

RESEARCH ARTICLE

Genome-wide transcriptome analysis of A β deposition on PET in a Korean cohort

Tamina Park^{1,2} | Jiyun Hwang³ | Shiwei Liu^{1,2} | Soumilee Chaudhuri^{1,2,4} | Sang Won Han⁵ | Dahyun Yi⁶ | Min Soo Byun^{7,8} | Yen-Ning Huang^{1,2} | Thea Rosewood^{1,2} | Gijung Jung⁷ | Min Jeong Kim⁷ | Hyejin Ahn⁷ | Jun-Young Lee⁹ | Yu Kyeong Kim⁹ | MinYoung Cho^{1,2} | Paula J. Bice^{1,2} | Hannah Craft^{1,2} | Shannon L. Risacher^{1,2,3} | Hongyu Gao^{10,11} | Yunlong Liu^{10,11,12} | SangYun Kim¹³ | Young Ho Park¹³ | Dong Young Lee^{6,7,8} | Andrew J. Saykin^{1,2,3,10} | Kwangsik Nho^{1,2,14}

¹Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, Indiana, USA

²Indiana Alzheimer's Disease Research Center, Indiana University School of Medicine, Indianapolis, Indiana, USA

³Genome and Health Big Data Laboratory Graduate School of Public Health, Seoul National University, Seoul, South Korea

⁴Medical Neuroscience Graduate Program, Stark Neurosciences Research Institute, Indiana University School of Medicine, Indianapolis, Indiana, USA

⁵Department of Neurology, Chuncheon Sacred Heart Hospital, Hallym University College of Medicine, Chuncheon-si, South Korea

⁶Institute of Human Behavioral Medicine, Medical Research Center, Seoul National University, Seoul, South Korea

⁷Department of Neuropsychiatry, Seoul National University Hospital, Seoul, South Korea

⁸Department of Psychiatry, Seoul National University College of Medicine, Seoul, South Korea

⁹Department of Psychiatry, Seoul National University Boramae Medical Center, Seoul, South Korea

¹⁰Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana, USA

¹¹Center for Medical Genomics, Indiana University School of Medicine, Indianapolis, Indiana, USA

¹²Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana, USA

¹³Department of Neurology, Seoul National University Bundang Hospital and Seoul National University College of Medicine, Seongnam-si, South Korea

¹⁴School of Informatics and Computing, Indiana University, Indianapolis, Indiana, USA

Correspondence

Kwangsik Nho and Andrew J. Saykin, Center for Neuroimaging, 355 W 16th Street, Goodman Hall, Ste 4100, Indianapolis, IN 46202, USA.
Email: knho@iu.edu and asaykin@iu.edu

Dong Young Lee, Department of Psychiatry, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul, Republic of Korea (South).
Email: selfpsy@snu.ac.kr

Abstract

INTRODUCTION: Despite the recognized importance of including ethnic diversity in Alzheimer's disease (AD) research, substantial knowledge gaps remain, particularly in Asian populations.

METHODS: RNA sequencing was performed on blood samples from the Korean Brain Aging Study for the Early Diagnosis and Prediction of Alzheimer's Disease (KBASE) to perform differential gene expression (DGE), gene co-expression network, gene-set enrichment, and machine learning analyses for amyloid beta (A β) deposition on positron emission tomography.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

Dong Young Lee, Andrew J. Saykin, and Kwangsik Nho directed the study.

Funding information

Ministry of Science and ICT, South Korea, Grant/Award Number: 2020R1C1C1013718; National Institute on Aging, Grant/Award Number: U01 AG072177; Korean government grants, Grant/Award Numbers: NRF-2014M3C7A1046042, HI18C0630, HI19C0149, RS-2022-00165636, RS-2023-KH136195; Foundation for the National Institutes of Health, Grant/Award Numbers: P30 AG010133, P30 AG072976, R01 AG019771, R01 AG057739, U19 AG024904, R01LM012535, R01LM013463, R01AG068193, T32AG071444, U01AG068057, U01AG072177, U19AG074879

RESULTS: DGE analysis identified 265 dysregulated genes associated with A β deposition and replicated three AD-associated genes in an independent Korean cohort. Network analysis identified two modules related to pathways including a natural killer (NK) cell-mediated immunity. Machine learning analysis showed the classification of A β positivity improved with the inclusion of gene expression data.

DISCUSSION: Our results in a Korean population suggest A β deposition-associated genes are enriched in NK cell-mediated immunity, providing a better understanding of AD molecular mechanisms and yielding potential diagnostic and therapeutic strategies.

KEYWORDS

Alzheimer's disease, amyloid beta, Asian population, diagnosis, Korean cohort, neuroimaging, transcriptomics

Highlights

- Dysregulated genes were associated with amyloid beta (A β) deposition on positron emission tomography in a Korean cohort.
- Dysregulated genes in Alzheimer's disease were replicated in an independent Korean cohort.
- Gene network modules were associated with A β deposition.
- Natural killer (NK) cell proportion in blood was associated with A β deposition.
- Dysregulated genes were related to a NK cell-mediated immunity.

1 | BACKGROUND

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by progressive cognitive decline and functional impairment.^{1,2} AD is the most widespread form of dementia and currently affects > 55 million people globally.³ Despite the increasing emphasis on ethnic diversity in AD research, there remains a substantial gap in understanding molecular mechanisms underlying the pathogenesis of AD in Asian populations.

A key pathological feature of AD is the accumulation of amyloid beta (A β) peptides, which aggregate into amyloid plaques within the brain.¹ A β aggregation is believed to be the initiating event in AD, accumulating in the brain 10 to 20 years before the first symptoms of dementia appear.^{4,5} Amyloid positron emission tomography (PET) has served as a key research tool for more than two decades and has gained approval for clinical use from various global regulatory authorities, including the US Food and Drug Administration and the European Medicines Agency.^{5,6}

Dysregulation of the immune response in both central and peripheral blood is recognized as a pathological characteristic of AD.⁷⁻⁹ Moreover, amyloid deposits can trigger chronic neuroinflammation, leading to the loss of synapses and neurons, which in turn contributes to brain atrophy and cognitive decline.^{10,11} Recent studies have shown that both peripheral innate immunity influences AD pathology by altering the function and number of peripheral immune cells in the blood,

facilitating their infiltration into the brain and aggravating neuropathy in AD.^{8,12,13}

Blood-based biomarkers are advantageous as they can be extracted more easily and affordably than other body fluids, such as cerebrospinal fluid, and are well suited for high-throughput testing.¹⁴ Blood biomarkers have emerged as a vital element in dementia research, providing key insights into AD pathology and offering prospects for clinical application.¹⁵⁻¹⁷ In high-throughput analysis, prior studies have established the significance of transcriptomic changes in brain tissue across various regions for identifying molecular subtypes of AD,¹⁸ directing target gene selection and validation through meta-analysis¹⁹ and pinpointing gene isoform alterations associated with AD.²⁰ Only a few studies investigated blood transcriptomes in AD research, finding significant changes in expression related to the dysregulated biological processes of AD.^{21,22}

In this study, we aim to identify significant gene expression changes associated with brain A β deposition in AD, focusing on identifying networks and their pathways in blood-based transcriptome and evaluating its potential for blood-based molecular biomarkers in a Korean cohort. We examined blood-based transcriptome data from 506 well-characterized participants, spanning various ages and cognitive states, from the Korean Brain Aging Study for the Early Diagnosis and Prediction of Alzheimer's Disease (KBASE) cohort to identify dysregulated genes as significantly associated with A β deposition on PET. To replicate these findings, we used blood-based transcriptome data from

an independent Korean cohort, Seoul National University Bundang Hospital (SNUBH). In addition, we conducted weighted co-expression network analysis, functional enrichment analysis, and immune cell deconvolution analysis to identify biological pathways and immune cell types associated with A β deposition.

2 | METHODS

2.1 | Participants and blood sample collection

The KBASE is a Korean cohort study that began in 2014.²³ The KBASE study description, along with the recruitment of participants, ethics approval, sampling, and neuroimaging, is detailed in Byun et al.²³ The cohort includes well-characterized participants ranging from young to elderly, covering various cognitive states such as young cognitively normal (CN), old CN, mild cognitive impairment (MCI), and AD dementia.²³ As described in an earlier KBASE cohort study,²³ genomic DNA was extracted from whole blood, and apolipoprotein E (APOE) genotypes were determined using established methods.²⁴ APOE ϵ 4 carrier status was defined based on the presence or absence of the APOE ϵ 4 allele, classifying participants as either APOE ϵ 4-positive or APOE ϵ 4-negative.

Blood samples from individuals were collected in PAXgene tubes at three different time points spanning the years 2014 to 2018, with 2-year intervals between each collection. Because some individuals only had one or two samples collected, we revised the collection time point labels to correspond to the visit occurrences, categorizing them as first (baseline), second, or third visits. Furthermore, participant age was adjusted to reflect the age at each visit, based on the age at enrollment and the corresponding sample collection year. In this study, we analyzed 506 gene expression profiles from baseline samples across the cognitive states of old CN, MCI, and AD to investigate early expression changes associated with A β deposition as measured by PET imaging.

Total RNA, including microRNA, was extracted at the National Centralized Repository for Alzheimer's Disease using the QIAcube and the PAXgene Blood miRNA kit (Qiagen), which included an on-column DNase treatment to remove residual DNA. RNA concentration was determined by measuring the absorbance at 260 nm (A260) on a NanoDrop 2000 UV spectrophotometer (Thermo Scientific). RNA purity was assessed using the A260/A280 and A260/A230 ratios. RNA integrity was evaluated using an Agilent 2100 Bioanalyzer with the Agilent RNA 6000 Nano Kit, which generated an RNA Integrity Number (RIN) for each sample. Aliquots of 1 μ g of DNase-treated RNA were sent to the Indiana University School of Medicine Center for Genomics and Bioinformatics, Genome Services Facility for library preparation and sequencing. The Xpose (Trinean) and TapeStation (Agilent) were used to conduct quality checks after the RNA samples were concentrated to 60 ng/ μ l using the Speedy vacuum.

