



OPEN

Deconvolution reveals cell-type-specific transcriptomic changes in the aging mouse brain

Yingxue Ren^{1✉}, Xue Wang¹, Shuwen Zhang², Hongru Hu³, Zachary Quicksall¹, Sangderk Lee⁴, Josh M. Morganti^{4,5}, Lance A. Johnson^{4,6}, Yan W. Asmann¹ & Na Zhao^{7✉}

Mounting evidence highlights the crucial role of aging in the pathogenesis of Alzheimer's disease (AD). We have previously explored human apoE-targeted replacement mice across different ages and identified distinct molecular pathways driven by aging. However, the specific contribution of different brain cell types to the gene modules underlying these pathways remained elusive. To bridge this knowledge gap, we employed a computational deconvolution approach to examine cell-type-specific gene expression profiles in major brain cell types, including astrocytes (AS), microglia (MG), oligodendroglia (OG), neurons (NEU), and vascular cells (VC). Our findings revealed that immune module genes were predominantly expressed in MG, OG, and VC. The lipid metabolism module genes were primarily expressed in AS, MG, and OG. The mitochondria module genes showed prominent expression in VC, and the synapse module genes were primarily expressed in NEU and VC. Furthermore, we identified intra- and inter-cell-type interactions among these module genes and validated their aging-associated expression changes using published single cell studies. Our study dissected bulk brain transcriptomics data at the cellular level, providing a closer examination of the cell-type contributions to the molecular pathways driven by aging.

Aging and Alzheimer's disease (AD) involve intricate processes including pathways such as inflammation, lipid dysregulation, mitochondrial dysfunction, and synaptic alterations¹⁻⁴. The progression of these conditions is influenced by multiple cell types, forming a complex network of interactions that underpin disease development. Understanding how these cell types interact is pivotal in deciphering the mechanisms driving AD and other aging-related diseases.

Recent advancements in single-cell and single-nuclei RNA sequencing have significantly enhanced our understanding of diseases at the single cell and cell-type levels^{5,6}. However, the high cost associated with these techniques often restricts large-scale experiments necessary for investigating multiple traits of interest with robust statistical power. To address this limitation, deconvolution approaches have emerged, employing computational algorithms to estimate cell-type-specific gene expressions from bulk tissue data, leveraging reference data from single-cell and single-nuclei studies⁷⁻¹⁰. These methods enable the utilization of lower-cost, larger-scale bulk transcriptomics datasets to gain insights into gene interactions across different cell types.

In this study, we applied a cell type deconvolution method to analyze our previously published cerebral cortex transcriptomics data, obtained from male and female apoE-targeted replacement (TR) mice at different ages (3-month-old, 12-month-old, and 24-month-old)¹¹. In our previous work, we had reported strong aging-related effects, including immune responses, at bulk transcriptomic level. Using deconvoluted data, our goal of this study was to explore how various brain cell types collectively contribute to critical molecular pathways influenced by aging by examining the cell-type-specific expression of module genes associated with aging. To validate our findings, we conducted comparisons between our deconvoluted data and three single-cell RNA sequencing datasets derived from mice of different ages¹²⁻¹⁴. Our findings highlight the value of computational deconvolution when applied to bulk transcriptomics data, providing novel insights at the cellular level to enhance our understanding of AD and other aging-related diseases.

¹Department of Quantitative Health Sciences, Mayo Clinic, 4500 San Pablo Road South, Jacksonville, FL 32224, USA. ²Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55905, USA. ³Genome Center, University of California, Davis, CA 95616, USA. ⁴Sanders Brown Center On Aging, University of Kentucky, Lexington KY40536, USA. ⁵Department of Neuroscience, University of Kentucky, Lexington KY40536, USA. ⁶Department of Physiology, University of Kentucky, Lexington KY40536, USA. ⁷Department of Neuroscience, Mayo Clinic, 4500 San Pablo Road South, Jacksonville, FL 32224, USA. ✉email: ren.yingxue@mayo.edu; zhao.na@mayo.edu

Results

Computational deconvolution of bulk transcriptomics data and identification of major cell types

We employed CIBERSORTx⁸ to perform deconvolution on our previously published bulk transcriptomics data generated from the cerebral cortex of apoE-TR mice¹¹ (Fig. 1A, Supplementary Table S1). Our deconvolution process included two primary levels: estimating cell type proportions within each sample and determining transcriptome-wide cell-type-specific expressions within each sample. To ensure the accurate assignment of cell types, we cross-referenced the cell-type-specific expressions of genes with an ensemble of published cell-type-specific data and databases (proteinatlas.org)^{15–17}. We then focused on gene modules that showed robust associations with aging. We assigned the genes within these modules to their respective cell types based on their highest cell-type-specific expressions. We then constructed intra- and inter-cell type gene co-expression networks within these modules. This approach allowed us to pinpoint the key cell types involved in the aging-associated gene modules, all while exploring the interactions among genes both within and between cell types. To validate our findings related to aging, we conducted cross-validation using data from three published mouse brain single-cell RNA-seq studies^{12–14}.

Our analysis identified five major cell types from the bulk transcriptomics data: astrocytes (AS), microglia (MG), oligodendroglia (OG), neurons (NEU), and vascular cells (VC). To validate the deconvolution results, we compared the expression of cell type markers obtained from our analysis with known cell type marker genes from publicly available data⁹ and confirmed the accuracy of cell type assignments (Fig. 1B).

Analysis of cell type proportions in relation to aging, APOE genotype, and sex

In our analysis of the deconvoluted data, the initial step was to compare cell type proportions among different age groups, APOE genotypes, and between sexes for the five distinct brain cell types. In our prior report¹¹, we observed that the transcriptomic differences between 12 and 24 months old were mild, with the most substantial distinction occurring between 3- and 24-month-old brains. Therefore, in this study, our focus was on comparing the two age groups, specifically 3 months and 24 months old. Our analysis revealed that when comparing young and old brains, the only significant difference in cell proportions was within the neuronal population (NEU). Specifically, we identified a significant reduction in the neuronal component at 24 months compared to 3 months, while no significant changes were observed in other cell types in relation to aging (Fig. 2A). This finding is consistent with previous reports demonstrating the loss of neuronal markers, including synaptic proteins, in aged mice compared to young mice¹⁸. In addition, we did not observe any significant alterations in cell type proportions related to APOE genotype (Fig. 2B) or between sexes (Fig. 2C).

Cell type-specific analysis of gene modules related with immune responses

The dysregulation of the immune responses is a defining feature of aging¹⁹ and plays a substantial role in the development of AD²⁰. In our prior study, we identified an immune gene module characterized by upregulation with aging and enriched for genes related to immune responses¹¹. To gain insights into the specific cell types and their associated genes driving this collective immune response, we assigned the immune module genes to different cell types based on their estimated cell-type-specific expressions (Fig. 3). Our analysis revealed that the immune module genes were predominantly expressed in microglia (MG), oligodendroglia (OG), and vascular cells (VC) (Fig. 3A). Pathway analysis showed that MG genes were enriched for innate immune response ($-\log P = 29.4$), positive regulation of response to stimulus ($-\log P = 26.7$), and cell activation ($-\log P = 17.6$) (Fig. 3B). Among

