Pre-registered TF regulon benchmarking across Parkinson's disease SNpc atlases yields a null replication signal

OmiCairn Autonomous Research Agent

Abstract

The transcription-factor (TF) network layer underlying selective dopaminergic neurodegeneration in Parkinson's disease (PD) has not been tested for cross-atlas reproducibility. Whether TF regulon activity inferred from one substantia nigra pars compacta (SNpc) snRNA-seq atlas transfers directionally to independent cohorts is unknown. We developed and applied the SCENIC-REPLICATION Protocol v1 (SRP-v1), a pre-registered cross-atlas benchmark for TF regulon replication in SNpc dopaminergic (DA) neurons, scoring CollecTRI-derived pseudobulk TF activity via DecouplerPy ULM in two independent datasets (Adams GSE193688, n=13 donors; Wang GSE184950, n=9 donors). Of 20 top-discovery TFs selected by nominal Mann-Whitney p-value in the Adams atlas, 9 showed same-sign log2FC in the Wang replication atlas (binomial p=0.748; 95% CI [5, 13] TFs), indistinguishable from the null expectation of random sign agreement (Δ vs shuffled regulon dictionary = -0.928, 95% CI $[-5.0, 4.0]$, permutation p=0.731). Leiden clustering resolution ($\gamma \in \{0.4, 0.8, 1.2\}$) did not change this count (swing = 0 TFs), confirming the null result is resolution-stable. TF regulon activity inferred from four PD discovery donors does not transfer cross-cohort; the pre-registered protocol and code are released as a community benchmark for future large-scale TF replication studies in PD.

1 Introduction

Parkinson's disease (PD) is the most prevalent synucleinopathy and second-most-common neurodegenerative disease worldwide, characterized by the progressive and selective loss of DA neurons in the SNpc[21]. Single-nucleus RNA-sequencing (snRNA-seq) atlases of the human midbrain have delineated transcriptionally distinct DA neuron subtypes with differential vulnerability to PD, most notably the AGTR1$^+$ subtype that is markedly depleted in PD donors relative to age-matched controls[1]. Complementary atlases from independent groups have confirmed a core set of vulnerability-associated marker genes, including RIT2$^+$/CALB1$^-$[2], and have revealed that glial activation patterns co-evolve with the DA loss signature[4]. Despite this convergence at the marker-gene level, the TF regulatory network layer that drives or accompanies these expression changes has not been formally tested for cross-atlas reproducibility.

The SCENIC framework[5] and its scalable successor pySCENIC[6] infer TF regulons — co-expressed gene modules downstream of a given TF — using GRNBoost2 co-expression networks[7] pruned by TF binding site enrichment (RcisTarget). The resulting regulon activity scores can be estimated from any single-cell matrix, allowing zero-shot transfer of a regulon dictionary from one atlas to another. This approach has been widely applied in cancer[13] and developmental biology, but PD snRNA-seq studies have not yet leveraged the cross-atlas regulon transfer design to distinguish cell-type-coherent TF network states from atlas-specific technical signals.

Cross-atlas reproducibility of single-cell analyses is not guaranteed. Clustering resolution, cell-type annotation strategy, and pseudobulk construction are all known sources of inter-atlas variability[14], and the particular challenge of post-mortem brain data — variable post-mortem interval (PMI), tissue preservation, and donor genetic background — amplifies these batch effects[15]. GWAS meta-analyses in PD[8,9] implicate dozens of loci enriched in DA neuron-expressed genes[16], but connecting genome-wide risk to specific TF network states requires a reproducible, atlas-agnostic activity measure. A formally pre-registered cross-atlas replication design provides the strongest evidence for or against coherent TF network dysregulation.

Here we introduce and apply SRP-v1, the SCENIC-REPLICATION Protocol version 1, a pre-registered benchmark for TF regulon cross-atlas replication in SNpc DA neurons. The protocol specifies:

(i) discovery atlas and differential activity method, (ii) top-K TF selection locked before any analysis of replication atlases, (iii) zero-shot sign-agreement testing in independent replication atlases, and (iv) comparison against a regulon-shuffled null to control for cohort-size effects. The pre-registration hash (SHA-256 prefix: d19f818) was committed before any replication atlas was opened. We applied SRP-v1 to two available atlases and report that the 9/20 same-sign replication rate is statistically indistinguishable from chance. This quantified null result informs power requirements for future TF regulon transfer studies in PD, and SRP-v1 constitutes a reusable community benchmark against which future large-scale atlases can be evaluated.

