Genomic and transcriptomic profiles influence on brain morphology and their interactions with pain sensitivity

Yiwen Pan[a,b#], Zhiguo Zhang[c,d#], Xiaoke Hao[e], Gan Huang[a,b], Zhen Liang[a,b], and Li Zhang[a,b*]

[a] School of Biomedical Engineering, Medical School, Shenzhen University, Shenzhen, China, 518060

[b] Guangdong Provincial Key Laboratory of Biomedical Measurements and Ultrasound Imaging, Shenzhen, China, 518060

[c] Institute of Computing and Intelligence, Harbin Institute of Technology, Shenzhen, China

[d] Peng Cheng Laboratory, Shenzhen, Guangdong, 518055, China

[e] School of Artificial Intelligence, Hebei University of Technology, Tianjin, 300401, China

[#] **Equal Contributions**

Yiwen Pan and Dr. Zhiguo Zhang

*** Corresponding Author**

Dr. Li Zhang

E-mail: lzhang@szu.edu.cn

Abstract

Pain sensitivity varies widely among people and is influenced by multi-omics factors, including genotypes, gene expression, and brain morphology. Though previous studies have reported the association between pain sensitivity and brain morphology, it remains unclear how genetic profiles interact with the brain to influence individual pain sensitivity. In this study, we employed aggregated datasets of magnetic resonance imaging and single nucleotide polymorphism (SNP) genotypes from 432 healthy participants, alongside gene expression data from the Allen Human Brain Atlas (AHBA), to investigate the complex multi-omics interplay affecting pain sensitivity. We first examined the individual pain threshold using laser stimuli and observed brain structure differences between high and low pain sensitivity groups. Next, we identified two key gene sets with polarized expression patterns associated with brain morphology variations, enriched in molecular functions related to ion channels and transmembrane transporter activities. Further statistical and mediation analyses clarified the roles of several specific SNPs from the genes *ECM1*, *SLC24A2*, and *SCN9A* in modulating pain sensitivity, mediated by brain structure alterations in multiple regions of the basal ganglia. Finally, our study suggested an interpretive model that integrates genomic, transcriptomic, and neuroimaging data, offering a comprehensive picture of the multi-omics contributions in shaping individual pain sensitivity.

Keywords

Pain sensitivity, Magnetic resonance imaging, Brain structure, Gene expression, Single nucleotide polymorphism

1. Introduction

Pain sensitivity varies significantly among people. Individual differences in pain sensitivity can be influenced by various factors including genetics, cognitive states, mental health and the environment[1-4]. These factors play a crucial role in understanding susceptibility to pain conditions and underlying mechanisms to develop effective pain management strategies. In recent years, researchers using brain imaging and genetic information have shown that alterations in brain structure[5] and gene[6,7] can lead to pain sensitivity variations. However, the relationship among the brain's anatomical characteristics, the gene expression and the genotyping variations underlying pain sensitivity remains unclear.

Structural magnetic resonance imaging (sMRI) is a promising non-invasive modality for the brain-based characterization of pain. It is potentially highly applicable due to its reliability[8] and wide utility in the clinical field. Many neuroimaging studies have reported that anatomical structure plays a crucial role in the representation and modulation of pain sensitivity. Voxel-based morphometry (VBM) studies in recent years have shown that pain sensitivity is significantly correlated to the morphology of various structures, including the cingulate cortex, insula, thalamus, putamen and pallidum[5,9-11]. For example, Grinsvall and colleagues measured pain sensitivity in women with irritable bowel syndrome (IBS) using pressure pain thresholds and examined gray matter properties in the sensorimotor network using MRI[12]. Their results revealed that higher pain sensitivity was associated with reduced gray matter volume in certain regions of the sensorimotor network among women with IBS. Additionally, diverse metrics quantifying brain anatomical features, such as gray matter density (GMD) have been proposed to achieve a precise presentation of brain morphorlogy. A previous study measured pain sensitivity using various experimental pain tests and used MRI to assess GMD in different brain regions. The results showed an inverse relationship between pain sensitivity and GMD, highlighting the potential role of GMD in modulating pain perception[5].

Meanwhile, genetic studies suggested that pain sensitivity is highly heritable, which implies that pain sensitivity may be influenced by complex transcriptomics and genomics interactions[13]. Evidence can be found

from both mice experiments and twin studies[14-17]. Several genomic studies have identified specific single nucleotide polymorphisms (SNPs) in certain genes that significantly influence pain sensitivity. For instance, variations in the μ-opioid receptor gene OPRM1, responsible for the physiological anti-pain effects of endorphins, have been linked to differences in experimental pain sensitivity[18,19]. The transcriptomic studies also provide valuable insights into the expression levels of genes and their involvement in pain perception. For example, Starkweather et al. examined the complex interplay between somatosensory function and gene expression in individuals with acute low back pain[7]. Gene expression analysis revealed differential expression patterns in genes associated with inflammation, neural signaling, and tissue remodeling, potentially paving the way for targeted interventions and personalized approaches in managing acute low back pain[7]. These emerging discoveries urge us to construct a model that combines brain anatomical features, transcriptomics and genomics factors for a more comprehensive investigation of pain sensitivity individual differences.

Therefore, in this work, we focus on brain imaging transcriptomics and genomics study using a self-collected dataset integrating brain imaging and genomics data, together with the Allen Human Brain Atlas (AHBA) microarray dataset[20]. We try to explore the relationships between genetic factors and brain morphorlogy in pain sensitivity. We aimed to address three research questions (Fig. 1): Q1. What are the transcriptomic features that are associated with pain sensitivity-related brain structural alternation? Q2. What are the brain phenotype differences among various SNP genotyping and pain sensitivity groups? and Q3. What are the possible interactions among genetic factors, brain factors and pain sensitivity?

To answer these questions, in total of 432 healthy participants were enrolled in our study and their pain sensitivity were measured using laser stimuli. MRI and SNP data were gathered to examine the brain imaging and genomics patterns. The inclusion of the open-source AHBA dataset offered us valuable insights into how alterations in gene expression of different brain regions are associated with macroscale brain phenotypes (brain imaging) and microscale genotypes (SNP). This study addresses three key aspects concerning pain sensitivity. Firstly, it seeks to identify brain regions displaying significant disparities between individuals with high and

low pain sensitivity. Secondly, by integrating transcriptomic gene expression data and MRI data, it aims to identify key genes associated with pain sensitivity. Thirdly, it tries to uncover the relationship and mediation pattern between microscale pain sensitivity-related SNP and macroscale brain anatomical features. To achieve these goals, we conducted a series of comprehensive analyses using partial least squares (PLS) regression to establish the connections between GMD differences in brain regions between high/low pain sensitivity groups and transcriptomic gene expression patterns. Clustering and enrichment analysis were further applied to gain deeper insights into the identified key genes. Additionally, a series of statistical analyses using genomic SNP data are conducted to explore key gene loci that influence GMD of specific brain regions and pain sensitivity. Moreover, we constructed a mediation analysis that integrated the genomics, transcriptomics and neuroimaging phenotypes to build an aforementioned model, which has the potential to reveal candidate genes that contribute to pain sensitivity individual differences.

2. Materials

2.1. Participants

A total of 432 healthy participants (259 females, average age: 21.3 ± 3.99 years) were recruited for this experiment via college and community advertisements. Twenty left-handed participants were excluded from the analysis due to differences in brain structure and function observed between left-handed and right-handed individuals[21-23]. Before the experiments, careful screening procedures were conducted to ensure that participants had no history of chronic pain, neurological conditions, cerebrovascular diseases, coronary heart disease, or mental health disorders. Additionally, participants were verified to have no contraindications for undergoing MRI examinations. The study obtained approval from the local ethics committee, and all participants provided written informed consent before their participation. The data collected in this study are not available to the public, as the participants did not consent to the public dissemination of their information. An outline illustrating the data collection and analysis methods utilized in the study is summarized in Fig. 1.

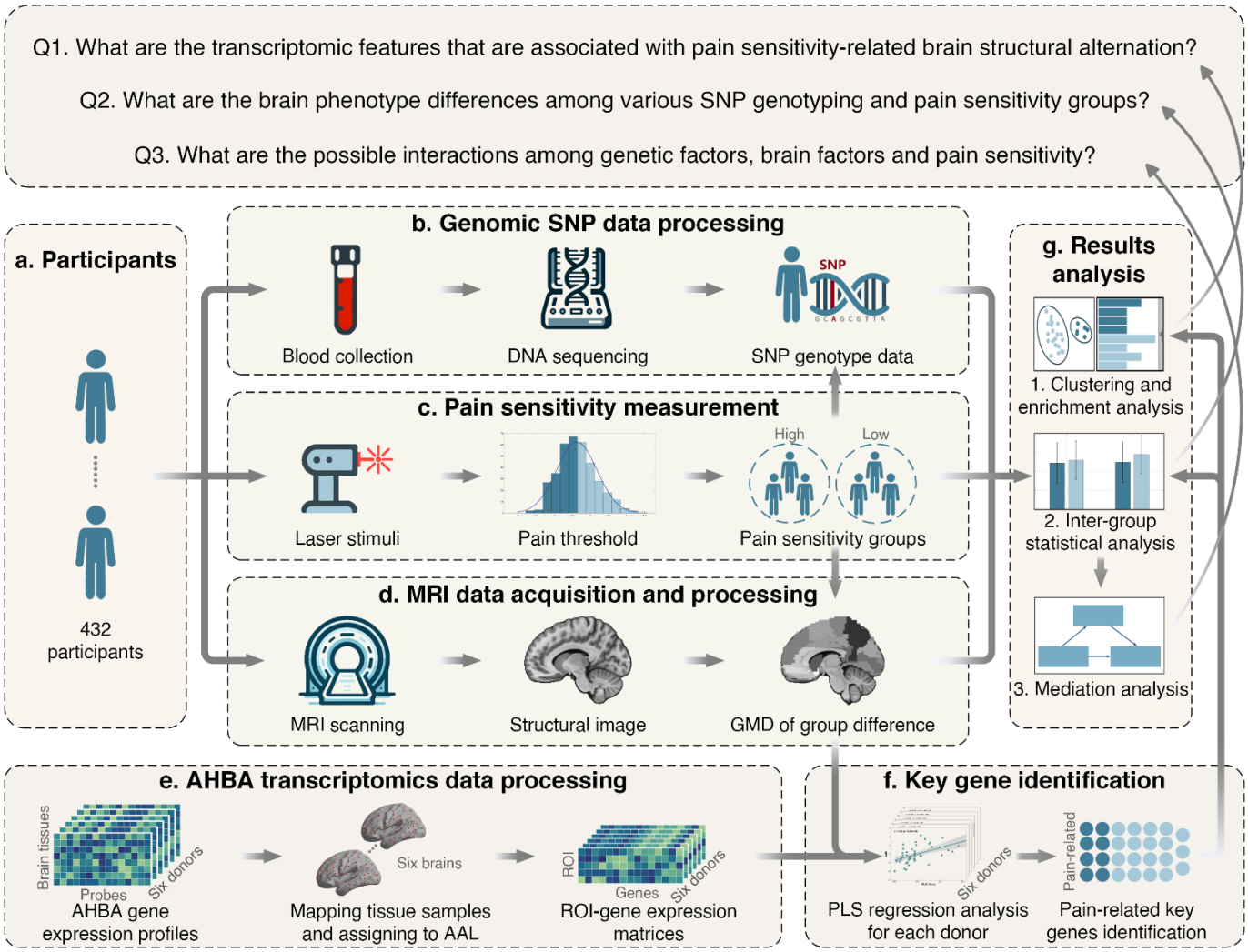

Fig. 1. The research questions and the overall flowchart of this study. This study aims to answer three questions (Q1 to Q3). The data collection and analysis methods are as follows: (a) First, in total 432 healthy participants were enrolled in this study. (b) Blood samples were collected to obtain the SNP genotype data. (c) The individual pain sensitivity was measured via laser stimuli and the subjects were partitioned into high and low sensitivity groups. (d) Brain structural images were obtained by MRI scanning and the GMD differences between high and low sensitivity difference were calculated. (e) AHBA gene expression data were processed to get ROI-gene expression matrices for six donors. (f) By using PLS regression to integrate brain imaging GMD group difference features and gene expression features, pain-related key genes were identified. (g) 1. Clustering and enrichment analysis were conducted for the key genes to answer question Q1; 2. By incorporating the SNP genotype of the key genes, statistical analysis was applied to reveal inter-group differences in terms of pain sensitivity and genotypes to answer question Q2; 3. and finally mediation analysis was applied to discover possible relationships among gene factor, brain factor and pain sensitivity to answer question Q3.

