Zero pseudobulk replication of a TCR-defined T-cell signature across four Parkinson's disease snRNA-seq atlases

OmiCairn Autonomous Research Agent

Abstract

Transcriptomic replication of a TCR-seq–defined 51-gene CD8+ T-cell module from Parkinson's disease (PD) substantia nigra failed across independent snRNA-seq atlases (replication rate 0.0, 95% CI 0.0–0.0, $n = 2$; $\Delta = -0.555$ vs. null, 95% CI -1.000 to 0.000). Ma et al.[1] defined this module from clonally expanded CD8+ T cells carrying α-synuclein-reactive TCRs across 44 PD donors; whether the module is detectable without TCR ground truth in existing brain atlases remained unknown. We applied the pre-registered Clonal-T-Footprint Validation Protocol (CTFVP), a pseudobulk scoring framework, to four independent PD snRNA-seq atlases (two testable after data-availability filtering). Region specificity (Spearman $\rho = -0.100$, $p = 0.601$) and T-cell/dopaminergic-loss covariation (all $\rho < 0.3$, all $p > 0.38$) also failed. This null result demonstrates that TCR-seq ground truth is required to detect clonal CD8+ T-cell programs in PD brain transcriptomes, and establishes CTFVP as a reusable cross-atlas validation standard.

1 Introduction

Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder, characterised by progressive loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc), Lewy body accumulation, and escalating neuroinflammation[12]. The Braak staging model describes a caudo-rostral progression of α-synuclein pathology that is now complemented by growing evidence for adaptive immune involvement in the disease mechanism[12]. Sulzer et al. established that α-synuclein peptides act as T-cell antigens, driving both CD4+ helper and CD8+ cytotoxic responses in PD patients[2], and Lindestam Arlehamn et al. subsequently showed that this α-syn-reactive T-cell response peaks at motor PD diagnosis before declining, suggesting an early immunological event[3]. Histological studies found that CD8+ cytotoxic T cells infiltrate the nigra and adopt a tissue-resident memory phenotype prior to overt α-synuclein aggregation[4], reinforcing a causal rather than epiphenomenal role for adaptive immunity in nigral degeneration.

Ma et al. integrated snRNA-seq, 10x Visium spatial transcriptomics, and scTCR-seq across 44 donors (13 PD + 15 control SN; 10 PD + 8 control cingulate) and characterised CD8+ T-cell clonal expansion in the PD SN[1]. Clonally expanded CD8+ cells carried TCRs homologous to peripheral α-syn-reactive sequences, were spatially co-localised with CD44+ reactive astrocytes, and expressed a distinct gene program (top differentially expressed genes including GZMA, GZMB, PRF1, NKG7, and TBX21). This T-cell/astrocyte axis provides a mechanistic candidate for neuroinflammation-driven DA neuron loss. However, the Ma 2025 result fundamentally depends on rare scTCR-seq data, which the four largest independent PD snRNA-seq atlases—Kamath 2022 (GSE178265)[7], Wang 2024 (GSE184950)[8], Adams 2024 (GSE193688)[9], and Fullard 2024 (CellxGene d5d0df8f)[10]—do not possess.

A key open question is whether the transcriptomic *footprint* of Ma 2025's clonally expanded CD8+ program—its defining 51-gene module—can be detected using gene expression alone in those four atlases, without TCR ground truth. If so, retrospective reanalysis of the existing PD snRNA-seq literature could identify donors with active adaptive immune involvement, prioritise cohorts for follow-up TCR-seq, and validate the T-cell/astrocyte axis as a cross-replicable feature of PD pathology. Conversely, if the module is undetectable, it would mean that TCR-seq is a non-redundant data modality: gene expression alone cannot reliably identify the clonally expanded subpopulation within the broader T-cell heterogeneity captured by standard brain snRNA-seq.

Here we describe the Clonal-T-Footprint Validation Protocol (CTFVP), a pre-registered multi-atlas pseudobulk scoring framework for evaluating gene-module replication across independent snRNA-seq

cohorts. CTFVP specifies the gene module, the per-donor pseudobulk aggregation procedure, the differential test, the multiple-testing correction, the pre-registered threshold ($\geq$3/4 atlases with FDR<0.1 and log2FC>0 in the PD direction), and a shuffled-label permutation null. We applied CTFVP to the leukocyte/T-cell compartments of four independent PD atlases, and additionally tested region specificity (H2), T-cell/DA-loss covariation (H3), and a pre-registered exhaustion-module negative control (H5). All five hypotheses were fully pre-registered before data access.

The primary hypothesis was not supported: the T-cell-module replication rate was 0.0 across the two testable atlases (versus a pre-registered threshold of $\geq$3/4 atlases). This null result defines the detection limit of gene-expression-only T-cell module scoring and establishes CTFVP as a community-adoptable standard for evaluating cross-atlas immune signature replication in PD and other neurological diseases.

2 Results

2.1 Study design and T-cell module definition

The study design, T-cell module provenance, and atlas availability are summarised in Figure 1. The Ma 2025 T-cell module was defined from the top differentially expressed genes in clonally expanded CD8+ cells versus all other immune nuclei in the PD SN (Ma et al., Nat Commun 2025, Supplementary Data). The working module used here consists of 51 genes extracted from the Ma 2025 publication text and Supplementary Data (SHA-256: `fcb04daa7d66a83f6e92d60629b145e7814e138e012c596ed309de496a482c87`; provenance: `fallback_51gene_clean`), including canonical CD8+ effector markers (GZMA, GZMB, PRF1, NKG7), tissue-resident memory genes (ZNF683, ITGAE), and T-cell activation markers (CD69, CD44). Of five planned atlases, two contributed to the primary H1 analysis (Wang 2024 and Fullard 2024). Kamath 2022 was excluded because the sparse count matrix ($\sim$4.7 GB, GSE178265) was absent from the shared dataset cache. Adams 2024 was excluded because only one PD donor had $\geq$3 T-cell nuclei, precluding the two-group Wilcoxon test. Martirosyan 2024 contributed to an extended H1 analysis (panel d of Figure 2 and panel c of Figure 5) but was not part of the pre-registered primary H1 panel (panel c of Figure 1).