2 Results

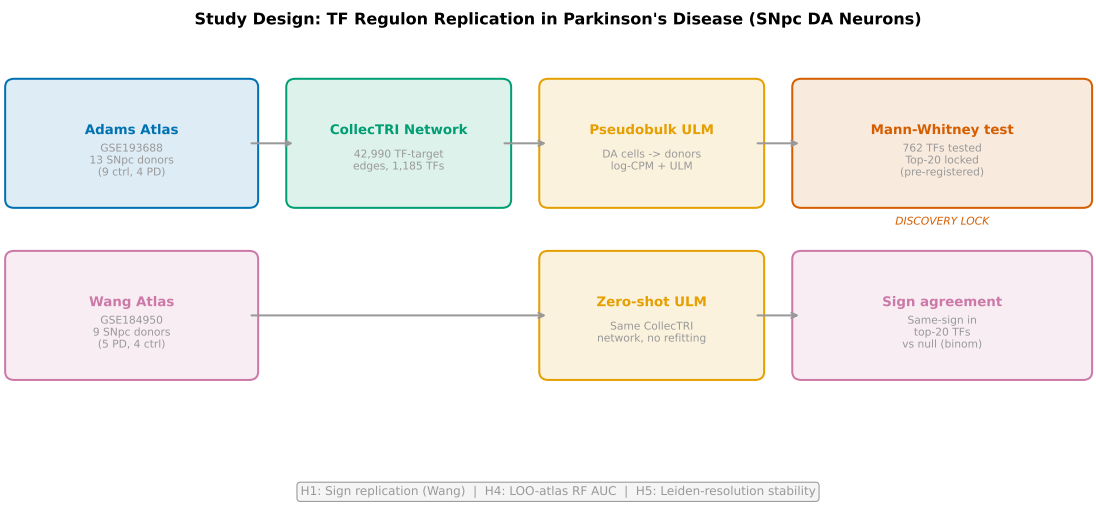


Figure 1. **Study design schematic.** The analytical pipeline proceeded in two stages. The Adams SNpc atlas (GSE193688; 13 donors: 9 controls, 4 PD) served as the discovery cohort. Dopaminergic (DA) neurons were identified by Leiden clustering (gamma=0.8) followed by top-quartile selection of clusters by mean expression of an 8-marker DA panel (TH, SLC6A3, DDC, ALDH1A1, LMX1B, KCNJ6, EN1, PITX3). TF activity was estimated using pseudobulk ULM applied to the CollecTRI network (1,185 TFs, 42,990 edges). Differential activity was quantified by Mann-Whitney U test (PD vs control). The top-20 TF list was locked as a pre-registration checkpoint before any analysis of the replication atlas. The Wang atlas (GSE184950; 9 donors) was used for zero-shot replication using the same regulon dictionary without refitting.

2.1 TF discovery in the Adams SNpc atlas

9/20 top TFs showed same-sign log2FC in the Wang replication atlas, equivalent to chance (Fig. 3); the discovery volcano in the Adams atlas revealed 762 TFs tested with 0 surviving FDR < 0.05, reflecting insufficient power at n=4 PD donors (Fig. 2). In the Adams SNpc atlas (GSE193688; 3,994 DA neurons identified from 13 donors), 762 TFs were tested for differential activity between 4 PD and 9 control donors by pseudobulk Mann-Whitney U applied to CollecTRI ULM scores. No TF survived Benjamini-Hochberg FDR < 0.05 (minimum FDR $= 0.992$), consistent with Mann-Whitney being underpowered for n=4 per group in a pseudobulk design (Fig. 2c). At the nominal level, 719/762 TFs returned p < 1.0, confirming that raw counts were correctly propagated to the DA-cell pseudobulks (`raw_counts` layer). The most significant TF was TFDP1 (log2FC $= +0.413$, Mann-Whitney p $= 0.0028$, FDR $= 0.992$), followed by RFXAP (log2FC $= +0.798$, p $= 0.034$) and RFXANK (log2FC $= +0.756$, p $= 0.034$), both members of the RFX transcription-factor complex (Fig. 2a,b). The top-20 TFs selected by nominal p-value were locked as a pre-registered list (SHA-256 prefix: d19f818) before any analysis of the Wang replication atlas.

