

Research Article

Identify Key Biomarkers Predicting Comorbidity of Alzheimer's Disease and Metabolic Syndrome Through Machine Learning, Single-cell Analysis, and Bioinformatics Strategies

Xu Deng¹, Kunpeng Du^{2*}

¹Department of Anesthesiology, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China

²Department of Radiation Oncology, Zhujiang Hospital of Southern Medical University, Guangzhou, Guangdong Province, China

*Correspondence to: Kunpeng Du, Department of Radiation Oncology, Zhujiang Hospital of Southern Medical University, Guangzhou, 253 Gongye Dadao Zhong, Haizhu District 510280, Guangdong Province, China; Email: dkp321098@smu.edu.cn

Abstract

This study investigates the interplay between metabolic syndrome (MetS) and Alzheimer's disease (AD) by exploring shared immunometabolic dysregulation. We aimed to identify diagnostic biomarkers and elucidate underlying mechanisms. Through integrated analysis of GEO microarray data, we identified differentially expressed genes and applied WGCNA to pinpoint disease-relevant gene modules. Machine learning algorithms refined the selection to 33 pivotal hub genes (e.g., *SARS*, *CREBBP*, *RBM25*). Functional enrichment analyses, supported by ssGSEA and scRNA-seq validation, revealed that these hub genes are significantly associated with alterations in immune infiltration and critical metabolic pathways, including fatty acid metabolism and oxidative phosphorylation. Our findings establish a novel diagnostic signature for MetS-related AD, offering valuable insights for clinical diagnosis and the development of targeted therapeutic interventions.

Keywords: hub genes, metabolic syndrome, Alzheimer's disease, machine learning, inflammation, metabolism, single-cell RNA sequencing

Received: June 26, 2025

Revised: August 20, 2025

Accepted: October 27, 2025

Published: November 10, 2025

Copyright © 2025 The Author(s).

Published by Innovation Forever Publishing Group Limited. This open-access article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, sharing, adaptation, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Deng X, Du K. Identify Key Biomarkers Predicting Comorbidity of Alzheimer's Disease and Metabolic Syndrome through Machine Learning, Single-cell Analysis, and Bioinformatics Strategies. *Innov Discov*, 2025; 2(4): 18.

1 INTRODUCTION

Alzheimer's disease (AD) is a prevalent neurodegenerative disorder, hallmarked by a progressive decline in cognitive faculties such as language, judgment, memory, and decision-making^[1,2]. As the global population continues to age, the prevalence of AD is on the rise, with current estimates suggesting that 475 million individuals are affected worldwide — a number anticipated to soar to 152 million by 2050^[3]. This escalating trend underscores the pressing public health challenge that AD presents. Despite ongoing research, the precise pathological mechanisms of AD remain elusive. However, a growing body of evidence suggests a correlation between AD and brain injuries associated with comorbid conditions such as obesity, diabetes, hypertension, and cardiovascular diseases^[4-6]. Metabolic syndrome (MetS),

characterized by a constellation of symptoms including abdominal obesity, hypertension, hyperglycemia, and dyslipidemia, is increasingly recognized as a significant comorbidity in the context of AD. The pathogenesis of MetS is intricately linked to dysregulated adipocyte differentiation and the accumulation of visceral fat, triggering the release of proinflammatory cytokines and culminating in peripheral insulin resistance^[7]. While previous research on MetS has primarily focused on peripheral tissues, the impact of MetS on brain function is gradually gaining attention.

Historically, AD and MetS were regarded as distinct clinical entities. However, several epidemiological studies have reported that MetS is an independent risk factor for dementia^[8]. A recent population-based investigation, centered on individuals presenting with vascular risk

factors-hypertension, diabetes, and dyslipidemia-has delineated a pronounced association between these conditions and an elevated risk of AD onset, attributed to the enhanced generation of amyloid-beta (A β) peptides^[9]. Preclinical research has further unveiled that cerebral ischemia, precipitated by hypoxic conditions, intensifies A β synthesis, thereby exacerbating both amyloidogenic and non-amyloidogenic cascades implicated in AD pathogenesis^[10]. Obesity, with repercussions beyond metabolic dysregulation, has been implicated in the neuropathology of AD. Obesity-induced systemic and neuroinflammation disrupts the blood-brain barrier (BBB) and stimulates A β (β -amyloid) production by activating amyloid precursor protein and β -secretase^[11]. Pugazhenth et al. have elucidated that MetS-induced microglial activation at the BBB interface results in mitochondrial dysfunction and heightened oxidative stress via the deacetylation of SIRT3 and the subsequent activation of pivotal metabolic enzymes and transcriptional regulators, a critical pathway in AD etiology^[12]. Animal studies have also revealed a correlation between MetS and cognitive dysfunction. Dinel et al. observed behavioral changes and increased central inflammatory markers in MetS mouse models^[13]. These findings collectively underscore the intricate relationship between MetS and cognitive dysfunction, highlighting the need for further elucidation of the underlying pathophysiological mechanisms. In conclusion, a robust correlation exists between MetS and AD, yet the underlying pathogenic mechanisms are not fully elucidated, warranting continued investigation.

In recent years, integrated bioinformatics analysis has been used to identify novel genes associated with various diseases, which can serve as diagnostic and prognostic biomarkers. Despite a surge in research exploring the nexus between AD and MetS, the pivotal molecules and the intrinsic mechanisms underpinning their interplay remain elusive. Therefore, we employed published RNA transcriptome data (GSE5281, GSE132903, and GSE98895) and single-cell sequencing data (GSE188545) to investigate diagnostic and prognostic biomarkers in both MetS and AD using bioinformatics approaches. This study harnessed bioinformatics methodologies to identify common differentially expressed genes (DEGs) between AD and MetS, subsequently employing Weighted Gene Co-expression Network Analysis (WGCNA) to cluster and prioritize these genes. We leveraged an array of machine learning algorithms, including random forest (RF), lasso regression, and Extreme Gradient Boosting (XGBoost), to discern hub genes from the DEGs and construct predictive models, with the goal of forecasting both diseases with an optimized subset of genes.

To further understand the relationship between immunity and metabolism in these two diseases, we employed single-sample gene set enrichment analysis (ssGSEA) to analyze immune infiltration and metabolism-related pathways. By linking hub genes to immune and

metabolic pathways, we aimed to explain the common pathogenic mechanisms of AD and MetS. To validate the accuracy of epigenetic mechanisms, we conducted cell-level validation using single-cell sequencing, revealing significant differences in immune cells and hub gene expression between AD and normal cells. Additionally, we explored metabolic and immune cell differences associated with hub genes at the cellular level, providing a comprehensive explanation of the metabolic and immune pathogenic mechanisms of both diseases. This study offers potential solutions for the joint treatment of AD and MetS.

2 MATERIALS AND METHODS

2.1 Data Collection

The research design is illustrated in the flowchart shown in Figure 1. We searched the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) using keywords "Alzheimer's disease" and "metabolic syndrome" to retrieve relevant gene datasets. The inclusion criteria were as follows: (1) array-based human expression profiles; (2) to ensure the accuracy of the results, the sample size in each group should be at least 5. Finally, three accessible mRNA datasets (test datasets: GSE5281 and GSE98895, and validation dataset: GSE132903) as well as one single-cell dataset GSE188545 were included. The GSE5281 dataset comprises RNA sequencing data from brain tissues of AD patients and normal aged controls. The GSE132903 dataset includes bulk RNA-seq data obtained from the middle temporal gyrus (MTG) of AD patients and non-demented controls. The GSE98895 dataset contains transcriptomic sequencing data comparing peripheral blood mononuclear cells between MetS patients and healthy controls. The GSE188545 dataset contains scRNA-seq data comparing the MTG tissues between AD patients and healthy controls, where the MTG represents a cortical region severely affected by AD pathology.

2.2. Data Processing and Identification of DEGs

To identify DEGs between healthy and diseased groups, the R package "LIMMA" was employed. Differential analysis of the processed gene expression matrix was conducted using the LIMMA R suite^[14]. The specific criteria for the analysis were: The cutoff criteria for the AD data sets in GSE5281 included adjusted p -value<0.05 and $|\log$ fold change (FC)|>1. Similarly, the threshold requirements for the MetS data sets in GSE98895 also included adjusted p -value<0.05 and $|\log$ fold change (FC)|>1. AD validation data sets GSE132903 also included adjusted p -value<0.05 and $|\log$ fold change (FC)|>1. The "ggplot2" package was used to visualize the results from the two datasets. Heatmaps and volcano plots were generated to intuitively represent the DEGs. This approach allowed for the identification and

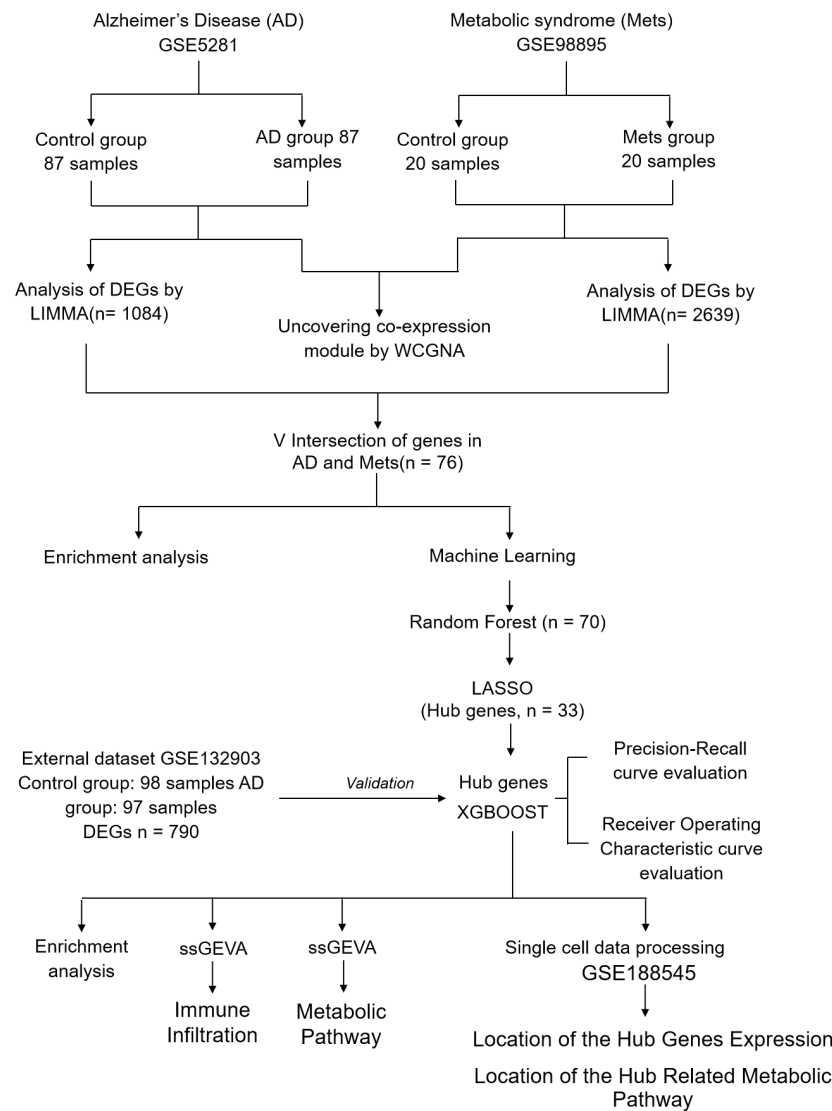


Figure 1. Flowchart of the Study Design.

visualization of genes that exhibited significant expression differences between diseased and healthy samples in both AD and MetS datasets.