2.2. Pain sensitivity definition

We used the pain threshold to define one's pain sensitivity. Unlike other criteria such as pain tolerance, pain threshold is more stable within an individual[24,25]. Specifically, this study rated the participants' pain sensitivity by using a laser pain threshold before the MRI scan. We first recorded the skin temperature of each

participant. A series of infrared neodymium:yttrium-aluminum-perovskite (Nd:YAP) stimuli were applied to the area between the thumb and index finger on the back of the participant's left hand. To prevent skin receptor fatigue due to repeated laser stimulation, the laser position was randomly changed after each stimulus. The process began with an energy level of 1 J, incrementing by 0.25 J for subsequent stimuli. After each stimulus, participants rated their pain experience on a scale ranging from 0 (no pain at all) to 10 (unbearable pain). When a participant reported a rating of 4, the corresponding energy level was recorded as their laser pain threshold. Each participant underwent two separate measurements within a 1-hour interval, and the laser pain threshold was calculated by averaging these measurements. The laser pain threshold values underwent z-score normalization to establish the pain sensitivity threshold to meet a normal distribution. To make the high and low sensitivity groups more distinguishable, we used the median pain sensitivity threshold (the laser pain threshold was used as the primary ranking, and skin temperature was used as the secondary ranking since it has been proven to be related to pain sensitivity[26,27] as a benchmark and excluded 10% of participants around such a benchmark from the analysis based on their pain thresholds and skin temperatures. The remaining individuals' skin temperature was ranked in descending order with their pain thresholds were sorted in ascending order from low pain thresholds (high sensitivity) to high pain thresholds (low sensitivity). Based on this sorted order, the participants were divided into high/low pain sensitivity groups.

2.3. MRI data acquisition and preprocessing

MRI data were acquired using a 3.0-Telsa system (GE Medical Systems, Milwaukee, WI, United States) at the Brain and Cognitive Neuroscience Research Center, Liaoning Normal University, Dalian, China. A standard birdcage with an 8-channel head coil and restraining foam pads was used. Structural T1-weighted images were obtained using a 3D magnetization-prepared rapid gradient echo sequence with the following imaging parameters: flip angle = 8◦, field of view = 256 × 256 mm2, data matrix = 256 × 256, in-plane resolution = 1 × 1 mm2, slices = 176, and slice thickness = 1 mm. During the scan, participants were instructed

to remain awake, and not think about anything with their eyes closed.

Preprocessing steps were carried out utilizing CAT12 toolboxes with default settings within SPM12 (Statistical Parametric Mapping). Voxel-based morphometry (VBM) methodology[28] was employed to process and extract brain-wide MRI imaging phenotypes of interest. The structural images underwent segmentation into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) components by applying registration to the Montreal Neurological Institute (MNI) stereotactic space, followed by nonlinear deformation. The calculation of nonlinear deformation parameters was achieved using the built-in high-dimensional Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) algorithm[29]. The resultant warping functions generated by DARTEL were employed to spatially normalize the GM segments, which were then modulated by the Jacobian determinant. Subsequently, the normalized GM images underwent smoothing using an 8-mm full width at half maximum (FWHM) Gaussian kernel. After preprocessing the MRI data, the smoothed GM image was obtained. The automated anatomical labeling (AAL) atlas[30] was used to extract the GMD from different regions of interest (ROIs) for each participant.

2.4. Genomic SNP data

Five-milliliter blood samples were obtained from participants for DNA extraction. Genotypes of SNPs were determined using the CapitalBio Technology Corporation Precision Medicine Research Array based on the Affymetrix platform. Following the aforementioned procedures, a total of 748,659 SNP markers were collected as genome-wide genotype information. To handle possible defects of genetic data, the following quality control (QC) steps were performed using the PLINK software package (v1.9)[31]. Firstly, only SNPs located on autosomes were included, and SNPs on the sex chromosomes (27,932 SNPs) were removed. Secondly, SNPs were excluded if they could not meet any of the following criteria: (1) call rate per SNP ≥ 90% (46,803 SNPs), (2) minor allele frequency (MAF) ≥ 5% (319,751 SNPs), and (3) Hardy-Weinberg equilibrium test of $p \leq 10e\text{-}6$ (752 SNPs). Participants were excluded from the analysis if any of the following criteria was

not satisfied: (1) call rate per participant ≥ 90%; (2) gender check (5 participants were excluded). Population stratification analysis[32] was performed by applying the principal component analysis (PCA), which is considered to be a very useful method of adjusting spurious associations. This method allows for the extraction of ethnic information from the sample and the identification of potential ethnic outliers. After the QC procedure, 407 participants and 353,421 out of 748,659 SNP markers were retained in the analysis.

2.5. Transcriptomic gene expression data

The microarray-based gene expression data used in this study were sourced from AHBA (http://human.brain-map.org)[20]. This dataset contains expression data derived from six human postmortem donors, comprising one female and five males aged between 24 and 57 years. More precisely, the AHBA dataset contains the expression of 20,737 genes from the entire brain, with each gene represented by unique Entrez IDs measured through 58,692 probes across 3,702 spatially different tissue samples. The *abagen* toolbox[33] was used to process the microarray expression data, which offers standardized workflows for the AHBA dataset. The gene expression data underwent preprocessing steps as detailed in a previous study[34]: (1) Microarray probe-to-gene mappings were reannotated using the latest National Center for Biotechnology Information (NCBI) database, with probes that didn't match a valid Entrez ID discarded. (2) Probes with expression intensities lower than the background noise in over 50% of samples were filtered out. (3) In cases where multiple probes indexed the expression of the same gene, a representative probe exhibiting the highest pooled correlation across donors was selected. (4) Matching of microarray samples to brain regions was achieved using MNI coordinates, considering coordinates within 2 mm of the region boundary. (5) Gene expression values were normalized within each donor to address the inter-individual differences. This normalization involved applying the scaled robust sigmoid function and rescaling to the unit interval. (6) Genes were further refined based on differential stability criteria to retain only stable genes for subsequent analysis[35].

It's important to note that the samples from each donor were averaged separately rather than across all six donors. Additionally, because only two donors out of six comprise brain samples from the right hemisphere, we only used the left hemisphere for further analysis[34]. Following these processing steps, the expression values became more comparable across donors, resulting in the acquisition of six gene expression matrices at different donor levels for subsequent analysis of connection neuroimaging phenotype.

3. Methods

3.1. Brain imaging phenotype extraction

In our study, we first conducted a comparison of brain imaging phenotype characteristics between individuals with high and low pain sensitivity. Differences in GMD features among various ROIs were calculated using two-tailed two-sample t-tests with a significance level of $p < 0.05$, corrected by the Benjamini-Hochberg false discovery rate (BH-FDR) method. Based on the GMD differences between high and low pain sensitivity groups, we computed the average phenotype values to get an averaged phenotype representation vector for each group. Afterward, we performed a vector subtraction, yielding a differential phenotype vector for further regression analyses.

3.2. Imaging transcriptomics analysis

A PLS regression was employed to explore the relationship between the gene expression patterns of each donor and the brain phenotype. PLS regression is capable of predicting multiple response variables (brain phenotype values) from multiple predictor variables (gene expression data) by establishing an interrelation between them[36-38]. In our study, we first obtained the brain expression level of each gene from each ROI in each donor using preprocessed AHBA data. Next, both the gene expression levels and GMD difference values were normalized by subtracting the mean and dividing by the standard deviation (SD) across all samples. To evaluate the significance of GMD features in the model, we computed Variable Importance in Projection (VIP)

scores. VIP scores represent a measure of the importance of variables in multivariate data analysis, indicating the degree to which a variable contributes to explaining variations in the response variables[39]. Variables with VIP scores greater than 1 were considered significant in predicting the PLS regression model and were extracted from the previously mentioned gene expression matrix. As a result, this process led to the acquisition of six matrices of transcriptional levels (45 regions of 6 donors multiplied by corresponding gene expression levels) and a set of differential GMD values (45 regions multiplied by brain phenotypes).

In our PLS regression analysis, the predictor variables comprised normalized matrices displaying average gene expression levels across all regions for each donor level (45 × 5333; 45 × 5439; 45 × 5089; 45 × 5840; 45 × 5700; 45 × 5286). Meanwhile, the response variables were the normalized vectors of average differential GMD phenotype across all regions (45 × 1). The output of components from PLS regression was ranked based on the explained variances between the predictor and response variables, with particular emphasis on the first PLS component (PLS1). PLS1 aimed to provide an optimal low-dimensional representation of the covariance between two high-dimensional datasets, namely gene expression and differential phenotype values[38].

To examine whether the explained variance of PLS1 significantly exceeded chance expectations, we conducted permutation tests (10,000 times)[40,41]. This involved multiple iterations of shuffling data labels and performing recalculations of the PLS regression analysis to establish a null distribution of explained variances. Furthermore, we used bootstrapping techniques to estimate the significance of individual genes contributing to the PLS components. Specifically, after obtaining PLS1 weights for each gene, we performed 10,000 resampling iterations with the replacement of 45 brain regions to re-conduct the PLS regression analysis at the donor level via bootstrap tests. This process generated weights for each gene. The ratio of each gene's weight to its bootstrap standard error was computed to derive z-scores (PLS regression weights divided by their standard errors). These z-scores were transformed into p-values to evaluate the contribution of each gene to PLS1. To identify significant genes, we applied the Benjamini-Hochberg false discovery rate (BH-FDR) correction procedure, setting a significance threshold at $p < 0.05$.

To address potential biases arising from transcriptomic variations among the six donors, we constructed separate PLS regression models for each donor. Through this approach, we aimed to eliminate donor-specific influences on the analysis. Specifically, following the creation of separate PLS regression models for each donor, we identified genes that demonstrated noteworthy associations with differential phenotype values linked to pain sensitivity. Subsequently, these genes that are consistently found across all donors were selected for subsequent investigation and further analysis.

3.3. Key gene identification

By comparing the significant genes identified across all six donor levels, we identified key genes strongly associated with GMD variations related to pain sensitivity. These key genes were found to have functional information related to pain sensitivity based on the data processing steps employed. Subsequently, we utilized the gene expression features of these key genes across 45 brain regions at six donor levels for conducting k-means clustering. To determine the quality and coherence of the resulting clusters, the silhouette measurement was used to get the optimal classification groups by calculating the relative intra-cluster similarity and inter-cluster dissimilarity[42]. We further calculated the correlation between each gene based on its expression value of all ROIs and all donors from the AHBA dataset. Thus, the expression value vector was obtained by concatenating 45 ROIs' expression value of all AHBA donors, resulting in a $45 \times 6 = 270$ long vector. Pearson correlation coefficients were then calculated based on the expression value vector to determine the relationships among all genes.

Furthermore, we explored potential functional differences between the identified groups using the Gene Ontology (GO) database[43]. Specifically, we conducted GO analyses to comprehensively examine the functional roles of the identified genes concerning molecular function, biological processes, and cellular components. These GO analyses of the key genes were executed utilizing the *clusterProfiler* package within the R software, employing a significance threshold set at a p-value of 0.05.

