Received: 19 March 2024 | Revised: 1 August 2024 | Accepted: 20 August 2024

DOI: 10.1002/dad2.70006

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Assessing Alzheimer's disease via plasma extracellular vesicle–derived mRNA

Le Hoang Phu Pham[1] | **Ching-Fang Chang**[1] | **Katherine Tuchez**[1] | **Fei Liu**[2] | **Yuchao Chen**[1]

[1]WellSIM Biomedical Technologies Inc., San Jose, California, USA

[2]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

Correspondence
Fei Liu, Department of medicine, Brigham and Women's Hospital, Harvard Medical School; 75 Francis Street, Boston, MA 02115, USA.
Email: fliu@bwh.harvard.edu

Yuchao Chen, WellSIM Biomedical Technologies Inc., 160 East Tasman Dr, San Jose, CA 95134, USA.
Email: yuchaochen@wellsimbiotech.com

Funding information
Healthy Aging and Senile Dementia, Grant/Award Number: P01 AG03991; Adult Children Study, Grant/Award Number: P01 AG026276; National Institute on Aging, Grant/Award Number: U24 AG021886; National Institutes of Health, Grant/Award Numbers: 1R41AG076098-01, 1R43AG080878-01, 2R44GM144009-02; Alzheimer's Disease Research Center, Grant/Award Number: P30AG066444

Abstract

INTRODUCTION: Alzheimer's disease (AD), the most prevalent neurodegenerative disorder globally, has emerged as a significant health concern. Recently it has been revealed that extracellular vesicles (EVs) play a critical role in AD pathogenesis and progression. Their stability and presence in various biofluids, such as blood, offer a minimally invasive window for monitoring AD-related changes.

METHODS: We analyzed plasma EV-derived messenger RNA (mRNA) from 82 human subjects, including individuals with AD, mild cognitive impairment (MCI), and healthy controls. With next-generation sequencing, we profiled differentially expressed genes (DEGs), identifying those associated with AD.

RESULTS: Based on DEGs identified in both the MCI and AD groups, a diagnostic model was established based on machine learning, demonstrating an average diagnostic accuracy of over 98% and showed a strong correlation with different AD stages.

DISCUSSION: mRNA derived from plasma EVs shows significant promise as a non-invasive biomarker for the early detection and continuous monitoring of AD.

KEYWORDS
AD diagnostics, Alzheimer's disease, extracellular vesicles, mRNA, non-invasive biomarker, RNA sequencing

Highlights

- The study conducted next-generation sequencing (NGS) of mRNA derived from human plasma extracellular vesicles (EVs) to assess Alzheimer's disease (AD).
- Profiling of plasma EV-derived mRNA shows a significantly enriched AD pathway, indicating its potential for AD-related studies.
- The AD-prediction model achieved a receiver-operating characteristic area under the curve (ROC-AUC) of more than 0.98, with strong correlation to the established Clinical Dementia Rating (CDR).

1 | BACKGROUND

Early diagnosis of Alzheimer's disease (AD) is crucial for timely intervention and symptomatic treatment to prevent irreversible brain damage.[1–3] Integrative diagnostic approaches combine cognitive and functional assessments, biomarker identification, and neuroimaging.[4] However, the high costs of positron emission tomography (PET) imaging limit its use, especially in under-resourced areas. Cerebrospinal fluid (CSF), despite being an optimal source for AD biomarkers, is difficult to use frequently due to the invasiveness of lumbar puncture.[5] Therefore, researchers are exploring non-invasive AD biomarkers in alternative body fluids, particularly blood.

Extracellular vesicles (EVs) have been recently recognized as significant contributors to AD pathology.[6] These membrane-bound nano vesicles encapsulate a range of cellular components including nucleic acids, proteins, and lipids, which reflect their originating cells.[7] In AD, EVs are known to carry hallmark biomolecules associated with the disease, indicating their potential role in spreading pathology within the brain.[8] Furthermore, these vesicles are implicated in the modulation of neuroinflammation, a key component of AD's pathological progression.[9] Beyond their role in disease mechanisms, EVs are gaining attention as potential diagnostic tools.[10,11] Their stability and presence in various biofluids, such as blood, offer a minimally invasive window into AD-related changes. Compared to free circulating nucleic acids in plasma, nucleic acids within EVs are usually more stable and can be released into bodily fluids by cells early in disease progression.[12,13] Therefore, analyzing the content of EVs could provide insights into the early stages and progression of AD, thereby highlighting their value in diagnosis.

Messenger RNA (mRNA) in EVs may play a crucial role in AD pathology. mRNA, carrying genetic information for protein synthesis, can be transferred to recipient cells via EVs, altering cellular behavior.[14] EV secretion occurs in both normal and pathological conditions, with recent cancer research highlighting the role of EV-carried mRNAs in tumor progression and the tumor microenvironment.[15] Similarly, mRNA in AD offers insights into gene expression patterns, though research on EV-derived mRNA in AD is limited. Sproviero et al. profiled mRNA in EVs from neurodegenerative disease patients but found no deregulated mRNAs in AD.[16] Luo et al. profiled long RNA in EVs from AD brain tissues, revealing differentially expressed genes linked to transcriptional changes in AD.[17] Given these insights, profiling EV-derived mRNA may provide novel diagnostic and therapeutic approaches for AD.

This study aims to further our understanding of AD and improve its early diagnostic capabilities through the comprehensive profiling of mRNA from plasma EVs, which were released by various cells and tissues. Considering that AD is a systemic disorder linked to a broad range of metabolic processes, EVs from various sources, including the central nervous system (CNS) and other non-CNS systems such as the immune, cardiovascular, and endocrine systems, could all exhibit significant changes in mRNA profiles as the disease progresses.