2.3. WGCNA and Module Gene Selection

WGCNA can be used to identify highly coordinated groups of genes, classify genes into modules by analyzing the association relationships between genes, and then find molecular characteristics of specific phenotypes through correlation analysis between specific modules and sample phenotypes^[15]. Firstly, the median absolute deviation of each gene is determined, and the bottom 50% of genes with the lowest values are removed. Secondly, poor-quality genes and samples are removed using the function of good samples' genes, and a scale-free co-expression network is established. Thirdly, an appropriate "soft" threshold power is determined to calculate the median proximity, and then the proximity values are transformed into a topological overlap matrix and the scale-free topology fit index and mean connectivity are determined. Fourthly, modules are detected using hierarchical clustering and the dynamic

tree cut function. Finally, gene significance and module membership are calculated, and the corresponding module gene information is extracted for further analysis.

We used the R package "WGCNA" throughout the analysis and constructed the network. Samples were organized into clusters to detect and eliminate important outliers. Then, the co-expression network was developed using the automatic network approach. Next, we determined the soft thresholding power for calculating adjacency in the co-expression network. To select important key modules as candidate modules, we performed hierarchical clustering and dynamic tree cut functions, while evaluating gene significance and module membership to associate modules with clinical features. Modules with the highest Pearson's correlation of module membership and $P < 0.05$ were selected as the matching modules, and the gene lists of each module were adopted for further study.

2.4. Enrichment Analysis

The Kyoto Encyclopedia of Genes and Genomes

(KEGG) and Gene Ontology (GO) are comprehensive knowledge resources for the analysis of gene functions and biological interactions^[16]. To further reveal the molecular mechanisms underlying the progression of MetS and AD comorbidities, GO and KEGG pathway enrichment analyses were performed on the intersected genes using the "clusterProfiler" R package (v 4.6.2). A statistical significance threshold of $P < 0.05$ was applied.

2.5 Screening Hub Genes and Establishing Prediction Models Using Various Machine Learning Algorithms

Machine learning is an important tool for artificial intelligence and has extensive applications in the biomedical field. To screen for important diagnostic biomarkers that predict the progression of AD related to MetS, RF recursive feature elimination is a feature selection method that utilizes RF to evaluate the importance of each feature and recursively removes the least important features until the desired number of features is left^[17]. Least Absolute Shrinkage and Selection Operator (LASSO) is a linear regression method that uses L1 regularization to shrink some unimportant coefficients to zero^[18]. XGBoost is a gradient boosting algorithm that utilizes decision trees as base learners. It is widely used in data mining, feature selection, and other fields^[19]. Therefore, we used the "RF," "lasso," and "XGBoost" algorithms in R software to develop corresponding deep machine learning analyses. The specific parameters are as follows: (1) RF (1,000 trees, seed=245083) ranked genes by MeanDecreaseGini importance, retaining those with scores > 0 ; (2) LASSO regression ($\alpha = 1$, 10-fold CV, seed=123456) with optimal lambda selection further refined the feature set. The final XGBoost model was trained using Bayesian-optimized hyperparameters and evaluated through stratified 10-fold cross-validation (seed=176). Through cross-validation, Bayesian optimization, and other methods for model establishment and optimization, we predicted the training sets of AD and MetS separately, observed the sensitivity, and simulated the AD training set to test the sensitivity.

2.6 The Receiver Operating Characteristic (ROC) Curve Analysis and Expression Analysis

We perform the ROC curve analysis to evaluate the diagnostic value of the hub genes. The ROC curve's proximity to the upper left corner increases the true positive rate and sensitivity of the test while decreasing the false positive rate and misdiagnosis rates, indicating a better diagnostic value of the genes for AD. The ROC curve was assessed and drawn using the "pROC" R package^[20]. Area under curve (AUC) is the region that the ROC curve encompasses, and axis and hub genes were considered helpful in screening the disease if their AUC was greater than 0.7.

2.7 Immune Landscape and Metabolic Pathways Analysis

The composition and abundance of 22 types of immune cells can be estimated from the transcriptome of AD and MetS compared with control samples in the GSE5281 and GSE98895 datasets using "ssGSEA." The correlations between the expression of AD and MetS-related hub genes with immune cell infiltrations were investigated, as well as their respective correlations. The same approach was used to explore the relationships between the samples, hub genes, and 10 metabolic pathways.

2.8 Single-cell Analysis

The single-cell RNA sequencing (scRNA-seq) data from the putamen was obtained from the GSE161045 dataset. To analyze this data, we employed the R package "Seurat" for quality control, normalization of expression data, dimension reduction, and clustering^[21]. Cell type annotation was automatically performed using the "SingleR" package^[22]. The results of the analysis were visualized using the R packages "ggplot2" and "cowplot".

Initially, we normalized the combined data through log-normalization and identified the top 2,000 highly variable genes using the FindVariableFeatures function, which relies on the variance stabilization transformation ("vst"). Additionally, all genes were scaled using the ScaleData function, and the RunPCA function was applied to reduce the dimensionality of PCA for the selected 2,000 highly variable genes. We chose a dimensionality of 50 and clustered the cells using the "FindNeighbors" and "FindClusters" functions (with a resolution of 0.1) to identify distinct cell clusters. Next, we selected the top 50 principal components and further reduced the dimensionality using the UMAP method. UMAP is a dimensionality reduction technique that assumes that the available data samples are uniformly distributed in a topological space (manifold). This method approximates and maps these limited data samples to a low-dimensional space. In essence, the UMAP algorithm is similar to t-SNE, which maps high-dimensional probability distributions to a low-dimensional space for dimensionality reduction. Based on the principles of manifold theory and topology algorithms, UMAP reduces the dimensionality of high-dimensional data to generate input features for other classification models.

Finally, we utilized the FindAllMarkers function to identify hub genes. Additionally, we explored the metabolic status of genes of interest.

2.10 Statistical Analysis

All statistical analyses of bioinformatics studies in this study were conducted using R (version 4.4.0) software. The differences between the groups were tested using

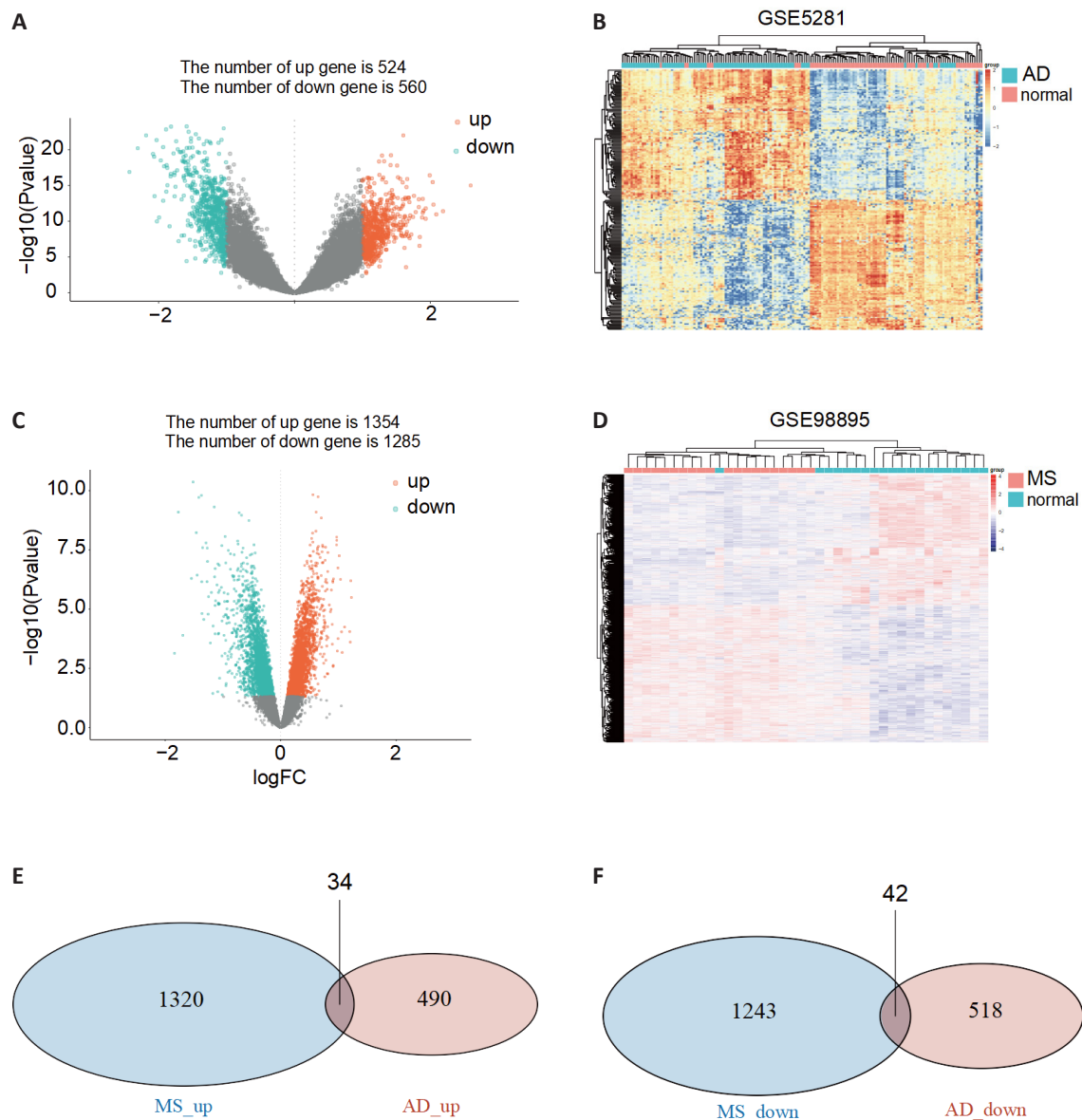


Figure 2. Identification of the DEGs Associated with AD and MetS Dataset (GSE5281, GSE98895). A: The volcano plot of AD dataset. B: The heatmap plot of up- and down- regulated DEGs of AD. C: The volcano plot of MetS dataset. D: The heatmap plot of up- and down-regulated DEGs of MetS. E: The intersection up-regulated DEGs of AD and MetS. F: The intersection down-regulated DEGs of AD and MetS.

