Received: 3 February 2026 | Revised: 2 May 2026 | Accepted: 3 May 2026

DOI: 10.1002/trc2.70273

Alzheimer's & Dementia
Translational Research & Clinical Interventions

RESEARCH ARTICLE

A diagnostic plasma omics-biomarker for Alzheimer's disease informed by microglial single-cell transcriptomics: A pilot study

Michael W. Lutz[1] | **Zhaohui Man**[1] | **Yifei Zheng**[2] | **Srilakshmi Venkatesan**[1] | **Ornit Chiba-Falek**[1,3]

[1]Division of Translational Brain Sciences, Department of Neurology, Duke University School of Medicine, Durham, North Carolina, USA

[2]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, North Carolina, USA

[3]Center for Genomic and Computational Biology, Duke University School of Medicine, Durham, North Carolina, USA

Correspondence
Ornit Chiba-Falek, Division of Translational Brain Sciences, Department of Neurology, Duke University Medical Center, 311 Research Drive, Box 2900, Bryan Research Building, Room 201A, Durham, NC 27710, USA.
Email: o.chibafalek@duke.edu

Funding information
National Institutes of Health/National Institute on Aging (NIH/NIA), Grant/Award Numbers: RF1 AG077695, R01 AG057522

Abstract

BACKGROUND: The current biomarker framework for the diagnosis and staging of Alzheimer's disease (AD) relies mainly on neuropathological features; thus, its performance for diagnosis is limited prior to the initiation of neurodegeneration. Here, we leveraged transcriptomic data to develop a new framework for omic-informed blood-based diagnostic biomarkers for AD from an early stage.

METHODS: Microglial gene expression from single nucleus RNA sequencing (snRNA-seq) data was analyzed via six statistical methods to identify candidate panels of genes predictive of AD. A total of 78 gene panels, 30 to 2000 genes in size, were selected and evaluated for their ability to distinguish AD patients from controls. Three top-ranked panels of 300, 50, and 30 genes were transferred to blood (monocyte) transcriptomic data obtained from living subjects via a graph-based mapping approach based on optimal transport statistics.

RESULTS: The 300-panel method resulted in an area under the curve (AUC) of 0.7 and moderate accuracy (75%) in classifying AD; however, the accuracy in predicting cognitively normal patients was lower (53%). While the 300 genes provided high accuracy, inspection of the distribution of *P* values for the gene set revealed that the panel could be greatly reduced in size to capture the most significant differences between AD patients and cognitively normal individuals. The accuracy and specificity of the 50 and 30 panels demonstrated similar AUC values but improved the balance between the prediction of AD patients and normal controls. Specifically, the 50-gene panel resulted in an AUC of 0.7, with 65% AD accuracy and 71% normal accuracy.

CONCLUSIONS: Integrating multiomics datasets into the AD biomarker discovery pipeline offers a powerful modality to increase precision and comprehensiveness in AD research and clinical applications.

KEYWORDS
Alzheimer's disease plasma-based omics biomarker, single-cell transcriptomics, microglia, preclinical Alzheimer's disease diagnosis

Highlights

- We developed a biomarker discovery framework based on statistical approaches to identify Alzheimer's disease (AD) brain microglial transcriptomic signatures and apply to blood samples for AD diagnosis.
- Panels of 30 and 50 genes provided moderate accuracy (70%) for disease diagnosis at both early (mild cognitive impairment) and progressed disease stages.
- Our new AD-biomarker discovery framework could be empowered for the development of additional omics biomarkers for other intended uses, including predicting disease risk, disease trajectories/prognosis, and monitoring treatment effects.
- Leveraging omics datasets is a powerful strategy for the development of omics-based biomarkers to increase accuracy and population-based precision of the new generation of AD biomarkers.

1 | INTRODUCTION

The current biomarker framework for the diagnosis and staging of Alzheimer's disease (AD), including the recent and first US Food and Drug Administration–approved AD blood test,[1,2] relies mainly on amyloid beta (Aβ) and tau (primarily phosphorylated tau [p-tau] fragments) pathologies as core markers. Thus, their performance is limited, particularly prior to the beginning of neurodegeneration and atrophy. While assays based on positron emission tomography (PET) imaging and cerebrospinal fluid (CSF) are invasive, expensive, and less accessible, newer plasma tests reduce these limitations and increase diagnostic accessibility.[1,2]

AD initiates decades before clinical symptoms and neuropathological hallmarks appear. The new AD blood test, which is based on the p-tau217/Aβ1-42 plasma ratio, can predict the presence of amyloid plaques and is intended for patients who are cognitively impaired. Thus, the accuracy and sensitivity of the available plasma tests are limited, especially in presymptomatic stages, highlighting the unmet clinical need for new diagnostic biomarkers enabling precise early diagnosis. Furthermore, predictive biomarkers to identify subjects at high risk of developing AD years and even decades before we notice symptoms would be beneficial for the design of the next generation of AD clinical trials; therefore, these biomarkers are crucial for shifting toward prevention treatments or disease-modifying therapies (DMTs) prior to the initiation of severe neurodegeneration processes, including the accumulation of Aβ and tau in the brain.

The role of neuroinflammation and the involvement of microglia in the pathogenesis of neurodegenerative diseases, including AD and related dementias, have been well established.[3–8] Genome-wide association studies (GWASs) have shown that microglia-expressed genes are involved in AD risk, and transcriptomic studies have demonstrated dysregulation of gene expression in microglia from the *post mortem* brains of patients compared to those from age- and sex-matched controls.[9–13] Moreover, it has been suggested that microglia constitute one of the earliest cell subtypes showing gene expression changes in AD.[14–17] We aim to translate this knowledge into the development of accurate, sensitive, and non-invasive diagnostic biomarkers for presymptomatic AD.

In this study, we developed a new biomarker framework and represented the first stage in the development pipeline of a transcriptomic-based AD blood biomarker for living individuals (Figure 1). First, we identified transcriptomic biomarkers based on AD-associated microglial differentially expressed genes (DEGs) using a publicly available single nucleus RNA sequencing (snRNA-seq) dataset.[18,19] We then applied a label transfer approach to translate the discovered microglial DEG signatures onto monocyte samples, both cell types of the myeloid lineage. Finally, we evaluated performance of the microglial transcriptomic signatures to assign phenotypic categories in the monocyte samples. Our approach is novel in the field of blood biomarker discovery for AD diagnosis even in early stages, as the biomarker is derived from brain microglia rather than by direct examination of the blood cells, for example, monocytes. This is a promising strategy because it captures a well-characterized core brain molecular signature associated with immune and inflammatory processes in microglia, an established disease-driving cell type that mediates pathogenic mechanisms contributing to AD onset.

2 | METHODS

2.1 | Study design and data sources

We developed a bioinformatics strategy to evaluate the transcriptional signatures (gene panels) of AD in brain microglia that could be transferred to peripheral blood monocytes for diagnostic applications. snRNA-seq data from microglia were obtained from the Religious Orders Study and Rush Memory and Aging Project (ROSMAP) cohort via the Alzheimer's Disease Knowledge Portal (syn31512863). These samples were derived from the dorsolateral prefrontal cortex tissue of participants with well-characterized cognitive status. Notably, samples from individuals with AD were, on average, 2.4 years older than those from individuals with normal cognition ($P < 0.0001$). Peripheral

blood monocyte (PBMC) bulk RNA-seq data were obtained from the same research program (syn22024498) to serve as the target domain for label transfer; all PBMC samples were obtained *ante mortem*. The PBMC samples were obtained an average of 1.6 years (standard deviation = 2.7) before death. Descriptive statistics for clinical and demographic variables are provided in Table 1. One of the study's aims was to examine predictive accuracy in monocyte samples that were independent from the microglial samples. There were 81 of the 236 monocytes samples that overlapped with the microglial samples, the remainder of the samples ($n = 155$) were unique; all statistical analyses accounted for overlapping samples by calculating predictive accuracy for the non-overlapping samples.