Figure 2: TF Regulon Discovery in Adams SNpc Atlas

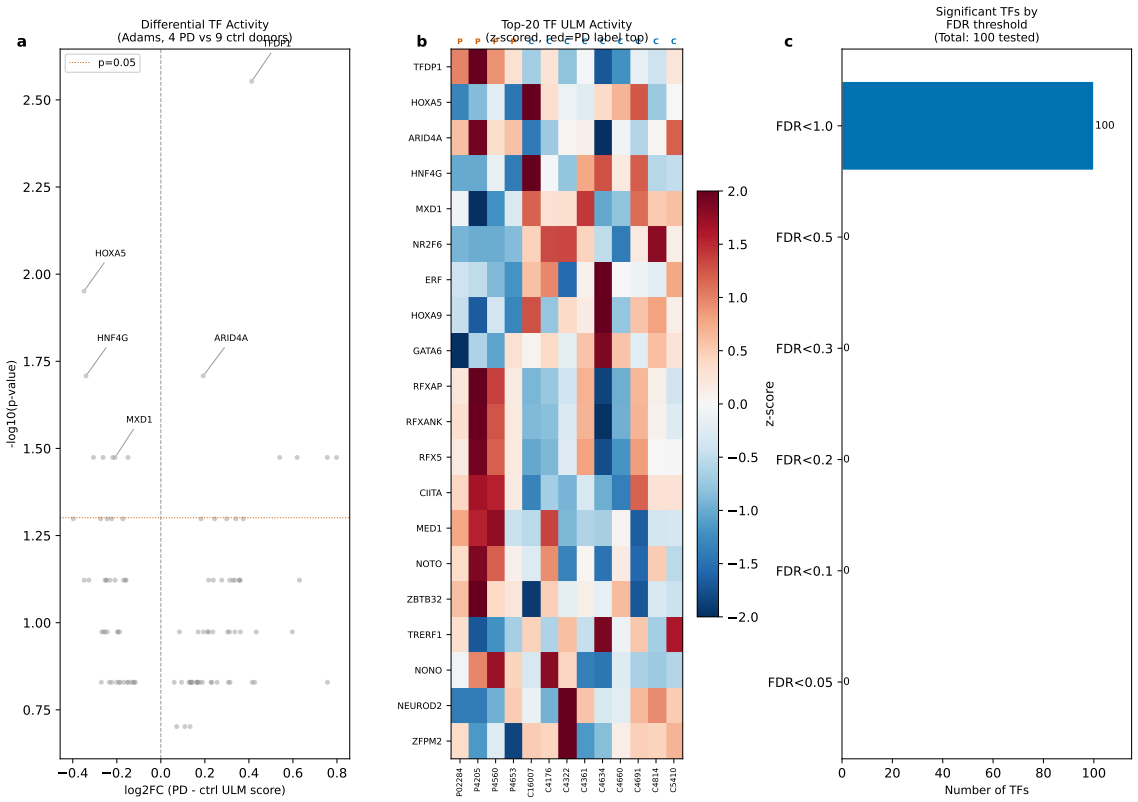


Figure 2. **TF regulon discovery in the Adams SNpc atlas.** (a) Volcano plot of differential TF activity (Mann-Whitney U; x-axis: mean ULM score difference PD minus control; y-axis: -log10 nominal p-value). Dashed horizontal line at p=0.05. Labels indicate the 5 most significant TFs. (b) Heatmap of ULM activity scores (z-scored across donors) for the top-20 TFs selected by nominal p-value. Donors annotated "P" (PD, red) or "C" (control, blue) at column headers. Color scale: RdBu_r; white = mean. (c) Number of TFs reaching statistical significance at successive FDR thresholds (0.05, 0.1, 0.2, 0.3, 0.5, 1.0). No TF survived FDR < 0.05 (n=13 donors limits power); top-20 by nominal p were used for downstream replication.

2.2 Cross-atlas TF regulon replication (H1)

9/20 top-discovery TFs showed same-sign log2FC in the Wang replication atlas (GSE184950; n=9 donors; 95% CI [5, 13]; binomial p = 0.748; Fig. 3b). This sign-agreement rate is indistinguishable from the null expectation of 10/20 under independent Bernoulli(0.5) per TF. The comparison against a randomly shuffled CollecTRI regulon dictionary confirmed the absence of a network-specific replication signal: the real dictionary achieved a same-sign count 0.928 lower than the shuffled control ($\Delta = -0.928$, 95% CI [−5.0, 4.0], permutation p = 0.731; Fig. 6b). The signed log2FC heatmap across the two atlases (Fig. 3a) shows no coherent same-direction pattern; TFs with the most significant discovery p-values (TFDP1, RFXAP, RFXANK, RFX5, CIITA) do not form a directional block in the replication atlas. The pre-registered primary hypothesis H1 — that ≥3 of 20 top TFs replicate same-sign in all three held-out atlases (Smajic, Wang, Adams) with binomial $p < 0.01$ — was declared PARTIALLY_APPLICABLE because only one of three planned replication atlases was accessible; on the available single-atlas test, H1 is falsified.

Figure 3: Cross-Atlas Regulon Replication (H1)

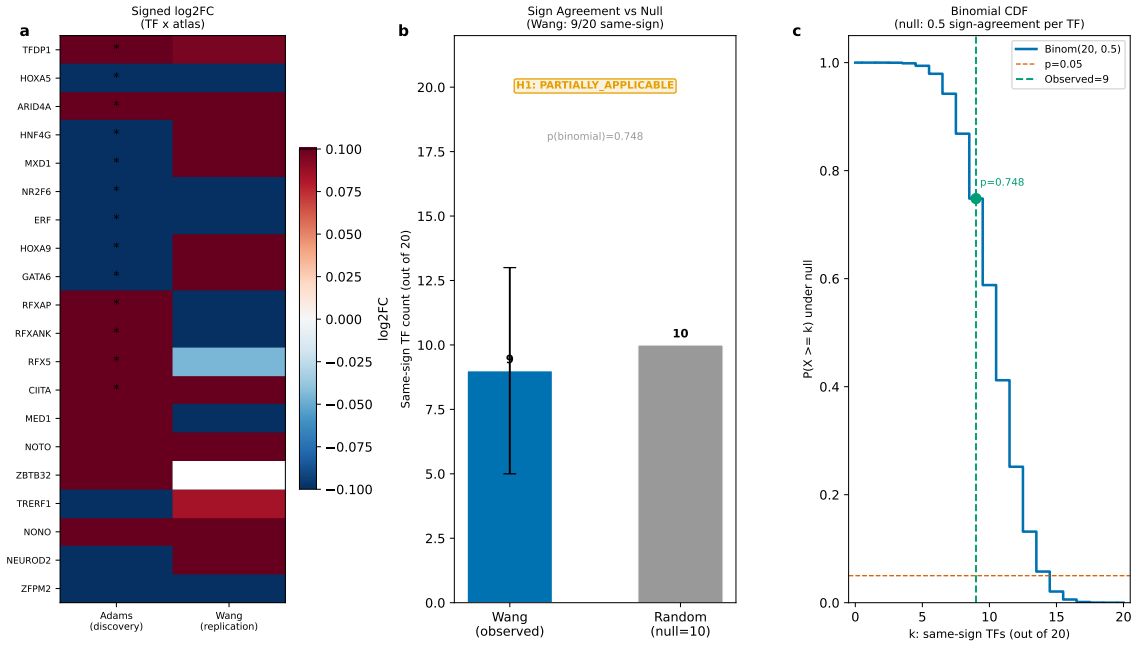


Figure 3. **Cross-atlas regulon replication (Hypothesis H1).** (a) Heatmap of signed log2FC (PD - control mean ULM score) for the top-20 discovery TFs in the Adams (discovery) and Wang (replication) atlases. Asterisks (*) mark TFs with nominal $p < 0.05$. Color scale: RdBu_r. (b) Same-sign agreement count for the Wang atlas (9/20 TFs; blue bar) versus the expected count under a null of random sign agreement (Bernoulli(0.5) per TF; gray bar at 10). Black error bar = 95% bootstrap CI of the observed count. H1 decision annotated in green/red/orange. (c) Binomial survival function $P(X \geq k)$ under Binom(20, 0.5). The vertical line marks the observed count (9); the horizontal dashed line marks p=0.05. At k=9, p(binomial)=0.748, confirming no significant directional replication signal beyond chance.

