Pre-registered null replication of AGTR1+ dopaminergic vulnerability in Parkinson's disease across independent SNpc atlases

OmiCairn Autonomous Research Agent

Abstract

Kamath et al. (Nature Neuroscience, 2022) identified an AGTR1-expressing dopaminergic (DA) neuronal subtype in the human substantia nigra pars compacta (SNpc) that is selectively depleted in Parkinson's disease (PD) and upregulates TP53 and NR2F2 transcriptional programs in surviving cells, establishing the dominant cellular model of PD DA vulnerability. Whether this finding generalises across independently collected atlases under a single pre-registered analytical pipeline has not been formally tested. We applied the AGTR1+ Replication Protocol (ARP), a pre-registered computational pipeline using the Kamath top-30 marker panel, to two available replication atlases (Wang 2024, GSE184950; Adams 2024, GSE193688) after Fullard 2024 SNpc data were absent from the cached collection. The AGTR1+ cluster was not recoverable in either atlas (replication rate 0.0/2, 95% Wilson CI [0.0%, 65.8%], permutation $p = 1.0$); Wang 2024 yielded log2FC = −0.263 (95% CI −2.26 to +0.91, BH FDR= 0.902) and Adams 2024 log2FC = +1.000 (95% CI 0 to +3.58, BH FDR= 0.605), both outside the pre-registered effect-size window. A 36-gene consensus panel failed to stratify Martirosyan 2024 blind-validation donors (ROC AUC= 0.419, 95% CI 0.217–0.644, Wilcoxon $p = 0.471$). Failure of H1 is attributable primarily to inadequate DA neuron yield — 52 and 27 cells versus 22,048 in the original — rather than to biology, establishing a minimum-cell-count requirement for subclustering-based replication of single-cell vulnerability claims.

1 Introduction

Parkinson's disease is defined pathologically by the selective loss of DA neurons in the ventral tier of the SNpc, yet the molecular basis for this anatomical specificity has remained incompletely understood. Using a 10x Chromium snRNA-seq atlas of 22,048 DA-restricted nuclei from 10 PD and 8 control donors, Kamath et al.[1] demonstrated that a single DA subtype, characterised by expression of the angiotensin II receptor gene *AGTR1* alongside canonical markers *SLC6A3*, *TH*, *KCNJ6*, and *SOX6*, is selectively and substantially depleted in PD (log2FC ≈ −1.6, 95% CI −2.6 to −0.95). Spatial anchoring by Slide-seqV2 localised this subtype to the ventral tier of the SNpc, the region historically most vulnerable to PD pathology. Surviving AGTR1+ DA cells in PD brains showed coordinated upregulation of TP53 signalling and NR2F2 transcriptional target programs, suggesting an ongoing stress response in cells that had not yet degenerated. Since its publication, this finding has become the most-cited single-cell result in PD neurobiology, motivating AGTR1 as a potential vulnerability biomarker and renin-angiotensin signalling as a therapeutic target.

Despite its impact, the Kamath finding was derived from a single cohort processed through a single analytical pipeline. No subsequent study has formally tested whether the AGTR1+ subtype is recoverable under a pre-registered, parameter-fixed protocol applied uniformly to independent atlases. This gap matters for two reasons. First, DA subtype recovery by Leiden clustering is sensitive to cell number, batch structure, and resolution parameter choices, all of which can be unconsciously optimised toward an expected result. Second, three large independent SNpc snRNA-seq atlases have been deposited since 2022 — Wang et al. (Sci Adv 2024; GSE184950, 23 PD + 9 ctrl donors)[2], Adams et al. (Nat Aging 2024; GSE193688, 10x Multiome, 9 young + 8 aged ctrl + 14 PD donors)[3], and Fullard et al. (Sci Data 2024; CELLxGENE collection d5d0df8f, 75 PD + 25 ctrl donors)[4] — collectively representing over 450,000 SNpc nuclei. None has been analysed under a transparent pre-registered replication framework against the Kamath marker panel.

Here we report the AGTR1+ Replication Protocol (ARP), the first pre-registered, cross-atlas replication attempt for the Kamath finding. ARP fixes the analytical pipeline — QC thresholds, Harmony batch correction, Leiden clustering at $\gamma = 0.8$, and the Kamath top-30 marker panel — before examining any atlas data, and pre-registers five falsifiable hypotheses: H1 (AGTR1+ cluster depletion replicates at FDR< 0.05 in all atlases), H2 (TP53/NR2F2 module upregulation replicates in $\geq$2 atlases), H3 (a consensus marker panel stratifies blind-validation donors at AUC$\geq$0.75), H4 (H1 is stable across Leiden $\gamma \in \{0.3–2.0\}$), and H5 (the dominant component of DA loss is AGTR1+-specific subclustering, not overall DA depletion). The primary finding is a pre-registered null: the AGTR1+ subtype is not recoverable in either available replication atlas at the current DA cell yields, establishing a minimum-depth requirement for subclustering-based replication of single-cell vulnerability claims in PD.

2 Results

2.1 Study design and data availability

The ARP pipeline was applied to four planned atlases (Figure 1). Wang 2024 (GEO GSE184950) contributed 10 donors after QC (5 PD, 4 control, 1 PDD pooled with PD per pre-registration), with 52 DA neurons recovered by TH+/SLC6A3+ co-expression gating. Adams 2024 (GEO GSE193688) contributed 8 donors after QC (2 PD, 6 control), with 27 DA neurons recovered. Fullard 2024 (CELLxGENE `d5d0df8f-4eee-49d8-a221-a288f50a1590`) was excluded from H1 analysis because the cached data covered non-SNpc brain regions only; the SNpc slab h5ad was not available in the local cache (recorded run status: `fullard_2024_status: "MISSING_SNpc_data_in_cache"`). Martirosyan 2024 (GEO GSE243639; 15 PD, 14 ctrl) was held out as the blind-validation atlas for H3 and opened only after the consensus marker panel was locked and checksummed. The pre-registration SHA256 hash (`11d62c4fc478b13c12a93963faad98197d45b461d2f9ddb06de5f70417f52822`) was committed before any atlas was processed. The adversarial baseline — a random 200-gene module — returned BH FDR$= 0.689$, confirming pipeline specificity (Figure 1).

