Cell-specific Gene Regulatory Networks in Alzheimer's Disease

A Differential Analysis of Regulatory Changes Across Cell Types

Ivo Haršáni[1]
Supervisors: Marcel Reinders[1]**, Bram Pronk**[1]**, Timo Verlaan**[1]
[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 21, 2026

Name of the student: Ivo Haršáni
Final project course: CSE3000 Research Project
Thesis committee: Marcel Reinders, Bram Pronk, Timo Verlaan, Sicco Verwer

Abstract

Alzheimer's disease (AD) does not affect all brain cells in the same way, and a cell's dysfunction is thought to involve changes in how its genes are regulated rather than only changes in their expression. Gene regulatory networks (GRNs) model these regulatory decisions, and recent single-cell methods can infer a separate network for each individual cell. How such cell-specific networks change in disease, and whether any changes are shared across cell types or unique to particular ones, remains largely unexplored. This work asks which regulatory relationships differ between AD and healthy cells, and to what extent these differences are cell-type specific. To address this, the cell-specific GRN method ScReNI is reimplemented in Python and applied to paired single-nucleus RNA and ATAC data from the SEA-AD cohort (27 donors), inferring one network per cell for microglia and an excitatory-neuron subclass (L2/3 IT). A differential pipeline then compares the networks against disease severity at the level of individual transcription-factor-to-target edges and of co-regulated gene modules, complemented by a module-preservation test, treating the donor as the unit of replication. The inferred networks change relatively little with disease: no edge survives stringent correction and weight-based filtering, the module co-regulation structure is preserved, and no module-specific shift is detected in either cell type. The one robust signal is a modest, severity-graded decline in overall regulatory activity in L2/3 IT neurons that persists after adjusting for sequencing depth and is absent in microglia. The results are best read as an absence of strong evidence rather than evidence of absence, motivating larger cohorts and broader gene panels in future work.

1 Introduction

Alzheimer's disease (AD) is the most common cause of dementia, and despite decades of research, only limited treatments exist that modestly slow progression. One reason is that AD does not affect the brain uniformly. Different cell types – neurons, microglia – play distinct roles, and the way each cell type becomes dysfunctional in disease is thought to involve changes in how its genes are *regulated* rather than simply changes in gene expression levels. To understand the disease, it is therefore needed to understand the regulatory machinery that decides, in each individual cell, which genes are turned on and off.

A gene regulatory network (GRN) is a model of those decisions: transcription factors act as regulators and the genes they control as targets, with edges weighted by the strength of regulation. Methods for inferring GRNs from bulk tissue have existed for years [6], but they average over millions of cells and miss the heterogeneity that defines a disease like AD. The arrival of single-cell RNA sequencing (scRNA-seq) and single-cell ATAC sequencing (scATAC-seq) made it possible to measure gene expression and chromatin accessibility one cell at a time, enabling a far finer view of regulation [7].

Several methods now exist for inferring GRNs from single-cell data. SCENIC [1] infers networks at the cell-type level from scRNA-seq alone. LINGER [12] integrates paired multiomic data with external atlases. Most recently, ScReNI [11] combines a nearest-neighbours approach with random forests to infer a separate regulatory network for every individual cell, using both scRNA-seq and scATAC-seq. ScReNI has been benchmarked extensively on healthy reference data such as peripheral blood mononuclear cells (PBMCs) and on retinal development. What remains largely unexplored is how these cell-specific networks change in disease: which regulatory edges or modules differ between AD and healthy cells, and whether those changes are shared across cell types or unique to specific ones.

This work addresses this gap. The ScReNI pipeline is reimplemented in Python and

applied to AD single-cell data to infer one regulatory network per cell. A differential framework then compares these networks against disease severity, both at the level of individual transcription-factor-to-target edges and at the level of modules of co-regulated genes, in two contrasting cell types: microglia, the immune population most strongly implicated in AD, and L2/3 IT, an abundant excitatory-neuron subclass.

2 Methodology

This section describes the data, the network-inference procedure, and the statistical tests used to compare the gene regulatory networks between AD and healthy donors. The comparison builds on a Python reimplementation of ScReNI [11], which infers one regulatory network per cell. Two complementary analyses are applied: a differential test on individual transcription-factor→target-gene (TF→TG) edges (Section 2.4) and a module-level analysis of co-regulated genes (Section 2.5).

2.1 Data: the SEA-AD paired multiome cohort

The data are drawn from the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD) [3], specifically the middle temporal gyrus profiled with the paired 10x multiome assay, in which snRNA-seq and snATAC-seq are measured from the same nuclei. This pairing is a requirement of wScReNI, which integrates the two modalities for each cell. The paired subset consists of 138,118 cells from only 28 donors; because the statistical tests treat the donor as the unit of replication, this donor count is the main constraint on the analysis design.

Disease status in SEA-AD is captured by the Overall AD Neuropathological Change (ADNC) score, a standardised post-mortem measure of amyloid-β and tau pathology graded on four ordinal levels: Not AD, Low, Intermediate, and High. Within the 28-donor paired multiome subset, these levels comprise 3, 4, 10, and 11 donors, respectively. For analyses that require a reference contrast, these four levels are collapsed into two groups: donors with ADNC Not AD or Low (lower-ADNC) and those with Intermediate or High (higher-ADNC).

The metadata additionally record two age-related co-pathologies that frequently accompany AD: limbic-predominant age-related TDP-43 encephalopathy (LATE), a TDP-43 proteinopathy of the medial temporal lobe, and Lewy body disease (LBD), defined by α-synuclein aggregates. Both are recorded as per-donor indicators.

2.2 Cell-type annotation and selection

Cell-type labels are taken directly from the annotations released with SEA-AD [3], where each nucleus was assigned the cortical cell type whose reference expression profile it most closely matched. The analysis uses the subclass level of this annotation and assigns no labels of its own.

Two subclasses are examined: Microglia-PVM, the resident immune population most strongly implicated in AD [4], and L2/3 IT, an abundant excitatory-neuron subclass that places a neuronal compartment alongside the immune one. For each donor, 50 cells are sampled per subclass, enough to form a stable per-donor aggregate while bounding the number of cells any single donor contributes. One of the 28 donors contributes no cells to either subclass and is excluded, leaving 27 donors and 2,552 cells (1,311 L2/3 IT, 1,241 Microglia-PVM).