a nonparametric Wilcoxon signed-rank test. Correlation analysis was performed using Spearman's correlation. In comparison, $P < 0.05$ was considered statistically significant ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$).

3 RESULTS

3.1 Identify the DEGs between AD and MetS

In Figure 1, we demonstrate the key processes of our overall study. To investigate the interaction between AD and MetS, we downloaded the corresponding sequencing data of samples from the GEO database. Using the "limma" package in R, we identified 1,084 DEGs in the AD dataset (GSE5281), including 524 up-regulated genes and 560 down-regulated genes, with the screening criteria of $|\text{Log}_2 \text{ Fold Change}| > 1$ and $|\text{adj.P.Val.}| < 0.05$ (Figure 2A and 2B). Similarly, in the MetS dataset (GSE98895), we identified 2,639 DEGs, including 1,354

upregulated genes and 1,285 downregulated genes, with the screening criteria of $|\text{Log}_2 \text{ Fold Change}| > 0$ and $|\text{adj.P.Val.}| < 0.05$ (Figure 2C and 2D). These findings are presented using volcano plots and heatmaps. Subsequently, we used a Venn diagram to showcase the 76 common genes between AD and MetS, consisting of 42 down-regulated genes and 34 up-regulated genes (Figure 2E and 2F).

3.2 Selecting Key Modules Related to Two Diseases through WGCNA

By combining the datasets (GSE5281 and GSE98895) using a batch correction method and setting $\text{cutHeight} = 150$ to filter out samples that are more discrete, the optimal soft-power value for the combined dataset was determined to be 5 based on scale independence and mean connectivity (Figure 3A). After merging the modules, a total of 8 co-expression modules were identified, each corresponding to a different color (Figure 3B and 3C).

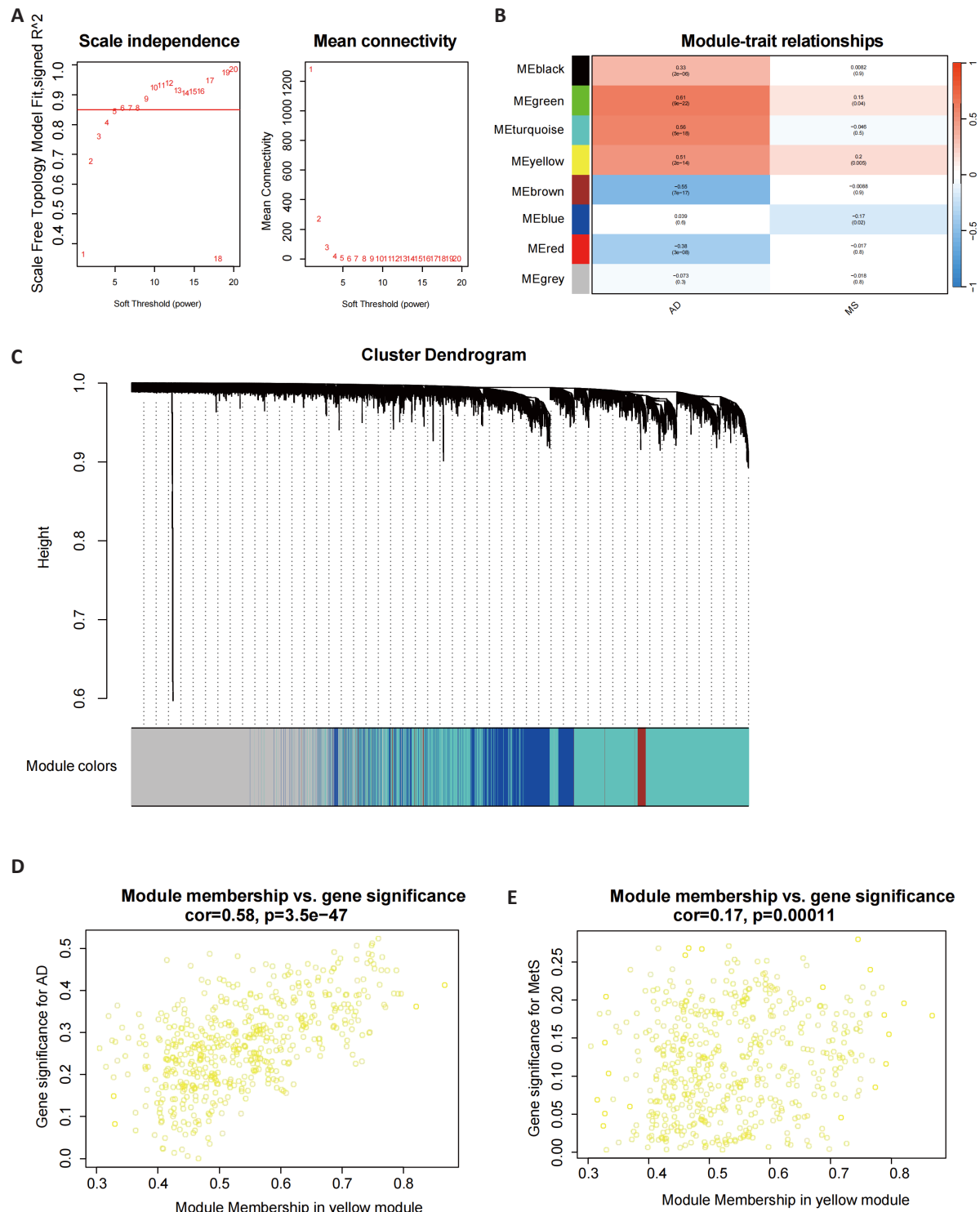


Figure 3. Selecting Key Modules Related to Two Diseases through WGCNA. A: Best soft- power value = 5 was regarded the according to scale independence and average connectivity. B: The heatmap to visualize the correlation between the different modules in the AD and MetS groups. C: A gene clustering tree with multiple divided modules. Different clusters were attached with different colours. D and E: The scatter plots of module membership and gene significance for the AD and MetS groups.

Given previous research findings indicating a positive correlation between the progression of AD and MetS, we considered the correlation analysis between modules and traits while setting the correlation coefficient threshold as module-Trait P value < 0.01 . Consequently, we selected the yellow module from these modules, which contained 510 genes (Table S1). The correlation coefficient of the module membership in yellow module and gene significance for AD is 0.58. The correlation coefficient of the module membership in yellow module and gene

significance for MetS is 0.17 (Figure 3D and 3E).

3.3 Functional Enrichment Analysis of DEGs

The Gene Set Enrichment Analysis (GSEA) of the differential gene expression profiles between the AD group and the normal group revealed significant enrichment in multiple pathways, including ECM-receptor interaction, Notch signaling pathway, Chronic myeloid leukemia, p53 signaling pathway, and Hippo signaling