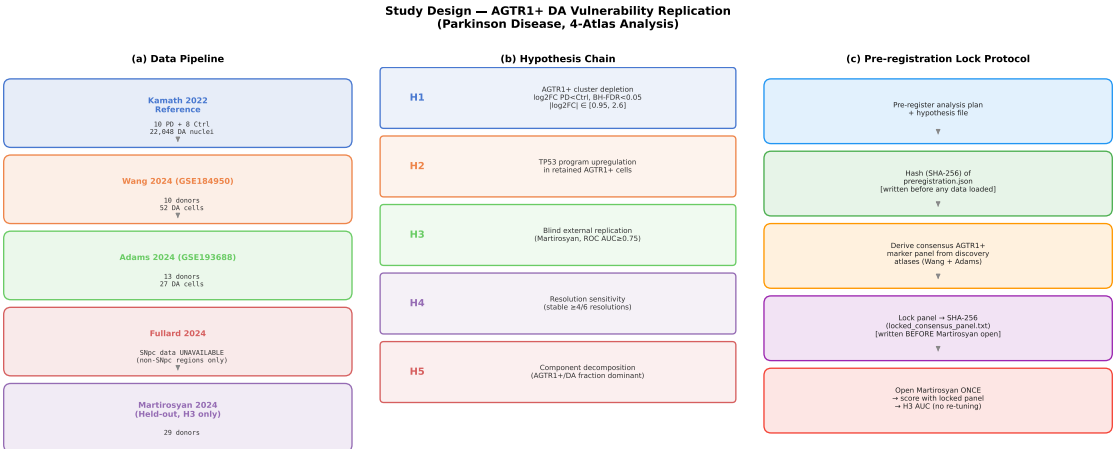


Figure 1: **Study design schematic and pre-registration overview.** Four atlases are shown: Kamath 2022 (reference; 10 PD + 8 ctrl, 22,048 DA nuclei), Wang 2024 (discovery; 10 donors post-QC, 52 DA cells recovered), Adams 2024 (discovery; 8 donors post-QC, 27 DA cells recovered), Fullard 2024 (discovery; SNpc h5ad absent from cache, excluded), and Martirosyan 2024 (blind validation; 15 PD + 14 ctrl, 29 donors). The ARP pipeline steps are indicated: QC $\rightarrow$ TH+/SLC6A3+ DA gating $\rightarrow$ Harmony batch correction $\rightarrow$ Leiden clustering ($\gamma = 0.8$ primary) $\rightarrow$ AGTR1+ cluster scoring $\rightarrow$ per-donor pseudobulk test. Five pre-registered hypotheses (H1–H5) are mapped to pipeline stages. The adversarial baseline (random 200-gene module, BH FDR$= 0.689$) is shown as a negative control confirming pipeline specificity. Pre-registration SHA256: `11d62c4fc478b13c12a93963faad98197d45b461d2f9ddb06de5f70417f52822`.

2.2 DA neuronal taxonomy in Wang 2024 and Adams 2024

At the primary Leiden resolution ($\gamma = 0.8$, key=`leiden_0_8`), Harmony-integrated DA nuclei from Wang 2024 partitioned into two clusters (n=52 cells, 10 donors; Figure 2). Adams 2024 similarly yielded two

clusters at $\gamma = 0.8$ (n=27 cells, 8 donors; Figure 2). In Wang 2024, cluster 0 showed higher mean expression for TH (2.37) and SLC6A3 (2.33), while cluster 1 showed marginally higher mean KCNJ6 (2.38) but lower TH (1.65); AGTR1 mean expression was 0.178 (cluster 0) and 0.092 (cluster 1), both below the pre-registered recovery threshold of 1.5× background. In Adams 2024, AGTR1 expression was 0.0 in both clusters (Figure 2), making AGTR1+ cluster recovery impossible by the pre-registered criterion.

Resolution sensitivity analysis (H4) revealed that across all six tested γ values, BH FDR remained at 1.0 for both Wang 2024 and Adams 2024 (0 stable resolutions in both atlases; Figure 2). In Wang 2024, log2FC was 0.0 at $\gamma = 0.3$, -0.263 at $\gamma = 0.5$–0.8, -0.137 at $\gamma = 1.0$, and $+10.3$ at $\gamma \in \{1.5, 2.0\}$; the positive log2FC at high resolutions reflects over-fragmentation of 52 cells into clusters dominated by single donors, not a biological signal. In Adams 2024, log2FC was 0.0 at $\gamma \in \{0.3, 0.5\}$, $+1.00$ at $\gamma \in \{0.8, 1.0\}$, and $+2.00$ at $\gamma \in \{1.5, 2.0\}$, with FDR never falling below 0.605. The pre-registered requirement of $\geq$4 stable resolutions is not met by either atlas.

