were described by Crespo-Piazuelo et al. (2022).

The SRA files downloaded from NCBI were converted to fastq format using fastq-dump in SRA Toolkit, followed quality control by Trimmomatic v0.39 [28]. Filtered reads were aligned to sus11.1 reference genome with HISAT2 v2.2.1 [29]. Gene expression matrix was returned by StringTie v2.1.2 in the form of transcripts per million [30]. PCA was performed by R package PCAtools v2.16.0 [31].

Genotypes (VCF format) of these 300 individuals were downloaded from GigaScience Database [32]; however, the sample information was missed, leading to unclear corresponding relationship between whole-genome sequencing (WGS) and RNA-seq data. For computational simplicity, SNPs were called for each tissue by RNA-seq data following best practices workflows in GATK4 v4.2.0.0 [33]. To match the sample identity of

two datasets, we combined the VCF files generated by WGS and RNA-seq. Then genetic distance (1-IBS) was calculated using Plink v1.9 [21] and visualized by ggtree v3.12.0 with the clustering method of ward.D2 [34]. We expected that every WGS sample was clustered with three RNA-seq samples, with one sample for each tissue.

ROH were called by Plink v1.9 with the same parameters in phenotypic association test [21]. The transcriptomic effects of ROH were detected at the levels of whole-genome and region. Briefly, the expression of each gene was fitted as the function of whole-genome (F_{ROH}) or regional ROH burden using the following model

$$\boldsymbol{y} = \boldsymbol{W\alpha} + \boldsymbol{x\beta} + \boldsymbol{g} + \boldsymbol{e}\ with\ \boldsymbol{g} \sim N(0, \sigma_g^2\boldsymbol{G}), \boldsymbol{e} \sim N(0, \sigma_e^2\boldsymbol{I}),$$

where $\boldsymbol{y}$ is gene expression, $\boldsymbol{\alpha}$ was covariates (fixed factors) of a column 1s, breed, sex and the first three PCs, $\boldsymbol{\beta}$ was effect size of ROH burden (F_{ROH} or the proportion of ROH segment for each 1-Mb non-overlapped window), $\mathbf{W}$ and $\mathbf{x}$ were corresponding design matrices, g was the genotypic additive value, and e was a random error. $\mathbf{G}$ was the kinship matrix and $\mathbf{I}$ was the identity matrix. Wald statistic was used to test the null hypothesis $\beta = 0$ in GEMMA v0.98.5 [25]. Multiple tests were corrected by the Bonferroni method. For significant window-expression associations, *cis*-regulation was defined as the associations in which the distance between the ROH window and gene was no more than 1 Mb, otherwise *trans*-regulation. R package clusterProfiler v4.10.0 was employed to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses, declaring significance at q-value < 0.05 [35].

Genomic Evolutionary Rate Profiling (GERP) scores

GERP scores based on the human reference genome version hg19 were downloaded (accessed from https://hgdownload.soe.ucsc.edu/gbdb/hg19/bbi/All_hg19_RS.bw). This file was generated by an alignment of 35 other mammalian species which did not contain the domestic pig. Given the biases resulted from inclusion of the focal reference genome, exclusion of domestic pig in alignment has been suggested to calculate conservation scores [36, 37]. Genome positions were remapped to pig reference genome version sus11.1 with liftOver and the chain file (accessed from https://hgdownload.cse.ucsc.edu/goldenPath/hg19/liftOver/hg19ToSusScr11.over.chain.gz).

We calculated the average GERP scores inside and outside ROH for each individual in the population with inbreeding depression. This was done for deleterious mutations (GERP > 0) in two models, including the additive model with both homozygotes and heterozygotes, and the recessive model only with homozygotes. The information of ancestral allele was downloaded from a previous study which was generated by aligning six suids [38].

Inference of lethal mutations

To identify depletion of ROH, we calculated the incidence of ROH for each SNP with detectRUNS v0.9.6 [39]. This analysis was initially done with all pigs, European and Asian samples and failed to identify evidence of lethality. Given the higher risk of inbreeding depression in inbred animals, we limited the analysis in ZXWIPIG of generations 17 to 22 [16], a new genetic resource of inbred pigs certified by the Chinese Association for Laboratory Animal Sciences in 2022 (NO: 00005). To mitigate genotyping technology effect, ROH were called using WGS data from one ZXWIPIG and six non-inbred Wuzhishan pigs with the same criteria, and were used to narrow the candidate regions. To pinpoint putative lethal sites, we calculated genotypes frequency and focused on variations devoid of altered homozygote in a reference panel of 937 genomes [19]. Variations classified as "stopgain", "stoploss" and frameshift substitution were specifically treated as putative lethal sites. Functional annotation of variations was performed with ANNOVAR (2020–06-07 version) [40]. The distribution of these sites was further confirmed by IAnimal database consisting of 1311 pigs across 65 populations [41].

Results

Global patterns of genetic diversity across wild boars and domestic pigs

The domestication of Eurasian domestic pigs from local wild boars occurred independently, approximately 9000 to 10,000 years ago [42]. To examine the genetic diversity of wild boars and domestic pigs, we merged genotype data from 12 previously published studies [16, 17, 26, 43–51]. This dataset encompassed wild and domestic individuals from Asia, Europe and Africa, commercial breeds, and feral populations, representing the global genetic diversity. PCA was employed to initially examine the relationship among samples. The PCA revealed a clear separation between Asian and European populations on the first principal component (Fig. 2), supporting the independent origins of Eurasian domestic pigs [42, 52, 53]. Similar patterns were observed in separate analyses of wild boars and domestic pigs (Fig. S1).

We then inferred the recent demographic history of commercial breeds and selected minipigs. A historical increase of Ne was evident across all breeds, except for ZXWIPIG (Fig. S2), indicating breed formation processes. The recent decrease of Ne reflected the loss of genetic diversity, likely attributable to recent inbreeding events. Lastly, we quantified the inbreeding coefficient of global populations using F_{ROH} based on the length of

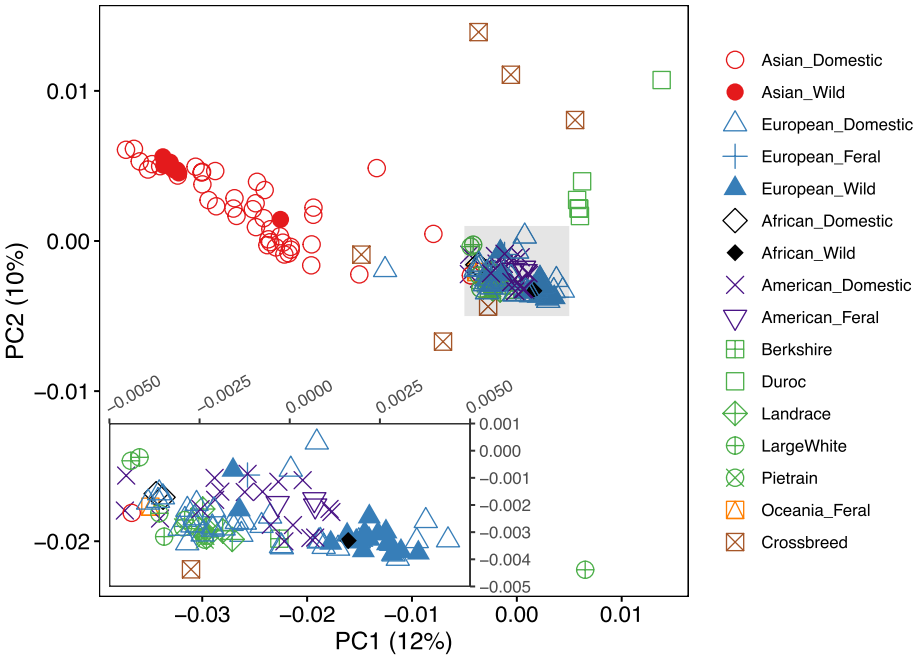

Fig. 2 Genetic relationship of global pigs. Each point denotes one breed/population by using the averages of the first and second principal components (PC1 and PC2). The numbers in parentheses represent the percentage of variance explained by PC1 and PC2