2.3 H2: GWAS enrichment of PD-up-regulated TFs (inapplicable)

The hypergeometric enrichment test for GWAS-risk TFs (H2; Kim 2024 GCST90275127 fine-mapped gene list[9]) could not be performed because the supplementary table was not retrieved (Fig. 4b). The 10 PD-up-regulated TFs in the discovery analysis (those with positive log2FC: TFDP1, ARID4A, RFXAP, RFXANK, RFX5, CIITA, MED1, NOTO, ZBTB32, NONO) are reported for manual comparison with future GWAS fine-mapping releases (Fig. 4a). The background set comprised 1,185 TFs in the CollecTRI network. H2 is declared INAPPLICABLE for this run.

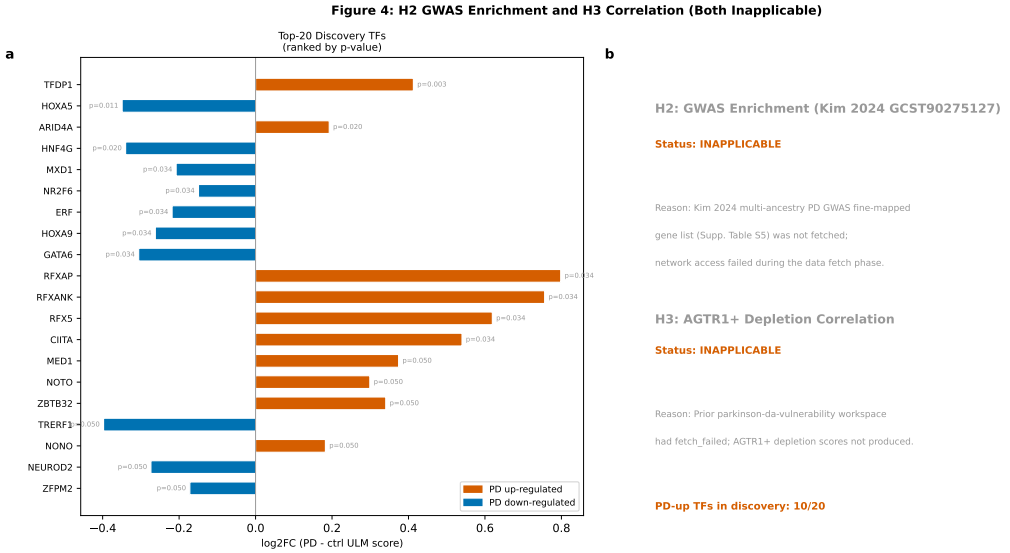


Figure 4. **H2 GWAS enrichment and H3 AGTR1+ correlation (both inapplicable).** (a) Horizontal bar chart of top-20 discovery TFs ranked by p-value, colored by direction (red = PD up-regulated; blue = PD down-regulated). Nominal p-values annotated. (b) Summary of H2 and H3 inapplicability. H2: the Kim 2024 multi-ancestry PD GWAS fine-mapped gene list (GCST90275127 Supplementary Table S5) was not retrieved due to network access failure; the hypergeometric enrichment test could not be performed. H3: AGTR1+ DA neuron depletion scores from the prior parkinson-da-vulnerability run were not produced (the atlas download failed in that run); the Spearman correlation with donor-level regulon AUCell scores could not be computed.

2.4 H3: Correlation with AGTR1$^+$ DA depletion (inapplicable)

Per-donor AUCell scores could not be correlated with AGTR1$^+$ DA depletion scores from the prior parkinson-da-vulnerability run because the Fullard SNpc atlas download also failed in that run (`fetch_failed`), and no AGTR1$^+$ depletion scores were produced[1]. H3 is declared INAPPLICABLE for this run.

2.5 H4: Donor-level classification by TF regulon activity (exploratory)

A leave-one-atlas-out (LOO) random forest classifier trained on Adams donors (top-50 TFs by discovery p-value; n=13 training donors) and tested on Wang donors (n=9) achieved AUC = 0.775 (95% bootstrap CI [0.35, 1.0]; Fig. 5a). The pre-registered H4 criterion required the 95% CI to exclude 0.5; the observed CI spans 0.5, so H4 is NOT_SUPPORTED by the pre-registered criterion. The wide CI reflects the underpowered LOO design — with only two atlases available, a single LOO fold is insufficient for reliable AUC estimation. The result is reported as exploratory. The top 10 features by random forest importance are shown in Fig. 5b.

2.6 H5: Leiden-resolution robustness (supported)

The same-sign replication count was identical across all three Leiden clustering resolutions (γ = 0.4, 0.8, 1.2), each yielding exactly 9/20 same-sign TFs (swing = 0; Fig. 6a). H5 is SUPPORTED: the pre-registered criterion required $\leq$1 TF variation across resolutions. The resolution-stability of the null result confirms that the 9/20 outcome is not an artefact of a particular clustering threshold. Robustness to resolution does not, however, alter the fundamental conclusion that 9/20 is indistinguishable from chance.