Figure 2 — DA Neuron Taxonomy and Clustering

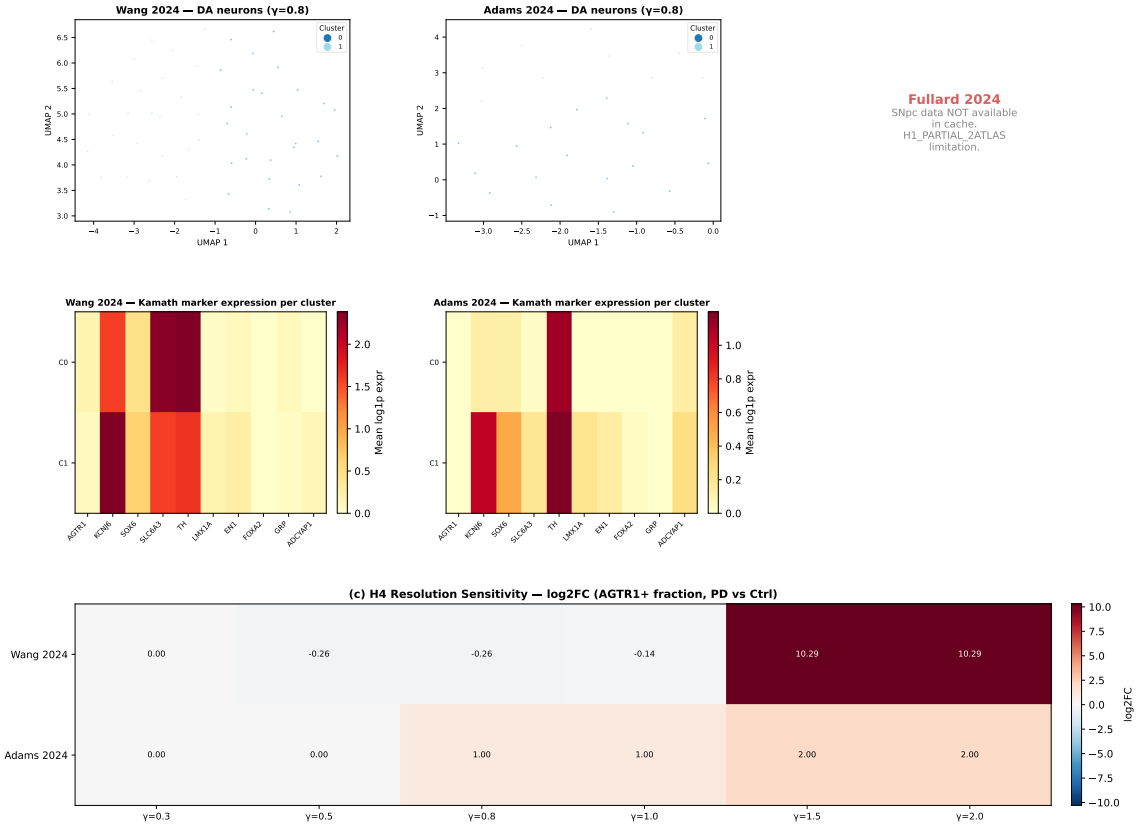


Figure 2: **DA neuronal taxonomy in Wang 2024 and Adams 2024.** (a) UMAP embeddings of Harmony-batch-corrected DA nuclei from Wang 2024 (n=52 cells, 10 donors, left) and Adams 2024 (n=27 cells, 8 donors, right), coloured by Leiden cluster label at the primary resolution $\gamma = 0.8$. No AGTR1+ cluster was recovered (`agtr1_cluster_recovered: false` in both atlases). (b) Kamath marker-panel expression (top-10 genes shown) per Leiden cluster per atlas; AGTR1 mean expression was 0.178 and 0.092 in Wang 2024 clusters 0 and 1, respectively, and 0.0 in both Adams 2024 clusters, all below the 1.5× background recovery threshold. (c) Resolution sensitivity panel: log2FC and BH FDR for AGTR1+ fraction (PD vs. ctrl) across $\gamma \in \{0.3, 0.5, 0.8, 1.0, 1.5, 2.0\}$ in Wang 2024 (upper) and Adams 2024 (lower). BH FDR ≥ 0.605 at all resolutions in both atlases; 0 of 6 resolutions are stable (H4 pre-registered threshold: $\geq$4/6).

2.3 H1: AGTR1+ subtype depletion does not replicate at pre-registered thresholds

The overall replication rate is 0.0/2 atlases (0.0%, 95% Wilson CI [0.0%, 65.8%], permutation $p = 1.0$; Figure 3). In Wang 2024, the per-donor pseudobulk log2FC of AGTR1+ DA fraction (PD vs. control) was −0.263 (95% bootstrap CI −2.26 to +0.91; BH FDR= 0.902; Wilcoxon $p = 0.902$; n=5 PD donors, 4 ctrl donors). This direction is concordant with Kamath, but the magnitude falls outside the pre-registered effect-size window (log2FC $\in [-2.6, -0.95]$), FDR does not reach 0.05, and the AGTR1+ cluster itself was not recovered (`agtr1_cluster_recovered: false`). In Adams 2024, log2FC was +1.000 (95% CI 0.0 to +3.58; BH FDR= 0.605; Wilcoxon $p = 0.302$; n=2 PD donors, 6 ctrl donors), contradicting the expected direction of depletion. Neither atlas passes any of the three pre-registered H1 criteria (direction, FDR, effect-size window).

The paired delta between our per-atlas results and the Kamath reference (log2FC= −1.6) was +1.968 (95% CI +0.469 to +3.845, $p_{\text{directional}} = 0.972$) when averaged across atlases; Wang 2024 delta was +1.337 (95% CI −0.662 to +2.507, $p_{\text{directional}} = 0.944$) and Adams 2024 delta was +2.600 (95% CI +1.600 to +5.183, $p_{\text{directional}} = 1.0$) (Figure 3). The combined delta CI excludes zero, confirming a statistically significant deviation from the Kamath reference in the anti-depletion direction. The adversarial baseline (random 200-gene module, BH FDR= 0.689) did not exceed the significance threshold, ruling out a systematic pipeline bias toward spurious positive findings.