autosomal ROH (Fig. S3). Of all populations, ZXWIPIG was characterized by the highest averaged F_{ROH} of 0.52, with a maximum up to 0.70. Asian boars, expect for Thailand and Korean populations, showed lower inbreeding compared to European boars. Relatively small estimates of Ne, especially in commercial breeds and minipigs, contributed to a substantial portion of their genomes comprising ROH.

Differential ROH landscapes in domestication and artificial selection of Eurasian pigs

To examine how domestication and artificial selection have shaped ROH landscape, we compared the incidence of ROH for each SNP between wild and domestic pigs, quantified as the ROH frequency difference ($\%\Delta_{domestic\ -\ wild}$; Fig. 3). Eurasian domesticated counterparts exhibited a significantly higher incidence of ROH compared to wild boars, as indicated by $\%\Delta_{domestic—wild} > 0$ (Mann–Whitney

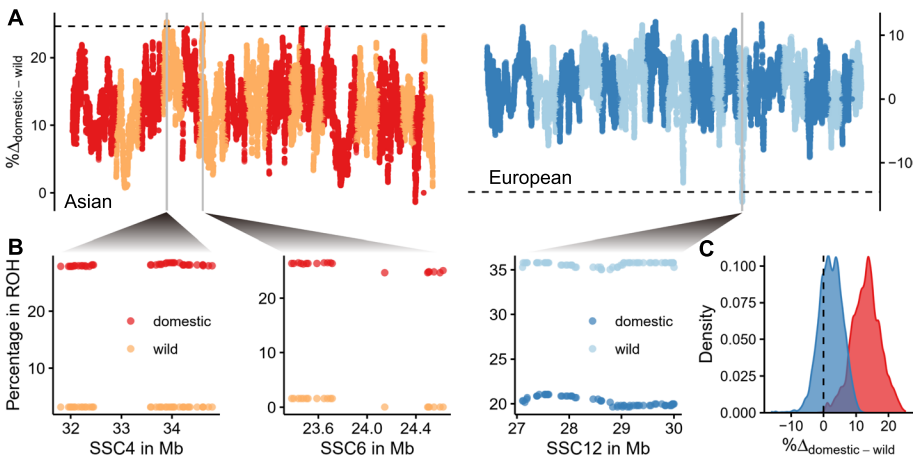

Fig. 3 ROH landscape in Eurasian pig domestication. **A** Manhattan plots of genome-wide ROH frequency difference (%Δ) between wild and domestic pigs. The *x*-axis is *Sus scrofa* chromosomes (SSC) 1–18. The dashed lines denote the thresholds of top 1‰ (right: 24.65 and left: -14.62). **B** The ROH frequency of top 1‰ sites in wild and domestic pigs. **C** Density distribution plots of %Δ indicates an increase of ROH in the domestication of Eurasian pigs

U test, $P=0$ for both Asia and Europe; Fig. 3). Notably, Asian animals showed a predominantly positive ROH frequency difference, influenced by the genetic diversity of Asian boars and the inclusion of minipigs (such as Xiang, Bamaxiang, and Congjiangxiang) and extremely inbred lines (ZXWIPIG and Banna minipig) in Asian domestic pigs.

To explore the relationship between domestication/artificial selection and the ROH landscape further, we assessed the linearity between genetic differentiation and the absolute value of $\%\Delta_{domestic-wild}$. Compared with Asian populations without significant correlation at the level of whole-genome, a positive correlation was observed in European populations ($r=0.02$, $P=4.4\times10^{-6}$), especially for SSC8, SSC12 and SSC14 (Fig. S4A–B). Focusing on SNPs with the top 1‰ absolute value of $\%\Delta_{domestic-wild}$, we identified distinct genetic signatures: two positive peaks on *Sus scrofa* chromosome 4 (SSC4) and SSC6 in Asian breeds and a negative signature on SSC12 in European populations (Fig. 3B). Notably, these signatures did not overlap with the top 1‰ domestication signatures of Eurasian pigs (top 61 SNPs). However, the genetic differentiation of the top 1‰ signature of European ROH frequency difference was significantly higher compared to other remaining SNPs ($P=2.6\times10^{-11}$, Mann–Whitney *U* test). Further investigation using Pig QTLdb release 53 (accessed July 2024) [54] revealed that these ROH signals mapped to quantitative trait loci (SSC4: teat number, feed intake, body circumference; SSC6: tuberculosis susceptibility, sperm motility; SSC12: lean meat percentage, teat number, hemoglobin), reflecting the artificial selection for growth, reproduction and health traits.

Associations of economic traits with inbreeding level

It is hypothesized that the ubiquity of ROH can be linked to both production performance and recessive genetic disease in domestic pigs (Fig. 1). To test this hypothesis, we utilized data from 5437 pigs in which phenotype records were available. Additionally, we incorporated data from 3286 animals across three other studies [55–57] in which genotypes were coded as 012 and could not be merged with others coded as AGCT. Association tests were performed for data from each research rather than across research due to differential phenotype measures. The traits tested are classified into reproduction, growth and disease. Non-significant association of whole-genome ROH burden was observed for most traits. However, the number of observed associations was significantly greater than what was expected under the null hypothesis of no associations (6 observed associations versus 1.7 expected associations under the null, $P=0.006$, binomial test; Fig. 4).

In an F_3 Eurasian population, scrotal hernia (SH) showed a significant association ($\beta=16.490$ and $P=0.023$), while in Sujiang pig (a Chinese synthesis breed), ROH burden was associated with hip width (HW; $\beta=$-8.980 cm and $P=0.002$), chest width (CWD; $\beta=$-8.376 cm and $P=0.018$), chest circumference (CC; $\beta=$-15.043 cm and $P=0.013$), body weight (BW; $\beta=$-39.167 kg and $P=3.9\times10^{-4}$), and body length (BL; $\beta=$-21.147 cm and $P=2.2\times10^{-4}$). For SH, affected individuals exhibited significantly higher inbreeding coefficients compared to unaffected counterparts (0.103 ± 0.010 versus 0.073 ± 0.002, $P<0.002$, Mann–Whitney *U* test). When using traditional linear regression, all six associations remained, with the emergence of reproduction traits, including average birth weight (AWB, $\beta=2.560$ kg, $P=0.003$) and piglet mortality at birth (MORT1: $\beta=0.032$, $P=0.010$; MORT2: $\beta=0.096$, $P=2.08\times10^{-7}$; Fig. S5 and Table S2). Therefore, estimating inbreeding depression was partly influenced by population structure.