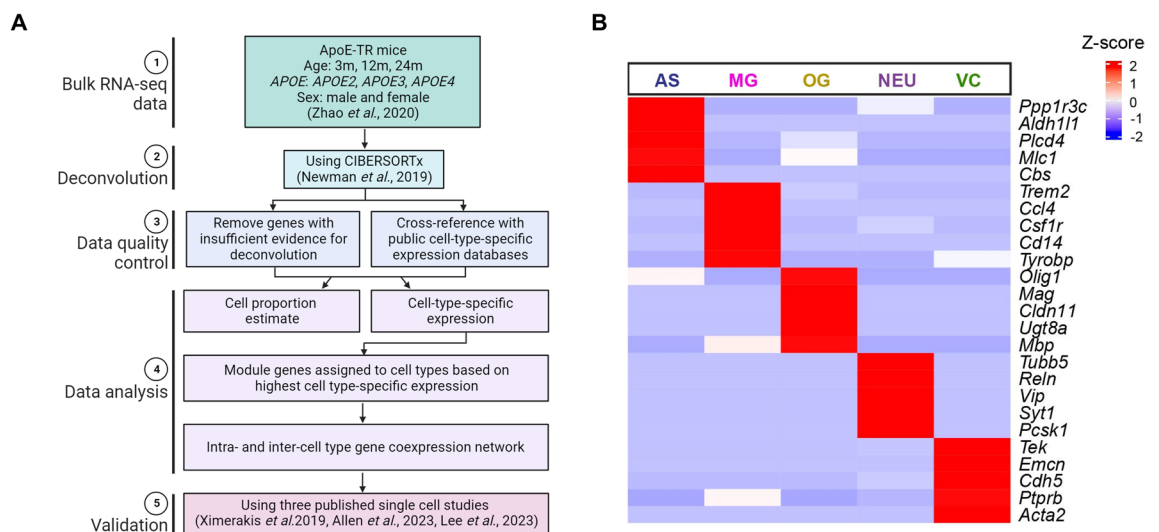


Figure 1. Computational deconvolution of bulk transcriptomics data and identification of major cell types. (A) Analytical workflow. (B) Heatmap showing the expression of known cell type marker genes deconvoluted from our previously published bulk RNA-seq data.

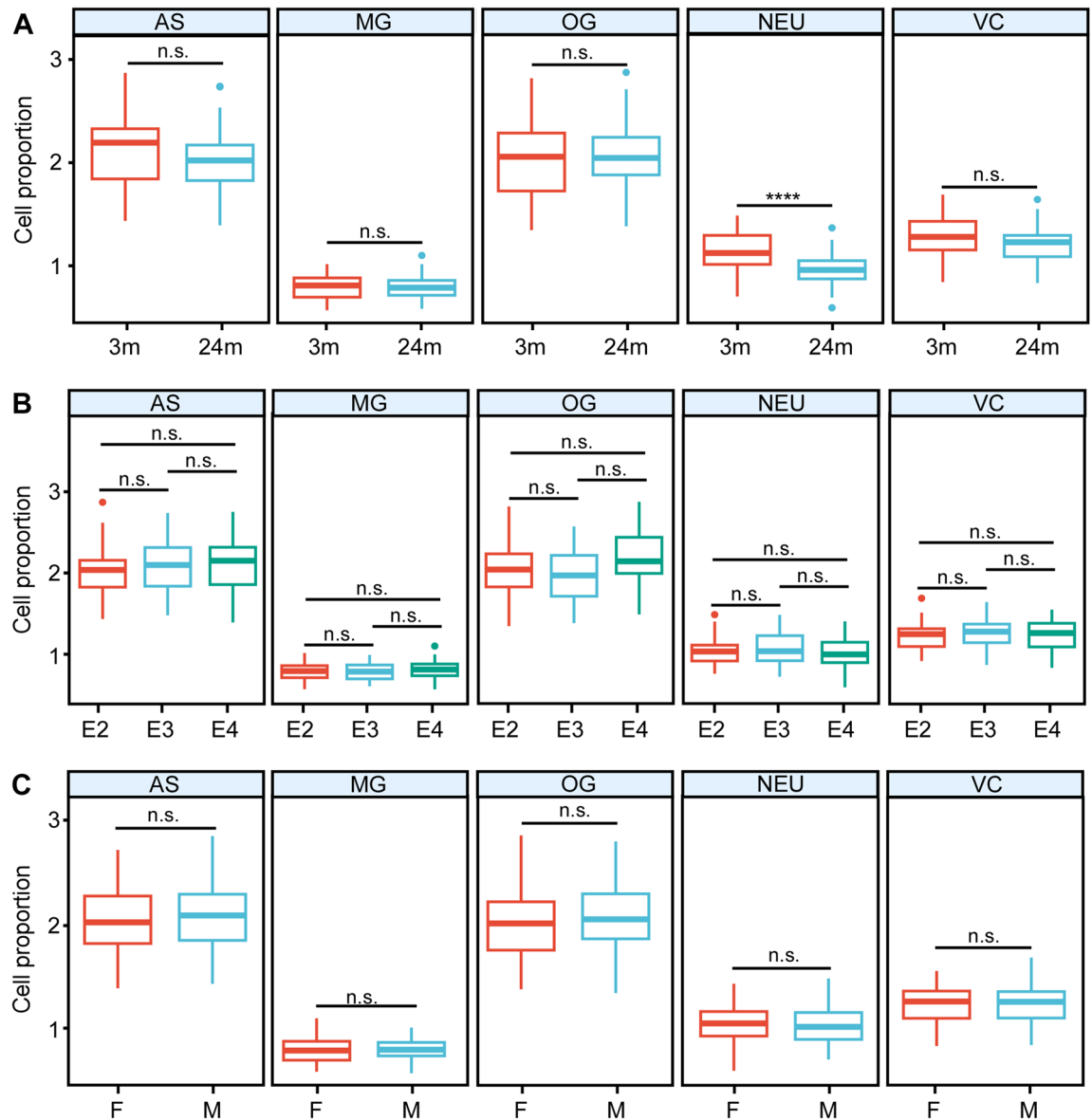


Figure 2. The proportion of the five major brain cell types in each age, *APOE* and sex group. (A) The estimated cell type proportions in 3 months (3 m) and 24 months (24 m) mice. (B) The estimated cell type proportions in *APOE2* (E2), *APOE3* (E3), and *APOE4* (E4) mice. (C) The estimated cell type proportions in female and male mice. T test was used to detect differences between groups. Bonferroni correction was applied to correct for multiple testing. **** $p < 0.0001$; n.s. not significant.

these pathways, well-known MG markers such as *Trem2*, *Tyrobp*, *Cd74*, *Cd14*, and *Csf1r* contributions to these pathways²¹. OG genes also displayed enrichment in immune system processes, sharing genes with MG such as *Stat1*, *Ifit1*, *Ifit3*, and *Appl2*, consistent with the recent findings suggesting that OG is active immunomodulator in response to disease^{22,23}. Additionally, OG genes were implicated in specialized pathways including glial cell development ($-\log P = 7.6$), myelination ($-\log P = 6.2$), and axon ensheathment ($-\log P = 6.1$) (Fig. 3C). Key genes in these pathways included *Smo*, *Tspan2*, *Itgb4*, *Ugt8a*, and *Sh3tc2*. Furthermore, numerous VC genes were co-expressed in MG and exhibited functions related to immune responses, with examples including *Gbp2*, *Gbp3*, *P2rx7*, *Trim30a*, and *Fcer1g*. VC-specific genes, on the other hand, displayed enrichment in processes such as hemostasis ($-\log P = 6$), blood vessel morphogenesis ($-\log P = 4.9$), and angiogenesis ($-\log P = 4$) (Fig. 3D). Notable genes in these pathways encompassed *Tgfa*, *Cmtm6*, and *Fmnl3*. The complete pathway analysis results can be found in Supplementary Table S2.

To gain a deeper understanding of the interactions among these genes, we constructed a gene correlation network encompassing the three cell types. Notably, we observed strong connections between MG genes (*Trem2*, *Stat1*, *Tyrobp*, *Trim30a*), OG genes (*Ugt8a*, *Itgb4*, *Car2*), and a VC gene (*Psmb9*) (Fig. 3E).