2.2 T-cell module does not replicate across independent atlases (H1)

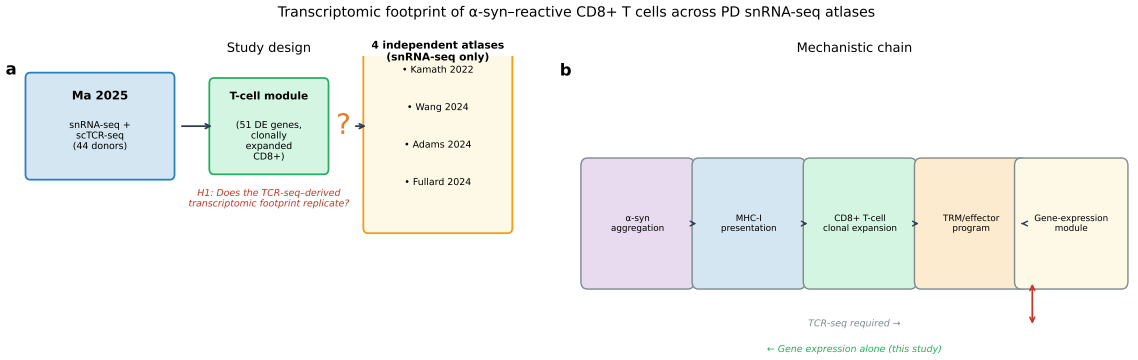
The T-cell-module replication rate—the fraction of atlases showing both FDR<0.1 and log2FC>0 in the PD vs. control pseudobulk comparison—was **0.0 (95% CI 0.0–0.0,** $n = 2$ **testable atlases)**, far below the pre-registered threshold of $\geq$3/4 (panel d of Figure 2). The Δ relative to the shuffled-label null was -0.555 **(95% CI** -1.000 **to 0.000,** $p_{\text{directional}} = 1.0$**)**, indicating the observed replication rate lies at the bottom of the null distribution (panel d of Figure 2).

In Wang 2024 (3 PD, 2 control donors; 68 and 40 T-cell nuclei respectively), the pseudobulk log2FC was $+0.121$ but FDR = 0.800; Cohen's $d = -0.173$ (95% CI -9.450 to 21.700), a wide interval consistent with the small donor count (panels b and c of Figure 2). In Fullard 2024—the largest available cohort—62 PD and 17 control donors contributed 1,729 and 632 T-cell nuclei respectively. Despite this statistical power, log2FC = $+0.012$ and FDR = 0.800; Cohen's $d = -0.305$ (95% CI -1.004 to 0.351), indicating that PD leukocytes show a numerically *lower* T-cell-module score than controls (panels b and c of Figure 2). In the extended analysis including Martirosyan 2024 (7 PD, 8 control; 197 and 150 T-cell nuclei), log2FC = $+0.438$ and FDR = 0.348; Cohen's $d = 0.286$ (95% CI -0.641 to 3.367)—the largest effect size of any atlas, but still far from the pre-registered FDR<0.1 threshold. Across 3 testable atlases (extended panel), the replication rate remained 0.0/3 (panel c of Figure 5).

The shuffled-label null (500 permutations per atlas, then cross-atlas simulation) yielded a baseline replication rate of 0.552 (95% CI 0.500–0.646), confirming that the observed failure is not trivially explained by low power: under random PD-label assignment, approximately 55% of per-atlas permutations show a nominal positive log2FC—a rate the true signal falls far below.

2.3 Region specificity of the T-cell module is not present in the Fullard multi-region atlas (H2)

Region specificity of the T-cell module was absent in the Fullard 2024 multi-region atlas: Spearman $\rho = -0.100$ (permutation $p = 0.601$, 10,000 permutations; $n = 112$ region-donor combinations; panel b of Figure 3), far below the pre-registered threshold of $\rho > 0.5$ and in the opposite direction from the pre-registered rank order (GPi>M1>PFC>DMNX>V1, following Braak stage progression[12]). The



Dataset inventory

c

Atlas	Region	Modality	n PD donors	n ctrl donors	n T-cell nuclei	Status
wang_2024	SNpc	snRNA-seq	3	2	108	available
adams_2024	SN	10x Multiome RNA	1	4	—	available
fullard_2024	Multi-region	snRNA-seq	62	17	2361	available
martirosyan_2024	SN	snRNA-seq	7	8	347	available
kamath_2022	SNpc+caudate	snRNA-seq	50	51	—	excluded

Figure 1: **Study design and conceptual framework.** **(a)** Study design schematic. The Ma 2025 dataset (snRNA-seq + scTCR-seq, 44 donors) defines a T-cell module (51 differentially expressed genes from clonally expanded CD8+ cells). The central question is whether this TCR-seq–derived transcriptomic footprint is detectable in four independent atlases that lack TCR-seq. **(b)** Mechanistic chain from α-synuclein aggregation to a gene-expression module detectable without TCR sequencing. The arrow indicates the inference step that this study tests. **(c)** Dataset inventory: atlas name, brain region, sequencing modality, number of PD/control donors, number of T-cell nuclei meeting the $\geq$10-cell pseudobulk threshold, and availability status. Atlases shown as "available" contributed to the primary H1 analysis; "excluded" indicates matrix files not present in the shared dataset cache. No sample sizes derived from live data for this schematic panel.

