This is a “preproof” accepted article for *Animal Nutriomics*.

This version may be subject to change during the production process.

10.1017/anr.2024.20

Identified Candidate Genes for Pig Mental Health: Insights from in Intensive Farming Systems

Short title: Pig Mental Health Genes

Xingyu Wei[1], Lingyao Xu[1], Jinyun Jiang[1], Jian Miao[1], Fen Wu[1], Zitao Chen[1], Zhe Zhang[1], Qishan Wang[1], Yuchun Pan[1], Zhen Wang[1, §]

[1] College of Animal Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China.

[§] Corresponding authors (Zhen Wang: wanghzen20@zju.edu.cn)

ABSTRACT

Understanding the genetic basis of porcine mental health (PMH)-related traits in intensive pig farming systems may promote genetic improvement animal welfare enhancement. However, investigations on this topic have been limited to a retrospective focus, and phenotypes have been difficult to elucidate due to an unknown genetic basis. Intensively farmed pigs, such as those of the Duroc, Landrace and Yorkshire breeds, have undergone prolonged selection pressure in intensive farming systems. This has potentially subjected genes related to mental health in these pigs to positive selection. To identify genes undergoing positive selection under intensive farming conditions, we employed multiple selection signature detection approaches. Specifically, we integrated disease gene annotations from three human gene–disease association databases (Disease, DisGeNET, and MalaCards) to pinpoint genes potentially associated with pig mental health, revealing a total of 254 candidate genes related to PMH. In-depth functional analyses revealed that candidate PMH genes were significantly overrepresented in signaling-related pathways (e.g., the dopaminergic synapse, neuroactive ligand–receptor interaction and calcium signaling pathways) or Gene Ontology (GO) terms (e.g., dendritic tree and synapse). These candidate PMH genes were expressed at high levels in the porcine brain regions such as the hippocampus, amygdala and hypothalamus, and the cell type in which they were significantly enriched was neurons in the hippocampus. Moreover, they potentially affect pork meat quality traits. Our findings make a significant contribution to elucidating the genetic basis of PMH, facilitating genetic improvements for the welfare of pigs and establishing pigs as valuable animal models for gaining insights into human psychiatric disorders.

Keywords: Intensive system; selection signatures; pig mental health genes; psychiatric disorders.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

INTRODUCTION

Pigs are an important source of human dietary protein worldwide. Intensive pig farming systems have gained popularity due to their ability to achieve higher productivity and streamlined management. However, intensive farming systems may cause stress in animals [1], behavioral signs of stress may include increased defecation, aggression, sow infanticide, lethargy, decreased appetite, and postpartum depression and can have adverse effects on productivity of pigs[2]. For example, when subjected to overcrowding and mixing, pigs experience decreased growth rates of 15.7% and 7.1%, respectively, while the feed-to-meat conversion ratio is estimated to decline by 15% and 10%, respectively [3]. Additionally, the reproductive performance of sows in close proximity to aggressive sows is compromised as a result of stress[4], which has a significant effect on the economic efficiency of pig farming. Aggression is recognized as a paramount concern in modern pig production systems because it poses significant challenges to the health, welfare and economic aspects of the pigs [5].

Selection signatures may be effective approaches for elucidating genetic mechanisms of porcine mental health (PMH), even in the absence of detailed phenotypic data [6]. Intensively farmed pig breeds, such as Duroc, Landrace, Yorkshire, Hampshire, and Pietrain pigs, have gone through long-term intensive production management [7], in which breeding pigs prone to mental health problems will be eliminated because of their reduced production performance. Under intense systemic selection pressure, pigs experiencing mental health issues are typically culled due to decreased production performance, meanwhile, genes related to mental health will also experience selection. In contrast, Chinese indigenous pig breeds such as Jinhua pigs and Meishan pigs experience less systemic selection pressure due to their lower level of intensification in farming. Therefore, by comparing the genomes of intensively and nonintensively farmed pig breeds using selection signature detection methods, candidate genes associated with PMH can be detected. Specifically, it was possible to construct intrapopulation selection signatures within intensively farmed pig breeds and interpopulation selection signatures between intensive and nonintensively farmed pig breeds. By employing various selection signature methods, it was feasible to detect candidate PMH genes.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Pigs are an excellent model animal species for gaining insights into the human brain and neurodevelopment [8]. The pig genome exhibits high sequence similarity and chromosomal synteny with the human genome, which is useful for genomic applications [9]. Additionally, Chen et al. [10] reported the identification of chromosomal regions associated with infanticidal behavior in pigs that contain genes comparable to those of humans and rodents that are involved in anxiety, bipolar affective disorder, or coping behaviors. Pigs exhibit anatomical similarities to humans, with the porcine brain having similarities to the human brain, exhibiting a gyrencephalic cranial structure [8]. In particular, the hippocampus and sensory cortex arrangement in pigs are more similar to those of humans than are those in mice [11]. In terms of practical considerations, pigs exhibit several advantages over primates and other livestock models. These features include a shorter generation time, bigger litter size, and suitability for straightforward genetic editing [12]. Since 2007, numerous robust and replicable genetic findings have been reported for psychiatric disorders. These findings have mostly been reported via genome-wide association (GWAS) and structural variation (SV) studies [13] and were systematically collected in disease gene databases (such as MalaCards [14], DISEASE [15], and DisGeNET [16]). Previous studies have suggested that human brain regions such as the amygdala, striatum and hippocampus are especially relevant to mental disorders [17]. For example, the hippocampus, which is part of the limbic system, is involved in emotion, memory formation, stress responses, emotional regulation, and memory processes, and damage to this region can lead to memory and learning impairments, as well as behavioral disorders such as anxiety, depression, and schizophrenia [18]. Since there are fewer studies on pig mental health trait, we consulted databases related to human mental disorders. By annotating genes from databases associated with human mental illnesses, we constructed a dataset for the PMH gene, thereby narrowing down the scope of candidate PMH genes.

In this study, we aimed to first compile a list of high-confidence psychiatric disease risk genes identified from human psychiatric disease databases and then further screen for selection signatures that reflect positive selection under an intensive pig farming system.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Finally, we systematically revealed the characteristics of candidate PMH genes, including their expression levels and changes in different types of tissues and cells, and their potential effects on pig production traits. Our findings provide novel insight into elucidating the genetic basis of PMH traits.

MATERIALS AND METHODS

Datasets and their sources

Human disease–gene data sets. From three databases, DISEASE [15], DisGeNET [16] and MalaCards & GeneCards [14], we collected a total of 1642 gene–disease records with gene–disease association scores (**Table S1**).