RESEARCH IN CONTEXT

- 1. Systematic review:** The authors reviewed the existing literature using traditional sources including PubMed and Google Scholar and found limited studies investigating blood-based transcriptomics data for Alzheimer's disease (AD) and amyloid beta (A β) deposition on positron emission tomography in a Korean population.
- 2. Interpretation:** Our study in a Korean cohort revealed significant changes in gene expression levels in blood associated with A β deposition in AD. The dysregulated genes were related to a natural killer (NK) cell-mediated immunity, suggesting an involvement of immune-related pathways in AD pathogenesis, providing a better understanding of AD molecular mechanisms, and yielding potential diagnostic and therapeutic strategies.
- 3. Future directions:** Future research should aim to investigate longitudinal RNA sequencing data to enhance our understanding of disease progression and provide potential prognostic and therapeutic strategies.

2.2 | Total RNA sequencing (RNA-Seq)

2.2.1 | Library preparation and sequencing

A total of 1140 RNAs were isolated from whole blood samples collected from participants in the KBASE cohort using the PAXgene Blood RNA kit (Qiagen). The RNA samples were subsequently treated with DNase (in-solution). One hundred nanograms of DNase-treated total RNA was used for library preparation with the Illumina Stranded Total RNA Prep, Ligation with Ribo-Zero Plus kit, following the manufacturer's instructions. The kit reagent and protocol included human globin RNA and rRNA depletion. Each resulting uniquely dual-indexed library was quantified and quality assessed by Qubit and Agilent TapeStation. Multiple libraries were pooled in equal molarity. The pooled libraries were sequenced on an Illumina NovaSeq 6000 sequencer to generate paired-end 150 bp reads. One hundred million paired-end reads were targeted for each sample.

2.2.2 | Preprocessing of RNA-Seq data

The raw 150 bp paired-end sequencing reads were processed through trimming steps to remove adapter sequence, low-quality (Q < 20), or short reads (length < 36 bp), using trimmomatic v0.39.²⁵ The cleaned reads were then aligned to the human reference genome using STAR v2.7.10a²⁶ with a two pass mode. Human reference genome release 44 (GRCh38.p14) and a primary assembly annotation file were

downloaded from Gencode.²⁷ To obtain the gene expression matrix, the aligned reads were quantified by counting gene IDs in the annotation file using featureCounts v2.0.3.²⁸

2.3 | Gene expression normalization

The longitudinal RNA-Seq data, consisting of 1140 samples from 538 participants, was divided into cognitive state groups: young CN ($n = 93$), old CN ($n = 571$), MCI ($n = 138$), AD ($n = 85$), and others ($n = 41$). Genes with low expression levels were excluded if their counts were below one count per million (CPM) in the smallest group, which is the young CN group ($n = 93$). To eliminate technical variability, such as differences in gene length and guanine-cytosine content (GC content), the gene expression matrix of a total of 1140 KBASE participants was normalized applying the conditional quantile normalization (CQN) approach.²⁹ Of these, 506 gene expressions at baseline were included in the analysis, excluding those from the young CN group, defined as ≤ 55 years old, as well as samples from individuals with unclear cognitive states (others) group.

2.4 | A β PET processing

The [¹¹C] Pittsburgh compound B (PiB) PET tracer was used as the amyloid neuroimaging biomarker. The specifics of the amyloid PET image acquisition and processing have been previously documented.²³ In summary, the PiB PET data underwent standard corrections and were reconstructed into a 256×256 image matrix using iterative techniques. During image processing, static PiB PET images were motion corrected and aligned with each participant's T1 structural images. These T1 images were normalized to standard Montreal Neurological Institute space, and the same transformation was applied to the PiB PET images. The final PET scans were intensity normalized to produce standardized uptake value ratio (SUVR) images using a cerebellar gray matter region of interest from the Centiloid project,³⁰ which were then smoothed. The A β positivity threshold was established in earlier research by applying a receiver operator characteristic (ROC) curve using global cortical A β deposition at baseline to classify young CN from AD individuals. The optimal cut-off was determined to be 1.2373, achieving 83% sensitivity and 100% specificity. We used a log-transformed global cortical PiB SUVR to identify dysregulated genes related to A β deposition, and a cut-off of 1.2373 to determine A β positivity.

2.5 | Differential gene expression analysis

A linear regression model was applied to evaluate the association between expression levels of genes and global cortical PiB SUVR. The model included biological covariates such as age, sex, and APOE $\epsilon 4$ carrier status, as well as technical covariates including batch effects and RIN values. The same sequencing date was used to define the batch. Only protein-coding genes were included for differential gene expression (DGE) analysis. After fitting the model, the regression coef-

ficients were used to interpret them as \log_2 (fold changes) because our expression data were \log_2 -transformed. The fold change means that for every one unit increase in A β deposition, gene expression is expected to double. The P values underwent a false discovery rate (FDR)-based multiple comparison adjustment with the Benjamini-Hochberg (BH) method.³¹ Genes were considered significant based on an FDR threshold of < 0.05 .

Next, we analyzed the association between expression levels of genes and compared AD to CN using a logistic regression model. The model included age, sex, APOE $\epsilon 4$ carrier status, batch, and RIN values as covariates. The \log_2 (fold changes) were considered by the regression coefficient. We investigated 947 potential therapeutic candidate genes identified by Agora³² to examine the overlap between significantly dysregulated genes and the nominated target genes.

2.6 | Replication analysis in an independent Korean cohort, SNUBH

2.6.1 | Participants and RNA-Seq

The SNUBH cohort consists of participants categorized into two diagnosis groups, CN ($n = 91$) and AD ($n = 102$). For the SNUBH cohort, a PAXgene Blood RNA kit (Qiagen) was used for RNA extraction. The library preparation with the TruSeq Stranded Total RNA Library Prep Globin Kit (Illumina) and RNA-Seq on the NovaSeq 6000 platform were performed at the Macrogen in South Korea. Raw sequencing read quality was assessed using FastQC v0.11.7 (www.bioinformatics.babraham.ac.uk/projects/fastqc/), and low-quality reads were filtered out using Trimmomatic v0.38.²⁵ The remaining reads were aligned to the GRCh38 human reference genome (NCBI, Annotation Release 109.20200522) with HISAT2 v2.1.0,³³ and gene and transcript abundances were quantified using StringTie2 v2.1.3b.³⁴

2.6.2 | Cross-cohort replication

In the SNUBH cohort, gene expression normalization and association analysis were conducted according to Methods 2.3 and 2.5. Overall, genes exhibiting low expression (CPM < 1) in the smallest group, CN ($n = 91$), were excluded and normalized using the CQN method.²⁹ To identify genes significantly dysregulated in AD compared to CN individuals in the SNUBH cohort, a logistic regression model was applied, considering sex, age, APOE $\epsilon 4$ carrier status, and RIN values as covariates.

To combine results from both the KBASE cohort (discovery) and the SNUBH cohort (replication), a fixed effect meta-analysis with an inverse variance weighted approach was performed using METAL.³⁵ A positive z score indicated upregulation in AD, while a negative z score indicated downregulation. The P value represents the statistical significance of the association between expression levels of a gene and AD, and it was adjusted for multiple comparisons using the BH method.³¹ We defined suggestive genes as those exhibiting the same direction of regulation across both independent cohorts and having an FDR of < 0.05 .

2.7 | Weighted gene co-expression network analysis (WGCNA)

The WGCNA³⁶ R package was used to construct network-based co-expression modules, using the CQN normalized gene expression table previously used in the DGE analysis. Given that WGCNA identifies modules of co-expressed genes based on expression correlation patterns, the presence of batch effects or RIN variations can generate inaccurate correlations. By regressing out these technical variables, we enhanced the accuracy and relevance of the co-expression networks generated through WGCNA. To filter out non-variable genes from all protein-coding genes, the upper quartile (Q3), representing 75% of genes with the highest standard deviations, was selected. Co-expression modules were constructed using a soft-thresholding power of 10, with a minimum requirement of 30 genes per module. The biweight midcorrelation method was used to calculate the similarity, and the network type to calculate adjacency was set as signed.

The association between the module eigenvalue (ME) and global cortical PiB SUVR was analyzed using a linear regression model, with age, sex, and *APOE* ϵ 4 carrier as covariates. All *P* values were adjusted for multiple comparisons using the FDR correction through the BH procedure.³¹ The significant modules were defined based on an FDR threshold of < 0.05 . We also checked the overlapping genes between significant modules and nominated target genes from Agora.³²

2.8 | Functional enrichment analysis

Over-representation analysis,³⁷ also known as functional enrichment analysis, was executed on the Gene Ontology (GO) biological process,³⁸ using the *enrichGO* function from the *clusterProfiler*³⁹ v4.12.2 packages in R v4.4.1. To identify GO biological pathways, dysregulated genes related to global cortical PiB SUVR were selected based on an FDR threshold of < 0.1 . Additionally, genes from network modules showing significant associations with global cortical PiB SUVR were included in the functional enrichment analysis. GO biological processes with BH-adjusted $P < 0.05$ were considered significant. To reduce complexity and enhance interpretability, highly similar GO biological processes with a similarity score > 0.7 were filtered using the *simplify* function, ensuring that only representative terms were retained. After this similarity assessment, the biological processes were clustered using the *APCluster*⁴⁰ method via the *SimplifyEnrichment*⁴¹ packages v1.14.0 in R v4.4.1. The representative terms for each cluster were selected based on the BH-adjusted P value. The network illustrating the enrichment between genes and the 10 biological pathways was visualized using *Cytoscape*⁴² v3.10.2.