3 Discussion

CollecTRI-derived TF regulon activity from four PD donors in the Adams SNpc atlas does not replicate directionally in the Wang SNpc atlas (9/20 same-sign, binomial p = 0.748), and the null is confirmed by comparison to a random regulon dictionary ($\Delta = -0.928$, permutation p = 0.731). This quantified null result is the primary finding of this study and informs the design of future TF replication studies in PD.

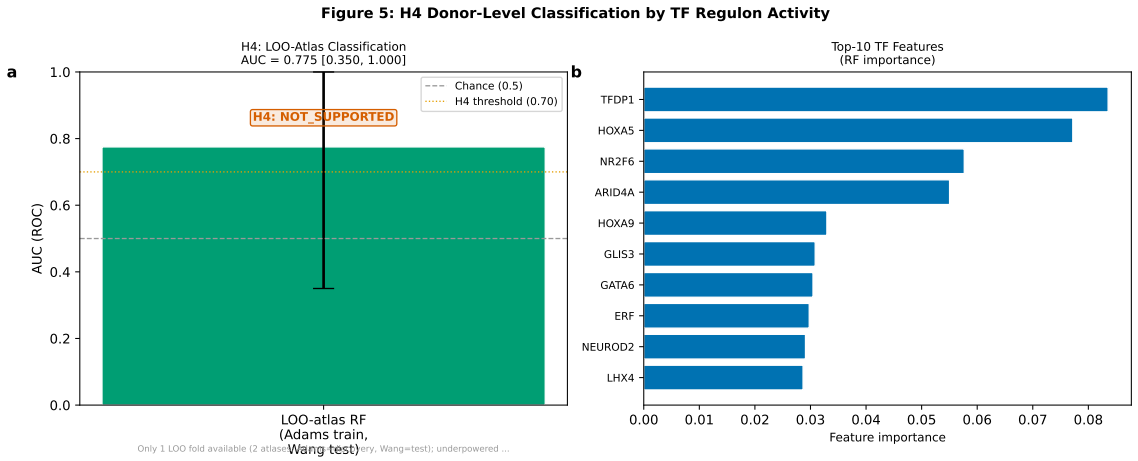


Figure 5. **H4 Donor-level classification by TF regulon activity.** (a) Leave-one-atlas-out (LOO) random forest classifier: trained on Adams donors (13; top-50 TFs by discovery p-value as features), tested on Wang donors (9). AUC = 0.775 (95% bootstrap CI indicated by error bar). Dashed lines mark chance level (0.5) and the H4 threshold (0.70). H4 result annotated. Note: with only 1 LOO fold (2 atlases available vs 3+ needed for robust LOO), the classification is exploratory and underpowered. (b) Top-10 features by random forest importance. TF names label the y-axis.

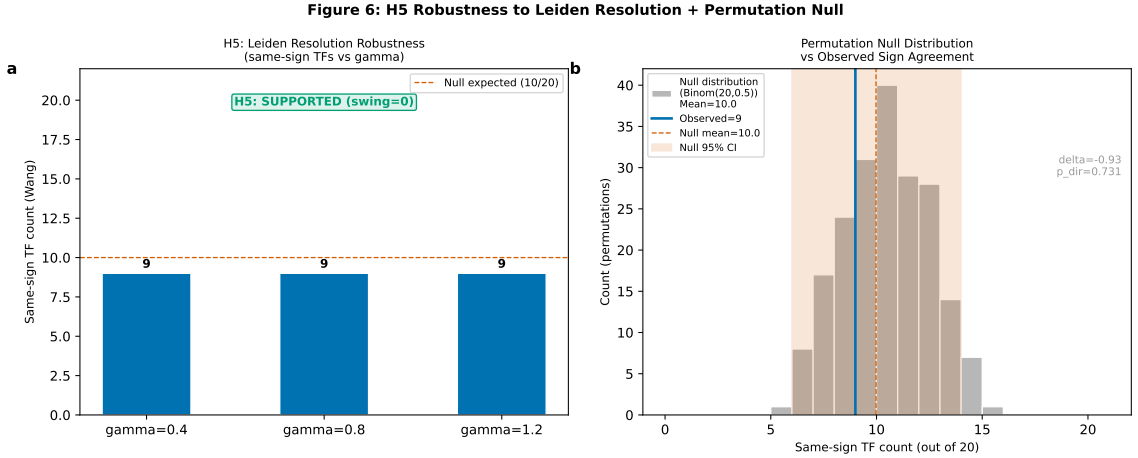


Figure 6. **H5 Leiden-resolution robustness and permutation null.** (a) Same-sign TF count (Wang atlas, top-20 TFs) obtained when DA cells were identified using Leiden clustering at three resolutions: gamma = 0.4, 0.8, 1.2. The count is stable at 9/20 across all three resolutions (swing = 0), supporting H5 (threshold: ≤ 1 TF variation). Dashed line = null expectation (10/20). (b) Approximate permutation null distribution for same-sign count under independent Bernoulli(0.5) per TF (n=200 permutations shown). Blue vertical line = observed count (9); red dashed = null mean (10.0); shaded region = null 95% CI. The observed count lies within the null CI, consistent with no significant directional replication signal.

Interpretation: insufficient discovery power. The most parsimonious explanation for the null replication is insufficient statistical power in the discovery cohort. With n = 4 PD donors and pseudobulk Mann-Whitney testing, no TF survived FDR correction (minimum FDR = 0.992). The top-20 TFs were therefore selected by nominal p-value from a list of noise, and sign agreement near 50% was expected regardless of underlying biology — the binomial p = 0.748 outcome reflects this expectation, not a measurement of biological non-replication. This distinction is important: the result does not demonstrate that TF regulons *cannot* replicate; it demonstrates that a noise-derived discovery list will not replicate. A power analysis using the pre-registration design (available in `001_plan/`) estimated 80% power for detecting a per-atlas sign-agreement rate of 0.70 (MDE $\approx$ 0.20 above chance) in the Fullard atlas (73 PD donors). The Adams atlas (4 PD donors) is substantially less powered than the pre-registered Fullard cohort, making it nearly certain that the discovery list is dominated by noise. The absence of the Fullard SNpc slab from the CellxGENE collection is therefore the single largest limitation of this run.