Genotyping dataset. We collected the genotypes of 348 intensively farmed pigs (including pigs of the Duroc, Landrace, Large White, Bershire, Hampshire, and Pietrain breeds) and 446 nonintensively farmed pigs (including 39 Chinese indigenous pig breeds, **Table S2**) from the PHARP v2 database [19], a pig genotyping source containing a total of 2,048 individuals of intensively farmed breeds, Chinese indigenous pig breeds, and wild boars. The genotypes contained approximately 23.1 million SNPs after removing those with minor allele frequencies of less than 0.05. All the statistical analyses were performed using the above genotype data unless otherwise specified.

Refinement of human psychiatric candidate genes.

Given the challenges in the past in ascertaining the phenotypes of mental health traits in pigs, studies on the genetic basis of these traits have been limited. Nevertheless, genes associated with human psychiatric conditions have been well identified and compiled in the aforementioned human disease–gene databases. Through homologous gene analysis between pigs and humans, we were able to identify potential mental health genes in pigs by refining human psychiatric genes. Here, we used three human disease–gene databases to assemble a list of human psychiatric genes. Briefly, we first classified the disease names in each database into the DSM-5 and ICD-10-CM mental disorders classification code [20] (criteria for the diagnosis of various mental disorders) utilizing UMLS

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Metathesaurus [21] and then manually reviewed uncategorized disease names in five databases—DSM-5-TR, OMIM [22], MalaCards [14], Disease Ontology [23], and UMLS Metathesaurus Browser [24]—to determine whether they were associated with mental disorders and were classified into the DSM5-TR code (**Table S3-S5**). According to the DSM5-TR and ICD10-CM codes, there are a total of 22 distinct categories for classified diseases. The above steps were used to classify the different diseases and determine whether a disease was related to a human mental disorder. We next filtered disease–gene items with disease–gene association scores below the top 2.5% to retain the high confidence disease–gene items for each database. Genes supported by at least two of these three disease–gene databases were considered relevant to human mental disorders. The Biomart tool [25] was used to analyze homologous genes between humans and pigs, thereby establishing a dataset of genes associated with PMH.

Population structure analysis.

We applied three methods to analyze the population genetic structure. First, we performed principal component analysis (PCA) using PLINK v1.90 [26] software to cluster individuals. Next, we used the ADMIXTURE program [27] to estimate the proportion of common ancestors among individuals. Finally, we built the NJ-tree with MEGA11 [28] using the genetic distance among individuals calculated by PLINK v1.90 and refined it with iTOL v6 [29].

Selection signature detection.

Here, we employed five approaches, including three interpopulation methods (F_{ST} [30], XP-CLR [31] and XP-EHH [32]) and two intrapopulation methods (CLR [33] and ROH [34]) to identify putative genomic regions undergoing positive selection. Considering that genes associated with pig mental health have undergone soft positive selection, we integrated selection signatures identified by five complementary approaches to identify potential PMH candidate genes.

F_{ST} test. F_{ST} is a method for detecting selection signatures based on population differentiation [30]. We screened the mean F_{ST} value for each genomic region with a

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

window size of 50 kb and a step size of 10 kb between intensively and nonintensively farmed pig populations by VCFtools v0.1.16 software [35] and considered the genomic regions with the top 1% mean F_{ST} values as putative positive selection signatures.

XP-CLR test. The cross-population compound likelihood ratio test (XP-CLR) is a method used to estimate the composite likelihood ratio (CLR) of a genomic region to detect genome-wide selection signatures [31]. We estimated the XP-CLR across each chromosome by using xpclr v1.1.2 [31] with the parameters ld 0.95 --maxsnps 200 --size 50000 --step 10000 and defined the genomic regions with XP-CLRs in the top 1% as putative positive selection signatures.

XP-EHH test. Cross-population extended haplotype homozygosity (XP-EHH) is a haplotype-based method for detecting selection signatures [32]. We used the intensively farmed pig population as the observed population and the nonintensively farmed population as the reference population using Selscanv1.3.0 [36] to calculate XP-EHH values and considered genomic regions with positive XP-EHH values in the top 2.5% as putative positive selection signatures.

CLR test. The CLR test is a method for detecting selection signatures based on the distribution of allele frequencies in a population. We calculated the CLR value for every 50 kb genomic region across the genome using SweeD software [37] and then ranked each genomic segment based on the CLR value. The genomic regions with 5% right-tail empirical *p* values were regarded as potential positive selection regions.

Runs of homozygosity detection. Runs of homozygosity (ROH) is an approach for detecting selection signatures by identifying genomic regions with reduced variation relative to the genome average. We estimated the ROH for each individual using genotypes after linkage disequilibrium (LD) pruning (with the parameter --indep-pairwise 50 10 0.2) by "*-homozyg*" in PLINK v1.90 software [26] with the parameters –homozyg-snp 100, –homozyg-window-snp 50, and –homozyg-window-missing 5. The criteria applied for ROH identification were as follows [38]: (1) 50 SNPs were contained in each sliding

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

window; (2) an ROH consisted of no less than 100 consecutive SNPs; and (3) the density was higher than one SNP per 50 kb [38]. We calculated the frequency of occurrence for each SNP within ROH segments across all individuals and defined SNPs with frequencies in the top 1% as significant regions.

Putative candidate genes under selection

We considered genes within the 50 kb region upstream and downstream of identified selection signatures as potentially selected candidate genes and utilized the GALLO package [39] with the pig reference genome Sus11.1 [40] to annotate selection signatures in the pursuit of potential selected candidate genes.

Characterization of candidate PMH genes

*Functional enrichment analysis***.** We used the clusterProfiler package [41] to identify Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO) terms enriched with candidate PMH genes. *P values* were adjusted for multiple tests using the BH method, and the adjusted *P value* was set to 0.05 as the significance threshold. Furthermore, we used the pig quantitative trait loci (QTL) database (https://www.animal-genome.org/cgi-bin/QTLdb/SS/index). [42] to annotate economic traits associated with the candidate genes.

Tissue gene expression analysis. To investigate gene-tissue expression specificity, we examined the expression profiles of candidate PMH genes across different tissues from Pig-GTEx [43] (**Table S6**). Using the Euclidean distance and Ward's clustering methods, we grouped the candidate PMH genes into different clusters according to their different expression profiles across tissues. A heatmap of gene expression was generated using the R 'gplots' package with the 'heatmap.2' function.

Weighted expression cell type enrichment analysis. Expression-weighted cell-type enrichment (EWCE) [44] was used for PMH candidate gene enrichment analysis with default parameters. This analysis was performed separately for each brain region dataset

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

[45] and for the hippocampus of pig single-cell data derived from our unpublished data. Then, the specificity matrix for each dataset was calculated by using the EWCE package. All the detected genes in each brain region dataset were used as background, and bootstrap resampling was performed 10000 times. The adjusted *P* value was corrected using the BH method, and 0.05 was used as the significance threshold.