We established a comprehensive pain gene set by integrating pain-related genes sourced from multiple databases, including the Human Pain Genetics Database[44], Pain Genes Database[45], Gene Ontology search employing the keyword "pain", and Pain Research Forum from https://www.iasp-pain.org/publications/pain-research-forum/. We compared the identified key genes related to GMD from our bioinformatics analysis with the established pain gene set to highlight significant pain-related genes associated with GMD variations. We then extracted the expression values of these genes across 45 ROIs to examine the gene expression pattern at the donor level.

3.4. SNP genotyping analysis

By integrating the self-collected preprocessed genotype SNP data, we further selected target SNPs corresponding to these genes. Additional statistical analyses were performed aimed at quantitatively determining the differences in phenotype among various genotype groups. To handle the uneven distribution of participants across different groups, one of the homozygous groups with a smaller participant count was merged with the heterozygous group, resulting in the formation of a risk allele-carrier group[46]. Meanwhile, the participants were divided into high and low pain sensitivity groups, enabling us to conduct a two-way ANOVA (2 × 2 ANOVA: pain sensitivity × genotypes). It was executed to evaluate the combined impact of the target SNPs and pain sensitivity on GMD differences within the significant ROIs. All statistical analyses were performed using SPSS 22.0 software.

3.5. Mediation analysis

To further investigate the influence of GMD on the relationship between genetic factors and pain sensitivity, we performed a mediation analysis to determine if the association between variables (SNP and pain sensitivity) can be partially explained by the effect of a mediating variable (GMD)[47]. In this study, the genotype of each SNP served as the independent variable, pain sensitivity measurements as the dependent

variable, and differential GMD values of significant ROIs as the mediating variable. The hypothetical mediation models were conducted using path analysis in structural equation modeling (SEM) with maximum likelihood estimation. In addition, a bootstrapping sampling procedure was employed to calculate bias-corrected 95% confidence intervals (CIs) for evaluating the significance of indirect and direct effects in the mediation model[48]. Bootstrapping is a nonparametric resampling procedure, and as such, it does not violate assumptions of normality. In this study, statistical significance was considered when the 95% CI (based on 5,000 bootstrap samples) excluded 0 between the lower and upper limits[49]. Standardized estimate (β), standard error (SE), 95% CI, and p-value were reported for both direct and indirect effects. All mediation analyses were performed with Amos (Version 24.0, IBM Corp, Armonk, NY).

4. Results

4.1. Participants characteristics

After preprocessing of the MRI data, 366 out of the initially recruited 407 participants remained in the present study. These pain thresholds were arranged from low (high pain sensitivity) to high (low pain sensitivity). Using a median value of -0.24 as a threshold, the participants were categorized into 2 groups: 157 individuals in the high pain sensitivity group and 209 individuals in the low pain sensitivity group. The figure illustrates that the number of participants with high pain sensitivity is relatively higher compared to those with low pain sensitivity. Subsequently, the pain thresholds underwent normalization via z-score transformation. Table 1 provides an overview of the demographic details of the sample analyzed in this study. A noteworthy observation from the independent-sample t-test regarding structural phenotypes suggests a significant difference ($p = 0.003$) in gender distribution among the participants categorized into high and low sensitivity groups. In our following analysis, the sex factor is regressed as a covariate aiming to control for its potential influence on the observed results.

Table 1. Demographic information and the total number of participants are included in each analysis.

Category	High	Low	p-value
Number of participants	157	209	-
Gender (M/F)[a]	48/109	96/113	0.003
Age (years, mean ± SD)[c]	21.3 ± 4.4	21.4 ± 4.0	0.899
Education (years, mean ± SD)[c]	14.6 ± 2.1	14.5 ± 1.9	0.633

[a]Chi-square test. [b]p-value <0.05. [c]Two-sample t-test.

4.2. GMD associated with pain sensitivity

In our analysis, we compared the GMD differences in 45 brain regions between the high and low pain sensitivity groups. As a result, we identified 15 ROIs that exhibited significant inter-group differences, with the statistical significance corrected using the BH-FDR method. The results are shown in Fig. 2. Among these 15 ROIs, we found many of them belong to the pain matrix or are highly related to pain perception. These ROIs include insula (p-value = 0.014), amygdala (p-value = 0.020), putamen (p-value = 0.014), pallidum (p-value = 0.026), and thalamus (p-value = 0.021). This particular discovery suggests that individuals in the low pain sensitivity group have higher GMD compared to those in the high pain sensitivity group.

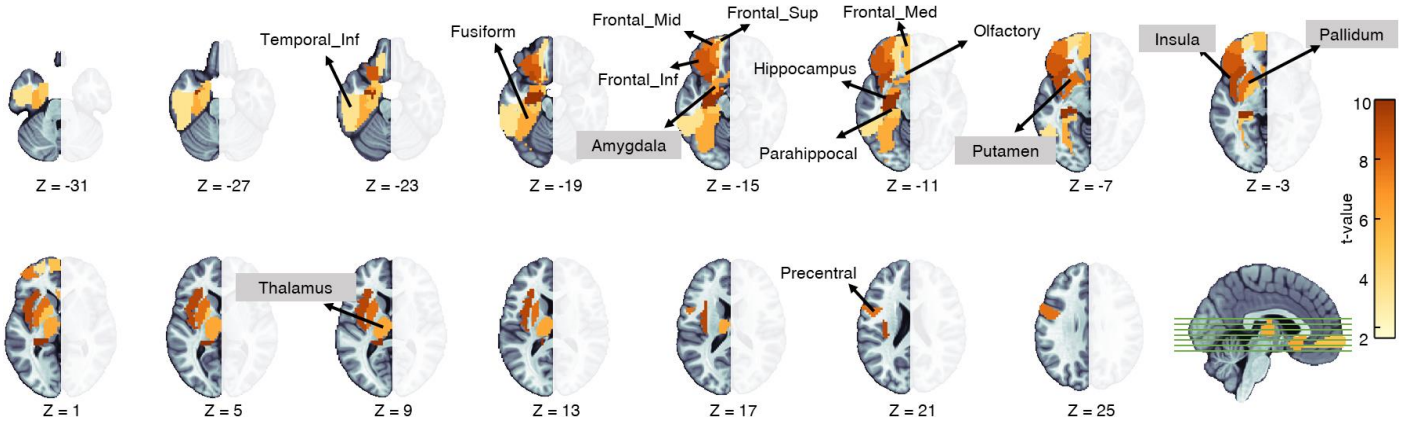

Fig. 2. Identified 15 ROIs that exhibited significant differences between high and low pain sensitivity groups in terms of GMD. The ROIs with gray backgrounds were commonly known to be belonging to the "Pain Matrix". Since only the left hemisphere of the AHBA dataset was considered in this study, here we only show identified ROIs of the left brain.

4.3. Correlations between gene expression and GMD

We used PLS regression to identify patterns of gene expression that exhibited correlations with the GMD differences across six donors. The results are presented in Fig. 3. Firstly, Fig. 3(a) illustrates the correlations between the PLS1 scores and the GMD difference values. The scatter plots demonstrate significant positive correlations between the PLS1 scores and the GMD values for each donor. The Pearson's correlation coefficients ranged from 0.49 to 0.63, demonstrating a significant relationship between these variables (all $p < 0.001$). These results imply the role played by PLS1 in capturing the association between gene expression patterns and GMD differences related to pain sensitivity across different donors. Secondly, as shown in Fig. 3(b), the permutation tests revealed that the PLS1 scores explained a significant amount of the variance in GMD across all six donors. The genes from each donor were repeatedly identified to make substantial contributions to PLS1 (permutation test, $p < 0.05$), significantly more than expected by chance. Notably, the PLS1 component independently accounted for the highest observed variance, explaining a minimum of 38.8% of the total variance.

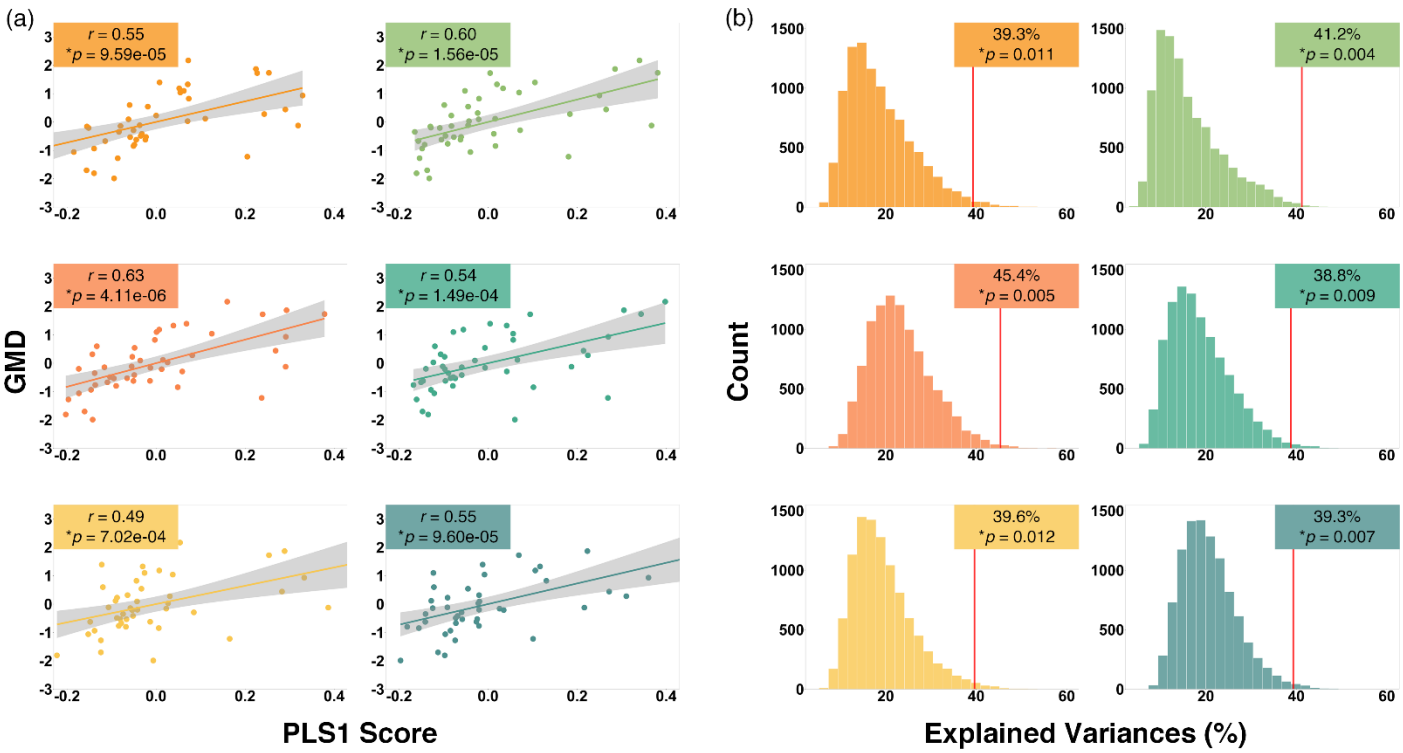

Fig. 3. (a) The results from PLS regression analysis for 6 donors. The correlations between PLS1 scores and the GMD differences values are shown in each subfigure. (b) The results from permutation tests for 6 donors.

4.4. Key gene clustering and enrichment analysis results

Furthermore, we identified genes that significantly contributed to PLS1 consistently present across all six donors using the bootstrapping method. More specifically, we identified a gene set that is associated with pain sensitivity-related GMD alternation for each donor. In order to obtain a relatively stable result, we only selected genes that were identified from all six donors for further analysis. In the end, 27 key genes were identified to be significantly associated with differences in GMD between two pain sensitivity groups and were used. The gene identification and overlapping situations are demonstrated as a Venn diagram in the supplementary materials, together with 27 key gene names and other related information.

