Uncovering the epigenetic regulatory clues of *PRRT1* in Alzheimer's disease: a strategy integrating multi-omics analysis with explainable machine learning

Fang Wang¹, Ying Liang^{3*} and Qin-Wen Wang^{2*}

Abstract

Background Alzheimer's disease (AD) is a complex neurodegenerative disorder with a largely unexplored epigenetic landscape.

Objective This study employs an innovative approach that integrates multi-omics analysis and explainable machine learning to explore the epigenetic regulatory mechanisms underlying the epigenetic signature of *PRRT1* implicated in AD.

Methods Through comprehensive DNA methylation and transcriptomic profiling, we identified distinct epigenetic signatures associated with gene *PRRT1* expression in AD patient samples compared to healthy controls. Utilizing interpretable machine learning models and ELMAR analysis, we dissected the complex relationships between these epigenetic signatures and gene expression patterns, revealing novel regulatory elements and pathways. Finally, the epigenetic mechanisms of these genes were investigated experimentally.

Results This study identified ten epigenetic signatures, constructed an interpretable AD diagnostic model, and utilized various bioinformatics methods to create an epigenomic map. Subsequently, the ELMAR R package was used to integrate multi-omics data and identify the upstream transcription factor MAZ for *PRRT1*. Finally, experiments confirmed the interaction between MAZ and *PRRT1*, which mediated apoptosis and autophagy in AD.

Conclusion This study adopts a strategy that integrates bioinformatics analysis with molecular experiments, providing new insights into the epigenetic regulatory mechanisms of *PRRT1* in AD and demonstrating the importance of explainable machine learning in elucidating complex disease mechanisms.

Keywords Alzheimer's disease, Multi-omics analysis, Interpretable machine learning, Biomarker, Epigenetic

*Correspondence: Ying Liang nbliangying@163.com Qin-Wen Wang 13486485004@163.com ¹Department of Pharmacy, Zhejiang Pharmaceutical University, Ningbo, China

²Zhejiang Provincial Key Laboratory of Pathophysiology, School of Medicine, Ningbo University, 818 Fenghua Road, Jiangbei District, Ningbo, China ³Ningbo Maritime Silk Road Institute, No.8, South Qianhu Road, Ningbo, China

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit [http://creati](http://creativecommons.org/licenses/by-nc-nd/4.0/) [vecommons.org/licenses/by-nc-nd/4.0/.](http://creativecommons.org/licenses/by-nc-nd/4.0/)

Introduction

Alzheimer's disease (AD) is the most common neurodegenerative disorder [[1\]](#page-20-0). Although the etiology of AD remains unclear, it is currently believed that the onset of AD involves complex interactions between genetic factors, environmental influences, and aging [[2\]](#page-20-1). Many studies have found that epigenetic regulation plays a role in the pathogenesis of AD $[3-5]$ $[3-5]$. The expression levels of many known key genes implicated in AD, such as *MAPT*, *BACE1, APOE,* are subject to epigenetic regulation [\[3](#page-20-2), [6](#page-20-4), [7\]](#page-20-5).

DNA methylation, a prevalent form of epigenetic modification, is significantly implicated in AD. The modulation of DNA methylation may offer potential diagnostic and therapeutic approaches for AD [[4\]](#page-20-6). Recent research indicates that methylation can act as a key regulator of tau protein aggregation in AD, and adjusting the DNA methylation status may alleviate the pathological features of AD [\[8](#page-20-7)]. Given the complexity of AD's etiology and the high heterogeneity in biological changes and disease progression, integrating multi-omics analyses such as genomics, epigenomics, transcriptomics, etc., aids in identifying biomarkers for AD and exploring their upstream regulatory mechanisms and downstream molecular effects [\[9\]](#page-20-8).

This study integrates transcriptomic and epigenomic data to investigate potential epigenetic drivers of AD. By performing multi-omics analysis on datasets from the same set of tissue samples through bioinformatics, we employed interpretable machine learning algorithms to construct an AD diagnostic model and identify potential AD epigenetic signatures (methylation-related differentially co-expressed genes).

For the selected epigenetic signatures, this study used the ELMER R package to integrate gene expression and methylation data, analyzing the potential motifs upstream regulated by DNA methylation and their corresponding transcription factors, which were verified through luciferase reporter assay and Chromatin Immunoprecipitation (ChIP) experiment. Subsequently, further experiments explored the roles of epigenetic signatures in AD and the regulatory mechanisms of transcription factors on them.

Methods

Feature selection

Data preprocessing

The GSE109627 dataset was downloaded from GEO database, which includes the methylation 450 K dataset of the middle temporal gyrus of the brain from subjects and their clinical information, comprising 45 samples from AD patients and 35 normal samples [\[10](#page-20-9)]. The 5-methylcytosine (5mC) methylation expression profile within the GSE109627 methylation 450 K dataset was extracted using the ChAMP R package $[11]$ $[11]$ (R version 4.0.2). For the beta matrix of the GSE109627 methylation dataset, filtering and normalization were performed using the ChAMP R package. The GSE109887 dataset [[10](#page-20-9)], encompassing mRNA expression profiles and clinical group information, was downloaded; this dataset includes temporal middle gyrus samples from 46 AD patients and 32 normal subjects. The study workflow is shown in Fig. [1](#page-2-0).

Identification and selection of differentially methylated positions in the GSE109627 dataset

Differentially methylated positions (dmCpGs) were analyzed and selected using the ChAMP R package based on the criteria $|\Delta \beta|$ > 0.02 and *P*<0.005.

Negative correlation analysis between differentially methylated positions and genes

In this study, the hm450.manifest.hg19 methylation 450 K chip annotation data provided by the ChAMP R package was used, followed by mapping differentially methylated positions to related genes using the tidyverse R package. These were then matched with the corresponding transcribed genes in the GSE109887 transcriptomic dataset, forming pairs of differentially methylated positions and transcribed genes. Spearman correlation analysis was conducted between the β values of the methylation positions and the transcriptional expression levels of the genes (*P*<0.05, *r* < -0.7).

Selection of differentially expressed genes and co-expressed genes

The Limma R package $[12]$ $[12]$ and the WGCNA R package [[13\]](#page-20-12) were employed to identify differentially expressed genes and co-expressed genes, respectively, from the GSE109887 dataset.

Identification of AD epigenetic signatures through joint analysis of epigenome and transcriptome

For each probe-gene pair, this study categorized genes into two types: hypermethylation with low-expression and hypomethylation with high-expression:

Up-regulated differentially methylated positions in the GSE109627 methylation dataset (selected based on criteria $Δβ > 0.02$ and $P < 0.005$) were matched with genes in the GSE109887 transcriptomic dataset that were downregulated (adj.P.Val<0.05, logFC < -0.2), AD co-expressed genes from the WGCNA purple module, and negatively correlated methylation-regulated genes. The intersection of these genes was used to identify potential epigenetic signatures, visualized using the VennDiagram R package [[14\]](#page-20-13).

Down-regulated differentially methylated positions in the GSE109627 methylation dataset (selected based on criteria $Δβ < -0.02$ and $P < 0.005$) were matched with

Fig. 1 Study flowchart

up-regulated genes in the GSE109887 transcriptomic dataset (adj.P.Val<0.05, logFC>0.2), AD co-expressed genes from the WGCNA purple module, and negatively correlated methylation-regulated genes. The intersection of these genes was used to identify potential epigenetic signatures, visualized using the VennDiagram R package [[14\]](#page-20-13).