2.3 Network inference with ScReNI

Per-cell networks are inferred with ScReNI [11] (Figure 1). In brief, ScReNI defines each cell's neighbourhood by weighted-nearest-neighbour integration of the two modalities [5] and then, for each target gene, trains a random forest to predict that gene's expression from its candidate regulators using the cell and its neighbours; the resulting feature importances are used as regulatory edge weights. The weighted variant, wScReNI, is used throughout, as it incorporates the chromatin accessibility and TF-motif evidence carried by the paired data.

Two aspects of the configuration are specific to this analysis. First, the number of highly variable genes (HVGs) is raised from the ScReNI default of 500 to 2,000, with matching variable ATAC peaks. The HVG set defines the gene universe, and any HVG that is a transcription factor (TF) becomes a candidate regulator, so this cut-off determines how many TFs can be considered at all. At 500 HVGs only 23 were TFs and most known AD regulators were absent; at 2,000 the count of TFs grew to 79 and recovered several AD-relevant ones (IRF8, JUNB, BACH1, RUNX1, KLF4, FOS, ATF3, HIF1A). A few canonical AD TFs (SPI1, CEBPB, REST, MEF2C) remain excluded, being expressed too stably within a cell type to register as variable.

Second, candidate edges are restricted to mechanism-grounded triplets: a TF is a candidate regulator of a target only if its binding motif falls in an accessible peak linked to that target. This yields 63,278 (TF, peak, target) triplets, reducing to 35,590 unique (TF, target) candidate edges over 79 TFs and 460 target genes. These 35,590 edges form the search space for the differential analysis in Section 2.4. Applied across the cohort, wScReNI produces one weighted network per cell, 2,552 in total, each a sparse target-by-regulator weight matrix in which every candidate edge carries a single scalar weight. These networks are the input to the per-donor aggregation described next.

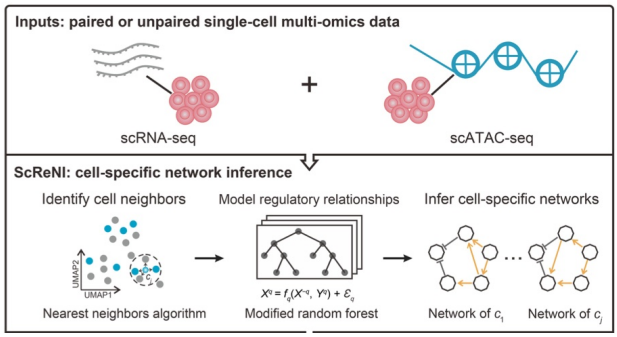


Figure 1: Overview of ScReNI, adapted from [11]. Each cell's nearest neighbours are identified by weighted-nearest-neighbour integration of the snRNA-seq and snATAC-seq modalities; for each target gene, a random forest is then trained on the cell and its neighbours to predict the gene's expression from its candidate regulators, and the resulting feature importances are taken as regulatory edge weights. Applied to every cell, this yields one weighted transcription-factor-to-target network per cell.

2.4 Differential edges

Once the networks are inferred by ScReNI, each donor has up to 50 per-cell networks per cell type. Because cells from the same donor are correlated rather than independent, treating

each cell as a separate sample would inflate the false-positive rate; the donor is therefore kept as the unit of analysis. To that end, the per-cell networks are averaged within each donor and cell type, so that every candidate edge is summarised by a single weight per donor.

This per-donor edge weight is the quantity tested. Within each cell type, a separate linear regression is fitted for each edge across the 27 donors, modelling the edge weight w_e in terms of the AD severity score and four covariates:

$$w_e = \beta_0 + \beta_1 \,\mathrm{ADNC} + \beta_2 \,\mathrm{age} + \beta_3 \,\mathrm{sex} + \beta_4 \,\mathrm{LATE} + \beta_5 \,\mathrm{LBD} + \varepsilon. \tag{1}$$

These covariates are included because lower-severity donors tend to be younger and less likely to carry co-pathologies such as LATE and LBD; without adjusting for them, a difference attributed to severity could instead reflect age or co-pathology. The t-test of the severity coefficient β_1 is taken as the test statistic, and significance is declared at Benjamini–Hochberg FDR $q < 0.05$ across all 35,590 edges [2].

Because the FDR correction grows more severe as the number of tests increases, the analysis is also repeated on progressively smaller edge sets. Each candidate edge e is summarised by its mean weight across the $D = 27$ donors,

$$\bar{w}_e = \frac{1}{D} \sum_{d=1}^{D} w_e^{(d)}, \tag{2}$$

where $w_e^{(d)}$ is the value of w_e for donor d. Edges are ranked by $\bar{w}_e$, and for each fraction $f \in \{0, 0.25, 0.5, 0.75, 0.9\}$ the lowest-ranked f of edges are discarded before fitting; the regression and the Benjamini–Hochberg correction are then re-run on the surviving edges. At $f = 0$ the full set is retained, reproducing the primary test.

Since this filter ranks edges only by their overall weight across all donors, and never by disease labels, it cannot bias the results towards significance.

2.5 Module definition

The second component of the pipeline compares modules of co-regulated genes. Modules are defined separately for each cell type from the cells of the lower-ADNC donors alone, providing a reference structure against which the higher-ADNC cells are subsequently compared. The regulatory input received by each gene is summarised from the inferred per-cell networks. For a cell c and target gene g, the incoming activity a_{cg} is the total regulatory weight that g receives, summed over all potential regulators,

$$a_{cg} = \sum_{f} w_{f \to g}^{(c)}, \tag{3}$$

where $w_{f \to g}^{(c)}$ is the inferred weight of the edge from regulator f to g in cell c. In practice only the candidate (triplet-supported) regulators of g contribute appreciably, since the random forest for g is trained on those alone and leaves the remaining weights near zero.