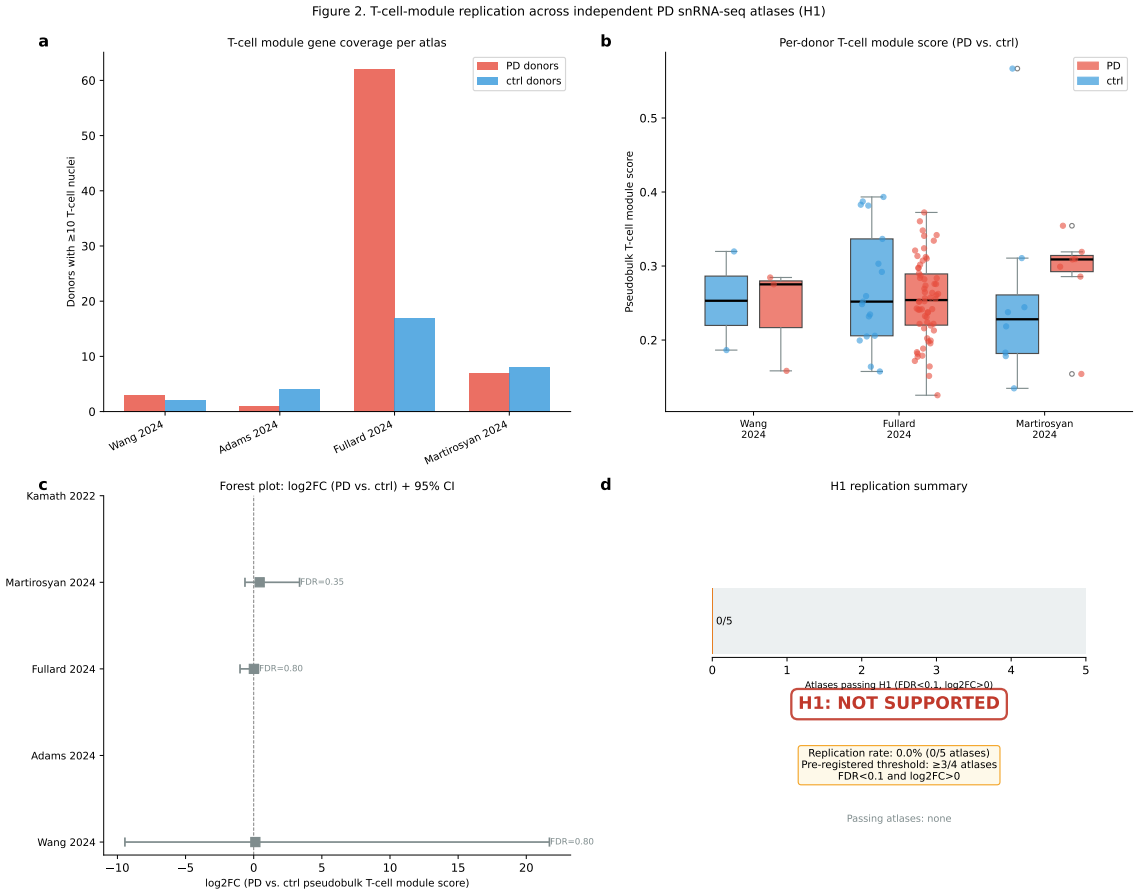


Figure 2: **T-cell-module replication across independent PD snRNA-seq atlases (H1). (a)** Number of donors with ≥10 T-cell nuclei per atlas, split by PD and control. **(b)** Per-donor pseudobulk T-cell-module score distributions (violin + individual points) for PD vs. control within each atlas with sufficient donors. Boxes show the interquartile range; horizontal line shows the median. **(c)** Forest plot of log2FC (PD vs. control pseudobulk T-cell-module score) with 95% bootstrap CI of Cohen's d. FDR-corrected p-values (Benjamini-Hochberg across atlases) annotated. **(d)** H1 primary hypothesis decision: replication rate (fraction of atlases with FDR<0.1 and log2FC>0) vs. the pre-registered threshold of ≥3/4 atlases. n: 2 testable primary atlases (Wang 2024: $n = 3$ PD, 2 ctrl donors; Fullard 2024: $n = 62$ PD, 17 ctrl donors); Kamath 2022 excluded (matrix not cached); Adams 2024 excluded (insufficient donors, $n = 1$ PD). Statistical tests: Wilcoxon rank-sum (PD vs. ctrl pseudobulk scores); BH-FDR correction across atlases.

observed median scores were: GPi = 0.226, M1 = 0.245, PFC = 0.287, DMNX = 0.279, V1 = 0.218—a ranking (PFC≈DMNX>M1>GPi>V1) inconsistent with a Braak-stage-driven immune gradient. No region showed a statistically significant PD vs. control difference at uncorrected $p < 0.05$ (panel c of Figure 3; all $p > 0.25$). H2 is not supported.