2.9 | Deconvolution analysis

Deconvolution analysis is a computational approach designed to estimate the relative abundance of cell types from bulk RNA sequencing data, aiding in the resolution of data complexity.⁴³ Immune cell

type proportions in blood-based RNA-Seq data were estimated using *CIBERSORTx*⁴⁴ with CQN normalized counts. For cell expression imputation, we applied the LM22 signature matrix, which includes 547 genes representing 22 immune cell phenotypes. We conducted 500 permutations to obtain a deconvolution P value. Immune cell types with an estimated mean proportion of < 0.01 were excluded, and the remaining 11 immune cells were used for the association analysis. A linear regression model was used to assess the relationship between the estimated cell type proportions and $A\beta$ deposition, adjusting for covariates such as age, sex, *APOE* ϵ 4 carrier status, batch, and RIN. The P values were corrected using the BH method, with significance determined by an FDR of < 0.05 .

2.10 | Machine learning classification

A normalized gene expression matrix, adjusted for batch and RIN values and used for WGCNA analysis, was subjected to STREAMLINE release Beta 0.3.4.^{45,46} STREAMLINE is an automated end-to-end machine learning (ML) platform that specializes in binary classification and evaluates various established ML algorithms.^{45,46} We conducted classification using 5-fold cross-validation with the stratified partition method, applying default settings for scaling.^{45,46} Feature importance (FI) was evaluated using the score calculated by the *multiSURF* method,^{45,46} and the top 30 genes were selected according to their median FI scores. We investigated three models to determine the improvement in classification power for $A\beta$ positivity with each successive model. The $A\beta$ -positive group included individuals with a high likelihood of $A\beta$ deposition, defined by a global cortical PiB SUVR in the brain exceeding a threshold of 1.2373. Model 1 was based on age and sex. Model 2 added *APOE* ϵ 4 carrier status to the Model 1. Model 3 further expanded by adding 30 selected genes associated with $A\beta$ deposition to Model 2. We compared 10 ML algorithms to identify which ML algorithm performs best: artificial neural networks, decision tree, elastic net, genetic programming, gradient boosting, k-nearest neighbors, logistic regression, random forest, support vector machines, and naive Bayes. To compare and evaluate the performance of each algorithm, we examined several metrics including balanced accuracy, accuracy, F1 score, sensitivity, specificity, and ROC area under the curve (AUC).

3 | RESULTS

3.1 | Participants characteristics

In this study, we used blood transcriptomic data from the KBASE cohort as a discovery dataset to identify dysregulated genes and their biological pathways associated with brain $A\beta$ deposition. The comprehensive workflow is depicted in Figure 1, encompassing a range of systems biology approaches, including DGE analysis, co-expression network module analysis, functional enrichment analysis, immune cell type deconvolution, and ML classification. Demographic

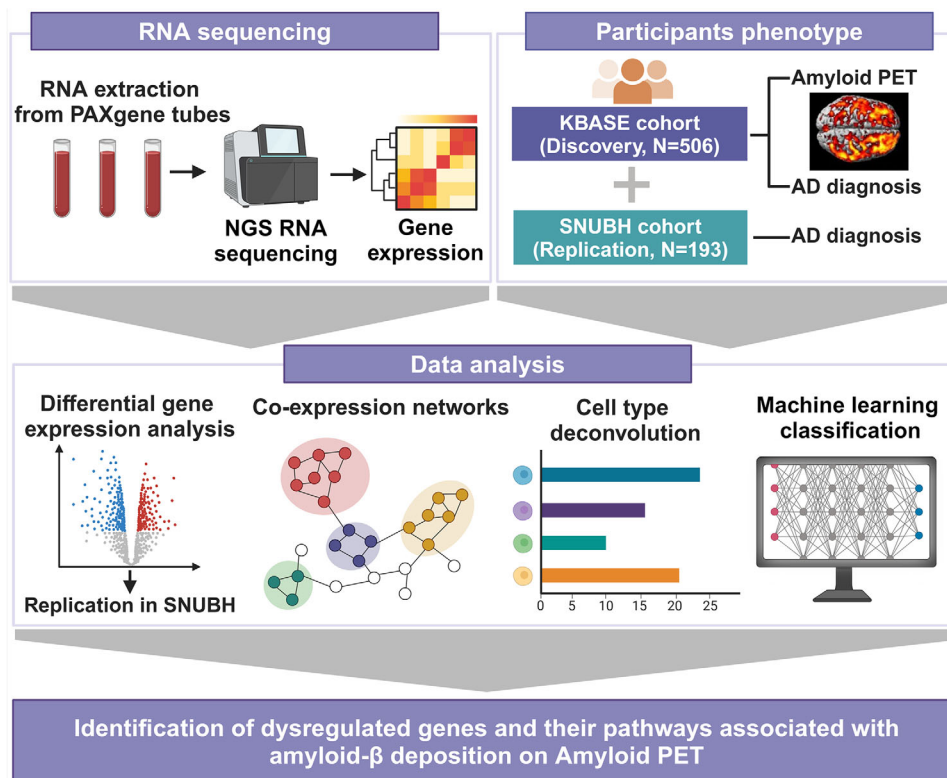


FIGURE 1 Overview workflow of the study. We generated blood-based transcriptomics data on participants from two Korean cohorts (KBASE as a discovery data and SNUBH as a replication data) and performed systems biology approaches to identify dysregulated genes and their pathways for global A β deposition on PET. Diagram was created with BioRender.com. A β , amyloid beta; AD, Alzheimer's disease; KBASE, Korean Brain Aging Study for the Early Diagnosis and Prediction of Alzheimer's Disease; NGS, next generation sequencing; PET, positron emission tomography; SNUBH, Seoul National University Bundang Hospital.

TABLE 1 Participant characteristics.

	KBASE (discovery, N = 506)			SNUBH (replication, N = 193)	
	CN (N = 274)	MCI (N = 138)	AD (N = 94)	CN (N = 91)	AD (N = 102)
Age (mean \pm SD)	69.2 \pm 8.34	73.6 \pm 6.93	73.6 \pm 7.48	72.4 \pm 5.39	79.4 \pm 5.14
Sex (female:male)	138:136	90:48	66:28	68:23	65:37
RNA integrity number (mean \pm SD)	7.31 \pm 2.06	6.99 \pm 2.05	6.22 \pm 2.75	7.73 \pm 0.56	7.71 \pm 0.56
APOE ϵ 4 carrier status (absence:presence)	218:51	86:52	29:55	76:15	50:52
Global cortical PiB SUVR (mean \pm SD)	0.07 \pm 0.05	0.13 \pm 0.08	0.20 \pm 0.09	-	-
A β positivity (negative:positive)	228:41	66:72	15:69	-	-

Abbreviations: A β , amyloid beta; AD, Alzheimer's disease; APOE, apolipoprotein E; AD, Alzheimer's disease; CN, cognitively normal older adults; KBASE, Korean Brain Aging Study for the Early Diagnosis and Prediction of Alzheimer's Disease; MCI, mild cognitive impairment; PiB, [11C] Pittsburgh compound B; SD, standard deviation; SNUBH, Seoul National University Bundang Hospital; SUVR, standardized uptake value ratio.

characteristics of all individuals in the present study are shown in Table 1. We used RNA-Seq data from 506 participants at baseline across three diagnostic groups of old CN, MCI, and AD to investigate dysregulated genes and their pathways associated with global cortical A β deposition measured by PET imaging. The SNUBH cohort, used as a replication cohort, includes a total of 193 participants: CN (N = 91) and AD (N = 102).

3.2 | DGE analysis for global cortical A β deposition

We performed DGE analysis to identify dysregulated genes as significantly associated with A β deposition using a linear regression model ($n = 506$, 13,603 protein-coding genes). The DGE results are described in Figure 2 and listed in Table S1 in supporting information. We identified 265 differentially expressed genes (DEGs) that were significantly