To further understand the genetic basis of inbreeding depression by the link of ROH burden and phenotypes, we first focused on the intersections of ROH and genome-wide association study (GWAS) signals in Sujiang pig. Of five body conformation traits with significant inbreeding depression, only CC has 13 GWAS variations [26], all of which fall into ROH of at least one sample. We hypothesize that inbreeding depression is not primarily caused by the ROH encompassing common GWAS variations. To test this, we reassessed the association by filtering out samples with ROH encompassing the aforementioned GWAS variations. Interestingly, this resulted in a more significant association ($P=0.005$), suggesting that inbreeding depression of CC was not primarily driven by common GWAS variations. Moreover, we compared the GERP scores between ROH and non-ROH regions. The GERP score is a widely used method based upon evolutionary conservation to predict deleterious mutations. For a variation, GERP ≤ 0 means non-deleteriousness and a positive GERP score indicates the degree of deleteriousness [58]. The average GERP score was significantly higher in ROH than in non-ROH regions for both additive and recessive models ($P=3.81\times10^{-8}$ and 2.98×10^{-23}, Mann–Whitney *U* tests). For the F_3 population with inbreeding depression on SH, we couldn't assign ancestral and derived alleles due to the numerical codes of genotypes. These results supported these recessive deleterious mutations rather than common GWAS variations underpinned inbreeding depression.

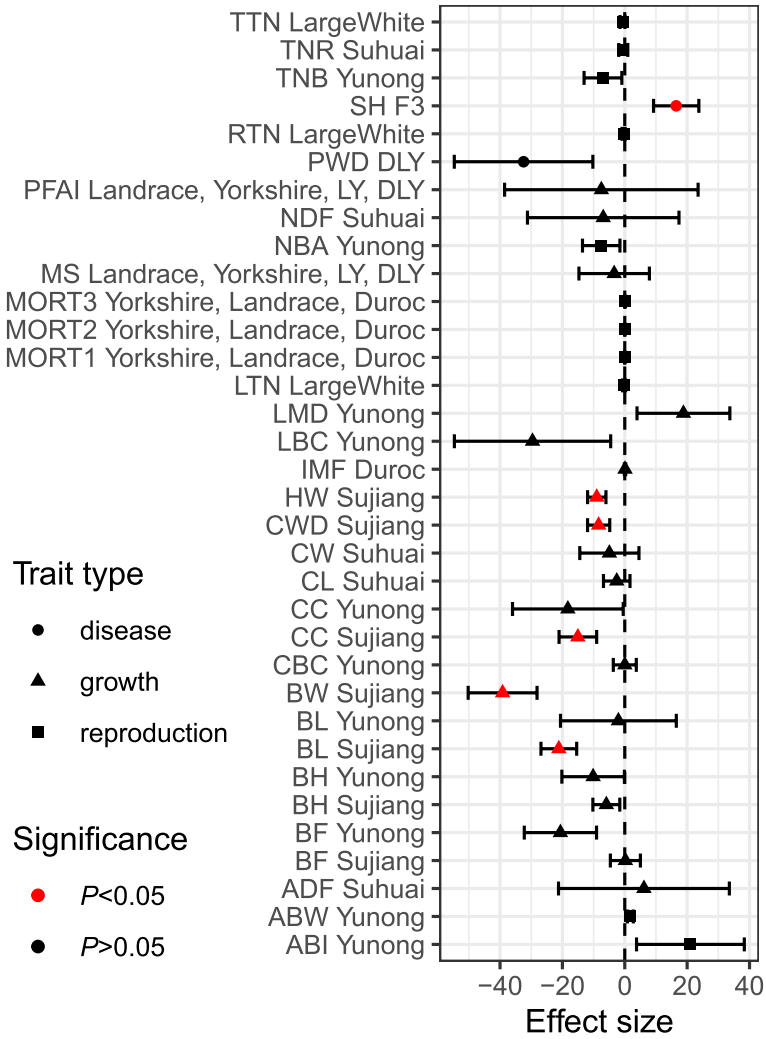

Fig. 4 Association of F_{ROH} with economic traits. ABI: the average birth interval of piglets, AWB: the average birth weight, ADF: acid detergent fiber, BF: backfat thickness, BH: body height, BL: body length, BW: body weight, CBC: cannon bone circumference, CC: chest circumference, CL: carcass length, CW: carcass weight, CWD: chest width, HW: hip width, IMF: intramuscular fat, LBC: legs buttocks circumference, LMD: loin muscle depth, LNT: left teat number, MS: marbling scores, NBA: number born alive, NDF: neutral detergent fiber, PFAI: proportion of fat areas in the image, PWD: post-weaning diarrhea, RNT: right teat number, SH: scrotal herni, TNB: total number of piglets born, TNR: total number of ribs, TTN: total teat number. MORT*i* denotes piglet mortality at birth for the i^{th} parity ($i = 1, 2, 3$)

Transcriptomic effects of ROH

To understand the effects of ROH on gene expression, we analyzed whole-genome sequencing and RNA sequencing data from three commercial breeds (Duroc, Landrace and Large White), sourced from GENE-SWitCH project [27]. Given the genetic background and tissue expression pattern across three herds (Fig. 5A and S6), we considered breed as a fixed effect to increase sample size and statistical power. Transcriptomic effects of ROH were detected at two levels. First, we checked the inbreeding depression on gene expression using inbreeding coefficient F_{ROH}. This resulted in 18, 71 and 32 genes of which the expression was affected by whole-genome ROH burden in liver, muscle and duodenum, respectively, without intersection across tissues (Fig. 5B). Function enrichment analysis revealed overrepresentation of pathways such as basal cell carcinoma and signaling pathways regulating pluripotency of stem cells in muscles, and ovarian steroidogenesis and SNARE interactions in vesicular transport in duodenum (Fig. 5C). These results indicated tissue-specific inbreeding depression on gene expression.

Second, we investigated the effects of regional ROH burden, quantified by the proportion of ROH in each 1-Mb window, on expression level for each gene. Unlike traditional expression quantitative trait loci (eQTL), a

