Integrated Analysis of Genes Related to Alzheimer's Disease Pathogenesis Using Gene Co-expression Networks

Xinghong he
Shanxi Medical University
Xiuya Zhou
Shanxi Medical University
Lifei Zhang
Shanxi Medical University
Ning Wang
Shanxi Medical University
Xiaochen Niu
Shanxi Medical University
Chongmin Zou
Shanxi Medical University
Yang Li
Shanxi Medical University
Li Wang
Shanxi Medical University
Xiaohui Wang (163.wangxh@163.com)
Shanxi Medical University https://orcid.org/0000-0002-5053-1008

Research Article

Keywords: Alzheimer's Disease, autophagy, STX17, WGCNA, LASSO

Posted Date: May 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1566317/v1

Abstract

Background: This study aims to identify key molecular targets in Alzheimer's disease (AD) occurrence and progression.

Methods: GSE5281 was obtained by screening and collection of the GEO database and denoted as set 1. Differential analysis was carried out on AD samples and healthy samples in set 1. Hippocampal tissues which were extracted from APP/PS1 double transgenic mice were used to measure the syntaxin 17 (STX17) gene and protein expression. Set 1 samples was divided into the STX17 high expression group and the low expression group and denoted as set 2. Differential analysis was carried out on set 2. The hippocampal sample expression matrix in GSE48350 was named as set 3 for weighted gene co-expression network analysis (WGCNA). GSE33000 was used to construct the Least absolute shrinkage and selection operator (LASSO) model for analysis and validation.

Results: 6151 differentially expressed genes (DEGs) were obtained in set 1. STX17 has significantly low expression in the hippocampal tissues of AD mice. 3651 DEGs were obtained from set 2. 2658 common DEGs were obtained in the overlap of the two sets. 22 co-expression modules were obtained from set 3. 401 genes were ultimately obtained from the overlap between genes in the WGCNA significance module and the common DEGs in the two sets and Cytoscape was used for further visualization analysis to obtain the PPI networks of 18 crucial genes. LASSO model construction and fitting was carried out to obtain seven genes common to AD and STX17 (AMPH, GAD2, GAP43, REPS2, SGIP1, STXBP1, SYN2).

Conclusion: Bioinformatics analysis was used to examine the intrinsic mechanisms of AD pathogenesis and determine the important role of STX17 in AD progression. Integrated analysis was used to obtain crucial molecular targets common to AD and STX17, which provides new ideas for future AD mechanistic studies and treatment.

Introduction

Alzheimer's disease (AD) is a neurodegenerative disease and is the most common form of dementia [1]. Its main clinical characteristics are progressive decline in cognitive function and learning memory, and abnormalities in personality and judgment ability [2]. Abnormal deposition of the B-amyloid protein (Aβ) is the most significant characteristic of AD pathogenesis and Aβ plaque deposition occurs earlier than another characteristic of AD, tau protein phosphorylation [3–5]. During formation of Aβ plaques, abnormal vesicle aggregation and axon swelling can be seen at the ends of synapses and these two effects result in vesicular recycling dysfunction, thereby resulting in an imbalance in Aβ plaque formation and clearance, accelerating abnormal Aβ plaque deposition and resulting in severe neuropathic changes [6–8].

Autophagy is an important step in vesicular recycling and is inhibited during AD progression [9]. Autophagy removes misfolded proteins and damaged vesicles to maintain cell homeostasis [10]. During autophagy, autophagy-related genes (ATG) regulate autophagosome formation [11, 12] and

autophagosomes transport phagocytosed substances to lysosomes for digestion and degradation [13, 14]. Autophagy is an important pathway for intracellular Aβ clearance [10, 15]. Recent studies found that impaired fusion between autophagosomes and lysosomes is a critical mechanism of abnormal Aβ plaque deposition in AD patients [16]. The intrinsic mechanism of impaired fusion between autophagosomes and lysosomes is still unclear and critical components and regulatory pathways remain to be elucidated.

In this study, bioinformatics analysis was employed and weighted gene co-expression network analysis (WGCNA) and differentially expressed gene (DEG) analysis was used to determine crucial genes that are intimately associated with AD. Least absolute shrinkage and selection operator (LASSO) was used to construct an AD prediction model. A receiver-operating characteristic (ROC) curve was used to assess the LASSO model. Integrated analysis will help to further determine the intrinsic mechanisms of AD occurrence and progression and examine crucial molecular targets that are common to AD and (STX17), which will provide new ideas for discovery of potential targets for AD treatment and confirmation of new biomarkers for AD diagnosis.

Materials And Methods

Data processing

The Gene Expression Omnibus database (GEO https://www.ncbi.nlm.nih.gov/geo/) contains sequencing data submitted by scientists globally and three AD-related datasets were selected in this study for research and validation. Among these datasets, GSE5281 was based on the GPL570 platform. This dataset includes brain tissue samples from 87 AD patients and 74 healthy individuals and is known as set 1. This set was used to validate the expression of DEGs in AD patients. The median expression level of STX17 in GSE5281 was used to form set 2 and used for screening of DEGs with significant correlation in STX17 expression. In this study, hippocampal tissue sequencing samples in GSE48350 were used to form set 3. This set includes brain tissue samples from 18 AD patients and 43 healthy individuals and the GPL570 platform was used to construct a co-expression network. The top 75% of genes by median absolute deviation were selected from the dataset. GSE33000 was based on the GPL4372 platform. This dataset includes brain tissue samples from 310 AD patients and 157 healthy individuals and was used to validate the fit between STX17 and its related genes in AD. The “justRMA” function in “Affy” was used for normalization of GSE5281 and GSE48350. The “limma” package was used for background correction and standardization of the GSE33000 dataset. HAMdb (http://hamdb.scbdd.com/) contains 797 autophagy-related genes. The datasets used in this study were obtained from the public access GEO database and ethics review and approval were not required.

Identification of DEGs

In this study, the GSE5281 dataset was selected and probes were converted to gene symbols for subsequent study. The “limma” package in R 4.0.3 was used to screen for differential expression results between AD and healthy tissues in set 1. “limma” was used to screen for differential expression results between STX17 high expression and STX17 low expression in set 2. The screening criteria were: corrected *P*-value (False discovery rate, FDR) ≤ 0.05 and $\log_2$FC ≤ -0.4 or $\log_2$FC ≥ 0.4 where FC is fold change. “pheatmap” in R software was used to plot differential heat maps. “ggplot2” in R software was used to plot volcano plots of DEGs.

Experimental animals

Nine-month-old male APP/PS1 mice weighing 20 ± 3 g were housed under suitable temperature and humidity and given *ad libitum* access to food and water. All animal experiments in this study were approved by the Ethics Committee of Shanxi Medical University and conformed to the National Experimental Animal Usage Regulations.

Primary reagents used in animal experiments

STX17 antibody (Proteintech, USA, 17815-1-P), GAPDH antibody (Beijing Zhongshan Golden Bridge Biotechnology, China, bsm-33033M), β-actin antibody (Beijing Zhongshan Golden Bridge Biotechnology, China, TA-09), horseradish peroxidase-labeled goat anti-mouse IgG (Beijing Zhongshan Golden Bridge Biotechnology, China, ZB-2305), SDS-PAGE horseradish peroxidase-labeled goat anti-rabbit IgG (Beijing Zhongshan Golden Bridge Biotechnology, China, ZB-2301), Prime Script RT Master Mix (TaKaRa, Japan, RR036A), and SYBR Premix Ex TaqTM II (TaKaRa, Japan, RR820A) were used.

Main equipment

Microplate reader (SoftMax, USA, SMP500-071 47-HLXU), low temperature high-speed centrifuge (Thermo Fisher, USA, LR56495), PCR machine (Stratagene, USA, MX3005P); thermal cycler (MJ Research, USA, MX3005P), vertical electrophoresis tank (BIO-RAD, USA, 042BR11805), semi-dry membrane transfer cell (BIO-RAD, USA, 221BR22693), and gel imaging system (UVP, USA, BioSpectrum 810) were used.

Western blot

BSA was used to measure protein concentration in the extracted tissue supernatant. After SDS-PAGE was used to resolve proteins, they were transferred to a PVDF membrane and blocked with 5% skimmed milk at room temperature for 2 h. After that, TBST was used to wash the membranes three times before primary antibodies were added and incubated at 4°C overnight. On the second day, the membranes were washed with TBST three times before secondary antibodies were added, and the membranes were incubated at 4°C for 2 h. The membranes were washed three times. Super ECL Plus (volume of solution

A: volume solution B = 1:1) was mixed evenly before added to PVDF membranes (around 200 μl). The BioSpectrum 810 Imaging System was used to acquire images. Exposure was performed in the gel imaging system and ImageJ software was used for analysis.

Real-time PCR

An appropriate amount of hippocampal tissue was cut and 300 μl Trizol was added for sufficient lysis before homogenization (60 s, 60 Hz). The lysate was then centrifuged (12000 rpm, 4°C, 5 min) and the supernatant was collected for total RNA extraction, RNA dissolution, and concentration tests. RNA was reverse transcribed into a cDNA template and SYBR green was used to amplify the cDNA product. All results were normalized to GAPDH expression in the control group. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression level of the target gene mRNA.