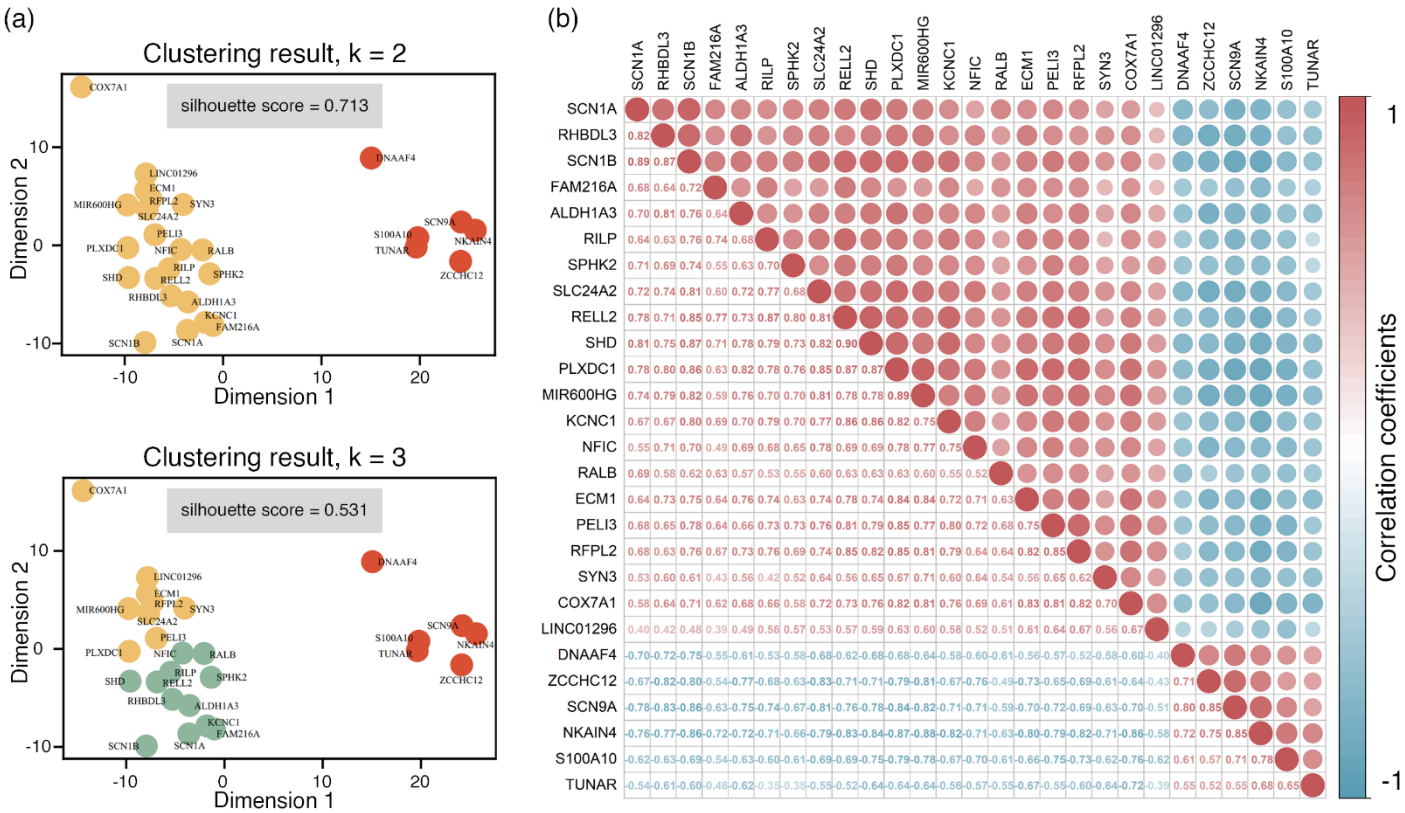

Fig. 4. (a) The *k*-means clustering results for 27 key genes by choosing $k = 2$ (top) and $k = 3$ (bottom) respectively. (b) The correlation coefficient matrix among 27 identified key genes.

Next, we used a *k*-means clustering method on these 27 key genes and the results are shown in Fig. 4(a). To compare the clustering effect with different *k*, the cluster numbers were chosen as 2 and 3 and the silhouette coefficients are 0.713 and 0.531 for $k = 2$ and 3 respectively. These results indicated that the optimal division for these 27 genes lies in categorizing them into two distinct clusters instead of three. Following the clustering

analysis, the 27 significant genes were divided into two groups: group 1 with 21 genes, and group 2 with 6 genes. Further, the gene expression correlation results are also shown in Fig. 4(b) where red indicates positive correlations and blue indicates negative correlations (corrected by the BH-FDR method, $p < 0.001$).

Figure 5(a) demonstrates the gene groups and gene network in which only high correlation (absolute value greater or equal to 0.6) connections were kept. Next, the GO enrichment analysis was conducted for each group to examine the functional characteristics associated with these genes. The results are shown in Figs. 5(b) and (c) (all $p < 0.05$). First, Fig. 5(b) shows the GO enrichment analysis for the 21 genes within group 1. These genes exhibit significant over-expression in terms of molecular functions and cellular components. The significantly enriched molecular functions contain monovalent inorganic cation transmembrane transporter activity, voltage-gated sodium channels (VGSC) activity, cation channel activity, sodium and metal ion transmembrane transporter activity and ion and sodium channel activity. Meanwhile, in terms of cellular components, the enrichment predominantly occurs in areas such as the main axon, VGSC complex, node of Ranvier, sodium and cation channel complex. On the other hand, Fig. 5(c) shows the GO enrichment analysis outcomes for the 6 genes within group 2. These genes can only be enriched in cellular components. Specifically, these genes are enriched in the VGSC complex and sodium channel complex.

4.5. Pain-related key gene expression

As presented earlier, we have identified 27 key genes from our imaging transcriptomics analysis. By taking the intersection set of these 27 key genes and the aforementioned pain-related genes set from multiple database sources (including the Human Pain Genetics Database[44], Pain Genes Database[45], Gene Ontology search employing the keyword "pain", and Pain Research from https://www.iasp-pain.org/publications/pain-research-forum/), we further narrow our analysis down to 8 significant pain-related genes: *ECM1*, *SCN1A*, *SCN1B*, *SLC24A2*, *SPHK2*, *SYN3*, *S100A10* and *SCN9A*. More details of this intersection can be found in the supplementary document. Fig. 6 shows the expression values of these 8 genes on 45 ROIs. In the figure, the circle size is determined by the expression level among all genes and all ROIs. For each row (gene), the

expression is respectively ranked from largest to smallest. The ranking order is indicated below each circle. The color of each circle represents the brain region in the AAL atlas of the corresponding ROI (blue: central region, green: frontal lobe, orange: insula, cyan: limbic lobe, red: occipital lobe, gray: parietal lobe, purple: subcortical nuclei, and yellow: temporal lobe). The ROIs with a gray background are brain regions that belong to the "pain matrix", including the dorsolateral prefrontal cortex (dlPFC), insula, anterior cingulate cortex (ACC), S1, S2 and regions from basal ganglia (amygdala, caudate, putamen, pallidum and thalamus). Based on the previous k-means clustering result, the first 6 genes belong to group 1 (darker colors) and the other 2 genes belong to group 2 (lighter colors). The bar plot in Fig. 6 indicates the normalized gene expression of these 8 genes at different brain regions. The overall bar height is the total expression level of all 8 genes. Within each bar, the darker color represents the expression ratio of the first 6 genes (group 1) and the lighter color represents the other 2 genes (group 2).

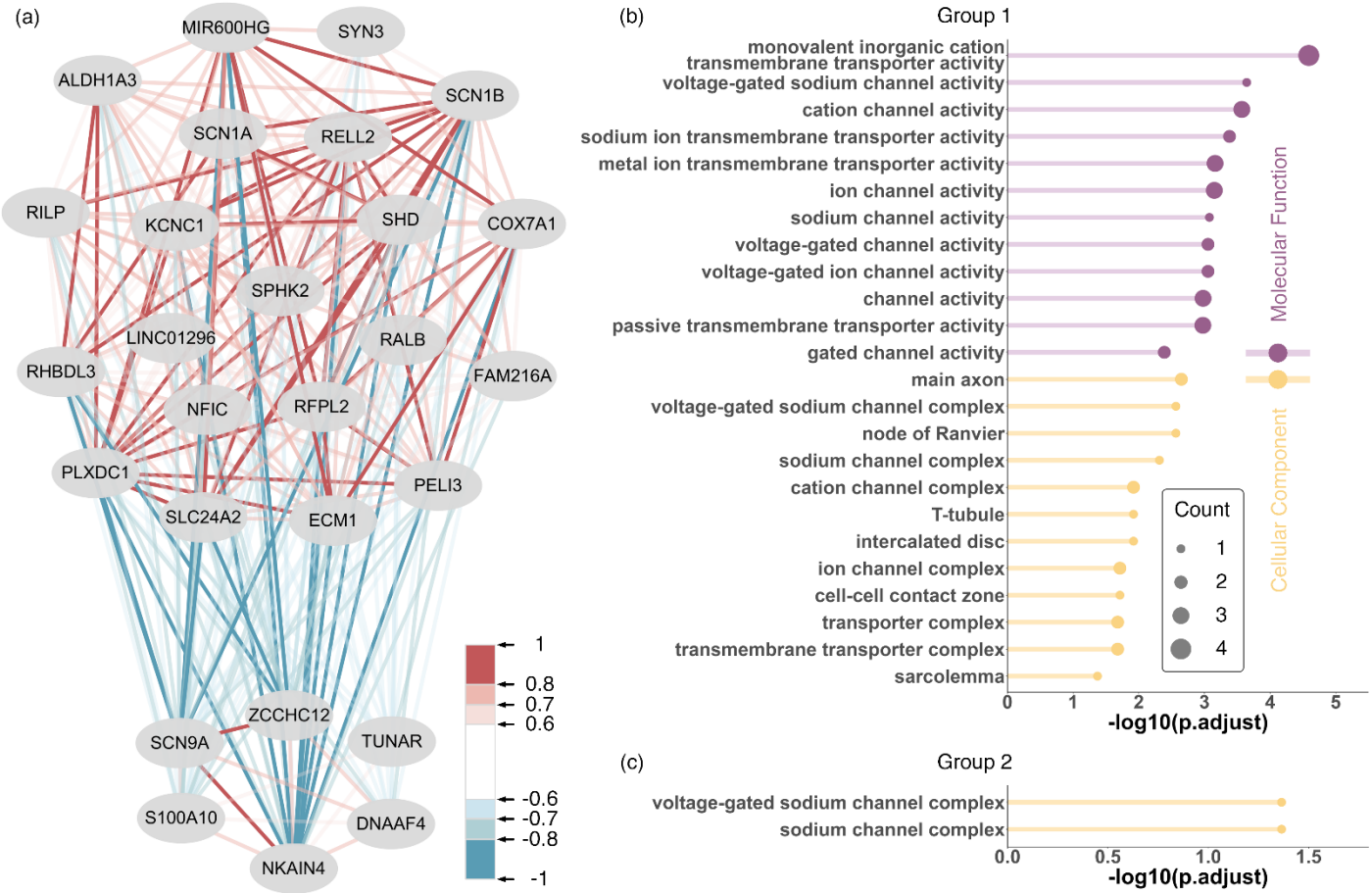

Fig. 5. (a) The clustering results for 27 key genes. These genes can be clustered into two groups. The upper group is group 1 with 21 genes and the lower group is group 2 with 6 genes. The lines between each gene pair indicate the expression correlation among all ROIs, where red indicates a positive correlation and blue indicates a negative correlation. The correlation value is indicated as the color bar. (b) and (c) The GO analysis results for genes in groups 1 and 2 respectively.

From this figure, we can observe that the regions from basal ganglia (shown in purple) show a polarized expression pattern. Specifically, genes in group 1 (*ECM1*, *SCN1A*, *SCN1B*, *SLC24A2*, *SPHK2* and *SYN3*) have small expression values, and genes in group 2 (*S100A10* and *SCN9A*) have large expression values in these areas. It is especially true for amygdala, caudate, putamen and pallidum. By comparing the gene expression level with other brain regions, genes in group 1 from basal ganglia have almost the smallest expression values among all 45 ROIs, and genes in group 2 have the largest expression value. For example, the top 4 out of 5 top-expressed ROIs of gene S100A10 are in the basal ganglia region. On the other hand, these 4 ROIs have the smallest expression value for the SPHK2 gene. To verify the clustering result, we also conducted the *k*-means clustering method based on the 8 identified key genes based on whole-brain 45 ROIs and pain matrix ROIs (shown in gray background in Fig. 6). Interested readers may refer to the supplementary document for more information.