Leveraging next-generation sequencing (NGS), we analyzed mRNA libraries derived from EVs in 82 plasma samples. By comparing the gene-expression profiles from both AD and mild cognitive impairment (MCI) cases with those from healthy controls, we identified a collection of gene biomarker candidates. These candidates underwent rigorous functional and correlation analyses, resulting in the establishment of a classification model based on a support vector machine (SVM) algorithm, showing high accuracy and strong correlation with the Clinical Dementia Rating (CDR) scale.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the literature through traditional sources (e.g., PubMed). It has been revealed that extracellular vesicles (EVs) play an important role in Alzheimer's disease (AD) pathogenesis and progression. The identification of plasma EV biomarkers holds promise for the development of non-invasive tests for early AD diagnosis. Although the discovery of RNA biomarkers is accelerating, the primary emphasis continues to be on micro-RNA (miRNA). Studies on messenger RNA (mRNA) biomarkers remain limited. A thorough investigation into mRNA biomarkers could contribute to the advancements in early AD diagnosis.
2. **Interpretation**: Through profiling of plasma EV-derived mRNA, our studies result in an AD prediction model with a receiver-operating characteristic area under the curve (ROC-AUC) of more than 0.98, based on a biomarker panel consisting of eight AD-related differentially expressed genes (DEGs): *INSR*, *ERN1*, *HSD17B10*, *PPIF*, *CHRM1*, *AKT2*, *IDE*, and *ATP2A2*.
3. **Future directions**: Future studies could focus on analyzing clinical samples from larger and more diverse cohorts, as well as conducting longitudinal studies and comparing AD samples with those from other neurodegenerative conditions.

2 | METHODS

2.1 | Clinical specimens

We analyzed a total of 82 ethylenediaminetetraacetic acid (EDTA)–treated plasma specimens, including 25 samples from healthy controls, 13 from individuals with MCI, and 44 from patients diagnosed with AD. These specimens were from three independent patient cohorts: Washington University in St. Louis (Knight Alzheimer Disease Research Center), Indiana University (National Centralized Repository for Alzheimer's Disease and Related Dementias), and PrecisionMed. The study was approved by the institutional review boards of all participating institutions, with written informed consent provided by all participants involved in the study. The demographics and

clinicopathological characteristics of these participants are detailed in Tables S1 and S2. AD diagnoses were based on either clinical evaluation (NCRAD and PrecisionMed) or a combination of clinical evaluation and PET biomarker (ADRC). The Control and MCI groups had a mean age of 67, compared to 74 in the AD group. CDR scale scores were available for all samples, with mean scores of 0.65 for the MCI group and 1.58 for the AD group. Upon receipt from the organizations, the samples were stored at −80°C. Before EV isolation, the samples were thawed at room temperature for 5 min and then kept on ice. Each sample underwent a maximum of three freeze–thaw cycles.

2.2 | EV isolation, library preparation, and RNA sequencing

For isolating EVs, 350 μL of each plasma sample was first diluted with 1 × PBS (phosphate-buffered saline) to a final volume of 15 mL. The diluted plasma was then passed through a 0.22 μm syringe filter to remove large particles and aggregates. The isolation of EVs was conducted using the EXODUS platform, which utilized 25 mm diameter exosome isolation devices (EIDs) according to the previously published method.[18] Previous studies have optimized and characterized EXODUS method for plasma EV isolation.[19,20] Once isolated, the EVs were resuspended in 400 μL of PBS. From this suspension, a 100 μL aliquot of sample was further diluted 4 times with PBS for nanoparticle tracking analysis (NTA) using NanoSight (Malvern Panalytical, LM14), equipped with 488 nm laser and a highly sensitive sCMOS camera. The diluted sample was introduced directly through a syringe to the instrument. The concentration and size distribution of EVs were quantified and generated automatically by NanoSight NTA 3.4 (Detection threshold: 3, number of captures: 5, and capture duration: 60 s). Typical results are shown in Figure S1.

Another 300 μL of EVs in PBS were then mixed with an equal volume of Dynabeads M-270 Streptavidin beads (ThermoFisher Scientific, 65305), conjugated previously with a biotinylated oligonucleotide that included a polymerase chain reaction (PCR) handle and a Poly(dT) sequence. These beads were suspended in a lysis buffer containing 0.4% Sarkosyl to facilitate the capture of mRNA from lysed EVs. Following this, solid-phase reverse transcription was performed using Maxima H Minus Reverse Transcriptase (ThermoFisher Scientific, EP0752). The complementary DNA (cDNA) synthesized on the beads underwent amplification through SMART PCR, using Kapa HiFi Hotstart Readymix (Roche, KK2601), and the resulting DNA was subsequently purified using AMPure XP beads (Beckman Coulter, A63881). The purified DNA was further subject to a second round of PCR and purification. The PCR products then underwent tagmentation using a Nextera XT Library Preparation Kit (Illumina, FC-131-1096). For fragmented mRNA within EVs, NEBNext Small RNA Library Prep Set (New England BioLabs, E7330L) was employed for library preparation according to manufacturer's instructions. Quality control of the NGS library preparation was accessed via a TapeStation 4200 with High Sensitivity D1000 ScreenTape (Agilent) as shown in Figure S1. Libraries that did not contain 1 ng of DNA in 5 μL were excluded. Finally, RNA sequencing was performed on an Illumina NovaSeq 6000 platform with at least 20 M of reads for each library.

2.3 | Data analysis

The data analysis procedure was detailed in the Supplementary Information. Briefly, sequencing data were trimmed and aligned to the Ensembl reference genome using the STAR aligner. Reads were annotated using the biomaRt package. Differential gene expression analysis was performed with DESeq2 and sva, using criteria of log2 fold change > |0.5| and adjusted p-value < 0.05. Functional profiling was done with g:Profiler. A diagnostic classifier was developed using a linear SVM model with nested cross-validation. Correlation analysis and clustering were performed using the WGCNA package, with statistical differences assessed using the Wilcoxon and Kruskal-Wallis tests. Various plots were created using GraphBio, BoxPlotR, and SRPlot.

3 | RESULTS

3.1 | Transcriptome and pathway analysis

As shown in Figure 1A, we isolated and purified plasma EVs from 82 individuals (see Table S1 for details). We then extracted mRNA from these EV samples for sequencing, identifying 19,495 genes, including 15,664 protein-coding genes (80.34%), 2391 long non-coding RNAs (12.26%), and 1440 genes of other biotypes (7.38%). The data set, rich in protein-coding genes, forms the basis of our study. Volcano plots of the MCI group (Figure 1B) showed 1097 significantly expressed protein-coding genes compared to the Control group, with 897 upregulated and 200 downregulated. In the AD group (Figure 1C), 1701 protein-coding genes were significantly differentially expressed compared to the Control group, with 1421 upregulated and 280 downregulated. Pathway analysis revealed significant enrichment in the AD pathway, with the MCI group showing a p-adjusted value of 1.36e-7 and 27 AD-related genes, and the AD group showing a p-adjusted value of 1.25e-13 and 44 AD-related genes. Heatmaps in Figure 1D,E displayed distinct expression patterns of these AD-related genes in MCI and AD cases compared to Control samples.