For validation, the prediction of mild cognitive impairment (MCI) was tested in an independent cohort, data were obtained from the Mayo Clinic Study of Aging (MCSA) obtained from the Alzheimer's Disease Knowledge Portal (syn22024998); descriptive statistics are provided in Table S1 in supporting information.

Details on the cohort composition, quality control procedures, and transcriptomic data processing are provided in supporting information.

In contrast to the PBMC samples, these samples are cell-free bulk RNA (cfRNA) from PAXgene blood tubes. While both label transfer performance evaluations provide statistics on predictive accuracy for disease diagnosis from the microglial transcriptome, the PBMC sample results are focused specifically on immune cell activity for later-stage AD while the cfRNA samples are centered on translation from the brain microglial signatures to earlier stage disease (MCI) using samples that would be more easily obtained during a clinical trial.

2.2 | Biomarker gene panel construction

We used a systematic approach to identify optimal gene signatures by testing six complementary statistical feature selection strategies on the microglial training data. Statistical details are described in the supporting information.

2.3 | Deep joint distribution optimal transport architecture

Cross-tissue label transfer uses a deep joint distribution optimal transport (DeepJDOT) neural network framework designed to align feature distributions between the source and target domains while preserving class-discriminative information. The architecture consisted of a shared encoder network with two fully connected layers (input → 2×hidden → hidden dimensions), incorporating batch normalization and dropout regularization (rates of 0.2–0.4) to prevent overfitting. Details of this approach, specifically optimization process, are in the supporting information.

2.4 | Label transfer evaluation and validation

Statistical details on the evaluation of the label transfer approach including receiver operating characteristic (ROC) analysis, uniform manifold approximation and projection (UMAP) visualization, quantification of spatial distributions of cell types and disease states and biological pathway analysis are in the Supporting Information.

RESEARCH IN CONTEXT

1. **Systematic review**: We reviewed existing literature including PubMed for publications on blood biomarker panels for Alzheimer's disease and related dementias (AD/ADRD) based on transcriptomics and proteomics signatures. We investigated statistical approaches for label transfer from tissue samples to blood.
2. **Interpretation**: This study represents a pilot study for a new AD biomarkers discovery framework. We constructed biomarker panels based on AD-associated microglial transcriptomic signatures and applied a label transfer approach to assign phenotypic categories from microglial samples to monocyte samples. Panels of 30 and 50 genes provided moderate accuracy (70%) to predict mild cognitive impairment in an independent cohort of blood samples. The developed transcriptomic-based AD biomarker panel is translational for the intended use as a diagnostic tool based on a blood test.
3. **Future directions**: Findings in other cohorts with peripheral blood monocyte transcriptomic datasets should be replicated. Extending the development of transcriptomic biomarkers using our novel omics biomarkers discovery framework for other intended uses such as predicting disease risk decades before clinical onset is also important, as is refining the transcriptomic biomarkers to capture AD heterogeneity across diverse populations.

3 | RESULTS

3.1 | Selection of gene panels of differentially expressed genes in microglia in AD

The first step identified high-resolution microglial transcriptomic signatures for distinguishing between individuals with and without AD. We obtained snRNA-seq data from the dorsolateral prefrontal cortex (DLPFC) of the ROSMAP cohort, which consisted of 145 samples with a cognitive diagnosis of AD and 125 samples with a diagnosis of normal cognition[18,19] (Table 1). The analysis used the microglial cell pseudobulk RNA-seq data extracted from this DLPFC snRNA-seq dataset and applied six different gene selection statistical methods (Figure 2, Extended Data Supplement in supporting information). It is noteworthy that our strategy was supported by prior publications reporting microglial differential gene expression in the analysis of ROSMAP snRNA-seq data.[20,21]

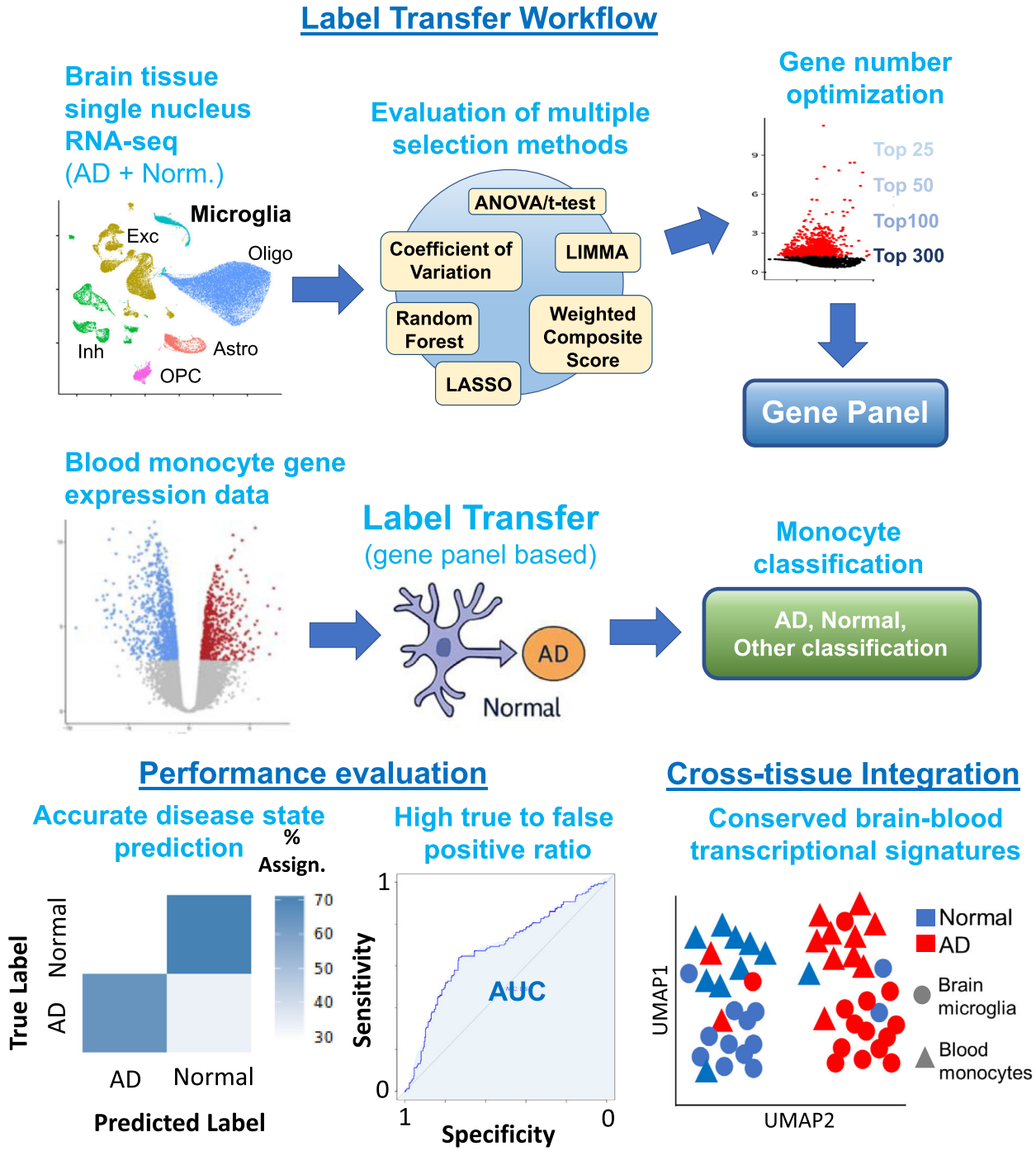

FIGURE 1 The workflow for the development of plasma transcriptomic biomarkers for AD prediction based on the microglial transcriptome. AD, Alzheimer's disease; ANOVA, analysis of variance; AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; OPC, oligodendrocyte progenitor cells; UMAP, uniform manifold approximation and projection.