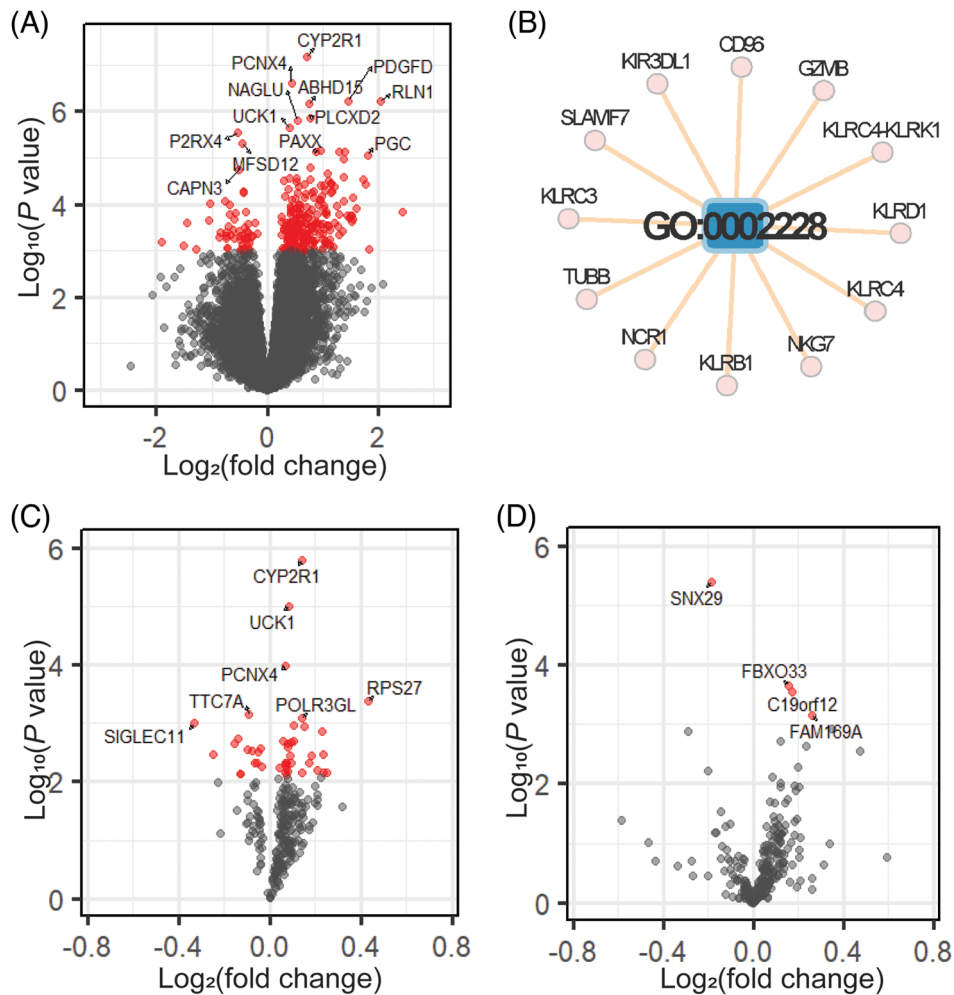


FIGURE 2 Differential gene expression analysis. A, The volcano plot shows 265 genes as significantly associated with global cortical deposition on PET in the KBASE cohort (Discovery data). B, The network represented a NK cell-mediated immunity pathway (GO:0002228) and involved 12 upregulated DEGs for global cortical deposition. C, The volcano plot shows 44 significantly dysregulated genes in AD compared to CN in the KBASE cohort. D, The volcano plot represents three replicated AD-associated genes (*SNX29*, *FBXO33*, and *C19orf12*) and one suggestive gene (*FAM169A*) in an independent Korean cohort, SNUBH (Replication data). In the volcano plots, red dots represented significant genes with FDR-corrected P value < 0.05 . AD, Alzheimer's disease; CN, cognitively normal older adults; DEG, differentially expressed gene; FDR, false discovery rate; KBASE, Korean Brain Aging Study for the Early Diagnosis and Prediction of Alzheimer's Disease; NK, natural killer; PET, positron emission tomography; SNUBH, Seoul National University Bundang Hospital.

associated with $A\beta$ deposition (FDR < 0.05 ; Figure 2A). Of these, 214 genes were upregulated, and 51 genes were downregulated in the presence of higher global cortical PiB SUVR values. We explored potential AD therapeutic target genes sourced from Agora³² and compared them to 265 DEGs associated with $A\beta$ deposition and found that seven genes overlapped (Figure S1 in supporting information). In addition, we identified 775 genes that were associated with $A\beta$ deposition (FDR-corrected P value < 0.05). Out of the 775 genes, 592 were found to be upregulated. From these upregulated genes, only one biological pathway, "natural killer cell mediated immunity" (GO:0002228), was identified as significant (FDR-corrected P value < 0.05). This pathway comprised 12 genes: *CD96*, *GZMB*, *KIR3DL1*, *KLRB1*, *KLRC3*, *KLRC4*, *KLRC4-KLRK1*, *KLRD1*, *NCR1*, *NKG7*, *SLAMF7*, and *TUBB* (Figure 2B).

3.3 | DGE analysis for AD

Using a logistic regression model, we investigated whether the 265 DEGs associated with $A\beta$ deposition also show differential expression in AD compared to CN. Among these genes, 44 exhibited differential expression between the AD and CN groups, with 30 genes upregulated and 14 genes downregulated (FDR < 0.05 ; Figure 2C). In an independent Korean cohort, SNUBH, as a replication dataset, 262 out of 265 genes were expressed and three DEGs (*SNX29*, *FBXO33*, and *C19orf12*) were replicated as significantly expressed in AD compared to CN (Figure 2D). One gene (*FAM169A*) was significantly expressed in AD compared to CN only in the SNUBH cohort. The meta-analysis of results from two cohorts showed that out of the 262 DEGs, 203 genes (167 upregulated and 36 downregulated) were

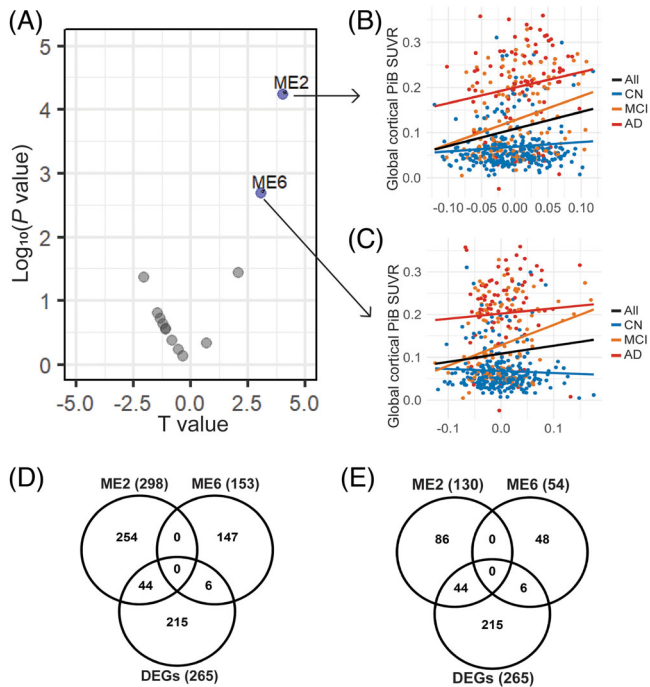


FIGURE 3 Weighted gene co-expression network analysis. Gene network analysis of RNA-Seq data in blood identified 13 network modules, where two modules (ME2 and ME6) were significantly associated with global cortical deposition on PET. A, The volcano plot represents the association of 13 network modules with global cortical deposition. The plots show the correlation between global cortical $A\beta$ deposition and eigengene values of ME2 (B) and ME6 (C). D, Venn diagram shows the overlapping genes between 256 DEGs associated with $A\beta$ deposition, genes in ME2, and genes in ME6. E, Venn diagram shows the overlapping genes between 256 DEGs associated with $A\beta$ deposition, $A\beta$ -associated genes in ME2, and $A\beta$ -associated genes in ME6. $A\beta$, amyloid beta; AD, Alzheimer's disease; All, total participants; CN, cognitively normal older adult; DEG, differentially expressed gene; FDR, false discovery rate; MCI, mild cognitive impairment; ME, module eigenvalue; RNA-Seq, RNA sequencing; PET, positron emission tomography; PiB, Pittsburgh compound B; SUVR, standardized uptake value ratio.

significantly dysregulated in AD compared to CN (Table S2 in supporting information).

3.4 | Weighted gene co-expression network analysis for global cortical $A\beta$ deposition

Network analysis identified 13 network modules, and the number of genes included in each module ranged from 32 to 588 (Table S3 in supporting information). Among the 13 network modules, two modules (ME2 and ME6) were significantly associated with global cortical $A\beta$ deposition (Figure 3). In Figure 3B and Figure 3C, the plot represents the association between the two modules (ME2 and ME6) and $A\beta$ deposition in three diagnostic groups. We investigated the overlap between genes within two significant modules and DEGs identified as being associated with global cortical $A\beta$ deposition (Figure 3D). The overlap-

ping genes include 44 genes in ME2, which consists of 298 genes, and 6 genes in ME6, which contains 153 genes. Association analysis was conducted to determine which genes within the two significant modules were associated with $A\beta$ deposition. After FDR correction, 130 genes in ME2 and 54 genes in ME6 were found to be significantly associated with $A\beta$ deposition. The number of genes overlapping with DEGs remained unchanged (Figure 3E). We explored potential AD therapeutic target genes and compared them to genes in two significant modules (Figure S1). This comparison revealed an overlap of 16 genes in ME2 and two genes in ME6 with potential AD target genes.

3.5 | Functional analysis between genes in network modules and their top 10 biological pathways

To explore the biological function, we analyzed the GO biological pathway on genes in two network modules (ME2 and ME6). Details on the representative GO biological processes and their cluster information are provided in Table S4 in supporting information, while the top 10 pathways, simplified and ordered by BH-adjusted P value, are illustrated in Figure S2 in supporting information. Pathway analysis with 298 genes in ME2 identified 33 pathways after similarity assessment, and nine clusters were generated (Table S4). The top 10 GO biological pathways were highlighted into five clusters: cell adhesion, natural killer (NK) cell-mediated immunity, cell killing, leukocyte mediated immunity, and response to lectin (Figure 4A and Table S4). The network revealed two primary clusters, representing cell adhesion and immune response, both linked by the *CADM1* gene (Figure 4A). Specifically, NK cell-mediated immunity pathway included 16 genes from ME2 (Table S4), 5 of which were DEGs (*GZMB*, *KLRD1*, *KLRC4*, *NCR1*, *KLRC4-KLRK1*) associated with $A\beta$ deposition, while 2 genes (*LAG3* and *CADM1*) had been previously nominated in Agora.³² After the similarity assessment, 20 ME6-related pathways were identified, which led to the formation of five clusters (Table S4). The top 10 GO biological pathways were grouped into five clusters, highlighting cytoplasmic translation, ribosome biogenesis, protein localization, nucleosome assembly, and rRNA processing (Figure 4B and Table S4). The most significant pathway was cytoplasmic translation, which included 42 genes from ME6. Notably, *RPS27* showed a significant association with $A\beta$ deposition (Table S1), while two genes, *S100A8* and *MRPL12*, had previously been nominated in Agora³² (Figure S1).