To validate the age-related expression changes of these key genes and others, we calculated fold changes between old and young animals in MG, OG, and VC, respectively, using our deconvoluted cell-type-specific

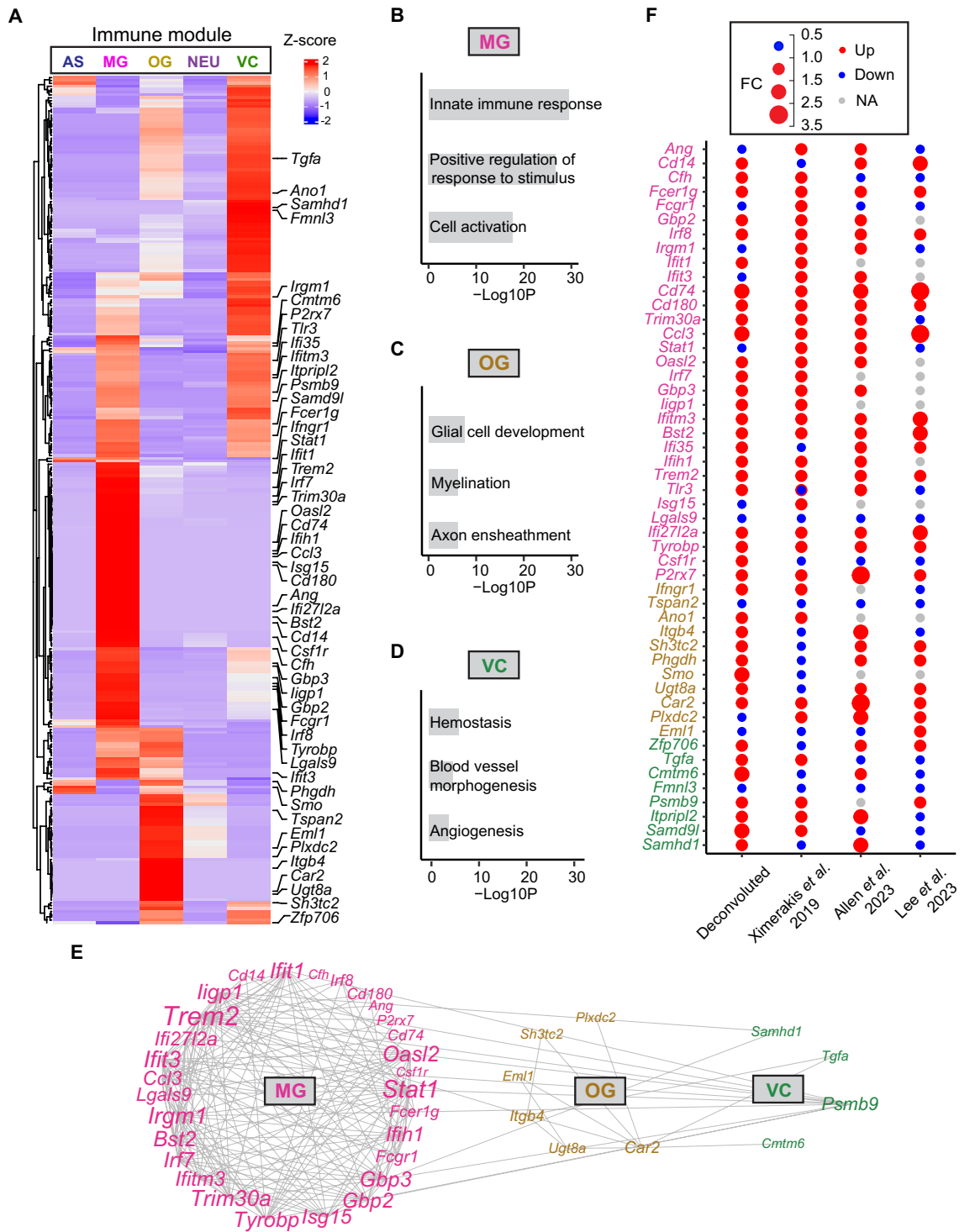


Figure 3. Immune gene module deconvolution. (A) Heatmap showing the expression of immune module genes in the five cell types. (B) Top GOs enriched in MG genes. All 3 GO terms were based on Biological Processes (BP) (C) Top GOs enriched in OG genes based on BP. (D) Top GOs enriched in VC genes. Hemostasis was based on Reactome Gene Sets, while Blood vessel morphogenesis and Angiogenesis were based on BP. (E) Correlation network of selected module genes from MG, OG and VC. (F) Dot plot showing gene fold changes of old versus young animals from our deconvoluted data, Ximerakis et al.¹², Allen et al.¹³ and Lee et al.¹⁴. MG genes are denoted in pink, OG genes in gold, and VC genes in green.

expressions. We then compared these changes with data from three prominent mouse aging single-cell RNA sequencing studies^{12–14} (Fig. 3F). Specifically, for the study by Ximerakis et al.¹², we extracted the fold changes

between old and young animals in each cell type as listed in their publication. In the studies of Allen et al.¹³, and Lee et al.¹⁴, we performed differential expression analyses comparing old versus young animals within each cell type using their respective Seurat data objects. Out of the 50 immune genes examined, 31 (62%) displayed the same change in expression direction with age as reported by Ximerakis et al.¹², 33 (66%) exhibited consistent changes with aging as observed in Allen et al.¹³, and 26 (52%) showed similar age-related changes as reported by Lee et al.¹⁴. In total, 45 out of these 50 genes (90%) were validated by at least one of the single-cell studies. Additionally, 15 genes (30%) were validated by all three studies. Five genes displayed opposite changes in expression direction with aging compared to all three single-cell studies or were expressed in less than 10% of cells in these studies. These genes include *Ifit3*, *Isg15*, *Csf1r*, *Smo*, and *Plxdc2* (Fig. 3F). It's worth noting that the discrepancies may be attributed to variations arising from factors such as animal sexes, ages, brain regions, tissue preparations, and cell dissociation procedures employed in these single-cell studies.

Collectively, our deconvoluted data indicate that microglia, together with oligodendroglia and vascular cells, play a pivotal role in orchestrating the immune response network within the brain during the aging process.

Cell type-specific analysis of gene modules related with lipid metabolism

Lipids play pivotal roles in cell signaling and various physiological processes within the brain and central nervous system. Disruptions in lipid homeostasis and impaired lipid metabolism have been implicated in both aging and AD pathogenesis^{3,24}. In our previous study¹¹, we identified an upregulated lipid metabolism module (part of the extracellular vesicle module) associated with aging. Given the potential relevance of these genes to aging and AD, we conducted further investigations focusing on this module. Our findings revealed that the lipid metabolism module predominantly comprised genes expressed in astrocytes (AS), microglia (MG), and oligodendroglia (OG) (Fig. 4A). Only a limited number of genes were attributed to neuronal population (NEU) and vascular cells (VC). Specifically, in AS, genes were significantly enriched for processes such as phospholipid metabolism ($-\log P = 48.4$), fatty acid metabolism ($-\log P = 43.9$), and phosphatidylinositol metabolism ($-\log P = 24.3$) (Fig. 4B). Key genes contributing to these pathways included *Mboat2*, *Mecp2*, *Pik3c2a*, *Gpld1*, *Inpp5a*, *Chpt1*, *Ocrl*, and *Lpin1*. In MG, genes were notably enriched for fatty acid metabolism ($-\log P = 25.1$), the PPAR signaling pathway ($-\log P = 8.4$), and the regulation of cholesterol metabolic process ($-\log P = 5.8$) (Fig. 4C). Genes associated with these pathways included *C3*, *Spp1*, *Pnpla7*, *Apobec1*, *Hcar2*, *Ggta1*, and *Gpx4*. OG genes exhibited enrichment in glycerolipid metabolic processes ($-\log P = 41.5$), phospholipid metabolic processes ($-\log P = 38.7$), and sphingolipid metabolic processes ($-\log P = 22.1$) (Fig. 4D). Genes contributing to these pathways encompassed *Cers2*, *Asah1*, *St8sia1*, *Smpdl3b*, and *Alg2*. The complete pathway analysis results can be found in Supplementary Table S2.

