

NEUROSCIENCE

Molecular profiling of human substantia nigra identifies diverse neuron types associated with vulnerability in Parkinson's disease

Qian Wang^{1,2,3,4,5,6†}, Minghui Wang^{1,5,6†}, Insup Choi^{2,3,4†}, Lily Sarrafha^{2,3,4}, Marianna Liang^{2,3,4}, Lap Ho^{1,5,6}, Kurt Farrell^{3,4,7}, Kristin G. Beaumont¹, Robert Sebra^{1,6,8}, Claudia De Sanctis⁹, John F. Cray^{3,4,7,9,10}, Tim Ahfeldt^{3,11,12}, Joel Blanchard^{3,11,12}, Drew Neavin¹³, Joseph Powell^{13,14}, David A. Davis¹⁵, Xiaoyan Sun¹⁵, Bin Zhang^{1,5,6,16*}, Zhenyu Yue^{2,3,4,17*}

Copyright © 2024 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Parkinson's disease (PD) is characterized pathologically by the loss of dopaminergic (DA) neurons in the substantia nigra (SN). Whether cell types beyond DA neurons in the SN show vulnerability in PD remains unclear. Through transcriptomic profiling of 315,867 high-quality single nuclei in the SN from individuals with and without PD, we identified cell clusters representing various neuron types, glia, endothelial cells, pericytes, fibroblasts, and T cells and investigated cell type–dependent alterations in gene expression in PD. Notably, a unique neuron cluster marked by the expression of *RIT2*, a PD risk gene, also displayed vulnerability in PD. We validated *RIT2*-enriched neurons in midbrain organoids and the mouse SN. Our results demonstrated distinct transcriptomic signatures of the *RIT2*-enriched neurons in the human SN and implicated reduced *RIT2* expression in the pathogenesis of PD. Our study sheds light on the diversity of cell types, including DA neurons, in the SN and the complexity of molecular and cellular changes associated with PD pathogenesis.

INTRODUCTION

The degeneration of dopaminergic (DA) neurons in the substantia nigra (SN) is a major pathological hallmark of Parkinson's disease (PD). DA neurons regulate movement, learning, reward, and addiction. The loss of DA neurons and other types of cells in PD causes motor symptoms and can lead to psychiatric complications (1). The molecular mechanisms underlying the loss of DA neurons in the SN remains poorly understood; several hypotheses—such as dopamine toxicity, iron burden, autonomous pace-making, and axonal arborization—have been proposed to explain their vulnerability in

PD (2). However, detailed molecular and cellular dissection of human DA neurons is needed to understand the underpinning of DA neuron degeneration.

Genome-wide association studies (GWAS) have identified many genetic variants and risk alleles of PD and begun to gain insight into the molecular mechanisms of the disease (3). However, the vast majority of PD cases have no known genetic cause, and their etiology remains unclear (4, 5). To fully understand the molecular mechanisms for DA neuron degeneration, particularly in idiopathic PD, post-GWAS research should investigate cell type–specific expression and functions of PD genes and GWAS variants.

Multiple studies in rodents have profiled DA neurons and demonstrated the heterogeneity of DA neurons in the midbrain. With single-cell transcriptomic analysis, they identified several molecularly distinct DA neuron subtypes in the midbrain, suggesting diverse functions and potentially differential vulnerability of different DA neuron types in PD (6–11). By integration of GWAS and single-cell transcriptomic data from mouse brains, one study revealed an unexpected role of oligodendrocytes in PD progression (12). While challenging due to PD sample scarcity and quality, few studies have performed single-nucleus RNA sequencing (snRNA-seq) in human postmortem midbrain and identified cell clusters representing DA neurons. Their results suggested an association of common risk for PD with DA neuron-specific expression (13, 14). A previous study further examined PD midbrains by snRNA-seq and reported a disease-specific DA neuron cluster and “pan-glial” activation (15). A more recent report performed snRNA-seq analysis of *NR4A2*-enriched neurons in the human SN and identified 10 subpopulations of DA neurons that displayed differential vulnerability in PD (16). However, whether neuron types beyond DA neurons in the SN show vulnerability in PD remains to be clarified.

Unlike rodents, human DA neurons contain neuromelanin (NM; dark pigment), which increases in concentration during aging (17) and is biosynthesized from L-3,4-dihydroxyphenylalanine, a precursor

¹Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. ²Department of Neurology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. ³Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA. ⁴Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA. ⁵Mount Sinai Center for Transformative Disease Modeling, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA. ⁶Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. ⁷Department of Pathology, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA. ⁸Sema4, a Mount Sinai venture, Stamford, CT 06902, USA. ⁹Neuropathology Brain Bank & Research CoRE, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA. ¹⁰Department of Artificial Intelligence & Human Health, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, NY 10029, USA. ¹¹Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. ¹²Ronald Loeb Alzheimer's Disease Center, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. ¹³Garvan-Weizmann Centre for Cellular Genomics, Garvan Institute for Medical Research, 384 Victoria Street, Sydney 2010, Australia. ¹⁴UNSW Cellular Genomics Futures Institute, University of New South Wales, Kensington, Sydney 2052, Australia. ¹⁵Department of Neurology, Evelyn F. McKnight Brain Institute, Brain Endowment Bank, University of Miami Miller School of Medicine, Miami, FL 33136, USA. ¹⁶Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. ¹⁷The Center for Parkinson's Disease Neurobiology, Icahn School of Medicine at Mount Sinai, 1470 Madison Avenue, New York, NY 10029, USA.

*Corresponding author. Email: zhenyu.yue@mssm.edu (Z.Y.); bin.zhang@mssm.edu (B.Z.)