3.2 | Evaluation of the accuracy and precision of the different microglial gene panels in the classification of AD diagnosis

Next, we evaluated the performance measures of each gene panel (total 78) to predict AD diagnosis versus normal cognition (Figure 3). The performance measures included accuracy (true positives + true negatives)/total predictions (Figure 3A); precision (true positives/[true positives + false positives]; Figure 3B); recall, that is, the proportion of actual positives (AD cases) that were correctly identified (or true positives/[true positives + false negatives]; Figure 3C); and the F1 score, the harmonic mean of precision and recall (or [2 x true positives]/[2 x true positives + false positives + false negatives]; Figure 3D). Overall, the 30-gene panel (Table S2 in supporting information) selected with the combined algorithm and the 50-gene panel (Table S2) based on the *t* test showed the best performance as demonstrated by F1 scores

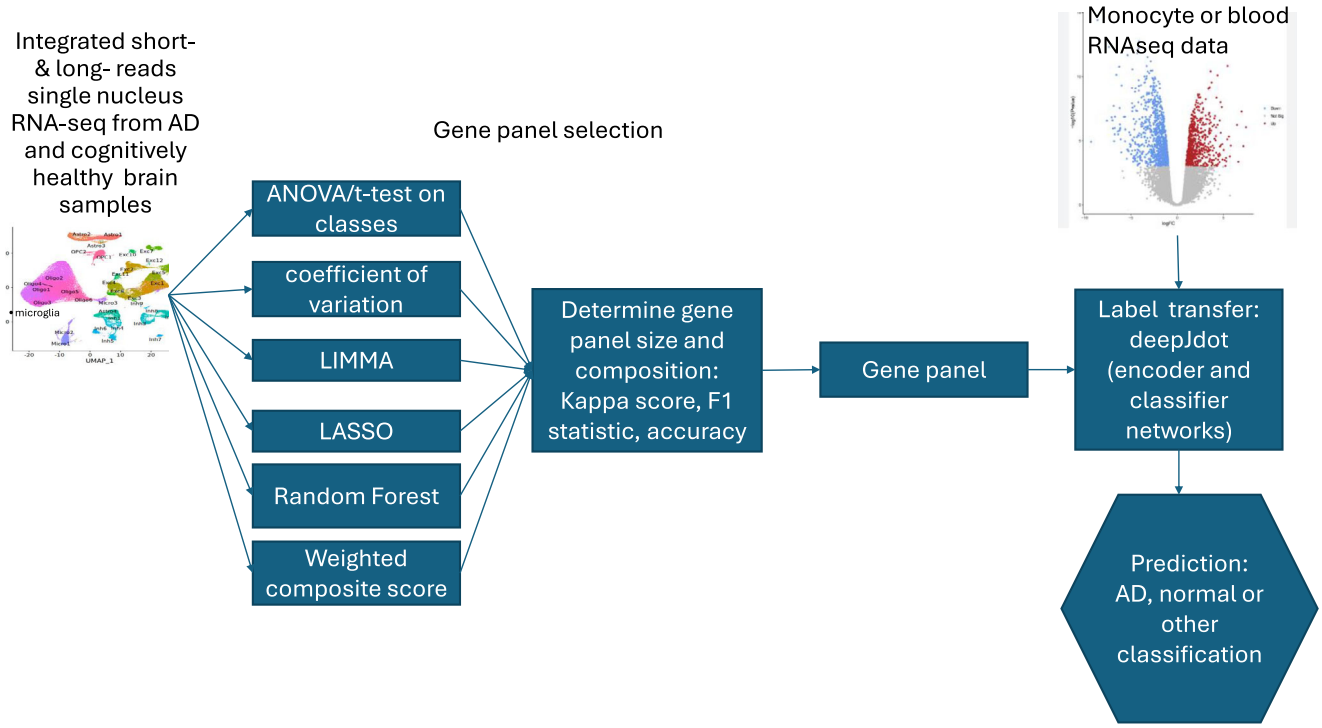

FIGURE 2 Overview of the statistical methods applied in the biomarker discovery pipeline: selection of gene panels based on single-nucleus data from microglia, label transfer from microglia to monocytes, and performance evaluation in monocytes. AD, Alzheimer's disease; ANOVA, analysis of variance; LASSO, least absolute shrinkage and selection operator.

TABLE 1 Clinical and demographic variables for Religious Orders Study and Rush Memory and Aging Project microglial and blood monocyte single nucleus RNA sequencing samples.

	Microglia		Monocytes	
Parameter	AD	NCI	AD	NCI
N	145	125	121	115
Age at death	88.2 (3.4)	85.8 (5.1)	86.8 (4.2)	84.9 (5.5)
N, % female	106, 73%	80, 64%	78, 64%	81, 70%
APOE alleles				
0	87, 60%	100, 80%	77, 65%	91, 81%
1	53, 37%	25, 20%	39, 33%	21, 19%
2	4, 3%	0, 0%	3, 2%	0, 0%
Braak score	4.1 (1.0)	3.1 (1.2)	4.2 (1.0)	3 (1.2)
CERAD score	1.7 (0.9)	2.7 (1.2)	1.8 (1.0)	2.8 (1.2)
PMI	7.3 (4.3)	7.3 (4.3)	7.9 (4.4)	9.5 (8.0)

Abbreviations: AD, Alzheimer's disease; *APOE*, apolipoprotein E; CERAD, Consortium to Establish a Registry for Alzheimer's Disease; NCI, not cognitively impaired; PMI, *post mortem* interval.

(Table S3 in supporting information). In addition, we examined six larger panels of genes (100–2000), and a panel of 300 genes selected by the limma algorithm (Table S2) demonstrated the best performance measures (Table S3). Comparisons between statistical methods for selection and gene panel sizes and specific performance measures are discussed in the Extended Data Supplement.

3.3 | Evaluation of the selected gene panels to classify AD versus normal subjects in monocyte transcriptomic data

In the second step, we used monocyte RNA-seq data[22,23] (Table 1) to test the label transfer performance of three selected panels of 300, 50, and 30 genes (Table S2) and their sensitivity, specificity, and area under the curve (AUC) to classify AD versus normal cognition.

The panel of 300 genes selected by the limma algorithm had an AUC of 0.7 for label transfer to monocytes, with strong accuracy in predicting the AD class (75%). However, the accuracy in the prediction of cognitively normal individuals was lower, at 53%. While the 300-gene panel provided high accuracy, inspection of the *P* value distribution for this gene panel (Figure S1 in supporting information) revealed that the gene panel size could be greatly reduced such that it would capture the most significant differences between the diagnostic criteria. Indeed, when a *P* value threshold of 0.05 ($-\log_{10}[0.05] = 1.3$) was used, the number of genes included in this panel was reduced to only eight genes.

Gene panels of 30 and 50 genes were selected because of the optimal F1 score and as they are potentially more feasible for implementation in clinical settings as transcriptomic biomarkers.