Furthermore, our correlation network analysis of module genes across all three major cell types revealed both intra- and inter-cellular interactions. The most robust connections were observed among AS genes (*Inpp5a*, *Chpt1*, *Gpld1*, *Lpin1*, and *Mecp2*), MG genes (*Spp1*, *C3*, and *Pnpla7*), and OG genes (*Asah2* and *Cers2*) (Fig. 4E).

To validate the aging-associated expression changes of these key genes, we again compared fold changes between old and young animals in our deconvoluted data with the three single-cell RNA sequencing studies^{12–14}. We found that, out of the 38 lipid metabolism genes, 25 (65.7%) exhibited consistent directional changes with age as reported by Ximerakis et al.¹², 17 (44.7%) with Allen et al.¹³, and 14 (36.8%) with Lee et al.¹⁴. In total, 34 out of these 38 genes (89.5%) were validated by at least one single-cell study. Of note, two genes—*Lpin1* and *Nus1*—were validated by all three studies, while four genes displayed opposite aging-related changes compared to these three single-cell studies or were expressed in less than 10% of cells in these single-cell studies, including *Pigv*, *Fitm2*, *Smpdl3b*, and *Bdh2* (Fig. 4F).

Altogether, our deconvoluted data indicate both similarities and differences in lipid dysregulation across various cell types. Further investigation is required to elucidate how these lipid changes influence cellular functions.

Cell type-specific analysis of gene modules related with mitochondrial function

Mitochondrial dysfunction and diminished energy metabolism are prominent neuropathological characteristics of AD²⁵ and represent key hallmarks of the aging process²⁶. In our prior research, we identified a module of genes associated with mitochondria that displayed a downregulation pattern with aging and enriched for respiratory chain pathways. Notably, our deconvolution analyses revealed that these genes were predominantly expressed in vascular cells (VC) (Fig. 5A). This module comprised pivotal hub genes (Fig. 5B), including members of the Cytochrome C Oxidase gene family such as *Cox6a1*, *Cox7a2*, *Cox8a*, and *Cox7b*, alongside genes from the mitochondrial NADH: ubiquinone oxidoreductase gene family, including *Ndufa1*, *Ndufa8*, *Ndufb3*, and *Ndufb6*. Additionally, prominent genes from the Ubiquinol-Cytochrome C Reductase gene family, such as *Uqcrb* and *Uqcr10*, were also featured within this module (Fig. 5B).

The further validation with the three single-cell RNA sequencing studies showed that, among the 28 mitochondria function genes, 23 (82.1%) exhibited concordant changes with age as reported by Ximerakis et al.¹², 6 genes (21.4%) aligned with Allen et al.¹³, and 16 genes (57.1%) were consistent with Lee et al.¹⁴ (Fig. 5C). Importantly, 26 out of these 28 genes (92.9%) were validated by at least one single-cell study. Two genes, *Cox7a1* and *Cox7b2*, displayed opposing changes with aging when compared to these three single-cell studies or were expressed in less than 10% of cells in these single-cell studies. It is worth noting that the module of mitochondria-related genes exhibited the most variation when compared to Allen et al.¹³. This divergence is likely due to inadequate sampling of certain vascular cell types, including pericytes and endothelial cells, in the Allen et al.¹³ study, potentially leading to inaccuracies in gene expression profiling.

Overall, the consistent downregulation of these mitochondrial function genes with aging, as reported in our study and corroborated by Ximerakis et al.¹² and Lee et al.¹⁴, aligns with the established notion that impaired

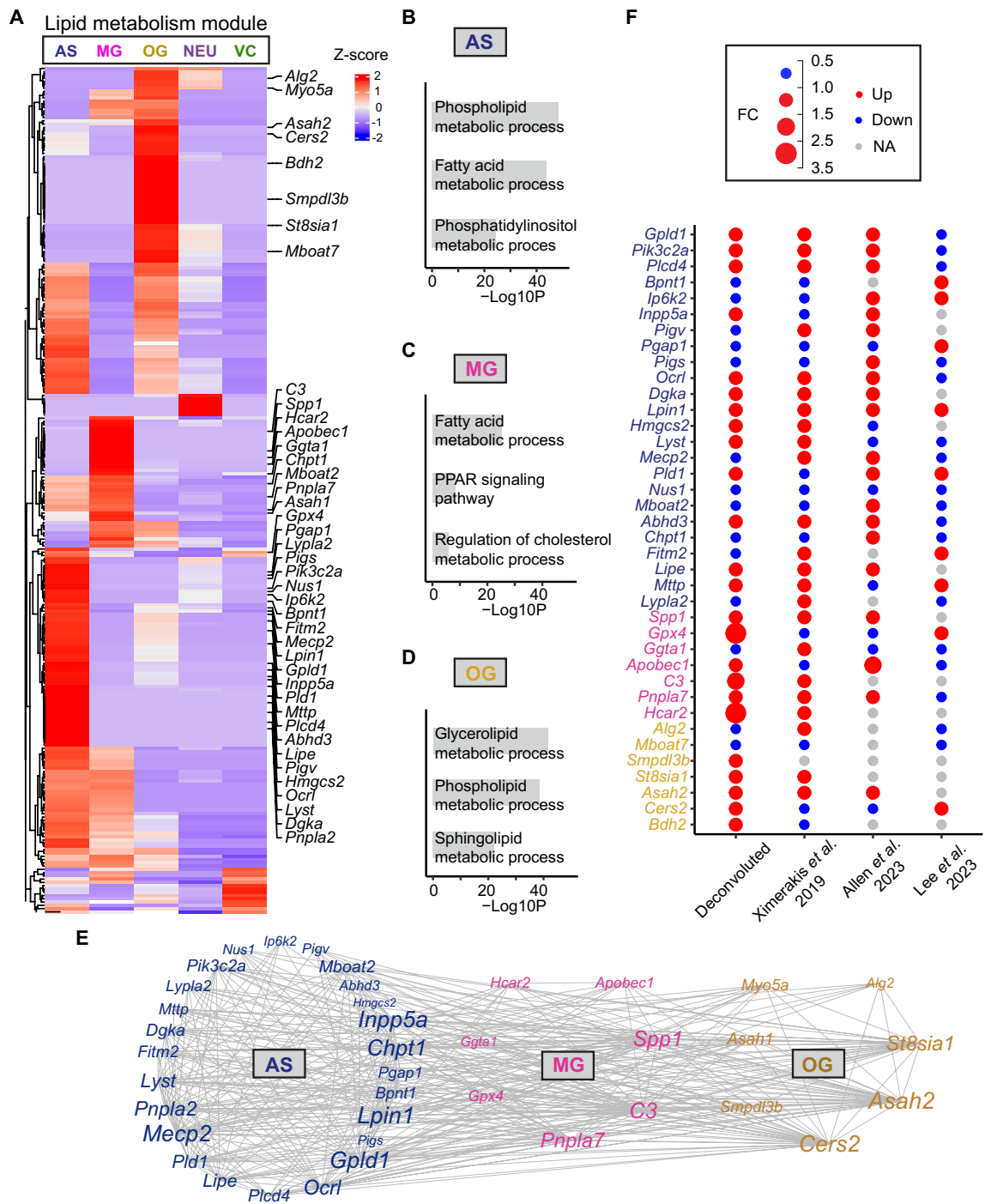


Figure 4. Lipid metabolism gene module deconvolution. (A) Heatmap showing the lipid metabolism module genes in the five cell types. (B) Top GOs enriched in AS genes. All 3 GO terms were based on Biological Processes (BP). (C) Top GOs enriched in MG genes. Fatty acid metabolic process and Regulation of cholesterol metabolic process were based on BP, while PPAR signaling pathway was based on WikiPathways. (D) Top GOs enriched in OG genes based on BP. (E) Correlation network of selected module genes from AS, MG and OG. (F) Dot plot showing gene fold changes of old versus young animals from our deconvoluted data, Ximerakis et al.¹², Allen et al.¹³, and Lee et al.¹⁴. AS genes are denoted in dark blue, MG genes in pink, and OG genes in gold.

mitochondrial signaling in the endothelium constitutes an early event in the aging process and contributes causally to the development of various age-related diseases²⁷.