Differentially expressed genes (DEGs). We used the 'scanpy.tl.rank_genes_groups' function in Scanpy (v1.9.1) [46] with the Wilcoxon rank-sum test to identify the differentially expressed genes among candidate PMH genes in the inhibitory neurons of the hippocampus in both Jinhua and Duroc pigs. The parameters were adjusted *P* value < 0.01 and |logfold change| >1.

PPI network. We constructed a network using the GeneMANIA database [47] online tool (https://genemania.org/) to explore the connections between genes.

GWAS signal enrichment analysis. To explore the potential effects of candidate PMH genes on economic traits, we performed GWAS signal enrichment analysis using our unpublished GWAS summary data (the effects considered were sex, parity, year-season, and three principal components) of six traits, including corrected backfat thickness of 115 kg (BFT115), corrected day of age to reach 115 kg (DAY115), feed conversion rate (FCR), feed intake (FI), loin muscle depth (LMD), off set intramuscular fat (OffsetIMF), and represent feed intake (RFI). We implemented a summation-based marker set significance test method [48] using the QGG software package [49] for the above six traits. We considered false discovery rate (FDR)-corrected *P* values < 0.05 as the threshold for identifying significantly correlated traits.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

RESULTS

Pig mental health candidate genes.

From the DisGeNET database, we assembled a list of 1,269 human mental disorder candidate genes related to 122 diseases and 2,786 gene–disease items (**Fig. 1A** and **Fig. S1A**). Using the disease database, we assembled a list of 2,467 human mental disorder candidate genes related to 108 mental diseases and 6,378 gene–disease items (**Fig. 1B** and **Fig. S1B**). Similarly, from the MalaCards database, we assembled a list of 2,661 human mental disorder candidate genes related to 377 diseases and 42,384 gene–disease items (**Fig. 1C** and **Fig. S1C**). Finally, we narrowed down a total of 1,642 genes identified as candidates for human mental disorders that appeared in at least two out of three databases (**Fig. 1D** and **Table S7**). Subsequently, we conducted a human-to-pig homologous gene search using Biomart for these candidate genes, resulting in 1549 pig genes (**Table S8**), which included 1,439 genes located on autosomes (**Fig. S2** and **Table S9**).

Population genetic structure analysis of the NJ-tree, PCA, and ADMIXTURE showed that individuals were well clustered into two groups (intensive and nonintensively farmed pig breeds, **Fig. 2A-C**). The PCA plot showed that PC1 could clearly separate individuals into intensively farmed pig breeds and nonintensively farmed pig breeds. PC1 and PC2 accounted for 51.24% and 12.68% of the total variance, respectively (**Fig. 2B**). Similarly, the intensively farmed pigs were clearly distinguished from the nonintensively farmed pigs when $K = 2$ in the admixture analysis **(Fig. 2C**). Next, we conducted selection signatures analysis on these intensive and nonintensively farmed pig breed. According to the selection criteria for each method (see details in Materials and Methods), the F_{ST}, XP-EHH and XP-CLR, CLR and ROH methods revealed selection signatures associated with 537, 922, 1,013, 912, and 622 genes, respectively (**Fig. 3 A-E** and **Table S10-S14**). A total of 2,844 unique positively selected genes (PSGs) were detected by these five selection signature detection methods (**Fig. 3F**).

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

We then compared PSGs with human mental disorders genes, resulting in 254 genes identified as candidate PMH genes (**Fig. 4A**, **Fig. S3-S7 and Table S15**). Functional enrichment analysis of the 254 candidate PMH genes revealed that they were overrepresented in 83 KEGG pathways and in 275 GO terms (**Fig. 4 B** and **C**). Many of these pathways or GO terms are highly relevant to signal transduction. The cellular components in which the candidate PMH genes were enriched were those such as dendrites, synapses, neuronal projections, and postsynapses, indicating their involvement in signal transmission (**Table S16**). The GO terms in which these genes were also enriched were related to the regulation of signal transmission (**Table S17**). KEGG pathway enrichment analysis revealed that the dopaminergic synapse, cAMP signaling pathway, neuroactive ligand–receptor interaction, calcium signaling pathway, glutamatergic synapse, GABAergic synapse, and serotonergic synapse genes were enriched (**Table S18**). The QTL enrichment analysis indicated that the loci encompassed by these genes were linked to meat quality traits such as meat color b*, intermuscular fat content, and drip loss (**Fig. 4D**). GWAS signal enrichment analysis revealed that porcine mental health candidate genes were significantly ($p < 0.05$) associated with corrected 115 kg backfat thickness, corrected age in days to reach 115 kg body weight, and offset intramuscular fat, suggesting that PMH candidate genes may influence routine production performance (**Fig. 4E**).

Spatial and temporal specificity of PMH candidate genes expression

Based on the expression profiles of PMH candidate genes in various brain tissues, we observed high expression of these genes in the frontal cortex, hippocampus, and hypothalamus tissues (**Fig. 4F and Fig. S8**). Genes, such as *DRD1*, *CACNA1I*, *CACNA1E*, *GABRB2*, and *SCN2A*, related to calcium channels, potassium channels, and GABA transporters exhibited high expression in the hippocampus. The *DRD1* gene is a dopamine receptor encoding a protein that initiates G-protein-coupled receptor activity and dopamine neurotransmitter receptor activity [50]. The *CACNA1I* and *CACNA1E* genes encode calcium channel proteins that mediate the entry of calcium ions into excitable cells and participate in various calcium-dependent processes [51,52]. The *GABRB2* gene

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

encodes the β-2 subunit, influencing the major inhibitory system in the brain—the GABA system [53]. The *SCN2A* gene encodes a sodium channel protein, and variations can lead to various neurological disorders, including benign epilepsy, epileptic encephalopathy, and autism spectrum disorders [54].