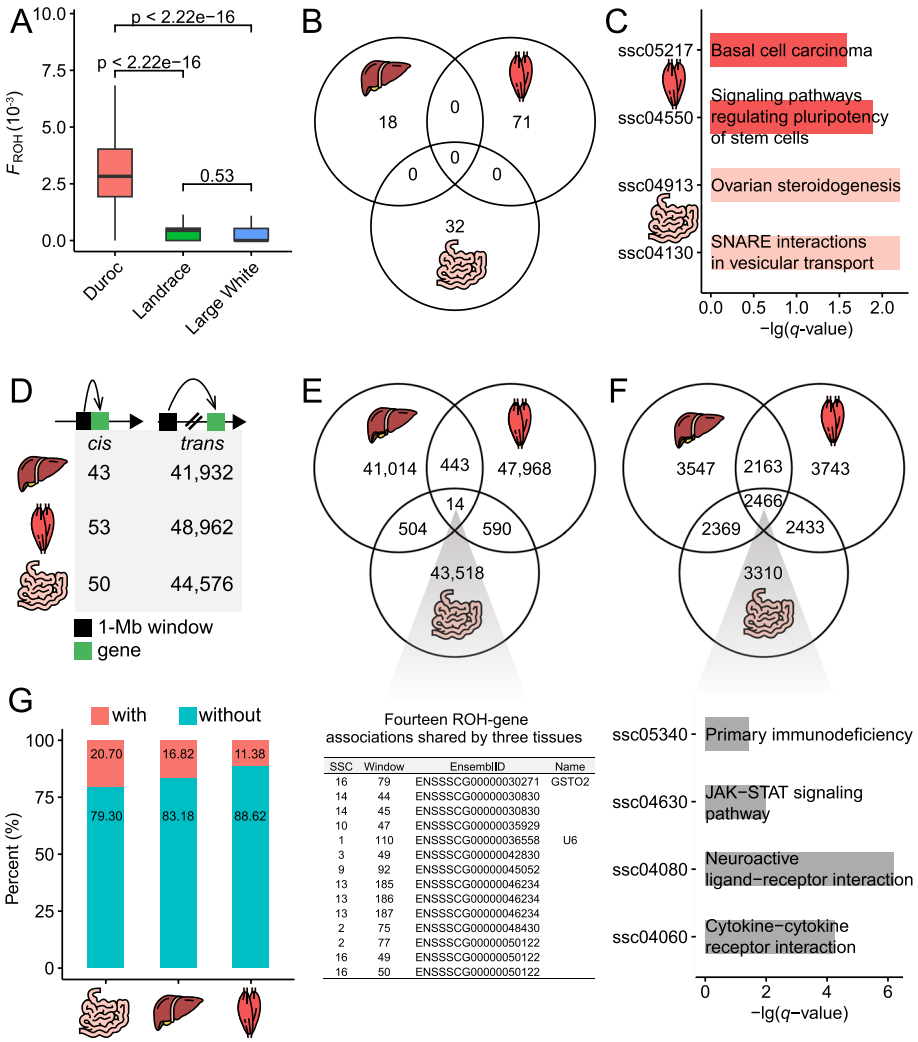

Fig. 5 The effects of ROH on gene expression in duodenum, liver, and muscle tissues. **A** Inbreeding coefficients of Duroc, Landrace, and Large White. **B** Genes of which the expression is influenced by F_{ROH} are not shared by three tissues. **C** Function enrichment of genes of which expression levels are linked to F_{ROH}. **D** Classification of the associations between regional ROH burden and gene expression. **E** significant associations between ROH burden in 1-Mb window and gene expression. *GSTO2*: glutathione S-transferase omega 2. **F** Genes of which expression levels are significantly affected by regional ROH burden. **G** The distribution of expression quantitative trait loci (eQTL) in the associations between regional ROH burden and gene expression

majority of genes influenced by regional ROH burden exhibited *trans*-effects rather than *cis*-effects (Fig. 5D). This difference was ascribed to the sharply reduced variation number due to increased length of markers when using regional ROH burden. Only 14 significant associations between regional ROH burden and gene expression were shared across these three tissues (Fig. 5E). The expression of 2466 genes (7.8%) were influenced by regional ROH burden in muscle, liver, and duodenum simultaneously (Fig. 5F). These shared genes were enriched to disease-associated and fundamental pathways (Fig. 5F), possibly suggesting the common biological pathways for inbreeding depression on gene expression. Comparing our findings to the eQTL list from original study [27], we found that 79%–89% associations identified here were novel and not previously reported by eQTL mapping (Fig. 5G), highlighting the potential of ROH as regulatory elements. For example, the shared association SSC16:79-*GSTO2* gene was identified in this study. *GSTO2* encodes an omega class glutathione S-transferase and is involved in some recessive genetic diseases, including hypothyroidism [59], primary open angle glaucoma [60] and asthma [61]. The associated window encompassed two Iroquois genes (*IRX2* and *IRX4*) influencing development of heart [62] and somites [63].

Recessive lethal mutations are inferred by ROH in inbred pigs

ROH represent homozygous regions in genome due to identical by descent and are enriched for deleterious variations [64, 65]. If the locus is recessive lethal, in principle we cannot observe ROH covering homozygotes of the lethal allele in any living individuals under the assumption that it is not effectively purged from focal populations. Therefore, we hypothesize that the depletion of ROH indicates the presence of recessive lethal loci (Fig. 1). We tested this hypothesis across global populations, including wild, feral and domestic individuals, and did not reveal evidence of recessive lethality. When restricted in European and Asian domestic pigs, no evidence was found. These results likely suggested differential purifying selection on lethal loci between breeds/populations, as indicated by previous studies on identifying lethal haplotype [12–14].

Given that this method was heavily depended on large sample size and that breeds with high F_{ROH} tended to carry more deleterious alleles in relatively few heterozygous sites (such as on average 1% heterozygosity in genome of F_{20} inbred individuals), we limited our analysis in 96 individuals from ZXWIPIG, characterized by the highest F_{ROH} in our study (Fig. S3). Consequently, cumulative 53.42 Mb autosomal genome (~2.4%) was identified with depletion of ROH, notably with extended runs on SSC8 and SSC13 (Table S3). To mitigate genotyping technology effect, we compared these regions with ROH using a whole-genome sequencing (WGS) variation dataset of one ZXWIPIG and six non-inbred Wuzhishan pigs, resulting in a refined region of 24.80 Mb after removing ROH segments on SSC8 and SSC13 (Table S4). The reason for including non-inbred individuals is that ZXWIPIG stems from normal non-inbred Wuzhishan pig so that they face the same risk in terms of a lethal mutation.

To finely map the putative lethal sites and to broaden the scope of our findings, we examined homozygous states of variations inside these regions using a WGS-based reference panel of 937 individuals [19], assuming that homozygotes of minor and altered alleles were lethal. We annotated variations lacking homozygotes of minor alleles and only considered functional variations of "stopgain", "stoploss" and frameshift substitution, revealing 34 functional variations. Cross-referencing with the IAnimal database containing 1311 pigs across 65 populations [41], we filtered out two mutations due to the presence of homozygote of altered alleles in other populations (SSC1:263222092G_C in Pietrain and wild boar; SSC13:160968909A_AT in Tibetan), and further confirmed 13 mutations, with 19 variations remaining unconfirmed (Table S5). Interestingly, a majority of mutations (21/32) and genes (15/22) were enriched to olfactory receptor, consistent with findings by haplotype approach in commercial populations [14]. Additionally, mutations were identified in genes related to metabolism and diseases such as *ASIC5* (recurrent pregnancy loss), *CMSS1* (autosomal recessive Robinow syndrome), and *FILIP1L* (ovarian cancer). These findings underscore the utility of extremely inbred populations in inferring recessive lethal mutations and provide insights into genetic mechanisms underlying disease susceptibility and other complex traits.

Discussion

The processes of recent inbreeding together with population bottlenecks and strong artificial selection in domestic pigs have substantially reduced their genetic diversity and recent Ne [66]. This genetic simplification has led to an increase in ROH, facilitating the persistence of deleterious variations within genomes. This provides us an opportunity to examine the effects of ROH on pig domestication and breeding, including ROH landscape during domestication, associations of ROH burden and economic traits/gene expression, and the inference of recessive lethal mutations.