PPI network construction and core gene screening

Key entries in DEG enrichment analysis were selected and genes were extracted. The STRING Database (https://cn.string-db.org/) provided construction and correlation evaluation for protein interaction networks. The Cytoscape 3.8.2 software was used for further analysis and visualization of interaction networks. To obtain the most significant modules, the MCODE plugin in Cytoscape was used for analysis and modules with scores greater than 4 were key modules.

Construction of weighted gene co-expression networks

To better understand AD gene expression and interactions, the “WGCNA” package in R software was used for weighted co-expression analysis of the GSE48350 dataset. Hierarchical clustering accordingly found that there were no abnormalities in the samples. Following that, a scale-free network was constructed based on gene expression levels and the “pickSoftThreshold” function was used for screening with a soft threshold of β = 7. Next, an adjacency matrix was constructed to determine the correlation between genes and the adjacency matrix was transformed into a topological overlap matrix (TOM) to describe the similarity of every node. The dynamic tree cut criteria was used to divide genes with similar expression spectrum into different modules and the minimum number of genes in each module was set as 30 (MinModuleSize = 30). The modules were visualized to obtain dendograms and the correlation coefficient matrix between module eigengenes (MEs) and clinical traits and the module with the highest AD correlation was known as the optimal module. The correlation between gene significance (GS) and module membership (MM) in the module was assessed and genes with GS > 0.2 and MM > 0.8 in the module were considered central genes.

Gene functional enrichment analysis

The “clusterProfiler” package in R 4.0.3 was used for enrichment analysis, including gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. GO is composed of molecular function (MF), cellular component (CC), and biological process (BP). The “ggplot2” package in R software was used for visualization. $P < 0.01$ and FDR < 0.05 were used as screening criteria.

LASSO model construction and ROC curve evaluation

The LASSO regression model has good simulation and prediction capabilities. The central genes in the GSE33000 dataset and the expression spectrum of core genes after intersection of DEGs were selected and used to construct the LASSO model using the “glmnet” package in the R software. The following formula was used to calculate the expression and regression coefficients of core genes: index = ExpGene1 × Coef1 + ExpGene2 × Coef2 + ExpGene3 × Coef3 +…+ ExpGeneN × CoefN. Where “Exp” is the gene expression level and “Coef” is the regression coefficient of the gene. To accurately identify the AD and healthy groups, we randomized the datasets: training set (70%) and test set (30%). At the same time, to accurately identify the correlation between genes and STX17, the median expression level of STX17 was used as a threshold value and the GSE33000 dataset was used to construct a new LASSO regression model. The “ROCR” package in the R software was used to construct receiver operating characteristic (ROC) curves to assess the stability and fit of the LASSO model.

Results

Importance of screening of significant DEGs in AD and confirmation of autophagy-related gene STX17 in AD

To determine genes with significant differences in AD patients, the GSE5281 dataset was downloaded from the GEO database and denoted as set 1 for differential analysis. 6151 significant DEGs were identified in set 1 (Additional file 1), of which 2364 were upregulated and 3787 were downregulated (Fig. 2a, b). To further examine the potential function of correlated genes, DEGs were selected for GO and KEGG enrichment analyses. KEGG showed that DEGs mainly participate in “Pathways of neurodegeneration-multiple diseases” and “Alzheimer’s disease” (Fig. 2c). GO enrichment analysis showed that DEGs in AD mainly participate in the biological process of “vesicle transport”, cellular component was mainly enriched for “Golgi vesicular transport”, and mainly participate in the molecular functions of “molecular adaptor activity” and “phosphatase binding” (Fig. 2d). To further explore the potential autophagy mechanism in AD, genes included in main entries were extracted and used for intersection analysis with autophagy-related genes in the autophagy database and only one intersecting gene, STX17, was obtained in the result (Fig. 2e). At the same time, we also found that the expression of STX17 is significantly decreased in the brain tissues of AD patients in set 1, $P = 0.00069$, $\log_2 FC = -0.44693$ (Fig. 2f), showing that low STX17 expression may be significant in autophagy in AD occurrence and progression.

Significant low expression of STX17 in hippocampal tissues in APP/PS1 double transgenic mice

To further prove the expression of STX17 in AD, 9-month-old APP/PS1 double transgenic mice were selected for study. Results showed that STX17 mRNA and protein expression levels in the hippocampal tissues of APP/PS1 double transgenic mice were significantly decreased compared with healthy mice ($P < 0.05$; Fig. 3a, b).

Screening of 2658 common DEGs in AD and STX17

To further determine significantly correlated genes in AD when STX17 expression was low, we used the median STX17 expression in set 1 as a threshold to construct set 2 and carried out differential expression correlation analysis of genes. 3651 significant DEGs were identified in set 2 (Additional file 1), of which 269 were upregulated and 3382 were downregulated (Fig. 4a, b). DEGs in sets 1 and 2 were preliminary selected based on differential analysis and Venn diagram was used to obtain the intersections of upregulated and downregulated genes in the 2 sets. 155 common upregulated DEGs and 2503 common downregulated DEGs were obtained. From the sum of these 2 intersections, 2658 common DEGs were obtained. These were known as hub genes and used for further analysis (Fig. 4c, d).

Weighted co-expression network analysis of set 3 and constructed of 22 AD-related modules

To better understand modules and genes with significant correlation to AD, we selected the GSE448350 dataset for screening of genes and modules with the most correlation to AD. The top 75% of genes by median absolute deviation (MAD) were selected and MAD > 0.01. Cluster analysis was carried out on the remaining genes to determine if there were significant outliers (Fig. 5a). The most crucial parameter in the scale-free network is the selection of the soft threshold. When the soft threshold was 7, the scale independence was close to 0.9 and good mean connectivity was obtained in the adjacency matrix (Fig. 5b). A total of 22 modules were identified based on TOM stratified clustering and the dynamic tree cut height was 0.25 (Fig. 5c, d). A heat map was used for visualization of gene networks (Fig. 5e).

Screening of modules and phenotypes in the weighted co-expression network and three modules with the most significant correlation with AD

We carried out correlation analysis of modules and phenotypes, of which the turquoise module has the highest negative correlation with AD (correlation coefficient = − 0.53, $P = 1e-05$), the red module has the highest positive correlation with AD (correlation coefficient = 0.52, $P = 2e-05$), followed by the positive

correlation between the black module and AD (correlation coefficient = 0.49, $P = 6e-05$; Fig. 6a). Following that, correlation analysis was carried out for the GS and MM of every gene in the above modules. The correlations of the turquoise module, red module, and black module were 0.75 ($P < 1e-200$; Fig. 6b), 0.7 ($P = 1.2e-141$; Fig. 6c), and 0.58 ($P = 1.9e-58$) (Fig. 6d), respectively. Genes in modules with significant AD correlation were selected for in-depth analysis. Genes selected based on thresholds of GS > 0.2 and MM > 0.8 were considered central genes, central genes were selected for GO and KEGG enrichment analyses (Additional file 2).

Construction of LASSO regression model and confirmation of seven significantly correlated genes common to AD and STX17

Central genes in the significant correlation module from WGCNA analysis (Additional file 3) and the hub gene in the differential analysis were extracted for intersection analysis to obtain 401 intersection genes (Fig. 7a). The 401 genes obtained were inputted into the STRING database to construct the PPI network (Additional file 4). Cytoscape was used for visualization analysis. The MCODE plugin was used for analysis of PPI networks and construction of multiple high correlation sub-networks to obtain 19 networks consisting of core genes (Fig. 7b). Following that, the GSE33000 dataset was used as the validation set for core genes and the expression matrices of core genes was extracted to construct the LASSO regression model (Fig. 7c). The optimal threshold lambda.min = 0.008200821 was selected and non-zero regression coefficient was used to identify 10 genes (AMPH, DNM1L, GAD2, GAP43, RAB3C, REPS2, SGIP1, STMN2, STXBP1, and SYN2). The following formula was constructed based on the gene model index: index = AMPH × 2.7966877 + DNM1L × (− 3.3706289) + GAD2 × (− 1.9793459) + GAP43 × (− 1.3960538) + RAB3C × (− 1.0899792) + REPS2 × 0.3138737 + SGIP1 × (− 6.1421089) + STMN2 × 0.6587228 + STXBP1 × (− 2.3881780) + SYN2 × 0.4154237. The ROC curve was used to validate the predicted accuracy of the LASSO regression model and the dataset was randomized into the test set and training set. The discrimination of samples in both datasets was good (Additional file 5). In the training set, the AUC of 10 genes was 0.95 (Fig. 7d). In the test set, the AUC of 10 genes was 0.93 (Fig. 7e), showing that the 10 genes above are intimately associated with AD occurrence and progression. In addition, the median value of STX17 in the GSE33000 dataset was also used as a threshold to construct the STX17 high-low expression dataset and sample discrimination was good (Additional file 5). The expression matrices of the 10 genes above were extracted from the LASSO regression model. The non-zero regression coefficient was used to identify seven genes (AMPH, GAD2, GAP43, REPS2, SGIP1, STXBP1, and SYN2) (lambda.min = 0.0144212; Fig. 7f). The construction of the following formula was based on the number of base models index = AMPH × (− 0.68484140) + GAD2 × 2.19059579 + GAP43 × 1.35726521 + REPS2 × 8.54811586 + SGIP1 × (− 7.33909761) + STXBP1 × 0.01621663 + SYN2 × 0.59866620. The ROC curve was used to validate the accuracy of the LASSO regression model. In the training set, the AUC of the seven genes was 0.81 (Fig. 7g). In the test set, the AUC of the seven genes

was 0.81 (Fig. 7h), showing that the seven genes above with significant low expression have good correlation with STX17 (Additional file 6).