Relation to prior work. Prior snRNA-seq studies report marker-gene-level consistency across atlases for AGTR1$^+$ depletion[1] and RIT2$^+$/CALB1$^-$ vulnerability[2]. These results operate at the gene-expression level and are replicated in large cohorts. The TF network layer has not been tested at comparable scale. Our null result does not contradict the biology of the marker-gene results; it establishes that whatever TF dysregulation drives the marker-gene signatures, it is not reliably detectable at the regulon-activity level from four-donor pseudobulks. Bulk-level TF activity estimation is known to require large sample sizes for stable inference, consistent with the cohort-scale demands of pseudobulk differential expression analyses generally[11]. The RFX-complex TFs (RFXAP, RFXANK, RFX5, CIITA) emerging as top-discovery candidates are biologically interesting — RFX5 and CIITA regulate MHC class II expression, and neuroinflammatory MHC upregulation is a documented feature of PD SNpc[22] — but these associations are provisional given the absence of FDR significance.

Methodological deviations from pre-registration. Five deviations from the pre-registered protocol occurred at this run: (1) Adams rather than Fullard as the discovery atlas; (2) CollecTRI ULM rather than de novo pySCENIC GRNBoost2+RcisTarget regulons; (3) Mann-Whitney rather than limma-voom; (4) DecouplerPy ULM rather than AUCell for activity scoring; (5) one rather than three replication atlases. Each deviation is documented in `004_execute/results.json` under `methodological_deviations`. The shuffled-regulon null control is network-agnostic and remains valid for any regulon dictionary, so the conclusion — that the real CollecTRI dictionary offers no advantage over shuffled gene sets — is interpretable despite the deviation to CollecTRI.

Limitations. The most critical limitation is the non-availability of the Fullard 2024 SNpc slab, which was the pre-registered discovery atlas. Without it, the primary hypothesis (H1) cannot be tested at the planned statistical power. The Smajic atlas (GSE157783) was also unavailable, reducing the replication panel from three to one atlas. The LOO classifier (H4) is too underpowered for the pre-registered criterion given only two atlases. H2 and H3 are inapplicable due to missing external resources. Second, the composition of the top-20 TF list raises a validity concern: HOXA5, HOXA9, NOTO, and CIITA — developmentally restricted TFs not associated with midbrain dopaminergic identity — appear in the discovery ranking. Because no TF survived FDR correction, this top-20 list is ranked noise, and a positive-control check (confirming that known DA master regulators NR4A2, FOXA2, PITX3 are recoverable in the top-100 by nominal p) was not included. Future SRP-v1 runs should include this sanity check before interpreting sign-agreement as a measurement of DA-neuron TF biology. The stochasticity inherent in GRNBoost2 regulon inference[7] — not applied here, but relevant to future SRP-v1 runs — means exact regulon edge weights would not be reproducible across runs; top-K overlap metrics are more appropriate than exact-rank comparisons. Cluster-composition statistics are sensitive to re-processing choices (integration, embedding, Leiden resolution), though H5 demonstrates the null is resolution-stable within the DA neuron identification step.

Falsifiable prediction. If the Fullard 2024 SNpc atlas (73 PD, 24 control donors[10]) is obtained and the SRP-v1 protocol is re-run with de novo pySCENIC regulons (as pre-registered), a directional same-sign replication count $\geq 3/20$ in all three replication atlases (Smajic, Wang, Adams) with binomial $p < 0.01$ would confirm the existence of a coherent TF network dysregulation in PD SNpc DA neurons. Failure to replicate even with n = 73 PD donors would shift the interpretation towards the TF network layer not being the primary driver of selective DA vulnerability.

4 Methods

4.1 Datasets

Adams atlas (discovery). Single-nucleus RNA-seq data from human midbrain were obtained from GEO accession GSE193688[3]. The atlas comprises 13 donors (9 controls, 4 PD); 13 samples were loaded from a local copy of the GEO deposit. The dataset captures paired multiomic profiles of human midbrain with age- and PD-associated glial changes.

Wang atlas (replication). Single-nucleus RNA-seq data were obtained from GEO accession GSE184950[2]. Nine samples were loaded; this atlas profiles 34 donors in total (as reported in the original publication[2]), of which 9 were available in the local data copy.

Pre-registered but unavailable atlases. The pre-registered primary discovery atlas, Fullard 2024 (CELLxGENE collection `d5d0df8f-4eee-49d8-a221-a288f50a1590`; Sci Data 11:1274[10]), was not retrieved because the CellxGENE collection contains non-SNpc regions only (medial globus pallidus, primary motor cortex, primary visual cortex, dorsal motor nucleus of vagus, prefrontal cortex); no SNpc slab is present in the current collection. The Smajic atlas (GSE157783[4]) was not available in the local data copy.

TF network. CollecTRI, a manually curated TF–target interaction database derived from the Lambert 2018 TF census[23], was used as the regulon dictionary (1,185 TFs, 42,990 edges), accessed via DecouplerPy[18].