3.6 | Cell-type deconvolution analysis of RNA-Seq data

Given that immune related pathways were associated with $A\beta$ deposition in both co-expression network analysis and GO biological pathway analysis, we conducted a deconvolution analysis of gene expression using CIBERSORTx⁴⁴ to estimate the relative cell-type proportions (Figure 5). In Figure 5A, deconvolution analysis showed neutrophils as the most prevalent cell type in blood, followed by resting NK cells, resting CD4 memory T cells, and CD8 T cells. Considering the

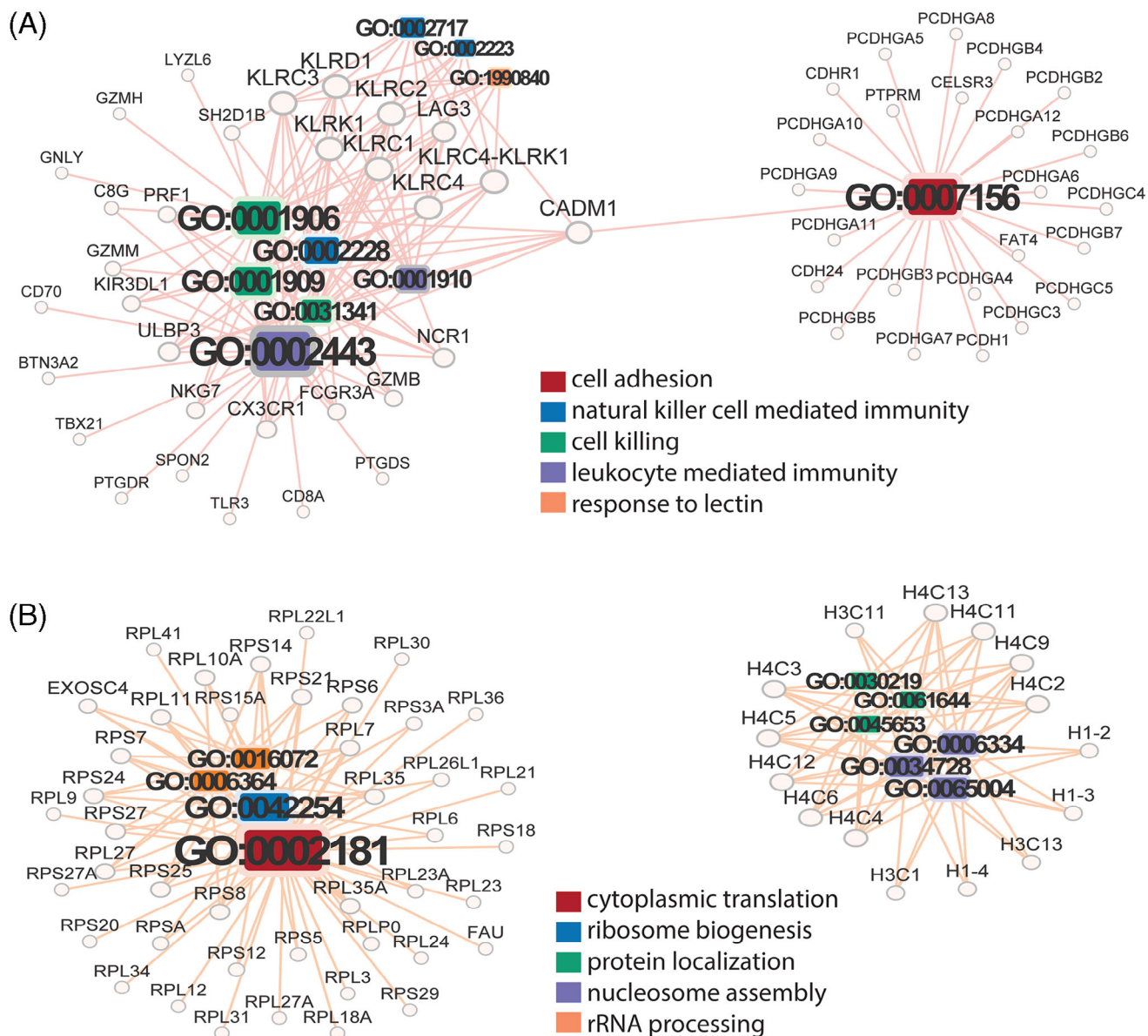


FIGURE 4 Functional enrichment analysis. The network included genes within ME2 (A) and ME6 (B), along with their top 10 significant biological pathways, which were ranked by BH-adjusted *P* value. The colors indicated the clusters of GO biological pathways and genes corresponding to their pathways. BH, Benjamini–Hochberg correction method; GO, Gene Ontology; ME, module eigenvalue.

inclusion of various T cell subtypes in our results, the overall proportion of T cells should be the second most abundant in blood. We subsequently assessed the association between the well-estimated immune cell proportions and global cortical A β deposition (Figure 5B). To focus on prevalent immune cells, only those with mean proportions > 0.01 were selected for association analysis using a linear regression model. Resting NK cells were the only cell type significantly associated with global cortical A β deposition. Figure 5C illustrates the relationship between the proportion of resting NK cells in blood and global cortical A β deposition across three diagnostic groups, with AD showing the strongest correlation between resting NK cells and A β deposition.

3.7 | ML-based classification of A β positivity

We evaluated whether incorporating gene expression data improved the classification performance of A β positivity using the STREAMLINE.^{45,46} We compared the classification performance of 10 ML algorithms across three models and found that elastic net showed better classification performance based on balanced accuracy, F1 score, sensitivity, and ROC AUC (Table S6 in supporting information). ROC curves depicting the performance of an elastic net model in classifying A β positivity are presented in Figure 6.

Figure 6A shows the classification performance of the Model 1 including only age and sex as features. The model yielded a mean

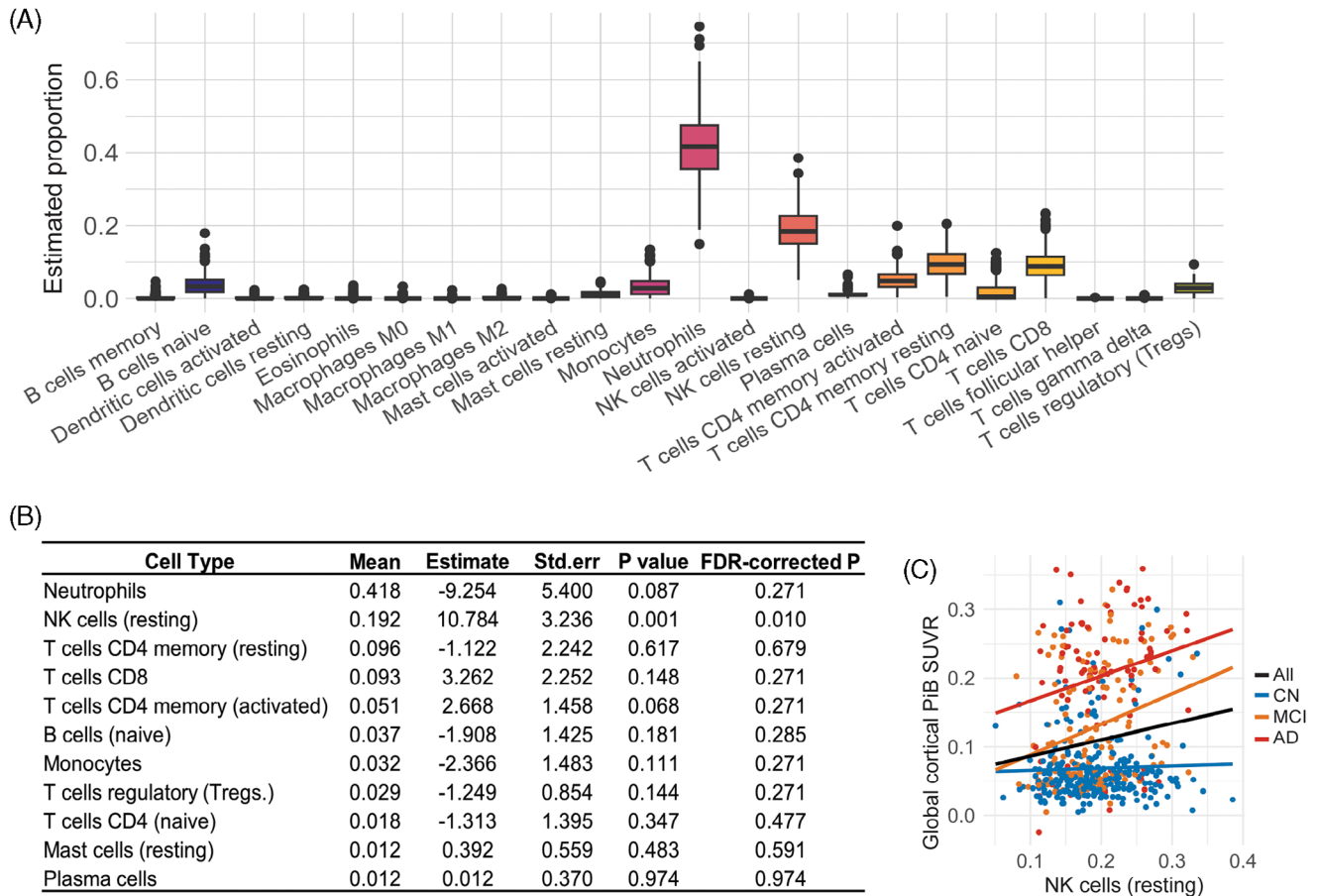


FIGURE 5 Deconvolution analysis. Immune cell type proportions were predicted by CIBERSORTx. A, Boxplots show estimated cell type proportions for 506 individuals in each of the 22 immune cell types. B, The table represents the mean value of proportions and the results of association analysis between immune cell type proportions and global cortical A β deposition. C, NK cells resting was significantly associated with global cortical A β deposition. A β , amyloid beta; AD, Alzheimer's disease; All, total participants; CN, cognitively normal older adults; FDR, false discovery rate; MCI, mild cognitive impairment; NK, natural killer; PIB, Pittsburgh compound B; Std.err, standard error; SUVR, standardized uptake value ratio.