Discussion

AD is a degenerative neurological disorder [17] and there is no suitable diagnostic and treatment method so far. FDA-approved drugs for AD treatment can only delay AD progression [18, 19]. Therefore, there is an urgent need to search for feasible AD therapeutic targets. Some researchers believe that autophagy is a critical pathway for homeostasis maintenance and has a huge potential as a target in future AD treatment [17, 20]. Studies found that impaired fusion between autophagosomes and lysosome can block autophagic flux [21], which may be a critical factor for ineffective Aβ clearance [10]. The critical molecules in related autophagy pathways are still unknown and the intrinsic mechanism of intermolecular crosstalk remains to be elucidated.

In this study, bioinformatics methods were used to search for AD-related potential intervention targets, particularly critical molecules in autophagy and related factors. Set 1 was used to compare differences and 6151 DEGs (2364 upregulated genes and 3787 downregulated genes) were obtained. GO enrichment analysis showed that DEGs in AD mainly participate in the biological process of “vesicle transport”, cellular component was mainly enriched for “Golgi vesicular transport”, and mainly participate in the molecular functions of “molecular adaptor activity” and “phosphatase binding”. To search for critical factors associated with autophagy in AD occurrence and progression, we extracted genes in main entries in this study and intersections with autophagy-related genes were obtained to obtain the only intersecting gene, STX17. This suggests that STX17 plays a crucial role in autophagy. In set 1, we found that STX17 expression was significantly low ($P = 0.00069$). To further validate STX17 expression in AD, we isolated hippocampal tissues from APP/PS1 double transgenic mice. The experimental results showed that STX17 mRNA and protein levels were significantly low in hippocampal tissues. STX17 plays a crucial role in fusion between autophagosomes and lysosomes [12, 17]. Studies showed that STX17 can form the SNARE complex with vesicle-associated membrane protein 8 (VAMP8), soluble NSF attachment protein 29 (SNAP29) to jointly mediate fusion between autophagosomes and lysosomes [12, 22]. STX17 downregulation will inevitably affect fusion between autophagosomes and lysosomes, demonstrating the significance of STX17 downregulation in AD occurrence and progression. To screen for genes intimately associated with STX17 in AD progression, we further constructed set 2 based on the median STX17 value and difference comparison of the groups was carried out to obtain 3651 DEGs (269 upregulated genes and 3382 downregulated genes). The intersection of DEG obtained by screening sets 1 and 2 was obtained and 2658 common DEGs (155 upregulated genes and 2503 downregulated genes) were obtained and known as hub genes. These significantly DEGs in AD enable us to better understand the potential mechanisms in AD and genes that are significantly correlated to STX17 expression could also provide new ideas for discovering feasible AD treatment regimens.

To better understand crucial genes in AD and to expand the range to decrease coincidence in DEGs in sets 1 and 2, the GSE448350 dataset was further selected for WGCNA analysis and construction of 22

AD-related co-expression modules. Three modules with the highest correlation were selected, including turquoise module (correlation coefficient = − 0.53, $P = 1e-05$), red module (correlation coefficient = 0.52, $P = 2e-05$), and black module (correlation coefficient = 0.49, $P = 6e-05$). Genes selected using GS > 0.2 and MM > 0.8 as thresholds were considered central genes. GO enrichment analysis of turquoise module genes mainly participates in oxidative stress. Studies have shown that mitochondrial dysfunction is one of the greatest risk factors for neurodegenerative diseases [23–25]. Mitochondrial dysfunction can be detected early in the course of AD. Mitochondrial dysfunction may be a pathogenic factor of AD [26]. Damaged mitochondria lead to increased production of reactive oxygen species (ROS), and binding to Aβ exacerbates AD progression [27].

To further screen related genes, central genes and hub genes obtained from the weighted co-expression module were used for overlapping to obtain 401 intersection genes. These genes were inputted into the STRING database for PPI network construction. The MCODE plugin in Cytoscape was used for core gene analysis and the core gene network of 18 genes was visualized. The expression matrices of 18 genes were extracted and LASSO regression was used for fit screening. ROC curves showed good fit. Finally, seven genes in the fit model (AMPH, GAD2, GAP43, REPS2, SGIP1, STXBP1, and SYN2) were considered significantly correlated core genes common to AD and STX17.

Studies have shown that amphiphysin (AMPH) is an important accessory protein in clathrin-mediated endocytosis (CME). CME is responsible for uptake of receptors or ligands on the cell surface and plays an important role in nutrient absorption, neurotransmission, and maintenance of cell homeostasis [28]. AMPH can bind to clathrin and the cell membrane and plays an important role in membrane fission [28, 29]. Studies have found that impaired synapse function due to increase in AMPH autoantibodies is an intrinsic mechanism causing chronic muscle stiffness in Stiff-Man syndrome (SMS) [30]. SH3GL interacting endocytic adaptor 1 (SGIP1) is a critical factor that regulates CME and its main function is in the assembly of clathrin-associated complexes [31]. In addition, SGIP1 is also important in energy regulation and mood control [32, 33]. This study found that AMPH and SGIP1 expression are significantly low in AD patients and their potential relationship with AD are still unclear. In particular, the roles of these genes in autophagy have not been reported and interactions between SGIP1, AMPH, and STX17 require further exploration.

γ-aminobutyric acid (GABA) is a major inhibitory neurotransmitter that is important in cognition, learning, depression, and drug addiction [34]. GABA synthesis is mainly controlled by glutamic acid decarboxylase (GAD) and GAD has two isoforms, GAD1 and GAD2. In mammals, GABA synthesis is mainly controlled by GAD1. GAD2 is mainly present in synapses and mediates GABA release in strong nerve excitation [35, 36]. Studies have pointed out that damage to the GABAergic system in AD becomes serious with age and is one of the risk factors for AD progression [37]. This study also found that expression of GAD2 is significantly low in AD and regulation of GAD2 expression in AD is worthy of further in-depth research. The release of neurotransmitters is mainly caused by fusion between synaptic vesicles and the cell membrane; SNAREs, unc-13 homolog B (Munc13), and syntaxin binding protein 1 (STXBP1) are indispensable for the fusion process. Studies have shown that STXBP1 and STX1 fold into a closed

conformation. Munc1 assists in opening the closed STX1 and forms a SNARE complex with STXBP1. After stimulation, the Ca^{2+} sensor synaptotagmin-1 jointly mediates the rapid fusion and release of neurotransmitters [38]. STXBP1 can accelerate the SNARE complex-mediated membrane fusion process and is important in neurotransmitter release [39]. Similarly, STX17 plays a crucial role in fusion between autophagosomes and lysosomes [12, 17] but the relationship between STXBP1 and STX17 have not been reported and should be explored in future research.

Neuron growth-associated protein 43 (GAP43) is considered to be a crucial regulatory factor of neuronal growth, axon regeneration, and synapse plasticity [40]. GAP43 is mainly phosphorylated by PKC before it performs its main function of extending synapse tips during nerve growth and regeneration [41]. Studies found that GAP43 can be used as a biomarker for diagnosis of asymptomatic AD patients and has significantly low expression in the cerebrospinal fluid of AD patients. This is consistent with this study that found that GAP43 has significantly low expression in the brain tissues of AD patients. The intrinsic relationship between GAP43 and STX17 in AD has not been reported.

The RALBP1 associated Eps domain containing 2 (REPS2) protein contains an EH domain, a proline-rich domain, and a coiled-coil domain [42]. Studies have shown that the EH domain mainly participates in protein-protein interactions and binding to related factors to regulate receptor-mediated internalization and signal transduction [42, 43]. The proline-rich domain can bind to proteins with the SH3 domain such as AMPH2 to inhibit EGFR-mediated internalization [44]. Studies have shown that uncontrolled internalization causes prostate cancer to worsen [45]. This study found that REPS2 expression is significantly low in AD patients and REPS2 is related to AD occurrence and progression. In particular, REPS2 acts as a crucial factor in internalization and its interaction with STX17 urgently requires further examination.

Synapses are important structures that are indispensable to nervous activity. The synapsin (SYN) family includes SYN1, SYN2, and SYN3. In particular, SYN2 is considered to be important for synapse formation and growth [46]. SYN2 is mainly located in the synapses of mature neurons and plays an important role in synaptic vesicle mobilization and recycling. Aberrant SYN2 expression may be a pathogenic factor of epilepsy [47, 48]. This study found that SYN2 expression is significantly low in AD, suggesting that it affects neuroregulatory networks in the brain. There are few reports of the effects of aberrant SYN2 expression on AD. SYN2 and STX17 are both crucial factors in vesicle recycling and their interactions deserve further study.