Machine learning algorithm for selection of epigenetic signatures

Python (v3.6.2) with the pandas library $[15]$ $[15]$ was used to read and merge clinical traits with methylation-related gene data from the GSE109887 dataset, followed by normalization. The LightGBM algorithm was applied with 10-fold cross-validation to extract feature genes (epigenetic signatures) along with clinical traits for subsequent development of an AD diagnostic model.

Modeling and evaluation

Construction and evaluation of AD diagnostic models

Using Python (v3.6.2) with the Scikit-learn library, various machine learning algorithms were implemented to build AD diagnostic models. These models were then evaluated through 5-fold cross-validation based on the ROC-AUC score to select the optimal algorithm for constructing the AD diagnostic model.

The data was split into a 70% training set and a 30% test set (random_state=3) using the Scikit-learn library to perform training and test set difference verification. Subsequently, the AD diagnostic model was built based on the optimal model parameters and evaluated using the ROC. Finally, the generalizability of the model was verified using the external validation set from the GSE132903 dataset, which includes expression profiles of 10 epigenetic signatures along with clinical data such as gender and age.

ROC curve analysis

Univariate ROC curve analysis was performed using pROC R package to assess the diagnostic performance of each epigenetic signature in the GSE109887 dataset, and its generalizability was also assessed using the external validation set (GSE132903).

Interpretation and visualization of the AD diagnostic model

The AD diagnostic model was interpreted and visualized using the SHAP library [[16](#page-20-15)].

Gene functional analysis

AlzData validation

AlzData database [\[17](#page-20-16)] was used to verify the expression differences of epigenetic signatures between AD and normal tissues.

GO, KEGG analysis and GSEA

In this study, GO and KEGG analyses were conducted on the identified epigenetic signatures using the clusterProfiler R package [\[18](#page-20-17)] (*P*<0.05). Additionally, AD samples were categorized from the GSE109887 dataset into "highexpression" and "low-expression" groups based on the median expression level of epigenetic signatures. Then GSEA [\[19,](#page-20-18) [20](#page-20-19)] was performed using the GSEA database "c2.cp.kegg.v7.4.symbols.gmt" (*P*<0.05).

The immune function analysis

The ssGSEA algorithm provided in the GSVA R package [[20,](#page-20-19) [21\]](#page-20-20) was used to comprehensively evaluate the immunological characteristics of each sample in the GSE109887 dataset.

Immune checkpoint analysis

Based on the median expression levels of epigenetic signatures, the AD samples within the GSE109887 dataset were divided into two groups: "high-expression group" and "low-expression group." Subsequently, the ggpubr R package was employed to conduct an analysis and visualization of the expression profiles of genes associated with immune checkpoints.

NetworkAnalyst analysis

The transcription factor (TF)-gene interaction analysis was performed by integrating common Alzheimer's disease risk genes [\[22,](#page-20-21) [23](#page-20-22)] (*APP*, *PSEN1*, *PSEN2*, *BDNF*, *SORL1*, *ABCA7*, *APOE*, *TREM2*, *PLCG2*) with epigenetic signatures using NetworkAnalyst [\[24](#page-20-23)]. Additionally, NetworkAnalyst was employed to develop a TF-miRNA coregulatory network and to pinpoint compounds that interact with the epigenetic signatures.

ELMER analysis

The ELMER R package [\[25](#page-20-24)] was utilized to analyze gene methylation (GSE109627 methylation dataset) and gene expression levels (GSE109887 transcriptomic dataset) based on the same sample cohort. By comparing the methylation levels of all enhancers and promoter regions in AD and normal samples, methylated sites were selected based on criteria *P*<0.005 and Δβ>0.02 between the two groups, specifically those that were hypermethylated in AD samples. Subsequently, an analysis was conducted to determine whether there was a negative correlation between the expression levels of target genes and these hypermethylated sites. Following this, the base sequences within the 250 bp region upstream and downstream of these sites were extracted to identify motifs enriched in epigenetic signatures, namely transcription factor binding motifs (TFBMs). Ultimately, by utilizing a database of transcription factors (TFs) binding to motifs, potential upstream driving factors were predicted.

Single-cell transcriptome analysis

SCAD-Brain [[26\]](#page-20-25) performs trajectory analysis to model the dynamic processes of cells based on single cell transcriptomics data which contains prefrotnal cortex of human (GSE157827) [[27\]](#page-20-26).

Single-cell analysis of the GSE157985 dataset [[28](#page-20-27)] which contains hippocampal tissues from wildtype mice and J20 AD mouse model, was performed using the SCAD-Brain database [\[26\]](#page-20-25).

Experimental validation Cell culture

SH-SY5Y cells were cultured in DMEM/f12 medium containing 10% fetal bovine serum and 1% Penicillin-Streptomycin, and HEK293T cells were cultured in DMEM medium with 10% fetal bovine serum. Both types of cells were incubated in a CO2 (5%) cell culture incubator at 37 °C.

Aβ oligomers preparation

Dissolve 1 mg of $A\beta_{1-42}$ (GL Biochem, Shanghai, China) in 400 µl of hexafluoroisopropanol (HFIP), and incubate at room temperature for 20 min. Then, transfer 100 µl of the solution into a clean eppendorf tube and add 900 µl of sterile deionized water, followed by incubation at room temperature for another 20 min. Next, centrifuge at 12,000 rpm for 15 min, and transfer the supernatant to another clean eppendorf tube. Subsequently, gently blow off the residual HFIP with high purity nitrogen for 10–15 min using a nitrogen blower. Stop when approximately $750 \mu l$ of liquid remains in the eppendorf tube. Finally, incubate at 500 rpm for 48 h using a magnetic stirrer before storing in a refrigerator at 4 °C.

Western blotting

The proteins of each group of cells were extracted (Beyotime, P0033, China). The extracted membrane proteins were separated via SDS-PAGE electrophoresis and transferred onto a PVDF membrane. The membranes were then incubated with primary antibodies including anti-PRRT1 (rabbit, 1:1000, 17261-1-AP, Proteintech) and internal reference antibodies Na, K-ATPase (rabbit, 1:5000, #3010, CST), anti-MAZ (rabbit, 1:1000, 21068-1- AP, Proteintech), anti-LC3 (rabbit, 1:1000, 14600-1-AP, Proteintech), anti-phosphorylated-tau (rabbit, 1:1000, 28866-1-AP, Proteintech), anti-tau (rabbit, 1:2000,10274- 1-AP, Proteintech), anti-actin (rabbit, 1:10000, TDY051, TDY Biotech), followed by incubation with secondary antibodies. Western blots were detected using an X-ray film, after which data analysis was conducted on the results.

Luciferase reporter assay

Prior to transfection, HEK293T cells were seeded at a density of 1×10^5 cells per well in a 24-well plate containing 500 µl of antibiotic-free culture medium. The cells were transfected when they reached a confluency of 30–50%.