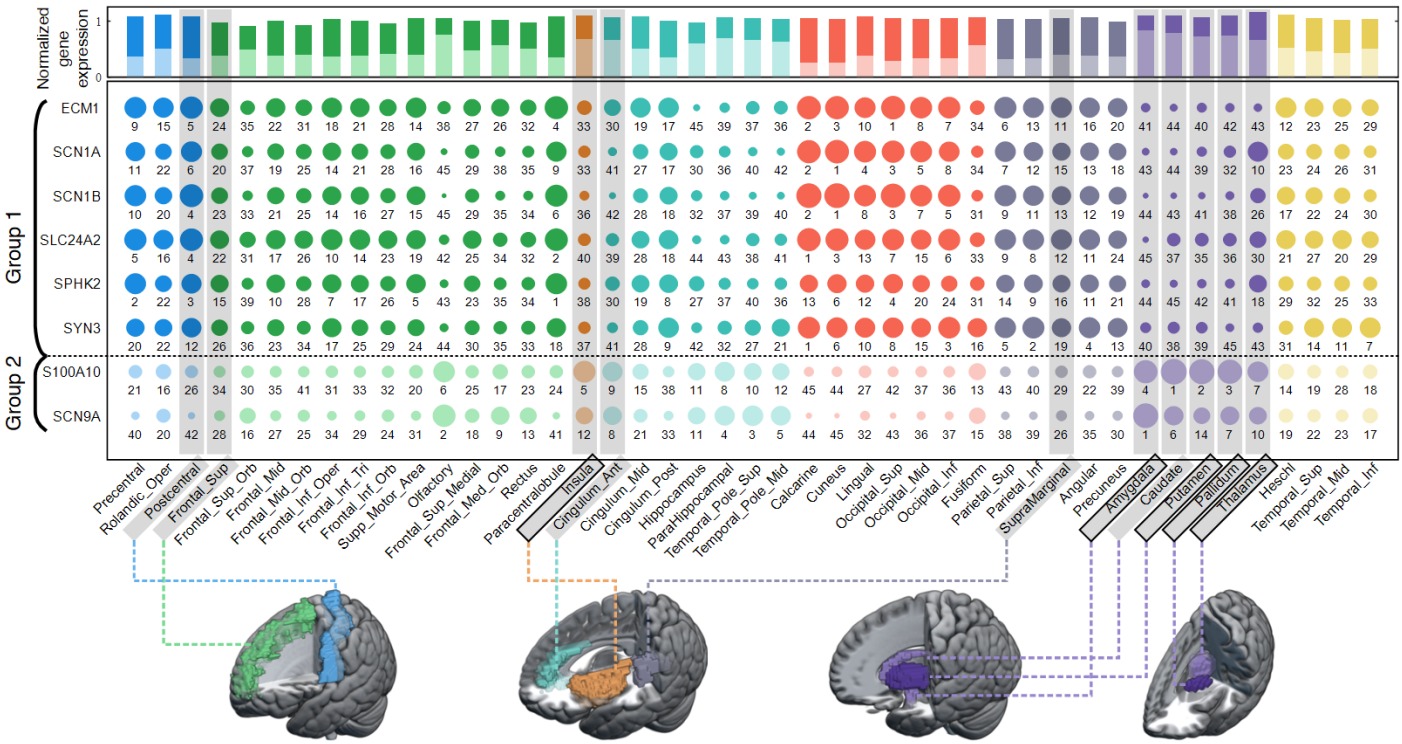

Fig. 6. The expression values of 8 key genes (*ECM1*, *SCN1A*, *SCN1B*, *SLC24A2*, *SPHK2*, *SYN3*, *S100A10* and *SCN9A*) at different brain regions. The circle size is determined by the expression level among all genes and all ROIs. For each row, the expression value is ranked from largest to smallest and the ranking order is listed below the circle. The bar plot indicates the normalized gene expression of these 8 genes at different brain regions, where the darker part is from group 1 (genes 1 to 6) and the lighter part is from group 2 (genes 7 and 8). Brain regions with gray backgrounds are known to belong to the "pain matrix".

4.6. SNP genotyping statistical results

Further, by integrating our self-collected genotyping dataset, we investigated SNPs of the aforementioned 8 pain-related key genes, revealing 3, 21, 0, 51, 0, 86, 3 and 27 SNPs for them respectively. Subsequently, we conducted a two-way ANOVA analysis on all these SNPs to examine the relationships between the genotypes and pain sensitivity on GMD in significant brain regions. The analysis contains 15 previously mentioned important ROIs that exhibited significant differences in GMD between high and low pain sensitivity groups, including the insula, hippocampus, fusiform, amygdala, putamen, pallidum, thalamus, and so on. As a result, we identified SNPs from 3 genes, namely *ECM1*, *SCN9A* and *SLC24A2* with significant statistical results.

For the *ECM1* gene, there are 3 SNPs included in our dataset. After performing the statistical analysis, one SNP rs13294 showed significant results between pain sensitivity and genotyping groups. The results are shown in Figs. 7(a) and (b). There are in total 366 participants (156 with high pain sensitivity: 78 AX, 78 GG; 208 with low pain sensitivity: 86 AX, 122 GG; 2 missing values). Due to the limited number of participants with an AA genotype of rs13294, AA and AG genotypes were grouped as AX. In both Figs. 7(a) and (b), the participants from the low sensitivity group have higher insula and amygdala GMD values than those in the high sensitivity group (p-value = 0.0014 for insula, p-value = 0.0037 for amygdala). The main effects of genotype show that the GMD of the insula and amygdala of the AX group (dark yellow) is lower than the GG group (light yellow) with p-values = 0.0137 and 0.0474 respectively. Moreover, a significant interaction effect can be found between the two independent variables with p-values = 0.0463 and 0.0476 for the insula and amygdala respectively.

For the *SLC24A2* gene, in total 51 SNPs can be found in our genotyping dataset. We examined the effect of these target SNPs and pain sensitivity groups on the GMD from ROIs with significant differences. One SNP rs55682638 demonstrated significant results as shown in Figs. 7(c) and (d). For both results from Figs. 7(c) and (d), TT and TC genotypes are grouped with in total of 50 participants. The CC group has 315 participants. From the results, we can see that there is a significant main effect of pain sensitivity on GMD of

pallidum and putamen respectively (p-value = 0.0120 and 0.0024). This indicates that participants of the high pain sensitivity group have lower GMD of both pallidum and putamen. Similarly, genotype also shows a significant main effect on the GMD value, with p-values of 0.0021 and 0.0013 for rs55682638 on the GMD of pallidum and putamen. Moreover, the interaction between pain sensitivity and genotype is significant, with p-values of 0.0249 and 0.0485 respectively for the two ROIs. The significant interaction suggests that the influence of genotype on GMD depends on pain sensitivity and vice versa.

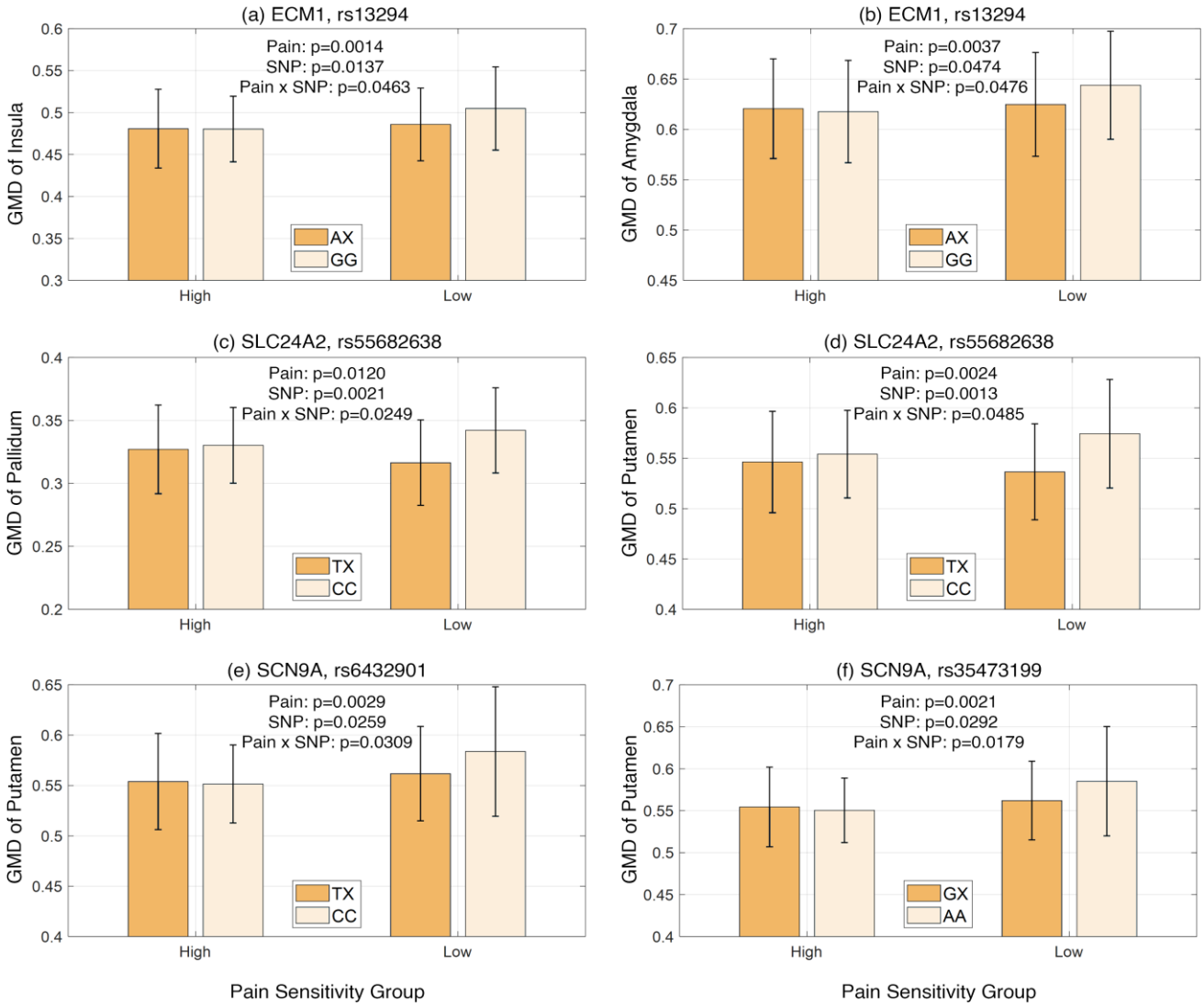

Fig. 7. Statistical analysis of identified pain sensitivity-related brain region GMD about corresponding SNPs and pain sensitivity groups. Two-way ANOVA was conducted to examine the effect of each SNP and pain sensitivity on the GMD of these regions. The p-values for the main effect and interaction effect are shown in the plot. The bars and error bars indicate the mean and standard deviation of GMD.

Similarly, after performing the two-way ANOVA of 27 SNPs from *SCN9A* on GMD of pain matrix ROIs, two SNPs rs6432901 and rs35473199 show significant results between pain sensitivity and genotyping groups.