Conclusions

In summary, this study employed differential analysis, WGCNA, and LASSO regression analysis to identify AD-related target genes and molecular pathways, the median STX17 expression was used for grouping, and genes intimately associated with STX17 were found, which provided new ideas for in-depth research on AD regulatory networks. The ROC curve also proved that the seven aforementioned genes have good specificity and accuracy in AD prediction and STX17-related expression models. The limitations of this

study are lack of attention on the entire AD progression process and the lack of analysis of the evaluation effectiveness of genes in different AD progression stages and therapeutic effects. In the future, we will study the interactions between STX17 and the seven aforementioned genes in AD and the regulatory effects of molecules. Molecular experiments and other methods will be used to explore the changes in these genes during AD progression and their effects in improving AD symptoms. Subsequently, the diagnostic and control effects of these genes in AD will also be assessed.

Declarations

Acknowledgements

We thank LetPub (https://www.letpub.com/) for its linguistic assistance during the preparation of this manuscript. This study was financially supported by the project grant from Natural Science Foundation of Shanxi Province Applied Basic Research Program (No: 201901D111197), Cultivate Scientific Research Excellence Programs of Higher Education Institutions in Shanxi (No: CSREP 2020KJ012), Shanxi “1331 Project” Key Subjects Construction (No: 1331KSC), Open Fund from Key Laboratory of Cellular Physiology (Shanxi Medical University), Ministry of Education, China, (No: KLMEC/SXMU-201904).

Authors’ Contributions

XH designed research and collected and analyzed data; XZ, LZ, and NW analyzed data; XH, XZ, LZ, NW, XN, CZ, YL, LW, XW participated in design discussions; XH, LW and XW, CZ wrote the paper. All authors read and approved the final manuscript.

Funding

This study was financially supported by the project grant from Natural Science Foundation of Shanxi Province Applied Basic Research Program (No: 201901D111197), Cultivate Scientific Research Excellence Programs of Higher Education Institutions in Shanxi (No: CSREP 2020KJ012), Shanxi “1331 Project” Key Subjects Construction (No: 1331KSC), Open Fund from Key Laboratory of Cellular Physiology (Shanxi Medical University), Ministry of Education, China, (No: KLMEC/SXMU-201904).

Data Availability

All data generated or analysed during this study are included in this published article (and its additional fles).

Ethics approval and consent to participate

This study was carried out in accordance with the Declaration of Helsinki and approved by the Medical Ethics Committee of Shanxi Medical University.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, Menon M, He L, Abdurrob F, Jiang X, et al: **Single-cell transcriptomic analysis of Alzheimer's disease.** *Nature* 2019, **570:**332-337.http://doi.org/10.1038/s41586-019-1195-2
2. Small G, Rabins P, Barry P, Buckholtz N, DeKosky S, Ferris S, Finkel S, Gwyther L, Khachaturian Z, Lebowitz B, et al: **Diagnosis and treatment of Alzheimer disease and related disorders. Consensus statement of the American Association for Geriatric Psychiatry, the Alzheimer's Association, and the American Geriatrics Society.** 1997, **278:**1363-1371.http://doi.org/10.1001/jama.278.16.1363
3. Busche MA, Hyman BT: **Synergy between amyloid-beta and tau in Alzheimer's disease.** *Nat Neurosci* 2020, **23:**1183-1193.http://doi.org/10.1038/s41593-020-0687-6
4. Ikonomovic MD, Buckley CJ, Abrahamson EE, Kofler JK, Mathis CA, Klunk WE, Farrar G: **Post-mortem analyses of PiB and flutemetamol in diffuse and cored amyloid-beta plaques in Alzheimer's disease.** *Acta Neuropathol* 2020, **140:**463-476.http://doi.org/10.1007/s00401-020-02175-1
5. Duara R, Barker W, Loewenstein D, Bain L: **The basis for disease-modifying treatments for Alzheimer's disease: the Sixth Annual Mild Cognitive Impairment Symposium.** *Alzheimers Dement* 2009, **5:**66-74.http://doi.org/10.1016/j.jalz.2008.10.006
6. Sardar Sinha M, Ansell-Schultz A, Civitelli L, Hildesjo C, Larsson M, Lannfelt L, Ingelsson M, Hallbeck M: **Alzheimer's disease pathology propagation by exosomes containing toxic amyloid-beta oligomers.** *Acta Neuropathol* 2018, **136:**41-56.http://doi.org/10.1007/s00401-018-1868-1
7. Sanchez-Varo R, Trujillo-Estrada L, Sanchez-Mejias E, Torres M, Baglietto-Vargas D, Moreno-Gonzalez I, De Castro V, Jimenez S, Ruano D, Vizuete M, et al: **Abnormal accumulation of autophagic vesicles correlates with axonal and synaptic pathology in young Alzheimer's mice hippocampus.** *Acta Neuropathol* 2012, **123:**53-70.http://doi.org/10.1007/s00401-011-0896-x
8. Sharoar MG, Hu X, Ma XM, Zhu X, Yan R: **Sequential formation of different layers of dystrophic neurites in Alzheimer's brains.** *Mol Psychiatry* 2019, **24:**1369-1382.http://doi.org/10.1038/s41380-019-0396-2
9. Yang C, Cai CZ, Song JX, Tan JQ, Durairajan SSK, Iyaswamy A, Wu MY, Chen LL, Yue Z, Li M, Lu JH: **NRBF2 is involved in the autophagic degradation process of APP-CTFs in Alzheimer disease models.** *Autophagy* 2017, **13:**2028-2040.http://doi.org/10.1080/15548627.2017.1379633
10. Heckmann B, Teubner B, Tummers B, Boada-Romero E, Harris L, Yang M, Guy C, Zakharenko S, Green DJC: **LC3-Associated Endocytosis Facilitates β-Amyloid Clearance and Mitigates Neurodegeneration in Murine Alzheimer's Disease.** 2019, **178:**536-551.e514.http://doi.org/10.1016/j.cell.2019.05.056

11. Carlsson SR, Simonsen A: **Membrane dynamics in autophagosome biogenesis.** *J Cell Sci* 2015, **128:**193-205.http://doi.org/10.1242/jcs.141036
12. Itakura E, Kishi-Itakura C, Mizushima N: **The hairpin-type tail-anchored SNARE syntaxin 17 targets to autophagosomes for fusion with endosomes/lysosomes.** *Cell* 2012, **151:**1256-1269.http://doi.org/10.1016/j.cell.2012.11.001
13. Klionsky D, Emr SJS: **Autophagy as a regulated pathway of cellular degradation.** 2000, **290:**1717-1721.http://doi.org/10.1126/science.290.5497.1717
14. Nixon RA: **The role of autophagy in neurodegenerative disease.** *Nat Med* 2013, **19:**983-997.http://doi.org/10.1038/nm.3232
15. Scrivo A, Bourdenx M, Pampliega O, Cuervo AM: **Selective autophagy as a potential therapeutic target for neurodegenerative disorders.** *The Lancet Neurology* 2018, **17:**802-815.http://doi.org/10.1016/s1474-4422(18)30238-2
16. Long Z, Chen J, Zhao Y, Zhou W, Yao Q, Wang Y, He GJA: **Dynamic changes of autophagic flux induced by Abeta in the brain of postmortem Alzheimer's disease patients, animal models and cell models.** 2020, **12:**10912-10930.http://doi.org/10.18632/aging.103305
17. Menzies FM, Fleming A, Caricasole A, Bento CF, Andrews SP, Ashkenazi A, Fullgrabe J, Jackson A, Jimenez Sanchez M, Karabiyik C, et al: **Autophagy and Neurodegeneration: Pathogenic Mechanisms and Therapeutic Opportunities.** *Neuron* 2017, **93:**1015-1034.http://doi.org/10.1016/j.neuron.2017.01.022
18. Mielke MM, Leoutsakos JM, Corcoran CD, Green RC, Norton MC, Welsh-Bohmer KA, Tschanz JT, Lyketsos CG: **Effects of Food and Drug Administration-approved medications for Alzheimer's disease on clinical progression.** *Alzheimers Dement* 2012, **8:**180-187.http://doi.org/10.1016/j.jalz.2011.02.011
19. Hampel H, Mesulam MM, Cuello AC, Farlow MR, Giacobini E, Grossberg GT, Khachaturian AS, Vergallo A, Cavedo E, Snyder PJ, Khachaturian ZS: **The cholinergic system in the pathophysiology and treatment of Alzheimer's disease.** *Brain* 2018, **141:**1917-1933.http://doi.org/10.1093/brain/awy132
20. Caccamo A, Ferreira E, Branca C, Oddo S: **p62 improves AD-like pathology by increasing autophagy.** *Mol Psychiatry* 2017, **22:**865-873.http://doi.org/10.1038/mp.2016.139
21. Chi C, Leonard A, Knight WE, Beussman KM, Zhao Y, Cao Y, Londono P, Aune E, Trembley MA, Small EM, et al: **LAMP-2B regulates human cardiomyocyte function by mediating autophagosome-lysosome fusion.** *Proc Natl Acad Sci U S A* 2019, **116:**556-565.http://doi.org/10.1073/pnas.1808618116
22. Corona AK, Jackson WT: **Finding the Middle Ground for Autophagic Fusion Requirements.** *Trends Cell Biol* 2018, **28:**869-881.http://doi.org/10.1016/j.tcb.2018.07.001
23. Lin MT, Beal MF: **Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases.** *Nature* 2006, **443:**787-795.http://doi.org/10.1038/nature05292