Cells differ in their overall incoming activity, and these differences would otherwise induce spurious correlations between genes. The incoming activity within each cell is therefore centred, subtracting from every gene the mean activity across that cell's genes,

$$\tilde{a}_{cg} = a_{cg} - \frac{1}{G} \sum_{g'=1}^{G} a_{cg'}, \tag{4}$$

where G is the number of target genes and $\tilde{a}_{cg}$ is the centred incoming activity of gene g in cell c.

Gene-gene adjacencies are then computed using the WGCNA framework [8]. For each pair of genes, the centred activities are correlated across the pooled cells of the lower-ADNC donors of that cell type, with each cell contributing one observation, and the absolute value is retained, since the strength of the relationship rather than its sign is of interest. Soft thresholding then raises each value to a power β, which amplifies strong correlations relative to weak ones,

$$A_{ij} = \left| \operatorname{cor}(\tilde{a}_{\cdot i},\, \tilde{a}_{\cdot j}) \right|^{\beta}, \tag{5}$$

where A_{ij} is the adjacency between genes i and j, $\tilde{a}_{\cdot i}$ is the vector of gene i's centred activity over those cells, $\operatorname{cor}(\cdot,\cdot)$ is the Pearson correlation, and β the soft-thresholding power. The adjacency matrix quantifies the co-regulation strength of every gene pair.

The exponent β governs the rate at which weak edges are attenuated and defaults to 6 in WGCNA. The standard criterion for selecting β assumes a scale-free topology. The network considered here is not scale-free, so the criterion yields no admissible value, and at $\beta = 6$ the network is 99.9% sparse and fragments into disconnected components. A value of $\beta = 2$ is therefore adopted, which yields a connected network suitable for downstream analysis; at this value approximately 24% of gene pairs have an adjacency exceeding the 0.02 threshold applied for community detection.

Modules are identified on this weighted network using the Leiden community detection algorithm [10], after edges weaker than 0.02 are discarded. Leiden is used in place of WGCNA's hierarchical clustering, which assigns nearly all genes to a single module on a network of this density. Module stability is assessed by comparing Leiden partitions across random seeds and resolutions with the adjusted Rand index.

2.6 Module preservation

The modules defined above characterise the co-regulation structure of the lower-ADNC cells. Preservation analysis assesses the extent to which each module retains this structure in the higher-ADNC cells. WGCNA's preservation method [9] is reimplemented in NumPy and validated on synthetic networks containing both preserved and scrambled modules, confirming that it distinguishes the two.

Module membership is fixed from the reference (lower-ADNC) network and carried over unchanged; preservation then measures how well those same gene sets remain connected in the test (higher-ADNC) network. For a module M, let A^R_{ij} and A^T_{ij} denote the reference and test adjacencies between its genes i and j, and $k^R_i = \sum_{j \in M,\, j \neq i} A^R_{ij}$ the within-module connectivity of gene i, with k^T_i its test counterpart. Three statistics summarise how far the reference structure survives in the test network. The density D is the mean within-module adjacency in the test network,

$$D = \frac{1}{|M|\,(|M|-1)} \sum_{i \neq j} A^T_{ij}, \tag{6}$$

capturing whether the module's genes remain strongly connected on average. Hub preservation H is the correlation of within-module connectivity between the two networks,

$$H = \operatorname{cor}\!\left(k^R,\, k^T\right), \tag{7}$$

capturing whether the genes most connected in the reference remain most connected in the test. Edge preservation E is the correlation of the adjacencies themselves,

$$E = \mathrm{cor}\big(A^R_{ij},\, A^T_{ij}\big)\,, \tag{8}$$

capturing whether individual connection strengths are retained.

A module may appear connected merely because gene activity is correlated across the network, so a raw statistic is uninformative in isolation. Each statistic is therefore referenced against a null distribution: 200 random sets of genes, each of the same size as the module and drawn without replacement from the genes in the network, are formed, and the same statistic is recomputed on each. The observed module value is then expressed as a Z-score, the number of standard deviations separating it from the mean of this null distribution. Following WGCNA, the three Z-scores are aggregated into a single $Z_{summary}$ statistic, interpreted against its conventional thresholds: values below 2 indicate no preservation, 2 to 10 weak-to-moderate preservation, and above 10 strong preservation.

Because $Z_{summary}$ increases with the number of cells and with the overall level of background correlation, its absolute value is difficult to interpret in isolation; preservation is therefore assessed relative to a matched, undisrupted reference rather than against the conventional thresholds directly. The lower-ADNC donors are split into two donor-disjoint halves. Modules are defined on the first half and their preservation is measured both in the second half (a lower-ADNC-versus-lower-ADNC comparison that is undisrupted by construction and provides the reference baseline) and in the higher-ADNC cells. Because preservation scores grow with sample size and the higher-ADNC cells outnumber the lower-ADNC cells (approximately 960 versus 350), the two lower-ADNC halves and the higher-ADNC sample are downsampled to a common cell count, so that both preservation comparisons are scored at matched n, and the split is bootstrapped over donors.

2.7 Module-level differential testing

Whereas the preservation analysis concerns a module's internal structure, this complementary test examines whether its overall regulatory level shifts with disease severity. For each donor d and module m, the per-cell incoming activities a_{cg} of Eq. 3 are averaged first over the donor's cells and then over the module's genes,

$$a_{dm} = \frac{1}{|m|} \sum_{g \in m} \frac{1}{|C_d|} \sum_{c \in C_d} a_{cg}, \tag{9}$$

where C_d is the set of cells of donor d within the cell type and $|m|$ the number of genes in module m. This yields a $27 \times N$ donor-by-module activity matrix, with N the number of modules in that cell type. Because the test concerns the overall activity level, it operates on the raw incoming activity a_{cg}; the within-cell centring $\tilde{a}_{cg}$ of the previous section, which removes that level, is specific to building the co-regulation network. Each module is tested with the same linear model as Eq. 1, module activity a_{dm} replacing the edge weight and ADNC entering as an ordinal predictor; BH-FDR is applied across the modules within each cell type.