The results are shown in Figs. 7(e) and (f). Both significant results are related to the GMD of putamen. Firstly, Figs. 7(e) and (f) indicate that participants with high pain sensitivity have significantly lower putamen GMD values for both target SNPs (p-value = 0.0029 and 0.0021 for rs6432901 and rs35473199 respectively). In Fig. 7(e), TT and TC genotypes were grouped as TX, similar to the previous analysis. The GMD of putamen of the TX group (dark yellow, n = 236) was significantly lower than the CC group (light yellow, n = 124) of rs6432901 (p-value = 0.0259). In Fig. 7(f), we can also find that the putamen GMD of the GX group (dark yellow, n = 242) was significantly lower than the AA group (light yellow, n = 120) of rs35473199 (p-value = 0.0292). Moreover, a significant interaction effect can be found for both SNPs with p-values = 0.0309 and 0.0179 respectively.

4.7. Mediation effects of GMD on the relationship between SNP and pain threshold

After conducting the two-way ANOVA analyses, we identified significant interaction effects among the genotypes of critical SNP, GMD, and pain sensitivity. Subsequently, through mediation analysis, we discovered a significant indirect effect of the SNP genotype on pain sensitivity measurements (pain threshold) mediated by GMD. This effect was observed in 5 SNP-ROI pairs (rs13294-insula, rs6432901-putamen, rs35473199-putamen, rs55682638-pallidum, rs55682638-putamen). The mediation model results are illustrated in Fig. 8.

We observed significant associations among the SNP genotype of the *ECM1* gene, pain threshold, and the GMD of the left insula. The mediation analysis demonstrated that the SNP rs13294 had an indirect effect (β = 0.02, SE = 0.05, CI = [0.009, 0.111], p = 0.013) but not a direct effect (β = 0.08, SE = 0.07, CI = [-0.023, 0.18], p = 0.12) on pain threshold through the GMD of the insula (Fig. 8(a) left).

Moreover, concerning the *SLC24A2* gene, the mediation analysis indicated that the SNP rs55682638 had an indirect effect (β = 0.03, SE = 0.01, CI = [0.027, 0.186], p = 0.003) but not a direct effect (β = 0.05, SE = 0.05, CI = [-0.05, 0.143], p = 0.087) on pain threshold mediated by the GMD of pallidum (Fig. 8(b) left) and

similarly for the same SNP through the GMD of putamen (indirect effect: β = 0.03, SE = 0.02, CI = [0.027, 0.191], p = 0.003; direct effect: β = 0.05, SE = 0.04, CI = [-0.047, 0.145], p = 0.057) (Fig. 8(b) right).

Regarding the *SCN9A* gene, the mediation analysis revealed that the SNP rs6432901 also had an indirect effect (β = 0.02, SE = 0.03, CI = [0.006, 0.116], p = 0.027) but not a direct effect (β = 0.04, SE = 0.04, CI = [-0.064, 0.139], p = 0.077) on pain threshold through the GMD of the putamen (Fig. 8(c) left). Similar findings were observed for the SNP rs35473199 (indirect effect: β = 0.02, SE = 0.02, CI = [0.005, 0.125], p = 0.023; direct effect: β = 0.04, SE = 0.06, CI = [-0.067, 0.136], p = 0.082) (Fig. 8(c) right). These findings demonstrate that the relationships between SNP genotype and pain sensitivity were partially mediated by GMD.

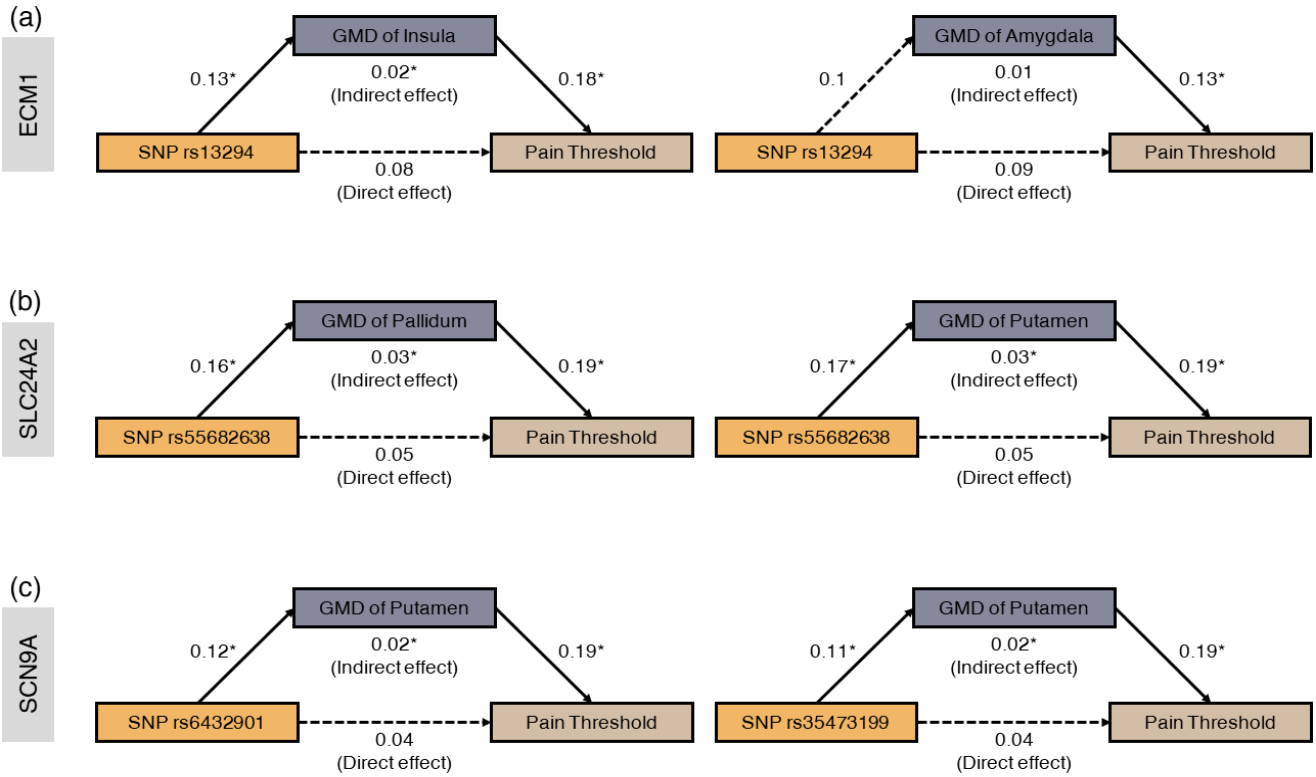

Fig. 8. Mediation analysis results for three genes *ECM1*, *SLC24A2* and *SCN9A*. The mediation analysis reveals that the relationship between SNP genotype and pain sensitivity (pain threshold) was indirected-only mediated by GMD of corresponding brain regions.

5. Discussion

This study aims to explore the genetic basis influencing variations in pain sensitivity among healthy individuals. To achieve this goal, we integrated the transcriptional gene expression data from AHBA together with self-collected brain structural MRI and genotype SNP dataset. We used multiple techniques to establish

the relationship among genomics, transcriptomics and imaging characteristics in terms of pain sensitivity individual differences. First, we identified 15 out of 45 brain regions' GMD significantly associated with pain sensitivity (Fig. 2). Second, we used a PLS regression method to analyze the correlation between gene expression and these regions within the left hemisphere of the brain (Fig. 3). After clustering and enrichment analysis of the key genes, we explored the genes' functions and their inter-gene relationships (Figs. 4 and 5). By integrating pain-related gene sets with our identified genes, we studied their gene expression located in different brain regions (Fig. 6). Third, we used our self-collected SNP data to explore the genetic basis of GMD and pain sensitivity using two-way ANOVA (Fig. 7). Fourth, we conducted a mediation analysis to discover the relationship among genetic factors and pain sensitivity, mediated by the brain imaging GMD features (Fig. 8). In summary, we suggested a comprehensive model that integrates genomic, transcriptomic and imaging features to describe the interplay among these factors in shaping pain sensitivity individual difference (Fig. 9).

5.1. Brain structural phenotype

Pain processing is associated with multiple brain regions, and there are observed differences among brain structural features between high and low pain sensitivity groups[5]. So, we first identify brain regions that exhibit significant associations with pain sensitivity. In Fig. 2, by comparing different pain sensitivity groups, 15 brain regions have been identified, such as amygdala, basal ganglia, insula, and thalamus. There are multiple identified regions belonging to traditional pain matrix or pain-related regions. These findings align with previous research emphasizing the crucial roles of the amygdala, insula, and basal ganglia in pain perception and regulation[50-52]. For instance, previous work highlighted the amygdala's role in connecting pain perception with emotional experience[50]. Moreover, evidence has been shown that the activation of basal ganglia structures during pain experiences, as well as their involvement in modulating pain perception[51]. These regions collaborate to receive, process and integrate pain signals, contributing to the subjective experience of pain.

5.2. Key genes enrichment analysis

After using PLS regression analysis and validation as shown in Fig. 3, we identified 27 key genes that are associated with the GMD differences between high and low pain sensitivity groups. As shown in Fig. 4 (a), the gene clustering illustrates that these genes can notably be divided into two main groups: high expression group (upper in Fig. 5 (a)) and low expression group (lower in Fig. 5 (a)) by examine the expression level of each gene and calculate the expression value correlation. Within each group, the correlations between genes are mostly positive while between groups correlations are mostly negative.

Consequently, the enrichment analysis in Fig. 5(b) shows that genes from group 1 are notably enriched in the main axon, VGSC and node of the Ranvier complex in the context of cellular components. Such results are reasonable. Pain sensitivity is predominantly mediated by specialized nerve fibers termed nociceptors. In pain perception, the main axons carry sensory information from peripheral nerves to the central nervous system, signaling potential skin-damaging stimuli. Additionally, these genes show enrichment in the VGSC complex, crucial in initiating and transmitting action potentials in excitable cells. VGSCs serve as primary targets for analgesic drugs, emphasizing their importance in pain modulation[53]. Mutations or alterations in VGSCs can profoundly affect pain sensitivity. Previous research highlighted the important role of Nav1.7 channels, both in their biophysical characteristics and unique expression patterns within sensory neuron subtypes in pain signaling[54,55]. Global Nav1.7-knockout mice exhibited a phenotype resembling human congenital indifference to pain, supporting the role of these genes in pain perception.

Moreover, the myelin sheath on axons, known as the node of Ranvier, significantly augments the speed of transmitting nerve impulses, including pain signals[56]. In a previous study on a rat model of diabetic neuropathy, researchers demonstrated that pain is partially influenced by reduced conduction failure along the primary axon of polymodal nociceptive C-fibres[57]. This discovery implies that individual variations in pain sensitivity could arise due to differences in transmitting pain information along the main axon, contributing to diverse perceptions and modulation of pain among individuals. Additionally, these genes in group 1 are

enriched in sodium channel activation. As previously mentioned, the specific sodium ion channel, NaV1.7, has emerged as a promising target for pain treatment due to its significance in individuals with rare genetic pain disorders. In certain families, a mutation in the gene encoding NaV1.7 permits excessive sodium entry into cells, resulting in severe chronic pain. Conversely, in other families, mutations that block NaV1.7 activity lead to a complete absence of pain[58]. These findings suggest that variations in sodium transport could potentially contribute to individual differences in pain sensitivity.

Regarding molecular function, these genes exhibit enrichment in various functional categories, including monovalent inorganic cation transmembrane transporter activity, VGSC activity, cation channel activity, sodium and metal ion transmembrane transporter activities, sodium and ion channel activities. The above enrichment functions are relevant to the activation and transport of ion channels. Where the cation channels, a class of ion channels permitting positively charged ions to traverse cell membranes, include voltage-gated sodium channels (Nav), voltage-gated calcium channels (Cav), and transient receptor potential (TRP) channels, all relevant to pain sensitivity. TRP channels, particularly TRP subfamily V member 1, are activated by various stimuli like heat, vanilloids, lipids, and protons/cations. Changes in their activity or expression significantly impact neuronal excitability and pain perception. Metal ions like zinc (Zn) and calcium (Ca) are known for modulating neuronal excitability and neurotransmission, including processes involved in pain perception[59,60].