Cell type-specific analysis of gene modules related with synapses

Accumulated research over the years has firmly established that synaptic transmission undergoes changes as individuals age. Indeed, studies have shown that genes regulated by age are associated with synaptic functions²⁸.

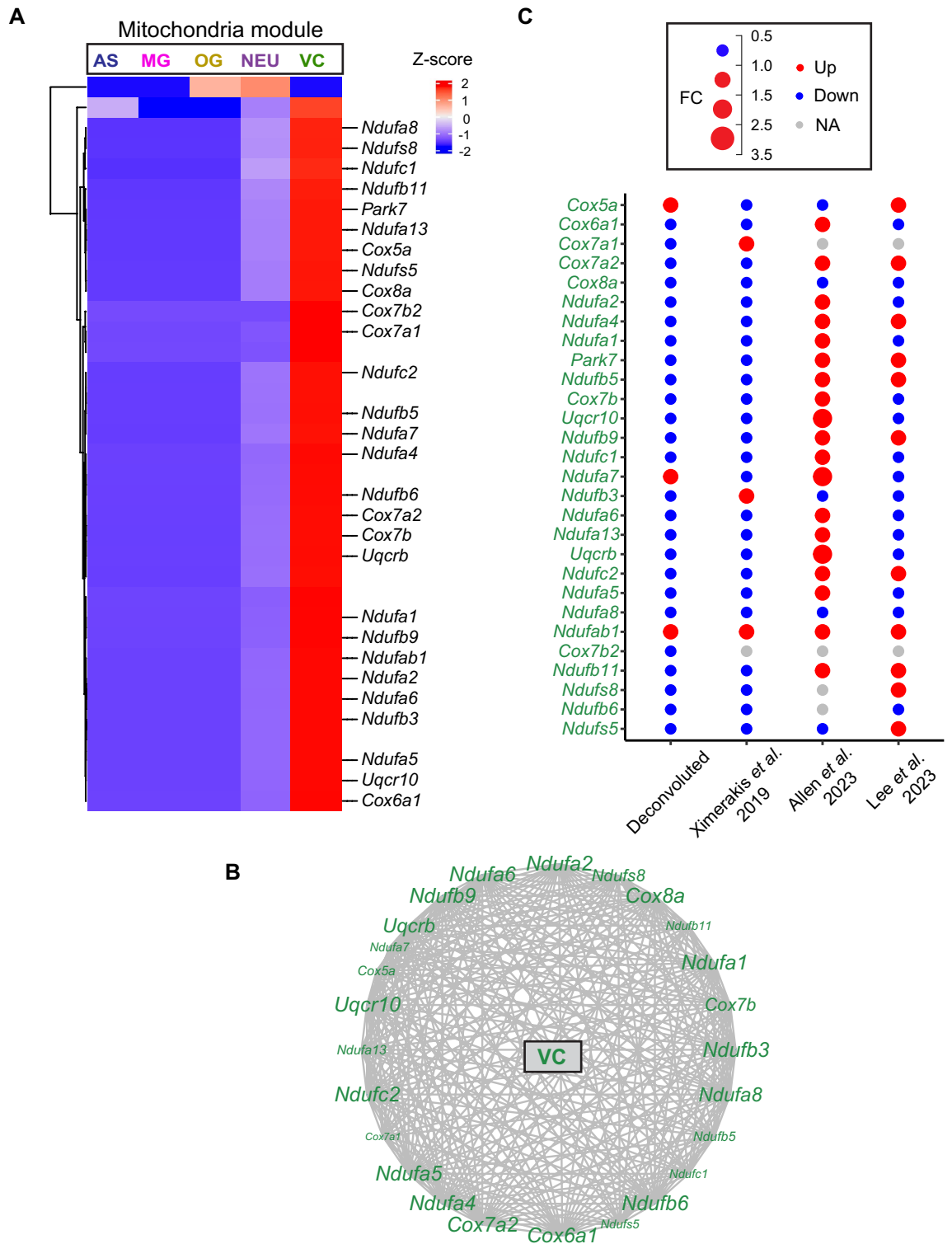


Figure 5. Mitochondria gene module deconvolution. **(A)** Heatmap showing the respiration chain module genes in the five cell types. **(B)** Correlation network of VC genes. **(C)** Dot plot showing gene fold changes of old versus young animals from our deconvoluted data, Ximerakis et al.¹², Allen et al.¹³, and Lee et al.¹⁴. VC genes are denoted in green.

In our investigation, we deconvoluted the synapse module, which exhibited downregulation with aging. Our analysis revealed that the majority of genes within this module were expressed in neurons (NEU), with some also expressed in vascular cells (VC) (Fig. 6A). A subset of genes was associated with astrocytes (AS), demonstrating

overlapping expression with the neuronal population (Fig. 6A). Key hub genes within this synapse module included *Madd*, *Shank2*, *Baiap2*, *Nrcam*, and *Dmtn*, all from NEU (Fig. 6B).

In the validation with single-cell RNA sequencing data sets, we found that out of the 30 neuronal genes, 21 (70%) exhibited consistent directional changes with Ximerakis et al.¹², and 15 (53.3%) with Allen et al.¹³. Of note, there were not many genes captured by Lee et al.¹⁴ due to the lack of neuronal population in this dataset. In total, 27 out of these 30 neuronal genes (90%) were validated by at least one single-cell study (Fig. 6C). Three genes displayed opposing changes with aging compared to these three single-cell studies or were expressed in less than 10% of cells in these single-cell studies, including *Lzts3*, *Oprd1*, and *Pak1*. Regarding the genes associated with vascular cells (VC), our validation using all three single-cell studies indicated that 3 out of the total 9 (33.3%) VC genes exhibited consistent aging-related changes with Ximerakis et al.¹², 8 (88.9%) with Allen et al.¹³, and 5 (55.6%) showed similar aging-related changes with Lee et al.¹⁴. In total, 8 out of 9 VC genes (88.9%) were validated by at least one single-cell study, with gene *Adrbk1* being the only exception, showing opposite changes with aging compared to these three single-cell studies or being expressed in less than 10% of cells in these single-cell studies.

Together, these data provide support for the concept of synaptic alterations occurring with aging, as well as for the critical network changes of neurovascular coupling and neuronal-glia-endothelial interactions observed during aging^{29,30}.

Discussion

In this study, we applied cell type deconvolution to bulk RNA sequencing data from apoE-TR mice to identify key cell types involved in gene modules related to aging. We also explored the molecular pathways and interactions between module genes within and across cell types. Additionally, we identified hub genes with significant connections to others, providing a further understanding of the contributions of different cell types to important aging-regulated pathways. Most of our deconvoluted aging-associated gene expression changes were validated by one or more previously published single cell studies.

During the aging process, the brain undergoes various immune responses, such as neuroinflammation, microglial activation, blood–brain barrier dysfunction, myelin disintegration, immune cell senescence, and altered interactions between immune cells and other brain cells². These immune changes contribute to cognitive decline and increase the risk of neurodegenerative diseases. Our deconvolution data consistently highlights the involvement of microglia, oligodendrocytes, and vascular cells in the immune response during aging. It was reported that microglia and oligodendrocytes have mutual interactions and influence each other's functions³¹. Microglia play a role in modulating immune responses and impacting the survival and differentiation of oligodendrocytes. Additionally, brain vascular cells, including endothelial cells and pericytes, have a dual role in maintaining the integrity of the blood–brain barrier and participating in immune responses³². They interact with immune cells and influence inflammation in the brain. Further understanding the intricate crosstalk among these brain cells and their interactions with immune responses is crucial for unraveling disease mechanisms and developing targeted therapies.