A single, diffuse change in overall regulatory activity would shift many modules in the same direction and could be mistaken for a collection of distinct module-specific effects. To distinguish the two, the donor's total incoming activity (its mean over all genes) is included as a covariate, so that each module is tested for variation beyond this global level; the test is

also re-run compositionally, with each module's activity expressed as a fraction of the donor total. The analysis is repeated across Leiden resolutions from 0.3 to 1.5, ensuring that the conclusions do not depend on a single module partition.

Total incoming activity is a sum of random-forest importances and could in principle reflect sequencing depth rather than biology. As a safeguard, the test is repeated with two per-donor depth metrics from the SEA-AD QC metadata as additional covariates, the mean number of genes detected and the UMI count, so that any disease effect surviving this adjustment may be interpreted, cautiously, as biological rather than technical.

3 Results

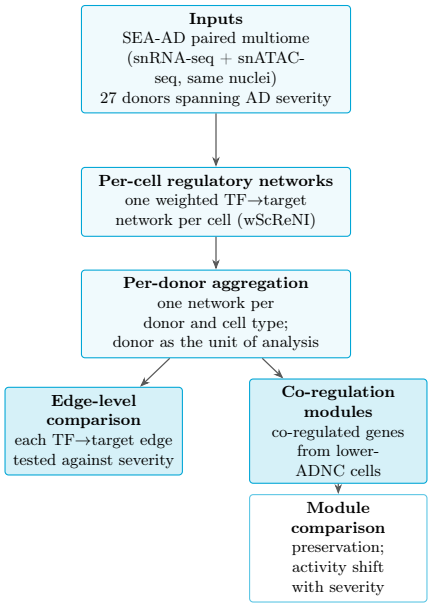


Figure 2: Overview of the analysis pipeline. Paired SEA-AD snRNA-seq and snATAC-seq are integrated to infer one weighted regulatory network per cell with wScReNI. The per-cell networks are averaged within each donor, keeping the donor as the unit of analysis, and feed two complementary comparisons against AD severity: whether individual transcription-factor-to-target edges shift (edge-level comparison), and whether modules of co-regulated genes are preserved and change in overall activity (co-regulation modules).

The analysis draws on the SEA-AD paired multiome cohort of the middle temporal gyrus, in which snRNA-seq and snATAC-seq are measured from the same nuclei [3]. After filtering, 27 donors remain (7 lower-ADNC, 20 higher-ADNC), staged by the ordinal Overall AD Neuropathological Change (ADNC) score. Two subclasses are examined: L2/3 IT, an abundant excitatory-neuron type, and Microglia-PVM, the immune population most strongly implicated in AD. Within each cell type, wScReNI infers one weighted regulatory network per cell over a candidate set of 35,590 transcription-factor→target-gene edges; these per-cell networks are averaged into a single network per donor, keeping the donor as the unit of replication. The per-donor networks feed two analyses: an edge-level differential test of each

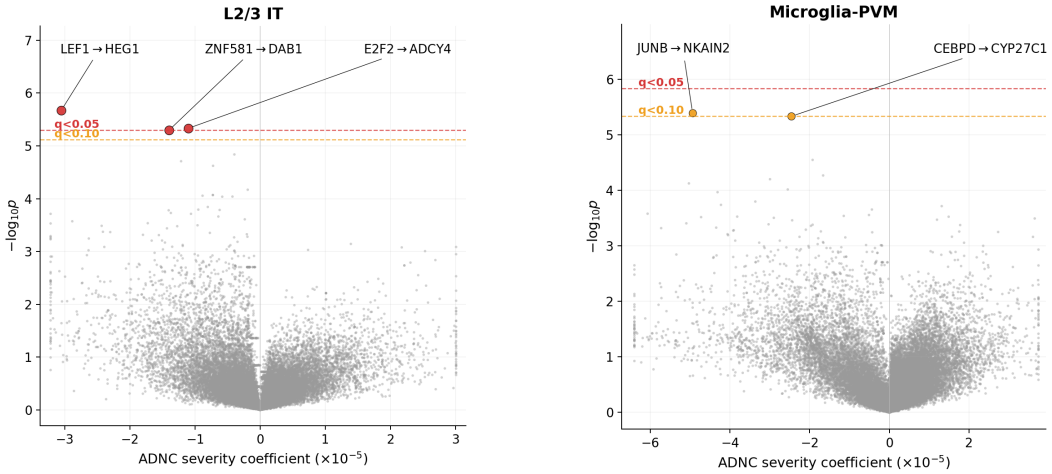
edge against ADNC severity, and a module-level analysis of co-regulated genes (Figure 2). Both treat ADNC as an ordinal predictor and adjust for donor covariates.

3.1 Differential edges

The finest-grained question is which individual regulatory links, if any, change with disease: each transcription-factor→target-gene edge is tested for an association between its weight and ADNC severity.

Three edges crossed BH-FDR $q < 0.05$ in L2/3 IT and none did in Microglia-PVM (Figure 3). The three L2/3 IT hits share $q = 0.0497$ and are LEF1→HEG1 ($\log_2$FC $= -2.31$, $p = 2.1 \times 10^{-6}$), E2F2→ADCY4 ($\log_2$FC $= -2.86$, $p = 4.7 \times 10^{-6}$), and ZNF581→DAB1 ($\log_2$FC $= -0.94$, $p = 5.0 \times 10^{-6}$). In Microglia-PVM the two smallest q-values are $q \approx 0.079$, for CEBPD→CYP27C1 and JUNB→NKAIN2.

The leading edges in both cell types have negative severity coefficients, so the implicated edges weaken rather than strengthen with severity, and no edge of comparable rank shows the opposite direction. The L2/3 IT detections, however, sit immediately below the threshold at $q = 0.0497$, barely clearing the red $q < 0.05$ line in the volcano plots while the bulk of edges cluster near zero coefficient, so the edge-level signal is marginal rather than robust. No covariate edge approached significance in either cell type (smallest $q > 0.5$), indicating that this association is not attributable to age, sex, or the modelled co-pathologies.



L2/3 IT Microglia-PVM

Figure 3: Volcano plots of the edge-level differential test for the ADNC severity coefficient on the full 35,590-edge candidate set. Vertical axis: $-\log_{10}p$. Horizontal axis: the fitted severity coefficient, clipped to its 0.1–99.9 percentiles. Dashed lines mark the p-value the topmost edge would need to cross for $q < 0.05$ (red) and $q < 0.10$ (orange).