The Firefly luciferase reporters, which include the *PRRT1* promoter constructs, and the pRL-TK Renilla luciferase reporter, were introduced into HEK293T cells via a transfection reagent (Lipofectamine 2000, Invitrogen). The expression vectors of transcription factors MAZ, VEZF1, TFAP2E, HMGA1 or control empty vector were transfected respectively. Dual Luciferase Reporter Gene Assay Kit (Beyotime, RG027, China) was used and the luciferase activity in cell lysates was detected. Relative luciferase activity was calculated by normalizing firefly luciferase activity to the luciferase activity of Renilla luciferase. ANOVA was employed for statistical analysis $(n=3, \dots, P<0.0001).$

ChIP-qPCR

HEK293T cells were cross-linked using formaldehyde to form a stable complex of proteins with DNA. Cell lysis was then performed to isolate the nuclei. The chromatin was cleaved into appropriately sized fragments using ultrasonic fragmentation. The cleaved chromatin was incubated with the specific antibody to bind the target protein to the antibody of MAZ (Thermo Fisher, A301-652A-T) or IgG (ABclonal, AC005). Protein A/G magnetic beads were then added to precipitate the protein-DNA complex by magnetic force. The precipitated genomic DNA was subjected to amplification through real-time PCR using the following specified primers: sense: 5'-CCCGACAGAGACTAACGTGAG-3'; antisense: 5'-GAGAATGGACCCCAAGGTCTT-3'.

Small interfering RNA knockdown and plasmid transfection

In order to silence PRRT1, SH-SY5Y cells were introduced to either siRNA or scrambled siRNA (ELK Biotechnology, China) utilizing Lipofectamine 2000.

In addition, in order to overexpression PRRT1 and MAZ, SH-SY5Y cells were transfected with plasmid or vectors. A $β_{1-42}$ was used 24 h after cell transfection. Cells were utilized in further experiments 48 h post-transfection to measure cell viability, apoptosis, and protein levels.

CCK8 assay

SHSY5Y cells in logarithmic growth phase were digested with trypsin and prepared into a cell suspension at a concentration of 1×10^5 cells/ml. The cells were seeded in 96-well plates at 10,000 cells/well, and 100 µl serum-free medium was added to each well. The cells were incubated

at 37 °C for 24 h to adhere to the wall. The medium of the control group was subsequently replaced with medium containing solvent, and the medium of 1/2 cell samples of the other groups was replaced with medium containing Aβ. After reaching the set time of different treatment in each group, cell viability was assessed by adding CCK8 reagents (Biosharp, BS350B), followed by a 2-hour incubation, and subsequently, absorbance at 490 nm was gauged using a 96-well plate reader.

Flow cytometric analysis

The cell apoptosis rate of each group was detected using the Annexin V-FITC Apoptosis Detection Kit (Sungene Biotech, AO2001-02P-G), following the guidelines provided by the manufacturer. The treated cells were immediately analyzed by flow cytometry on the BD FAC-SCalibur (Becton, Dickinson and Company).

Results

Feature selection

Identification and selection of differentially methylated positions

In the GSE109627 methylation dataset, 2220 differentially methylated positions were identified (Supplementary Fig. 1), with 2188 being upregulated and 32 downregulated.

Negative correlation analysis between differentially methylated positions and genes

A total of 79 negatively correlated differentially methylated position-transcript expression gene pairs were selected.

Selection of differentially expressed genes and co-expressed genes

Differentially expressed genes were selected based on $|logFC| > 0.2$ and adj.P.Val<0.05 criteria; within the GSE109887 dataset, there were 4151 differentially expressed genes, with 2037 upregulated and 2114 downregulated (Supplementary Fig. 2). WGCNA results indicated that the purple module in GSE109887 had the smallest p-value $(p=2e-06)$ and the highest correlation with AD (correlation coefficient $r = -0.51$). Therefore, 3119 co-expressed genes from the purple module were selected for subsequent analysis (Supplementary Fig. 3).

Identification of AD epigenetic signatures through joint analysis of epigenome and transcriptome

The study identified 10 epigenetic signatures among hypermethylated low-expression genes: *BAIAP2*, *SEMA4F*, *PRRT1*, *GNB1*, *ENC1*, *R3HDM1*, *CDK5*, *RIMS3*, *SRRM4*, *PRDM8* (Fig. [2](#page-6-0)A); no epigenetic signatures were found among hypomethylated high-expression genes (Fig. [2](#page-6-0)B).

Machine learning algorithm for selection of epigenetic signatures

Using the LightGBM algorithm (threshold=0.9999) with 10-fold cross-validation, importance scores for each feature were obtained. Based on these results, all epigenetic signatures were retained for the subsequent construction of an AD diagnostic model (Fig. [2C](#page-6-0) and D, Supplementary Table 1).

Modeling and evaluation

Construction and evaluation of an AD diagnostic model

Using Python (v3.6.2) with the Scikit-learn library, various machine learning algorithms were employed to construct an AD diagnostic model, which was then validated through 5-fold cross-validation to determine the model's ROC-AUC (Supplementary Tables 2–3, Fig. [3A](#page-7-0)). Among these, the model built using the random forest algorithm achieved the highest average ROC-AUC of 0.801. Consequently, the random forest algorithm was chosen for the development of the AD diagnostic model. Validation results for differences between the training and test sets confirmed that there were no statistical differences between the data of the split training and test sets (Supplementary Table 4). Ultimately, the AD diagnostic model constructed using the random forest algorithm showed good diagnostic performance on the test set with a ROC-AUC=0.829. The model also demonstrated good generalization ability on the external validation set GSE132903 with a ROC-AUC=0.708 (Supplementary Table 5).

Sample size and power analysis are key steps in research design, crucial for ensuring the validity and reliability of study results. A larger sample size can provide more stable and reliable outcomes. To ensure the reliability of the model, this study used an external validation dataset (GSE132903) with a larger sample size to assess the generalizability of the model. Given that the external validation set GSE132903 has a larger sample size, it could potentially influence the model's predictive results.

ROC curve analysis

In the GSE109887 dataset, nine genes exhibited an AUC greater than 0.7 (Fig. [3](#page-7-0)B). This finding was corroborated in the external validation set, GSE132903, where five genes maintained an AUC above 0.7 (Fig. [3](#page-7-0)C), underscoring their commendable diagnostic efficacy and adaptability.

Interpretation and visualization of the AD diagnostic model

The SHAP library was utilized to interpret and visualize the AD diagnostic model. Referring to Fig. [4](#page-8-0)A, the scatter points on the right side of the X-axis represent a positive driving effect on the occurrence and development of AD, while the scatter points on the left side of the X-axis

Fig. 2 Identification of epigenetic signatures in Alzheimer's disease. (**A**) Recognition of hypermethylated genes with downregulated expression; (**B**) Recognition of hypomethylated genes with upregulated expression. (**C**) Features required for 0.9999 of cumulative importance. (**D**) Feature Importances

represent a negative driving effect. The color of the scatter points represents feature values (red indicates high values, blue indicates low values). The comprehensive interpretation of the AD diagnostic model, derived from the SHAP algorithm, reveals that alterations in these ten epigenetic signatures significantly impact the onset and progression of AD. For example, the low expression of *BAIAP2*, *SEMA4F*, *R3HDM1*, *GNB1*, *ENC1*, *RIMS3*, and *PRRT1* increases the risk of AD.

In the SHAP model (Fig. [4A](#page-8-0)), it is clearly evident that there are potential confounding factors in the samples, such as 10 epigenetic signatures, age, and gender. In Fig. [4A](#page-8-0), the color of the scatter points for the "age" feature represents the feature values (red for high, blue for low). In this study, the majority of samples show an increased risk of AD with advancing age, although there are individual confounding factors present. The color of the scatter points for the "sex" feature in Fig. [4A](#page-8-0) represents the category (red for male, blue for female), with the study showing that most samples indicate a higher risk of AD in females, although there are individual confounding factors as well. The SHAP model provides a new perspective on the heterogeneity of AD.