The accuracy and specificity of the 30-gene panel selected by the combined algorithm and the 50-gene panel selected by the *t* test in the label transfer analysis of the monocyte RNA-seq data demonstrated a good balance between the predictions of the AD and normal classes. As shown by the ROC curves, the 50-gene panel had an AUC of 0.7 (Figure 4A), 65% AD accuracy (Figure 4C), and 71% normal

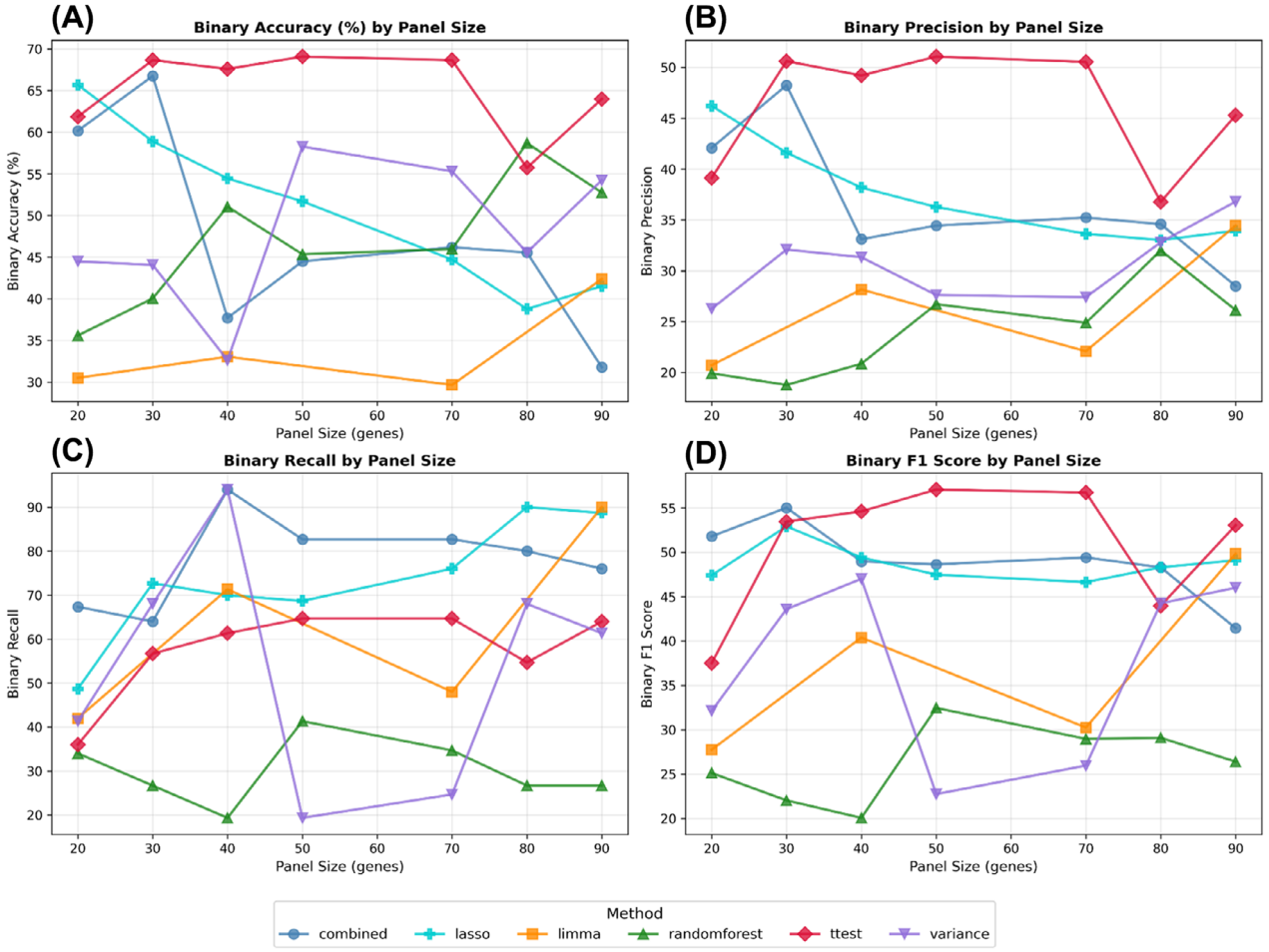

FIGURE 3 Performance measures for each of the six gene selection methods. Each plot shows a statistical measure of the performance of the gene panel differentiating Alzheimer's disease cases from controls for panels of different sizes. A, Accuracy. B, Precision. C, Recall. D, F1 score.

class accuracy (Figure 4C), and the 30-gene panel had an AUC of 0.7 (Figure 4B), 64% AD accuracy (Figure 4D), and 68% normal class accuracy (Figure 4D).

Integrated UMAP diagrams revealed strong clustering by disease for both the microglial samples and the monocyte samples according to UMAP1 scores (Figure 5). The 50-gene panel, formed into two distinct clusters of AD samples for monocytes and one cluster for microglia, was largely separated by UMAP2 values (Figure 5A). In the two AD monocyte clusters, there were only a few normal microglial samples; however, in the AD microglial cluster, we observed some overlap between the normal and AD samples. The normal samples presented one cluster of monocyte samples and one cluster of microglial samples, which were primarily separated by UMAP2 values. The normal monocyte cluster covered a large range of UMAP2 values, with few misclassified AD microglia or monocyte samples. The normal microglial cluster contained predominantly normal samples, with only a few AD samples. The 30-gene panel revealed that most of the microglia and monocyte samples were in one large cluster (Figure 5B). AD microglial and monocyte samples were separable from normal samples in terms of UMAP1 and UMAP2 values but to a lesser extent than the 50-gene panel (Figure 5B).

Regional composition analysis of the UMAP clusters was performed by overlaying a 10×10 grid on the UMAP plots and counting the number of samples labeled by cell type and disease state. The 30-gene panel annotations revealed significant intermingling between cell types and disease states (Figure 5C). There were distinct monocyte-enriched (left cluster) and microglia-enriched (upper/right cluster) regions, but many regions contained both cell subtypes (regions 3, 8, 9, 13, 14, and 16). Disease state (AD/normal) classification was shown primarily within monocyte populations (regions 18, 19, 23, 24, and 25 showed > 85% monocyte–AD purity). However, mixed regions, including region 13 (41% monocyte–normal, 39% monocyte–AD), were indicative that the separation by disease state was limited. In contrast, the 50-gene panel regional composition analysis revealed strong distinct separation by cell type and disease state (Figure 5D). For the cell type, there was clear stratification by UMAP1 values, with monocyte-enriched regions (1, 2, 6–17, 20) showing tight clustering and microglia-enriched/mixed regions (18–26, 33) on the right and dramatically reduced intermingling, in contrast to the 30-gene panel results. For disease status stratification, the monocyte–AD regions (27–32, 34, 35) formed a compact, highly pure cluster (> 93% monocyte–AD purity in regions 27–30, 35), whereas the monocyte–normal regions (1–17) were similarly