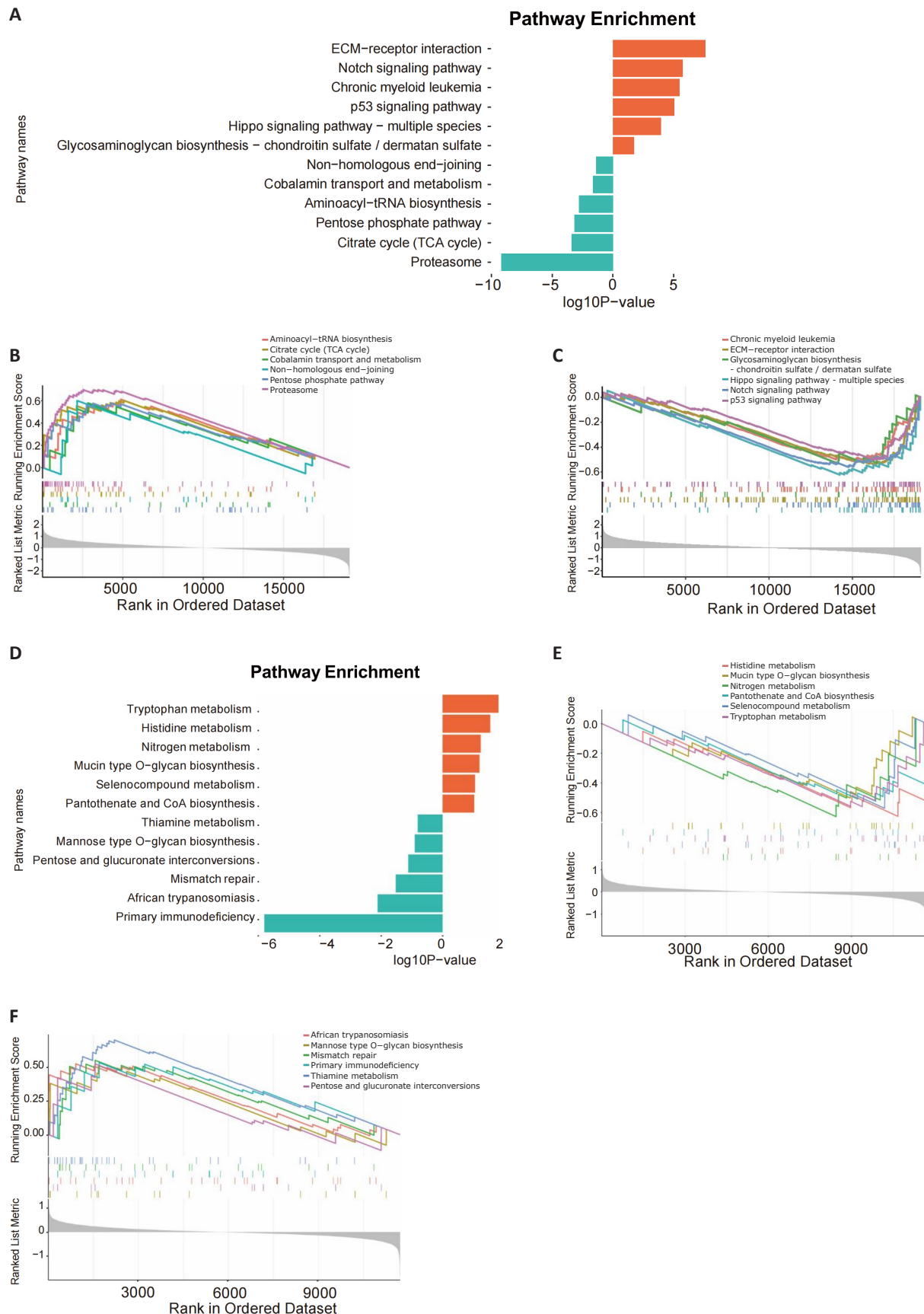


Figure 4. Functional Enrichment Analysis. A-C: The GSEA of the differential gene expression profiles between the AD group and the normal group revealed significant enrichment in multiple pathways. D-F: The GSEA of the differential gene expression profiles between the MetS group and the normal group revealed significant enrichment in multiple pathways.

pathway (Figure 4A). Similarly, enrichment analysis of the differential genes between the MetS and normal groups identified Tryptophan metabolism, Histidine metabolism, Nitrogen metabolism, Mucin type O-glycan

biosynthesis, Selenocompound metabolism (Figure 4D). We also performed enrichment analysis separately for up-regulated and down-regulated genes (Figure 4B, 4C, 4E and 4F).

3.4 Identifying Candidate Core Genes through Multi-machine Learning Approaches

Three machine learning strategies, including RF, LASSO regression, and the machine learning algorithm based on the Gradient Boosting framework (XGBoost), were used to screen candidate core biomarkers and establish disease prediction models. The RF algorithm was applied to select disease-specific genes by constructing multiple decision trees and integrating them to improve prediction accuracy and reduce the risk of overfitting. Genes with importance scores greater than 0 were selected, resulting in a total of 70 genes (Figure 5A and 5B). For LASSO regression, we observed that 33 of these 70 genes achieved the lowest binomial deviation on the curve (Figure 5E and 5F). The gene set includes: *SARS*, *CREBBP*, *RBM25*, *SMC3*, *TNFRSF10B*, *SMA5*, *SAMD4A*, *SMOX*, *ANKRD12*, *C1orf123*, *PHB*, *ANXA6*, *DDIT4*, *PPARD*, *RBBP4*, *LSM7*, *CD44*, *DDX50*, *COMMD4*, *CORO1A*, *ANKDD1A*, *WDR70*, *CMTM7*, *POGK*, *CDC37*, *CCT7*, *PSMC4*, *TMEM123*, *EXOSC5*, *CLPP*, *MIF*, *EPAS1*, and *COPE*. To further validate the precision and recall rates, we scored the model using a Precision-Recall Curve (PR Curve) and evaluated the sensitivity and specificity of the model, constructing a ROC Curve evaluation. Some of the genes are now presented. Finally, we used XGBoost to predict disease models based on the results of the first two machine learning algorithms, adopting cross-validation, Bayesian optimization, and parameter tuning to validate the model predictions within the datasets of two diseases. The optimal solution in 100 rounds of predictions was selected in the 88th round, and the final established model was assessed as follows: AD training set: ROC AUC: 0.989; PR AUC: 0.991 (Figure 5E), MetS training set: ROC AUC: 1; PR AUC: 1 (Figure 5F). To further validate the generalizability of our diagnostic model, we performed external validation using an independent dataset (GSE132903, AD cohort). The results demonstrated robust performance with an ROC AUC of 0.74 and PR AUC of 0.746 (Figure S1), indicating both high accuracy and satisfactory generalizability of our diagnostic model.

3.5 Immune Landscape Analysis

Based on the results of enrichment analysis, the ultimate core genes are likely to be involved in the immune-related mechanisms of AD and MetS progression. Therefore, the correlation between these 33 diagnostic marker genes and immune cell infiltration in the diseases merits further exploration. To provide further insights into the pathogenesis of these two diseases, we utilized the "GSVA" package to calculate immune cell infiltration, observing the correlation and significance between samples and 28 types of immune cells. "SSGSEA" was used to assess the proportions of 28 immune cell types. In AD samples, Activated B cell, Effector Memory CD8 T cell, Immature B cell, Memory B cell, Type 17 T helper cell, and Myeloid-Derived Suppressor Cell (MDSC) were significantly upregulated; however, Activated CD4 T cell, Activated CD8

T cell, Eosinophil, CD56dim natural killer cell, and Effector Memory CD4 T cell levels were significantly decreased (Figure 6A and 6B). In MetS samples, Activated B cell, Activated CD8 T cell, CD56dim natural killer cell, Central Memory CD4 T cell, Immature B cell, Type 1 T helper cell, and MDSC were significantly upregulated; whereas, Activated CD4 T cell, Activated dendritic cell, CD56bright natural killer cell, Central Memory CD8 T cell, Gamma delta T cell, Immature dendritic cell, Memory B cell, Plasmacytoid dendritic cell, Type 17 T helper cell, and Type 2 T helper cell levels were significantly decreased (Figure 6D and 6E).

Next, the correlation between these 33 diagnostic marker genes and immune cell types was explored (Figure 6C and 6F). For instance, the expression of ANKRD12 in the AD dataset was negatively correlated with the proportion of Activated CD8 T cell, Activated CD4 T cell, Eosinophil, CD56dim natural killer cell, and Effector Memory CD4 T cell, while positively correlated with Activated B cell, Effector Memory CD8 T cell, Immature B cell, Memory B cell, Type 17 T helper cell, and MDSC. In the MetS dataset, its expression was positively correlated with CD56bright natural killer cell, Effector Memory CD4 T cell, Immature dendritic cell, Activated CD4 T cell, Type 17 T helper cell, and Type 2 T helper cell, while negatively correlated with Activated B cell, Activated CD8 T cell, Effector Memory CD8 T cell, Immature B cell, Natural killer cell, Monocyte, and MDSC (Figure 6F).

3.6 The Relationship between Shared DEGs and Metabolism

To further explore the correlation between diseases and metabolism, we used the GSVA package to analyze the metabolic pathways of shared DEGs. We focused on analyzing the relationship between nine metabolic pathways, including hypoxia, cholesterol homeostasis, adipogenesis, xenobiotic metabolism, fatty acid metabolism, oxidative phosphorylation, glycolysis, heme metabolism, and bile acid metabolism, and 33 core genes (Figure 7). In the AD dataset, the core genes had varying degrees of influence on the nine metabolic pathways (Figure 7A). Among them, DDIT4, TNFRSF10B, and SMOX were positively correlated with adipogenesis, xenobiotic metabolism, hypoxia, cholesterol homeostasis, and heme metabolism, and negatively correlated with oxidative phosphorylation and fatty acid metabolism. In the MetS dataset, the metabolic pathways affected by the core genes were mainly focused on adipogenesis, xenobiotic metabolism, fatty acid metabolism, and oxidative phosphorylation (Figure 7B). Among them, PPARD, ANKRD12, ANKDD1A, and DDIT4 were negatively correlated, while MIF, PHB, EXOSC5, PSMA5, COPE, COMMD4, and C1orf123 were positively correlated.