The independent domestication events of Eurasian domestic pigs have been well-documented [42, 52, 53], which is supported by our findings of PCA and the differential ROH landscape of Eurasian pig domestication. A significant difference in ROH abundance and size was reported between European and Asian wild boars, in which large ROH in Asian pigs appeared to stem from recently reduced population size, whereas European wild boars exhibited a more uniform distribution of ROH, attributed to glacial bottlenecks and prolonged periods of a low Ne [66, 67]. Consequently, low genetic diversity of European wild boars caused a relatively small skewed distribution in domestication of European pigs compared to their Asian counterparts. Differences in the correlation between ROH frequency differences and genetic differentiation can be partially explained by multiple domestication events that have increased heterogeneity among Asian samples [52]. It is important to note that genetic drift and imputation accuracy of wild boars may influence our results. Additionally, the potential effects of ascertainment bias should be acknowledged in that the PorcineSNP60 BeadChip used in this study was predominantly designed based on European breeds [68].

Various genomic inbreeding coefficients have been developed to measure inbreeding level and depression, leaving the controversy on selection of inbreeding coefficients which is depended on the history of population [69]. F_{ROH}, in particular, is proposed to provide accurate estimates in populations with low effective size where reduced selection pressure may expose deleterious

mutations [70]. Our analyses of inbreeding depression are under (partial) dominance hypothesis in which deleterious mutations with recessive effects contribute to reduced fitness in individuals with high homozygosity [71]. We considered the effects of non-lethal and lethal recessive mutations on economic traits/gene expression and individual survival, respectively.

Although different statistical models were used to estimate inbreeding depression, detrimental effects of ROH burden on complex traits were widely documented in humans and agricultural animals [7, 11, 72–75]. Comparing classical linear model with linear mixed model incorporating genomic relationship matrix, we found that some estimates were influenced by population structure, in line with recent findings [69]. Of note, we observed a significantly unfavorable effect of ROH burden on SH in an F_3 Eurasian population (White Duroc boar × Erhualian sow) across both two models. Given the recessive inheritance of SH [76], this observation supports our hypothesis that the ubiquity of ROH could be linked to recessive genetic disease. This finding echoes similar observations in domestic dogs, where ROH burden correlates with diseases such as cranial cruciate ligament disease, elbow dysplasia and lymphoma [11]. Together with widespread inbreeding depression affecting growth traits in an F_2 Eurasian population (Large White boar × Min sow) [6], the finding of F_3 population highlighted the necessity of inbreeding monitoring in crossed herds. Additionally, we identified negative effects of ROH burden on multiple body conformation traits in Sujiang pig, a synthetic breed developed by crossing Chinese Jiangquhai, Fengjing and Western Duroc (62.5%) [26]. The observed depression in Sujiang likely stems from founder effects and subsequent selection within a closed population. The role of recessive deleterious alleles in causing inbreeding depression [11, 71, 77] is supported by our data, with negligible effects of GWAS variations within ROH in association tests and the higher deleteriousness, indicated as the average GERP scores, of ROH compared to non-ROH regions.

To elucidate which genes and pathways are influenced by inbreeding, differentially expressed genes (DEGs) have been explored by comparing individuals with varying inbreeding coefficients across multiple species. The number of identified genes varies widely among species: from zero in the brain of zebra finches, to tens in the hypothalamus of chickens, to hundreds in chickens and *Drosophila*, to thousands in Pacific oyster [78–81]. These results would be affected by gene quantification technology, tissue type, developmental stage, and statistical thresholds. In our study with pigs, results were comparable to other species, revealing tens DEGs associated with whole-genome ROH burden, with tissue-specific patterns, and thousands associated with regional ROH burden. These genes were enriched in pathways related to signal transduction, reproduction, and immunity, showing good agreement with previous findings on metabolism, stress and defense in chickens and *Drosophila* [79, 82]. Especially, *GSTO2*, encoding a glutathione transferase, emerged prominently across all three tissues studied, in broad concordance with findings in *Drosophila* [81]. Biologically, this association suggests that functionally recessive mutations in genes such as *IRX2* and *IRX4*, located in the associated window, influence *GSTO2* transcription. Compared to the identification of DEGs between inbred and non-inbred groups, our strategy of linking gene expression to ROH burden offers flexibility and could be easily extended to other species, irrespective of their specific inbreeding levels. However, it's important to note that regions lacking ROH polymorphisms within a population cannot be assessed for associations. Furthermore, the transcriptomic effects of ROH provide an approach to discover novel regulatory elements, a large fraction of which may not be captured by traditional eQTL studies. It was challenging to link our findings at the level of phenotypes and gene expression in this study, owning to the absence of an appropriate dataset. Colocalization analyses based on ROH burden in future research will help to understand the genetics of inbreeding depression.

The impact of inbreeding depression extends beyond individual fitness to include agricultural profitability through maternal prolificacy. While the genetic architecture of postnatal piglet mortality (often at birth and weaning) has been investigated by GWAS focusing on litter size alive [56], identifying loci associated with pregnancy loss and perinatal mortality remains challenging due to the absence of lethal genotypes in living animals available for sampling. In human and animal genetics, inference of recessive lethal mutations has traditionally required ultra-large sample size for assessing deficits in homozygosity and haplotypes [9, 10, 12–14]. Such studies have been constrained in commercial pig populations due to limited sample size and high research costs. Alternatively, we present a method to infer recessive lethal mutations using the depletion of ROH in inbred populations with relatively small sample size. This method capitalizes on the higher risk of survival fitness depression in inbred populations compared to non-inbred populations. After inspection with WGS data, we identified an enrichment of putative lethal mutations in genes related to olfactory receptor, in line with the discovery and validation from haplotype analyses in commercial populations [14]. The weak purifying selection observed in olfactory receptor genes facilitates the discovery of lethal mutations [83]. The highly developed olfactory system in wild pigs contributes significantly to their survival and reproductive success, contrasting with the reduced and irreversible

olfactory capabilities in domestic pigs [84]. We propose that the putative lethal mutations and olfactory system with attenuated function showcase the genetic cost of pig domestication [85]. However, we couldn't rule out the effects of copy number variation in olfactory receptors [86]. The other genes identified are involved in metabolism and disease. For example, *ASIC5* functioning in recurrent pregnancy loss and *FILIP1L* promoting synthesis of estradiol are closely related to perinatal mortality [87, 88]. Another example is *CMSS1* playing a role in autosomal recessive Robinow syndrome (GeneCards database, accessed July 2024), reflecting the recessive inheritance of ROH. Adopting graph-based reference genomes or more sophisticated alignment algorithms could help address the issue of mapping bias in traditional variation calling. We appreciate that the results inferred by deficit of ROH, homozygosity and haplotype will benefit from confirmation with larger sample size in future studies, and that the reproductive effects of putative lethal alleles are needed to be validated by scrutinizing the litter size of carrier-by-carrier matings.