24. Vandoorne T, De Bock K, Van Den Bosch L: **Energy metabolism in ALS: an underappreciated opportunity?** *Acta Neuropathol* 2018, **135:**489-509.http://doi.org/10.1007/s00401-018-1835-x
25. Kapogiannis D, Mattson MP: **Disrupted energy metabolism and neuronal circuit dysfunction in cognitive impairment and Alzheimer's disease.** *The Lancet Neurology* 2011, **10:**187-198.http://doi.org/10.1016/s1474-4422(10)70277-5
26. Nortley R, Korte N, Izquierdo P, Hirunpattarasilp C, Mishra A, Jaunmuktane Z, Kyrargyri V, Pfeiffer T, Khennouf L, Madry C, et al: **Amyloid beta oligomers constrict human capillaries in Alzheimer's disease via signaling to pericytes.** *Science* 2019, **365**.http://doi.org/10.1126/science.aav9518
27. Sivanesan S, Chang E, Howell MD, Rajadas J: **Amyloid protein aggregates: new clients for mitochondrial energy production in the brain?** *FEBS J* 2020, **287:**3386-3395.http://doi.org/10.1111/febs.15225
28. Mettlen M, Chen PH, Srinivasan S, Danuser G, Schmid SL: **Regulation of Clathrin-Mediated Endocytosis.** *Annu Rev Biochem* 2018, **87:**871-896.http://doi.org/10.1146/annurev-biochem-062917-012644
29. Takei K, Slepnev V, Haucke V, De Camilli PJNcb: **Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis.** 1999, **1:**33-39.http://doi.org/10.1038/9004
30. Petzold G, Marcucci M, Butler M, van Landeghem F, Einhäupl K, Solimena M, Valdueza J, De Camilli PJAon: **Rhabdomyolysis and paraneoplastic stiff-man syndrome with amphiphysin autoimmunity.** 2004, **55:**286-290.http://doi.org/10.1002/ana.10841
31. Lee SE, Cho E, Jeong S, Song Y, Kang S, Chang S: **SGIP1alpha, but Not SGIP1, is an Ortholog of FCHo Proteins and Functions as an Endocytic Regulator.** *Front Cell Dev Biol* 2021, **9:**801420.http://doi.org/10.3389/fcell.2021.801420
32. Dvorakova M, Kubik-Zahorodna A, Straiker A, Sedlacek R, Hajkova A, Mackie K, Blahos J: **SGIP1 is involved in regulation of emotionality, mood, and nociception and modulates in vivo signalling of cannabinoid CB1 receptors.** *Br J Pharmacol* 2021, **178:**1588-1604.http://doi.org/10.1111/bph.15383
33. Hajkova A, Techlovska S, Dvorakova M, Chambers JN, Kumpost J, Hubalkova P, Prezeau L, Blahos J: **SGIP1 alters internalization and modulates signaling of activated cannabinoid receptor 1 in a biased manner.** *Neuropharmacology* 2016, **107:**201-214.http://doi.org/10.1016/j.neuropharm.2016.03.008
34. Deng PY, Xiao Z, Yang C, Rojanathammanee L, Grisanti L, Watt J, Geiger JD, Liu R, Porter JE, Lei S: **GABA(B) receptor activation inhibits neuronal excitability and spatial learning in the entorhinal cortex by activating TREK-2 K+ channels.** *Neuron* 2009, **63:**230-243.http://doi.org/10.1016/j.neuron.2009.06.022
35. Tao R, Davis KN, Li C, Shin JH, Gao Y, Jaffe AE, Gondre-Lewis MC, Weinberger DR, Kleinman JE, Hyde TM: **GAD1 alternative transcripts and DNA methylation in human prefrontal cortex and hippocampus in brain development, schizophrenia.** *Mol Psychiatry* 2018, **23:**1496-1505.http://doi.org/10.1038/mp.2017.105
36. Patel AB, de Graaf RA, Martin DL, Battaglioli G, Behar KL: **Evidence that GAD65 mediates increased GABA synthesis during intense neuronal activity in vivo.** *J Neurochem* 2006, **97:**385-

396.http://doi.org/10.1111/j.1471-4159.2006.03741.x

37. Huang Y, Mucke L: **Alzheimer mechanisms and therapeutic strategies.** *Cell* 2012, **148:**1204-1222.http://doi.org/10.1016/j.cell.2012.02.040

38. Ma C, Su L, Seven AB, Xu Y, Rizo J: **Reconstitution of the vital functions of Munc18 and Munc13 in neurotransmitter release.** *Science* 2013, **339:**421-425.http://doi.org/10.1126/science.1230473

39. Shen C, Rathore SS, Yu H, Gulbranson DR, Hua R, Zhang C, Schoppa NE, Shen J: **The trans-SNARE-regulating function of Munc18-1 is essential to synaptic exocytosis.** *Nat Commun* 2015, **6:**8852.http://doi.org/10.1038/ncomms9852

40. Verge V, Tetzlaff W, Richardson P, Bisby MJTJontojotSfN: **Correlation between GAP43 and nerve growth factor receptors in rat sensory neurons.** 1990, **10:**926-934.http://doi.org/10.1523/jneurosci.10-03-00926.1990

41. Gauthier-Kemper A, Igaev M, Sundermann F, Janning D, Bruhmann J, Moschner K, Reyher HJ, Junge W, Glebov K, Walter J, et al: **Interplay between phosphorylation and palmitoylation mediates plasma membrane targeting and sorting of GAP43.** *Mol Biol Cell* 2014, **25:**3284-3299.http://doi.org/10.1091/mbc.E13-12-0737

42. Badway JA, Baleja JD: **Reps2: a cellular signaling and molecular trafficking nexus.** *Int J Biochem Cell Biol* 2011, **43:**1660-1663.http://doi.org/10.1016/j.biocel.2011.08.014

43. Morinaka K, Koyama S, Nakashima S, Hinoi T, Okawa K, Iwamatsu A, Kikuchi A: **Epsin binds to the EH domain of POB1 and regulates receptor-mediated endocytosis.** *Oncogene* 1999, **18:**5915-5922.http://doi.org/10.1038/sj.onc.1202974

44. Tomassi L, Costantini A, Corallino S, Santonico E, Carducci M, Cesareni G, Castagnoli L: **The central proline rich region of POB1/REPS2 plays a regulatory role in epidermal growth factor receptor endocytosis by binding to 14-3-3 and SH3 domain-containing proteins.** *BMC Biochem* 2008, **9:**21.http://doi.org/10.1186/1471-2091-9-21

45. Oosterhoff JK, Kuhne LC, Grootegoed JA, Blok LJ: **EGF signalling in prostate cancer cell lines is inhibited by a high expression level of the endocytosis protein REPS2.** *Int J Cancer* 2005, **113:**561-567.http://doi.org/10.1002/ijc.20612

46. Corradi A, Fadda M, Piton A, Patry L, Marte A, Rossi P, Cadieux-Dion M, Gauthier J, Lapointe L, Mottron L, et al: **SYN2 is an autism predisposing gene: loss-of-function mutations alter synaptic vesicle cycling and axon outgrowth.** *Hum Mol Genet* 2014, **23:**90-103.http://doi.org/10.1093/hmg/ddt401

47. Longhena F, Faustini G, Brembati V, Pizzi M, Benfenati F, Bellucci A: **An updated reappraisal of synapsins: structure, function and role in neurological and psychiatric disorders.** *Neurosci Biobehav Rev* 2021, **130:**33-60.http://doi.org/10.1016/j.neubiorev.2021.08.011

48. Medrihan L, Cesca F, Raimondi A, Lignani G, Baldelli P, Benfenati F: **Synapsin II desynchronizes neurotransmitter release at inhibitory synapses by interacting with presynaptic calcium channels.** *Nat Commun* 2013, **4:**1512.http://doi.org/10.1038/ncomms2515