Furthermore, Fig. [4B](#page-8-0) and C illustrate the epigenetic signatures in individual samples that influence the development and advancement of the disease. This model can be used to identify high-risk patients for early intervention and treatment. Lastly, this study constructed an interactive SHAP panel (red denotes risk factors, blue represents protective factors) based on all samples from the GSE109887 dataset (Fig. [4](#page-8-0)D), which enhances the interpretability of the model, allowing doctors to better understand and trust the model's predictions, thereby providing a basis for precision medicine.

Gene functional analysis

AlzData validation

The expression of all 10 epigenetic signatures were significantly down-regulated in AD compared with the normal group (Supplementary Fig. 4, Supplementary Table 6).

GO, KEGG analysis and GSEA of epigenetic signatures

The results of GO, KEGG analysis, and GSEA, which were conducted on the 10 identified epigenetic signatures, are shown in Supplementary Figs. 5–6.

Fig. 3 Modeling and evaluation. (**A**) 5-fold cross-validation ROC-AUC for Machine learning models (boxplot). (**B**) ROC analysis of the 10 epigenetic signatures for GSE109887. (**C**) ROC analysis of the 10 epigenetic signatures for GSE132903

The immune function analysis

The results revealed statistically significant differences in immune infiltration between AD and normal groups (Supplementary Fig. 7A). In AD samples, the 10 epigenetic signatures showed varying degrees of significant correlation with both immune cell infiltration and immune-related functional scores (Supplementary Fig. 7B).

Immune checkpoint analysis

In AD samples, the gene expression of immune checkpoints exhibited varying statistical disparities between groups with high and low expression groups of 10 epigenetic signatures (Supplementary Fig. 8).

NetworkAnalyst analysis

The results of TF-gene Interactions, TF-miRNA Coregulatory Network and Protein-chemical Interactions were shown in Supplementary Fig. 9.

ELMER analysis

Utilizing the ELMER R package to integrate gene expression and methylation data, we analyzed potential transcription factors upstream of epigenetic signatures. A total of four hypermethylated site-gene pairs were selected (FDR<0.05): cg14270302-PRRT1, cg24138857- PRRT1, cg23145336-PRRT1, cg20636526-PRRT1 (Supplementary Table 7). The chromosomal location of each methylation site and its target gene are shown in Fig. [5A](#page-9-0) to D, respectively.

Fig. 4 SHAP summary plots. (**A**) SHAP model interpretation. (**B**) SHAP model interpretation of AD individual sample GSM2973309. (**C**) SHAP model interpretation of AD individual sample GSM2973332. (**D**) SHAP interactive panel of AD diagnosis (according to sample similarity)

Fig. 5 Location of methylation site in chromosomes and target gene. (**A**) Location of methylation site cg14270302 in chromosomes and target gene *PRRT1.* (**B**) Location of methylation site cg24138857 in chromosomes and target gene *PRRT1.* (**C**) Location of methylation site cg23145336 in chromosomes and target gene *PRRT1.* (**D**) Location of methylation site cg20636526 in chromosomes and target gene *PRRT1*

This study identified two motifs (TFBM) that are coregulated by these four methylation sites (Fig. [6A](#page-10-0)), where the first three transcription factors for MAZ_HUMAN. H11MO.0. A are TFAP2E, ATOH1, HMGA1 (Fig. [6](#page-10-0)B). For VEZF1_HUMAN.H11MO.0.C, the first three transcription factors are TFAP2E, ZNF335, HMGA1 (Fig. [6C](#page-10-0)).

Experimental validation Luciferase reporter assay

The results demonstrate that MAZ markedly elevates the expression level of luciferase under the control of the *PRRT1* gene promoter segment (Fig. [7A](#page-10-1)).

Fig. 6 Identification of motif co-regulated by 4 methylation sites in *PRRT1.* (**A**) The four colored ellipses in the figure represent the motif regulated by four high methylation site-gene pairs (cg14270302-PRRT1, cg23145336-PRRT1, cg24138857-PRRT1, cg20636526-PRRT1). (**B**) Transcription factor near Motif MAZ. (**C**) Transcription factor near Motif VEZF1

Fig. 7 Luciferase reporter assay and ChIP-qPCR analysis of PRRT1 promoter region in HEK293T cells. (**A**) Luciferase activity was detected after co-transfected with PGL6-PRRT1 and transcription factors overexpression plasmid or control vector (*n*=3). (**B**) ChIP-qPCR of PRRT1 promoter region with anti-MAZ antibody or IgG control in HEK293T cells

ChIP-qPCR

The results reveal that the IgG group shows virtually no enrichment, while the IP group exhibits about 3.8% of the input sample's enrichment (Fig. [7](#page-10-1)B).

Expression of PRRT1 in AD cell models

An AD cell model was established by treating SH-SY5Y cells with $A\beta_{1-42}$. CCK8 assay results indicate that SH-SY5Y cells exposed to varying concentrations of $A\beta_{1-42}$ (0.25µM, 0.5µM, 1µM, 2µM, and 5µM) exhibited viabilities of 91.79%, 79.67%, 51.73%, 36.96% and 28.34%. We selected a concentration of IC50 to establish an AD cellular model. Subsequently, the protein expression levels of PRRT1 was examined. The results revealed a notable reduction in PRRT1 protein levels in the AD group when contrasted with the control group (Supplementary Fig. 10A-10B).

Targeted regulation of PRRT1 effects viability in AD cells

We employed siRNA to suppress the expression of PRRT1 in SH-SY5Y cells and evaluated the efficacy of three distinct PRRT1 siRNAs using immunoblotting (Fig. [8A](#page-11-0)). The results demonstrated that si-PRRT1-3 had the highest efficiency; consequently, this study utilized si-PRRT1-3 for subsequent experiments. CCK8 assay was used to explore the viability and IC_{50} concentration of SH-SY5Y cells treated with different $A\beta_{1-42}$ concentrations (Supplementary Fig. 11). Then, IC_{50} concentration

Fig. 8 Targeted regulation of PRRT1 can influence cell viability, phosphorylated tau production, and autophagy in Aβ₁₋₄₂-treated SH-SY5Y cells. (**A**) Western blot analysis to assess the effects of the plasmid and three types of siRNAs on PRRT1 expression. (**B**) Comparison of cell viability among the groups in SH-SY5Y cells. (**C**) (**D**) Comparison of flow cytometry apoptosis analysis among different groups of SH-SY5Y cells. (**E**) (**F**) Western blot analysis of p-Tau, Tau, and LC3 expression in each group of SH-SY5Y cells (*n*=3)

of $A\beta_{1-42}$ was used to construct an AD cell model for subsequent experiments.

The impact of intervention targeting PRRT1 on Aβ-treated SH-SY5Y cells was assessed using CCK8 assays. The results (Fig. [8](#page-11-0)B) revealed a substantial rise in cell viability within the control group relative to the Aβ group ($p < 0.0001$), and the viability in the A β +PRRT1 group was notably higher than in the Aβ+vector group (p <0.01). Additionally, the viability in the $\text{A}\beta$ +si-PRRT1 group was significantly lower compared to the $A\beta + si-NC$ group $(p < 0.0001)$.

Overexpression of PRRT1 significantly reduces apoptosis rate in AD cells

The effect of overexpressing PRRT1 on apoptosis in AD cells was evaluated through flow cytometry apoptosis assays. The findings indicated that overexpression of PRRT1 could reduce the apoptosis rate in AD cells. Figure [8](#page-11-0)C illustrates a notable reduction in the apoptosis rate within the Aβ+PRRT1 group when compared to the Aβ+vector group ($p < 0.0001$).