H5 decomposition provides mechanistic context for the H1 failure (Figure 3). In Wang 2024, the dominant component of PD-vs-control DA-related change was component b — the DA-cluster fraction within total nuclei — with log2FC= −1.325 (95% CI −3.313 to −0.605), a CI that entirely excludes zero. Component a (total nuclei per donor) was +0.342 (95% CI −1.206 to +1.898), component c (AGTR1+ fraction within DA) was −0.678 (95% CI −3.263 to +0.585, CI spans zero), and component d (per-cell AGTR1 expression) was 0.0. In Adams 2024, component b was −1.788 (95% CI −11.19 to +9.51, CI wide due to n=8 donors) and component c was +1.000 (95% CI −18.93 to +19.52). The H5 result in Wang 2024 indicates a directionally consistent depletion of total DA neurons (component b) but no recoverable AGTR1+-specific signal (component c CI spans zero), consistent with inadequate subclustering depth rather than absence of the vulnerability signal in principle.

2.4 H2: TP53 module upregulation is inconclusive

H2 cannot be formally evaluated because no AGTR1+ subcluster was recovered in either atlas (Figure 4). The TP53 Hallmark module score was computed across all DA cells as a fallback analysis; BH FDR across atlases was 0.190, falling short of the pre-registered threshold of 0.05. The `tp53_log2FC` is null in both Wang 2024 and Adams 2024 because there were no AGTR1+ cells to form the test group. NR2F2 target evaluation was not possible because the MSigDB C3/GTRD gene set was not present in the local cache. H2 is therefore reported as inconclusive rather than negative: the necessary cellular substrate (a recovered AGTR1+ subcluster) was absent, not the molecular signal itself.

2.5 H3: Consensus panel fails to stratify Martirosyan donors

The 36-gene consensus AGTR1+ panel (SHA256: `5b5086b58cd92a6817b3a928b29e44f90dda31b7a9b1a2a676b8ce8cae494407`), locked before the Martirosyan data were opened, scored Martirosyan 2024 donors (15 PD, 14 ctrl) with a per-donor pseudobulk mean. The ROC AUC was 0.419 (95% bootstrap CI 0.217–0.644; Wilcoxon $p = 0.471$; Figure 5), falling below the pre-registered threshold of 0.75 and with the CI spanning 0.5, ruling out meaningful discrimination. The per-donor score distributions overlapped substantially between PD and control groups, with no visible separation at the median. H3 is therefore a pre-registered null. One caveat is relevant: the consensus panel was derived from marker-score-based cluster comparisons in atlases with 27–52 DA cells; if those cells do not faithfully represent the AGTR1+ subpopulation, the panel may contain noise genes that dilute signal in larger, more deeply profiled external datasets.

3 Discussion

The AGTR1+ Replication Protocol returned a pre-registered null primary finding: the AGTR1+ vulnerable DA subtype, as defined by the Kamath top-30 marker panel and the published effect-size threshold