3.7 Single-cell Analysis

This study utilized single-cell RNA sequencing data

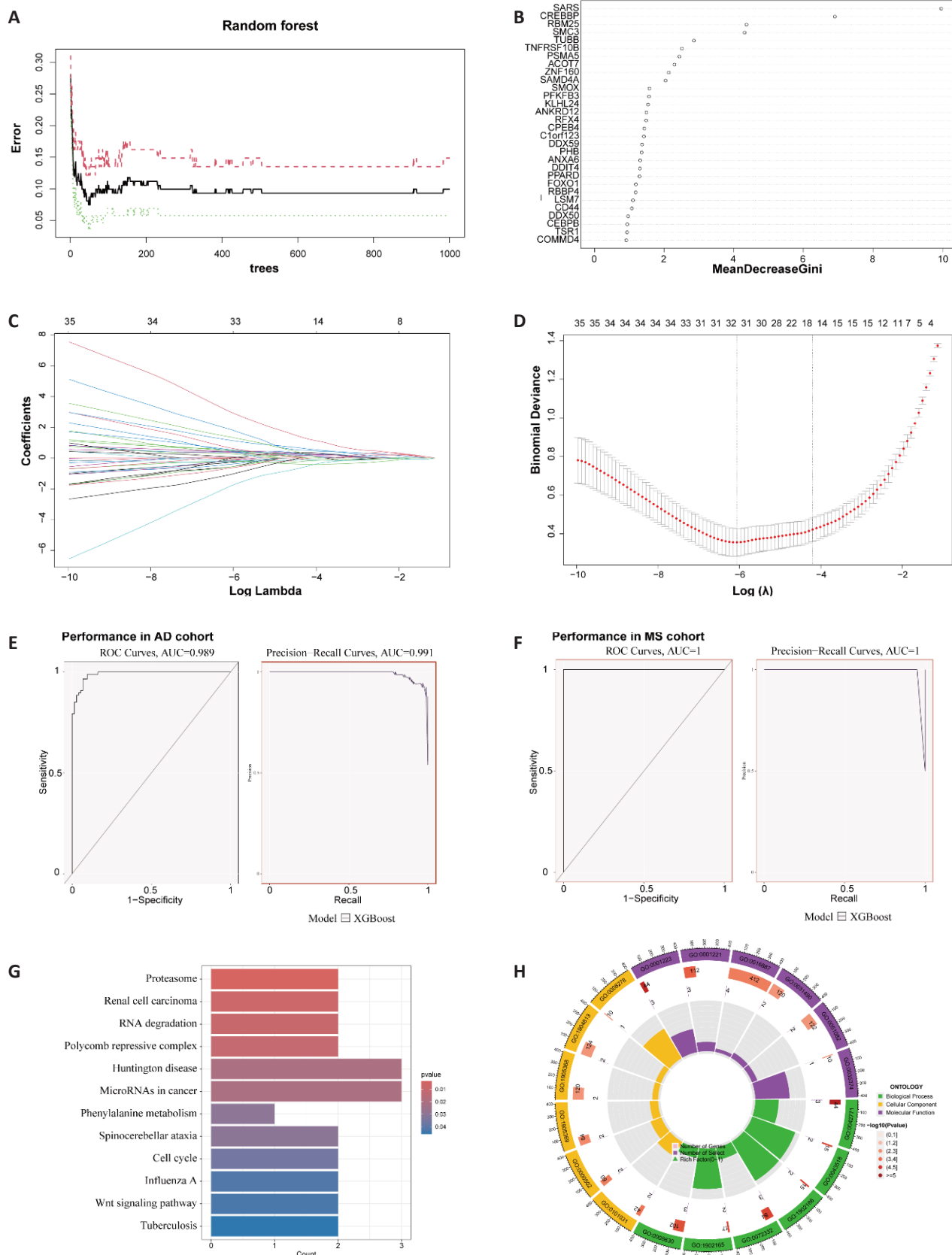


Figure 5. Selection of Candidate Hub DEGs Using Multi-machine Learning Algorithms. A and B: Gene selection via RF computational strategy. Genes with importance scores greater than 0 were selected. C and D: Features screening in the LASSO regression model based on 10-fold cross-validation, 33 genes corresponding to the lowest point of the curve were considered the most superior biomarkers for the forecast of AD progression in patients with MetS (E and F) Evaluate the accuracy and sensitivity of XGBoost prediction model with ROC and AUC curves. G: KEGG pathway analysis of the hub genes. H: GO functional analysis of the hub genes.

(GSE188545), which was collected from the temporal middle gyrus of 12 individuals with and without AD.

We employed Seurat for single-nucleus analysis and integrated the brain snRNA-seq data from six AD patients

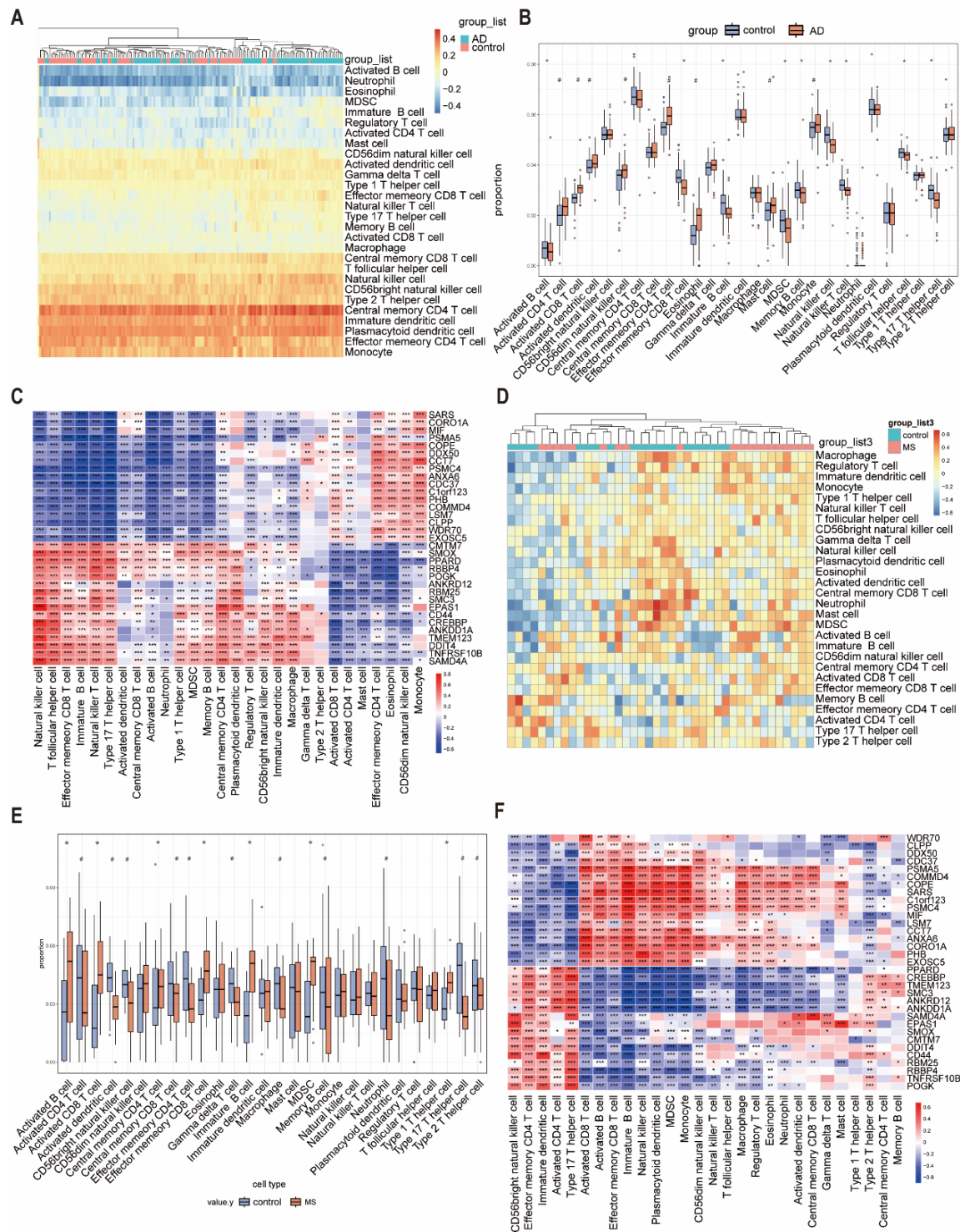


Figure 6. The Immune Cell Infiltration Levels in AD and MetS. A and D: Heatmap of the correlation between samples and immune cell types of AD and MetS. B and E: Sample grouping and immune cell box plot of AD and MetS. C and F: Heatmap of the correlation between hub genes and immune cell types of AD and MetS. * $P < 0.05$, ** $P < 0.001$, *** $P < 0.0001$.

($n=6$) and six age-matched healthy controls (HC, $n=6$). After quality control and data filtering, we analyzed 64,845 single-nucleus transcriptomes and clustered all the cells together (Figure S2A). The allowed gene counts and ribosomal ratios were set as minGene=200, maxGene=2,500, pctMT=10, and pctHB=3. We identified the top 2,000 most variable genes using the FindVariableFeatures function and annotated the top 10, which included OOE-AS1, FAM189A2, AC089983.1,

FLT1, CHI3L1, CCL4, CLDN5, RELN, CCL3 and HTR2C (Figure S2B). After dimensionality reduction and clustering, the combined tSNE and UMAP visualizations displayed 25 clusters of cells (Figure S2C and S2D).

Subsequently, we annotated the 5 clusters into biological cell types using the "SingleR" package. These clusters represented T_{cell}:CD4⁺_effector_memory, MetSC, Endothelial_cells: lymphatic: TNFa_48h,

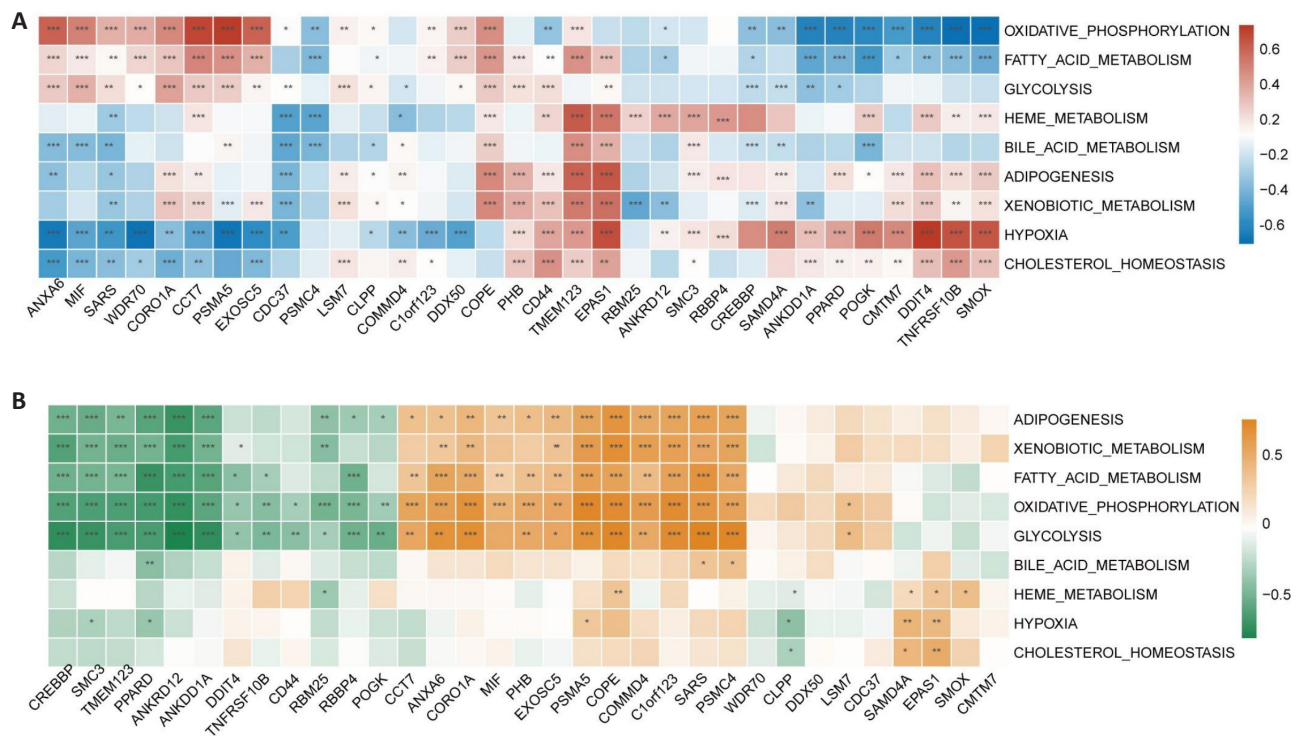


Figure 7. The Relationship between Shared DEGs and Metabolism. A: Heatmap of the correlation between samples and metabolism of AD. B: Heatmap of the correlation between samples and metabolism of MetS. * $P < 0.05$, ** $P < 0.001$, *** $P < 0.0001$.