Overexpression of PRRT1 reduces phosphorylated tau production and affects autophagy in AD cells

Changes in phosphorylated tau and autophagy in AD cells were detected (Fig. [8](#page-11-0)E). The results revealed that overexpression of PRRT1 reduced the production of phosphorylated tau in Aβ-treated SH-SY5Y cells compared to the Aβ+vector group (p <0.001) (Fig. [8F](#page-11-0)). In comparison to the control group, the Aβ group exhibited a significant increase in LC3 I and LC3 II expression levels, whereas the Aβ+PRRT1 group showed a notable decrease compared to the $Aβ$ +vector group (Fig. [8F](#page-11-0)).

Overexpression of MAZ upregulates PRRT1 expression, improves viability, and reduces apoptosis rate in AD cells

Immunoblotting detection showed that overexpression of MAZ in Aβ-treated SH-SY5Y cells also upregulated PRRT1 expression (Fig. [9](#page-13-0)A). CCK8 assays (Fig. [9](#page-13-0)B) showed a significant decrease in cell viability within the Aβ group relative to the control group (p <0.0001); however, the viability in the $A\beta+MAZ$ group was significantly improved when compared to the Aβ+vector group $(p<0.001)$.

Flow cytometry apoptosis detection indicated that overexpression of MAZ could reduce the apoptosis rate in AD cells. Figure [9](#page-13-0)C shows a significant decrease in the apoptosis rate in the Aβ+MAZ group compared to the Aβ+vector group ($p < 0.0001$).

Overexpression of MAZ reduces phosphorylated tau production and affects autophagy in AD cells

The impact of overexpressing MAZ on phosphorylated tau production in Aβ-treated SH-SY5Y cells was examined by immunoblotting (Fig. [9](#page-13-0)E). The results indi-

cated that the $A\beta+MAZ$ group could reduce phosphorylated tau production compared to the Aβ+vector group (Fig. [9F](#page-13-0)).

Additionally, relative to the control group, the Aβ group exhibited a significant rise in LC3 I and LC3 II expression levels, along with an increased LC3 II/LC3 I ratio. In contrast, the Aβ+MAZ group showed notably reduced expression of LC3 I and LC3 II compared to the Aβ+vector group (Fig. [9F](#page-13-0)).

MAZ improves AD cell viability and reduces apoptosis rate via PRRT1

CCK8 assays showed that the Aβ+MAZ group had a significantly increased cell viability over the Aβ group, whereas the $A\beta$ +si-PRRT1 group exhibited a markedly decreased viability compared to the Aβ group. Furthermore, the viability in the $A\beta+MAZ$ group was notably higher than in the $A\beta + MAZ + si-PRRT1$ group (Fig. [10](#page-14-0)B).

Flow cytometry analysis for apoptosis indicated that the Aβ+MAZ group had a significantly reduced apoptosis rate when compared to both the Aβ group and the Aβ+MAZ+si-PRRT1 group. The apoptosis rate in the $A\beta+MAZ+si-PRRT1$ group was also significantly lower than in the $Aβ + si-PRRT1$ group (Fig. [10](#page-14-0)C).

MAZ reduces phosphorylated tau production and affects autophagy via PRRT1

Immunoblotting detection (Fig. [10](#page-14-0)F) showed that the Aβ+si-PRRT1 group had significantly elevated levels of phosphorylated tau compared to the Aβ group (p <0.01). It also showed a notable difference in LC3 I expression levels between the Aβ group and the Aβ+si-PRRT1 group $(p<0.0001)$. Additionally, the LC3 I expression level in the Aβ+MAZ+si-PRRT1 group was significantly higher than in the $Aβ+MAZ$ group ($p < 0.05$) but lower than in the Aβ+si-PRRT1 group (p <0.001). Moreover, the LC3 II expression in the $Aβ + si-PRRT1$ group was significantly increased when compared to the Aβ group (p <0.001), and the LC3 II level in the A β +MAZ+si-PRRT1 group was significantly reduced compared to the Aβ+si-PRRT1 group (p <0.01).

Single-cell transcriptome analysis

Figure [11](#page-15-0) presents the cell trajectory analysis of *PRRT1* expression in the GSE157827 dataset, including facet plots, pseudo-time plots, and cell state plots. The *PRRT1* is highly expressed in each cell state (Fig. [11C](#page-15-0)).

In the GSE157985 dataset, cell annotations are displayed in Fig. [12](#page-16-0)A. Figure [12B](#page-16-0) shows that the expression of *Prrt1* is higher in Ex_Neuron and In_Neuron. In both AD and control groups, the expression of *Prrt1* was notably different across three cell subsets: Astrocytes, Ex_Neuron, In_Neuron (Fig. [12](#page-16-0)C). We found that these

Fig. 9 Overexpression of MAZ enhances cell viability, diminishes phosphorylated tau production, and modulates autophagy in Aβ₁₋₄₂-treated SH-SY5Y cells by upregulating PRRT1 expression. (**A**) Western blot evaluation of the impact of MAZ overexpression on PRRT1 expression. (**B**) Comparison of cell viability among the groups in SH-SY5Y cells. (**C**) (**D**) Comparison of flow cytometry apoptosis analysis among different groups of SH-SY5Y cells. (**E**) (**F**) Western blot analysis of p-Tau, Tau, and LC3 expression in each group of SH-SY5Y cells (*n*=3)

cell types exhibit significant heterogeneity within the AD brain tissue. We hypothesize that this heterogeneity among cell types may lead to diverse responses across different cell types during the disease process, thereby affecting disease progression and therapeutic outcomes. Consequently, interventions targeting specific cell types could be potential therapeutic strategies for improving AD.

Discussion

Advantages of multi-omics analysis in identifying AD biomarkers

Research has shown that aging is often accompanied by DNA methylation imbalance, and epigenetic reprogramming can reverse aging [\[29](#page-20-28)]. In a multi-omics study based on DNA methylome, transcriptome, and metabolome, researchers found that natural aging mice undergoing single-cycle transient reprogramming could drive changes in epigenetics, transcriptome, and metabolomics, improving aging phenotypes [[30](#page-20-29)]. As a

Fig. 10 MAZ regulates cell viability, phosphorylated tau production, and autophagy in SH-SY5Y cells treated with Aβ₁₋₄₂ in a PRRT1-dependent manner. (**A**) Western blot analysis of MAZ and PRRT1 expression in each group of SH-SY5Y cells. (**B**) Comparison of cell viability among the groups in SH-SY5Y cells. (**C**) (**D**) Comparison of flow cytometry apoptosis analysis among different groups of SH-SY5Y cells. (**E**) (**F**) Western blot analysis of p-Tau, Tau, and LC3 expression in each group of SH-SY5Y cells (*n*=3)

neurodegenerative disease, the pathogenesis of AD is also closely related to methylation [\[4](#page-20-6)].