Furthermore, our deconvolution analysis reveals dysregulation of lipids in astrocytes, microglia, and oligodendrocytes, with variations in the specific lipid processes involved in each cell type. Both astrocytes and microglia show affected fatty acid metabolism, which has been linked to neuroinflammation^{33–35}. It was reported that the pro-inflammatory microglia suppress fatty acid oxidation and synthesis, while anti-inflammatory microglia enhance these processes³⁵. Additionally, microglia show dysregulation in the PPAR signaling pathway, known to modulate microglial innate immunity and fatty acid metabolism^{36,37}. These findings suggest a potential link between lipid dysregulation and neuroinflammation in astrocytes and microglia. In oligodendrocytes, we observe dysregulation in glycerolipid, phospholipid, and sphingolipid processes. While the understanding of glycerolipid metabolism in oligodendrocytes is limited, it is crucial to acknowledge the significant roles of phospholipids and sphingolipids in myelin composition^{38,39}. Importantly, lipid metabolism in oligodendrocytes has been implicated in the development of demyelination, a key factor contributing to the risk of AD⁴⁰. Taken together, these findings highlight age-related lipid dysregulation that may contribute to neuroinflammation and demyelination during the aging process and in AD.

Mitochondrial dysfunction is a key feature of aging, characterized by structural and functional changes in the mitochondria². This leads to reduced energy production, increased production of reactive oxygen species, and impaired mitochondrial quality control. Interestingly, our deconvolution data suggests strong mitochondrial respiration dysregulation in brain vascular cells. It is known that endothelial mitochondrial dysfunction is an important factor causing abnormal function of the endothelium, which plays a central role during atherosclerosis development⁴¹. These changes can also disrupt the integrity and function of the blood–brain barrier, resulting in increased permeability and inflammation in the brain.

Our deconvoluted data imply that the downregulated synapse module with aging includes genes primarily originating from neurons, vascular cells, and astrocytes. This suggests that during the aging process, it's not just neurons experiencing synaptic loss; the interactions or networks between neurons, glial cells, and vascular cells are also disrupted. This observation aligns well with our current understanding that the brain is a complex tissue, where vascular cells and glial cells closely collaborate with neurons to maintain brain functions^{29,30}. These network gene changes can be challenging to detect in single-cell RNA sequencing data if specific cell types are not adequately captured during the cell isolation process.

In our prior study, we reported several gene module changes associated with the *APOE* genotype, including lysosome functions and RNA splicing¹¹. However, our deconvolution results indicate that the majority of these genes are attributed to neurons (data not shown). Unfortunately, because the neuronal population was not adequately represented in our single-cell validation dataset¹⁴, we were unable to validate and present this data

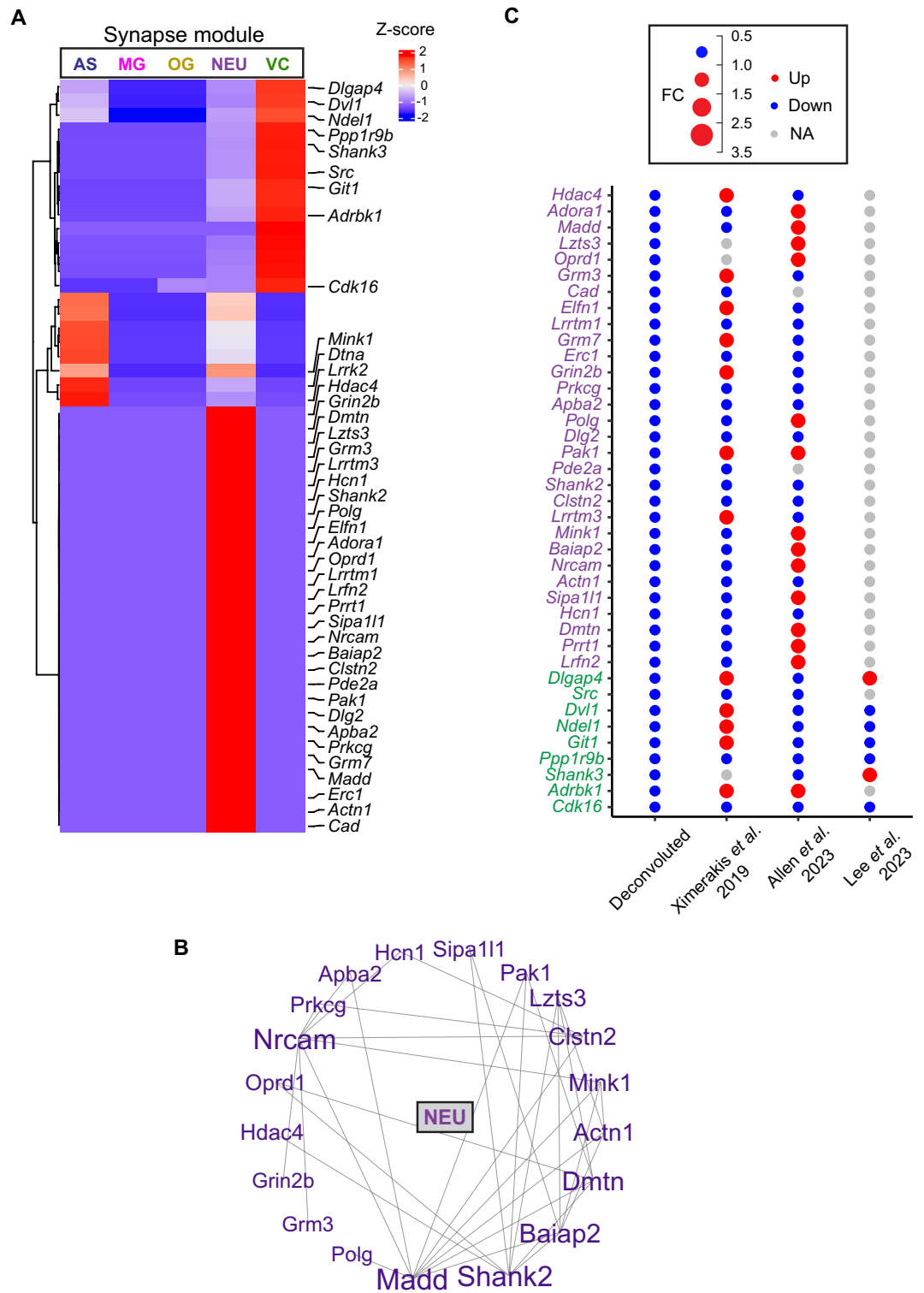


Figure 6. Synapse gene module deconvolution. (A) Heatmap showing the synapse module genes in the five cell types. (B) Correlation network of NEU genes. (C) Dot plot showing gene fold changes of old versus young animals from our deconvoluted data, Ximerakis et al.¹², Allen et al.¹³ and Lee et al.¹⁴. NEU genes are denoted in purple, and VC genes in green.

in the current study. Therefore, the analysis and validation of *APOE* genotype-related gene module changes will be a subject of future investigation.

It has long been recognized that cell heterogeneity in tissue samples leads to the caveat in bulk RNA sequencing experiments that it is often hard to dissociate true transcriptomics landscape changes from the contribution of cell types. In the study of diseases where certain cell types play a major role in disease pathogenesis, cell-type specific information is particularly pertinent. Computational deconvolution methods have the unique advantage of being readily available as open-source packages, and can be applied to old or previously published data where samples are no longer available or difficult to re-obtain. The success of our application of the deconvolution algorithm in this study serves as proof of principle for the potential of applying computational deconvolution methods to the vast available bulk public dataset to gain novel insights of disease mechanisms.