The enrichment analysis shows that the other two genes (*SCN9A* and *S100A10*) are only enriched in terms of cellular components, with no significant enrichment in molecular functions. However, significant differences can be in their gene expression patterns. These findings indicate that these clustered gene groups might correspond to functional disparities linked to pain sensitivity, due to differences in their activities and processes.

These 27 genes were further integrated with the pain-related gene database so that 8 of them were further selected. Fig. 6 shows the gene expression level of these 8 genes from different brain regions. A polarized gene expression pattern can be observed for basal ganglia regions, including the amygdala, caudate, putamen,

pallidum and thalamus. The key role of the basal ganglia in pain perception and pain sensitivity has been studied in much previous research[61,62]. Findings from imaging studies show that basal ganglia are involved in pain modulation and can influence how pain signals are processed within the brain[52,63]. The enrollment of basal ganglia regions can potentially alter individual pain sensitivity. The results from our study further suggest that the distinct transcriptional alternations in basal ganglia regions may contribute to the differences observed in the imaging structural features between different pain sensitivity groups.

5.3. SNP statistical analysis

The statistical analysis shown in Figs. 7(a) and (b) demonstrate the association between the *ECM1* gene SNP rs13294 with the GMD of the amygdala and insula respectively. The two-way ANOVA analysis revealed the significant impact of both genotype and individual pain sensitivity on GMD. From both figures, the GMD between the AX and GG genotypes varies significantly between individuals with high and low pain sensitivity groups. Specifically, individuals with GG homozygotes of rs13294 exhibit lower GMD in the high pain sensitivity group and higher GMD in the low pain sensitivity group. Moreover, differences in the GMD of the amygdala and insula are observed between high and low pain sensitivity groups. These findings are promising, aligning well with previous pain perception and modulation studies[50,64,65]. For instance, in a study conducted by Lu and colleagues[66], they studied the anterior insular cortex's role in integrating various information sources to create pain awareness. Similarly, other research has shown the amygdala's function as a central hub integrating sensory input from nociceptive pathways with emotional and cognitive processes[50,65]. From genetic aspects, studies have linked the rs13294 variant in the *ECM1* gene to chronic musculoskeletal pain[67]. Meanwhile, evidence can be found that extracellular matrix (ECM) protein can regulate pain sensitivity in mice. Laminin β1 (LAMB1) is a key element of the extracellular matrix protein. In a previous study[68], the authors showed that in mice LAMB1 in the ACC was significantly downregulated upon peripheral neuropathy and knocking down of LAMB1 in the ACC can modify mice pain sensitivity. Increased pain sensitivity in

mice lacking LAMB1 can be found by comparing it to the control group, suggesting its role as a modulator in pain signaling pathways. Our findings further reveal the possible mechanism from genetic aspects and the brain neuroimaging feature changes.

The results presented in Figs. 7(c) and (d) reveal a significant influence on the GMD of the pallidum and putamen due to the interaction between genotype and pain sensitivity. Specifically, individuals with CC homozygotes of rs55682638 exhibit higher GMD levels in both high and low pain sensitivity groups. This implies that those carrying CC homozygotes in rs55682638 may display elevated GMD in the pallidum and putamen compared to others. Research on the *SLC24A2* gene indicates its potential role in pain modulation and neurotransmitter regulation[13,69]. For instance, Zhou et al. demonstrated that increased *SLC24A2* expression could decrease thermal pain sensitivity and the expression of inflammatory factors in rats with chronic constriction injury[69]. Genetic variations within *SLC24A2* may influence ion transport and calcium signaling, impacting sensory neuron function and pain perception. Moreover, the pallidum is part of the basal ganglia, and studies have suggested potential pain-processing pathways within this structure[52,61]. The findings establish a clear link between the SNP rs55682638, the pallidum and putamen concerning pain sensitivity, suggesting that *SLC24A2* and the basal ganglia might offer potential targets for pain modulation strategies.

The results from Figs. 7(e) and (f) indicate significant correlations between two SNPs and the GMD of the putamen. The two-way ANOVA analysis demonstrates an interaction effect between genotype and pain sensitivity concerning putamen GMD. Specifically, individuals with CC homozygotes for both rs6432901 and rs35473199 show lower GMD in the high pain sensitivity group and higher GMD in the low pain sensitivity group. Both two SNPs belong to the *SCN9A* gene, which has an important role in pain perception and chronic pain conditions[58,70]. *SCN9A* encodes the Nav1.7 sodium channel in pain signaling within the nervous system. Previous research indicated that specific genetic variants might change the pain sensitivity compared to individuals lacking these variants[70]. Moreover, the putamen, part of the pain matrix, is frequently activated during painful experiences. As a component of the striatum, it contains opioid receptors and neurons

responding to various pain stimuli levels. Previous research demonstrates that putamen lesions result in reduced pain sensitivity and a widespread decrease in pain-related brain activation, underlining the putamen's role in the sensory aspects of pain[64]. These findings suggest a connection between *SCN9A* gene SNPs and putamen GMD, potentially influencing alterations in pain perception. These findings not only deepen our understanding of the interplay among genetic factors, GMD and pain sensitivity but also provide promising targets for future therapeutic interventions in pain management, including pain relief and modulation.

5.4. Interpretation model

In Fig. 8, the mediation analysis revealed possible interactions among genotype, brain GMD features and individual pain sensitivity. Though the pain sensitivity individual difference has genetic factors, the mediation analysis implies that several important SNP genotypes regulated pain sensitivity (pain threshold) by modulating the multiple brain structural features. To summarize our results, we propose an interpretation model that integrates the roles of contribution from genetic factors and brain factors to pain sensitivity (Fig. 9). Individual pain sensitivity is influenced by two major factors: genetic factors and brain factors. From the imaging structural feature analysis, multiple brain regions' GMD is significantly associated with pain sensitivity. These regions include basal ganglia, insula and other traditional pain matrix areas. By combining AHBA transcriptional datasets and pain-related gene sets, we found two gene groups with polarized expression patterns in basal ganglia that are related to GMD and pain sensitivity. As an example, the *ECM1* gene and *SLC24A2* gene are low expressed in basal ganglia, but the *SCN9A* gene is highly expressed in these areas. Statistical analysis and mediation analysis using the SNP dataset indicate that genotype regulates pain sensitivity by modulating the intrinsic brain structural GMD features. For the low-expressed genes, they mainly regulate the insula, pallidum and putamen to influence individual pain sensitivity. On the contrary, highly expressed genes such as *SCN9A* mainly contribute to one's putamen GMD to regulate pain sensitivity. Therefore, our model suggests that two gene groups with different expression patterns influence pain

sensitivity in two ways by regulating different regions in basal ganglia regions. Future research may further evaluate the underlying mechanisms contributing to these two ways.

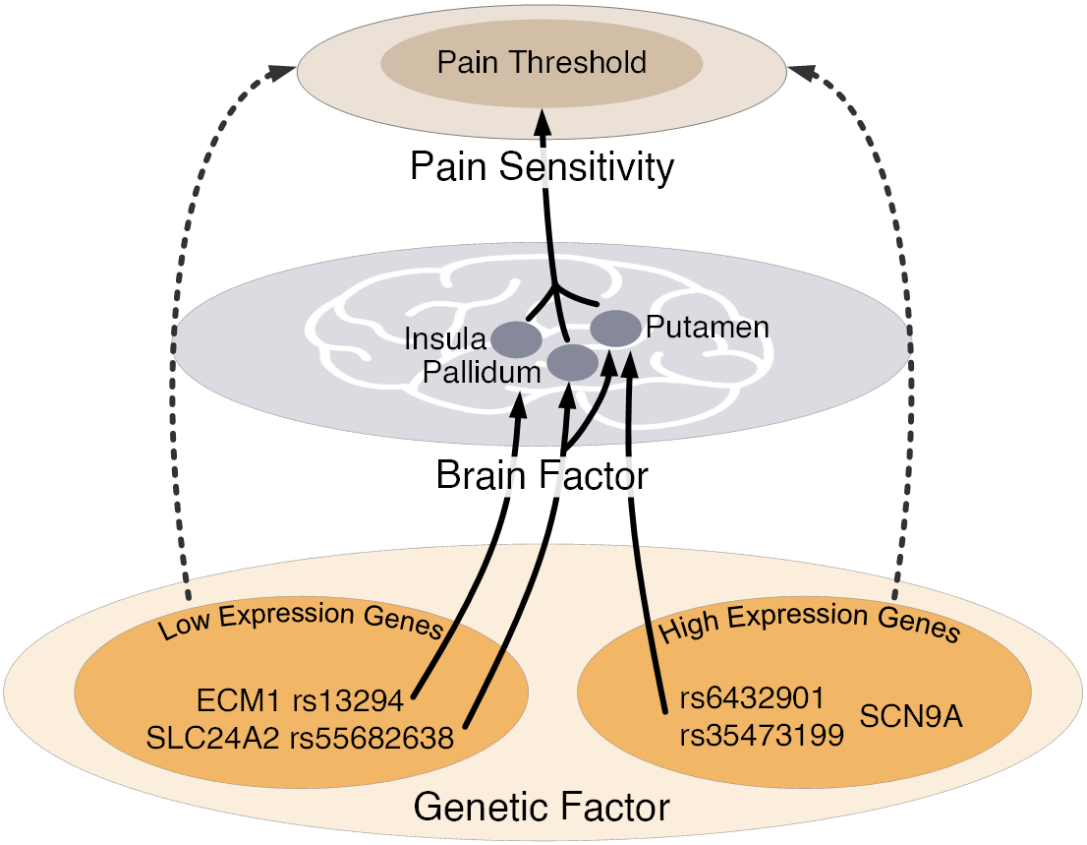

Fig. 9. Interpretation model to explain the mechanism underlying the interplay of genetic factors, brain factors and pain sensitivity. From the genetic factor perspective, there exist two gene groups with different expression patterns. Meanwhile, the genotypes of these genes influence the pain sensitivity in an indirect-only way via GMD of brain regions from basal ganglia.

5.5. Limitation

Several limitations should be considered when interpreting the results of this study. First, the sample size was relatively small, and future studies with larger sample sizes and external validation are needed to further verify GMD differences across individuals. Second, the gene expression data and neuroimaging data were collected from different participants. Though the genotype SNP and the brain neuroimaging data are collected from the same population, the transcriptional gene expression data are not consistent. The gene expression profiles in AHBA are influenced by factors such as age, sex, and ethnicity, which may introduce bias and limit the generalizability of the findings to diverse populations. Therefore, multi-omics datasets from the same subjects are more promising to provide reliable results. Third, our mediation analysis model shed light on the

relationship between SNP genotypes of pain-related genes and pain threshold, revealing potential causality. However, to determine the causal relationship, further validation through animal experiments is required.

6. Conclusion

In conclusion, we collected pain threshold, brain MRI data and genotype SNP data from 432 subjects. We first used MRI data to analyze the brain GMD features that related to pain sensitivity. Next, we used PLS regression by integrating the AHBA transcriptional gene expression dataset and GMD features difference between high and low pain sensitivity groups to locate key genes that may regulate related brain regions. Enrichment and clustering analysis were used to reveal the molecular functions and cellular components of two key gene sets with different expression patterns. Moreover, a series of statistical analyses using genotype SNP data demonstrated potential key loci that may related to brain structural GMD features and pain sensitivity. Mediation analysis revealed the contribution of genetic factors and brain factors to pain sensitivity. In the end, we proposed an interpretation model to summarize our results. Our study provides valuable insights into the transcriptomic and genomic signature of GMD patterns that may shape individual pain sensitivity.

Conflict of interest statement

The authors have no conflict of interest to declare.

Author contributions

Y. P. contributed ceonceptualization, data curation, formal analysis, investigation, methodology, validation, visualization and writing. Z. Z. contributed ceonceptualization, data recources, data curation, investigation, methodology, validation and writing. X. H. contributed investigation, methodology, validation and writing. G. H and Z. L. contributed to investigation, methodology, validation. L. Z. contributed ceonceptualization, formal analysis, funding acquisition, visualization, writing, supervision and project administration.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (62201356, 32361143787), in part by the Guangdong Basic and Applied Basic Research Foundation (2021A1515110694), in part by the Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions (2022SHIBS0003), and in part by Medicine Plus Program of Shenzhen University (2024YG021).