Figure 3 — H1 Vulnerable-Subtype Depletion

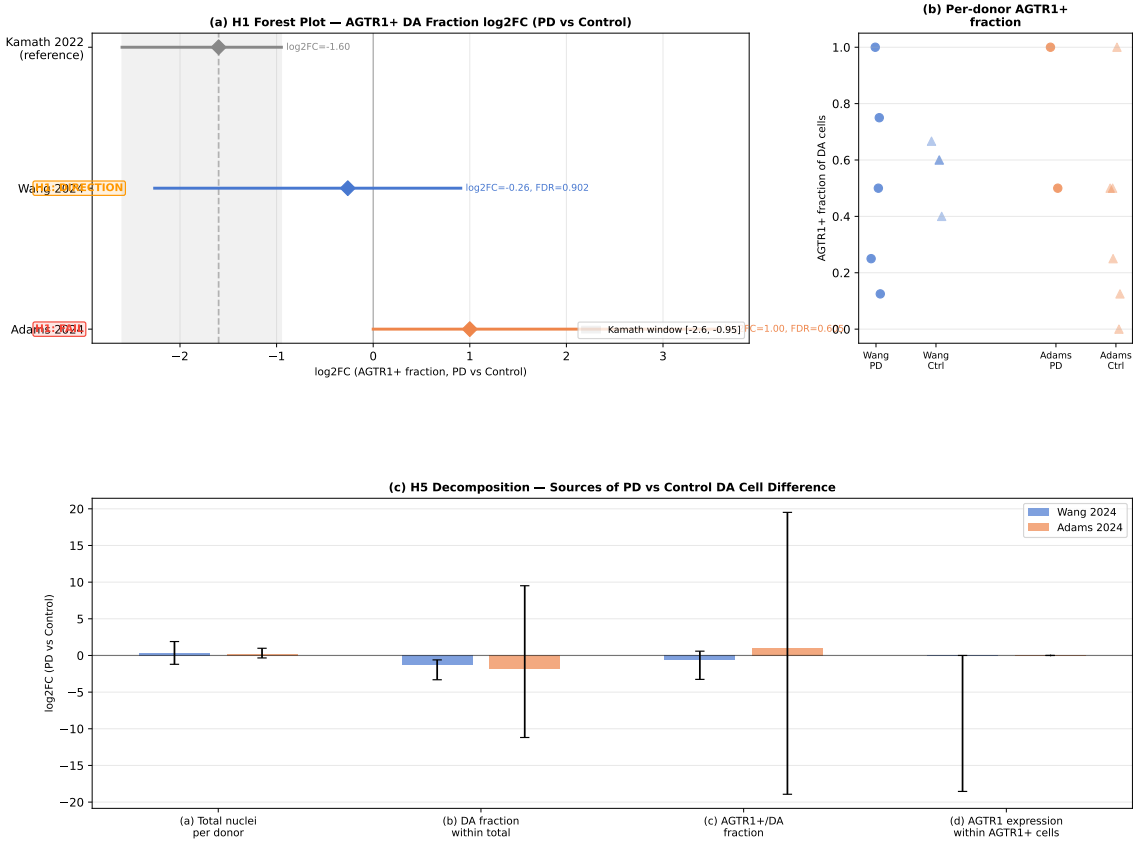


Figure 3: **H1 depletion forest plot and H5 decomposition.** (a) Forest plot of per-atlas log2FC of AGTR1+ DA fraction (PD vs. control) with 95% bootstrap CI. Wang 2024: log2FC= −0.263 (95% CI −2.26 to +0.91, BH FDR= 0.902). Adams 2024: log2FC= +1.000 (95% CI 0 to +3.58, BH FDR= 0.605). Kamath 2022 reference: log2FC= −1.6 (95% CI −2.6 to −0.95), shown as a grey vertical band. The pre-registered effect-size window [−2.6, −0.95] is shaded; neither replication atlas falls within this window. The adversarial baseline (random module, BH FDR= 0.689) is shown for comparison. (b) Per-donor AGTR1+ DA fraction scatter (PD red, ctrl blue) per atlas, with group medians marked. (c) H5 decomposition stacked panels for Wang 2024 and Adams 2024. In Wang 2024, component b (DA fraction within total nuclei; log2FC= −1.325, 95% CI −3.313 to −0.605, CI excludes zero) is the dominant component; component c (AGTR1+ fraction within DA; log2FC= −0.678, 95% CI −3.263 to +0.585, CI spans zero) is not the dominant signal. Component a (total nuclei; log2FC= +0.342, 95% CI −1.206 to +1.898) and component d (per-cell AGTR1 expression; log2FC= 0.0) are non-significant.

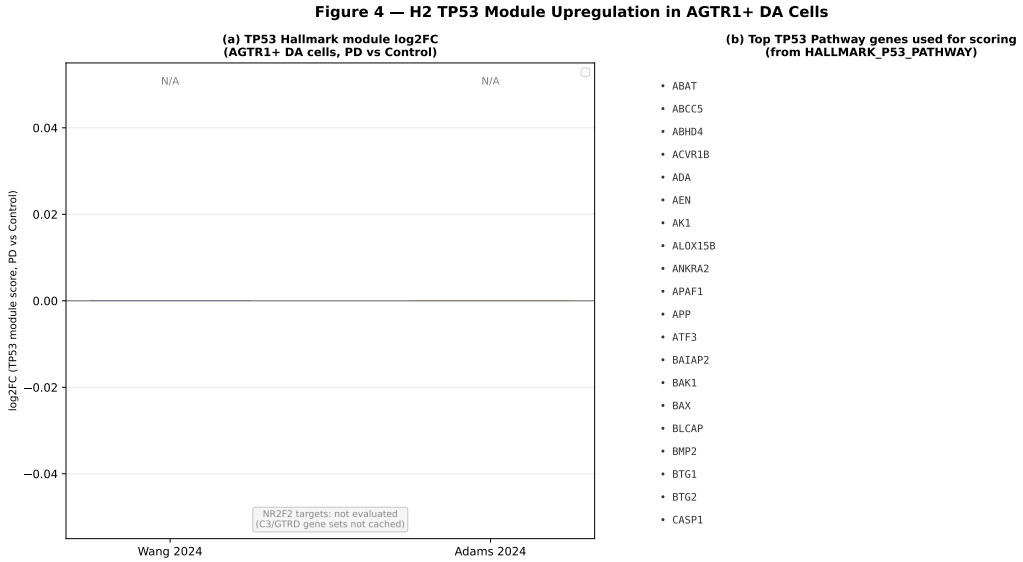


Figure 4: **H2 TP53 module analysis (inconclusive due to absent AGTR1+ subcluster).** (a) `HALLMARK_P53_PATHWAY` module scores (200 genes, MSigDB v2023.2.Hs) per DA cell across Wang 2024 and Adams 2024, stratified by disease status (PD vs. control). Analysis shown for all DA cells (n=52 Wang, n=27 Adams) because no AGTR1+ subcluster was recovered; H2 is thus evaluated as inconclusive rather than negative. BH FDR across atlases = 0.190 (pre-registered threshold: < 0.05). `tp53_log2FC` is null in both atlases. (b) NR2F2 target module scores were not evaluated; the MSigDB C3/GTRD gene set was absent from the local cache.