We compared Gene Ontology (GO) annotations for AD-related DEGs in the MCI and AD groups (Figure 1F-G), focusing on the top 10 significant annotations. In the Biological Processes category, MCI DEGs are involved in the Wingless/Integrated (Wnt) signaling pathway and cellular response to oxygen-containing compounds, whereas AD DEGs focus on energy metabolism pathways. In the Cellular Components domain, MCI DEGs are associated with protein complexes and organelle membranes, whereas AD DEGs are linked to cellular structures like organelle membranes and catalytic complexes. In the Molecular Functions category, MCI genes are involved in binding

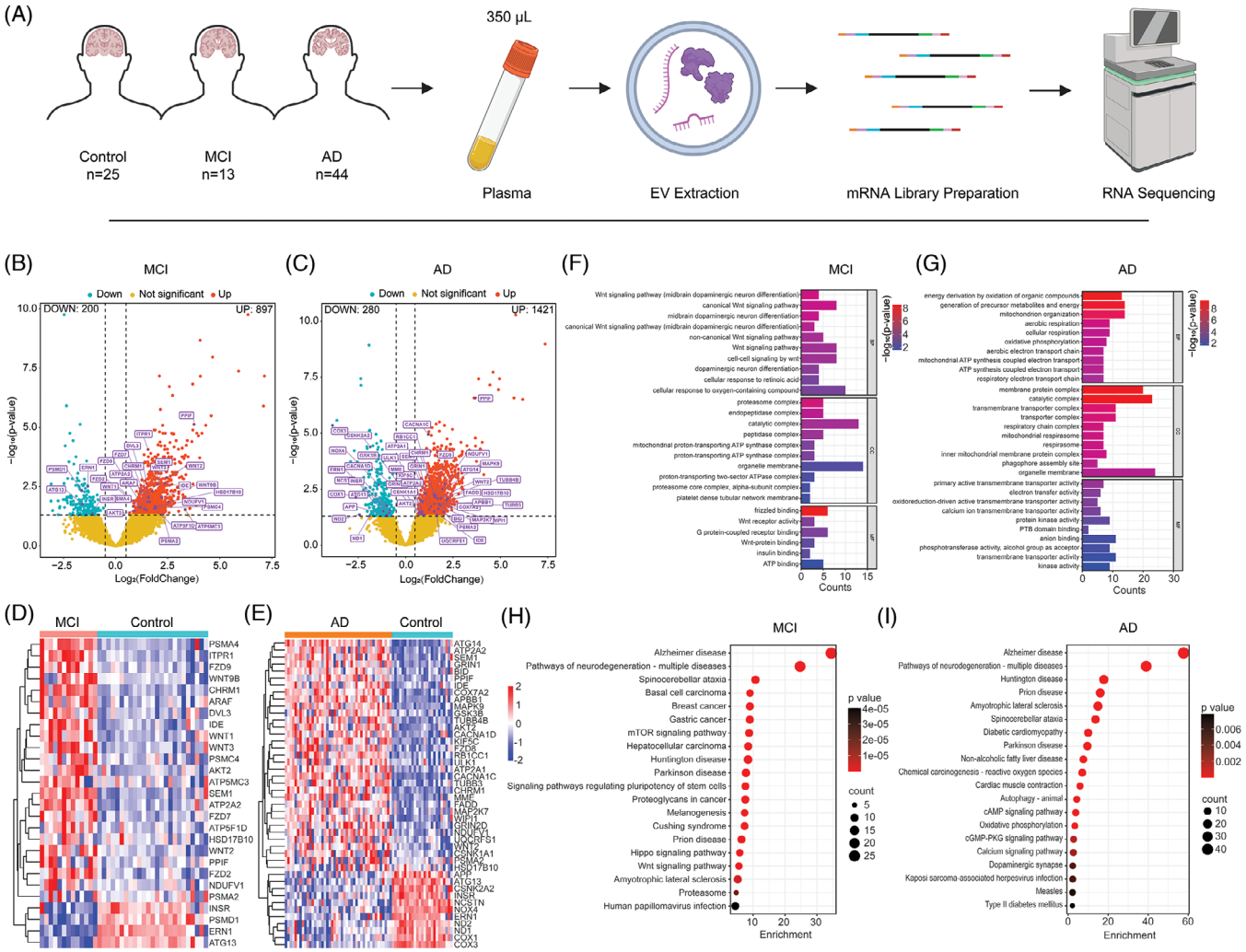

FIGURE 1 Analysis of mRNA from plasma-derived EVs highlighting key AD-related genes in the MCI and AD groups. (A) Illustration of sample-preparation workflow. (B,C) Volcano plots showing differentially expressed protein-coding genes in the MCI and AD groups compared to the Control group (p-adjusted < 0.05 and log2 fold change > |0.5|). The labeled genes are AD related. (D,E) Heatmaps display expression patterns of the AD-related DEGs in the MCI and AD groups compared to the Control group. Gene clustering in the heatmaps was based on hierarchical clustering using the Euclidean distance metric. (F,G) GO annotations for the AD-related DEGs in the MCI and AD groups. (H, I) KEGG pathway enrichment analysis for AD-related DEGs in MCI and AD groups. AD, Alzheimer's Disease; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MCI, Mild Cognitive Impairment.

activities, whereas AD genes focus on enzymatic functions, reflecting a shift in molecular characteristics from MCI to AD.

We subsequently conducted a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of AD-related DEGs in both MCI and AD groups (Figure 1H,I), revealing similarities and distinctions. Both groups shared pathways related to AD, neurodegeneration, and other neurologically linked diseases. Notably, the AD pathway showed higher enrichment scores in the AD group (57.69) compared to the MCI group (34.50), indicating a progressive gene-expression shift. MCI was linked to pathways such as mammalian target of rapamycin (mTOR) signaling and pluripotency of stem cells, whereas AD was associated with metabolic and systemic processes, such as diabetic cardiomyopathy and oxidative phosphorylation. The AD group consistently showed higher gene counts within shared pathways, indicating greater genetic complexity.