Conclusions

This study demonstrated how the population history of pigs has elevated the homozygous regions carried in ROH, influencing economic traits, gene expression and fitness. Firstly, the significant associations between ROH burden and economic traits provide insights into the genetic architecture of complex traits, suggesting the roles of recessive variations in some phenotypes, especially diseases. However, more attention has been paid to additive effects in previous GWAS of pigs and other farm animals. Secondly, the associations of ROH burden and gene expression aid us to identify genes and pathways involved in inbreeding depression, and also provide an approach to search novel regulatory elements. Lastly, inference of recessive lethal mutations using inbred lines opens up an avenue for identification of lethal loci with relatively sample size, and also has implications in enhancing reproductive performance. These results contribute to the genetic understanding of inbreeding depression in farm animals.

Abbreviations

ABI	Average birth interval
ABW	Average birth weight
BF	Backfat thickness
BW	Body weight
CC	Chest circumference
CWD	Chest width
DEGs	Differentially expressed genes
eQTL	Expression quantitative trait loci
GERP	Genomic Evolutionary Rate Profiling
GWAS	Genome-wide association study
HW	Hip width
IMF	Intramuscular fat content
LMD	Loin muscle depth
MORT	Piglet mortality at birth
MS	Marbling scores
Ne	Effective population size
PCA	Principal component analysis
RNA-seq	RNA sequencing
ROH	Runs-of-homozygosity
SH	Scrotal hernia
SNP	Single nucleotide polymorphism
SSC	*Sus scrofa* chromosome
WGS	Whole-genome sequencing
ZXWIPIG	ZhongXu Wuzhishan minipig inbred line

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11189-y.

Additional file 1: Table S1 Sample information of SNP chip data. Table S2 Association between different phenotypes and the amount of the genome in ROH. Table S3 Regions devoid of ROH identified by SNP chip in ZXWIPIG. Table S4 Inferred regions with recessive lethal mutations narrowed by sequencing data. Table S5 Putative recessive lethal mutations.

Additional file 2: Fig. S1 Principal component analysis of wild boars (including feral pigs; left) and domestic pigs (right). Each point denotes one breed/population by using the averages of the first and second principal components (PC1 and PC2). The numbers in parentheses represent the percentage of variance explained by the first two principal components. Fig. S2 Effective population size (Ne) of minipigs (left) and commercial pigs (right) using GONE. Breed codes were listed in Table S1. Fig. S3 Quantification of inbreeding level by F_{ROH}. Breed codes were listed in Table S1. Fig. S4 Relationship between ROH frequency difference and genetic differentiation (F_{ST}). (A) Asian pigs. (B) European pigs. Correlation coefficients are shown for each chromosome. ns, $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$. Fig. S5 Associations of F_{ROH} with economic traits using classical linear model (LM) to linear mixed model (LMM). Trait codes and details of fitted models were available in Table S2. Fig. S6 Relationship inference of samples based on whole-genome sequencing (A) and RNA sequencing (B).

Acknowledgements

Not applicable.

Authors' contributions

Y.P.Z. and H.B.X. conceived the study. S.T.F. developed the ZXWIPIG. L.T. and H.L were responsible for bioinformatic analysis. L.T. wrote the manuscript. L.T., H.L., A.C.A., H.B.X., S.T.F. and Y.P.Z. revised the manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFF1000602 and 2022YFA1105401), the Chinese Academy of Sciences (XDA24010107), the National Natural Science Foundation of China (32388102), the Yunnan Provincial Science and Technology Department (202405AC350090), the Spring City Plan: the High-level Talent Promotion and Training Project of Kunming (2022SCP001), the Southwest Research Center for Pig Molecular Breeding and Translational Medicine, and the Animal Branch of the Germplasm Bank of Wild Species, Chinese Academy of Sciences (the Large Research Infrastructure Funding).

Data availability

No new sequencing data is generated in this study. All genotypes and phenotypes are publicly available. Detailed information about the datasets utilized can be found in Table S1 and original articles.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 20 September 2024 Accepted: 27 December 2024
Published online: 06 January 2025