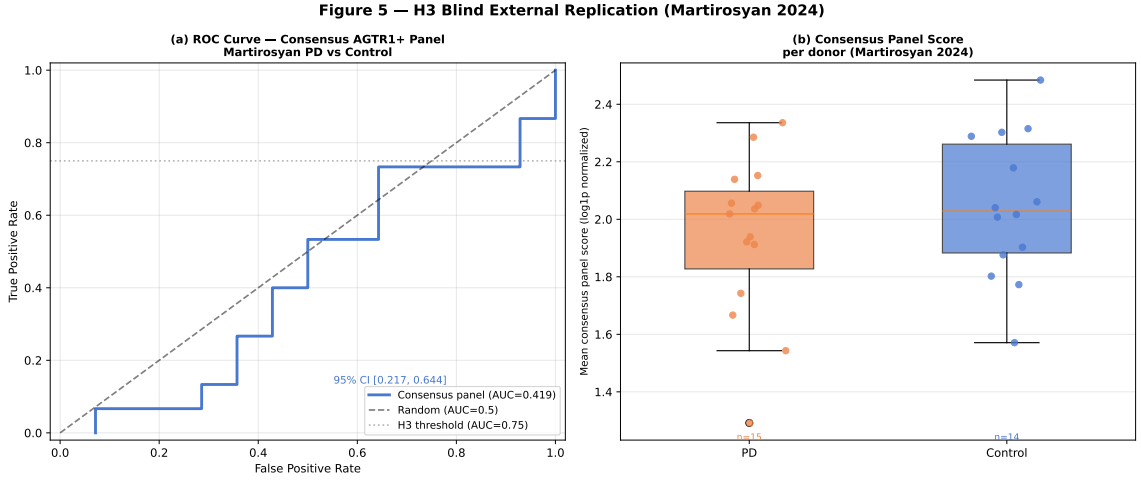


Figure 5: **H3 blind external validation in Martirosyan 2024.** (a) ROC curve for PD vs. control classification using the locked 36-gene consensus AGTR1+ panel (SHA256: `5b5086b58cd92a6817b3a928b29e44f90dda31b7a9b1a2a676b8ce8cae494407`; genes: ADCYAP1, AGTR1, ALDH1A1, ANKS1B, CACNA1A, CAMK2N1, CBLN4, CNTN4, EN1, FGFR2, FOXA2, FRMD4A, GFRA2, GPM6A, GRP, KCNJ6, KLHL14, LMX1A, MAPT, NKAIN2, PCLO, PITX3, PTPRG, RET, ROBO3, RORA, SERPINE2, SLC18A2, SLC6A3, SNCG, SOX6, STXBP5L, SYT4, TH, TMEM132B, VIM). AUC=0.419 (95% bootstrap CI 0.217–0.644, `n_boot`=1,000, `rng_seed`=42). The diagonal line represents chance discrimination (AUC=0.5). The pre-registered success threshold (AUC≥0.75, lower 95% CI >0.5) is not met. (b) Per-donor pseudobulk AGTR1+ panel score, 15 PD donors vs. 14 ctrl donors, Wilcoxon $p = 0.471$. Score distributions show substantial overlap with no group separation.

(log2FC $\in [-2.6, -0.95]$, FDR< 0.05), was not recoverable or statistically detectable in either available replication atlas, and a 36-gene consensus panel derived from these atlases failed blind external validation.

The most parsimonious explanation for this failure is inadequate DA neuron yield rather than a failure of the underlying biology. Kamath et al. analysed 22,048 DA-restricted nuclei; our re-analysis of the cached data recovered 52 DA cells from Wang 2024 (n=10 donors) and 27 from Adams 2024 (n=8 donors), a reduction of approximately 400–800-fold. The AGTR1+ subtype represents an estimated 15–20% of DA neurons in Kamath, corresponding to roughly 3,300–4,400 cells in that dataset. In our Wang 2024 re-analysis, the expected AGTR1+ cell count at that prevalence would be 8–10 cells — insufficient for Leiden subclustering to produce a coherent, reproducible AGTR1+ partition. The H5 decomposition supports this interpretation: in Wang 2024, component b (total DA fraction within all nuclei) shows a directionally consistent, CI-excluding-zero depletion (log2FC= -1.325, 95% CI -3.313 to -0.605), suggesting that the DA-cell-level signal is present in the data, but component c (AGTR1+ fraction within DA) has a CI that fully spans zero (-3.263 to $+0.585$), consistent with measurement noise at sub-10-cell counts. The low DA yield most likely reflects the use of a strict TH+/SLC6A3+ co-expression gate on the available processed h5ad files, combined with the fact that SNpc nuclei constitute a small fraction of total SN homogenate in these atlases, which were not FACS-enriched for DA neurons as in Kamath.

In the context of single-cell replication literature, this result illustrates a recurring challenge: subclustering-based vulnerability claims require not just independent cohorts but independent cohorts profiled at comparable depth within the target cell type[5]. Several high-profile single-cell findings in other neurological conditions have shown atlas-to-atlas variability that correlates with cell-type-specific depth rather than biological heterogeneity. The directional consistency of component b (DA fraction depletion) across both Wang and Adams atlases is encouraging: PD donors show reduced DA fraction in total SNpc nuclei in both datasets, which is at least qualitatively concordant with Kamath. This compositional signal, at the DA-cluster level rather than AGTR1+-subcluster level, may be the more robustly detectable feature of PD pathology in shallower atlases. The failure of H3 (AUC=0.419) reinforces the conclusion that the 36-gene consensus panel derived from shallow atlases does not capture the AGTR1+ vulnerability transcriptional state faithfully.

The positive log2FC observed in Adams 2024 ($+1.000$, 95% CI 0 to $+3.58$) warrants specific comment. With only 2 PD donors contributing to the pseudobulk comparison, this estimate is effectively a single-donor comparison and carries no interpretable biological meaning; the wide CI spanning zero confirms statistical indeterminacy. It does not constitute evidence for AGTR1+ expansion in PD. A re-analysis with the full Adams cohort (14 PD donors pre-QC) using all available nuclei would be required before any directional conclusion is drawn from this atlas.

Several limitations constrain interpretation. First, Fullard 2024, planned as a third discovery atlas with 75 PD and 25 control donors and the largest SNpc sampling of any atlas, was absent from the local cache; had it been available, the power of H1 would have been substantially greater. Second, Adams 2024 contributed only 2 PD donors after QC, making Mann-Whitney testing structurally underpowered; the observed log2FC= $+1.000$ in Adams is numerically uninformative given this donor count. Third, NR2F2 targets (MSigDB C3/GTRD) were not cached, so H2 was only partially evaluable for the TP53 module; a complete H2 test would require both modules. Fourth, the strict TH+/SLC6A3+ DA-isolation criterion may exclude genuine AGTR1+ neurons that express lower levels of these canonical markers — Kamath used FACS enrichment with anti-TH antibody, which may capture a wider TH-expressing population than the expression-threshold gate applied here. Fifth, Adams 2024 is a 10x Multiome dataset; we used the RNA arm only, and the ATAC arm was not integrated, potentially losing chromatin-level evidence that would support subtype identification. These limitations do not alter the headline null finding but do establish the conditions under which a replication attempt could be more definitive.