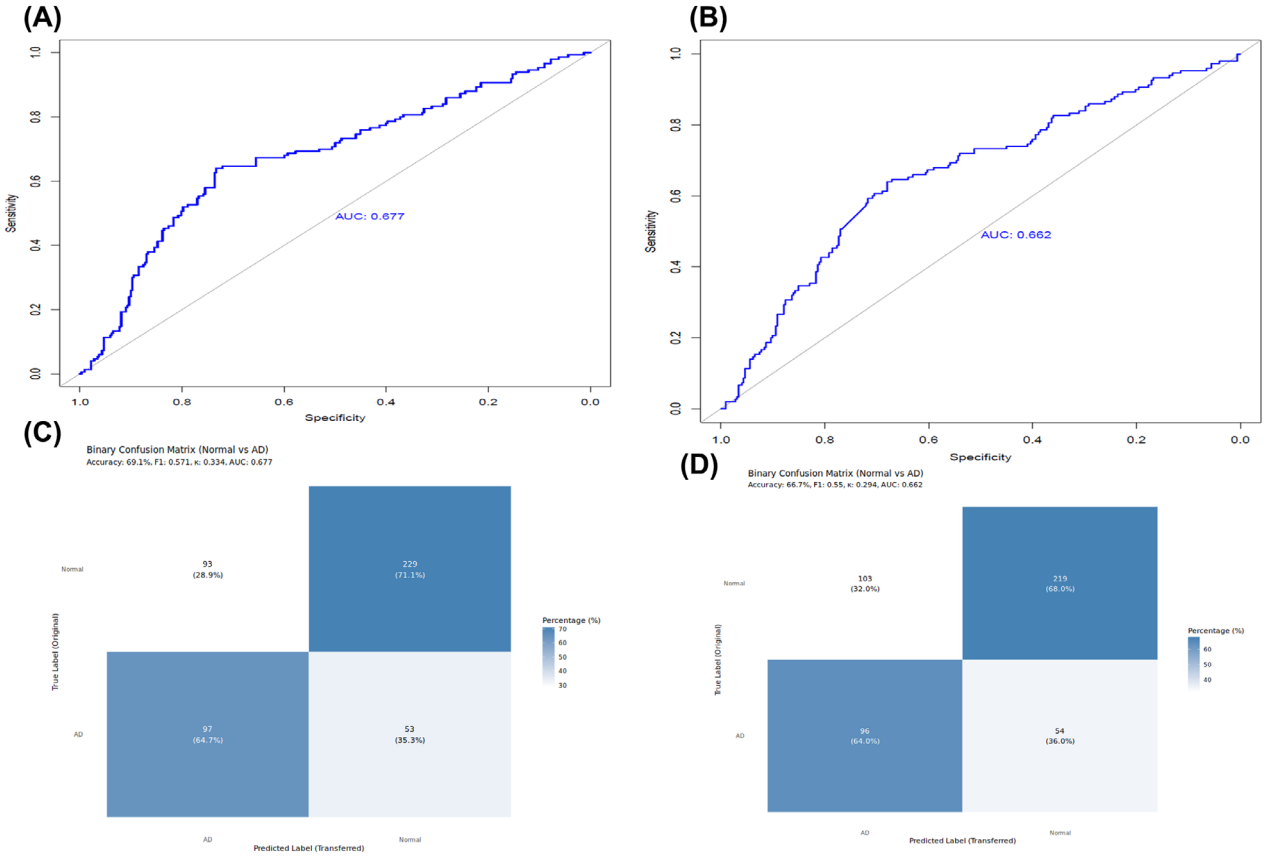

FIGURE 4 Performance of the microglial-based gene panels predicting AD from monocyte data. A, ROC curve for the 50-gene panel, *t* test selection algorithm, (B) ROC curve for the 30-gene panel, combined algorithm for selection. C, Confusion matrix of accuracy values for the 50-gene panel, *t* test selection algorithm. D, Confusion matrix of accuracy values for the 30-gene panel, combined selection algorithm. AD, Alzheimer's disease; AUC, area under the curve; ROC, receiver operating characteristic.

compact and pure, indicating superior AD/normal discrimination in the monocyte cell type. There were fewer regions with mixed compositions than with the 30-gene panel, suggesting more effective feature-based separation of both the cell type and disease status simultaneously.

The 50-gene panel demonstrated substantially improved AD versus normal classification within monocytes; the 30-gene panel results revealed that 8/17 monocyte-containing regions (47%) achieved > 80% AD or normal purity, whereas the 50-gene panel revealed that 17/32 monocyte-containing regions (53%) achieved > 80% purity. Notably, the 50-gene panel resulted in highly enriched monocyte–AD clusters (regions 27–35) with 93% to 100% purity, whereas the 30-gene panel resulted in fragmented AD classification (regions 18–25 with 46%–95% heterogeneity).

Collectively, the results demonstrated that disease status determined clustering across the cell types. The multiple clusters for microglial AD samples suggested that the sample cohort consisted of several AD molecular subtypes. The UMAP plots, most notably for the 50-gene panel, showed strong separation between AD and cognitively normal samples with clusters containing both microglial and monocyte cell subtypes, pointing to conserved transcriptional signatures between the brain and peripheral monocytes, thus supporting our biomarker development strategy.

We compared the gene panels to the results previously reported for snRNA-seq analysis of PBMCs.[24] These AD-associated DEG sets were significant for 1, 8, and 71 genes (false discovery rate < 0.05) from our 30, 50, and 300 gene panels, respectively. The cell subtypes most represented for these DEGs included CD4+, CD8+, and CD56 NK cells.

3.4 | Evaluation of the selected microglial transcriptomic signatures to classify MCI versus normal subjects

We tested the performance of the selected microglial DEG panels as biomarkers for MCI prediction via an independent dataset, the MCSA. This dataset consists of blood samples from diagnoses of cognitively normal individuals and patients with MCI. In this analysis we used a balanced sample of 44 patients with MCI and 44 cognitively normal individuals selected at random out of 374 samples (Table S1). Consistent with the results of the AD diagnosis analysis, the 30-gene panel selected via the combined (composite score) algorithm presented the best performance for panels with < 100 genes, demonstrating an optimal balance between the accuracies to predict both MCI and normal cognition. The prediction of patients with MCI versus

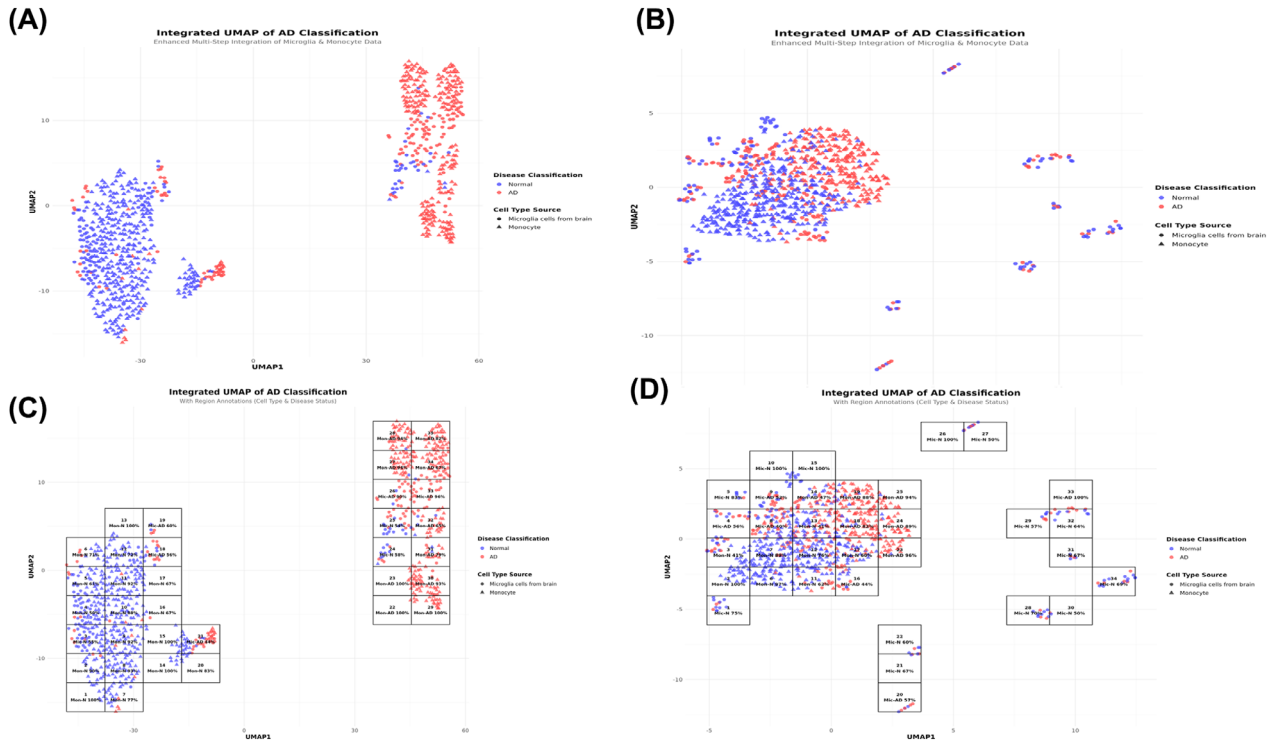

FIGURE 5 Integrated UMAP plots showing clustering of microglial and monocyte data for samples from AD and cognitively normal individuals. UMAP visualization illustrates a low-dimensional graph of the high-dimensional gene expression data. A, UMAP for the 50-gene panel, *t* test selection algorithm. B, UMAP for the 30-gene panel selected with the combined algorithm. C, Cluster mapping and calculations of separation by cell subtype and disease state for the 50-gene panel, *t* test selection algorithm. D, Cluster mapping and calculations of separation by cell subtype and disease state for the 30-gene panel selected by the combined algorithm. AD, Alzheimer's disease; UMAP, uniform manifold approximation and projection.