3.2 | Identification of potential AD biomarkers

In the study to identify biomarkers for AD, we focused on a subset of DEGs shared by the MCI and AD groups (Figure 2A), which made up 22.5% of the total AD-related gene pool. This indicates a notable genetic overlap between MCI and AD, suggesting that these genes are potentially critical in AD progression. The expression patterns of these genes, analyzed through variance stabilizing transformation (VST) Related Dementias showed distinct trends: 10 genes were upregulated and 3 were downregulated in the MCI and AD stages compared to the Control group. These trends were quantified statistically as shown in Figure 2B. Generally, the upregulated genes showed higher median VST values in diseased states, whereas downregulated genes had lower median VST values, indicating a reduction in expression as the disease advances. An exception was *PSMA2*, which did not exhibit

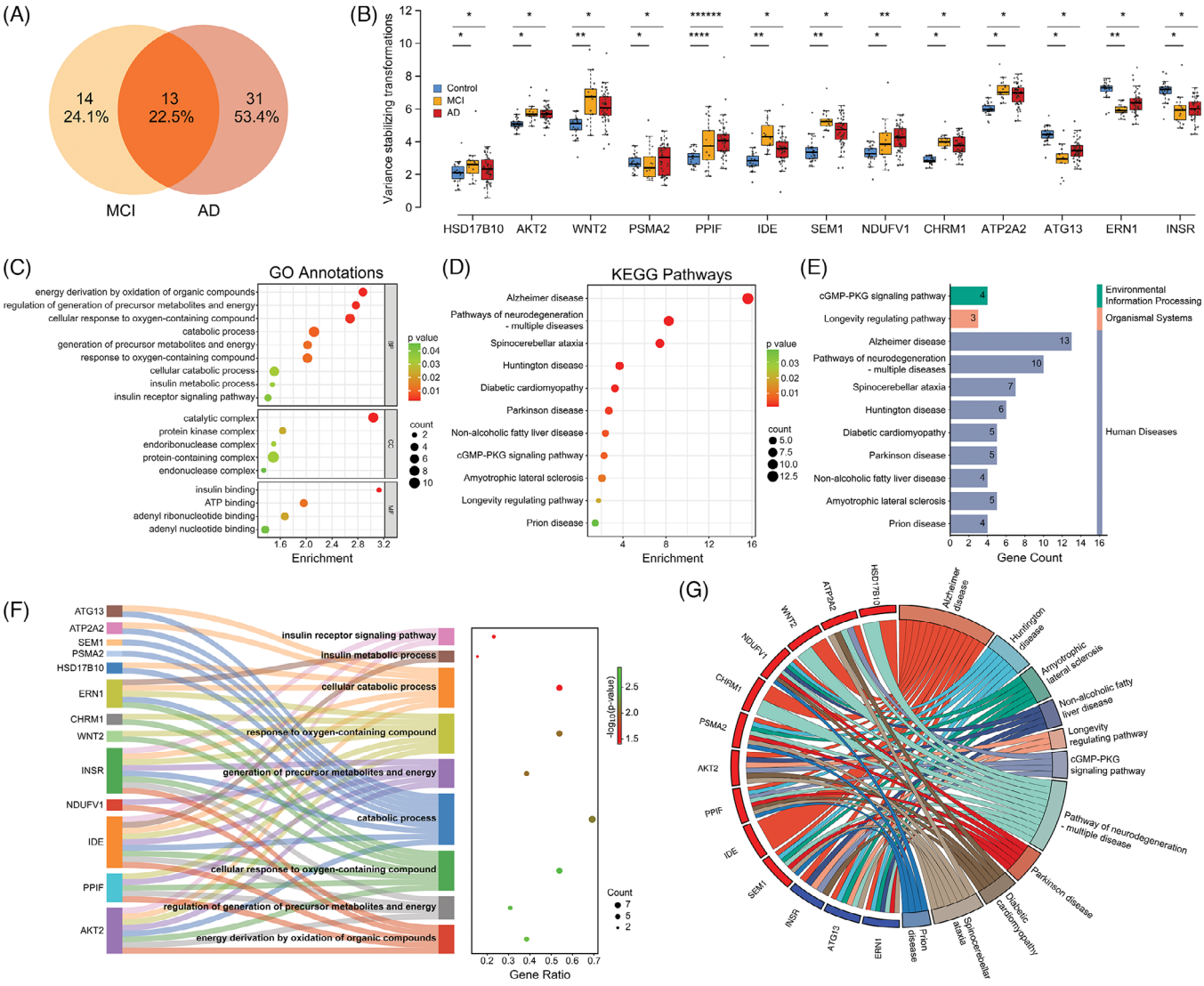

FIGURE 2 Integrative analysis of 13 shared DEGs. (A) Venn diagram showing the overlap of differentially expressed AD-related genes between the MCI and AD groups. (B) Box plot showing variations in gene expression across the 13 shared AD genes through different disease stages: Control, MCI, and AD (*$p < 0.05$, **$p < 0.005$, ****$p < 0.00005$, ******$p < 0.0000005$). (C) Enrichment bubble plot illustrating the GO annotations for the 13 genes. (D) Enrichment bubble plot illustrating the KEGG pathways enriched by the 13 genes. (E) Summary of pathway enrichment listing the number of genes involved in each pathway. (F) Sankey diagram mapping the 13 genes to their associated biological processes. (G) Chord plot showing the relationships between the 13 genes and their linked KEGG pathways.

a clear upregulation or downregulation trend when MCI and AD were compared with the Control group. The statistical significance of these differential expressions was marked with asterisks above the boxplots (Figure 2B).

3.3 | Functional enrichment and pathway analysis of common genes

We further analyzed GO annotations for 13 AD-related DEGs shared between the MCI and AD groups (Figure 2C). Significant enrichments were found in cellular metabolism processes, with five genes involved in energy derivation and seven genes in oxidative stress responses. In the Cellular Components domain, protein-containing complexes, including the catalytic and protein kinase complexes, were prominent. For Molecular Functions, insulin-binding and nucleotide-binding functions were significant. The Sankey diagram highlighted the connection of these genes to critical biological processes (Figure 2F), such as energy derivation through oxidation of organic compounds and regulation of precursor metabolites, with notable genes such as *AKT2*, *PPIF*, and *INSR* indicating disruptions in metabolic and oxidative stress pathways in AD.