AUC of 0.631 ± 0.056 . The inclusion of APOE $\epsilon 4$ carrier status into Model 1 improved the classification performance with a mean AUC of 0.774 ± 0.053 (Figure 6B and Figure 6D). Incorporating the top 30 genes (Table S5 in supporting information) ranked by feature importance into Model 2 resulted in the best AUC curve performance among the three models with a mean AUC of 0.807 ± 0.022 (Figure 6C and Figure 6D).

4 | DISCUSSION

We performed total RNA sequencing on blood samples from the KBASE cohort recruited in Korea²³ and analyzed blood-based transcriptomic data to identify dysregulated genes and their biological pathways for global cortical A β deposition measured from PET scans. Amyloid PET A β deposition⁴⁷ is a widely recognized biomarker for AD.⁴⁻⁶

We identified 256 dysregulated genes as significantly associated with A β deposition. Furthermore, our findings were replicated in

an independent Korean cohort, SNUBH. Pathway analysis of the upregulated genes revealed the enrichment of the NK cell-mediated immunity pathway. Some of these upregulated genes including GZMB, KLRB1, KLRD1, NCR1, and NKG7, are recognized as markers for NK cells,^{48,49} while others including CD96, KLRC3, KLRC4, and KLRK1 are known to be expressed in NK cells.^{50,51} SLAMF7 expression on NK cells was elevated in the bone marrow of multiple myeloma patients, and higher SLAMF7 expression on blood NK cells was associated with shorter progression-free survival.⁵² Additionally, 44 genes showed differential expressions between AD and CN individuals in the KBASE cohort. Among the 44 genes analyzed, CYP2R1 (cytochrome P450 family 2 subfamily R member 1) showed the most significant association with both A β and AD. CYP2R1 enzymes, which exhibit 25-hydroxylase activity, initiate the formation of active vitamin D by producing 25-hydroxyvitamin D (25[OH]D), the main circulating form of vitamin D in the blood and an essential indicator for evaluating vitamin D status.^{53,54} Vitamin D plays a role in the clearance of A β aggregates, and its deficiency is associated with cognitive impairment.^{54,55} In addition, the Vitamin D receptor interacts with SMAD3, a transcription

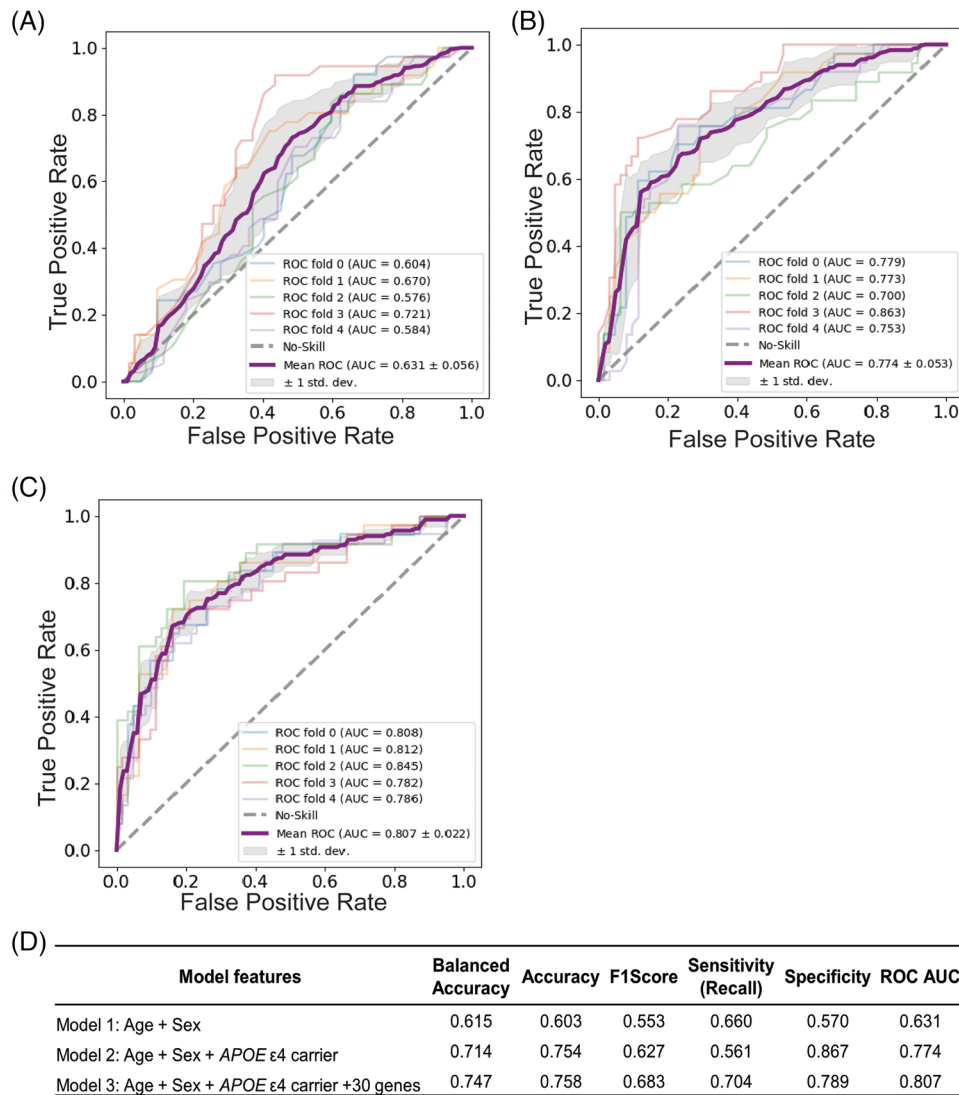


FIGURE 6 Machine learning-based classification analysis of A β positivity. The ROC and AUC for the classification performance of A β positivity from elastic net were evaluated for Model 1 (A), Model 2 (B), and Model 3 (C). A, Model 1 included age and sex as features. B, Model 2 included age, sex, and APOE ϵ 4 carrier status. C, Model 3 included age, sex, APOE ϵ 4 carrier status, and top 30 genes selected based on feature importance scores. The purple line represented the mean ROC curve obtained from 5-fold cross-validation, while the background lines displayed the ROC curve across five folds. The gray area represents the standard deviation of the mean ROC curve. D, The table shows average elastic net performance metrics across 5-fold cross-validation. A β , amyloid beta; APOE, apolipoprotein E; AUC, area under the curve; ROC, receiver operating characteristic.

factor involved in amyloid precursor protein (APP) processing through transforming growth factor beta (TGF- β) signaling.⁵⁶

Three dysregulated genes in AD (*SNX29*, *FBXO33*, and *C19orf12*) were replicated in an independent Korean cohort, SNUBH. The expression of sorting nexin 29 (*SNX29*) was downregulated in AD. The sorting nexins (SNXs) were broadly investigated in AD and other nervous system pathology.⁵⁷ *SNX9*, *SNX17*, and *SNX33* have been reported to play roles in regulation of APP trafficking.⁵⁷ In a study of the Han Chinese population, *SNX29* was found to share genetic risk factors across major psychiatric disorders, including schizophrenia, bipolar disorder, and major depressive disorder and was also found to be involved in TGF- β signaling,⁵⁸ indicating a potential role for *SNX29* in AD, particularly in relation to A β deposition. There is currently no

well-established connection between *FBXO33* (F-box only protein 33) and A β or AD. However, *FBXO33* is predominantly associated with the modulation of protein ubiquitination and solubility, particularly within the context of polyglutamine disorders such as spinocerebellar ataxia type 3 (SCA3).⁵⁹ SCA3 is a progressive neurodegenerative disorder characterized by ataxia, with amyloid aggregation of the ataxin-3 protein being a hallmark of the disease.^{57,59} Given that both pathologies involve protein aggregation leading to neurodegeneration, understanding the mechanism of ataxin-3 aggregation in SCA3 could offer valuable insights into the process of A β deposition in AD. While the function of *C19orf12* (chromosome 19 open reading frame 12), located in the mitochondria, remains unclear, mutations in this gene have been linked to neurodegeneration with brain iron

accumulation (NBA), a group of inherited neurological disorders, as well as widespread α -synucleinopathy.^{60,61}

The correlation-based network analysis showed more explanatory power than DEG analysis. Two network modules, ME2 (immune response-related genes) and ME6 (cytoplasmic translation-related genes), showed significant associations with $A\beta$ deposition. Notably, NK cell-mediated immunity was identified in both the DEG analysis and the co-expression network analysis. In the ME2-related network, *CADM1* (cell adhesion molecule 1), a marker gene for disease-associated microglia, was enriched in both immune response and cell adhesion-related pathways. One study using spatial sequencing analysis revealed a specific pattern in which *CADM1* expression in cortical microglia increased within $\approx 20 \mu\text{m}$ of $A\beta$ deposits, while expression sharply decreased near the deposits.⁶² This pattern was not observed in hippocampal microglia.