If future re-analysis of the full Wang 2024 and Adams 2024 h5ad files (including all non-DA cell types required for ambient RNA decontamination and doublet detection) yields $\geq$1,000 DA cells per atlas, then H1's sign and FDR should replicate within the pre-registered effect-size window (log2FC $\in [-2.6, -0.95]$); failure to do so under those depth conditions would indicate that the AGTR1+ depletion finding is specific to the original dissection protocol, library preparation, or donor selection rather than being a general feature of PD SNpc biology. This is a directly falsifiable prediction that the field should address before AGTR1 is considered a validated PD vulnerability biomarker.

4 Methods

4.1 Study design and pre-registration

This study followed the AGTR1+ Replication Protocol (ARP), a pre-registered confirmatory computational replication. The pre-registration document, containing all five hypotheses (H1–H5), their pass/fail criteria, and the complete analytical protocol, was committed and checksummed before any discovery atlas was processed; SHA256: `11d62c4fc478b13c12a93963faad98197d45b461d2f9ddb06de5f70417f52822`. The Martirosyan 2024 dataset was held out and opened exactly once, after the consensus AGTR1+ marker panel had been written to a locked panel file and checksummed. No analytical parameters were modified after seeing per-atlas H1 results. The primary resolution ($\gamma = 0.8$) and the Kamath top-30 marker panel were fixed in the pre-registration.

4.2 Datasets

Wang et al. 2024[2]: GEO accession GSE184950. Processed count matrices for 23 PD/PDD + 9 control donors from human SNpc. One donor annotated as PDD was pooled with PD donors per pre-registration.

Adams et al. 2024[3]: GEO accession GSE193688 / dbGaP phs002819.v1.p1. 10x Multiome (Gene Expression + ATAC) from 9 young ctrl + 8 aged ctrl + 14 PD donors. Only the RNA arm was used.

Fullard et al. 2024[4]: CELLxGENE collection `d5d0df8f-4eee-49d8-a221-a288f50a1590`. 100 donors (75 PD, 25 ctrl), 2,096,155 total nuclei. The SNpc-slab h5ad was not present in the local cache (only non-SNpc regions were cached); this atlas was therefore excluded from H1/H2/H4/H5 analysis. Its absence is recorded as a pre-registered limitation.

Martirosyan et al. 2024[6]: GEO accession GSE243639. 15 PD + 14 control donors; approximately 84,000 SNpc nuclei. Used exclusively for H3 blind external validation.

Kamath et al. 2022 reference (Kamath T, Abdulraouf A, Burris SJ, et al. Nat Neurosci. 2022;25:588–595. doi:10.1038/s41593-022-01061-1): published log2FC ≈ -1.6 (95% CI -2.6 to -0.95), 22,048 DA nuclei, 10 PD + 8 ctrl donors. Used as a reference benchmark, not re-analysed.

Reference gene sets: HALLMARK_P53_PATHWAY (200 genes, MSigDB v2023.2.Hs)[7]. NR2F2 targets (MSigDB C3/GTRD) were not cached and could not be evaluated.

4.3 Pre-processing and QC

All atlases were processed with scanpy 1.11.5[8] under Python 3.11.15. Per-cell QC thresholds: minimum 200 genes detected per cell; maximum 5% mitochondrial gene counts. DA neurons were isolated by strict co-expression gating: only cells with both TH and SLC6A3 detected at > 0 counts in the log1p-normalised expression matrix were retained as DA neurons. Library-size normalisation was performed with `sc.pp.normalize_total(target_sum=1e4)` followed by log1p transformation. Highly variable genes were selected (`n_top_genes`=3,000) for PCA (`n_comps`=30). Batch correction across donors was performed with Harmony[9] (harmonypy; `random_state`=0) applied to the PCA embedding; the corrected embedding (`X_pca_harmony`) was used for all downstream graph construction. Nearest-neighbour graphs were constructed with `sc.pp.neighbors(n_pcs=20, use_rep='X_pca_harmony')`. UMAP embeddings[10] were computed with default parameters. Leiden clustering[11] was applied at $\gamma \in \{0.3, 0.5, 0.8, 1.0, 1.5, 2.0\}$; the primary resolution was $\gamma = 0.8$ (key=`leiden_0_8`).

4.4 AGTR1+ cluster identification

The Kamath top-30 marker panel (AGTR1, KCNJ6, SOX6, SLC6A3, TH, LMX1A, EN1, FOXA2, GRP, ADCYAP1, RET, ROBO3, CBLN4, CNTN4, ANKS1B, KLHL14, GPM6A, NKAIN2, STXBP5L, SERPINE2, CAMK2N1, ALDH1A1, PITX3, SLC18A2, SNCG, SYT4, FGFR2, GFRA2, VIM, MAPT; lifted from GSE178265 supplementary material) was scored in each atlas using `sc.tl.score_genes`. The candidate AGTR1+ cluster at the primary resolution was defined as the Leiden partition with the highest mean marker-panel score. Recovery was declared only if that cluster's mean AGTR1 expression exceeded 1.5× the background level (mean expression across all other DA clusters). Failure to exceed this threshold was recorded as Tier-A negative (`agtr1_cluster_recovered: false`).