The KEGG pathway analysis of 13 AD-related DEGs (Figure 2D,E) highlighted the AD pathway as the most significant, with all 13 genes involved. This pathway as illustrated in Figure S2, had a high enrichment score and low *p*-adjusted value, emphasizing a common genetic

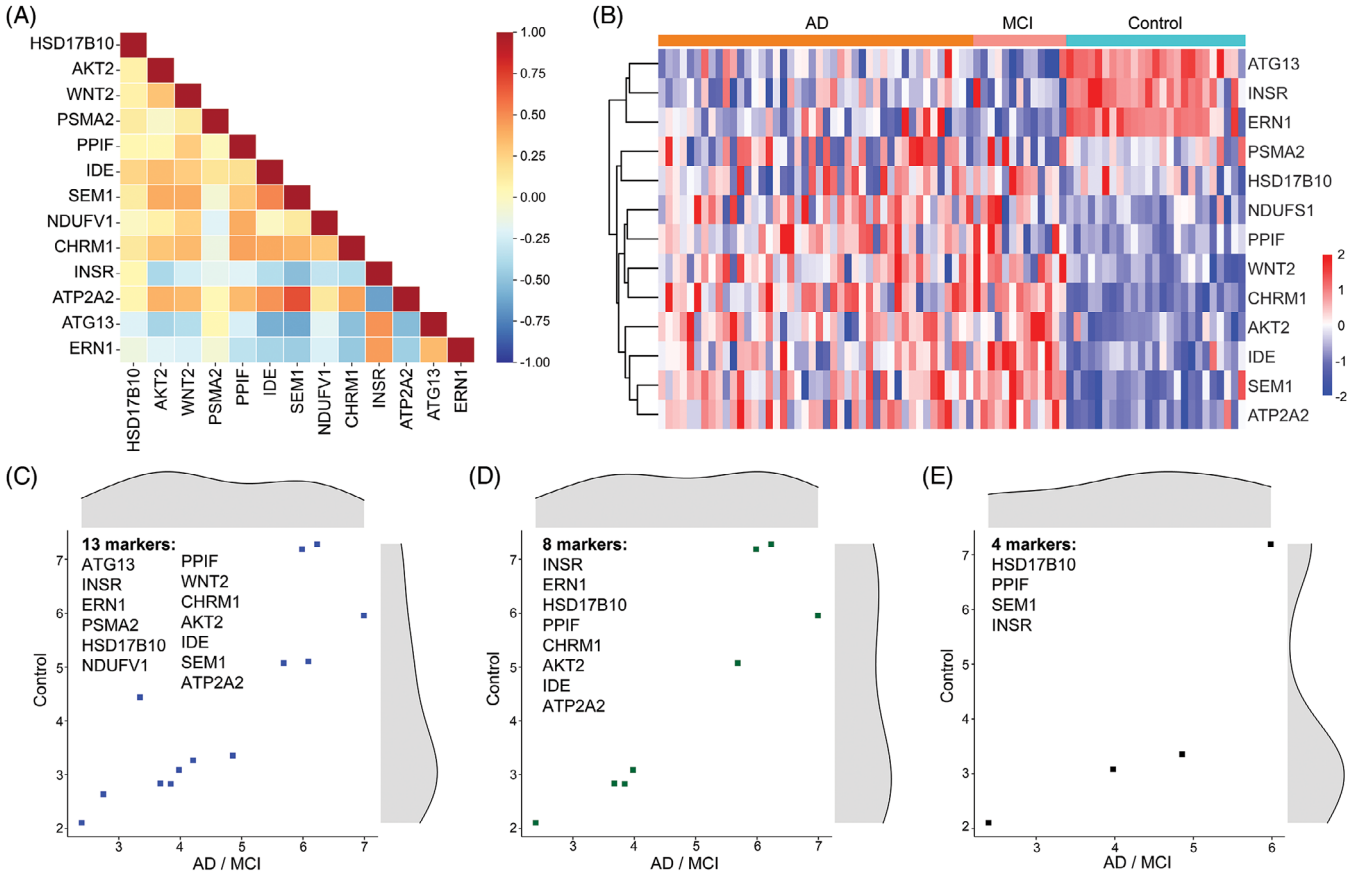

FIGURE 3 Biomarker selection through correlation analysis and clustering. (A) Bi-weight correlation matrix showing genes with strong correlations. Color intensity represents the correlation strength with darker hues indicating stronger correlations. (B) Heatmap with a dendrogram constructed using Euclidean distance, showing gene expression similarities among the 13 AD genes. (C–E) Scatter plots showing the median expression distributions of the Control and AD/MCI groups across three selected gene sets.

basis from MCI to AD. Other significant pathways included those related to multiple neurodegenerative diseases, such as spinocerebellar ataxia and Huntington's disease, and metabolic diseases such as diabetic cardiomyopathy. The analysis also identified signaling pathways crucial for vascular and neuronal functions and aging. A chord diagram (Figure 2G) showed the extensive interactions of genes such as *ATP2A2*, *NDUFV1*, *PSMA2*, and *SEM1* with various pathways, suggesting their broad impact on neurodegenerative and metabolic diseases. Genes such as *AKT2*, *PPIF*, *INSR*, *ATG13*, and *ERN1* demonstrated connections to pathways involving metabolic regulation, aging, and neurodegeneration, highlighting their roles in these processes.

3.4 | Biomarker selection through correlation analysis and clustering

We refined the 13 AD-related DEGs using biweight midcorrelation analysis (Figure 3A), identifying gene pairs with a correlation coefficient above |0.65|, such as *ATP2A2* and *SEM1*. This helped maintain data set integrity and avoid redundancy. A comprehensive clustering analysis using Euclidean distance produced a heatmap (Figure 3B), clustering genes based on their expression patterns across AD, MCI, and Control groups. Closely related gene pairs identified in the dendrogram were marked for potential exclusion to ensure a unique, non-overlapping set for further validation.