cognitively normal controls resulted in an AUC of 0.52, with the accuracy of normal class prediction showing the best performance of 64% accuracy, whereas the prediction of patients with MCI was modest, at 41% accuracy. The 50-gene panel selected with the *t* test showed an improved AUC of 0.60, with 66% accuracy in detecting MCI; however, the accuracy in detecting the cognitively normal cases decreased to 18% compared to that of the 30-gene panel. The larger gene panels of 600 genes selected by maximum variance resulted in an AUC of 0.63, with 91% accuracy in detecting the cognitively normal cases and 34% accuracy in detecting MCI. Collectively, the results of this MCI prediction analysis supported the premise that the discovered transcriptomic panels enabled the detection of early-stage cognitive decline associated with AD. Remarkably, the substantially improved performance of the larger gene panel relative to the smaller panels indicated that various transcriptomic signatures may offer different predictive accuracies depending on the stage of the disease (i.e., pre-clinical, early, late) through the course of AD development.

3.4.1 | Biochemical pathway analysis and gene panel overlap with AD GWAS

The network of enriched biological pathways (Figure S2 in supporting information) contained a set of highly connected pathways relevant to neuronal processes and neurodegeneration (Figure S3 in supporting information). For the 300-gene panel, we identified 13 AD GWAS genes. These genes are involved in multiple signaling pathways, including RHO GTPases, RAS GTPases, and MAP kinases. Further details are in Extended Data Supplement.

4 | DISCUSSION

This study is the first step in a novel framework for AD biomarker discovery informed by microglial transcriptomic signatures and based on a blood test. Here, we identified microglial transcriptomic-informed biomarkers and evaluated their performance to diagnose AD at both early and progressed disease stages by leveraging microglial and monocyte RNA-seq datasets, respectively. The results of our pilot study demonstrated the strong *translational* potential of transcriptomic biomarkers as the new generation of blood-based diagnostic biomarkers for AD at early stages in living subjects.

The strategy to develop blood-based biomarkers for early-stage AD diagnosis via microglial transcriptomic data is supported by the following rationale: (1) Microglia play a key role in the etiology of AD, including the contributions of neuroinflammation and immunomodulation pathways to AD pathogenesis.[3,20] Specifically, disease-associated microglia (DAMs) have been identified in several studies and charac-

terized by enhanced phagocytic activity and lipid metabolism changes associated with the disease state, localization around Aβ plaques, and increased expression of *TREM2* and apolipoprotein E (*APOE*).[25–27] (2) snRNA-seq studies identified multiple DEGs in AD microglial cell types and subtypes,[9,28–30] and microglial transcriptomic signatures of AD have been reported.[5,8,20] For example, compared to non-AD brains, 643 deregulated genes were identified in human brains with early amyloid pathology. (3) Microglia are involved in the early stages of AD development. (4) Microglia and PBMCs are both part of the myeloid lineage. Microglia are brain immune cells that are derived primarily from primitive macrophages and can be replenished from circulating monocytes (a type of PBMC) throughout life. Furthermore, these genes share gene expression profiles. Our results revealed some overlap between the genes in the biomarker panels and AD GWAS genes, with the 300-gene panel enriched for genes involved in endocytosis, RNA splicing, and neuron projection development, which are all critical pathways involved in AD. Overall, the results of this study support the strategic plan for biomarker development. Specifically, UMAP analysis revealed clusters that overlapped monocyte and microglia cell populations and were stratified by disease status, confirming the concept of deducing the disease state across these cell types. Furthermore, pathway analysis revealed that biological processes are likely affected in early-stage AD, suggesting their utility for early-stage diagnosis.

The accuracy of our selected gene panels in predicting AD in monocyte samples should be considered in the context of previous blood-based AD biomarker development studies. The performance of six AD blood biomarkers for incident all-cause and AD dementia showed strong predictive performance of 77.5% to 82.6% for p-tau181, p-tau217, neurofilament light chain (NfL), and glial fibrillary acidic protein (GFAP), with high negative predictive values (> 90%) but low positive predictive values (12%–34%).[31] Notably, for the prediction of AD dementia, the AUCs ranged from 0.71 for NfL to 0.77 for p-tau217.[31] Studies of blood-based gene expression data have identified genes related to AD and their ability to predict early AD.[32,33] For external validation experiments (cross-cohort), the AUCs for the prediction of AD were on the order of 0.70 to 0.66,[32] when the Alzheimer's Disease Neuroimaging Initiative (ADNI) was used as the input to predict AddNeuroMed1 (ANM1).[32] Using 1604 DEGs from ANM1 for training yielded AUCs of 0.62 and 0.79 for AD prediction in the ADNI and AddNeuroMed2 cohorts, respectively.[32] Integrative analysis of single-cell and cell-free plasma RNA transcriptomics was used to derive a 340-gene panel for the prediction of AD, with reported AUCs ranging from 0.58 to 0.89 in different validation sets, via several prediction algorithms, including random forest, support vector machine, and logistic regression.[34] A study of plasma proteomics identified panels of proteins that differentiate AD samples from non-AD samples with AUC values of 0.93,[35,36] while these proteins were not specifically linked to markers of neuroinflammation. An evaluation of our 30- and 50-gene panels of monocytes revealed ≈ 70% accuracy in detecting disease (AD and MCI) in an independent cohort, which was comparable to the findings of previous studies. Thus, these two panels are strong candidates as blood-based transcriptomic biomarkers for AD diagnosis. These results provide the proof of concept needed for advancing the development of the identified transcriptomic biomarkers to the next stage in the pipeline, including validation studies using blood samples from living subjects.

AD is a heterogeneous disease manifested by clinical comorbidities, diverse rates of progression, ages of onset, and co-pathologies.[37] The heterogeneity of AD poses challenges to the pipeline of AD biomarker discovery. Moreover, recent multiomics studies, including CSF proteomics,[38,39] brain transcriptomics,[40] and multiomic integration,[41] have identified molecular subtypes of AD. The characterization of multimodal molecular subtypes of AD has provided the premise for identifying AD biomarkers via multiomic data integration.[42] While the present study focused on binary classification for AD to develop a gene panel based on < 50 genes, the approach could be generalized to more strata, including phenotype-based subgroups, that reflect the heterogeneity of AD.

A recent study used an optimal transfer approach to map DEGs from ROSMAP monocyte data to PBMC samples from two independent cohorts, ADNI and ANMerge.[22] This work aimed to predict AD disease subtypes and molecular signatures from blood samples from living individuals based on precise neuropathological diagnoses of *post mortem* brains. Similar to our study, the authors used a single cell type based on *post mortem* neuropathological staging or diagnosis to perform a label transfer for the diagnosis of living patients. However, there are major differences between our study and that of Huang et al.:[22] (1) We based the label transfer process on specific, microglial transcriptomic signatures from the brain to identify an optimal statistical method for developing predictive gene panels for blood-based biomarkers, whereas Huang et al. transferred subtype labels from ROSMAP monocyte samples to PBMC samples within the ADNI and ANMerge cohorts; (2) we used the diagnostic labels of normal cognition, MCI, and AD, whereas Huang et al. used labels of control, asymptomatic AD, typical AD, and low neurofibrillary tangle AD; (3) Huang reported 40% to 80% label transfer accuracy for the blood samples: using their optimal transport approach[22,43] resulted in the highest accuracy (80%); Ingest,[44] the Monocle,[45] and maximum mean discrepancy manifold alignment[46] approaches yielded lower accuracy values. The accuracy of our approach using deep joint distribution optimal transport was ≈ 70%, which is close to the higher accuracies reported by Huang et al.[22]

The study design was rigorous; however, several limitations should be noted. First, the ability to predict AD and/or MCI status warrants replication in other independent cohorts. Nonetheless, the independent MCSA cohort provided valuable information related to the prediction of MCI as an indication of early disease stages. The construction of the gene panel was based on microglial cell types as a whole and did not consider the granulated subtypes of microglia, such as activated microglia and/or disease-associated migroglia. It is possible that microglia signatures could be perturbed by other inflammatory processes related to neurodegeneration but not specific to AD.[26,47,48] Another limitation concerns the disease stage at the time the PBMC samples were collected. The PBMC samples were taken at different times during the time course that the individual participants were followed, years before death in many cases, and therefore it is difficult to ascertain the impact of this time lag on the results. Finally, although

sample sizes had adequate statistical power to detect moderate effect sizes, larger sample sizes would enable the construction of alternative biomarker panels and the testing of panels with smaller effect sizes.