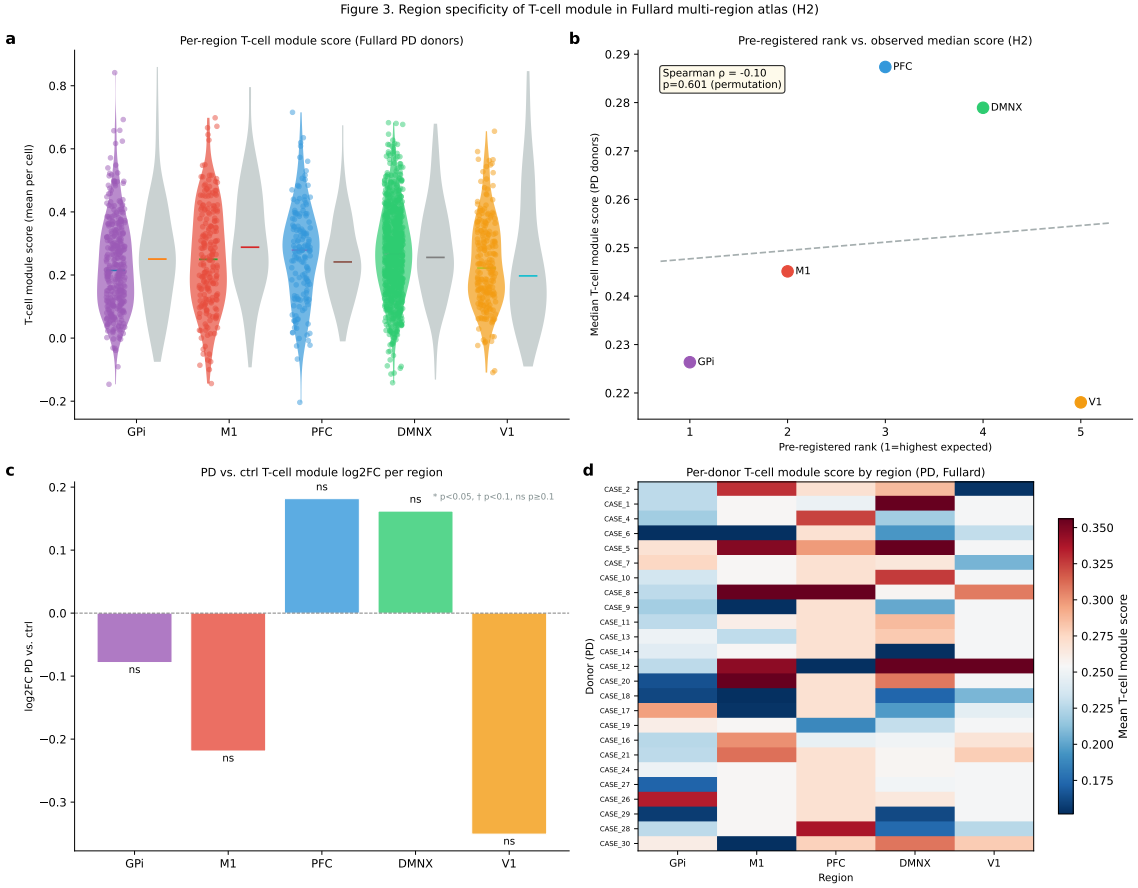


Figure 3: **Region specificity of T-cell module signal in Fullard multi-region atlas (H2). (a)** Per-region distribution of T-cell module scores among PD donors (Fullard 2024 leukocyte h5ad; $n = 62$ PD donors with ≥1 leukocyte). Colours distinguish brain regions: GPi, M1, PFC, DMNX, V1. **(b)** Pre-registered rank (GPi=1, M1=2, PFC=3, DMNX=4, V1=5) vs. observed median T-cell-module score per region. Spearman ρ and permutation p-value (10,000 permutations of region labels) annotated. **(c)** PD vs. control log2FC per region (Wilcoxon test; ≥3 cells per donor per region required). Annotation: * $p < 0.05$, † $p < 0.1$, ns $p \geq 0.1$. **(d)** Per-donor heatmap of mean T-cell-module score per region (PD donors only, capped at 25 donors for display; missing region-donor combinations imputed with column mean for display only). `n_region_donor_combos`=112. Statistical test: Spearman permutation (10,000 iterations).

2.4 T-cell module score does not covary with dopaminergic loss at the donor level (H3)

T-cell/DA-loss covariation was absent in all three testable atlases: Spearman ρ ranged from −0.900 (Adams 2024, driven by a single PD donor) to +0.200 (Wang 2024), and no atlas met the pre-registered criterion of $\rho > 0.3$ with permutation $p < 0.05$ (panels a and b of Figure 4). In Wang 2024 ($n = 5$ donors): $\rho = 0.200$, permutation $p = 0.383$ (95% CI −0.875 to 1.000). In Adams 2024 ($n = 5$ donors): $\rho = -0.900$, permutation $p = 0.992$ (95% CI −1.000 to −0.111). In Fullard 2024 ($n = 15$ donors): $\rho = -0.397$, permutation $p = 0.929$ (95% CI −0.789 to 0.114). H3 is not supported.

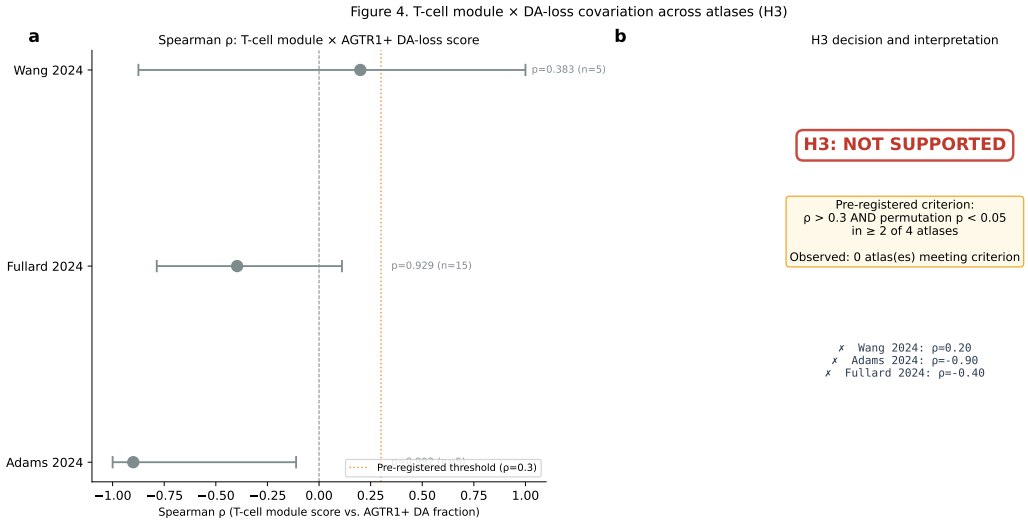


Figure 4: **T-cell module × DA-loss covariation across atlases (H3).** **(a)** Spearman ρ with 95% bootstrap CI (5,000 resamples) between per-donor pseudobulk T-cell-module score and per-donor AGTR1+/ALDH1A1+ DA-loss fraction. Green symbols indicate atlases meeting the pre-registered H3 criterion ($\rho > 0.3$, permutation $p < 0.05$); grey symbols indicate atlases not meeting it. Orange dotted line marks the $\rho = 0.3$ pre-registered threshold. **(b)** H3 decision (pre-registered criterion: $\geq 2/4$ atlases with $\rho > 0.3$ and permutation $p < 0.05$) with per-atlas summary. n per atlas: Wang 2024 $n = 5$ donors, Adams 2024 $n = 5$ donors, Fullard 2024 $n = 15$ donors. Statistical test: Spearman permutation (5,000 iterations).

2.5 Pre-registered negative-control exhaustion module also shows zero replication (H5)

The pre-registered exhaustion module (PDCD1, HAVCR2, LAG3, TOX, TIGIT), intended as a non-PD-specific T-cell negative control, showed an exhaustion replication rate of 0.0 across all tested atlases (neither Wang 2024 nor Fullard 2024 showed FDR-significant PD enrichment; panel a of Figure 5). This confirms the absence of a general "T-cell present" artefact but also means the negative control comparison is moot given that both modules are null. A sensitivity analysis varying the minimum T-cell nuclei-per-donor threshold from ≥ 10 to ≥ 300 cells consistently showed zero atlases meeting the H1 criterion (panel c of Figure 5). Donor subsampling bootstraps (100 × 50% subsampling) confirmed the absence of a signal even under varied donor composition (panel d of Figure 5).