To investigate the candidate PMH genes spatiotemporal specificity expression pattern in pigs of different ages, we subsequently compared the transcriptomic expression data between days 38 and 56 in the prefrontal cortex region of the porcine brain. During this period, pigs have just experienced weaning and regrouping, which may induce some stress responses. (**Fig. 4G**). We detected an increase in the expression of *GABRA4* at the 56-day time point. These two genes encode the β subunits of GABA-A receptors, which are inhibitory neurotransmitter receptors. The expression of the *SLC17A7* gene was also significantly higher on day 56 than on day 38. The *SLC17A7* gene is primarily expressed in neurons and plays a crucial role in the packaging and transport of the neurotransmitter glutamate into synaptic vesicles. The *SLC1A1* gene encodes a protein that removes excess glutamate from the synaptic gap to maintain neurotransmitter homeostasis. The *DRD1* gene functions as a type of dopamine receptor. The *SCN2A* and *SCN1A* genes also encode sodium channel proteins, while the *KCNQ2* and *KCNC2* genes are related to potassium channels. The *DTNBP1* gene encodes a membrane-associated protein. Notably, the expression of these genes was higher at 56 days than at 38 days. Calcium ion-related genes, *CACNA1I, CACNA1A*, and *CACNA1E*, did not change significantly between these two periods. Compared with the expression of the HSPA5 gene on day 38, that on day 56 significantly decreased.

High expression of candidate PMH genes in inhibitory neurons in the hippocampus

Intriguingly, we observed that the expression of PMH genes was significantly enriched in inhibitory neurons in the hippocampus of JH pigs (a nonintensively farmed pig breed) (**Fig. 5A**), whereas the enrichment was not significant in Duroc pigs (an intensively farmed pig breed) (**Fig. 5B**). Therefore, we compared the expression of candidate PMH genes in inhibitory neurons in the hippocampus between the two breeds (**Fig. 5C** and **Table S19**). In Duroc pigs, there were 11 genes whose expression was greater than that in

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

JH pigs and eight genes whose expression was lower than that in JH pigs. Moreover, protein–protein interaction network (PPI) analysis of these eight downregulated genes revealed that the *HSPA5* and *HSPA8* genes were hub genes (**Fig. 5D**). The differences in hippocampal gene expression patterns between the two breeds led us to focus on specific cell types for studying PMH traits in the future. Moreover, we conducted enrichment analyses of candidate PMH genes using human brain data, and the candidate PMH genes were also enriched in neuron cell types in the human brain (**Fig. 5E** and **5F**). This finding suggested a significant association between PMH genes and neurons in the hippocampus of pigs.

DISCUSSION

With the intensification of pig farming in China, large-scale and intensive production management methods have imposed significant stress on pigs, leading to decreased production performance and substantial economic losses in the swine industry. However, there has been limited research focusing on porcine mental health, as these traits are challenging to measure and identify. Currently, there is a scarcity of studies and a lack of phenotype data in this area. Given the high homology between humans and pigs, we aim to leverage information on human diseases to identify candidate genes related to porcine mental health through comparative genomics. Our findings are beneficial for conducting molecular design breeding and genetic improvement to adapt pigs to intensive environments, reduce stress, and enhance economic trait performance.

In our study, we identified 254 candidate PMH genes that underwent positive selection during the transition from nonintensive to intensive farming. The KEGG pathways and GO terms in which these genes were enriched were related to neurotransmission, such as the dopaminergic synapse, neuroactive ligand–receptor interaction, calcium signaling pathway, glutamatergic synapse, and GABAergic synapse. Dysregulation in the dopamine system is one of the factors contributing to the development of schizophrenia, and current antipsychotic drugs mainly act by antagonizing dopamine receptors [53]. A previous study also showed that imbalances in glutamate and GABA (the primary

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

inhibitory neurotransmitter) can directly disrupt endocrine activity, leading to behavioral abnormalities [54].

Furthermore, PMH candid ate genes are significantly expressed in brain regions involved in mood disorders, such as the prefrontal cortex and hippocampus. Dysfunction in the prefrontal cortex and its connectivity with other brain regions, particularly the limbic system, has been observed in patients with schizophrenia [55]. The hippocampus, another important brain structure for cognition and emotion [56]. These findings highlighted that the prefrontal cortex and hippocampus might be important brain regions associated with PMH.

Interestingly, the candidate genes are significantly expressed in inhibitory neurons in the hippocampus, suggesting a potential role in regulating neural activity and maintaining the balance between excitatory and inhibitory signals. Imbalances in excitatory and inhibitory neurotransmission have been implicated in mental disorders, including major depressive disorder [59]. Stress during development has been shown to affect neuronal proliferation and differentiation in the mammalian (e.g., rats and monkeys) hippocampus [60].

In addition, we found that the expression levels of certain candidate genes were lower in individuals from the intensively farmed pig breed (Duroc) than in those from the nonintensively farmed pig breed (Jinhua). These genes, including *HSPA5*, *HSPA8*, *KCNC2*, etc., have been implicated in genetic risk factors for bipolar disorder and schizophrenia [61]. The downregulation of these genes may contribute to mental health in individuals exposed to intensive farming conditions.

Our findings highlighted that PMH might affect economic traits. First, we observed that PMH candidate genes were significantly enriched in meat quality and carcass traits in pigs. Additionally, through differential gene expression and protein interaction network analysis, *HSPA5* and *HSPA8* were identified as hub genes, indicating their close association with heat stress, which can effect meat quality and the immune response [62]. Furthermore, the GWAS signal enrichment analysis of PMH candidate genes (**Fig. 4D**)

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

revealed a significant association between these candidate genes and important production traits for meat quality, such as corrected 115 kg backfat thickness, corrected day of age to reach 115 kg, and offset intramuscular fat, both of which are important production traits for meat quality. Among these traits, backfat thickness is recognized as an essential indicator of fattening performance in commercial pigs. Notably, we identified seven genes (*ASS1, MC4R, C3, INSR, LMNB2, TUBB4A*, and *SCARB2*) with a total of 64 specific loci that were significantly associated with backfat thickness. These findings provide novel valuable insights into the relationship between PMH genes and meat quality traits, specifically focusing on backfat thickness and off set intramuscular fat.

CONCLUSION

Overall, in this study, we systematically identified a candidate gene list related to PMH at the genome level for the first time. Moreover, we provided comprehensive functional annotations of these putative genes. Our results will aid in the genomic selection of mental health traits in pigs, improve animal welfare, and facilitate the use of pigs as models of human psychiatric disorders.

ACKNOWLEDGEMENTS

We thank C. Cao and Q. Xie for help in collecting the brain tissues of pigs. We thank Z. Wang and Z. Zhang et al. for providing the whole-genome data for pigs.

Author contributions: Z. Wang designed the research framework of the paper and directed the writing of the manuscript. X. Wei wrote the manuscript. X. Wei., L. Xu. and J. Jiang performed the bioinformatics analyses. Z.Z., Q.S.W. and Y.C.P. provided useful suggestions for the analysis. Z. Chen, F. Wu and J. Miao conducted a rigorous review of the manuscript.

Funding: This work was supported by the Natural Science Foundation of China (32172691) and the National Key Research and Development Program of China (2022YFF1000500 and 2021YFD1200802).