Following the analysis, we derived three potential biomarker panels for further examination, each characterized by scatterplots of median expression values for the Control group and a combined group of MCI and AD. Panel 1 includes all 13 genes (Figure 3C). Panel 2, with eight genes, was created by removing one gene from each redundant pair (Figure 3D). An alternative to Panel 2, also with eight genes, selected the alternate genes from the redundant pairs (Figure S3A). Panel 3 selected one gene from the downregulated branch and three from the upregulated branch of the dendrogram (Figure 3E). The scatterplots showed similar gene expression ranges, ensuring comprehensive biomarker coverage and minimal redundancy for subsequent validation.

3.5 | Development and characterization of AD prediction model

For each biomarker panel, we developed an AD prediction model by training a linear SVM classifier with the data set. A nested

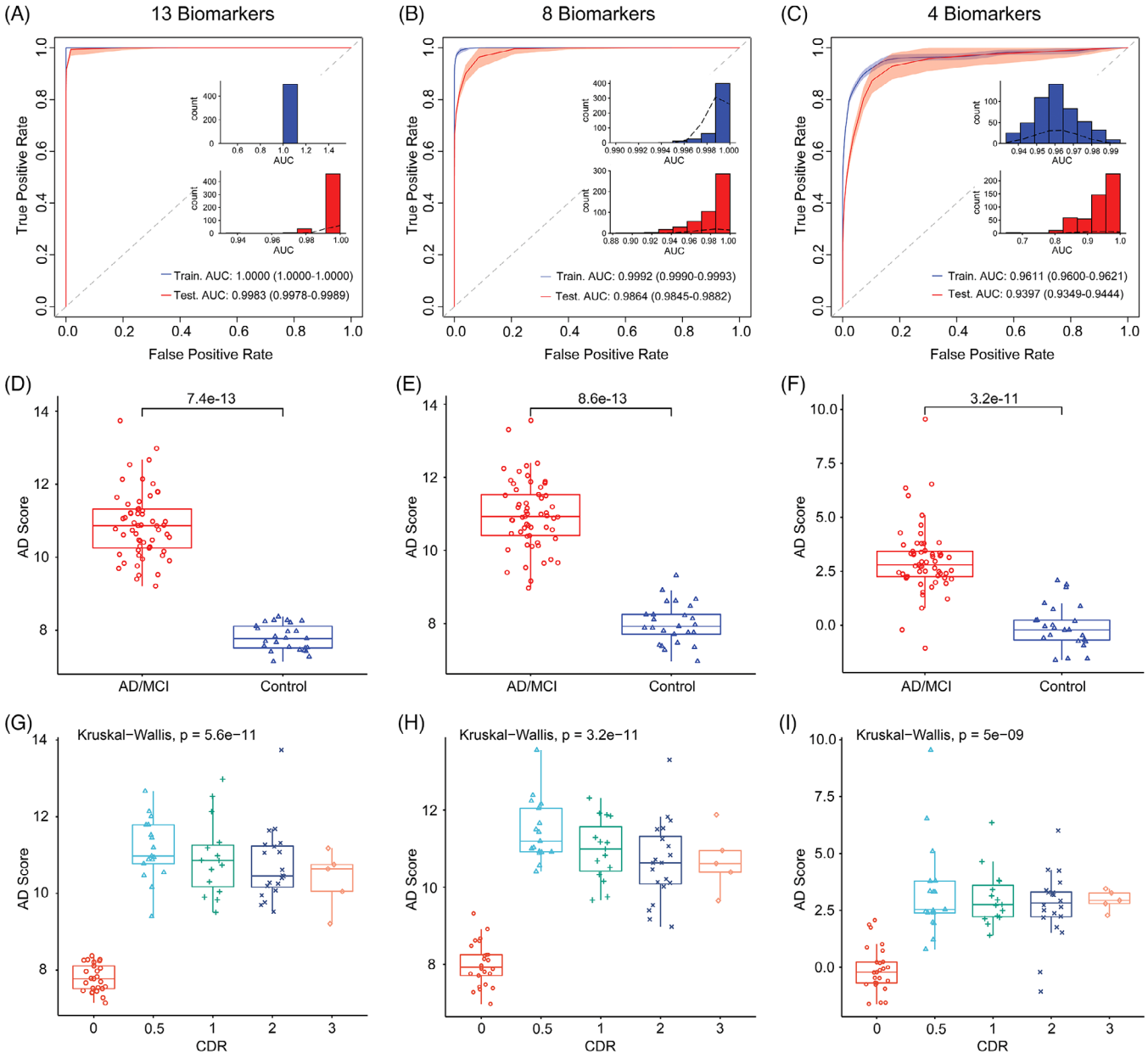

FIGURE 4 Model characterization for AD evaluation. (A–C) Training and testing of an SVM classification model for each biomarker panel with the performance characterized through ROC curves, highlighting the AUC and 95% confidence interval. Inset figures present the distribution of AUC values over 500 iterations, illustrating the model's performance consistency. (D–F) Comparison of AD scores for each sample as generated by the prediction models, differentiating between AD/MCI and control groups. The significance of differences was characterized using the Wilcoxon test. (G–I) Evaluation of correlation between AD scores and CDR scale values. Statistical differences among groups were assessed using the Kruskal-Wallis test. AUC: area under the curve

cross-validation (5-fold inner and outer cross-validation) was first employed to determine the optimal hyperparameters for the SVM model. Subsequently, both the AD/MCI and control groups were divided randomly into a training set (80%) and a testing set (20%) each time for 500 trials of model training and testing, in order to characterize model performance in terms of ROC-AUC value. Figure 4A-C illustrate the mean ROC curves for each biomarker panel. The shading of the curves indicates the standard deviations (SDs) across the 500 training and testing iterations, reflecting the consistency of the models' performance. Additionally, the distribution of AUC values (inset of Figure 4A-C) provides insights into the variation of the models' prediction outcomes over these iterations.