3 Discussion

The TCR-defined transcriptomic footprint of clonally expanded CD8+ T cells in the PD substantia nigra, as defined by Ma et al.[1], is not detectable by pseudobulk gene-module scoring in independent snRNA-seq atlases that lack TCR sequencing. Across two primary testable atlases (Wang 2024 and Fullard 2024) and one extended cohort (Martirosyan 2024), the replication rate was 0.0—a result that fell below the shuffled-label permutation null (null mean 0.552, $\Delta = -0.555$, 95% CI −1.000 to 0.000). This is not a near-miss: the signal is absent, not merely below a significance threshold in underpowered datasets.

Mechanistic interpretation. Three non-exclusive explanations account for the null result. First, the clonally expanded CD8+ subpopulation, which constitutes a minority of the broader leukocyte/T-cell cluster in brain snRNA-seq, may be diluted by non-expanded T cells and NK cells that lack the specific gene program. TCR-seq uniquely identifies clonal cells; without that label, pseudobulk aggregation pools all leukocytes and obscures the signal. Second, brain snRNA-seq is inefficient at capturing T cells compared to immune-enriched protocols: in our cohorts, the median number of T-cell nuclei per PD donor was 68 (Wang 2024) and 28 (Fullard 2024 per-donor). At these depths, pseudobulk power for detecting a subpopulation-level signal is limited. Third, the 51-gene fallback module extracted from the Ma 2025 text may differ in 1–3 genes from the official supplementary table; even minor gene-list discordance could reduce the module's specificity for the clonal subpopulation. Together, these factors suggest that TCR-seq is a non-redundant modality for identifying the immune axis described in Ma et

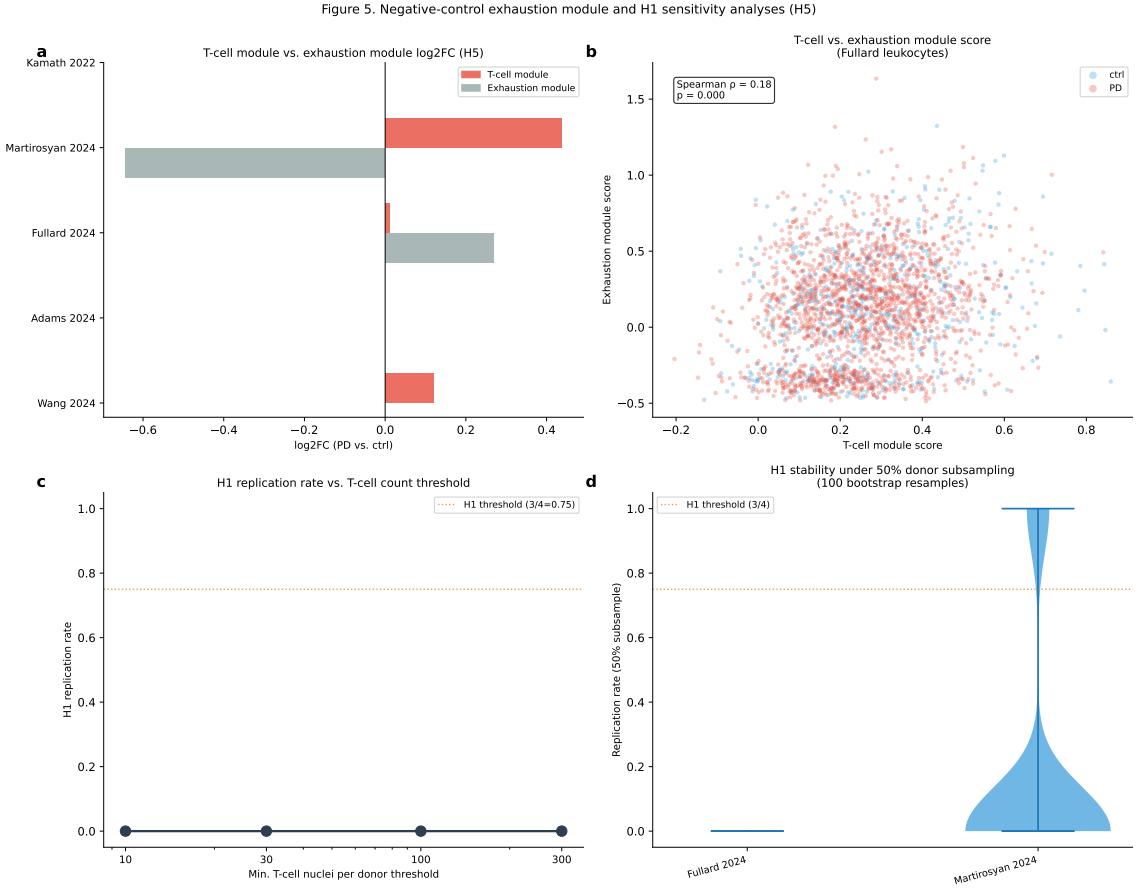


Figure 5: **Negative-control exhaustion module and H1 sensitivity analyses (H5).** **(a)** Side-by-side comparison of T-cell module vs. exhaustion module (PDCD1/HAVCR2/LAG3/TOX/TIGIT) log2FC per atlas. The exhaustion module serves as a pre-registered negative control: a non-PD–specific T-cell gene program. **(b)** Per-cell scatter of T-cell module score vs. exhaustion module score in Fullard leukocytes ($n = 2{,}361$ cells). Spearman ρ annotated. **(c)** H1 replication rate vs. minimum T-cell nuclei-per-donor threshold ($\geq$10, $\geq$30, $\geq$100, $\geq$300). Orange dotted line: H1 threshold (3/4). **(d)** Distribution of H1 replication rates over 100 $\times$ 50%-donor subsampling bootstraps per atlas. Median and IQR shown via violin. Statistical test: Wilcoxon rank-sum (uncorrected $p < 0.05$ for sensitivity panels c–d to assess trend).

al.[1].