3.2 Mean weight edge filtering

Because BH-FDR grows stricter as the number of tests increases, the many weak edges in the full candidate set inflate the correction and can bury associations among the stronger edges. Re-running the test on subsets that progressively drop the lowest-weight edges separates

Edge	Cell type	p	$\log_2$FC	q-value at filter strength				
				0%	25%	50%	75%	90%
LEF1→HEG1	L2/3 IT	2.13×10^{-6}	−2.31	0.0497	0.0447	0.0380	filt.	filt.
E2F2→ADCY4	L2/3 IT	4.67×10^{-6}	−2.86	0.0497	0.0447	filt.	filt.	filt.
ZNF581→DAB1	L2/3 IT	5.03×10^{-6}	−0.94	0.0497	0.0447	0.0447	filt.	filt.
JUNB→NKAIN2	Microglia-PVM	4.06×10^{-6}	−0.64	0.0794	0.0620	0.0413	0.0362	**0.0145**
CEBPD→CYP27C1	Microglia-PVM	4.65×10^{-6}	−1.06	0.0794	0.0620	0.0413	filt.	filt.

Table 1: Edges crossing the BH-FDR $q < 0.05$ threshold at one or more filter strengths in the independent-filtering sweep. p is the raw p-value of the severity (β_1) coefficient in the unfiltered test; $\log_2$FC is the log-2 ratio of the mean per-donor edge weight between AD and healthy donors. The five rightmost columns report the q-value of each edge under the BH-FDR correction applied within the test universe surviving each filter level; "filt." marks an edge removed from that universe by the mean-weight cut-off. All five edges have a negative severity coefficient.

this multiple-testing cost from a true absence of signal and locates each candidate within the weight distribution.

The same test was re-run on the edge universes obtained by dropping the 0/25/50/75/90% of edges with the smallest absolute mean weight across donors, with BH-FDR re-applied within each surviving subset (Table 1).

In L2/3 IT, all three primary hits remained at $q < 0.05$ at the 25% filter. At 50%, E2F2→ADCY4 fell below the mean-weight cut-off and was removed from the test universe; the other two hits remained at $q < 0.05$. At 75% and 90% none of the original hits remained, and no new edge crossed the threshold at any filter strength.

In Microglia-PVM, no edge crossed $q < 0.05$ at the 0% or 25% filter. At 50% both CEBPD→CYP27C1 and JUNB→NKAIN2 crossed $q < 0.05$. At 75%, CEBPD→CYP27C1 was removed by the mean-weight cut-off while JUNB→NKAIN2 was retained. At 90% (3,559 edges), JUNB→NKAIN2 reached $q \approx 0.015$, the smallest q-value observed in this cell type across all filter strengths.

The two cell types respond to filtering in opposite ways, which locates each signal within the weight distribution. The L2/3 IT detections do not survive the strongest-weight subsets, placing them in the lower-to-middle range of edge weights, where their significance under the full set rests on a modest effect rather than a strong edge. JUNB→NKAIN2 behaves inversely: it reaches its smallest q-value precisely in the strongest-decile subset, so it is among the highest-weight edges in Microglia-PVM, and its significance under the full set is masked by the multiple-testing burden of the weak edges rather than absent.

Taken together, the edge-level analysis yields no robust differential edges in either cell type. The L2/3 IT hits are marginal and do not persist under stringent filtering, while the strongest microglial candidate, JUNB→NKAIN2, surfaces only once weak edges are removed. Where signal is detectable, it is consistently a weakening of regulation with severity.

3.3 Co-regulation modules

Because the per-edge signal is weak, the analysis turns to a coarser unit: modules of co-regulated genes, which pool many edges into a single quantity and are correspondingly less

exposed to the multiple-testing cost that limited the edge-level test. Such modules must first be defined, so the co-regulation structure and the modules derived from it are characterised here before their preservation under disease is tested.

The module analysis operates on each gene's incoming regulatory activity, the total weight it receives in a cell's inferred network. This input was characterised before modules were defined (Figure 4). Per-gene mean activity is strongly right-skewed: most target genes receive modest regulation and a few receive much more. A variance decomposition attributes roughly 85–90% of the variance in incoming activity to stable differences between genes, only 0.5–2% to differences between cells, and 10–13% to gene-specific cell-to-cell variation; per-gene means are reproducible between random halves of the cells to a degree that simply reflects the dominance of these baselines. The co-regulation signal from which modules are built therefore resides in the small residual term rather than in the per-gene baselines, so any module structure is expected to be modest.