Figures

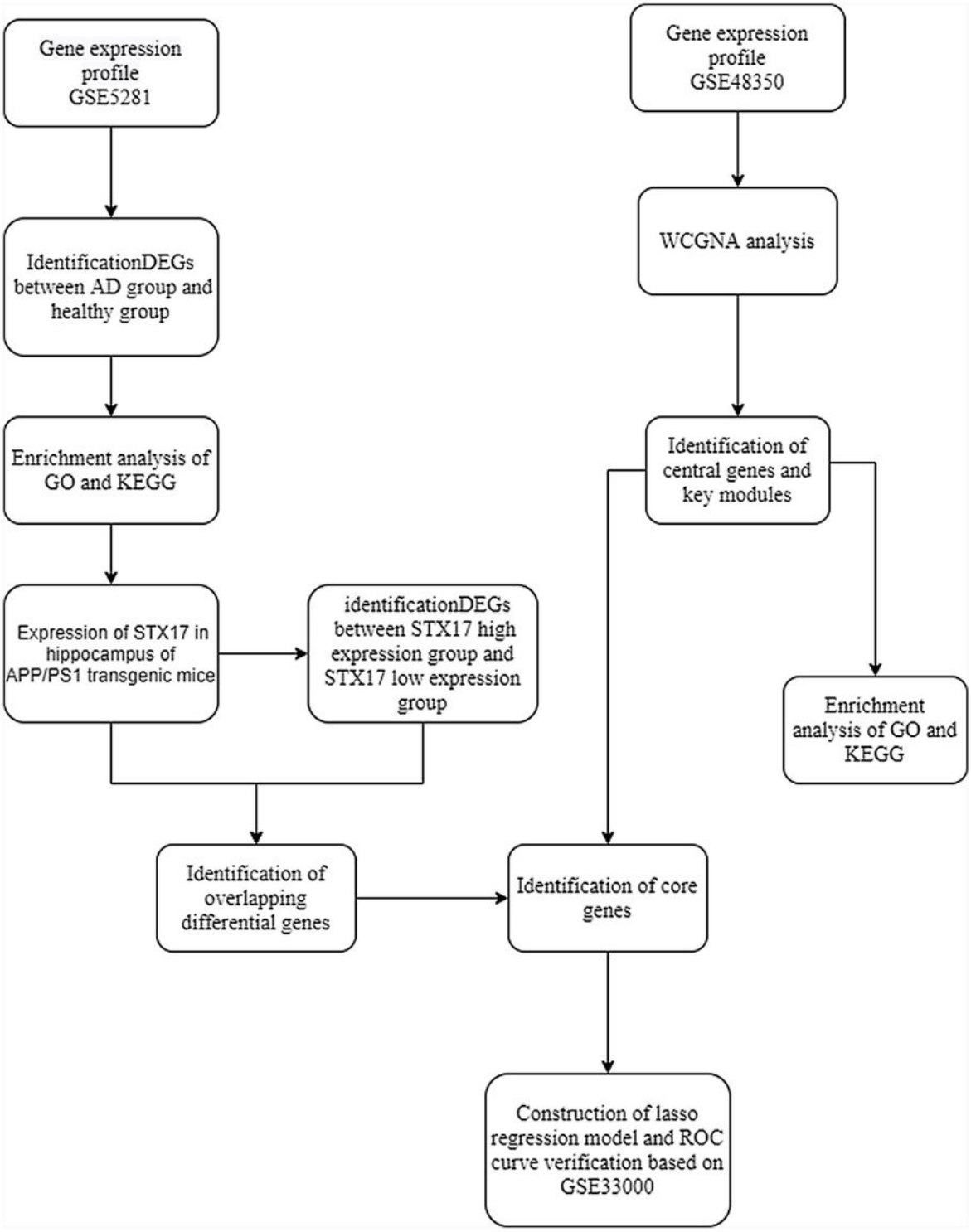

Figure 1

Study procedure flowchart.

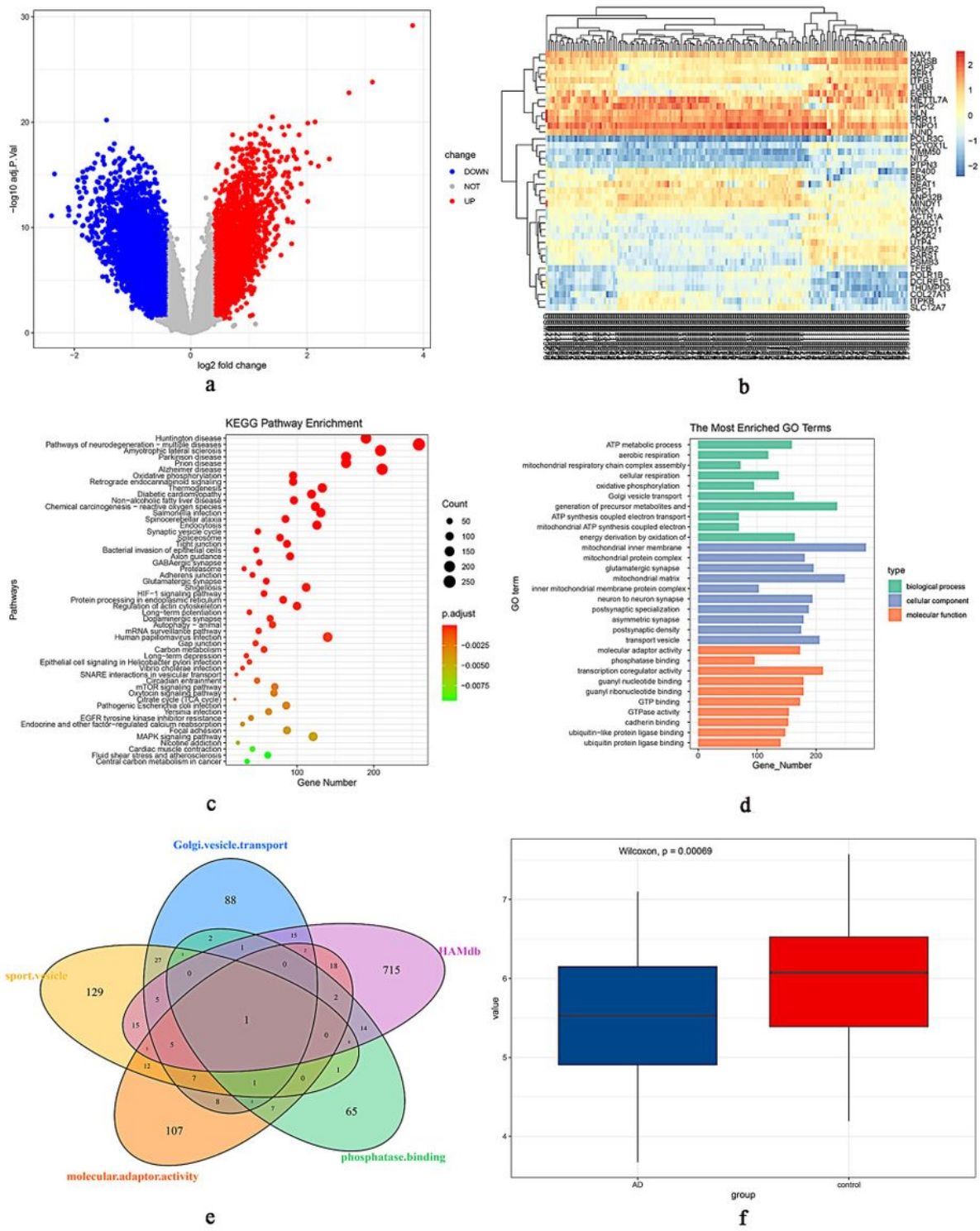

Figure 2

Screening of differentially expressed genes (DEGs). **a** Volcano plot of DEGs in Alzheimer's disease (AD) and healthy groups. Red are upregulated genes, blue are downregulated genes, and grey are genes without significant difference. **b** DEGs between the AD and health groups, of which 20 are significantly upregulated genes and 20 are significantly downregulated genes. **c** KEGG enrichment analysis of DEGs. **d**

GO enrichment analysis of DEGs. **e** Intersection between genes in 4 major entries and autophagy database. **f** STX17 expression is significantly reduced in AD patients.

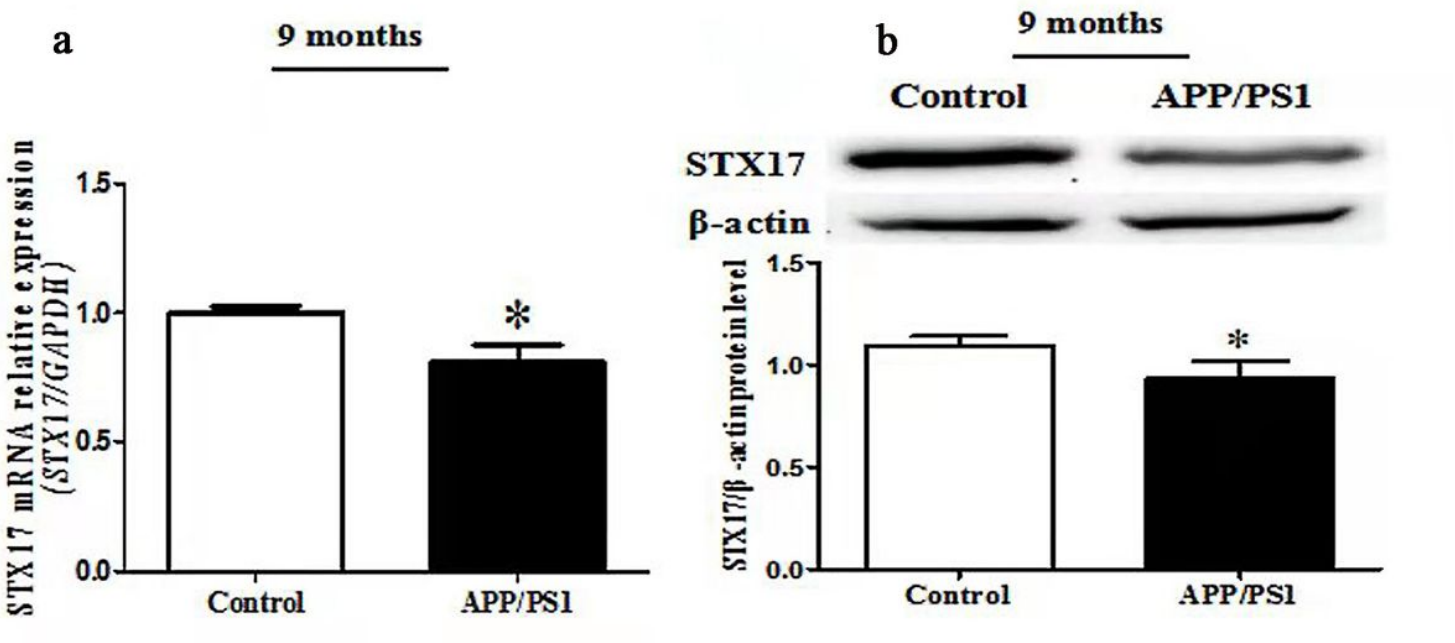

Figure 3

Expression of STX17 in hippocampal tissues in APP/PS1 mice. **a** STX17 mRNA expression in hippocampal tissues from APP/PS1 double transgenic mice and healthy mice. **b** STX17 protein expression in hippocampal tissues from APP/PS1 double transgenic mice and healthy mice.