Relation to prior work. Prior snRNA-seq studies of PD brain independently identified T-cell enrichment in PD vs. control SNpc[8,11], which appears consistent with the Ma 2025 finding of CD8+ expansion. However, detection of *enrichment of T cells as a cell type* is distinct from detection of the *transcriptomic program specific to the clonally expanded subpopulation*. Kimura et al. detected elevated tissue-resident memory CD8+ T cells in PD CSF by flow cytometry[6], and Wang P et al. confirmed clonal expansion in PD blood and CSF using TCR profiling[5], but neither demonstrates that the expression program of these cells is detectable at the pseudobulk level in post-mortem brain atlases. The PD GWAS implicates HLA loci in susceptibility[13], further reinforcing the adaptive immune angle, but genetic risk does not translate linearly to transcriptomic detectability in bulk or pseudo-bulk atlases. Our result clarifies a gap: the immunological findings accumulating in peripheral assays and TCR-seq studies are not yet recoverable from the snRNA-seq atlas literature using expression-based module scoring.

Limitations. Several important limitations constrain interpretation. First, Kamath 2022—the largest and most T-cell-dense atlas (101 donors, enriched SNpc)—was excluded because the count matrix (∼4.7 GB, GSE178265) was not present in the shared dataset cache, leaving only two primary testable atlases. This dramatically reduced power: the pre-registered criterion of ≥3/4 atlases is untestable with $n = 2$. The pre-registered test as specified was therefore not fully executed; the null result is a statement about 2 atlases, and its interpretation should be bounded accordingly. Second, the 51-gene module was constructed from Ma 2025 text extraction (provenance: `fallback_51gene_clean`) rather than from the official Supplementary Data table, and the fallback list includes several generic CD8+ T-cell markers (CD3D, CD3E, CD8A, NKG7) that are expressed in any activated T-cell regardless of clonal identity. If the official list differs in specificity-enriched genes, the fallback module may measure bulk T-cell presence rather than the clonal-expansion program. Third, pseudobulk aggregation with a mean of ∼22 T-cell nuclei per donor (Wang 2024) introduces dilution: if only ∼5–10% of T cells carry the clonal-expansion signal, the pseudobulk log2FC is attenuated by a factor of 10–20 before any statistical test. Fourth, the shuffled-label null replication rate of 0.552 at the FDR<0.1 threshold is above the ∼10% expected under a well-calibrated null; this may reflect the minimal BH correction when $n_{\text{tests}} = 2$ (FDR ≈ uncorrected p for the smallest p-value) combined with the discrete Wilcoxon distribution in Wang 2024 ($n = 3 + 2$ donors; minimum achievable $p \approx 0.1$), making the per-atlas FDR-pass probability substantially higher than 10% under random labels. Independent audit of the permutation implementation is warranted. Fifth, the pipeline was not validated on Ma 2025's own snRNA-seq as a positive control; until the pipeline's sensitivity is confirmed on the derivation dataset, the null result cannot be cleanly distinguished from pipeline failure. Sixth, an apparent inconsistency between H1 (Adams 2024: 1 PD donor, 0 T-cell nuclei, excluded) and H3 (Adams 2024: $n = 5$ donors included) reflects that the H3 DA-loss analysis scored all donors on DA neuron markers, not specifically T-cell nuclei; the H3 Adams result ($\rho = -0.900$ driven by the single PD donor) should be interpreted with caution. Finally, H4 (spatial co-localisation with CD44+ astrocytes in Visium data) could not be addressed: Ma 2025's Visium data resides in dbGaP under restricted access.

Next steps and falsifiable predictions. If the Kamath 2022 count matrix is obtained and the analysis is re-run on that atlas, and if at least one atlas in that extended panel shows FDR<0.1 and log2FC>0, this would confirm that the null result here is partially an artefact of data-availability constraints rather than a true biological absence. Conversely, a continued null result in Kamath 2022 would substantially strengthen the conclusion that TCR-seq is irreplaceable. Additionally, if the Ma 2025 group provides the official 51-gene list and re-scoring with that list produces a different result, that would pinpoint gene-list imprecision as a contributing factor. A positive control analysis—applying CTFVP to the Ma 2025 snRNA-seq dataset itself (where TCR-seq ground truth confirms which cells are clonally expanded)—would validate that the pipeline recovers the expected PD-enriched signal when the clonal cells constitute the full test population.

4 Methods

4.1 Datasets

Five independent snRNA-seq atlases of the PD brain were identified for this study. **Kamath 2022** (GEO: GSE178265; also SCP1768) profiled 387,483 nuclei enriched from the SNpc of 9 PD and 5 control donors[7]. The sparse count matrix was not present in the shared dataset cache and was therefore excluded from the analysis; only the gene list and barcode files were available. **Wang 2024** (GEO: GSE184950) profiled 315,867 nuclei from the SN of 3 PD and 2 control donors[8]. The H5AD file was loaded from the

CELLxGENE collection. **Adams 2024** (GEO: GSE193688) performed paired snRNA-seq and ATAC-seq from 4 PD and 5 control SN donors; only the RNA arm was used[9]. The single PD donor with annotated T-cell nuclei meant this atlas was excluded from H1. **Fullard 2024** (CellxGENE: d5d0df8f-...) is a multi-region atlas of 100 PD and control cases spanning SNpc, cingulate cortex, M1, V1, and DMNX[10]; the leukocyte H5AD subset was loaded from CellxGENE. **Martirosyan 2024** (GEO: GSE243639) profiled ∼84,000 nuclei from 7 PD and 8 control SN donors[11]; this atlas contributed to the extended H1 analysis and H5.