Although single-omics analysis methods can provide information on biological processes that differ between disease and normal groups, there are still some limitations. Multi-omics analysis helps to find upstream driving factors and downstream regulatory mechanisms that affect key disease factors, providing new ideas for basic research and precision medicine of diseases. Compared with single-omics, multi-omics has shown advantages in identifying AD biomarkers. In this study, we first used epigenomics and transcriptomics to dig deeper into the epigenetic signatures related to AD from a deeper level,

using differential methylation sites, differential methylation sites and transcriptome gene negative correlation analysis, differential expression gene analysis, WGCNA and other comprehensive analysis to screen AD epigenetic signatures. Through the integration of epigenomic and transcriptomic multi-omics analysis, hypermethylation low-expression genes and hypomethylation highexpression genes were analyzed respectively, and 10 epigenetic signatures were identified. Then, this study used the AlzData database to verify that compared with the normal group, the expression profile of these genes in the human brain transcriptome samples showed a significant decline in AD, and these analysis results showed that

Fig. 11 Cell trajectory plot displayed by *PRRT1* expression in GSE157827. (**A**) Faceted cell trajectory plots displayed by *PRRT1* expression. (**B**) Plots of *PRRT1* expression changes with pseudo-time. (**C**) Plots of *PRRT1* expression changes with cell state

Fig. 12 Single cell analysis of *Prrt1* expression in hippocampus of AD mouse model. (**A**) The expression of *Prrt1* in each cell type (scatter plot). (**B**) The expression of *Prrt1* in each cell type (violin plot). (**C**) The expression of the *Prrt1* in all cell types between AD and control groups

these 10 genes were likely to be involved in the pathogenesis of AD. Furthermore, the results of univariate ROC curve analysis showed that the ROC-AUC of *BAIAP2*, *ENC1*, *PRRT1*, *SEMA4F*, *R3HDM1* in multiple GEO data sets were all greater than 0.7, suggesting that these genes have good clinical diagnostic value.

Explainable machine learning-driven biomarker identification

In this study, we employed an optimized random forest algorithm to construct an AD diagnostic model using 10 epigenetic signatures. The model achieved an ROC-AUC of 0.829 in the test set and 0.708 in the external validation set (GSE132903), indicating its promising clinical application prospects. We then utilized SHAP to interpret and visualize the AD diagnostic model, which clearly presented the primary factors driving and inhibiting AD progression in each subject, as well as showcasing interactions between epigenetic signatures, thereby aiding in the identification of features playing significant roles in AD development [\[16](#page-20-15), [31,](#page-20-30) [32](#page-20-31)]. This study found that alterations in the expression of these 10 genes are key factors in the development and advancement of AD.

Through the SHAP explainable AD diagnostic model established based on the random forest algorithm, AD can be effectively predicted and important influencing factors of the disease can be identified. According to this model, AD treatment plans can be accurately guided, and by influencing or changing epigenetics, such as dietary adjustments and other interventions, the occurrence of the disease can be delayed or prevented [[33–](#page-20-32)[35](#page-20-33)]. This strategy can provide accurate guidance and individualized treatment for the prevention, diagnosis, and treatment of the disease. Since DNA methylation plays a role in the etiology of AD, it is suggested that epigenetic treatment strategies be adopted for AD patients or high-risk populations. Research has discovered that some natural compounds can regulate DNA methylation status, thereby alleviating pathological characteristics of AD, such as epigallocatechin gallate (EGCG) and quercetin, which can competitively inhibit DNMT1 and lead to the re-expression of genes silenced by DNMT1-mediated methylation [\[36\]](#page-20-34). A systematic review and meta-analysis of AD proposed several intervention suggestions, including controlling diabetes, hypertension, and other medical histories, advocating physical exercise, improving lifestyle or diet [\[37\]](#page-20-35), providing guidance for the prevention of AD.

Furthermore, we developed an interactive SHAP panel for AD diagnosis that includes all samples, which can be used for disease diagnosis and visual analysis of clinical sample subjects. Based on sample similarity, it is possible to analyze which disease-affecting features are present in similar patients and whether these features are responsible for the identification of AD, thus helping medical professionals gain a deeper understanding of disease factors and provide corresponding interventions. Therefore, employing machine learning algorithms to establish an explainable AD diagnostic model by SHAP can more effectively identify disease biomarkers and provide a decision-making basis for precision treatment of the disease.

To apply the research findings to clinical practice, the following steps need to be completed: (1) Validate the model's performance on a larger independent dataset. (2) Conduct clinical trials to assess the model's effectiveness in real clinical settings. (3) Collaborate with clinical experts to optimize model parameters to better suit clinical needs. In the process of translating research findings into clinical applications, the challenges we may face include: (1) Data quality and availability: ensuring the model's robustness across data from different hospital sources. (2) Patient privacy and ethical issues: ensuring compliance with relevant laws and regulations when collecting and using patient data.

Epigenetic mechanism of *PRRT1*

To further investigate the potential upstream transcription factors of the 10 epigenetic signatures discussed in this study, we utilized the ELMER R package to integrate gene expression and methylation data, analyzing the potential upstream transcription factors of these genes. We specifically identified four hypermethylated site-gene pairs upstream of *PRRT1*, with Motifs shared by all four methylation sites. Notably, two Motifs were found to be commonly regulated by these four methylation sites. For MAZ_HUMAN.H11MO.0.A, the top three transcription factors were TFAP2E, ATOH1 and HMGA1, while for VEZF1_HUMAN.H11MO.0.C, the top three transcription factors were TFAP2E, ZNF335 and HMGA1. Subsequent luciferase reporter assays suggested an interaction between the transcription factor MAZ and the promoter region of *PRRT1*. Further ChIP-qPCR indicated an interaction between MAZ and *PRRT1*, implying that the transcription factor MAZ may participate in AD pathogenesis by promoting the transcriptional expression of *PRRT1*. Given the novelty of the mechanism involving *PRRT1* and the transcription factor MAZ, which has not been previously reported in the literature, our findings point toward a new direction for exploring the regulatory role between the transcription factor MAZ and its target gene *PRRT1*, offering innovative application value.

In conjunction with existing literature, bioinformatics analysis from our study, and experimental results, we analyzed the potential biological functions of the epigenetic signature *PRRT1* as follows.

PRRT1 (Proline-Rich Transmembrane Protein 1) is a protein-coding gene known to regulate basal and

Fig. 13 Diagram of the transcriptional regulation of *PRRT1* by MAZ in SH-SY5Y cells. In the AD cell model, the transcription factor MAZ mediates apoptosis and autophagy by binding to the target gene *PRRT1*

plasticity-induced AMPA receptor trafficking [\[38](#page-20-36)]. Transmembrane proteins play crucial roles in cell signaling $[39]$, substance transport $[40]$ $[40]$, and cell adhesion [[41\]](#page-20-39), and are key in the pathogenic mechanisms of many diseases [[42](#page-20-40), [43](#page-20-41)]. The role of transmembrane proteins in neurological disorders has gained attention in recent years, with studies indicating that *PRRT1* is essential for the development of excitatory synapses and cognitive function plasticity [[44\]](#page-20-42). Our explainable AD diagnostic model suggests that *PRRT1* is a significant factor influencing the progression of AD. Univariate ROC curve analysis showed an ROC-AUC of 0.786 for *PRRT1* in the GSE109887 dataset, indicating its potential clinical diagnostic value. Given the unclear role of *PRRT1* in AD pathogenesis, we assessed its expression in an AD cell model. Our experimental results showed significantly lower membrane protein expression levels of *PRRT1* in the AD cell model compared to controls, corroborating the SHAP model interpretation that low expression of *PRRT1* increases the risk of AD. Therefore, we conducted further experiments to explore the role of *PRRT1* in AD and the regulatory mechanism of transcription factor MAZ on *PRRT1*. Firstly, we demonstrated that *PRRT1* is involved in the pathogenesis of AD. Our experimental results revealed that knockdown of *PRRT1* affects AD cell viability, while overexpression of *PRRT1* improves cell viability and reduces apoptosis. Moreover, overexpression of *PRRT1* influences autophagy and decreases the

production of phosphorylated tau. Next, our study found that overexpression of MAZ in AD cells also increased the protein expression level of *PRRT1*. Additionally, overexpression of MAZ improved AD cell viability, reduced apoptosis rate, and decreased the production of phosphorylated tau, as well as affecting autophagy. Finally, our rescue experiments revealed that MAZ regulates cell viability, phosphorylated tau production, and autophagy in SH-SY5Y cells treated with $A\beta_{1-42}$ in a PRRT1-dependent manner.