Conflicts of interest： The authors declare that they have no competing interests.

Data availability： The genotype datasets generated and/or analyzed during the current study are available at PHARP (http://alphaindex.zju.edu.cn/PHARP/index.php) and at the SRA repository, https://www.ncbi.nlm.nih.gov/sra. See the ' MATERIALS AND METHODS ' section above for their availability. Computer code for data processing is available from the authors upon request.

REFERENCS

1. Barnett, J. K. L., Hemsworth, P., Cronin, G., Jongman, E. & Hutson, G. D. A review of the welfare issues for sows and piglets in relation to housing. *Crop and Pasture Science* **52**, 1–28 (2000).
2. Jarvis, S. *et al.* Programming the offspring of the pig by prenatal social stress: Neuroendocrine activity and behaviour. *Hormones and Behavior* **49**, 68–80 (2006).
3. Hyun, Y., Ellis, M. & Johnson, R. W. Effects of feeder type, space allowance, and mixing on the growth performance and feed intake pattern of growing pigs. *Journal of Animal Science* **76**, 2771–2778 (1998).
4. Stookey, J. M. & Gonyou, H. W. The effects of regrouping on behavioral and production parameters in finishing swine1. *Journal of Animal Science* **72**, 2804–2811 (1994).
5. Kongsted, A. G. Stress and fear as possible mediators of reproduction problems in group housed sows: a review. *Acta Agriculturae Scandinavica, Section A - Animal Science* (2004) doi:10.1080/09064700410032031.
6. Burgos-Paz, W. *et al.* Porcine colonization of the Americas: a 60k SNP story. *Heredity* **110**, 321–330 (2013).
7. Petrelli, S. *et al.* Population genomic, olfactory, dietary, and gut microbiota analyses demonstrate the unique evolutionary trajectory of feral pigs. doi:10.1111/mec.16238.
8. Jelsing, J. *et al.* The postnatal development of neocortical neurons and glial cells in the Göttingen minipig and the domestic pig brain. *Journal of Experimental Biology* **209**, 1454–1462 (2006).

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

9. Chen, K., Baxter, T., Muir, W. M., Groenen, M. A. & Schook, L. B. Genetic Resources, Genome Mapping and Evolutionary Genomics of the Pig (Sus scrofa). *Int J Biol Sci* **3**, 153–165 (2007).
10. Chen, C. *et al.* A Genome Wide Detection of Quantitative Trait Loci on Pig Maternal Infanticide Behavior in a Large Scale White Duroc × Erhualian Resource Population. *Behav Genet* **39**, 213–219 (2009).
11. ZHANG, X.-L. *et al.* Experimental primates and non-human primate (NHP) models of human diseases in China: current status and progress. *Dongwuxue Yanjiu* **35**, 447–464 (2014).
12. Lunney, J. K. *et al.* Importance of the pig as a human biomedical model. *Science Translational Medicine* **13**, eabd5758 (2021).
13. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* **13**, 537–551 (2012).
14. Rappaport, N. *et al.* MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Res* **45**, D877–D887 (2017).
15. Grissa, D., Junge, A., Oprea, T. I. & Jensen, J. DISEASES 2.0: a weekly updated database of disease–gene associations from text mining and data integration.
16. Piñero, J. *et al.* DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database (Oxford)* **2015**, bav028 (2015).
17. Matthies, S. *et al.* Small amygdala – high aggression? The role of the amygdala in modulating aggression in healthy subjects. *The World Journal of Biological Psychiatry* **13**, 75–81 (2012).
18. Rastegar-Moghaddam, S. H., Mohammadipour, A., Hosseini, M., Bargi, R. & Ebrahimzadeh-Bideskan, A. Maternal exposure to atrazine induces the hippocampal cell apoptosis in mice offspring and impairs their learning and spatial memory. *Toxin Reviews* **38**, 298–306 (2019).
19. Wang, Z. *et al.* PHARP: a pig haplotype reference panel for genotype imputation. *Sci Rep* **12**, 12645 (2022).
20. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. (American Psychiatric Association, Washington, D.C, 2013).
21. Schuyler, P. L., Hole, W. T., Tuttle, M. S. & Sherertz, D. D. The UMLS Metathesaurus: representing different views of biomedical concepts. *Bull Med Libr Assoc* **81**, 217–222 (1993).

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

22. Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res* **43**, D789–D798 (2015).
23. Schriml, L. M. *et al.* Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Res* **40**, D940–D946 (2012).
24. Bodenreider, O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* **32**, D267–D270 (2004).
25. Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440 (2005).
26. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* **81**, 559–575 (2007).
27. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246 (2011).
28. Miura, S. *et al.* A new method for inferring timetrees from temporally sampled molecular sequences. *PLoS Comput Biol* **16**, e1007046 (2020).
29. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49**, W293–W296 (2021).
30. Weir, B. S. & Cockerham, C. C. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358–1370 (1984).
31. Chen, H., Patterson, N. & Reich, D. Population differentiation as a test for selective sweeps. *Genome Res* **20**, 393–402 (2010).
32. Sabeti, P. C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).
33. Nielsen, R. *et al.* Genomic scans for selective sweeps using SNP data. *Genome Res* **15**, 1566–1575 (2005).
34. Gibson, J., Morton, N. E. & Collins, A. Extended tracts of homozygosity in outbred human populations. *Human Molecular Genetics* **15**, 789–795 (2006).
35. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

36. Szpiech, Z. A. & Hernandez, R. D. selscan: An Efficient Multithreaded Program to Perform EHH-Based Scans for Positive Selection. *Mol Biol Evol* **31**, 2824–2827 (2014).
37. Pavlidis, P., Živković, D., Stamatakis, A. & Alachiotis, N. SweeD: Likelihood-Based Detection of Selective Sweeps in Thousands of Genomes. *Mol Biol Evol* **30**, 2224–2234 (2013).
38. Fang, Y. *et al.* Genome-Wide Detection of Runs of Homozygosity in Laiwu Pigs Revealed by Sequencing Data. *Front. Genet.* **12**, 629966 (2021).
39. Fonseca, P. A. S., Suárez-Vega, A., Marras, G. & Cánovas, Á. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *Gigascience* **9**, giaa149 (2020).
40. Warr, A. *et al.* An improved pig reference genome sequence to enable pig genetics and genomics research. *Gigascience* **9**, giaa051 (2020).
41. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141 (2021).
42. 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 Research* **50**, D956–D961 (2022).
43. Consortium, T. F.-P. *et al.* A compendium of genetic regulatory effects across pig tissues. 2022.11.11.516073 Preprint at https://doi.org/10.1101/2022.11.11.516073 (2022).
44. Skene, N. G. & Grant, S. G. N. Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. *Frontiers in Neuroscience* **10**, (2016).
45. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A* **112**, 7285–7290 (2015).
46. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15 (2018).
47. Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* **38**, W214–W220 (2010).