References

1. Lunney JK, Van Goor A, Walker KE, Hailstock T, Franklin J, Dai C. Importance of the pig as a human biomedical model. Sci Transl Med. 2021;13(621):eabd5758.
2. Harlizius B, Mathur P, Knol EF. Breeding for resilience: new opportunities in a modern pig breeding program. J Anim Sci. 2020;98(Supplement_1):S150–S154.
3. Shatokhin KS. Problems of mini-pig breeding. Vavilovskii Zhurnal Genet Selektsii. 2021;25(3):284–91.
4. Peripolli E, Munari DP, Silva M, Lima ALF, Irgang R, Baldi F. Runs of homozygosity: current knowledge and applications in livestock. Anim Genet. 2017;48(3):255–71.
5. Sun S, Wang B, Li C, Xu G, Yang J, Hufford MB, Ross-Ibarra J, Wang H, Wang L. Unraveling prevalence and effects of deleterious mutations in maize elite lines across decades of modern breeding. Mol Biol Evol. 2023;40(8):msad170.
6. Tao L, Wang LG, Adeola AC, Zhang LC, Li LW, Li QL, Cen DJ, Yan C, Ma ZS, Wang LX, et al. Associations of autozygosity with economic important traits in a cross of Eurasian pigs. J Genet Genomics. 2023;50(3):216–20.
7. Zhang Y, Zhuo Y, Ning C, Zhou L, Liu JF. Estimate of inbreeding depression on growth and reproductive traits in a Large White pig population. G3: Genes, Genomes, Genet.2022;12(7):jkac118.
8. Zhao F, Zhang P, Wang X, Akdemir D, Garrick D, He J, Wang L. Genetic gain and inbreeding from simulation of different genomic mating schemes for pig improvement. J Anim Sci Biotechnol. 2023;14(1):87.
9. Oddsson A, Sulem P, Sveinbjornsson G, Arnadottir GA, Steinthorsdottir V, Halldorsson GH, Atlason BA, Oskarsson GR, Helgason H, Nielsen HS et al. Deficit of homozygosity among 1.52 million individuals and genetic causes of recessive lethality. Nat Commun. 2023;14(1):3453.
10. Kostyunina O, Yaryshkin A, Bykova O, Stepanov A, Shevkunov O, Shakirov V, et al. Identification of lethal recessive genetic variants in Holstein cattle. BIO Web Conf. 2024;108:01002.
11. Mooney JA, Yohannes A, Lohmueller KE. The impact of identity by descent on fitness and disease in dogs. Proc Natl Acad Sci U S A. 2021;118(16):e2019116118.
12. Derks MFL, Gjuvsland AB, Bosse M, Lopes MS, van Son M, Harlizius B, Tan BF, Hamland H, Grindflek E, Groenen MAM, et al. Loss of function mutations in essential genes cause embryonic lethality in pigs. PLoS Genet. 2019;15(3):e1008055.
13. Häggman J, Uimari P. Novel harmful recessive haplotypes for reproductive traits in pigs. J Anim Breed Genet. 2017;134(2):129–35.
14. Derks MFL, Megens HJ, Bosse M, Lopes MS, Harlizius B, Groenen MAM. A systematic survey to identify lethal recessive variation in highly managed pig populations. BMC Genomics. 2017;18(1):858.
15. Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MFW, Fisher EMC. Genealogies of mouse inbred strains. Nat Genet. 2000;24(1):23–5.
16. Wang L, Mu Y, Xu L, Li K, Han J, Wu T, Liu L, Gao Q, Xia Y, Hou G, et al. Genomic analysis reveals specific patterns of homozygosity and heterozygosity in inbred pigs. Animals. 2019;9(6):314.
17. Huo JL, Zhang LQ, Zhang X, Wu XW, Ye XH, Sun YH, Cheng WM, Yang K, Pan WR, Zeng YZ. Genome-wide single nucleotide polymorphism array and whole-genome sequencing reveal the inbreeding progression of Banna minipig inbred line. Anim Genet. 2022;53(1):146–51.
18. Hofmeister RJ, Ribeiro DM, Rubinacci S, Delaneau O. Accurate rare variant phasing of whole-genome and whole-exome sequencing data in the UK Biobank. Nat Genet. 2023;55(7):1243–9.
19. Tong X, Chen D, Hu J, Lin S, Ling Z, Ai H, Zhang Z, Huang L. Accurate haplotype construction and detection of selection signatures enabled by high quality pig genome sequences. Nat Commun. 2023;14(1):5126.
20. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867–73.
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
22. Santiago E, Novo I, Pardiñas AF, Saura M, Wang J, Caballero A. Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. Mol Biol Evol. 2020;37(12):3642–53.
23. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156–8.
24. Chen H, Wang C, Conomos Matthew P, Stilp Adrienne M, Li Z, Sofer T, Szpiro Adam A, Chen W, Brehm John M, Celedón Juan C, et al. Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. Am J Hum Genet. 2016;98(4):653–66.
25. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44(7):821–4.
26. Xu P, Ni L, Tao Y, Ma Z, Hu T, Zhao X, Yu Z, Lu C, Zhao X, Ren J. Genome-wide association study for growth and fatness traits in Chinese Sujiang pigs. Anim Genet. 2020;51(2):314–8.
27. Crespo-Piazuelo D, Acloque H, González-Rodríguez O, Mongellaz M, Mercat MJ, Bink M, Huisman AE, Ramayo-Caldas Y, Sánchez JP, Ballester M. Identification of transcriptional regulatory variants in pig duodenum, liver, and muscle tissues. Gigascience. 2022;12:giad042.
28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
29. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
30. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.
31. Blighe K, Lun A: PCAtools: everything principal components analysis. In.; 2024.
32. Daniel C-P, Hervé A, González-Rodríguez O, Mongellaz M, Marie-José M, Marco B, et al. Supporting data for "Identification of transcriptional regulatory variants in pig duodenum, liver and muscle tissues." GigaScience Database. 2023. https://doi.org/10.5524/102388.
33. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303.
34. Yu G. Using ggtree to visualize data on tree-like structures. Curr Protoc Bioinf. 2020;69(1):e96.
35. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7.
36. Fu W, Gittelman RM, Bamshad MJ, Akey JM. Characteristics of neutral and deleterious protein-coding variation among individuals and populations. Am J Hum Genet. 2014;95(4):421–36.
37. Grossen C, Guillaume F, Keller LF, Croll D. Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. Nat Commun. 2020;11(1):1001.
38. Yan C, Xie HB, Adeola AC, Fu Y, Liu X, Zhao S, Han J, Peng MS, Zhang YP. Inference of ancestral alleles in the pig reference genome. Anim Genet. 2023;54(5):649–51.
39. Biscarini F, Cozzi P, Gaspa G, Marras G. detectRUNS: detect runs of homozygosity and runs of heterozygosity in diploid genomes. In.; 2019.
40. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164–e164.
41. Fu Y, Liu H, Dou J, Wang Y, Liao Y, Huang X, Tang Z, Xu J, Yin D, Zhu S, et al. IAnimal: a cross-species omics knowledgebase for animals. Nucleic Acids Res. 2022;51(D1):D1312–24.
42. Groenen MAM. A decade of pig genome sequencing: a window on pig domestication and evolution. Genet Sel Evol. 2016;48(1):23.