We estimated immune cell types in blood bulk RNA-Seq data through deconvolution analysis to assess alterations in circulating immune cell populations. The analysis confirmed a dominant proportion of neutrophils and resting NK cells in the blood. Recent studies have found that neutrophils in the brains of AD were observed near or at $A\beta$ plaques, as well as in the peripheral blood.^{63–66} In a study examining myeloperoxidase, which is expressed at high levels in neutrophils, researchers observed a significant increase in neutrophils in the brains of AD patients, along with notable neutrophil accumulation around amyloid plaques.⁶³ Recent studies have demonstrated that changes in neutrophil levels can serve as an indicator of disease progression and cognitive decline.⁶⁷ Additionally, a prospective cohort study of 101,582 individuals found that low counts of blood monocytes and eosinophils, combined with high counts of blood leukocytes and neutrophils, were associated with an increased risk of AD.⁶⁵ The neutrophil-to-lymphocyte ratio (NLR) is an emerging biomarker for impaired immunity calculated as a simple ratio between the neutrophil and lymphocyte counts measured in peripheral blood.⁶⁶ A study on myeloproliferative neoplasms (MPNs) using blood cell counts showed elevated NLR, indicating that neutrophils circulate in an active state, leading to increased inflammatory signaling pathway activation.⁶⁶ Notably, the proportion of resting NK cells was significantly associated with $A\beta$ deposition. Circulating NK cells generally remain in a resting state, but their activation by cytokines prompts infiltration into pathogen-infected tissues or malignant cells.^{68,69} NK cells are implicated in a range of central nervous system disorders, including AD, and their infiltration into the brain is frequently observed in neurodegenerative diseases.^{70,71} A study on NK cell infiltration into the brain in Parkinson's disease revealed that NK cells contribute to the clearance of α -synuclein aggregates in a mouse model.⁷² The observed reduction in $A\beta$ plaques in *Rag2*^{-/-} mice crossed with APP/PS1 mice, which lack T and B cells but have sufficient NK cells, suggests that NK cells may play a crucial role in the clearance of $A\beta$ plaques, potentially surpassing the contributions of T and B cells.⁷³ While NK cells have been demonstrated to infiltrate the brains of APP/PS1 mice, contributing to $A\beta$ pathology, there is a paucity of research on their infiltration into the AD brain.⁷⁴ The mechanisms underlying NK cell activation and brain infiltration remain unclear. Specifically, whether peripheral $A\beta$ stimula-

tion is required to activate circulating NK cells and whether $A\beta$ serves as a chemoattractant for these cells needs further investigation.⁷⁴ The finding that NK cells are significantly associated with $A\beta$ deposition suggests that they might influence or reflect the progression of amyloid pathology, potentially acting as early responders in the peripheral immune response.

Elastic net was selected as the optimal ML algorithm for classification of $A\beta$ positivity. Our results suggest that including gene expression data makes the classification model more effective and provides valuable insights into the classification power of gene expression data.

Peripheral blood samples do not directly describe the physiological environment as much as brain tissue in AD. Although brain tissue can only be analyzed *post mortem*, blood samples are more accessible for collection and enable high-throughput testing and longitudinal studies, which are vital for diagnosing and monitoring the progression of diseases.^{14–17} Additionally, multiple studies have reported that the immune response in peripheral blood can elucidate the pathological characteristics of AD.^{7–9,12,13} Our results suggest that by establishing an association between brain imaging data and blood transcriptome profile data, significant gene expression changes could potentially act as diagnostic blood-based molecular biomarkers.

The study investigated the association between gene expression levels and global cortical $A\beta$ deposition in the AD continuum using blood-based RNA-Seq data from a Korean cohort (KBASE), replicated in an independent Korean cohort (SNUBH). Dysregulated genes related to immune pathways including NK cell-mediated immunity exhibit a strong association with $A\beta$ deposition. Future work will involve longitudinal transcriptome analysis in the KBASE cohort to comprehensively understand the effects of pathophysiological characteristics over time in AD. Our findings highlight that blood-based transcriptomics data may provide a better understanding of AD molecular mechanisms and yield potential diagnostic and therapeutic strategies.

ACKNOWLEDGMENTS

The authors extend our gratitude to (1) the KBASE participants for their generous contribution of time to this study, (2) the committed clinical coordinators for their pivotal role in data collection and study facilitation, and (3) the cross-continental team responsible for data generation, storage, and project management for KBASE. Kwangsik Nho, PhD, receives support from NIH grants (R01 LM012535, U01 AG072177, and U19 AG074879). Andrew J Saykin, PsyD, receives support from multiple NIH grants (P30 AG010133, P30 AG072976, R01 AG019771, R01 AG057739, U19 AG024904, R01 LM013463, R01 AG068193, T32 AG071444, U01 AG068057, U01 AG072177, and U19 AG074879). Dong Young Lee, MD, PhD, receives support from Korean government grants (NRF-2014M3C7A1046042, HI18C0630, and HI19C0149) and a NIA grant (U01 AG072177). Min Soo Byun, MD, PhD, receives support from New Faculty Startup Fund from Seoul National University, a Korean government grant (RS-2022-00165636 and RS-2023-KH136195), and a NIA grant (U01 AG072177). Young Ho Park, MD, PhD, receives support from the National Research

Foundation of Korea grant funded by the Korean government (Ministry of Science and ICT; No. 2020R1C1C1013718).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the [Supporting information](#).

DATA AVAILABILITY STATEMENT

The datasets used and analyzed in this study will be provided by the corresponding author upon approval request.

CONSENT STATEMENT

All participants provided informed consent.

REFERENCES

- DeTure MA, Dickson DW. The neuropathological diagnosis of Alzheimer's disease. *Molecular Neurodegeneration*. 2019;14. doi:10.1186/s13024-019-0333-5
- Breijyeh Z, Karaman R. Comprehensive review on Alzheimer's disease: causes and treatment. *Molecules*. 2020;25. doi:10.3390/molecules25245789
- 2024 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2024;20:3708-3821. doi:10.1002/alz.13809
- Thal DR, Rüb U, Orantes M, Braak H. Phases of A β -deposition in the human brain and its relevance for the development of AD. *Neurology*. 2002;58:1791-1800. doi:10.1212/WNL.58.12.1791
- Ruan D, Sun L. Amyloid- β PET in Alzheimer's disease: a systematic review and Bayesian meta-analysis. *Brain and Behavior*. 2023;13. doi:10.1002/brb3.2850
- Chapleau M, Iaccarino L, Soleimani-Meigooni D, Rabinovici GD. The Role of amyloid PET in imaging neurodegenerative disorders: a review. *J Nucl Med*. 2022;63. doi:10.2967/jnumed.121.263195
- Walker KA, Ficek BN, Westbrook R. Understanding the role of systemic inflammation in Alzheimer's disease. *ACS Chem Neurosci*. 2019. doi:10.1021/acchemneuro.9b00333
- Shi M, Chu F, Zhu F, Zhu J. Peripheral blood amyloid- β involved in the pathogenesis of Alzheimer's disease via impacting on peripheral innate immune cells. *J Neuroinflamm*. 2024;21. doi:10.1186/s12974-023-03003-5
- Shi M, Chu F, Tian X, et al. Role of adaptive immune and impacts of risk factors on adaptive immune in Alzheimer's disease: are immunotherapies effective or off-target?. *Neuroscientist*. 2021. doi:10.1177/1073858420987224.-02-03
- Spangenberg EE, Green KN. Inflammation in Alzheimer's Disease: lessons learned from microglia-depletion models. *Brain Behav Immun*. 2017;61. doi:10.1016/j.bbi.2016.07.003
- Solana C, Tarazona R, Solana R. Immunosenescence of natural killer cells, inflammation, and Alzheimer's disease. *International Journal of Alzheimer's Disease*. 2018;2018. doi:10.1155/2018/3128758
- Bettcher BM, Tansey MG, Dorothée G, Heneka MT. Peripheral and central immune system crosstalk in Alzheimer disease — a research prospectus. *Nat Rev Neurol*. 2021;17. doi:10.1038/s41582-021-00549-x
- Hou J-H, Ou Y-N, Xu W, et al. Association of peripheral immunity with cognition, neuroimaging, and Alzheimer's pathology. *Alzheimer's Research & Therapy*. 2022;14. doi:10.1186/s13195-022-00968-y
- Thambisetty M, Lovestone S. Blood-based biomarkers of Alzheimer's disease: challenging but feasible. *Biomark Med*. 2010;4. doi:10.2217/bmm.09.84
- Teunissen CE, Verberk IMW, Thijssen EH, et al. Blood-based biomarkers for Alzheimer's disease: towards clinical implementation. *The Lancet Neurology*. 2022;21. doi:10.1016/S1474-4422(21)00361-6
- Hansson O, Edelmayer RM, Boxer AL, et al. The Alzheimer's Association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimer's & Dementia*. 2022;18. doi:10.1002/alz.12756
- Hampel H, O'Bryant SE, Molinuevo JL, et al. Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. *Nature Reviews Neurology* 2018 14:11. 2018;14. doi:10.1038/s41582-018-0079-7
- 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. doi:10.1126/sciadv.abb5398.-01
- Wan Y-W, Al-Ouran R, Mangleburg CG, et al. Meta-Analysis of the Alzheimer's disease human brain transcriptome and functional dissection in mouse models. *Cell Rep*. 2020;32. doi:10.1016/j.celrep.2020.107908
- Twine NA, Janitz K, Wilkins MR, Janitz M. Whole transcriptome sequencing reveals gene expression and splicing differences in brain regions affected by Alzheimer's Disease. *PLoS One*. 2011;6. doi:10.1371/journal.pone.0016266
- Zhong H, Zhou X, Uhm H, et al. Using blood transcriptome analysis for Alzheimer's disease diagnosis and patient stratification. *Alzheimer's & Dementia*. 2024;20. doi:10.1002/alz.13691
- Song L, Yang YT, Guo Q, Consortium TZ, Zhao X-M. Cellular transcriptional alterations of peripheral blood in Alzheimer's disease. *BMC Medicine*. 2022;20. doi:10.1186/s12916-022-02472-4
- Byun MS, Yi D, Lee JH, et al. Korean brain aging study for the early diagnosis and prediction of alzheimer's disease: methodology and baseline sample characteristics. *Psychiatry Investigation*. 2017;14. doi:10.4306/pi.2017.14.6.851
- Wenham P, Price W, Blundell G. Apolipoprotein E genotyping by one-stage PCR. *Lancet North Am Ed*. 1991;337. doi:10.1016/0140-6736(91)92823-K
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30. doi:10.1093/bioinformatics/btu170
- Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29. doi:10.1093/bioinformatics/bts635
- Frankish A, Diekhans M, Ferreira A-M, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res*. 2019;47. doi:10.1093/nar/gky955
- Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30. doi:10.1093/bioinformatics/btt656
- Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics*. 2012;13. doi:10.1093/biostatistics/kxr054
- Klunk WE, Koeppe RA, Price JC, et al. The centiloid project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimer's & dementia : the journal of the Alzheimer's Association*. 2015;11. doi:10.1016/j.jalz.2014.07.003
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 1995;57. doi:10.1111/j.2517-6161.1995.tb02031.x
- Greenwood AK, Gockley J, Daily K, et al. Agora: an open platform for exploration of Alzheimer's disease evidence. *Alzheimer's & Dementia*. 2020;16. doi:10.1002/alz.046129
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 2019;37. doi:10.1038/s41587-019-0201-4