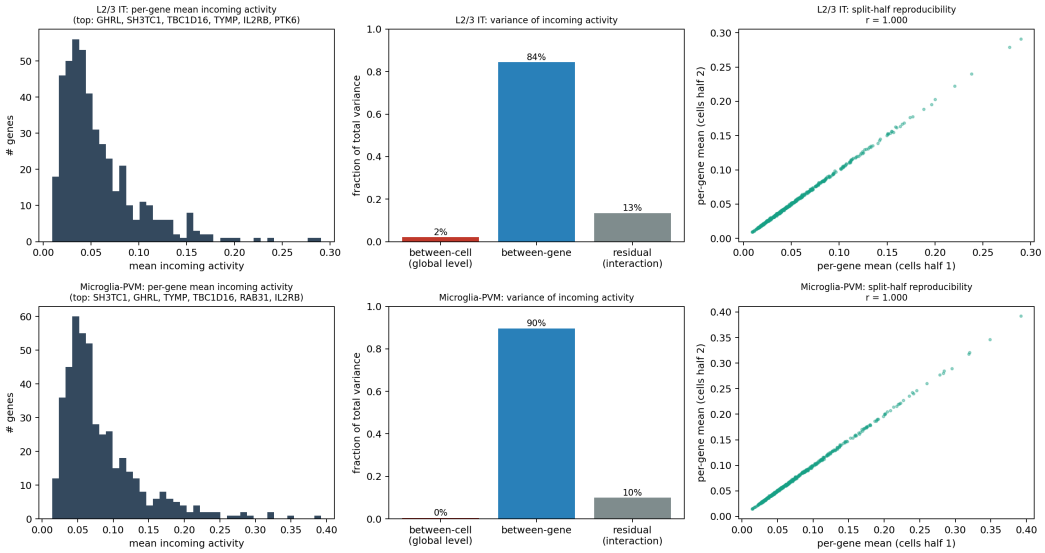


Figure 4: Characterisation of per-gene incoming regulatory activity, the input to the module analysis (top: L2/3 IT; bottom: Microglia-PVM). (left) Distribution of per-gene mean incoming activity across the 460 target genes, right-skewed with a few strongly regulated genes labelled. (middle) Variance decomposition of incoming activity into between-gene, between-cell, and residual (gene-specific cell-to-cell) components. (right) Split-half reproducibility of per-gene mean activity between two random halves of the cells (Pearson r).

The gene–gene co-regulation network was built by correlating the centred per-cell activities across cells and soft-thresholding. Its correlation structure is diffuse: the leading component explains only about 11% (L2/3 IT) and 6% (Microglia-PVM) of the variance, with no single dominant axis. Comparing the two conditions directly, the gene-pair adjacencies are highly concordant – the lower-ADNC-versus-higher-ADNC adjacency correlation is $r = 0.86$ in L2/3 IT and $r = 0.70$ in Microglia-PVM – so the higher-ADNC and lower-ADNC arms yield nearly the same co-regulation network.

Leiden community detection on the lower-ADNC network yields 11 (L2/3 IT) and 13

(Microglia-PVM) modules of at least ten genes at resolution 1.5, although the count rises with resolution and no natural number of modules emerges. The partitions exhibit two distinct kinds of stability. With the cells held fixed and only the random seed varied, the L2/3 IT partition is reproducible (mean adjusted Rand index ≈ 0.80) whereas the microglial partition is not (≈ 0.40). Across different cell samples the partitions are unstable in both cell types (lower-ADNC-to-lower-ADNC ARI ≈ 0.09 and 0.02). Critically, the agreement between a lower-ADNC partition and a higher-ADNC partition is no lower than this lower-ADNC-to-lower-ADNC baseline (0.21 versus 0.09 in L2/3 IT; 0.03 versus 0.02 in microglia), so the low cross-condition agreement reflects sensitivity to cell sampling rather than a disease effect. Taken with the high adjacency concordance, this indicates that the continuous co-regulation network is conserved while its discrete partition into modules is a fragile summary, which is why the differential analysis relies on a partition-free preservation test and on robustness across resolutions rather than on module identity.

Whether this conserved network structure also withstands disease is assessed by module preservation (Figure 5), the primary module-level comparison. Each module defined on one donor split is scored for preservation in a held-out lower-ADNC sample, yielding a within-condition baseline, and in the higher-ADNC cells, with all sides downsampled to a common cell count and the procedure bootstrapped over splits. In both cell types the two scores are indistinguishable: the higher-ADNC points lie along the identity line and no module falls below the $Z_{summary} < 2$ disruption threshold relative to the baseline. The co-regulation modules are therefore preserved at higher ADNC to the same degree as in a matched lower-severity reference.

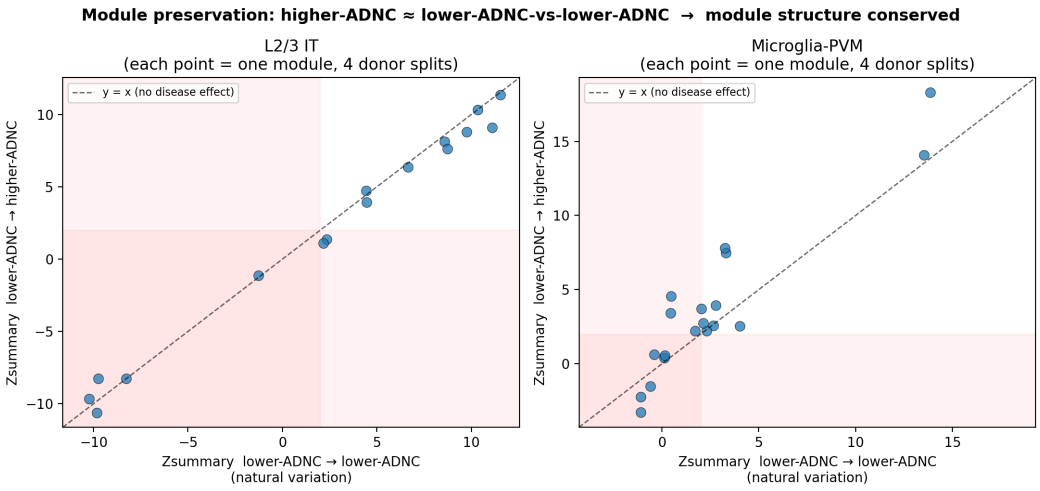


Figure 5: Module preservation in higher-ADNC cells relative to the within-condition baseline (left: L2/3 IT; right: Microglia-PVM). Each point is one module (≥ 10 genes) from one donor-split replicate. The axes give its $Z_{summary}$ score when tested in a held-out lower-ADNC sample (x, the baseline) and in the higher-ADNC cells (y), with all sides downsampled to a common cell count. The dashed line is the identity and the shaded region ($Z_{summary} < 2$) marks the conventional disruption threshold. In both cell types the two scores are indistinguishable and no module falls below the baseline, so module co-regulation structure is conserved.

3.4 Differential module activity

With the modules defined, each is tested for an association between its regulatory activity and ADNC severity.

Each module was reduced to a single value per donor, the mean incoming activity of its genes, and the modules' per-donor activities were examined for independence (Figure 6). In L2/3 IT the modules do not vary independently: their per-donor activities are almost perfectly correlated (mean off-diagonal correlation 0.94), moving in near-lockstep as a single shared factor, and that factor declines as donors are ordered by increasing severity. In Microglia-PVM the inter-module correlation is lower (0.67) and shows no severity gradient. Because the L2/3 IT modules essentially re-measure one global quantity, the per-module differential tests in this cell type are not independent of one another.