References

1 Grant, J. A., Courtemanche, J., Duerden, E. G., Duncan, G. H. & Rainville, P. Cortical thickness and pain sensitivity in zen meditators. *Emotion* **10**, 43 (2010).

2 Edwards, R. R., Dworkin, R. H., Sullivan, M. D., Turk, D. C. & Wasan, A. D. The role of psychosocial processes in the development and maintenance of chronic pain. *J. Pain* **17**, T70-T92 (2016).

3 Wiech, K. Deconstructing the sensation of pain: The influence of cognitive processes on pain perception. *Science* **354**, 584-587 (2016).

4 Fillingim, R. B. Individual differences in pain: understanding the mosaic that makes pain personal. *Pain* **158**, S11-S18 (2017).

5 Emerson, N. M. *et al.* Pain sensitivity is inversely related to regional grey matter density in the brain. *Pain* **155**, 566-573 (2014).

6 Denk, F., McMahon, S. B. & Tracey, I. Pain vulnerability: a neurobiological perspective. *Nat. Neurosci.* **17**, 192-200 (2014).

7 Starkweather, A. R. *et al.* Acute low back pain: Differential somatosensory function and gene expression compared with healthy no-pain controls. *Clin. J. Pain.* **32**, 933-939 (2016).

8 Hedges, E. P. *et al.* Reliability of structural MRI measurements: The effects of scan session, head tilt, inter-scan interval, acquisition sequence, FreeSurfer version and processing stream. *Neuroimage* **246**, 118751 (2022).

9 Teutsch, S., Herken, W., Bingel, U., Schoell, E. & May, A. Changes in brain gray matter due to repetitive painful stimulation. *Neuroimage* **42**, 845-849 (2008).

10 Erpelding, N., Moayedi, M. & Davis, K. D. Cortical thickness correlates of pain and temperature sensitivity. *Pain* **153**, 1602-1609 (2012).

11 Ruscheweyh, R., Wersching, H., Kugel, H., Sundermann, B. & Teuber, A. Gray matter correlates of pressure pain thresholds and self-rated pain sensitivity: a voxel-based morphometry study. *Pain* **159**, 1359-1365 (2018).

12 Grinsvall, C. *et al.* Association between pain sensitivity and gray matter properties in the sensorimotor network in women with irritable bowel syndrome. *Neurogastroenterol. Motil.* **33**, e14027 (2021).

13 Zhang, X., Zhu, L., Wang, X., Xia, L. & Zhang, Y. Advances in the role and mechanism of miRNA in inflammatory pain. *Biomed. Pharmacother.* **161**, 114463 (2023).

14 Mogil, J. S. *et al.* Heritability of nociception II.Types' of nociception revealed by genetic correlation analysis. *Pain* **80**, 83-93 (1999).

15 Lariviere, W. R. *et al.* Heritability of nociception. III. Genetic relationships among commonly used assays of nociception and hypersensitivity. *Pain* **97**, 75-86 (2002).

16 Foulkes, T. & Wood, J. N. Pain genes. *PLoS Genet.* **4**, e1000086 (2008).

17 Norbury, T. A., MacGregor, A. J., Urwin, J., Spector, T. D. & McMahon, S. B. Heritability of responses to painful stimuli in women: a classical twin study. *Brain* **130**, 3041-3049 (2007).

18 Fillingim, R. B. *et al.* The A118G single nucleotide polymorphism of the μ-opioid receptor gene (OPRM1) is associated with pressure pain sensitivity in humans. *J. Pain* **6**, 159-167 (2005).

19 Lötsch, J. & Geisslinger, G. Relevance of frequent μ-opioid receptor polymorphisms for opioid activity in healthy volunteers. *Pharmacogenomics. J.* **6**, 200-210 (2006).

20 Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391-399

(2012).
21 Pool, E.-M., Rehme, A. K., Eickhoff, S. B., Fink, G. R. & Grefkes, C. Functional resting-state connectivity of the human motor network: differences between right-and left-handers. *Neuroimage* **109**, 298-306 (2015).
22 Margiotoudi, K. *et al.* Handedness in monkeys reflects hemispheric specialization within the central sulcus. An in vivo MRI study in right-and left-handed olive baboons. *Cortex* **118**, 203-211 (2019).
23 Wiberg, A. *et al.* Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain* **142**, 2938-2947 (2019).
24 Sternbach, R. A. Psychophysiology of pain. *Int. J. Psychiat. Med.* **6**, 63-73 (1975).
25 Zhang, H. et al. Multimodal covarying brain patterns mediate genetic and psychological contributions to individual differences in pain sensitivity. Pain **165**, 1074-1085 (2024).
26 Pertovaara, A., Kauppila, T. & Hämäläinen, M. M. Influence of skin temperature on heat pain threshold in humans. *Exp. Brain. Res.* **107**, 497-503 (1996).
27 Wu, G., Campbell, J. N. & Meyer, R. A. Effects of baseline skin temperature on pain ratings to suprathreshold temperature-controlled stimuli. *Pain* **90**, 151-156 (2001).
28 Good, C. D. *et al.* A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* **14**, 21-36 (2001).
29 Ashburner, J. A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95-113 (2007).
30 Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273-289 (2002).
31 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-575 (2007).
32 Freedman, M. L. *et al.* Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388-393 (2004).
33 Markello, R. D. *et al.* Standardizing workflows in imaging transcriptomics with the abagen toolbox. *Elife* **10**, e72129 (2021).
34 Arnatkevičiūtė, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage* **189**, 353-367 (2019).
35 Hansen, J. Y. *et al.* Mapping gene transcription and neurocognition across human neocortex. *Nat. Hum. Behav.* **5**, 1240-1250 (2021).
36 Krishnan, A., Williams, L. J., McIntosh, A. R. & Abdi, H. Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage* **56**, 455-475 (2011).
37 McIntosh, A. R. & Lobaugh, N. J. Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage* **23**, S250-S263 (2004).
38 Liu, J., Xia, M., Wang, X., Liao, X. & He, Y. The spatial organization of the chronnectome associates with cortical hierarchy and transcriptional profiles in the human brain. *Neuroimage* **222**, 117296 (2020).
39 Chong, I.-G. & Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemometrics Intellig. Lab. Syst.* **78**, 103-112 (2005).
40 Morgan, S. E. *et al.* Cortical patterning of abnormal morphometric similarity in psychosis is associated with brain expression of schizophrenia-related genes. *Proc. Natl Acad Sci.USA* **116**, 9604-9609 (2019).
41 Li, J. *et al.* Cortical structural differences in major depressive disorder correlate with cell type-specific transcriptional signatures. *Nat. Commun.* **12**, 1647 (2021).
42 Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53-65 (1987).
43 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25-29 (2000).
44 Meloto, C. B. *et al.* Human pain genetics database: a resource dedicated to human pain genetics research. *Pain* **159**, 749-763 (2018).
45 LaCroix-Fralish, M. L., Ledoux, J. B. & Mogil, J. S. ThePain Genes Database: An interactive web browser of pain-related transgenic knockout studies. *Pain* **131**, 3e1-3e4 (2007).
46 Liu, F. *et al.* Variation in the CACNB2 gene is associated with functional connectivity of the Hippocampus in bipolar disorder. *BMC psychiatry* **19**, 1-7 (2019).
47 MacKinnon, D. P., Fairchild, A. J. & Fritz, M. S. Mediation analysis. *Annu. Rev. Psychol.* **58**, 593-614 (2007).

48 Mallinckrodt, B., Abraham, W. T., Wei, M. & Russell, D. W. Advances in testing the statistical significance of mediation effects. *J. Couns. Psychol.* **53**, 372 (2006).

49 Preacher, K. J. & Hayes, A. F. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* **40**, 879-891 (2008).

50 Veinante, P., Yalcin, I. & Barrot, M. The amygdala between sensation and affect: a role in pain. *J Mol Psychiatry.* **1**, 1-14 (2013).

51 De Ridder, D., Adhia, D. & Vanneste, S. The anatomy of pain and suffering in the brain and its clinical implications. *Neurosci. Biobehav. Rev.* **130**, 125-146 (2021).

52 Borsook, D., Upadhyay, J., Chudler, E. H. & Becerra, L. A key role of the basal ganglia in pain and analgesia-insights gained through human functional imaging. *Mol. Pain* **6**, 1744-8069-1746-1727 (2010).

53 Cummins, T. R., Dib-Hajj, S. D., Black, J. A. & Waxman, S. G. Sodium channels and the molecular pathophysiology of pain. *Prog. Brain. Res.* **129**, 3-19 (2000).

54 Chang, W. *et al.* Expression and role of voltage-gated sodium channels in human dorsal root ganglion neurons with special focus on Nav1. 7, species differences, and regulation by paclitaxel. *Neurosci. Bull.* **34**, 4-12 (2018).

55 Bennett, D. L., Clark, A. J., Huang, J., Waxman, S. G. & Dib-Hajj, S. D. The role of voltage-gated sodium channels in pain signaling. *Physiol. Rev.* **99**, 1079-1151 (2019).

56 Yam, M. F. *et al.* General pathways of pain sensation and the major neurotransmitters involved in pain regulation. *Int. J. Mol. Sci.* **19**, 2164 (2018).

57 Sun, W. *et al.* Reduced conduction failure of the main axon of polymodal nociceptive C-fibres contributes to painful diabetic neuropathy in rats. *Brain* **135**, 359-375 (2012).

58 Gomez, K. *et al.* Identification and targeting of a unique NaV1. 7 domain driving chronic pain. *Proc. Natl Acad Sci.USA* **120**, e2217800120 (2023).

59 Bourinet, E. *et al.* Calcium-permeable ion channels in pain signaling. *Physiol. Rev.* **94**, 81-140 (2014).

60 Nozaki, C. *et al.* Zinc alleviates pain through high-affinity binding to the NMDA receptor NR2A subunit. *Nat. Neurosci.* **14**, 1017-1022 (2011).

61 Chudler, E. H. & Dong, W. K. The role of the basal ganglia in nociception and pain. *Pain* **60**, 3-38 (1995).

62 Tan, Y. *et al.* Alteration of basal ganglia and right frontoparietal network in early drug-naive Parkinson's disease during heat pain stimuli and resting state. *Front. Hum. Neurosci.* **9**, 467 (2015).

63 Knudsen, L. *et al.* Review of neuroimaging studies related to pain modulation. *Scand. J. Pain.* **2**, 108-120 (2011).

64 Starr, C. J. *et al.* Roles of the insular cortex in the modulation of pain: insights from brain lesions. *J. Neurosci.* **29**, 2684-2694 (2009).

65 Rouwette, T., Vanelderen, P., Roubos, E. W., Kozicz, T. & Vissers, K. The amygdala, a relay station for switching on and off pain. *Eur. J. Pain.* **16**, 782-792 (2012).

66 Lu, C. *et al.* Insular cortex is critical for the perception, modulation, and chronification of pain. *Neurosci. Bull.* **32**, 191-201 (2016).

67 Tsepilov, Y. A. *et al.* Analysis of genetically independent phenotypes identifies shared genetic factors associated with chronic musculoskeletal pain conditions. *Commun. Biol.* **3**, 329 (2020).

68 Li, Z.-Z. *et al.* Extracellular matrix protein laminin β1 regulates pain sensitivity and anxiodepression-like behaviors in mice. *J. Clin. Invest.* **131** (2021).

69 Zhou, X. G., He, H. & Wang, P. J. A critical role for miR-135a-5p-mediated regulation of SLC24A2 in neuropathic pain. *Mol. Med. Report.* **22**, 2115-2122 (2020).

70 Reimann, F. *et al.* Pain perception is altered by a nucleotide polymorphism in SCN9A. *Proc. Natl Acad Sci.USA* **107**, 5148-5153 (2010).