Macrophage: monocyte-derived:IL-4/Dex/TGFb, and Astrocyte:Embryo nic_stem_cell-derived (Figure S3). The feature plot exhibited the expression patterns of 33 key biomarker genes across cell types and groups, revealing increased expression of SARS, PHB, ANXA6, LSM7, CCT7, MIF, EPAS1, and COPE in the AD group, while RBM25, PSMA5, ANKRD12, DDIT4, CD44, CREBBP, CORO1A, and CDC37 showed decreased expression (Figure 8A). Among these, we focused on the distribution of the genes of interest (CREBBP and ANKRD12) in each cluster, finding that both CREBBP and ANKRD12 were upregulated in T_cell:CD4+_effector_memory and T_cell:CD8+ in the AD group, with a lower number of T_cell:CD4+_effector_memory cells in AD compared to CT (Figure 8A). Figure 8B displays the expression levels of these 33 hub genes across different samples (6 AD group vs. 6 control group).

To further explore the relationship between these genes, immune cells, and metabolism, we selected metabolic pathways significantly correlated with the genes of interest from previous studies and searched for expression differences in single cells. Fatty acid metabolism was upregulated in the AD dataset ($W=9,033$, $P=0.00545$), and both CREBBP and ANKRD12 were upregulated in CD4+ effector memory T cells (Figure 9A-D). By connecting these three factors, we found that in AD patients, CREBBP and ANKRD12 promoted the expression of fatty acid metabolism in CD4+ effector memory T cells. Similarly, the oxidative phosphorylation pathway was downregulated in AD cells ($W=8,576$, $P=0.0006899$), and CREBBP and ANKRD12 in CD4+ effector memory T cells promoted the

expression of the oxidative phosphorylation pathway in AD patients (Figure 9E-G). Therefore, we can infer that downregulation of CREBBP and ANKRD12 promotes the expression of CD4+ effector memory T cells, fatty acid metabolism, and oxidative phosphorylation pathways, leading to the occurrence and progression of AD (Figure 9). These findings are consistent with the results obtained from our previous immune infiltration algorithms.

4 DISCUSSION

AD and MetS are formidable adversaries in the realm of global public health, with research indicating a complex interplay that influences their mutual progression, despite the mechanisms underlying this relationship remaining enigmatic. Through differential analysis and WGCNA to identify DEGs and important module genes, this investigation identified 76 genes that are shared between AD and MetS, subsequently subjecting these to an enrichment analysis. Furthermore, RF and Lasso regression methods were used to identify 33 diagnostic marker genes, which were corroborated by the XGBoost algorithm's diagnostic prowess in both AD and MetS training datasets.

To delve deeper into the related mechanisms of immune and metabolic interactions, we used the ssGEVA method to determine the relationships between samples, diagnostic marker genes, 28 immune cell types, and 9 metabolic pathways. This approach identified differences in immune cell infiltration and cell expression between disease and control samples, which regulate and affect metabolism, especially fatty acid metabolism. In addition

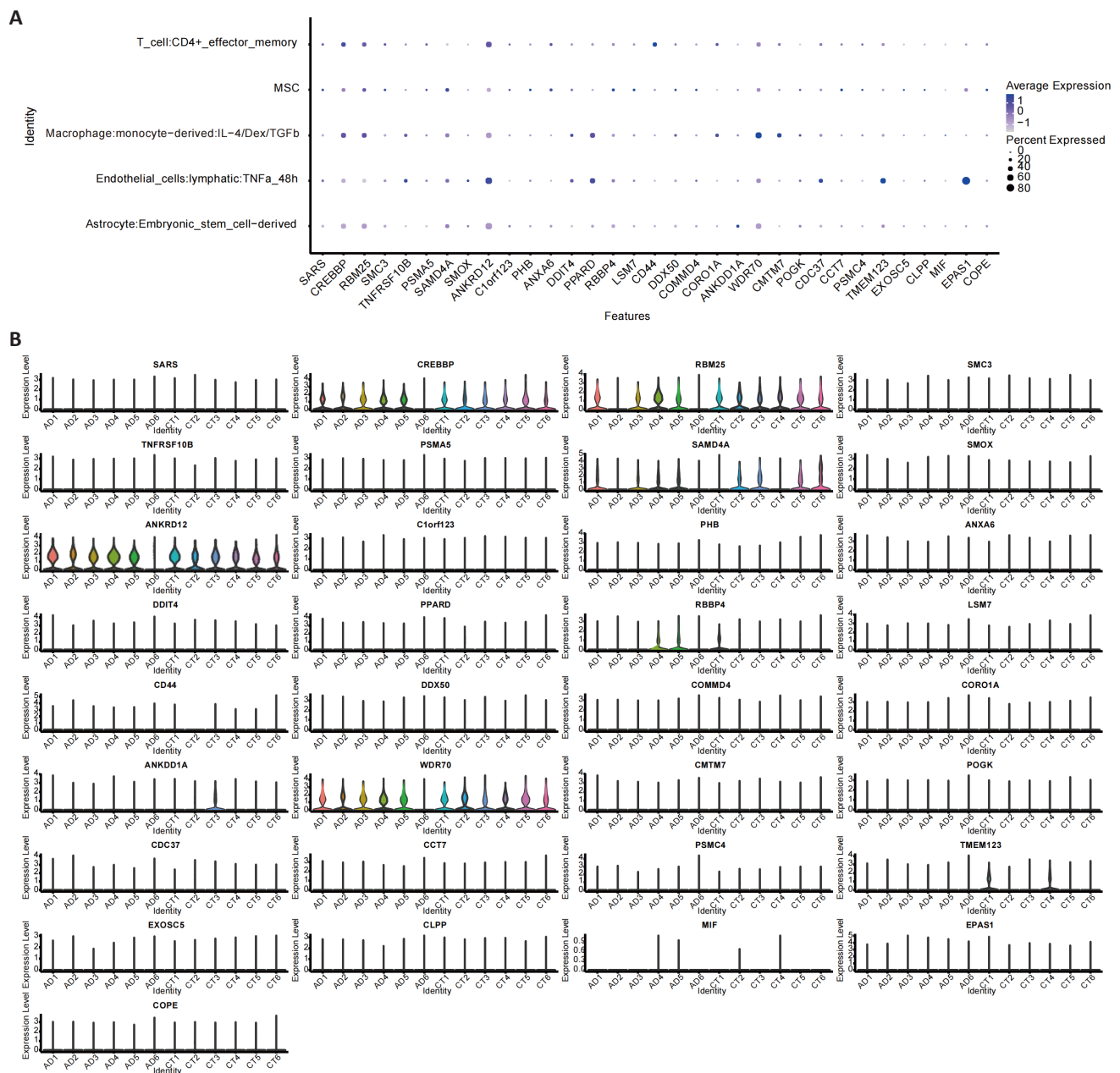


Figure 8. Overview of Single Cells from AD Samples and Normal Samples. A: Violin plot of ribosomal gene ratio after data quality control. B: The top 2,000 most variable genes using the *Find Variable Features* function and annotated the top 10.

to validation at the epigenetic level, we collected single-cell data for analysis. Implementing a rigorous protocol of quality control, filtering, dimensionality reduction clustering, cell type identification, and scoring, we authenticated the transcriptomes of 64,845 single-cell nuclei, affirming the consistency of our findings with prior conclusions.

In our study, we conducted a comprehensive series of screenings and analyses, successfully identifying 33 diagnostic marker genes pivotal for the progression assessment of MetS-associated AD. KEGG pathway enrichment analysis revealed that the 33 marker genes are associated with pathways such as Proteasome, Polycomb repressive complex, Phenylalanine metabolism, and Spinocerebellar ataxia, suggesting these genes may be involved in cerebellar ataxia, proteasome biosynthesis,

phenylalanine metabolism, and related biological processes. GO enrichment analysis indicated that the 33 marker genes are associated with molecular functions such as transcription coactivator binding (GO:0001223), unfolded protein binding (GO:0051082), and chromatin DNA binding (GO:0031490), suggesting that these genes may be involved in the process of transcriptional regulation, chromatin and DNA binding. To ascertain the diagnostic efficacy of these biomarkers in clinical practice, we meticulously constructed ROC curves and PR curves, facilitating a robust quantitative evaluation. Our immune cell infiltration analysis elucidated significant correlations between these 33 genes and distinct immune cell populations, with a notable inverse relationship observed with CD4+ T cells and a positive association with macrophages. Additionally, we analyzed nine metabolic pathways associated with the 33 genes based