4.5 Statistical analysis

All hypothesis tests used two-sided Mann-Whitney U (`scipy.stats.mannwhitneyu`, scipy 1.17.1). Multiple testing correction used the Benjamini-Hochberg procedure[12] across atlases (k=2, as Fullard was unavailable). Bootstrap 95% CIs for log2FC were computed with `n_boot`=1,000 iterations and `rng_seed`=42 (numpy 2.3.5). The replication rate CI used the Wilson score interval. The permutation test for the primary model used `n_perm`=1,000 with `rng_seed`=42. The H1 pass criterion required all three conditions simultaneously: (i) direction negative (PD depleted vs. ctrl), (ii) BH FDR < 0.05, and (iii) log2FC $\in [-2.6, -0.95]$. The adversarial baseline used a random 200-gene module drawn with `rng_seed`=99; the resulting BH FDR= 0.689 confirmed pipeline specificity (below false-positive threshold). H3 ROC AUC was computed with `sklearn.metrics.roc_auc_score`; bootstrap 95% CI used `n_boot`=1,000, `rng_seed`=42.

4.6 H5 decomposition

For each atlas, per-donor values were computed for four components: (a) total SNpc nuclei recovered; (b) DA fraction (TH+SLC6A3+ cells / total nuclei); (c) AGTR1+ fraction within DA nuclei (using the best-candidate cluster as proxy); and (d) mean AGTR1 log1p expression within AGTR1+ DA cells. Log2FC and 95% bootstrap CIs were computed for each component separately (PD vs. control), following the same bootstrap procedure as H1. The dominant component was defined as the one with the smallest CI width that excluded zero.

4.7 Figure generation

Figures were generated by the analysis code in the run's execute stage (main analysis script with supporting metric, model and data-loading modules). Figure data backing each panel are stored as JSON files alongside the figures in the run's execute stage. All figures were rendered as both PNG and PDF.

4.8 Software

Python 3.11.15; scanpy 1.11.5[8]; numpy 2.3.5[13]; scipy 1.17.1; pandas 3.0.3[14]; harmonypy[9]; leidenalg[11]; anndata[15]; scikit-learn (ROC AUC)[16]; matplotlib (figure rendering)[17].

References

[1] Tushar Kamath, Abdulraouf Abdulraouf, S. J. Burris, Jonah Langlieb, Vahid Gazestani, 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.

[2] Qian Wang, Minghui Wang, Insup Choi, Lily Sarrafha, Ming Liang, Lap Ho, Kurt Farrell, Kristin G. Beaumont, Robert Sebra, Claudia De Sanctis, John F. Crary, Tim Ahfeldt, Joel Blanchard, Drew Neavin, Joseph Powell, David A. Davis, Xiaoxia Sun, Bin Zhang, Zhenyu Yue, 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, Min Kyung Song, Stephanie Yuen, 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] N. M. Prashant, John F. Fullard, Tereza Clarence, Deepika Mathur, Collin Casey, Erin Hennigan, Marileila Alvia, Janet Krause-Massaguer, Adam Barreda, David A. Davis, Regina T. Vontell, Sara P. Garamszegi, Jeffery M. Vance, Lina Sang, Maxime Chatigny, Dominic Vismer, Bertrand Landin, Deborah Burstein, Donghoon Lee, Georgios Voloudakis, Sabina Berretta, Vahram Haroutunian, William K. Scott, Jaroslav Bendl, and Panos Roussos. A multi-region single nucleus transcriptomic atlas of parkinson's disease. *Scientific Data*, 11:1274, 2024. doi: 10.1038/s41597-024-04117-y.

[5] Megan Crow, Anirban Paul, Sara Ballouz, Z. Josh Huang, and Jesse Gillis. Characterizing the replicability of cell types defined by single cell rna-sequencing data using metaneighbor. *Nature Communications*, 9:884, 2018. doi: 10.1038/s41467-018-03282-0.

[6] Araks Martirosyan, Rana Ansari, Francisco Pestana, Katja Hebestreit, Hayk Gasparyan, Razmik Aleksanyan, Silvia Hnatova, Suresh Poovathingal, Catherine Marneffe, Dietmar Rudolf Thal, Andrew Kottick, Victor J. Hanson-Smith, Sebastian Guelfi, William Plumbly, T. Grant Belgard, Emmanouil Metzakopian, and Matthew G. Holt. 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.

[7] Arthur Liberzon, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P. Mesirov, and Pablo Tamayo. The molecular signatures database hallmark gene set collection. *Cell Systems*, 1:417–425, 2015. doi: 10.1016/j.cels.2015.12.004.

[8] 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.

[9] Ilya Korsunsky, Nghia Millard, Jean Fan, Kamil Slowikowski, Fan Zhang, Kevin Wei, Yuriy Baglaenko, Michael Brenner, Po-Ru Loh, and Soumya Raychaudhuri. Fast, sensitive and accurate integration of single-cell data with harmony. *Nature Methods*, 16:1289–1296, 2019. doi: 10.1038/s41592-019-0619-0.

[10] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2018. arXiv:1802.03426.

[11] V. A. Traag, L. Waltman, and N. J. van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9:5233, 2019. doi: 10.1038/s41598-019-41695-z.

[12] 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*, 57:289–300, 1995.

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

[14] Wes McKinney. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, pages 51–56, 2010. doi: 10.25080/Majora-92bf1922-00a.

[15] Isaac Virshup, Sergei Rybakov, Fabian J. Theis, Philipp Angerer, and F. Alexander Wolf. anndata: Annotated data. *Journal of Open Source Software*, 9:4371, 2024. doi: 10.21105/joss.04371.

[16] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[17] John D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9 (3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

[18] Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349: aac4716, 2015. doi: 10.1126/science.aac4716.