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

48. Fang, L. *et al.* Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Res* **30**, 790–801 (2020).

49. Rohde, P. D., Fourie Sørensen, I. & Sørensen, P. qgg: an R package for large-scale quantitative genetic analyses. *Bioinformatics* **36**, 2614–2615 (2020).

50. Li J. *et al.* Research Progress of DRD1 in Neurological Disorders and Functions. *kmykdxxb* **45**, 1–7 (2024).

51. El Ghaleb, Y. *et al.* CACNA1I gain-of-function mutations differentially affect channel gating and cause neurodevelopmental disorders. *Brain* **144**, 2092–2106 (2021).

52. Helbig, K. L. *et al.* De Novo Pathogenic Variants in CACNA1E Cause Developmental and Epileptic Encephalopathy with Contractures, Macrocephaly, and Dyskinesias. *The American Journal of Human Genetics* **103**, 666–678 (2018).

53. GABRB2, a key player in neuropsychiatric disorders and beyond - ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0378111921006168?via%3Dihub.

54. Sanders, S. J. *et al.* Progress in Understanding and Treating SCN2A-Mediated Disorders. *Trends in Neurosciences* **41**, 442–456 (2018).

55. Amato, D., Vernon, A. C. & Papaleo, F. Dopamine, the antipsychotic molecule: A perspective on mechanisms underlying antipsychotic response variability. *Neuroscience & Biobehavioral Reviews* **85**, 146–159 (2018).

56. Nian, H. *et al.* Effect of Noise and Music on Neurotransmitters in the Amygdala: The Role Auditory Stimuli Play in Emotion Regulation. *Metabolites* **13**, 928 (2023).

57. Lewis, D. A., Hashimoto, T. & Volk, D. W. Cortical inhibitory neurons and schizophrenia. *Nat Rev Neurosci* **6**, 312–324 (2005).

58. Abrous, D. N., Koehl, M. & Moal, M. L. Adult Neurogenesis: From Precursors to Network and Physiology. *Physiological Reviews* (2005) doi:10.1152/physrev.00055.2003.

59. Duman, R. S., Sanacora, G. & Krystal, J. H. Altered Connectivity in Depression: GABA and Glutamate Neurotransmitter Deficits and Reversal by Novel Treatments. *Neuron* **102**, 75–90 (2019).

60. Lemaire, V., Lamarque, S., Le Moal, M., Piazza, P.-V. & Abrous, D. N. Postnatal Stimulation of the Pups Counteracts Prenatal Stress-Induced Deficits in Hippocampal Neurogenesis. *Biological Psychiatry* **59**, 786–792 (2006).

61. Wedenoja, J. *et al.* Replication of Association Between Working Memory and Reelin, a Potential Modifier Gene in Schizophrenia. *Biological Psychiatry* **67**, 983–991 (2010).

62. Bejaoui, B. *et al.* Physicochemical Properties, Antioxidant Markers, and Meat Quality as Affected by Heat Stress: A Review. *Molecules* **28**, 3332 (2023).

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

FIGURES AND LEGENDS.

Fig. 1. Summary of human disease gene databases. Summary of the raw data and subset mental data records, disease, and gene numbers of A: the disease database, B: the DisGeNET database, and C: the MalaCards database. Each bubble point's three numbers correspond to entries for mental disorders, the number of diseases, and the number of genes. Red: raw data, Blue: human mental disorder related data, Purple: mental disorder related data after F class filtration, Green: mental disorder related data after DSM-5 filtration. D. The UpSet plot showing genes overlapping between the three databases.

Fig. 2. Population structure. A: Nj-tree, B: PCA, and C: admixture structure at K=2 for intensive and nonintensive pigs. The x-axis represents individuals, and the y-axis indicates the percentage of pedigree purity.

Fig. 3. Selection signatures across the autosomes of pigs. A: The distribution of selection signatures between intensively farmed pigs and nonintensively farmed pigs detected by the F_{ST} method. The x-axis represents the chromosome location of SNPs, the y-axis represents the mean F_{ST} value of SNPs, and B represents the distribution of selection signatures between intensive and nonintensively farmed pigs detected by the XP-CLR method. The x-axis represents the chromosome location of SNPs, and the y-axis represents the XP-CLR rank score of SNPs; C: the distribution of selection signatures between intensive and nonintensively farmed pigs detected by the XP-EHH method. The x-axis represents the chromosome location of SNPs, and the y-axis represents the XP-EHH values of SNPs. D: The distribution of selection signatures in nonintensively farmed pigs detected by the CLR method. The x-axis represents the chromosome location of SNPs, and the y-axis represents the CLR rank score of SNPs; E: the distribution of selection signatures in nonintensively farmed pigs detected by the ROH method. The x-axis represents the chromosome location of SNPs, and the y-axis represents the percentage of SNPs; F, UpSet plot showing selected overlapping genes identified by these five methods.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Fig. 4. Functional annotations of candidate PMH genes. A: Venn diagram showing gene overlap between PSGs and HPGs; B: GO enrichment results for candidate PMH genes; C: KEGG enrichment results for candidate PMH genes; D: QTL database enrichment results for candidate PMH genes; E: GWAS enrichment results for six economic traits; F: Expression of candidate PMH genes in different porcine brain tissues. The x-axis is the different brain regions and y-axis is the PMH genes, see Fig. S8 for details; G: Expression of candidate PMH genes in the hippocampus at days 38 and 56.

Fig. 5. The expression pattern of Pig Mental Health candidate genes in single cells. A: Cell-type enrichment analysis using EWCE method for candidate PMH genes in the hippocampus of JinHua; B: Cell-type enrichment analysis using EWCE method for candidate PMH genes in the hippocampus of Duroc; C: Cell-type enrichment analysis of candidate PMH genes in human brain; D: Cell-type enrichment analysis of candidate PMH genes in human cortex; E: the volcano plot of differentially expressed genes (DEGs) in the inhibitory neuron; F: the PPI network of the eight (center) downregulated genes in the inhibitory neuron in Duroc pig compared to Jinhua pig. Ast: astrocytes, EN: excitatory neurons, IN: inhibitory neurons, Mic: microglia, OPC: oligodendrocyte progenitor cells, Oli: oligodendrocytes, End: endothelial cells, Neu: neuron, Gran: granulocyte cell, Per: perineuronal cells, Purk1: Purkinje cells 1, Purk2: Purkinje cells 2.