5 | CONCLUSION

This study represents a new frontier in AD biomarker discovery by integrating single-cell omics data, such as transcriptomic data, into the pipeline. The outcomes provide a proof of concept for continuing the development of the 30- and 50-gene panels as transcriptomic biomarkers for AD administered via a blood test. Further development of our investigational new omic-based biomarker will enable accurate and sensitive diagnosis of AD at early presymptomatic stages before the clinical onset of symptoms. Moreover, it is designed to be a blood test and is thus non-invasive and accessible. Here we showed the diagnostic utility of omics biomarkers in clinical settings. In addition, our new AD-biomarker discovery framework could be empowered for the development of additional omics biomarkers for other intended uses, including predicting disease risk, disease trajectories/prognosis, monitoring treatment effects, and so on.[49] These omics biomarkers accurate for intended use will facilitate AD drug discovery and clinical trials through the enablement of precise selection of subjects at preclinical stages and evaluation of patients who are responsive versus non-responsive to disease-modifying therapies.

To conclude, multicomponent biomarker (MCB) approaches[50,51] have become increasingly important for their ability to characterize and dissect disease complexity (for example, diagnosis, staging, progression, and stratification). Integrating omics-based biomarkers into MCB frameworks offers a powerful modality to increase precision and comprehensiveness in AD research and clinical applications.

AUTHOR CONTRIBUTIONS

Conceptualization: Ornit Chiba-Falek and Michael W. Lutz. *Methodology*: Ornit Chiba-Falek, Michael W. Lutz, and Zhaohui Man. *Investigation*: Ornit Chiba-Falek, Michael W. Lutz, Zhaohui Man, Srilakshmi Venkatesan, and Yifei Zheng. *Visualization*: Ornit Chiba-Falek, Michael W. Lutz, Zhaohui Man, and Yifei Zheng. *Funding acquisition*: Ornit Chiba-Falek. *Project administration*: Ornit Chiba-Falek. *Supervision*: Ornit Chiba-Falek. *Writing–original draft*: Ornit Chiba-Falek and Michael W. Lutz. *Writing–review and editing*: Ornit Chiba-Falek and Michael W. Lutz.

ACKNOWLEDGMENTS

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org). The data available in the AD Knowledge Portal would not be possible without the participation of research volunteers and the contribution of data by collaborating researchers. Study data were provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding from NIA grants P30AG10161 (ROS), R01AG15819 (ROSMAP; Genomics and RNA-seq), RC2AG036547 (H3K9Ac), R01AG36836 (RNA-seq), R01AG48015 (monocyte RNA-seq) RF1AG57473 (single nucleus RNA-seq), the Illinois Department of Public Health (ROSMAP), and the Translational Genomics Research Institute (genomic). The snRNAseq data were generated from *post mortem* brain tissue provided by the Religious Orders Study and Rush Memory and Aging Project (ROSMAP) cohort at Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. This work was funded by NIH grants U01AG061356 (De Jager/Bennett), RF1AG057473 (De Jager/Bennett), and U01AG046152 (De Jager/Bennett) as part of the AMP-AD consortium, as well as NIH grants R01AG066831 (Menon) and U01AG072572 (De Jager/St George-Hyslop). This work was funded in part by the National Institutes of Health/National Institute on Aging (NIH/NIA) RF1 AG077695 (OCF) and R01 AG057522 (OCF).

CONFLICT OF INTEREST STATEMENT

Drs. Chiba-Falek and Lutz are coinventors of the related IP. Duke University filed a patent application for the technology developed in this study. Zhaohui Man, Yifei Zheng, and Srilakshmi Venkatesan report no conflicts of interest. Author disclosures are available in the supporting information. Author disclosures are available in the Supporting Information.

DATA AVAILABILITY STATEMENT

The data used for this study are available from the Alzheimer's Disease Knowledge Portal, for the single-nucleus RNAseq data, syn31512863; for the monocyte data, syn22024498; and for the MCI cohort, syn22024998. All the code used for analysis is available at the GitHub repository, https://github.com/NCTrailRunner/microglia-biomarker

ETHICAL APPROVAL

The Religious Order Study (ROS) and Memory and Aging Project (MAP), collectively known as ROSMAP, adhere to a strict ethics statement centered on informed consent, participant autonomy, and responsible data sharing. Both studies received approval from the Rush University Medical Center Institutional Review Board (IRB) and comply with the Declaration of Helsinki and local laws. All participants provided written informed consent and consented to the sharing of their de-identified data and biospecimens with qualified researchers through the RADC Sharing Hub and the AD Knowledge Portal, following IRB procedures. For the MCSA data, all study procedures and ethical aspects were reviewed and approved by Mayo Clinic's IRBs and Olmsted Medical Center. All participants were fully informed about the MCSA study's scope and signed consent forms.

REFERENCES

1. Wang J, Huang S, Lan G, et al. Diagnostic accuracy of plasma p-tau217/Abeta42 for Alzheimer's disease in clinical and community cohorts. *Alzheimers Dement*. 2025;21:e70038.
2. Palmqvist S, Whitson HE, Allen LA, et al. Alzheimer's Association Clinical Practice Guideline on the use of blood-based biomarkers in the diagnostic workup of suspected Alzheimer's disease within specialized care settings. *Alzheimers Dement*. 2025;21:e70535.
3. Leng F, Edison P. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here?. *Nat Rev Neurol*. 2021;17:157-172.
4. Aleman-Villa KM, Armienta-Rojas DA, Camberos-Barraza J, et al. Neuroinflammation across the spectrum of neurodegenerative