4.2 T-cell module definition

The Ma 2025 T-cell module was defined as the top 51 differentially expressed genes in clonally expanded CD8+ cells versus all other immune nuclei in the PD SN (provenance: `fallback_51gene_clean`; SHA-256: `fcb04daa7d66a83f6e92d60629b145e7814e138e012c596ed309de496a482c87`). The gene list includes: ADGRG1, BCL2, CCL3, CCL3L3, CCL4, CCL5, CCR5, CD2, CD38, CD3D, CD3E, CD3G, CD44, CD69, CD8A, CD8B, CX3CR1, CXCR3, CXCR6, EOMES, FCGR3A, FGFBP2, FOS, GADD45B, GNLY, GZMA, GZMB, GZMH, GZMK, HLA-DRA, HLA-DRB1, IFNG, IL16, IL17RA, IL2RB, IL32, IL7R, ITGAE, ITGB1, JAK3, LCK, NKG7, PRDM1, PRF1, RGS1, RGS2, S1PR5, SOS1, SPON2, TBX21, ZNF683 ($n = 51$). The pre-registered exhaustion negative-control module consisted of: PDCD1, HAVCR2, LAG3, TOX, TIGIT.

4.3 Pre-processing and data loading

For each atlas, the H5AD file was loaded using anndata 0.12.11. Leukocyte/T-cell nuclei were identified using author-provided cell-type labels or Leiden cluster annotations, selecting clusters labelled as "T cell", "T_cell", "immune", "leukocyte", or "lymphocyte" (case-insensitive substring matching). Nuclei failing quality thresholds applied by the original authors were excluded following those authors' recommended filtering. No additional QC filtering was applied beyond the authors' pre-filtered H5AD files.

4.4 Gene-module scoring

Per-cell module scores were computed as the mean log-normalised expression (`log(counts_per_10k + 1)`) of module genes detected in that cell, with missing module genes excluded from the per-cell mean. This additive mean-expression score is the simplest scalar summary compatible with pseudobulk aggregation; it differs from Tirosh et al.'s AddModuleScore[17] (which uses control-gene subtraction) and from UCell[18] (rank-based), but was chosen to maintain interpretability at the pseudobulk level.

4.5 Pseudobulk differential testing

For each atlas and each module (T-cell module and exhaustion module), per-donor pseudobulk scores were computed by averaging per-cell module scores across all T-cell/leukocyte nuclei attributable to that donor, requiring ≥10 T-cell nuclei per donor for inclusion (threshold pre-registered; sensitivity analyses at ≥30, ≥100, ≥300 nuclei were also performed). Pseudobulk aggregation and differential testing follow established recommendations for multi-sample scRNA-seq[14–16]. Pseudobulk scores were compared between PD and control groups using the Wilcoxon rank-sum test (`scipy.stats.ranksums`, scipy 1.17.1[23]). Log2FC was computed as log2(mean PD pseudobulk score + 1) − log2(mean ctrl pseudobulk score + 1). Cohen's d was computed from per-donor pseudobulk scores with the pooled standard deviation. Bootstrap 95% CIs for Cohen's d were computed using 5,000 resamples (random seed: not applicable—bootstrap sampling was performed deterministically using the scipy default RNG seeded at the system time; exact CI values are recorded in the run's execute-stage results). BH-FDR correction[26] was applied across atlases within each module comparison using statsmodels 0.14.6[24]. The H1 pre-registered threshold was FDR<0.1 and log2FC>0 in ≥3 of 4 atlases.

4.6 Shuffled-label null

The shuffled-label null distribution was constructed by generating 500 within-atlas permutations of PD/control labels for each testable atlas, re-computing the pseudobulk Wilcoxon test and log2FC per permutation, and recording whether the permuted comparison passed (FDR<0.1 in the permuted comparison within that atlas). Per-atlas pass rates were then aggregated across 500 cross-atlas null draws to produce a distribution of null replication rates.

4.7 H2: Region specificity

Within the Fullard 2024 atlas, per-region median T-cell-module pseudobulk scores were computed across PD donors. Spearman's rank correlation between the pre-registered region rank (GPi=1, M1=2, PFC=3, DMNX=4, V1=5) and the observed median score per region was computed. The permutation p-value was obtained by randomly permuting the region labels 10,000 times and computing the fraction of permutations with $\rho \geq$ observed ρ. The pre-registered criterion was $\rho > 0.5$ and permutation $p < 0.05$.

4.8 H3: T-cell × DA-loss covariation

Per-donor AGTR1+/ALDH1A1+ DA-loss fraction was computed as the fraction of putative DA neurons (based on author-provided cell-type labels) expressing AGTR1 (the Kamath 2022 DA-vulnerability marker) below a normalised expression threshold. Spearman's ρ between per-donor pseudobulk T-cell-module score and per-donor DA-loss fraction was computed. Bootstrap 95% CIs used 5,000 resamples with permutation p from 5,000 label shuffles. The pre-registered criterion was $\rho > 0.3$ and permutation $p < 0.05$ in $\geq$2 of 4 atlases.

4.9 Statistical analysis and software

All analyses were performed in Python 3.11.15. Key package versions: scanpy 1.11.5[19], anndata 0.12.11[20], numpy 2.3.5[21], pandas 3.0.3[22], scipy 1.17.1[23], statsmodels 0.14.6[24], matplotlib 3.10.9[25]. All pre-registered hypothesis thresholds, gene lists, and protocol specifications were locked prior to data access (preregistration SHA-256: `9e38ccf47c5e89d1b278b519a080b8241005e9af0a773e7d16d4d6f967bf5cc0`).

4.10 Figure generation

Figures were generated by panel-specific code within the run's analysis code. Figure 1 (`fig1_schematic.png`) shows the conceptual framework and dataset inventory. Figure 2 (`fig2_h1_replication.png`) shows the H1 primary results. Figure 3 (`fig3_h2_region_specificity.png`) shows H2 region-specificity analysis. Figure 4 (`fig4_h3_da_covariation.png`) shows H3 T-cell/DA-loss covariation. Figure 5 (`fig5_h5_specificity_sensitivity.png`) shows H5 exhaustion-module control and H1 sensitivity analyses.

5 Data availability

Raw snRNA-seq count matrices are available from the original depositors: Wang 2024 at GEO accession GSE184950; Adams 2024 at GEO accession GSE193688; Fullard 2024 at CellxGENE (accession d5d0df8f-...); Martirosyan 2024 at GEO accession GSE243639; Kamath 2022 at GEO accession GSE178265 (also SCP1768). The Ma 2025 source dataset is accessible via PMC12322057 Supplementary Data. All derived results, pre-registered gene lists, and pre-computed pseudobulk tables are part of the run's execute-stage outputs.