The first biomarker panel (Figure 4A), incorporating all 13 genes, showed outstanding training performance with an AUC of 1.00. This high accuracy continued in testing, achieving a mean AUC of 0.9983 with a tight 95% confidence interval (CI) of 0.9978–0.9989, indicating strong predictive reliability. The second biomarker panel (Figure 4B), refined to eight biomarkers, recorded a mean training AUC of 0.9992

and a testing mean AUC of 0.9864, with a 95% CI of 0.9845 to 0.9882, suggesting a slight reduction in classification capability but still retaining high diagnostic precision. An alternative version of this panel (Figure S3B) had slightly lower AUCs in both training and testing. The final biomarker panel (Figure 4C), with only four genes, exhibited a training AUC of 0.9611. Although lower than the other panels, it still demonstrated strong differentiation capacity. The testing phase for this panel yielded a mean AUC of 0.9397 with a wider 95% CI of 0.9349–0.9444, indicating greater variability in predictive accuracy.

The AD classifier for each marker panel was developed by averaging weights and bias terms from the linear SVM model over 500 iterations. This produced an AD score for each sample, compared between the AD/MCI and Control groups (Figures 4D–F and S3C). As the number of genes in the panels decreased from 13 to 4, the overlap in interquartile ranges (IQRs) between the groups increased, confirmed by a higher *p*-value from the Wilcoxon test. Despite this, the model maintained significant discriminative ability with four gene markers (Figure 4F), shown by a *p*-value of 3.2e-11. Although the AD group has a higher average age, the model can distinguish between the MCI and Control groups (Figure S4) independent of age, with both groups having average ages between 67 and 68 years (Table S2). The AD and MCI groups also have significant differences using 13 and 8 gene markers, showing the model's potential to distinguish between early and late AD stages.

The prediction model effectively distinguished between the Control and AD/MCI groups and correlated strongly with the CDR scale score. Samples were subdivided by CDR scale values to compare AD scores (Figures 4G-I and S3D). Median AD scores increased significantly from the Control group (CDR 0) to the AD/MCI group (CDR 0.5–3), with a gradual decrease as CDR scale values increased from 0.5 to 3. This trend was consistent across the three biomarker panels.

4 | DISCUSSION

In this study, mRNA extracted from plasma EVs of 82 subjects, including AD patients, MCI individuals who later progressed to AD, and healthy controls, was analyzed using NGS (Figure 1). The increase in DEGs from MCI to AD suggests disease progression. KEGG analysis confirmed the clinical significance of these genes as potential AD biomarkers by showing enrichment in the AD pathway. Shared pathways related to neurodegeneration between MCI and AD groups indicate a consistent pathology, supporting the view that MCI is an early stage of AD, crucial for early diagnosis and monitoring. GO analysis revealed distinct biological activities and molecular functions for the MCI and AD groups, with Wnt signaling prevalent in MCI and metabolic pathways emphasized in AD. These results suggest a progression in cellular complexity and potential changes in cellular dynamics from MCI to AD, consistent with the view of AD as a systemic disease.

The identification of shared AD-related DEGs as potential biomarkers offers new opportunities for early detection and monitoring of AD (Figure 2). The 22.5% gene overlap between the MCI and AD groups underscores their potential in understanding the disease. Functional enrichment and pathway analyses reveal key themes of disrupted cellular metabolism and oxidative stress. The involvement of protein complexes and disruptions in insulin- and nucleotide-binding signaling pathways suggests significant contributions to AD pathology. KEGG pathway analysis highlights the genetic basis of AD and suggests its systemic impact, whereas gene-pathway interactions emphasize the roles of specific genes in both neurodegenerative and metabolic disorders, suggesting the complexity of AD.

The ROC analysis of the AD-prediction model demonstrates outstanding performance in both training and testing, indicating the biomarkers' effectiveness in distinguishing patients with AD from healthy individuals (Figure 3). The first biomarker panel, including all 13 shared genes, shows the highest accuracy (AUC >0.99), but the similarity in gene expression suggests potential overfitting. The second biomarker panel, with eight genes, maintains an impressive AUC greater than 0.98, making it more suitable for practical use. Although the third panel with four genes is simpler, it offers less genetic scope and accuracy. Therefore, the second panel, including *INSR*, *ERN1*, *HSD17B10*, *PPIF*, *CHRM1*, *AKT2*, *IDE*, and *ATP2A2*, balances genetic comprehensiveness and expression clarity, thereby optimizing predictive accuracy and minimizing overfitting risks.

The analysis of AD scores derived from the prediction model provides an insightful perspective on AD progression in relation to CDR scale evaluations (Figure 4). The model shows a significant increase in AD scores from CDR scale score of 0 to 0.5, highlighting its sensitivity to early AD stages, likely reflecting early neuropathological changes associated with MCI. This may involve the regulation of specific genes as a compensatory mechanism or response to initial neuronal stress. Conversely, the model shows a gradual decline in AD scores from CDR 0.5 to 3, indicating a reduction in gene regulation as AD severity increases, although the trend is not significant. This could be due to neuronal cell loss and diminished activity in surviving neurons.

In summary, this study analyzed mRNA from plasma EVs in 82 AD samples and healthy controls using NGS, revealing detailed gene-expression profiles and genetic progression from MCI to AD. Functional enrichment and pathway analyses highlighted disruptions in metabolic processes and oxidative stress in neurodegenerative diseases. The research identified potential AD biomarkers and differentiated the molecular functions unique to MCI and AD. The ROC analysis of the AD prediction model showed high accuracy and strong correlation with CDR evaluation, suggesting mRNA profiling from plasma EVs as a promising tool for early detection and monitoring of AD.

The study has several limitations. First, the statistical power was approximately 0.54 with a significant level of 0.05, whereas clinical research prefers a power of 0.80 or higher, typically requiring larger sample sizes.[21,22] Second, the participants were drawn from three different cohorts, leading to unavoidable pre-analytical variations that could also bias the results. Third, there was potential RNA contamination from plasma, which may not have been completely eliminated during purification. To further validate the reliability of the identified markers and optimize the prediction model, future studies should focus on analyzing more clinical samples from larger and more diverse cohorts with different ethnicities, genetic backgrounds, and environmental exposures.

ACKNOWLEDGEMENTS

We thank Knight Alzheimer's Disease Research Center (ADRC) of Washington University in St. Louis and their grants (Healthy Aging and Senile Dementia [P01 AG03991]), ADRC [P30 AG066444], Adult Children Study [P01 AG026276]). We also thank Indiana University for providing samples from the National Centralized Repository for Alzheimer's Disease and Related Dementias (NCRAD), which receives government support under a cooperative agreement grant (U24 AG021886) awarded by the National Institute on Aging (NIA). This work was supported by grants from the National Institutes of Health (NIH) (1R41AG076098-01, 1R43AG080878-01, and 2R44GM144009-02)

CONFLICT OF INTEREST STATEMENT

Le Hoang Phu Pham, Ching-Fang Chang, Katherine Tuchez, and Yuchao Chen are employees of WellSIM Biomedical Technologies, Inc., with pending patent applications (63/609,346). Fei Liu has no conflicts of interest to disclose. Author disclosures are available in the Supporting Information.