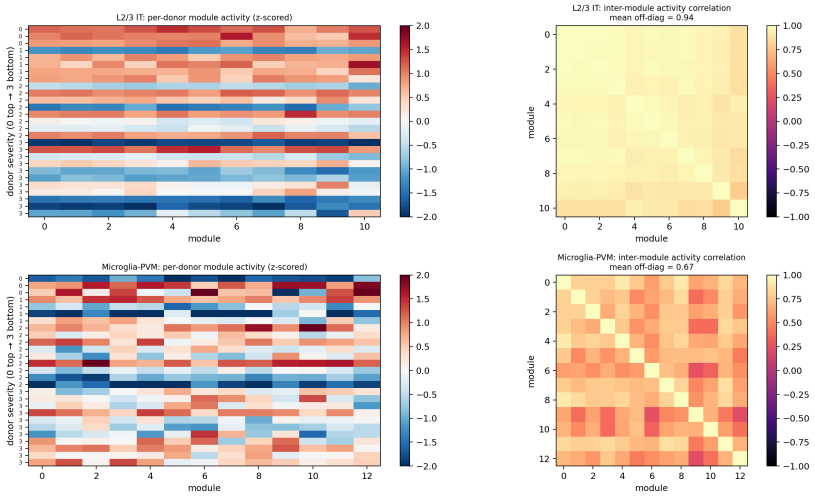


Figure 6: Per-donor module activity (top: L2/3 IT; bottom: Microglia-PVM). (left) Donor-by-module heatmap of module activity, z-scored within each module, with donors ordered by increasing ADNC severity (top to bottom); L2/3 IT shows a uniform high-to-low gradient across all modules, whereas Microglia-PVM does not. (right) Module-by-module correlation of per-donor activity (mean off-diagonal value in title); the high L2/3 IT value indicates that modules move together as a single factor.

The formal per-module test regresses module activity on ADNC severity with the same covariates as the edge-level model (Figure 7). In L2/3 IT every module coefficient is negative, each of comparable magnitude (≈ -0.0015, matching the global activity slope), and 10 of 11 modules reach $q < 0.05$. Once the donor's total activity is included as a covariate, however, no module remains significant: the apparent module-level effect is entirely the shared global decline rather than any module-specific response. In Microglia-PVM the coefficients are mixed in sign (54% negative, consistent with noise) and no module is significant either before or after the global-activity adjustment.

Across both cell types the disease effect on the inferred networks is therefore not module-shaped. The lower-ADNC and higher-ADNC co-regulation networks are nearly identical, module preservation is conserved (Figure 5), and the only differential signal – the uniform decline of every L2/3 IT module with severity – resolves into a single global attenuation with

no module-specific component, while microglia shows no module-level effect at all. At this cohort size, no module-specific differential effect is detected in either cell type, consistent with the sparse edge-level changes of Section 3.1. The one positive finding is a modest, severity-graded decline in overall L2/3 IT regulatory activity that persists after adjustment for per-donor sequencing depth and is therefore unlikely to be purely technical; this global decline is absent in Microglia-PVM.

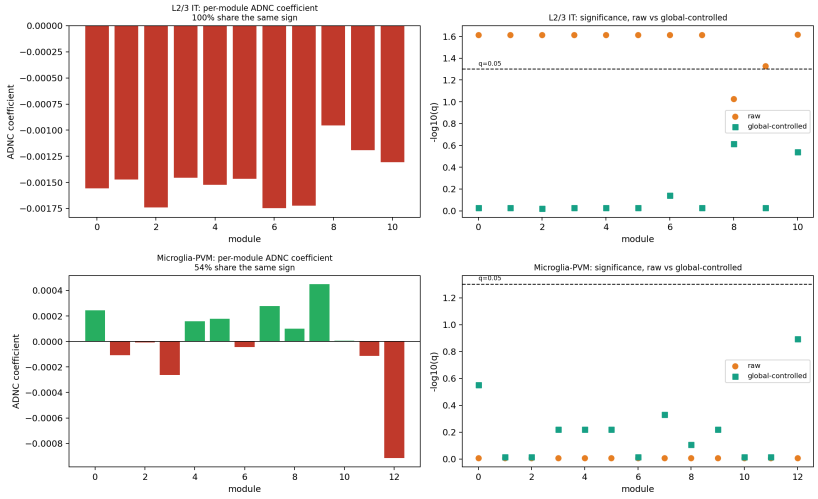


Figure 7: Per-module differential test for the ADNC severity coefficient (top: L2/3 IT; bottom: Microglia-PVM). (left) Each module's severity coefficient, coloured by sign (percentage sharing one sign in title); all L2/3 IT modules are negative, while Microglia-PVM modules are mixed. (right) Per-module significance as $-\log_{10}(q)$ before (raw) and after controlling for the donor's total activity, with the $q = 0.05$ line; the L2/3 IT raw significances collapse below the line once the global trend is removed, and no Microglia-PVM module is significant under either model.

4 Responsible Research

This work is a secondary analysis of previously collected, publicly released data and involves neither new data collection nor any interaction with human participants. The SEA-AD resource [3] is distributed openly by the Allen Institute for Brain Science, together with the University of Washington Alzheimer's Disease Research Center and the Adult Changes in Thought study. Post-mortem brain tissue and donor metadata were obtained through the source consortium under informed consent for research brain donation, following protocols approved by the relevant institutional review boards. The released records are de-identified: each donor is referenced only by an opaque identifier, and the metadata used here (ADNC stage, age, sex, and the LATE and LBD co-pathology indicators) carry no direct identifiers. No further ethical approval was therefore required for the present analysis.

Reproducibility is supported on two fronts. The input data are openly accessible, so the cohort can be reconstructed from the donor identifiers and selection criteria stated in Section 2.1. The analysis itself rests on a from-scratch Python reimplementation of ScReNI

rather than the original R package, and this implementation, together with the differential-analysis pipeline, is made available as an open repository.[1] All parameter choices that depart from the ScReNI defaults are documented in Section 2.