34. Kovaka S, Zimin AV, Perteu GM, Razaghi R, Salzberg SL, Perteu M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* 2019;20. doi:10.1186/s13059-019-1910-1
35. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26. doi:10.1093/bioinformatics/btq340
36. Langfelder P, Horvath S, Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008 9:1. 2008;9. doi:10.1186/1471-2105-9-559
37. Boyle EI, Weng S, Gollub J, et al. GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics.* 2004;20. doi:10.1093/bioinformatics/bth456
38. Consortium TheGene. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* 2019;47. doi:10.1093/nar/gky1055
39. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *The Innovation.* 2021;2. doi:10.1016/j.xinn.2021.100141
40. Bodenhofer U, Kothmeier A, Hochreiter S. APCluster: an R package for affinity propagation clustering. *Bioinformatics.* 2011;27. doi:10.1093/bioinformatics/btr406
41. Gu Z, Hübschmann D. SimplifyEnrichment: a bioconductor package for clustering and visualizing functional enrichment results. *Genomics Proteomics Bioinformatics.* 2023;21. doi:10.1016/j.gpb.2022.04.008
42. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13. doi:10.1101/gr.1239303
43. Momeni K, Ghorbian S, Ahmadvan E, et al. Unraveling the complexity: understanding the deconvolutions of RNA-seq data. *Translational Medicine Communications* 2023 8:1. 2023;8. doi:10.1186/s41231-023-00154-8
44. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology* 2019 37:7. 2019;37. doi:10.1038/s41587-019-0114-2
45. Urbanowicz RJ, Bandhey H, Keenan BT, et al. STREAMLINE: an automated machine learning pipeline for biomedicine applied to examine the utility of photography-based phenotypes for OSA prediction across international sleep centers. *arXiv preprint arXiv:231205461.* 2023. doi:10.48550/arXiv.2312.05461
46. Urbanowicz R, Zhang R, Cui Y, Suri P. STREAMLINE: A simple, transparent, end-to-end automated machine learning pipeline facilitating data analysis and algorithm comparison. *Genetic and Evolutionary Computation.* 2023. doi:10.1007/978-981-19-8460-0_9
47. Klunk WE, Engler H, Nordberg A, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol.* 2004;55. doi:10.1002/ana.20009
48. Qi C, Liu Q. Natural killer cells in aging and age-related diseases. *Neurobiol Dis.* 2023;183:106156. doi:10.1016/j.nbd.2023.106156
49. Qi C, Liu F, Zhang W, et al. Alzheimer's disease alters the transcriptomic profile of natural killer cells at single-cell resolution. *Front Immunol.* 2022;13. doi:10.3389/fimmu.2022.1004885
50. Wei L, Xiang Z, Zou Y, Wei L, Xiang Z, Zou Y. The Role of NKG2D and its ligands in autoimmune diseases: new targets for immunotherapy. *International Journal of Molecular Sciences* 2023, Vol 24, Page 17545. 2023;24. doi:10.3390/ijms242417545
51. Hikami K, Tsuchiya N, Yabe T, et al. Variations of human killer cell lectin-like receptors: common occurrence of NKG2-C deletion in the general population. *Genes & Immunity* 2003 4:2. 2003;4. doi:10.1038/sj.gene.6363940
52. Pazina T, MacFarlane AW, Bernabei L, et al. Alterations of NK cell phenotype in the disease course of multiple myeloma. *Cancers* 2021, Vol 13, Page 226. 2021;13. doi:10.3390/cancers13020226
53. Bivona G, Sasso BL, Gambino CM, et al. The role of vitamin D as a biomarker in Alzheimer's disease. *Brain Sciences.* 2021;11. doi:10.3390/brainsci11030334
54. Bivona G, Lo Sasso B, Iacolino G, et al. Standardized measurement of circulating vitamin D [25(OH)D] and its putative role as a serum biomarker in Alzheimer's disease and Parkinson's disease. *Clin Chim Acta.* 2019;497:82-87. doi:10.1016/j.cca.2019.07.022
55. Ghahremani M, Smith EE, Chen H-Y, Creese B, Goodarzi Z, Ismail Z. Vitamin D supplementation and incident dementia: effects of sex, APOE, and baseline cognitive status. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring.* 2023;15. doi:10.1002/dad2.12404
56. Yanagisawa J, Yanagi Y, Masuhiro Y, et al. Convergence of transforming growth factor- β and vitamin D signaling pathways on SMAD Transcriptional Coactivators. *Science.* 1999;283. doi:10.1126/science.283.5406.1317
57. Vieira N, Rito T, Correia-Neves M, et al. Sorting out sorting nexins functions in the nervous system in health and disease. *Molecular Neurobiology* 2021 58:8. 2021;58. doi:10.1007/s12035-021-02388-9
58. Chen J-H, Zhao Y, Khan RAW, et al. SNX29, a new susceptibility gene shared with major mental disorders in Han Chinese population. *The World Journal of Biological Psychiatry.* 2021. doi:10.1080/15622975.2020.1845793
59. Chen ZS, Wong AKY, Cheng TC, Koon AC, Chan HYE. FipoQ/FBXO33, a Cullin-1-based ubiquitin ligase complex component modulates ubiquitination and solubility of polyglutamine disease protein. *J Neurochem.* 2019;149. doi:10.1111/jnc.14669
60. Nguyen V, Garcia D, Setthavongsack N, et al. Secondary tauopathy in a genetic synucleinopathy, mitochondrial protein-associated neurodegeneration (MPAN). *Alzheimer's & Dementia.* 2020;16. doi:10.1002/alz.046690
61. Angelini C, Durand CM, Fergelot P, et al. Autosomal dominant MPAN: mosaicism expands the clinical spectrum to atypical late-onset phenotypes. *Mov Disord.* 2023;38. doi:10.1002/mds.29576
62. Scoyni F, Giudice L, Väänänen M-A, et al. Alzheimer's disease-induced phagocytic microglia express a specific profile of coding and non-coding RNAs. *Alzheimer's & Dementia.* 2024;20:954-974. doi:10.1002/alz.13502
63. Smyth LCD, Murray HC, Hill M, et al. Neutrophil-vascular interactions drive myeloperoxidase accumulation in the brain in Alzheimer's disease. *Acta Neuropathologica Communications* 2022 10:1. 2022;10. doi:10.1186/s40478-022-01347-2
64. Aries ML, Hensley-McBain T. Neutrophils as a potential therapeutic target in Alzheimer's disease. *Front Immunol.* 2023;14. doi:10.3389/fimmu.2023.1123149
65. Luo J, Thomassen JQ, Nordestgaard BG, Tybjaerg-Hansen A, Frikke-Schmidt R. Blood Leukocyte Counts in Alzheimer Disease. *JAMA Netw Open.* 2022. doi:10.1001/jamanetworkopen.2022.35648
66. Larsen MK, Skov V, Kjær L, et al. Neutrophil-to-lymphocyte ratio and all-cause mortality with and without myeloproliferative neoplasms—a Danish longitudinal study. *Blood Cancer Journal* 2024 14:1. 2024;14. doi:10.1038/s41408-024-00994-z
67. Dong Y, Lagarde J, Xicota L, et al. Neutrophil hyperactivation correlates with Alzheimer's disease progression. *Ann Neurol.* 2018;83. doi:10.1002/ana.25159
68. Mandal A, Viswanathan C. Natural killer cells: in health and disease. *Hematology/Oncology and Stem Cell Therapy.* 2015;8:47-55. doi:10.1016/j.hemonc.2014.11.006
69. Glas R, Franksson L, Une C, et al. Recruitment and activation of natural killer (Nk) cells in vivo determined by the target cell phenotype: an adaptive component of Nk Cell-Mediated Responses. *J Exp Med.* 2000;191. doi:10.1084/jem.191.1.129
70. Poli A, Kmiecik J, Domingues O, et al. NK cells in central nervous system disorders. *J Immunol.* 2013;190. doi:10.4049/jimmunol.1203401
71. Lu Y, Li K, Hu Y, Wang X. Expression of immune related genes and possible regulatory mechanisms in Alzheimer's disease. *Front Immunol.* 2021;12. doi:10.3389/fimmu.2021.768966

72. Guan Q, Liu W, Mu K, et al. Single-cell RNA sequencing of CSF reveals neuroprotective RAC1+ NK cells in Parkinson's disease. *Front Immunol.* 2022;13. doi:[10.3389/fimmu.2022.992505](https://doi.org/10.3389/fimmu.2022.992505)
73. Späni C, Suter T, Derungs R, et al. Reduced β -amyloid pathology in an APP transgenic mouse model of Alzheimer's disease lacking functional B and T cells. *Acta Neuropathologica Communications.* 2015;3. doi:[10.1186/s40478-015-0251-x](https://doi.org/10.1186/s40478-015-0251-x)
74. Shi M, Chu F, Zhu F, Zhu J. Peripheral blood amyloid- β involved in the pathogenesis of Alzheimer's disease via impacting on peripheral innate immune cells. *J Neuroinflam.* 2024;21. doi:[10.1186/s12974-023-03003-5](https://doi.org/10.1186/s12974-023-03003-5)

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: Park T, Hwang J, Liu S, et al. Genome-wide transcriptome analysis of A β deposition on PET in a Korean cohort. *Alzheimer's Dement.* 2024;20:8787–8801. <https://doi.org/10.1002/alz.14348>