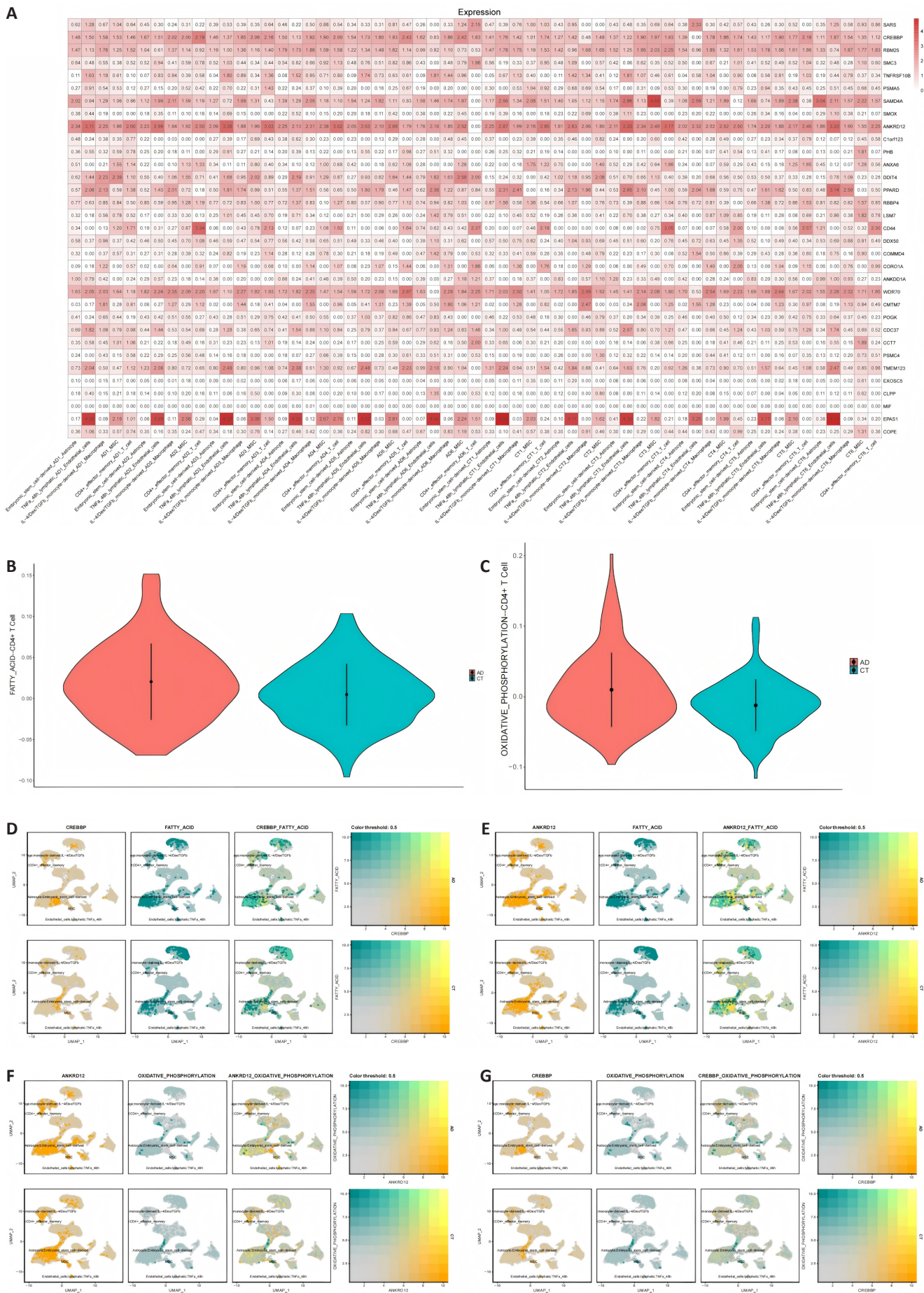


Figure 9. Overview of Single Cells between Hub Genes and 5 Clusters. A: Hub gene expression heatmap across single-cell clusters. B and C: Violin plot of expression levels of fatty acid metabolism and oxidative phosphorylation in CD4+ T cells from AD and normal samples by wilcox test. D and E: Feature Plot of fatty acid metabolism along with CREBBP and ANKRD12 Expression in CD4+ T Cells. F and G: Feature Plot of oxidative phosphorylation along with CREBBP and ANKRD12 Expression in CD4+ T Cells.

on metabolic gene databases. Among the two genes we focused on, both CREBBP and ANKRD12 are related to immunity and metabolism. Single-cell level analysis indicated that the downregulation of these genes in the brain tissues of AD patients exhibits a positive correlation with the presence of CD4+ effector memory T cells, as well as alterations in fatty acid metabolism and oxidative phosphorylation pathways, suggesting their integral role in the disease's immune and metabolic landscape.

CREBBP (cAMP-response element binding protein), a eukaryotic protein, belongs to the transcription co-activator and histone acetyltransferase family. This multifaceted protein plays a pivotal role in various cellular processes, encompassing gene transcription regulation, cell cycle governance, cell differentiation, and the orchestration of apoptosis. Abnormal expression of CREB family proteins and their signaling pathways are linked to neurological disorders, including Alzheimer's, Parkinson's, and depression. Research indicates that CREB's transcriptional activity is reduced in Alzheimer's patients, potentially causing synaptic dysfunction and memory impairment^[23]. Reduced CREB phosphorylation may contribute to the degeneration of dopaminergic neurons, with CREBBP potentially modulating this through its effects on CREB phosphorylation^[24].

ANKRD12 (ankyrin repeat domain 12) could suppress the transcriptional activity of nuclear receptors through the recruitment of histone deacetylases. The encoded protein interacts with p160 co-activators and also inhibits transcription mediated by co-activator-associated arginine methyltransferase 1 (CARM1) / activator of transcription 3, altered/deficient in renal dysplasia and genitourinary tumors, and lack of differentiation (ADA3)^[25]. Chia-Yen Chen et al. identified a correlation between eight genes, including ANKRD12, and adult cognitive function in a large-scale exome study^[26]. Additionally, in a study where exome sequencing identified genes associated with sleep-related traits, ANKRD12 was enriched in circadian rhythm and central nervous system neurons, and phenotypic association analysis showed that ANKRD12 is associated with cognitive and inflammatory traits^[27]. However, the specific impact of ANKRD12 as a rare protein-coding variant on adult cognitive function remains unclear and requires further research.

RF is an ensemble learning method that combines the prediction results of multiple decision trees to improve the accuracy of classification or regression. It works by constructing multiple decision trees and integrating their outputs, thus reducing the risk of overfitting and often being robust to noise and outliers. Lasso achieves sparsity and feature selection by constructing a penalty function that compresses some coefficients and sets some coefficients to zero. XGBoost is an optimized distributed gradient boosting library that is efficient, flexible,

and portable. Leveraging multiple machine learning algorithms for the optimization of core biomarker genes facilitates the precise identification of shared genetic factors and the construction of robust disease prediction models. Consequently, our model enables a preliminary estimation of the risk for MetS patients to develop AD by assessing the expression levels of diagnostic marker genes in peripheral blood. Despite the scarcity of related studies, our work presents a reliable diagnostic model for the prediction of MetS-associated AD and uncovers a range of novel potential therapeutic targets.

Previous studies have underscored the pivotal role of immune and metabolic responses in the progression of AD. A substantial body of research confirms that neuroinflammatory processes are significantly implicated in both the genesis and progression of AD. The primary immune cells implicated in neuroinflammation encompass microglia, macrophages, mast cells, and T cells^[28]. In individuals with AD, the heightened permeability of the BBB due to neuroinflammation permits peripheral macrophages to infiltrate the brain, engaging in the phagocytosis and clearance of amyloid-beta (A β) and tau proteins^[29]. The integrity of meningeal function is essential for preserving immune equilibrium within the CNS. Neuroinflammation-induced alterations can result in heightened permeability of the BBB^[30]. Elevation of both CD8+ and CD4+ effector memory and late effector T cells in the cerebrospinal fluid of AD patients correlates inversely with cognitive performance and positively with microstructural damage in the parahippocampal region^[31,32]. In patients with early symptoms of mild cognitive impairment, the development of tau deposition and atrophy of the temporal and parietal lobes lead to reduced brain metabolism. The energy required is metabolized by neurons, glial cells, and other cells in the central nervous system at different rates through glycolysis and mitochondrial oxidative phosphorylation (OXPHOS) in the blood. Changes in oxidative phosphorylation and mitochondrial bioenergetics defects may be related to the pathogenesis and progression of AD^[33]. Lipid metabolism, particularly fatty acid and cholesterol metabolism, is intricately linked to cerebral energy balance, oxidative stress, and neuroinflammatory processes. In multiple studies, changes in the abundance of unsaturated fatty acids have been found to be related to AD. In these studies, the levels of unsaturated fatty acids in the brain (cortex, middle frontal gyrus, and inferior temporal gyrus) and plasma of AD patients, including ω -3 PUFAs and MUFA (oleic acid), were lower than those in the cognitively normal control group^[34]. Furthermore, research has indicated that ApoE4 expression in AD patients diminishes the astrocyte's capacity for fatty acid clearance, leading to lipid droplet aggregation. Neurons accumulate fatty acids and generate peroxides under oxidizing conditions, leading to oxidative stress damage

and neuroinflammation^[35]. The positions of various metabolic and immune cells in the progression of AD, as well as the various manifestations of AD in different immune environments, provide a theoretical basis for further studying the relationship between immunity and AD. Both of these metabolic pathways are major therapeutic approaches for MetS. David Sancho and colleagues have explored the regulatory role of OXPHOS in tissue macrophage homeostasis and functionality across various organs, crucial for sustaining the energy demands of M2 macrophages and astrocytes^[36]. Loss of OXPHOS can affect fatty acid oxidation and cholesterol efflux pathways in adipose tissue macrophages, which is also a cause of lipid accumulation^[37].

The same role of immune cells and metabolism also exists in MetS diseases. Studies have shown that high-fat and high-sugar diets promote metabolic diseases by depleting Th17-inducing microorganisms, while restoring symbiotic Th17 cells can restore protective effects^[38]. Stefanie K Wculek et al. found through RNA sequencing analysis that under steady-state conditions, both in mice and humans, OXPHOS-related transcriptome changes vary across different tissues. Tfam deletion leads to a reduction in alveolar macrophages due to impaired lipid processing ability, resulting in increased cholesterol content and cellular stress, causing cell cycle arrest in the body, which leads to obesity^[36].

This is similar to our research findings, where CREBBP and ANKRD12 are positively correlated with the number of CD4+ effector memory T cells, as well as fatty acid metabolism and oxidative phosphorylation pathways. Consequently, a deeper understanding of these biomarkers, along with the implicated immune cells and metabolic pathways, sheds light on the intrinsic mechanisms driving AD progression in patients with MetS.