Existing evidence suggests that impaired autophagy in AD can be improved by enhancing mitochondrial autophagy, reversing AD pathological features and memory functions in AD animal models [[45\]](#page-20-43). Overexpression of *PRRT1* was found to improve neuronal viability and reduce apoptosis, impacting autophagy. Thus, we speculate that *PRRT1* may improve AD conditions by affecting autophagy. Excessive phosphorylated tau leading to the formation of neurofibrillary tangles is another major pathological feature of AD, making targeted Tau protein therapy a hotspot in AD research [[46](#page-20-44)]. Studies have identified the transmembrane protein TREM2 as a significant risk gene implicated in AD pathogenesis [\[47](#page-20-45)]. A study recently showed that soluble TREM2 activates transgelin-2 in AD, improving Tau protein phosphorylation and cognitive deficits [[48\]](#page-20-46). Alector Therapeutics has developed the monoclonal antibody AL002 aiming to activate the TREM2 signal and improve cell viability,

thereby ameliorating AD conditions; however, the drug is still in phase II clinical trials, and its clinical benefits remain uncertain [\[49\]](#page-20-47). In the cell trajectory analysis of single-cell transcriptomics, we can observe that *PRRT1* is highly expressed in different cell states, indicating that this gene is an important participant in the progression of AD disease. Our research demonstrates that overexpression of *PRRT1* reduces the production of phosphorylated tau and improves AD conditions, providing new insights into the molecular pathological mechanisms of AD and offering novel avenues for targeting the transmembrane protein *PRRT1* in AD diagnosis and treatment.

Single-cell analysis showed that *Prrt1* was predominantly localized in neurons within the hippocampus of an AD mouse model. Furthermore, there were notable differences in its gene expression between the AD and control groups in cell types of neurons. The lack of animal experimentation is a limitation of this study. Consequently, this study requires further experimental validation of the molecular mechanisms in animal models. Additionally, according to our research findings, we speculate that in AD, the hypermethylation of DNA methylation sites cg14270302, cg23145336, cg24138857 and cg20636526 affects the interaction between transcription factor MAZ and target gene *PRR*T1, thereby mediating phenotypes such as apoptosis and autophagy. Therefore, exploration of the fine-tuned regulation on the interaction between MAZ and target gene *PRRT1* by the coordinated action of histone modification and DNA methylation would also be conducted.

Conclusion

This study identified several AD epigenetic signatures through multi-omic analyses and explainable machine learning, mapped the epigenetic landscape of DNA methylation and transcription factor regulation on epigenetic signature of *PRRT1*, and experimentally explored the transcriptional regulatory mechanisms of *PRR*T1. It was revealed that the interaction between transcription factor MAZ and target gene *PRR*T1 mediates apoptosis and autophagy in AD, a previously unreported mechanism (Fig. [13\)](#page-18-0). Based on our research findings, we speculate that in AD, the hypermethylation of DNA methylation sites cg14270302, cg23145336, cg24138857 and cg20636526 affects the interaction between transcription factor MAZ and target gene *PRR*T1, thereby mediating phenotypes such as apoptosis and autophagy. This discovery enriches the AD biomarker map and broadens our understanding of potential pathogenic mechanisms in AD.

Meanwhile, the explainable machine learning models ensure transparency and biological coherence in our findings. In the SHAP interaction panel for AD individual patients (Fig. [4](#page-8-0)C), an interactive relationship between

PRRT1 and CDK5 can be observed. The abnormal activation of CDK5 is closely related to the excessive phosphorylation of tau, which is one of the important pathological features of AD. This interaction synergistically promotes disease progression. Such an interactive relationship is of significant importance for understanding the onset and development of the disease. Machine learning explains the heterogeneity of the pathogenesis in AD patients through the interactions among genes between different patients, pointing out directions for finding potential therapeutic targets and precise treatment.

This integrative strategy offers a robust framework for future studies aiming at epigenetic targets in AD and demonstrates the power of combining multi-omics data with advanced computational analyses in elucidating disease mechanisms.

Supplementary Information

The online version contains supplementary material available at [https://doi.or](https://doi.org/10.1186/s13195-024-01646-x) [g/10.1186/s13195-024-01646-x.](https://doi.org/10.1186/s13195-024-01646-x)

Supplementary Material 1 Supplementary Material 2 Supplementary Material 3

Acknowledgements

We thank Dr. Yuzhou Xue from the Department of Cardiology, Peking University Third Hospital for generously sharing his experience.

Author contributions

Study design: FW, YL, QW. Data analysis: FW and YL. Data interpretation: FW and YL. Experimental operation: FW. Drafting manuscript: FW. Revising manuscript content: YL and QW. All authors approved the final version of the Manuscript.

Funding

This research was funded by the Zhejiang Province Key program (No. 2024C03101), Zhejiang Provincial Natural Science Foundation of China (No. LQ23H090003), Zhejiang Province Medical and Health Technology Project (2023KY1135), National Natural Science Foundation of China (No. 32171035).

Data availability

The datasets analysed in this study is publicly available in the GEO database (<https://www.ncbi.nlm.nih.gov/gds/>) and SCAD-Brain database ([http://www.b](http://www.bioinform.cn/SCAD/) [ioinform.cn/SCAD/](http://www.bioinform.cn/SCAD/)).

Declarations

Ethical approval

There were no human or animal studies in this research, so this declaration is not applicable.

Consent for publication

All authors have verified every detail in the article and agreed to submit it to your journal.

Competing interests

The authors declare no competing interests.

Received: 10 September 2024 / Accepted: 11 December 2024Published online: 07 January 2025