CONSENT STATEMENT

The study was approved by the institutional review boards of all participating institutions with written informed consent provided by all participants involved in the study.

REFERENCES

1. Morris JC, Storandt M, Miller JP, et al. Mild cognitive impairment represents early-stage Alzheimer disease. *Arch Neurol*. 2001;58(3):397-405. doi:10.1001/archneur.58.3.397
2. Almkvist O, Basun H, Bäckman L, et al. Mild cognitive impairment–an early stage of Alzheimer's disease? *J Neural Transm Suppl*. 1998;54:21-29. doi:10.1007/978-3-7091-7508-8_3
3. Sheng J, Shao M, Zhang Q, Zhou R, Wang L, Xin Y. Alzheimer's disease, mild cognitive impairment, and normal aging distinguished by multi-modal parcellation and machine learning. *Scientific Reports*. 2020;10(1):5475. doi:10.1038/s41598-020-62378-0
4. Porsteinsson AP, Isaacson RS, Knox S, Sabbagh MN, Rubino I. Diagnosis of early Alzheimer's disease: clinical practice in 2021. *J Prev Alzheimers Dis*. 2021;8(3):371-386. doi:10.14283/jpad.2021.23
5. Evans RW. Complications of lumbar puncture. *Neurol Clin*. 1998;16(1):83-105. doi:10.1016/s0733-8619(05)70368-6
6. Watson LS, Hamlett ED, Stone TD, Sims-Robinson C. Neuronally derived extracellular vesicles: an emerging tool for understanding Alzheimer's disease. *Mol Neurodegener*. 2019;14(1):22. doi:10.1186/s13024-019-0317-5
7. Tian C, Stewart T, Hong Z, et al. Blood extracellular vesicles carrying synaptic function- and brain-related proteins as potential biomarkers for Alzheimer's disease. *Alzheimers Dement*. 2023;19(3):909-923. doi:10.1002/alz.12723
8. Ruan Z, Pathak D, Venkatesan Kalavai S, et al. Alzheimer's disease brain-derived extracellular vesicles spread tau pathology in interneurons. *Brain*. 2021;144(1):288-309. doi:10.1093/brain/awaa376
9. Spiers JG, Vassileff N, Hill AF. Neuroinflammatory modulation of extracellular vesicle biogenesis and cargo loading. *Neuromolecular Med*. 2022;24(4):385-391. doi:10.1007/s12017-022-08704-3
10. Li TR, Wang XN, Sheng C, et al. Extracellular vesicles as an emerging tool for the early detection of Alzheimer's disease. *Mech Ageing Dev*. 2019;184:111175. doi:10.1016/j.mad.2019.111175
11. Lee S, Mankhong S, Kang JH. Extracellular vesicle as a source of Alzheimer's biomarkers: opportunities and challenges. *Int J Mol Sci*. 2019;20(7):1728. doi:10.3390/ijms20071728
12. Garcia-Silva S, Gallardo M, Peinado H. DNA-loaded extracellular vesicles in liquid biopsy: tiny players with big potential? *Front Cell Dev Biol*. 2020;8:622579. doi:10.3389/fcell.2020.622579
13. Thakur K, Singh MS, Feldstein-Davydova S, Hannes V, Hershkovitz D, Tsuriel S. Extracellular vesicle-derived DNA vs. CfDNA as a biomarker for the detection of colon cancer. *Genes (Basel)*. 2021;12(8):1171. doi:10.3390/genes12081171
14. O'Brien K, Breyne K, Ughetto S, Laurent LC, Breakefield XO. RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nat Rev Mol Cell Biol*. 2020;21(10):585-606. doi:10.1038/s41580-020-0251-y
15. Prieto-Vila M, Yoshioka Y, Ochiya T. Biological functions driven by mRNAs carried by extracellular vesicles in cancer. *Front Cell Dev Biol*. 2021;9:620498. doi:10.3389/fcell.2021.620498
16. Sproviero D, Gagliardi S, Zucca S, et al. Extracellular vesicles derived from plasma of patients with neurodegenerative disease have common transcriptomic profiling. *Front Aging Neurosci*. 2022;14:785741. doi:10.3389/fnagi.2022.785741
17. Luo D, Liu H, Liu H, et al. Long RNA profiles of human brain extracellular vesicles provide new insights into the pathogenesis of Alzheimer's disease. *Aging Dis*. 2023;14(1):229-244. doi:10.14336/AD.2022.0607
18. Chen Y, Zhu Q, Cheng L, et al. Exosome detection via the ultrafast-isolation system: EXODUS. *Nat Methods*. 2021;18(2):212-218. doi:10.1038/s41592-020-01034-x
19. Zhu Q, Xu H, Huang L, et al. Identification and detection of plasma extracellular vesicles-derived biomarkers for esophageal squamous cell carcinoma diagnosis. *Biosens Bioelectron*. 2023;225:115088. doi:10.1016/j.bios.2023.115088
20. Zhu Q, Luo J, Li HP, et al. Robust acute pancreatitis identification and diagnosis: RAPIDx. *ACS Nano*. 2023;17(9):8564-8574. doi:10.1021/acsnano.3c00922
21. Brookmeyer R, Abdalla N. Design and sample size considerations for Alzheimer's disease prevention trials using multistate models. *Clin Trials*. 2019;16(2):111-119. doi:10.1177/1740774518816323
22. Ard MC, Edland SD. Power calculations for clinical trials in Alzheimer's disease. *J Alzheimers Dis*. 2011;26(3):369-377. doi:10.3233/JAD-2011-0062

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pham LHP, Chang CF, Tuchez K, Liu F, Chen Y. Assessing Alzheimer's disease via plasma extracellular vesicle–derived mRNA. *Alzheimer's Dement*. 2024;16:e70006. https://doi.org/10.1002/dad2.70006