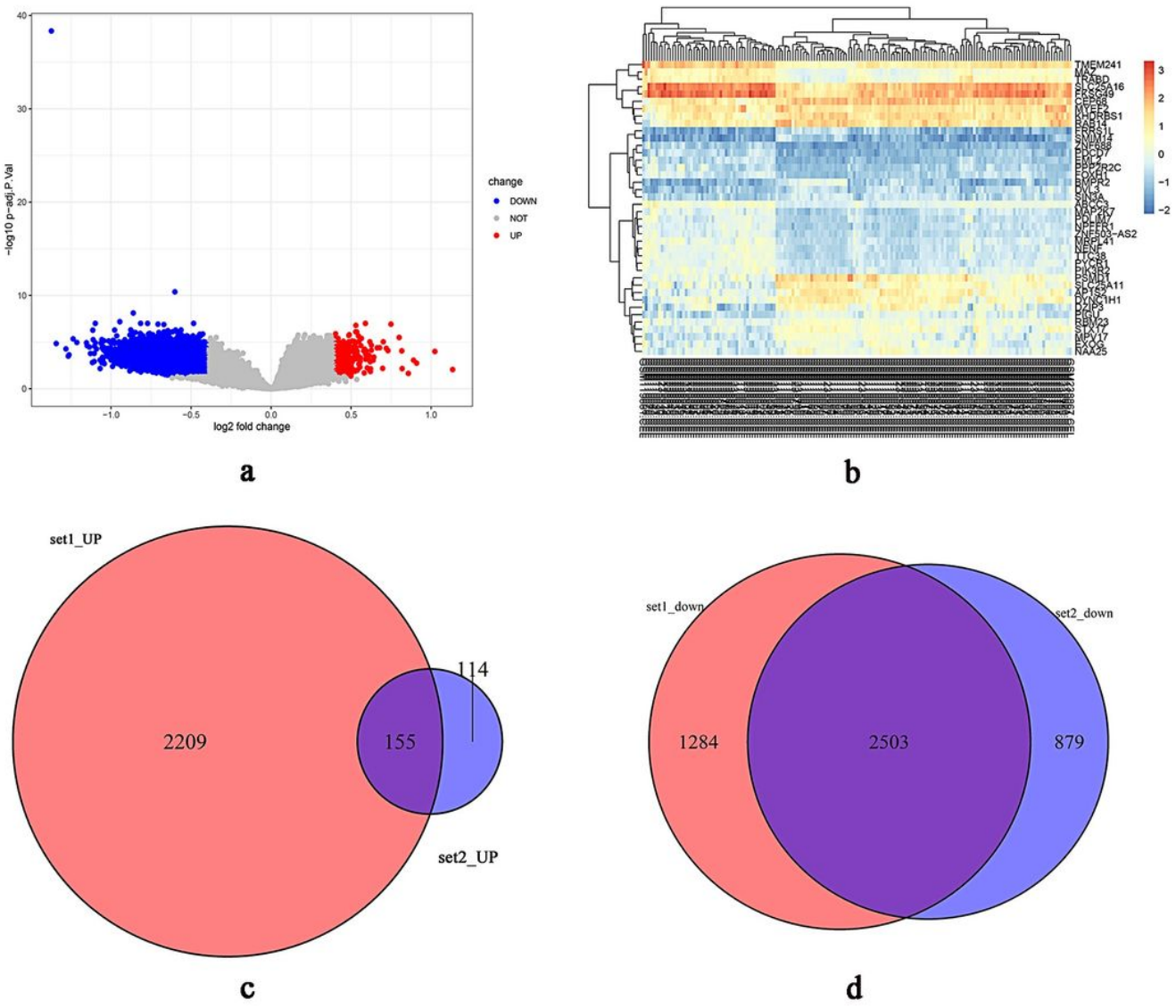

Figure 4

Differentially expressed genes (DEGs) in screening co-expression. **a** Volcano plot of DEGs in set 2. Red are upregulated genes, blue are downregulated genes, and grey are genes without significant difference. **b** DEGs in set 2, of which 20 are significantly upregulated genes and 20 are significantly downregulated genes. **c** There are 155 highly expressed genes in the intersection between sets 1 and 2. **d** There are 2503 genes with low expression in the intersection between sets 1 and 2.

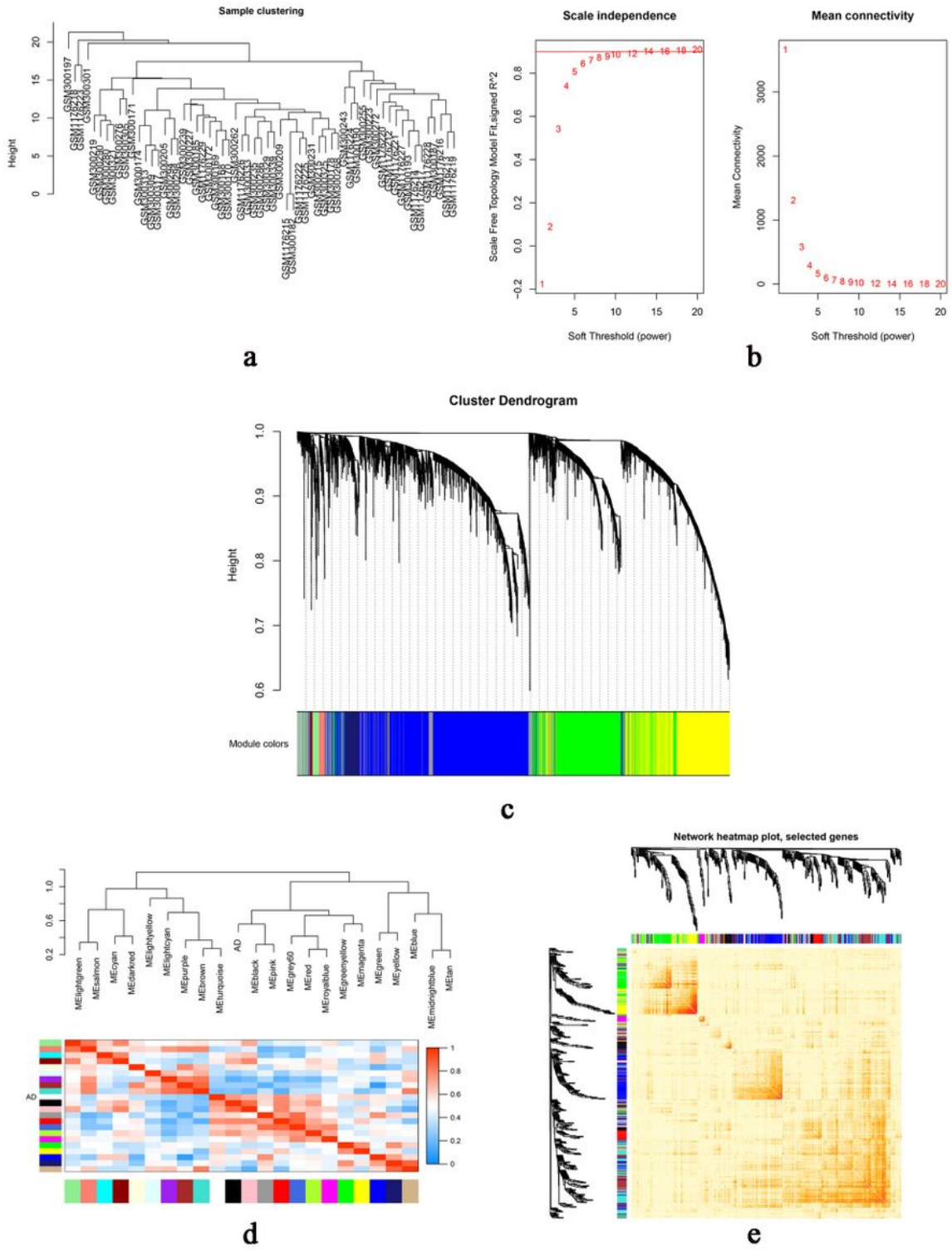

Figure 5

Construction of weighted gene co-expression network. **a** Cluster 3 of 18 Alzheimer's disease (AD) samples and 43 healthy samples. **b** Fit indexes of various soft thresholds to the scale-free network, mean connectivity of various soft thresholds. **c** Cluster trees of different measurement modules. **d** Heat map of characteristic gene trees and module adjacency Red shows high adjacency and blue show slow

adjacency. **e** Visualizable heat map of gene network (400 genes can be randomly selected), dark colors represent high gene overlap and light colors represent low gene overlap.