References

- 1. Scheltens P, Blennow K, Breteler MM, et al. Alzheimer's disease[J]. Lancet. 2016;388(10043):505–17.
- 2. Bettens K, Sleegers K, Van Broeckhoven C. Genetic insights in Alzheimer's disease[J]. Lancet Neurol. 2013;12(1):92–104.
- 3. Xiao X, Liu X, Jiao B, Epigenetics. Recent Advances and Its Role in the Treatment of Alzheimer's Disease[J]. Front Neurol. 2020;11:538301.
- 4. Qazi TJ, Quan Z, Mir A, et al. Epigenetics in Alzheimer's Disease: Perspective of DNA Methylation[J]. Mol Neurobiol. 2018;55(2):1026–44.
- 5. Han M, Liu Z, Xu Y, et al. Abnormality of m6A mRNA Methylation Is Involved in Alzheimer's Disease[J]. Front Neurosci. 2020;14:98.
- 6. Coupland KG, Kim WS, Halliday GM, et al. Effect of PSEN1 mutations on MAPT methylation in early-onset Alzheimer's disease[J]. Curr Alzheimer Res. 2015;12(8):745–51.
- 7. Tulloch J, Leong L, Thomson Z, et al. Glia-specific APOE epigenetic changes in the Alzheimer's disease brain[J]. Brain Res. 2018;1698:179–86.
- 8. Balmik AA, Chinnathambi S. Methylation as a key regulator of Tau aggregation and neuronal health in Alzheimer's disease[J]. Cell Commun Signal. 2021;19(1):51.
- 9. Hampel H, Nistico R, Seyfried NT, et al. Omics sciences for systems biology in Alzheimer's disease: State-of-the-art of the evidence[J]. Ageing Res Rev. 2021;69:101346.
- 10. Lardenoije R, Roubroeks Ja, Pishva Y. Alzheimer's disease-associated (hydroxy)methylomic changes in the brain and blood[J]. Clin Epigenetics. 2019;11(1):164.
- 11. Tian Y, Morris TJ, Webster AP, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips[J]. Bioinformatics. 2017;33(24):3982–4.
- 12. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies[J]. Nucleic Acids Res. 2015;43(7):e47.
- 13. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis[J]. BMC Bioinformatics. 2008;9:559.
- Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R[J]. BMC Bioinformatics. 2011;12:35.
- 15. W M. Data Structures for Statistical Computing in Python. In: Proc of the 9th Python in Science Conf: 2010; SCIPY 2010;[J], 2010.
- 16. Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery[J]. Nat Biomed Eng. 2018;2(10):749–60.
- 17. Xu M, Zhang DF, Luo R, et al. A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease[J]. Alzheimers Dement. 2018;14(2):215–29.
- 18. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters[J]. OMICS. 2012;16(5):284–7.
- 19. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes[J]. Nat Genet. 2003;34(3):267–73.
- 20. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles[J]. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
- 21. Fu XW, Song CQ. Identification and Validation of Pyroptosis-Related Gene Signature to Predict Prognosis and Reveal Immune Infiltration in Hepatocellular Carcinoma[J]. Front Cell Dev Biol. 2021;9:748039.
- 22. Choi SH, Bylykbashi E, Chatila ZK et al. Combined adult neurogenesis and BDNF mimic exercise effects on cognition in an Alzheimer's mouse model[J]. Science, 2018, 361(6406).
- 23. Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer's disease[J]. Lancet. 2021;397(10284):1577–90.
- 24. Zhou G, Soufan O, Ewald J, et al. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis[J]. Nucleic Acids Res. 2019;47(W1):W234–41.
- 25. Silva TC, Coetzee SG, Gull N, et al. ELMER v.2: an R/Bioconductor package to reconstruct gene regulatory networks from DNA methylation and transcriptome profiles[J]. Bioinformatics. 2019;35(11):1974–7.
- 26. Li XW, Duan TT, Chu JY, et al. SCAD-Brain: a public database of single cell RNA-seq data in human and mouse brains with Alzheimer's disease[J]. Front Aging Neurosci. 2023;15:1157792.
- 27. Lau SF, Cao H, Fu A, K Y, et al. Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease[J]. Proc Natl Acad Sci U S A. 2020;117(41):25800–9.
- 28. Xu P, Chang JC, Zhou X et al. GSAP regulates lipid homeostasis and mitochondrial function associated with Alzheimer's disease[J]. J Exp Med, 2021, 218(8).
- 29. Lu Y, Brommer B, Tian X, et al. Reprogramming to recover youthful epigenetic information and restore vision[J]. Nature. 2020;588(7836):124–9.
- 30. Chondronasiou D, Gill D, Mosteiro L, Urdinguio RG, Berenguer-Llergo A, Aguilera M, et al. Multi-omic rejuvenation of naturally aged tissues by a single cycle of transient reprogramming[J]. Aging Cell. 2022;21(3):e13578.
- 31. Hathaway QA, Roth SM, Pinti MV, et al. Machine-learning to stratify diabetic patients using novel cardiac biomarkers and integrative genomics[J]. Cardiovasc Diabetol. 2019;18(1):78.
- 32. Deshmukh F, Merchant SS. Explainable machine learning model for predicting GI bleed mortality in the intensive care unit[J]. Am J Gastroenterol. 2020;115(10):1657–68.
- Zwergel C, Valente S, Mai A. DNA Methyltransferases Inhibitors from Natural Sources[J]. Curr Top Med Chem. 2016;16(7):680–96.
- 34. Pimenova AA, Raj T, Goate AM. Untangling Genetic Risk for Alzheimer's Disease[J]. Biol Psychiatry. 2018;83(4):300–10.
- 35. Killin LO, Starr JM, Shiue IJ, et al. Environmental risk factors for dementia: a systematic review[J]. BMC Geriatr. 2016;16(1):175.
- 36. Lee WJ, Shim JY, Zhu BT. Mechanisms for the inhibition of DNA methyltransferases by tea catechins and bioflavonoids[J]. Mol Pharmacol. 2005;68(4):1018–30.
- 37. Yu JT, Xu W, Tan CC, et al. Evidence-based prevention of Alzheimer's disease: systematic review and meta-analysis of 243 observational prospective studies and 153 randomised controlled trials[J]. J Neurol Neurosurg Psychiatry. 2020;91(11):1201–9.
- 38. Troyano-Rodriguez E, Mann S, Ullah R, et al. PRRT1 regulates basal and plasticity-induced AMPA receptor trafficking[J]. Mol Cell Neurosci. 2019;98:155–63.
- Pang S, Liu J, Li T, et al. Folding and Unfolding of a Fully Synthetic Transmembrane Receptor for ON/OFF Signal Transduction[J]. J Am Chem Soc. 2023;145(38):20761–6.
- 40. Shen J, Ren C, Zeng H. Membrane-Active Molecular Machines[J]. Acc Chem Res. 2022;55(8):1148–59.
- 41. Jimenez-Amilburu V, Stainier DYR. The transmembrane protein Crb2a regulates cardiomyocyte apicobasal polarity and adhesion in zebrafish[J]. Development, 2019, 146(9).
- 42. Fu YX, Li F. [Research Progress of Transmembrane Protein Abnormality in Non-Hodgkin's Lymphoma --Review][J]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2022;30(5):1617–21.
- 43. Liu Y, Zheng Q, He G, et al. Transmembrane protein 215 promotes angiogenesis by maintaining endothelial cell survival[J]. J Cell Physiol. 2019;234(6):9525–34.
- 44. Matt L, Kirk LM, Chenaux G, et al. SynDIG4/Prrt1 Is Required for Excitatory Synapse Development and Plasticity Underlying Cognitive Function[J]. Cell Rep. 2018;22(9):2246–53.
- 45. Fang EF, Hou Y, Palikaras K, et al. Mitophagy inhibits amyloid-beta and tau pathology and reverses cognitive deficits in models of Alzheimer's disease[J]. Nat Neurosci. 2019;22(3):401–12.
- 46. Congdon EE, Ji C, Tetlow AM, et al. Tau-targeting therapies for Alzheimer disease: current status and future directions[J]. Nat Rev Neurol. 2023;19(12):715–36.
- 47. Huang W, Huang J, Huang N, et al. The role of TREM2 in Alzheimer's disease: from the perspective of Tau[J]. Front Cell Dev Biol. 2023;11:1280257.
- 48. Zhang X, Tang L, Yang J, et al. Soluble TREM2 ameliorates tau phosphorylation and cognitive deficits through activating transgelin-2 in Alzheimer's disease[J]. Nat Commun. 2023;14(1):6670.
- 49. Wang S, Mustafa M, Yuede CM et al. Anti-human TREM2 induces microglia proliferation and reduces pathology in an Alzheimer's disease model[J]. J Exp Med, 2020, 217(9).

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.