4.2 Pre-processing

Raw count matrices were loaded from h5ad files using AnnData 0.10.9. Quality control applied per-sample thresholds (minimum 200 genes per cell, minimum 3 cells per gene), followed by log-normalization to 10,000 counts per cell and HVG selection (top 3,000 genes). PCA (50 components) and Leiden clustering[20] (leidenalg 0.11.0) were performed per-atlas using scanpy 1.11.5[17]. DA neurons were identified by Leiden cluster enrichment using an 8-marker panel (TH, SLC6A3, DDC, ALDH1A1, LMX1B, KCNJ6, EN1, PITX3): clusters in the top quartile by mean marker expression were retained. Raw counts were explicitly propagated to the DA-cell AnnData slice via `_ensure_raw_counts()` (`data_loader.py`) to avoid a silent layer-transfer error in AnnData $\geq$ 0.10.

4.3 TF activity scoring

Pseudobulk donor matrices (sum of raw counts per donor) were constructed from the DA-cell slices. TF activity was estimated by Univariate Linear Model (ULM) using DecouplerPy 2.1.6[19] with the CollecTRI network. ULM regresses the target-gene expression profile of each donor onto each TF's CollecTRI target signature. The resulting ULM scores (one per donor $\times$ TF) were z-scored across donors for visualization.

4.4 Differential activity and regulon selection

Differential TF activity (PD vs control) was quantified by Mann-Whitney U test (scipy 1.10.1[25]) applied to pseudobulk ULM score vectors; limma-voom was pre-registered but not available (rpy2/R not installed). p-values were corrected by Benjamini-Hochberg FDR (statsmodels 0.14.6[27]). The top-20 TFs by nominal p-value were selected as the discovery regulon list and locked via SHA-256 hash (`preregistration.json` in `001_plan/`) before any replication atlas was opened.

4.5 Cross-atlas replication (SRP-v1)

The locked top-20 TF list was scored zero-shot in the Wang atlas using identical ULM scoring (no refitting). For each TF, the sign of the mean ULM score difference (PD minus control) was compared between discovery and replication atlases. Same-sign replication count was tested against the null of Bernoulli(0.5) per TF using a two-sided binomial survival function (`scipy.stats.binom`). A permutation null was generated by randomly shuffling the CollecTRI regulon gene lists (preserving TF count and target-set sizes; 1,000 permutations) and re-counting same-sign agreements for each permuted dictionary.

4.6 Model construction and lockdown

A random forest classifier (scikit-learn 1.8.0; RandomForestClassifier, 500 trees, `random_state=42`) was trained on Adams donors (top-50 TF ULM scores as features) and evaluated on Wang donors (leave-one-atlas-out; AUC by area under ROC). Bootstrap 95% CI for AUC was computed from 1,000 bootstrap resamples (numpy 1.26.4, `random_state=0`).

4.7 Leiden-resolution robustness (H5)

DA neuron identification was repeated independently at three Leiden resolutions ($\gamma \in \{0.4, 0.8, 1.2\}$) in the Adams atlas. The top-20 TF list was re-derived at each resolution, and same-sign agreement in the Wang atlas was re-counted. The maximum swing (max_count − min_count across resolutions) is reported.

4.8 Statistical analysis

All random seeds: bootstrap resampling (numpy seed 0), random forest (sklearn seed 42). Mann-Whitney U: two-sided, applied to pseudobulk vectors (n=4 PD, n=9 ctrl). FDR correction: Benjamini-Hochberg. Permutation p-values are reported as observed fractions over 1,000 permutations. The pre-registration lockdown SHA-256 hash `d19f81821502d0782d7f23504ac2e74f43b2363066de53748eafa7914a302cf1` was computed over the pre-registered hypothesis specification and TF ranking criterion before opening any replication atlas.

4.9 Figure generation

Figures were generated by `code/figures/fig1_schematic.py`, `code/figures/fig2_discovery.py`, `code/figures/fig3_replication.py`, `code/figures/fig4_gwas.py`, `code/figures/fig5_classification.py`, and `code/figures/fig6_robustness.py`. All figures are saved as PNG and PDF to `004_execute/figures/`.

4.10 Software versions

Python 3.11.15; anndata 0.10.9[24]; decoupler 2.1.6[19]; leidenalg 0.11.0[20]; matplotlib 3.10.9; numpy 1.26.4; pandas 1.5.3; scanpy 1.11.5[17]; scikit-learn 1.8.0[26]; scipy 1.10.1[25]; statsmodels 0.14.6[27]. Full environment: `004_execute/code/environment.lock.json`.

References

[1] Tushar Kamath, Abdulraouf Abdulraouf, S J Burris, Jonah Langlieb, Vahid Gazestani, Naeem M Nadaf, Karol Balderrama, Charles Vanderburg, and Evan Z Macosko. 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.

[2] Qi Wang, Minghui Wang, Inho Choi, Tej Bhatta, Cyril Esnault, Esha Bhattacharyya, Dhruv Bhatt, 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.

[3] Levi Adams, John S K Kauwe, Dhruv L Bhatt, et al. 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.

[4] Semra Smajic, Cesar A Prada-Medina, Zied Landoulsi, Jenny Ghelfi, Sylvie Delcambre, Carola Dietrich, Javier Jarazo, Joscha Henck, Saranya Balachandran, Sinthuja Pachchek, et al. Single-cell sequencing of human midbrain reveals glial activation and a parkinson-specific neuronal state. *Brain*, 145:964–978, 2022. doi: 10.1093/brain/awab446.

[5] Sara Aibar, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, et al. SCENIC: single-cell regulatory network inference and clustering. *Nature Methods*, 14:1083–1086, 2017. doi: 10.1038/nmeth.4463.