Fig. 6. Flowchart of PMH candidate gene construction and functional annotation. Briefly, using three disease gene databases, we compiled a gene list of 1,642 genes associated with human mental disorders, identified their homologous genes in pigs, and then used genomic data to detect a total of 2,844 PSGs. Finally, we narrowed the number of PMH candidate genes to 254 and systematically performed functional annotation of these genes (see details in the Methods section).

SUPPLEMENTARY MATERIALS

Supplementary figures

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Fig. S1. Summary of the disease database. Violin plot of the score distributions of the raw gene–disease records and subset mental data from the A: Disease database, B: DisGeNET database, and C: MalaCards.

Fig. S2. Distribution of porcine mental health candidate genes across chromosomes.

Fig. S3. Candidate genes for pig mental health detected by the F_{ST} test. Scatter plot of candidate PMH genes with F_{ST} scores and gene–disease scores from the A: DisGent, B: DISEASE, and C: MalaCards databases. D. The UpSet plot of candidate PMH genes across three disease databases and selected genes detected by the F_{ST} approach.

Fig. S4. Pig mental health candidate genes detected by the XPCLR method. Scatter plot of candidate PMH genes with XPCLR scores and gene–disease scores from the A: DisGent, B: DISEASE, and C: MalaCards databases. D. UpSet plot of candidate PMH genes across three disease databases and selected genes detected by the XPCLR approach.

Fig. S5. Pig mental health candidate genes detected by the XPEHH method. Scatter plot of candidate PMH genes with XPEHH values and gene–disease scores from the A: DisGent, B: DISEASE, and C: MalaCards databases. D. UpSet plot of candidate PMH genes across three disease databases and selected genes detected by the XPEHH approach.

Fig. S6. Pig mental health candidate genes detected by the CLR test. Scatter plot of candidate PMH genes with CLRs and gene–disease scores from the A: DisGent, B: DISEASE, and C: MalaCards databases. D. UpSet plot of candidate PMH genes across three disease databases and selected genes detected by the CLR approach.

Fig. S7. Pig mental health candidate genes detected by the ROH method. Scatter plot of candidate PMH genes with F_{ST} scores and gene–disease scores from the A: DisGent,

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

B: DISEASE, and C: MalaCards databases. D. The UpSet plot of candidate PMH genes across three disease databases and selected genes detected by the ROH approach.

Fig. S8. Expression of candidate PMH genes in different porcine brain tissues.

Fig. S9. Schematic diagram of the annotation and classification of gene–disease in Disease database.

Fig. S10. Schematic diagram of the annotation and classification of gene–disease in DisGeNET database.

Fig. S11. Schematic diagram of the annotation and classification of gene–disease in the MalaCards and GeneCards databases.

Supplementary tables

Table S1. Summary of three human gene–disease databases.

Table S2. Samples and sources used for select signatures analysis.

Table S3. Disease database annotation and classifications in DSM5.

Table S4. DisGeNET database annotation and classifications in DSM5.

Table S5. MalaCards database annotation and classifications in DSM5.

Table S6. Heatmap of pig GTEX samples.

Table S7. Human mental disorder genes and sources.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Table S8. Human-pig homologous gene conversion.

Table S9. Distribution of pig mental health genes on chromosomes.

Table S10. The genes selected by the F_{ST} method.

Table S11. The genes selected by the XPCLR method.

Table S12. The genes selected by the XPEHH method.

Table S13. The genes selected by the CLR method.

Table S14. The genes selected by the ROH method.

Table S15. Pig Mental Health candidate genes.

Table S16. GO cellular component enrichment of the PMH candidate genes.

Table S17. GO biological processes enriched in the PMH candidate genes.

Table S18. KEGG pathways enriched for the PMH candidate genes.

Table S19. Differentially expressed genes in inhibitory neurons of the hippocampus.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Figure 1:

Figures

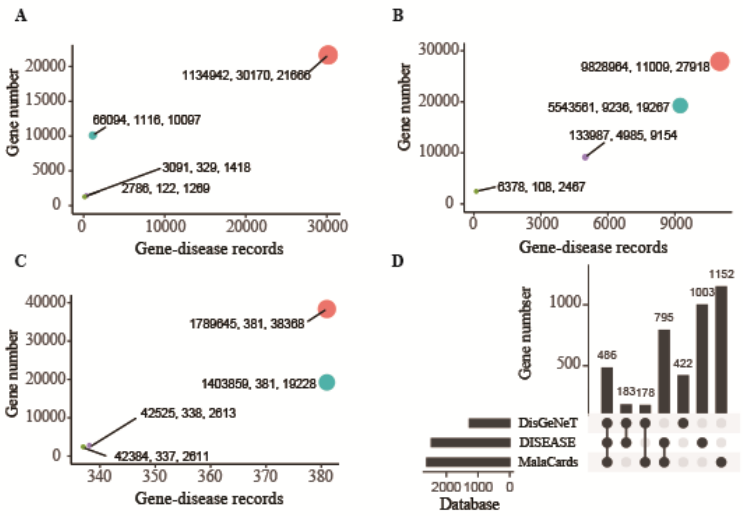

Fig. 1.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Figure 2:

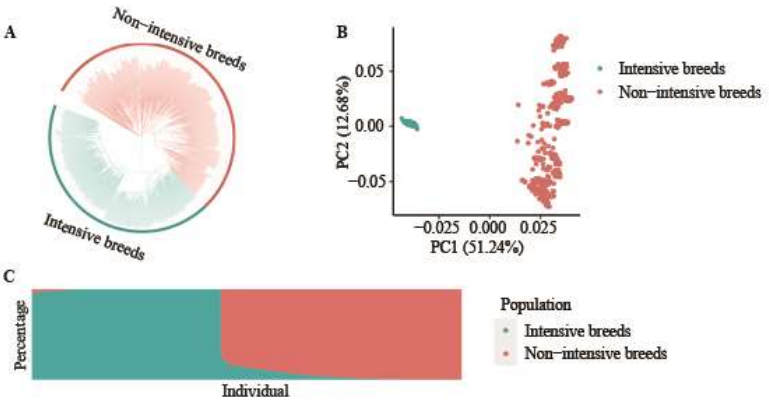

Fig. 2.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Figure 3:

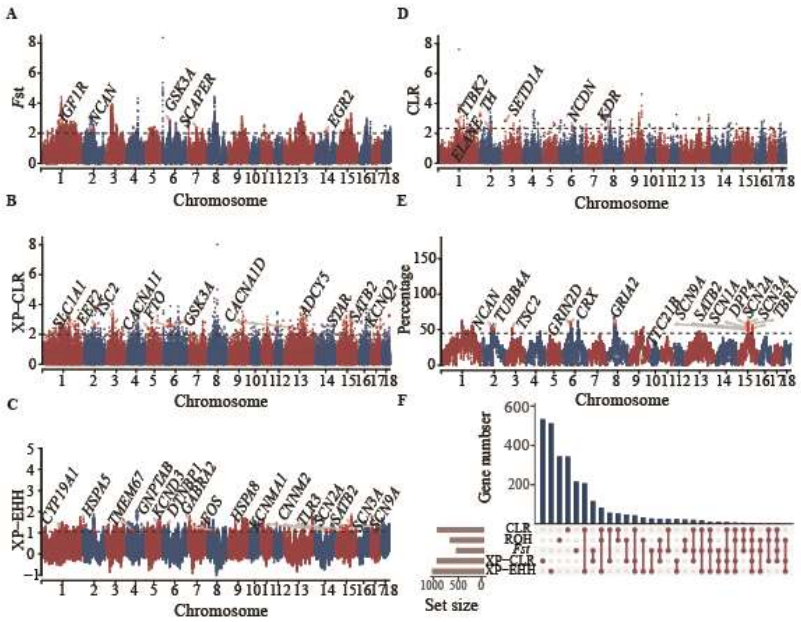

Fig. 3.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Figure 4:

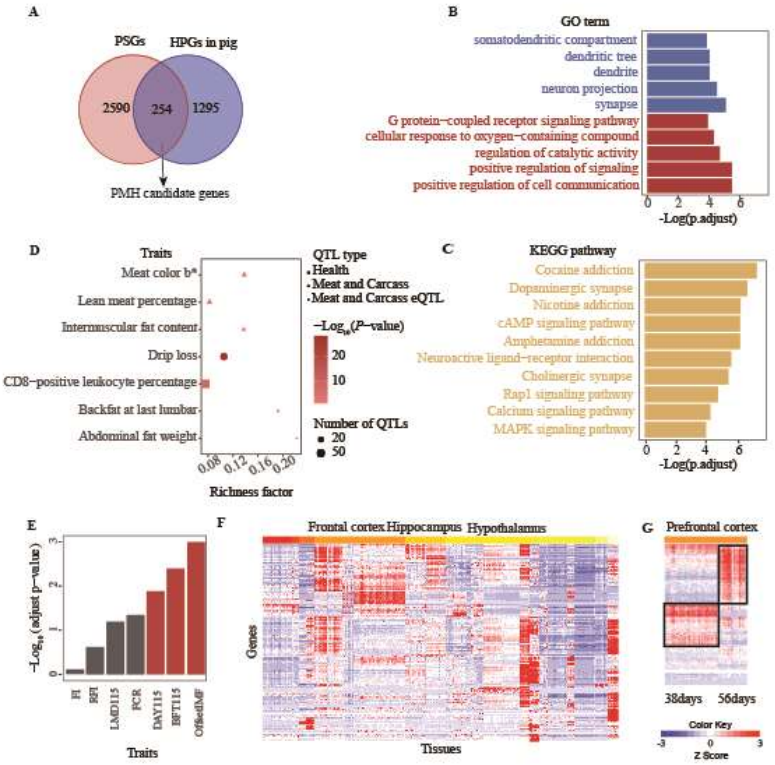

Fig. 4.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Figure 5:

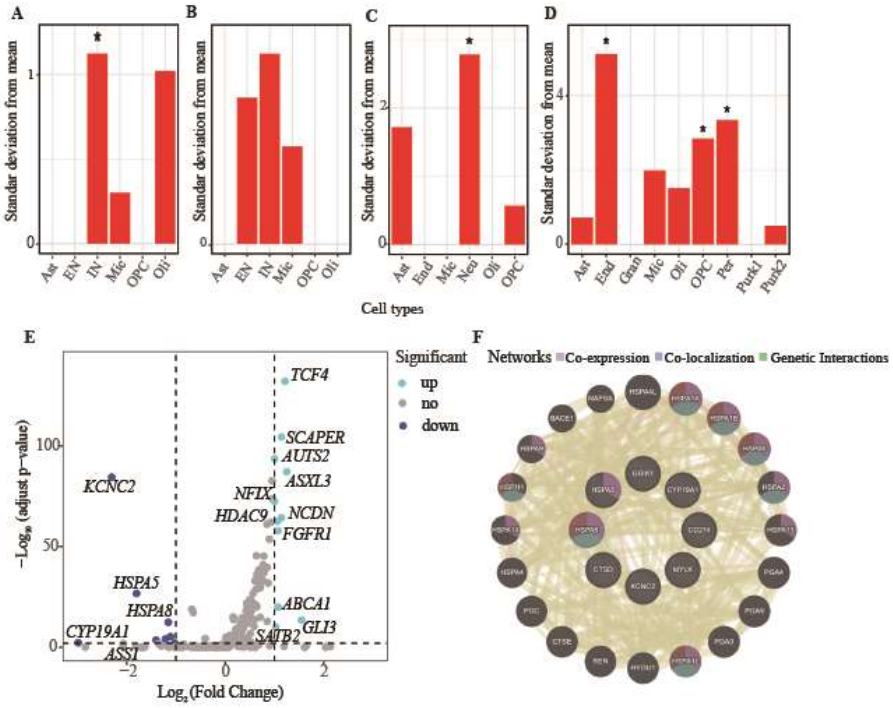

Fig. 5.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press

Figure 6:

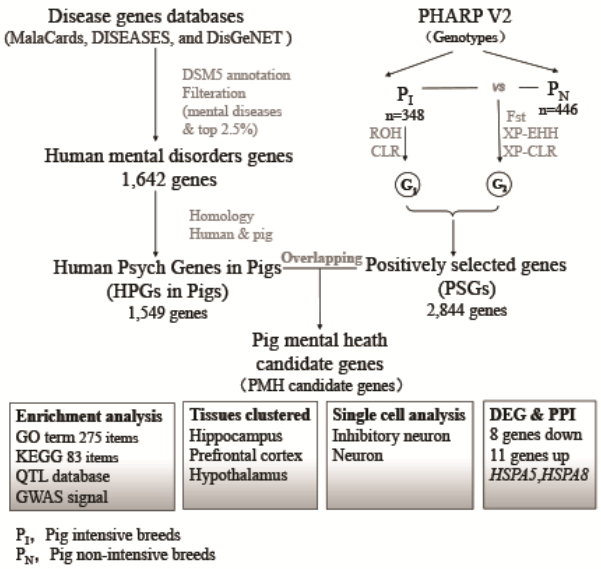

Fig. 6.

https://doi.org/10.1017/anr.2024.20 Published online by Cambridge University Press