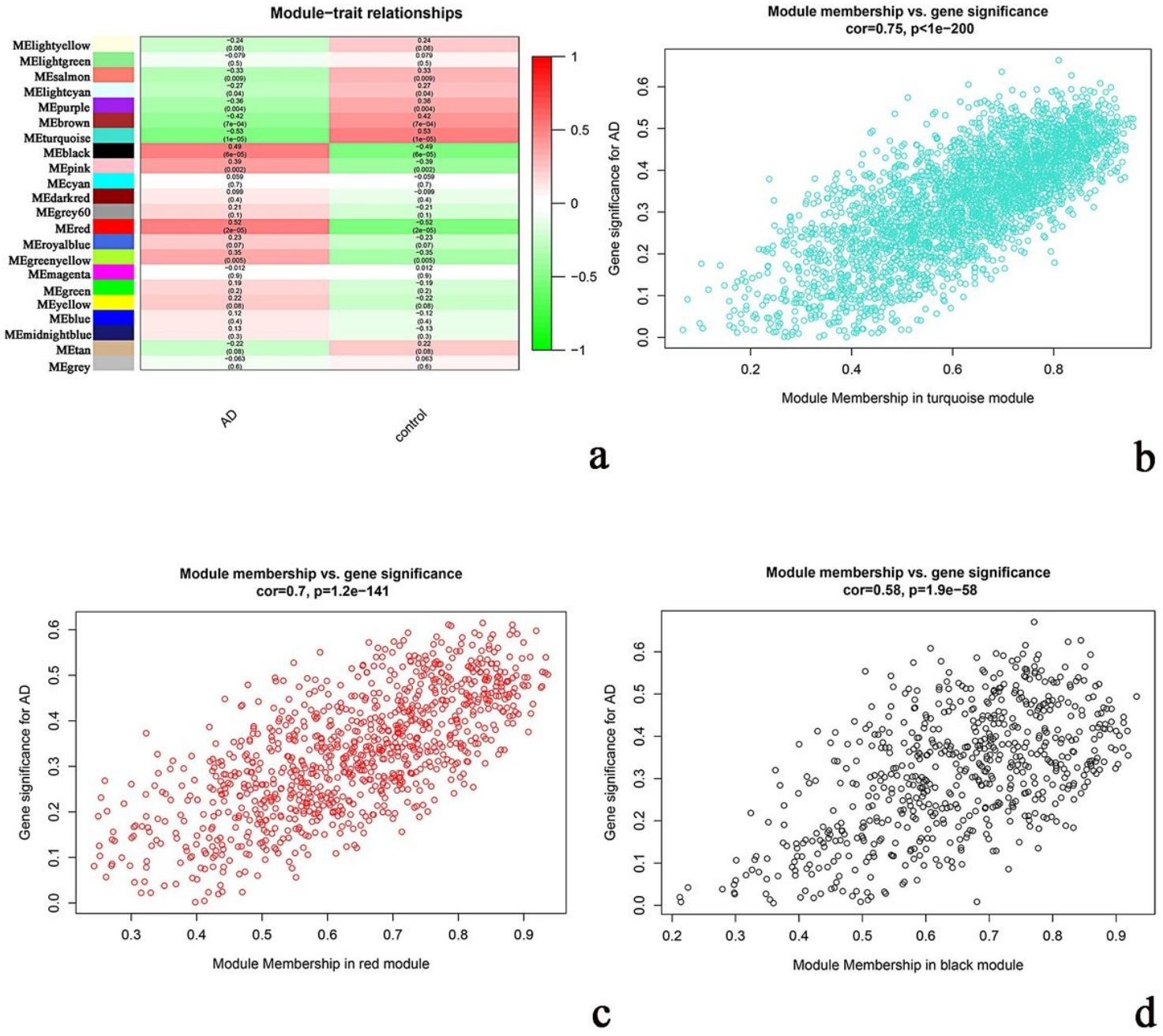

Figure 6

Construction of modules with most significantly correlation. **a** Correlation heat map of modules and traits, red represents positive correlation, green represent negative correlation. The darker the color, the stronger the correlation. The lighter the color, the weaker the correlation. **b** Correlation of turquoise module gene significance and module membership. **c** Correlation of red module gene significance and module membership. **d** Correlation of black module gene significance and module membership.

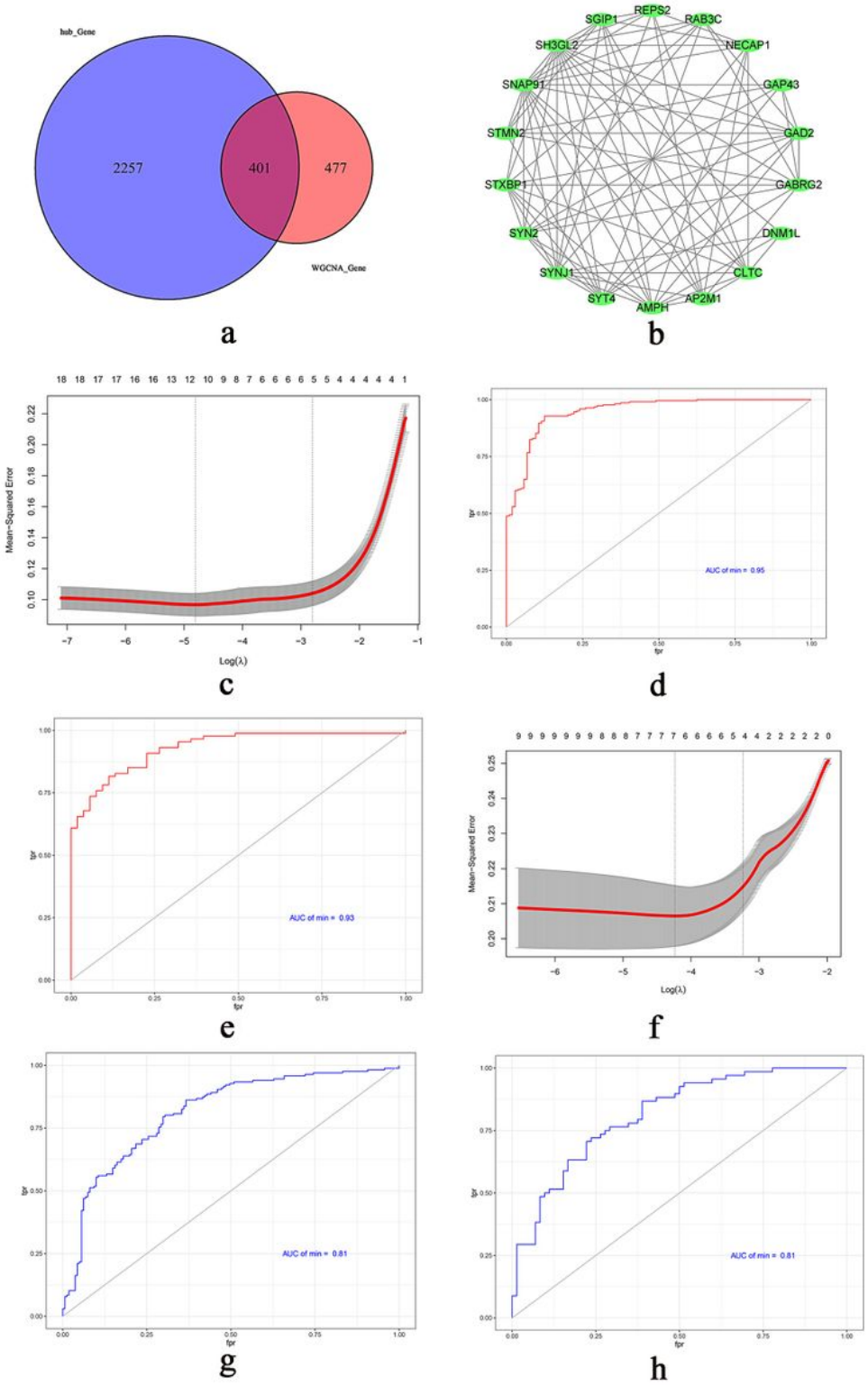

Figure 7

Core gene screening and construction of the LASSO regression model. **a** Intersection of central genes and hub genes in the WGCNA significance module. **b** 18 core genes were obtained through PPI network construction analysis. **c** Construction of Alzheimer's disease (AD)-related LASSO model. **d** ROC curve of genes in the AD training set. **e** ROC curve of genes in the AD test set. **f** Construction of STX17 high-low

expression-related LASSO model. **g** ROC curve of genes in the STX17 high-low expression training set. **h** ROC curve of genes in the STX17 high-low expression test set.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xlsx
- Additionalfile2Fig.S1..tif
- Additionalfile3.xlsx
- Additionalfile4Fig.S2..tif
- Additionalfile5Fig.S3.tif
- Additionalfile6Fig.S4.tif