Although the current study provides some enlightening insights, there are still some limitations. Firstly, the analysis relied solely on a single dataset from the GEO repository, potentially introducing bias. Secondly, while the established model has high predictive performance, further exploration is needed to determine whether the number of key biomarkers can be further optimized and streamlined. In the later stage, we need to measure the diagnostic efficacy of each biomarker to optimize their number to the maximum extent. Thirdly, despite the high diagnostic value of the established model, further experimental validation is needed to determine whether the key biomarkers identified in this study can be applied in clinical practice. Therefore, we will conduct more research on key biomarkers.

5 CONCLUSION

Leveraging bioinformatics approaches and a suite of machine learning algorithms, our study has successfully pinpointed 33 pivotal biomarkers and crafted a diagnostic

model for the concurrent identification of MetS and AD. Additionally, our research uncovered an immune system imbalance in MetS-related AD patients, potentially impacting the infiltration levels of immune cells. These identified genes regulate these immune cells and simultaneously influence metabolic pathways including fatty acid and bile acid metabolism. These discoveries have been further validated at the single-cell analysis level, reinforcing the validity of our findings.

Acknowledgements

Not applicable.

Conflicts of Interest

The authors declared no conflict of interest.

Data Availability

The data that support the findings of this study are available in the TCGA (The Cancer Genome Atlas) dataset and GEO repository. These data were derived from the following resources available in the public domain: The TCGA dataset is available using the following link: <https://portal.gdc.cancer.gov/>. The GEO dataset is available using the following link: <https://www.ncbi.nlm.nih.gov/geo/>.

Ethical Statement

The data used in this study were provided by public source and collected from publicly available sequencing data. Therefore, no ethics committee approval was required for this study.

Author Contribution

Du K conceived and designed the research; Deng X collected the data and analyzed and interpreted the results; Deng X and Du K drafted the manuscript; All authors have read and approve the final version of the manuscript.

Abbreviation List

AD, Alzheimer's disease
AUC, Area under curve
BBB, Blood-brain barrier
DEGs, Differentially expressed genes
GEO, Gene Expression Omnibus
GO, Gene Ontology
KEGG, The Kyoto Encyclopedia of Genes and Genomes
LASSO, Least Absolute Shrinkage and Selection Operator
MDSC, Myeloid-derived suppressor cell
MetS, Metabolic syndrome
MTG, Middle temporal gyrus
RF, Random forest
ROC, Receiver operating characteristic
scRNA-seq, Single-cell RNA sequencing
ssGSEA, Single-sample Gene Set Enrichment Analysis
WGCNA, Weighted Gene Coexpression Network Analysis
XGBoost, Extreme Gradient Boosting

References

- [1] Reitz C, Brayne C, Mayeux R. Epidemiology of Alzheimer disease. *Nat Rev Neurol*, 2011; 7: 137-152.[\[DOI\]](#)
- [2] Lopez JAS, González HM, Léger GC. Alzheimer's disease. Handbook of clinical neurology, 2019; 167: 231-255.[\[DOI\]](#)
- [3] Meng X, Su J, Li H et al. Effectiveness of caregiver non-pharmacological interventions for behavioural and psychological symptoms of dementia: An updated meta-analysis. *Ageing Res Rev*, 2021; 71: 101448.[\[DOI\]](#)
- [4] White LR, Edland SD, Hemmy LS et al. Neuropathologic comorbidity and cognitive impairment in the Nun and Honolulu-Asia Aging Studies. *Neurology*, 2016; 86: 1000-1008.[\[DOI\]](#)
- [5] Montine TJ, Sonnen JA, Montine KS et al. Adult Changes in Thought study: dementia is an individually varying convergent syndrome with prevalent clinically silent diseases that may be modified by some commonly used therapeutics. *Curr Alzheimer Res*, 2012; 9: 718-723.[\[DOI\]](#)
- [6] Kawas CH, Kim RC, Sonnen JA et al. Multiple pathologies are common and related to dementia in the oldest-old: The 90+ Study. *Neurology*, 2015; 85: 535-542.[\[DOI\]](#)
- [7] Fahed G, Aoun L, Bou Zerdan M et al. Metabolic syndrome: updates on pathophysiology and management in 2021. *Int J Mol Sci*, 2022; 23: 786.[\[DOI\]](#)
- [8] Forti P, Pisacane N, Rietti E et al. Metabolic syndrome and risk of dementia in older adults. *J Am Geriatr Soc*, 2010; 58: 487-492.[\[DOI\]](#)
- [9] Villapol S, Janatpour ZC, Afram KO et al. The renin angiotensin system as a therapeutic target in traumatic brain injury. *Neurotherapeutics*, 2023; 20: 1565-1591.[\[DOI\]](#)
- [10] Salminen A, Kauppinen A, Kaarniranta K. Hypoxia/ischemia activate processing of Amyloid Precursor Protein: impact of vascular dysfunction in the pathogenesis of Alzheimer's disease. *J Neurochem*, 2017; 140: 536-549.[\[DOI\]](#)
- [11] AlAnazi FH, Al-Kuraishy HM, Alexiou A et al. Primary hypothyroidism and Alzheimer's disease: a tale of two. *Cell Mol Neurobiol*, 2023; 43: 3405-3416.[\[DOI\]](#)
- [12] Pugazhenth S. Metabolic syndrome and the cellular phase of Alzheimer's disease. *Prog Mol Biol Transl*, 2017; 146: 243-258.[\[DOI\]](#)
- [13] Dinel AL, Andre C, Aubert A et al. Cognitive and emotional alterations are related to hippocampal inflammation in a mouse model of metabolic syndrome. *PLoS One*, 2011; 6: e24325.[\[DOI\]](#)
- [14] Diboun, I.; Wernisch, L.; Orengo, C.A.; Koltzenburg, M. Microarray Analysis after RNA Amplification Can Detect Pronounced Differences in Gene Expression Using Limma. *BMC Genomics*, 2006; 7: 252.[\[DOI\]](#)
- [15] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 2008; 9: 559.[\[DOI\]](#)
- [16] Wu T, Hu E, Xu S et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation*, 2021; 2: 100141.[\[DOI\]](#)
- [17] Paul A, Mukherjee DP, Das P et al. Improved random forest for classification. *IEEE T Image Process*, 2018; 27: 4012-4024.[\[DOI\]](#)
- [18] Vasquez MM, Hu C, Roe DJ et al. Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. *BMC Med Res Methodol*, 2016; 16: 154.[\[DOI\]](#)
- [19] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; ACM: San Francisco, CA, USA, 2016; 785-794.[\[DOI\]](#)
- [20] Robin X, Turck N, Hainard A et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 2011; 12: 77.[\[DOI\]](#)
- [21] Butler A, Hoffman P, Smibert P et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*, 2018; 36: 411-420.[\[DOI\]](#)
- [22] Aran D, Looney AP, Liu L et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol*, 2019; 20: 163-172.[\[DOI\]](#)
- [23] Grochowska KM, Gomes GM, Raman R et al. Jacob-induced transcriptional inactivation of CREB promotes A β -induced synapse loss in Alzheimer's disease. *EMBO J*, 2023; 42: e112453.[doi: 10.15252/embj.2022112453](#).[\[DOI\]](#)
- [24] Mantamadiotis T, Lemberger T, Bleckmann S C, et al. Disruption of CREB function in brain leads to neurodegeneration. *Nat Genet*, 2002; 31: 47-54.[\[DOI\]](#)
- [25] Fagerberg L, Hallström B M, Oksvold P et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*, 2014; 13: 397-406.[\[DOI\]](#)
- [26] Chen C Y, Tian R, Ge T et al. The impact of rare protein coding genetic variation on adult cognitive function. *Nat Genet*, 2023; 55: 927-938.[\[DOI\]](#)
- [27] Fei CJ, Li ZY, Ning J et al. Exome sequencing identifies genes associated with sleep-related traits. *Nat Hum Behav*, 2024; 8: 576-589.[\[DOI\]](#)
- [28] Wang H, Shen Y, Chuang H et al. Neuroinflammation in Alzheimer's disease: microglia, molecular participants and therapeutic choices. *Curr Alzheimer Res*, 2019; 16: 659-674.[\[DOI\]](#)
- [29] Mammanna S, Fagone P, Cavalli E et al. The role of macrophages in neuroinflammatory and neurodegenerative pathways of Alzheimer's disease, amyotrophic lateral sclerosis, and multiple sclerosis: pathogenetic cellular effectors and potential therapeutic targets. *Int J Mol Sci*, 2018; 19: 831.[\[DOI\]](#)
- [30] Solár P, Zamani A, Kubičková L et al. Choroid plexus and the blood-cerebrospinal fluid barrier in disease. *Fluids Barriers CNS*, 2020; 17: 35.[\[DOI\]](#)
- [31] Oberstein TJ, Taha L, Spitzer P et al. Imbalance of circulating Th17 and regulatory T cells in Alzheimer's disease: a case control study. *Front Immunol*, 2018; 9: 1213.[\[DOI\]](#)
- [32] Hu WT, Chen-Plotkin A, Grossman M et al. Novel CSF biomarkers for frontotemporal lobar degenerations. *Neurology*, 2010; 75: 2079-2086.[\[DOI\]](#)
- [33] Bano D, Ehninger D, Bagetta G. Decoding metabolic signatures in Alzheimer's disease: a mitochondrial perspective. *Cell Death Discov*, 2023; 9: 432.[\[DOI\]](#)
- [34] Zhang X, Chen C, Liu Y. Navigating the metabolic maze: anomalies in fatty acid and cholesterol processes in Alzheimer's astrocytes. *Alzheimers Res Ther*, 2024; 16: 63.[\[DOI\]](#)
- [35] Yin F. Lipid metabolism and Alzheimer's disease: clinical evidence, mechanistic link and therapeutic promise. *FEBS J*, 2023; 290: 1420-1453.[\[DOI\]](#)
- [36] Wculek SK, Heras-Murillo I, Mastrangelo A et al. Oxidative phosphorylation selectively orchestrates tissue macrophage homeostasis. *Immunity*, 2023; 56: 516-530.E9.[\[DOI\]](#)
- [37] Houston S. Tissue differences in macrophage metabolism. *Nat Immunol*, 2023; 24: 378-378.[\[DOI\]](#)
- [38] Kawano Y, Edwards M, Huang Y et al. Microbiota imbalance induced by dietary sugar disrupts immune-mediated protection from metabolic syndrome. *Cell*, 2022; 185: 3501-3519.e20.[\[DOI\]](#)