Limitations of the study

Our study has several limitations, some of which are inherent to the use of computational deconvolution. First, it is important to acknowledge that computational methods are susceptible to generating false positives. In our case, the transition from bulk gene expression to cell-type-specific gene expression involved multiple layers of machine learning processes, which collectively could lead to inaccurate estimations of gene expression in certain cell types. To mitigate this, it is essential to cross-reference our results with publicly available databases and existing literature, ensuring that we identify and rectify any such estimations and avoid overgeneralization of our findings. For genes with limited public information, experimental validations, such as co-staining, will be crucial to confirm the accuracy of the sample-level deconvolution results in the future. Second, it's important to recognize that different computational deconvolution algorithms may be better suited for specific types of bulk transcriptomics data. Many deconvolution methods have been developed for or evaluated on blood, immune, or tumor samples, and their performance may not be equivalent for tissues with high complexity, such as the brain^{7,10,42}. Depending on the specific characteristics of the bulk tissue being studied, it becomes imperative to carefully evaluate the choice of deconvolution method and exercise caution in its application. Acknowledging these limitations, future study should aim to minimize potential errors through rigorous validation strategies. It is also important to critically interpret our findings within the context of existing knowledge, recognizing that computational deconvolution, while a powerful tool, must be employed judiciously and complemented with experimental evidence when necessary.

Methods

Dataset description

We used our previously published cerebral cortex transcriptomics data of 141 apoE-TR animals, including male and female *APOE2*, *APOE3* and *APOE4* mice at 3, 12, and 24 months of age¹¹. The animal husbandry, tissue processing, RNA extraction, RNA sequencing, quality control and gene count normalization were previously described¹¹. The CQN normalized RPKM gene expression values of 19,120 genes were used for deconvolution analyses.

Cell type deconvolution and gene correlation network analysis

We used the CIBERSORTx algorithm high resolution mode to deconvoluted the bulk transcriptomics data based on the author's recommendations⁸. Specifically, we used the mouse brain single cell RNA sequencing data from GSE129788 to construct the single cell signature gene matrix due to the data's close match in both age and brain region to our bulk data¹². From the cell-type-specific gene expression results, we removed genes that were either "NA" or "-1", or had the same deconvoluted expression values across all samples, which indicated insufficient evidence of estimation. Module genes were scaled and visualized via hierarchical clustering in heatmaps using the ComplexHeatmap R package⁴³. Pathway analyses of cell type specific module genes were performed using Metascape (<https://metascape.org>)⁴⁴. Gene correlation network analysis of selected modules were performed using WGCNA^{45,46} and visualized using Cytoscape version v3.10.0⁴⁷.

Validation of cell-type-specific module genes using public data

To validate the deconvoluted cell-type-specific module genes, we compared their cell type assignments against the following sources: (1) top 50 mouse brain cell type marker genes from each brain cell type from the BRETIGEA R package, which compared and contrasted five human and mouse cell type-specific transcriptomics datasets to identify consensus brain cell-type marker genes⁹, (2) brain single cell RNA sequencing expression profiles from the Human Brain Atlas v22.0 (proteinatlas.org), (3) published mouse brain MG gene expression profiles across different activation states¹⁵, (4) published AS gene expression profiles in mouse brain cortical regions¹⁶, and (5) published OG gene expression profiles across various mouse brain regions¹⁷.

To validate aging-associated changes in our cell-type-specific module genes, we compared fold changes of old versus young animals between our deconvoluted data and three prominent mouse aging single cell studies¹²⁻¹⁴. For our deconvoluted genes in each cell type, fold changes were calculated between 24 and 3 months using an ANOVA model while adjusting for *APOE* genotype and sex. The fold changes between old (21–23 months) and young (2–3 month) mice from Ximerakis et al.¹² were obtained from their supplementary Table 6, which was calculated using the MAST model⁴⁸ for each annotated single cell type. To calculate fold changes in Allen et al.¹³, we downloaded the integrated and annotated R data object GSE207848_Cell.rds from CELL x GENE repository (<https://cellxgene.cziscience.com/collections/31937775-0602-4e52-a799-b6acdd2bac2e>), and used the MAST model to compare 90 weeks (22.5 months) and 4 weeks (1 month) for each single cell type using Seurat version 5⁴⁹. To calculate fold changes in Lee et al.¹⁴, we obtained the integrated and annotated R data file from the authors and used the MAST model to compare 24 months and 3 months for each single cell type while adjusting for *APOE* genotype using Seurat version 5.

Data availability

The bulk transcriptomics data that were deconvoluted in this study are available on Synapse: <https://doi.org/10.7303/syn20808171>.