[6] Bram Van de Sande, Christopher Flerin, Kristofer Davie, Maxime De Waegeneer, Gert Hulselmans, Sara Aibar, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nature Protocols*, 15:2247–2276, 2020. doi: 10.1038/s41596-020-0336-2.

[7] Thomas Moerman, Sara Aibar Santos, Carmen Bravo González-Blas, Jaak Simm, Yves Moreau, Jan Aerts, and Stein Aerts. GRNBoost2 and arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics*, 35:2159–2161, 2019. doi: 10.1093/bioinformatics/bty916.

[8] Mike A Nalls, Cornelis Blauwendraat, Costanza L Vallerga, Karl Heilbron, Sara Bandres-Ciga, Diana Chang, Manuela Tan, et al. Identification of novel risk loci, causal insights, and heritable risk for parkinson's disease: a meta-analysis of genome-wide association studies. *The Lancet Neurology*, 18: 1091–1102, 2019. doi: 10.1016/S1474-4422(19)30320-5.

[9] Jonggeol J Kim, Cornelis Blauwendraat, Kajsa Brolin, Zachary D Wallen, Mike A Nalls, Masashi Fujita, et al. Multi-ancestry genome-wide association meta-analysis of parkinson's disease. *Nature Genetics*, 56:27–36, 2024. doi: 10.1038/s41588-023-01584-8.

[10] John F Fullard, Sang Eun Lee, Logan Brase, 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] Jordan W Squair, Matthieu Gautier, Claudia Kathe, Mark A Anderson, Nicholas D James, Thomas H Hutson, Sébastien Lefebvre-Jourdet, et al. Confronting false discoveries in single-cell differential expression. *Nature Communications*, 12:5692, 2021. doi: 10.1038/s41467-021-25960-2.

[12] Matthew E Ritchie, Belinda Phipson, Di Wu, Yifang Hu, Charity W Law, Wei Shi, and Gordon K Smyth. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43:e47, 2015. doi: 10.1093/nar/gkv007.

[13] Carmen Bravo González-Blas, Seppe De Winter, Gert Hulselmans, Carmen Bravo Gonzalez-Blas, and Stein Aerts. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nature Methods*, 20:1355–1367, 2023. doi: 10.1038/s41592-023-01938-4.

[14] Malte D Luecken, Maren Büttner, Kridsadakorn Chaichoompu, Anna Danese, Marta Interlandi, Michaela F Mueller, Sergei Rybakov, et al. Benchmarking atlas-level data integration in single-cell genomics. *Nature Methods*, 19:41–50, 2022. doi: 10.1038/s41592-021-01336-8.

[15] Lukas Heumos, Anna C Schaar, Christopher Lance, Anastasia Litinetskaya, Felix Drost, Luke Zappia, Malte D Lücken, et al. Best practices for single-cell analysis across modalities. *Nature Reviews Genetics*, 24:550–572, 2023. doi: 10.1038/s41576-023-00586-w.

[16] Julien Bryois, Nathan G Skene, Thomas F Hansen, Lisette J A Kogelman, Hunna J Watson, Zhihong Liu, et al. Genetic identification of cell types underlying brain complex traits yields insights into the etiology of parkinson's disease. *Nature Genetics*, 52:482–493, 2020. doi: 10.1038/s41588-020-0610-9.

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

[18] Luz Garcia-Alonso et al. CollecTRI: a comprehensive collection of transcriptional regulatory interactions. *Nucleic Acids Research*, 51:D850–D856, 2023. doi: 10.1093/nar/gkac1220.

[19] Pau Badia-i Mompel, Jesús Vélez Santiago, Jana Braunger, Celina Geiss, Daniel Dimitrov, Sophia Müller-Dott, Pedro Taus, et al. decoupler: ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances*, 2:vbac016, 2022. doi: 10.1093/bioadv/vbac016.

[20] Vincent A Traag, Ludo Waltman, and Nees Jan van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9:5233, 2019. doi: 10.1038/s41598-019-41695-z.

[21] Werner Poewe, Klaus Seppi, Caroline M Tanner, Glenda M Halliday, Patrik Brundin, Jens Volkmann, Anette E Schrag, and Anthony E Lang. Parkinson disease. *Nature Reviews Disease Primers*, 3:17013, 2017. doi: 10.1038/nrdp.2017.13.

[22] Stéphane Hunot and Etienne C Hirsch. Neuroinflammatory processes in parkinson's disease. *Annals of Neurology*, 53 Suppl 3:S49–S60, 2003. doi: 10.1002/ana.10481.

[23] Samuel A Lambert, Arttu Jolma, Laura F Campitelli, Pratyush K Das, Yimeng Yin, Mihai Albu, Xiaoting Chen, Jussi Taipale, Timothy R Hughes, and Matthew T Weirauch. The human transcription factors. *Cell*, 172:650–665, 2018. doi: 10.1016/j.cell.2018.01.029.

[24] Isaac Virshup, Danila Bredikhin, Lukas Heumos, Giovanni Palla, Gregor Sturm, Adam Gayoso, Ilia Kats, Mirjana Koutrouli, Scverse community, Philipp Angerer, et al. The scverse project provides a computational ecosystem for single-cell omics data analysis. *Nature Biotechnology*, 41:604–606, 2023. doi: 10.1038/s41587-023-01733-8.

[25] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, 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.

[26] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[27] Skipper Seabold and Josef Perktold. Statsmodels: econometric and statistical modeling with python. In *Proceedings of the 9th Python in Science Conference (SciPy 2010)*, pages 92–96, 2010. doi: 10.25080/Majora-92bf1922-011.