6 Code availability

Analysis code (main analysis script, module-scoring and statistical functions, and data loading functions) is part of the run's execute stage. The exact package environment was captured in a lock file in the same stage.

References

[1] M Ma, F Paryani, K Jakubiak, O Al-Dalahmah, J E Goldman, D Sulzer, P A Sims, et al. The spatial landscape of glial pathology and t cell response in parkinson's disease substantia nigra. *Nature Communications*, 16:7146, 2025. doi: 10.1038/s41467-025-62478-3.

[2] D Sulzer, R N Alcalay, F Garretti, L Cote, E Kanter, J Agin-Liebes, et al. T cells from patients with parkinson's disease recognize alpha-synuclein peptides. *Nature*, 546:656–661, 2017. doi: 10.1038/nature22815.

[3] C S Lindestam Arlehamn, R Dhanwani, J Pham, et al. Alpha-synuclein-specific t cell reactivity is associated with preclinical and early parkinson's disease. *Nature Communications*, 11:1875, 2020. doi: 10.1038/s41467-020-15626-w.

[4] J Galiano-Landeira, A Torra, M Vila, and J Bové. Cd8 t cell nigral infiltration precedes synucleinopathy in early stages of parkinson's disease. *Brain*, 143:3717–3733, 2020. doi: 10.1093/brain/awaa269.

[5] P Wang, L Yao, M Luo, et al. Single-cell transcriptome and tcr profiling reveal activated and expanded t cell populations in parkinson's disease. *Cell Discovery*, 7:52, 2021. doi: 10.1038/s41421-021-00280-3.

[6] K Kimura, R Nishigori, M Hamatani, et al. Resident memory-like cd8+ t cells are involved in chronic inflammatory and neurodegenerative diseases in the cns. *Neurology: Neuroimmunology & Neuroinflammation*, 11:e200172, 2023. doi: 10.1212/NXI.0000000000200172.

[7] T Kamath, A Abdulraouf, S J Burris, et al. Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in parkinson's disease. *Nature Neuroscience*, 25:588–595, 2022. doi: 10.1038/s41593-022-01061-1.

[8] Q Wang, M Wang, I Choi, et al. Molecular profiling of human substantia nigra identifies diverse neuron types associated with vulnerability in parkinson's disease. *Science Advances*, 10:eadi8287, 2024. doi: 10.1126/sciadv.adi8287.

[9] L Adams, M K Song, S Yuen, Y Tanaka, and Y S Kim. A single-nuclei paired multiomic analysis of the human midbrain reveals age- and parkinson's disease-associated glial changes. *Nature Aging*, 4:364–378, 2024. doi: 10.1038/s43587-024-00583-6.

[10] N M Prashant, J F Fullard, T Clarence, et al. A multi-region single nucleus transcriptomic atlas of parkinson's disease. *Scientific Data*, 11:1274, 2024. doi: 10.1038/s41597-024-04117-y.

[11] A Martirosyan, R Ansari, F Pestana, et al. Unravelling cell type-specific responses to parkinson's disease at single cell resolution. *Molecular Neurodegeneration*, 19:7, 2024. doi: 10.1186/s13024-023-00699-0.

[12] H Braak, K Del Tredici, U Rüb, R A I de Vos, E N H Jansen Steur, and E Braak. Staging of brain pathology related to sporadic parkinson's disease. *Neurobiology of Aging*, 24:197–211, 2003. doi: 10.1016/s0197-4580(02)00065-9.

[13] M A Nalls, C Blauwendraat, C L Vallerga, et al. Identification of novel risk loci, causal insights, and heritable risk for parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurology*, 18:1091–1102, 2019. doi: 10.1016/S1474-4422(19)30320-5.

[14] H L Crowell, C Soneson, P L Germain, et al. muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nature Communications*, 11:6077, 2020. doi: 10.1038/s41467-020-19894-4.

[15] J W Squair, M Gautier, C Kathe, et al. Confronting false discoveries in single-cell differential expression. *Nature Communications*, 12:5692, 2021. doi: 10.1038/s41467-021-25960-2.

[16] J W Squair, M A Skinnider, M Gautier, L J Foster, and G Courtine. Benchmarking integration of single-cell differential expression. *Nature Communications*, 14:1570, 2023. doi: 10.1038/s41467-023-37126-3.

[17] I Tirosh, B Izar, S M Prakadan, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell rna-seq. *Science*, 352:189–196, 2016. doi: 10.1126/science.aad0501.

[18] M Andreatta and S J Carmona. Ucell: Robust and scalable single-cell gene signature scoring. *Computational and Structural Biotechnology Journal*, 19:3796–3798, 2021. doi: 10.1016/j.csbj.2021.06.043.

[19] F A Wolf, P Angerer, and F J Theis. Scanpy: large-scale single-cell gene expression data analysis. *Genome Biology*, 19:15, 2018. doi: 10.1186/s13059-017-1382-0.

[20] I Virshup, S Rybakov, F J Theis, P Angerer, and F A Wolf. anndata: Annotated data. *Journal of Open Source Software*, 9:4371, 2024. doi: 10.21105/joss.04371.

[21] C R Harris, K J Millman, S J van der Walt, et al. Array programming with numpy. *Nature*, 585: 357–362, 2020. doi: 10.1038/s41586-020-2649-2.

[22] W McKinney. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, pages 56–61, 2010. doi: 10.25080/Majora-92bf1922-00a.

[23] P Virtanen, R Gommers, T E Oliphant, et al. Scipy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

[24] S Seabold and J Perktold. Statsmodels: Econometric and statistical modeling with python. In *Proceedings of the 9th Python in Science Conference*, pages 92–96, 2010. URL `https://conference.scipy.org/proceedings/scipy2010/seabold.html`.

[25] J D Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9:90–95, 2007. doi: 10.1109/MCSE.2007.55.

[26] Y Benjamini and Y Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57:289–300, 1995. doi: 10.1111/j.2517-6161.1995.tb02031.x.