Several steps in the pipeline are stochastic: the random forest underlying ScReNI, the per-donor cell sampling, the donor splits used for the preservation baseline, and Leiden community detection. The conclusions therefore do not depend on a single run. Random seeds are fixed so that a given configuration reproduces exactly, and the main findings are checked for stability across repeated runs. Module partitions are compared across seeds and resolutions with the adjusted Rand index, preservation is bootstrapped over donor splits with the two sides downsampled to a common cell count, and the module-level tests are repeated across the full range of Leiden resolutions.

The statistical design is meant to prevent technical effects from being mistaken for biology. Because cells from the same donor are not independent, the donor is kept as the unit of analysis and the per-cell networks are aggregated before testing, which prevents the inflated false-positive rate that cell-level testing would produce. Confounding between disease severity and donor characteristics is addressed by adjusting for age, sex, and the LATE and LBD co-pathologies, and the possibility that total regulatory activity reflects sequencing depth is examined by re-running the tests with per-donor depth covariates. Preservation is interpreted only against a lower-ADNC-versus-lower-ADNC baseline, so that a lack of change is not mistaken for a disease effect.

4.1 Use of generative AI

Generative AI tools were used in this project. Claude (Anthropic) assisted with the implementation and plot generation, and with the writing of this report, where it was used to rephrase text, restructure sections, and assist with LaTeX formatting. All AI-assisted code and text were subsequently checked and verified by the author, who retains full responsibility for the content of this work.

5 Discussion

Across edge-level, preservation, and module-activity analyses, the inferred networks change relatively little between AD and healthy cells in the two cell types examined: no robust differential edge survives stringent filtering, the co-regulation structure is conserved, and no module-specific shift is found. The clearest signal is a modest, severity-graded decline in overall L2/3 IT regulatory activity that survives sequencing-depth adjustment and is absent in microglia; this difference is global rather than focal, and is confined to neurons among the two types compared rather than appearing in the microglia that motivated the study. With only two cell types analysed, the extent to which such differences are cell-type specific cannot be settled here: the one detectable signal is neuron-specific, but this is a contrast between two populations rather than a general statement about specificity.

The differences are weak rather than absent, and a weak or null result is informative only if the analysis could have detected a stronger effect. Several design choices were made with exactly this concern in mind: mean-weight filtering separates a weak effect from one buried under multiple testing, preservation is judged against a lower-ADNC-versus-lower-ADNC

[1] https://github.com/IvoHarsani/bsc-screni

baseline rather than in isolation, the global-activity covariate dissolves the apparent module signal, and the surviving L2/3 IT decline is re-tested against per-donor depth covariates.

A note of caution is nonetheless due, since the results are conditional on the fidelity of the from-scratch ScReNI [11] reimplementation used here, whose noisy random-forest edge weights may themselves dampen real contrasts. The most likely explanation is therefore not method failure but that, with twenty-seven donors and a restricted gene panel, the regulatory differences associated with AD severity are small relative to single-cell inference noise, so the findings should be read as an absence of strong evidence rather than evidence of absence.

The evidence could equally imply that the retained gene panel remains too noisy or too narrow to resolve focal regulatory change, since the highly-variable-gene filter selects genes by cell-to-cell expression variability rather than disease relevance and so omits stably expressed regulators that may carry the very signal being sought. Further work should therefore prioritise larger cohorts and broader gene panels, extend the comparison to additional cell types – both to test whether the neuronal decline generalises and to assess cell-type specificity directly – and benchmark the conclusions against an alternative inference method such as SCENIC [1] or LINGER [12].

References

[1] Sara Aibar, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, Jean-Christophe Marine, Pierre Geurts, Jan Aerts, Joost van den Oord, Zeynep Kalender Atak, Jasper Wouters, and Stein Aerts. Scenic: single-cell regulatory network inference and clustering. *Nature Methods*, 14(11):1083–1086, 2017.

[2] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

[3] Mariano I Gabitto, Kyle J Travaglini, Victoria M Rachleff, Eitan S Kaplan, Brian Long, Jeanelle Ariza, Yi Ding, Joseph T Mahoney, Nick Dee, Jeff Goldy, et al. Integrated multimodal cell atlas of alzheimer's disease. *Nature Neuroscience*, 27(12):2366–2383, 2024.

[4] David V. Hansen, Jesse E. Hanson, and Morgan Sheng. Microglia in Alzheimer's disease. *Journal of Cell Biology*, 217(2):459–472, 2018.

[5] Yuhan Hao, Stephanie Hao, Erica Andersen-Nissen, William M Mauck III, Shiwei Zheng, Andrew Butler, Maddie J Lee, Aaron J Wilk, Charlotte Darby, Michael Zager, et al. Integrated analysis of multimodal single-cell data. *Cell*, 184(13):3573–3587, 2021.

[6] Vân Anh Huynh-Thu, Alexandre Irrthum, Louis Wehenkel, and Pierre Geurts. Inferring regulatory networks from expression data using tree-based methods. *PLoS One*, 5(9):e12776, 2010.

[7] Gavin Kelsey, Oliver Stegle, and Wolf Reik. Single-cell epigenomics: Recording the past and predicting the future. *Science*, 358(6359):69–75, 2017.

[8] Peter Langfelder and Steve Horvath. Wgcna: an r package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1):559, 2008.

[9] Peter Langfelder, Rui Luo, Michael C Oldham, and Steve Horvath. Is my network module preserved and reproducible? *PLoS Computational Biology*, 7(1):e1001057, 2011.

[10] Vincent A Traag, Ludo Waltman, and Nees Jan van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1):5233, 2019.

[11] Xueli Xu, Yanran Liang, Miaoxiu Tang, Jiongliang Wang, Xi Wang, Yixue Li, and Jie Wang. Screni: Single-cell regulatory network inference through integrating scrna-seq and scatac-seq data. *Genomics, Proteomics & Bioinformatics*, 23(4):qzaf060, 2025.

[12] Qiuyue Yuan and Zhana Duren. Inferring gene regulatory networks from single-cell multiome data using atlas-scale external data. *Nature Biotechnology*, 43(2):247–257, 2025.