diseases: mechanisms and therapeutic frontiers. *Neuroimmunomodulation*. 2025;*32*(1):278-305.
5. Sobue A, Komine O, Yamanaka K. Neuroinflammation in Alzheimer's disease: microglial signature and their relevance to disease. *Inflamm Regen*. 2023;43:26.
6. Colonna M, Butovsky O. Microglia Function in the Central Nervous System During Health and Neurodegeneration. *Annu Rev Immunol*. 2017;35:441-468.
7. Calsolaro V, Edison P. Neuroinflammation in Alzheimer's disease: current evidence and future directions. *Alzheimers Dement*. 2016;12:719-732.
8. Sobue A, Komine O, Hara Y, et al. Microglial gene signature reveals loss of homeostatic microglia associated with neurodegeneration of Alzheimer's disease. *Acta Neuropathol Commun*. 2021;9:1.
9. Gamache J, Gingerich D, Shwab EK, et al. Integrative single-nucleus multi-omics analysis prioritizes candidate cis and trans regulatory networks and their target genes in Alzheimer's disease brains. *Cell Biosci*. 2023;13:185.
10. Friedman BA, Srinivasan K, Ayalon G, et al. Diverse Brain Myeloid Expression Profiles Reveal Distinct Microglial Activation States and Aspects of Alzheimer's Disease Not Evident in Mouse Models. *Cell Rep*. 2018;22:832-847.
11. Griciuc A, Patel S, Federico AN, et al. TREM2 Acts Downstream of CD33 in Modulating Microglial Pathology in Alzheimer's Disease. *Neuron*. 2019;103:820-835 e7.
12. You SF, Brase L, Filipello F, et al. MS4A4A modifies the risk of Alzheimer disease by regulating lipid metabolism and immune response in a unique microglia state. Preprint. *medRxiv*. 2023;2023.02.06.23285545.
13. Zhou Y, Song WM, Andhey PS, et al. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med*. 2020;26:131-142.
14. Gazestani V, Kamath T, Nadaf NM, et al. Early Alzheimer's disease pathology in human cortex involves transient cell states. *Cell*. 2023;186:4438-4453.
15. Fan Z, Brooks DJ, Okello A, Edison P. An early and late peak in microglial activation in Alzheimer's disease trajectory. *Brain*. 2017;140:792-803.
16. Prater KE, Green KJ, Mamde S, et al. Human microglia show unique transcriptional changes in Alzheimer's disease. *Nat Aging*. 2023;3:894-907.
17. Sarlus H, Heneka MT. Microglia in Alzheimer's disease. *J Clin Invest*. 2017;127:3240-3249.
18. Zeng L, Fujita M, Gao Z, et al. A single-nucleus transcriptome-wide association Study implicates novel genes in depression pathogenesis. *Biological Psychiatry*. 2024;96(1):34–43. doi:10.1016/j.biopsych.2023.12.012
19. Green GS, Fujita M, Yang HS, et al. Cellular dynamics across aged human brains uncover a multicellular cascade leading to Alzheimer's disease. Preprint. *bioRxiv*. 2023;2023.03.07.531493..
20. Sun N, Victor MB, Park YP, et al. Human microglial state dynamics in Alzheimer's disease progression. *Cell*. 2023;186:4386-4403.
21. Li B, Wang S, Kerman B, et al. Microglial states are susceptible to senescence and cholesterol dysregulation in Alzheimer's disease. *Aging Cell*. 2025;24:e70189.
22. Huang X, Jannu AJ, Song Z, et al. Alzheimer's disease neuroimaging I, et al. predicting Alzheimer's disease subtypes and understanding their molecular characteristics in living patients with transcriptomic trajectory profiling. *Alzheimers Dement*. 2025;21:e14241.
23. Mustafa Y, Main LR, Mews M, et al. Monocyte inflammation and resilience to Alzheimer's disease: novel genetic risk genes identified by transcriptome-wide association study. *Transl Psychiatry*. 2025;15:481.
24. Grandke F, Fehlmann T, Kern F, et al. A single-cell atlas to map sex-specific gene-expression changes in blood upon neurodegeneration. *Nat Commun*. 2025;16:1965.
25. Keren-Shaul H, Spinrad A, Weiner A, et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell*. 2017;169:1276-1290.
26. Cheng YH, Ho MS. Disease-associated microglia in neurodegenerative diseases: friend or foe?. *PLoS Biol*. 2025;23:e3003426.
27. Miao J, Ma H, Yang Y, et al. Microglia in Alzheimer's disease: pathogenesis, mechanisms, and therapeutic potentials. *Front Aging Neurosci*. 2023;15:1201982.
28. Morabito S, Miyoshi E, Michael N, et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat Genet*. 2021;53:1143-1155.
29. Mathys H, Davila-Velderrain J, Peng Z, et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*. 2019;570:332-337.
30. Fujita M, Gao Z, Zeng L, et al. Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain. *Nat Genet*. 2024;56(4):605-614.
31. Grande G, Valletta M, Rizzuto D, et al. Blood-based biomarkers of Alzheimer's disease and incident dementia in the community. *Nat Med*. 2025;31:2027-2035.
32. Lee T, Lee H. Prediction of Alzheimer's disease using blood gene expression data. *Sci Rep*. 2020;10:3485.
33. Cooper YA, Nachun D, Dokuru D, et al. Progranulin levels in blood in Alzheimer's disease and mild cognitive impairment. *Ann Clin Transl Neurol*. 2018;5:616-629.
34. Wu L, Zhang R, Wang Y, Dai S, Yang N. Integrative single-cell and cell-free plasma RNA transcriptomics identifies biomarkers for early non-invasive AD screening. *Front Aging Neurosci*. 2025;17:1571783.
35. Huang YN, Liu S, Park T, et al. Plasma proteomic analysis identifies proteins and pathways related to Alzheimer's risk. *Alzheimers Dement*. 2025;21:e70579.
36. Kuchenbecker LA, Thompson KJ, Hurst CD, et al. Novel set of plasma proteins classifies Alzheimer's dementia in African American individuals with high accuracy. *Alzheimers Dement*. 2025;21:e70505.
37. Shwab EK, Pathak GA, Harvey J, et al. Leveraging multiomic approaches to elucidate mechanisms of heterogeneity in Alzheimer's disease: neuropsychiatric symptoms, co-pathologies, and sex differences. *Alzheimers Dement*. 2025;21:e70549.
38. Tijms BM, Vromen EM, Mjaavatten O, et al. Cerebrospinal fluid proteomics in patients with Alzheimer's disease reveals five molecular subtypes with distinct genetic risk profiles. *Nat Aging*. 2024;4:33-47.
39. Ali M, Timsina J, Western D, et al. Multi-cohort cerebrospinal fluid proteomics identifies robust molecular signatures across the Alzheimer disease continuum. *Neuron*. 2025;113:1363-1379.
40. Neff RA, Wang M, Vatansever S, et al. Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets. *Sci Adv*. 2021;7(2):eabb5398. doi: 10.1126/sciadv.abb5398
41. Eteleeb AM, Novotny BC, Tarraga CS, et al. Brain high-throughput multi-omics data reveal molecular heterogeneity in Alzheimer's disease. *PLoS Biol*. 2024;22:e3002607.
42. Lerch O, Levine SZ, Sivakumaran S, et al. Editorial: amplifying efficiency and accuracy in dementia drug development. *J Prev Alzheimers Dis*. 2024;11:1180-1182.
43. Flamary R, Courty N, Gramfort A, et al. POT python optimal transport library. *Journal of Machine Learning Research*. 2021;22:1-8.
44. Weinreb C, Rodriguez-Fraticelli A, Camargo FD, Klein AM. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science*. 2020;367(6479):eaaw3381.
45. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566:496-502.
46. Liu J, Huang Y, Singh R, Vert JP, Noble WS. Jointly embedding multiple single-cell omics measurements. *Algorithms Bioinform*. 2019;;143:10.
47. Mirarchi A, Albi E, Arcuri C. Microglia signatures: a cause or consequence of microglia-related brain disorders?. *Int J Mol Sci*. 2024:25(20):10951.

48. Krasemann S, Madore C, Cialic R, et al. The TREM2-APOE pathway drives the transcriptional phenotype of dysfunctional microglia in neurodegenerative diseases. *Immunity*. 2017;47:566-581 e9.
49. Brady LS, Mahinrad S, Edelmayer RM, et al. Beyond Alzheimer's disease—translating biomarker insights across CNS diseases. *Sci Transl Med*. 2025;17(823). doi:10.1126/scitranslmed.adr2511
50. In: Posey Norris SM, Childers E, Carter S, eds. *Multimodal Biomarkers for Central Nervous System Disorders: Development, Validation, and Clinical Integration: Proceedings of a Workshop*. National Academies Press (US); 2023 Aug 29.
51. Agyeman AS, Bandukwala A, Bouri K, et al. US FDA public meeting: identification of concepts and terminology for multicomponent biomarkers. *Biomark Med*. 2023;17:523-531.

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: Lutz MW, Man Z, Zheng Y, Venkatesan S, Chiba-Falek O. A diagnostic plasma omics-biomarker for Alzheimer's disease informed by microglial single-cell transcriptomics: A pilot study. *Alzheimer's Dement*. 2026;12:e70273. https://doi.org/10.1002/trc2.70273