43. Xue Y, Li C, Duan D, Wang M, Han X, Wang K, Qiao R, Li XJ, Li XL. Genome-wide association studies for growth-related traits in a crossbreed pig population. Anim Genet. 2021;52(2):217–22.
44. Li C, Duan D, Xue Y, Han X, Wang K, Qiao R, Li XL, Li XJ. An association study on imputed whole-genome resequencing from high-throughput sequencing data for body traits in crossbred pigs. Anim Genet. 2022;53(2):212–9.
45. Li M, Li X, Liu C, Han J, Han X, Wang K, Qiao R, Yang F, Li XL, Li XJ. Genome-wide association study reveals the candidate genes for reproduction traits in Yunong black pigs. Anim Genet. 2023;54(3):403–7.
46. Deng M, Qiu Z, Liu C, Zhong L, Fan X, Han Y, Wang R, Li P, Huang R, Zhao Q. Genome-wide association analysis revealed new QTL and candidate genes affecting the teat number in Dutch Large White pigs. Anim Genet. 2024;55(2):206–16.
47. Liu K, Hou L, Yin Y, Wang B, Liu C, Zhou W, Niu P, Li Q, Huang R, Li P. Genome-wide association study reveals new QTL and functional candidate genes for the number of ribs and carcass length in pigs. Anim Genet. 2023;54(4):435–45.
48. Ibragimov E, Eriksen EØ, Nielsen JP, Jørgensen CB, Fredholm M, Karlskov-Mortensen P. Towards identification of new genetic determinants for post-weaning diarrhea in piglets. Anim Genet. 2024;55(3):387–95.
49. Li K, Hou L, Du T, Pu G, Liu C, Liu G, Zhou J, Huang R, Zhao Q, Li P. Genome-wide association study identified five quantitative trait loci and two candidate genes for digestive traits in Suhuai pigs. Anim Genet. 2024;55(3):471–4.
50. Ding R, Yang M, Quan J, Li S, Zhuang Z, Zhou S, Zheng E, Hong L, Li Z, Cai G. Single-locus and multi-locus genome-wide association studies for intramuscular fat in Duroc pigs. Front Genet. 2019;10:619.
51. Iacolina L, Scandura M, Goedbloed DJ, Alexandri P, Crooijmans RPMA, Larson G, Archibald A, Apollonio M, Schook LB, Groenen MAM, et al. Genomic diversity and differentiation of a managed island wild boar population. Heredity. 2016;116(1):60–7.
52. Wu GS, Yao YG, Qu KX, Ding ZL, Li H, Palanichamy MG, Duan ZY, Li N, Chen YS, Zhang YP. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. Genome Biol. 2007;8(11):R245.
53. Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, et al. Analyses of pig genomes provide insight into porcine demography and evolution. Nature. 2012;491(7424):393–8.
54. Hu Z-L, Park CA, Reecy JM. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. Nucleic Acids Res. 2021;50(D1):D956–61.
55. Xie L, Qin J, Rao L, Tang X, Cui D, Chen L, Xu W, Xiao S, Zhang Z, Huang L. Accurate prediction and genome-wide association analysis of digital intramuscular fat content in longissimus muscle of pigs. Anim Genet. 2021;52(5):633–44.
56. An M, Zhou G, Li Y, Xiang T, Ma Y, Liu X, Li X, Zhao S, Zhu M. Characterization of genetic fundamentals for piglet mortality at birth in Yorkshire, Landrace, and Duroc sows. Anim Genet. 2022;53(1):142–5.
57. Xu W, Chen D, Yan G, Xiao S, Huang T, Zhang Z, Huang L. Rediscover and refine QTLs for pig scrotal hernia by increasing a specially designed F_3 population and using whole-genome sequence imputation technology. Front Genet. 2019;10:890.
58. Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 2005;15(7):901–13.
59. Piacentini S, Monaci PM, Polimanti R, Manfellotto D, Fuciarelli M. GSTO2*N142D gene polymorphism associated with hypothyroidism in Italian patients. Mol Biol Rep. 2013;40(2):1967–71.
60. Sobot V, Stamenkovic M, Simic T, Jerotic D, Djokic M, Jaksic V, Bozic M, Milic J, Savic-Radojevic A, Djukic T. Association of *GSTO1*, *GSTO2*, *GSTP1*, *GPX1* and *SOD2* polymorphism with primary open angle glaucoma. Exp Eye Res. 2022;214:108863.
61. Piacentini S, Polimanti R, Iorio A, Cortesi M, Papa F, Rongioletti M, Liumbruno GM, Manfellotto D, Fuciarelli M. GSTA1*-69C/T and GSTO2* N142D as asthma-and allergy-related risk factors in Italian adult patients. Clin Exp Pharmacol Physiol. 2014;41(3):180–4.
62. Christoffels VM, Keijser AGM, Houweling AC, Clout DEW, Moorman AFM. Patterning the embryonic heart: identification of five mouse *Iroquois* homeobox genes in the developing heart. Dev Biol. 2000;224(2):263–74.
63. Freese NH, Lam BA, Staton M, Scott A, Chapman SC. A novel gain-of-function mutation of the proneural *IRX1* and *IRX2* genes disrupts axis elongation in the Araucana rumpless chicken. PLoS ONE. 2014;9(11):e112364.
64. Szpiech ZA, Xu J, Pemberton TJ, Peng W, Zöllner S, Rosenberg NA, Li JZ. Long runs of homozygosity are enriched for deleterious variation. Am J Hum Genet. 2013;93(1):90–102.
65. Zhang Q, Guldbrandtsen B, Bosse M, Lund MS, Sahana G. Runs of homozygosity and distribution of functional variants in the cattle genome. BMC Genomics. 2015;16(1):542.
66. Bosse M, Megens H-J, Madsen O, Paudel Y, Frantz LA, Schook LB, Crooijmans RP, Groenen MA. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. PLoS Genet. 2012;8(11):e1003100.
67. Scandura M, Iacolina L, Crestanello B, Pecchioli E, Di Benedetto M, Russo V, Davoli R, Apollonio M, Bertorelle G. Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? Mol Ecol. 2008;17(7):1745–62.
68. Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, et al. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS ONE. 2009;4(8):e6524.
69. Lavanchy E, Weir BS, Goudet J. Detecting inbreeding depression in structured populations. Proc Natl Acad Sci U S A. 2024;121(19):e2315780121.
70. Caballero A, Villanueva B, Druet T. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. Evol Appl. 2021;14(2):416–28.
71. Charlesworth D, Willis JH. The genetics of inbreeding depression. Nat Rev Genet. 2009;10(11):783–96.
72. Ceballos FC, Hazelhurst S, Clark DW, Agongo G, Asiki G, Boua PR, Xavier Gómez-Olivé F, Mashinya F, Norris S, Wilson JF, et al. Autozygosity influences cardiometabolic disease-associated traits in the AWI-Gen sub-Saharan African study. Nat Commun. 2020;11(1):5754.
73. Johnson EC, Evans LM, Keller MC. Relationships between estimated autozygosity and complex traits in the UK Biobank. PLoS Genet. 2018;14(7):e1007556.
74. Tao L, He X, Wang X, Di R, Chu M. Litter size of sheep (*Ovis aries*): Inbreeding depression and homozygous regions. Genes. 2021;12(1):109.
75. Zhao G, Liu Y, Niu Q, Zheng X, Zhang T, Wang Z, Xu L, Zhu B, Gao X, Zhang L, et al. Runs of homozygosity analysis reveals consensus homozygous regions affecting production traits in Chinese Simmental beef cattle. BMC Genomics. 2021;22(1):678.
76. Carman GM. Hernia and its heredity. Iowa State Univ Vet. 1952;14(3):143–6.
77. Stoffel M, Johnston S, Pilkington J, Pemberton JM. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. Nat Commun. 2021;12(1):2972.
78. Hansson B, Naurin S, Hasselquist D. Does inbreeding affect gene expression in birds? Biol Lett. 2014;10(9):20140648.
79. Xue Q, Li G, Cao Y, Yin J, Zhu Y, Zhang H, Zhou C, Shen H, Dou X, Su Y, et al. Identification of genes involved in inbreeding depression of reproduction in Langshan chickens. Anim Biosci. 2021;34(6):975–84.
80. Fang J, Xu C, Li Q. Transcriptome analysis of inbreeding depression in the Pacific oyster *Crassostrea gigas*. Aquaculture. 2022;557:738314.
81. Ayroles JF, Hughes KA, Rowe KC, Reedy MM, Rodriguez-Zas SL, Drnevich JM, Cáceres CE, Paige KN. A genomewide assessment of inbreeding depression: gene number, function, and mode of action. Conserv Biol. 2009;23(4):920–30.
82. Paige KN. The functional genomics of inbreeding depression: a new approach to an old problem. Bioscience. 2010;60(4):267–77.
83. Kang M, Ahn B, Youk S, Jeon H, Soundarajan N, Cho E-S, Park W, Park C. Individual and population diversity of 20 representative olfactory receptor genes in pigs. Sci Rep. 2023;13(1):18668.
84. Maselli V, Polese G, Larson G, Raia P, Forte N, Rippa D, Ligrone R, Vicidomini R, Fulgione D. A dysfunctional sense of smell: the irreversibility of olfactory evolution in free-living pigs. Evol Biol. 2014;41(2):229–39.
85. Lu J, Tang T, Tang H, Huang J, Shi S, Wu C-I. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. Trends Genet. 2006;22(3):126–31.
86. Paudel Y, Madsen O, Megens H-J, Frantz LAF, Bosse M, Crooijmans RPMA, Groenen MAM. Copy number variation in the speciation of

pigs: a possible prominent role for olfactory receptors. BMC Genomics. 2015;16(1):330.
87. Liu Y, Yang Y, Li W, Ao H, Zhang Y, Zhou R, Li K. Effects of melatonin on the synthesis of estradiol and gene expression in pig granulosa cells. J Pineal Res. 2019;66(2):e12546.
88. Al Qahtani NH, AbdulAzeez S, Almandil NB, Fahad Alhur N, Alsuwat HS, Al Taifi HA, Al-Ghamdi AA, Rabindran Jermy B, Abouelhoda M, Subhani S, et al. Whole-genome sequencing reveals exonic variation of *ASIC5* gene results in recurrent pregnancy loss. Front Med. 2021;8:699672.

Publisher's Note