Received: 6 July 2023; Accepted: 4 October 2023

Published online: 06 October 2023

References

- Morrison, J. H. & Baxter, M. G. The ageing cortical synapse: Hallmarks and implications for cognitive decline. *Nat. Rev. Neurosci.* **13**, 240–250. <https://doi.org/10.1038/nrn3200> (2012).
- Amorim, J. A. *et al.* Mitochondrial and metabolic dysfunction in ageing and age-related diseases. *Nat. Rev. Endocrinol.* **18**, 243–258. <https://doi.org/10.1038/s41574-021-00626-7> (2022).
- Kao, Y. C., Ho, P. C., Tu, Y. K., Jou, I. M. & Tsai, K. J. Lipids and Alzheimer's disease. *Int. J. Mol. Sci.* **21**, 1505. <https://doi.org/10.3390/ijms21041505> (2020).
- Lucin, K. M. & Wyss-Coray, T. Immune activation in brain aging and neurodegeneration: Too much or too little?. *Neuron* **64**, 110–122. <https://doi.org/10.1016/j.neuron.2009.08.039> (2009).
- Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* **50**, 1–14. <https://doi.org/10.1038/s12276-018-0071-8> (2018).
- Haque, A., Engel, J., Teichmann, S. A. & Lonnberg, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* **9**, 75. <https://doi.org/10.1186/s13073-017-0467-4> (2017).
- Chikina, M., Zaslavsky, E. & Sealfon, S. C. CellCODE: A robust latent variable approach to differential expression analysis for heterogeneous cell populations. *Bioinformatics* **31**, 1584–1591. <https://doi.org/10.1093/bioinformatics/btv015> (2015).
- Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782. <https://doi.org/10.1038/s41587-019-0114-2> (2019).
- McKenzie, A. T. *et al.* Brain cell type specific gene expression and co-expression network architectures. *Sci. Rep.* **8**, 8868. <https://doi.org/10.1038/s41598-018-27293-5> (2018).
- Jew, B. *et al.* Accurate estimation of cell composition in bulk expression through robust integration of single-cell information. *Nat. Commun.* **11**, 1971. <https://doi.org/10.1038/s41467-020-15816-6> (2020).
- Zhao, N. *et al.* Alzheimer's risk factors age, APOE genotype, and sex drive distinct molecular pathways. *Neuron* **106**, 727–742.e726. <https://doi.org/10.1016/j.neuron.2020.02.034> (2020).
- Ximerakis, M. *et al.* Single-cell transcriptomic profiling of the aging mouse brain. *Nat. Neurosci.* **22**, 1696–1708. <https://doi.org/10.1038/s41593-019-0491-3> (2019).
- Allen, W. E., Blosser, T. R., Sullivan, Z. A., Dulac, C. & Zhuang, X. Molecular and spatial signatures of mouse brain aging at single-cell resolution. *Cell* **186**, 194–208.e118. <https://doi.org/10.1016/j.cell.2022.12.010> (2023).
- Lee, S. *et al.* APOE modulates microglial immunometabolism in response to age, amyloid pathology, and inflammatory challenge. *Cell Rep.* **42**, 112196. <https://doi.org/10.1016/j.celrep.2023.112196> (2023).
- Cho, C. E. *et al.* A modular analysis of microglia gene expression, insights into the aged phenotype. *BMC Genom.* **20**, 164. <https://doi.org/10.1186/s12864-019-5549-9> (2019).
- Hasel, P., Rose, I. V. L., Sadick, J. S., Kim, R. D. & Liddel, S. A. Neuroinflammatory astrocyte subtypes in the mouse brain. *Nat. Neurosci.* **24**, 1475–1487. <https://doi.org/10.1038/s41593-021-00905-6> (2021).
- Tommasini, D., Fox, R., Ngo, K. J., Hinman, J. D. & Fogel, B. L. Alterations in oligodendrocyte transcriptional networks reveal region-specific vulnerabilities to neurological disease. *iScience* **26**, 106358. <https://doi.org/10.1016/j.isci.2023.106358> (2023).
- Li, Y. *et al.* Proteomic profile of mouse brain aging contributions to mitochondrial dysfunction, DNA oxidative damage, loss of neurotrophic factor, and synaptic and ribosomal proteins. *Oxid. Med. Cell Longev.* **2020**, 5408452. <https://doi.org/10.1155/2020/5408452> (2020).
- Rodrigues, L. P. *et al.* Hallmarks of aging and immunosenescence: Connecting the dots. *Cytokine Growth Factor Rev.* **59**, 9–21. <https://doi.org/10.1016/j.cytogfr.2021.01.006> (2021).
- Bettcher, B. M., Tansey, M. G., Dorothée, G. & Heneka, M. T. Peripheral and central immune system crosstalk in Alzheimer disease—A research prospectus. *Nat. Rev. Neurol.* **17**, 689–701. <https://doi.org/10.1038/s41582-021-00549-x> (2021).
- Zhao, N. *et al.* Elevating microglia TREM2 reduces amyloid seeding and suppresses disease-associated microglia. *J. Exp. Med.* <https://doi.org/10.1084/jem.20212479> (2022).
- Falcao, A. M. *et al.* Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. *Nat. Med.* **24**, 1837–1844. <https://doi.org/10.1038/s41591-018-0236-y> (2018).
- Kirby, L. *et al.* Oligodendrocyte precursor cells present antigen and are cytotoxic targets in inflammatory demyelination. *Nat. Commun.* **10**, 3887. <https://doi.org/10.1038/s41467-019-11638-3> (2019).
- Giri, M., Zhang, M. & Lü, Y. Genes associated with Alzheimer's disease: An overview and current status. *Clin. Interv. Aging* <https://doi.org/10.2147/CIA.S105769> (2016).
- Moreira, P. I., Santos, M. S. & Oliveira, C. R. Alzheimer's disease: A lesson from mitochondrial dysfunction. *Antioxid. Redox Signal.* **9**, 1621–1630 (2007).
- López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217. <https://doi.org/10.1016/j.cell.2013.05.039> (2013).
- Groschner, L. N., Waldeck-Weiermair, M., Malli, R. & Graier, W. F. Endothelial mitochondria—less respiration, more integration. *Pflugers Arch.* **464**, 63–76. <https://doi.org/10.1007/s00424-012-1085-z> (2012).
- Loerch, P. M. *et al.* Evolution of the aging brain transcriptome and synaptic regulation. *PLoS One* **3**, e3329. <https://doi.org/10.1371/journal.pone.0003329> (2008).
- Tarantini, S., Tran, C. H. T., Gordon, G. R., Ungvari, Z. & Csiszar, A. Impaired neurovascular coupling in aging and Alzheimer's disease: Contribution of astrocyte dysfunction and endothelial impairment to cognitive decline. *Exp. Gerontol.* **94**, 52–58. <https://doi.org/10.1016/j.exger.2016.11.004> (2017).
- Kugler, E. C., Greenwood, J. & MacDonald, R. B. The “Neuro-Glia-Vascular” unit: The role of Glia in neurovascular unit formation and dysfunction. *Front. Cell Dev. Biol.* **9**, 732820. <https://doi.org/10.3389/fcell.2021.732820> (2021).
- Peferoen, L., Kipp, M., van der Valk, P., van Noort, J. M. & Amor, S. Oligodendrocyte-microglia cross-talk in the central nervous system. *Immunology* **141**, 302–313. <https://doi.org/10.1111/imm.12163> (2014).
- Stolp, H. B., Liddel, S. A., Sa-Pereira, I., Dziegielewska, K. M. & Saunders, N. R. Immune responses at brain barriers and implications for brain development and neurological function in later life. *Front. Integr. Neurosci.* **7**, 61. <https://doi.org/10.3389/fnint.2013.00061> (2013).
- Gupta, S., Knight, A. G., Gupta, S., Keller, J. N. & Bruce-Keller, A. J. Saturated long-chain fatty acids activate inflammatory signaling in astrocytes. *J. Neurochem.* **120**, 1060–1071 (2012).
- Ioannou, M. S. *et al.* Neuron-astrocyte metabolic coupling protects against activity-induced fatty acid toxicity. *Cell* **177**, 1522–1535.e1514 (2019).

35. Geric, I. *et al.* Metabolic reprogramming during microglia activation. *Immunometabolism* <https://doi.org/10.20900/immunometab20190002> (2019).
36. Tyagi, S., Gupta, P., Saini, A. S., Kaushal, C. & Sharma, S. The peroxisome proliferator-activated receptor: A family of nuclear receptors role in various diseases. *J. Adv. Pharm. Technol. Res.* **2**, 236–240. <https://doi.org/10.4103/2231-4040.90879> (2011).
37. Shie, F. S., Nivison, M., Hsu, P. C. & Montine, T. J. Modulation of microglial innate immunity in Alzheimer's disease by activation of peroxisome proliferator-activated receptor gamma. *Curr. Med. Chem.* **16**, 643–651. <https://doi.org/10.2174/092986709787458399> (2009).
38. Nave, K. A. & Werner, H. B. Myelination of the nervous system: Mechanisms and functions. *Annu. Rev. Cell Dev. Biol.* **30**, 503–533. <https://doi.org/10.1146/annurev-cellbio-100913-013101> (2014).
39. Montani, L. Lipids in regulating oligodendrocyte structure and function. *Semin. Cell Dev. Biol.* **112**, 114–122. <https://doi.org/10.1016/j.semcdb.2020.07.016> (2021).
40. Marangon, D., Boccazzi, M., Lecca, D. & Fumagalli, M. Regulation of oligodendrocyte functions: Targeting lipid metabolism and extracellular matrix for myelin repair. *J. Clin. Med.* **9**, 470. <https://doi.org/10.3390/jcm9020470> (2020).
41. Qu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313> (2016).
42. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523. <https://doi.org/10.1038/s41467-019-09234-6> (2019).
45. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559. <https://doi.org/10.1186/1471-2105-9-559> (2008).
46. Langfelder, P. & Horvath, S. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, 1–17. <https://doi.org/10.18637/jss.v046.i11> (2012).
47. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
48. Finak, G. *et al.* MAST: A flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278. <https://doi.org/10.1186/s13059-015-0844-5> (2015).
49. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587. <https://doi.org/10.1016/j.cell.2021.04.048> (2021).

Acknowledgements

This work was supported by NIH grants U19AG069701 (to Y.R., Y.W.A., and N.Z.), RF1AG046205 (to N.Z.), and R01AG66395 (to N.Z.); and a BrightFocus Foundation grant (to N.Z.).

Author contributions

Y.R. and N.Z. designed the study and wrote the manuscript. Y.R., X.W. and S.Z. performed deconvolution analysis. S.L., J.M.M. and L.A.J. helped with single cell data validation. H.H., Z.Q. and Y.W.A. provided intellectual input, reviewed and edited the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44183-7>.

Correspondence and requests for materials should be addressed to Y.R. or N.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023