†These authors contributed equally to this work.

of DA. The loss of melanin-containing DA neurons in the SN has long been recognized in PD, but it is underappreciated that a subpopulation of DA neurons in the SN of PD persisted through many years after the onset of motor symptoms, suggesting resilience (18). Whether or not other types of neurons in the SN degenerate remains unclear. How gene expression and cellular functions are altered in the remaining DA neurons and other cell types in the SN of patients with PD is largely unknown.

In this study, we performed molecular profiling of human SN and midbrain organoids and identified a unique neuron subtype, marked by the expression of *RIT2*, a PD risk gene, and spatially concentrated in the SN with vulnerability in PD. Our observations of *RIT2* neuron population are validated in independent cohorts and human midbrain organoids. We also delineated cell type-specific gene expression changes in the SN in PD. Our transcriptomics data will be an important resource for the elucidation of the cellular heterogeneity in the SN and molecular mechanisms underlying the complexity of PD symptoms.

RESULTS

Cell-type composition and diversity in the human SN

We have collected the SN samples from postmortem brains of 32 donors, including 23 idiopathic PD and 9 controls with an average age of 81, and processed them for snRNA-seq analysis (sequencing cohort, Table 1 and table S1). By using 10x Genomics Chromium Single Cell 3' Solution, we obtained 457,453 droplet-based snRNA-seq profiles from these brains. Using an established Seurat-based data preprocessing (19) and clustering analysis pipeline that includes quality control, data integration by Harmony (20), cluster stability assessment, and doublets detection (detailed in Materials and Methods and fig. S1), we obtained 315,867 high-quality nuclei (248,245 of PD and 67,622 of the control) and identified 12 cell clusters (c0 to c11) (Fig. 1B), ranging in number from 134,011 (c0) to 1384 (c11). To discern the identity of each cluster, we used two complementary strategies: (i) examining the expression pattern of known gene markers of major brain cell types, such as astrocytes (*AQP4*), neurons (*SLC17A6*, *GAD1*, and *RBFOX3*), microglia (*C3* and *CSF1R*), oligodendrocytes (*MOG*), oligodendrocyte progenitor cells (*VCAN*), endothelial cells (*FLT1*), and pericytes (*PDGFRB*) (Fig. 1C) and (ii) comparing de novo cluster-specific marker gene signatures (fig. S2 and table S2) with a large-scale collection of cell-type markers curated from more than 1054 single-cell experiments (fig. S3) (21). Together, our annotation identified nine different cell types with fractions as follows: oligodendrocytes (c0 and c3; 51.3%), neurons (c6, c7, and c9; 13.1%), microglia (c1; 9.4%), astrocytes (c2; 8.4%), endothelial cells (c4; 7.0%), oligodendrocyte progenitor cells (OPC)

(c5; 6.5%), pericytes (c8; 3.1%), fibroblast-like cells (c10; 0.8%), and T cells (c11; 0.4%). Thus, in our data, we observed the largest cell type in the human SN is oligodendrocytes, followed by neurons, microglia, astrocytes, and the rest (Fig. 1D).

Identification of *RIT2*-enriched neurons in the human SN that show vulnerability in PD

Our initial clustering analysis of the combined control and PD samples identified three distinct neuron clusters, c6, c7, and c9 (Fig. 1C). First, we noticed that the relative fractions of the control versus PD samples are proportional to the ratio of their sample size (9 versus 23) in all clusters, except neuron cluster c9 (Fig. 1E). c9 displayed a disproportionate distribution of the cell fractions between the control and PD (2741 control versus 1479 PD) (Fig. 1E), suggesting a reduction of c9 cell number in PD samples. To test the loss of c9 neurons, we calculated the odds ratio and performed an immunohistochemistry (IHC) study by using two separate cohorts (table S1). In the first cohort, the control brains presented a significantly higher proportion of c9 neurons than PD brains [mean of 3% versus 0.6%, overall odds ratio (OR) = 6.6, *P* = 0.0073 by Wilcoxon rank sum test of odds] (Fig. 2, A and B) despite the smaller number of control samples sequenced. Omission of one control sample, which appears to contribute >30% of the cells in the control cluster (fig. S1B), has little effect in the outcome (mean of 2.7% versus 0.6%, overall OR = 4.3, Wilcoxon rank sum test *P* value = 0.016). We noticed that *RIT2* is a marker gene for c9 (Fig. 1C). *RIT2* encodes a neuronal guanosine triphosphatase (GTPase) and was previously identified as a PD susceptible gene based on at least two large-scale GWAS reports (3, 22). A fraction of c9 from the control (7.1%, 194 of 2741) expressed tyrosine hydroxylase (*TH*), a marker gene for DA neurons.

We then took an independent human cohort and performed IHC analysis of human SN with anti-*RIT2* or anti-*TH* antibodies. The distribution of *RIT2*⁺ neurons seem to resemble that of *TH*⁺ neurons in the SN (Fig. 2C). Furthermore, we found that ~80% fraction of *RIT2*⁺ neurons contain NM (yellow arrows and arrowheads), while ~20% (blue arrowhead) are negative for NM (Fig. 2, D and E). The *RIT2*⁺*NM*⁺ neurons can be separated into *TH*⁺ (70%, yellow arrows) and *TH*[−] (~30%, yellow arrowheads) two subpopulations (Fig. 2, D and E). Consistently, a previous study reported that 7 to 30% of *NM*⁺ neurons are negative for *TH* staining (23). We confirmed that the total number of *NM*⁺ neurons is reduced in the SN of PD (81.58% reduction; Fig. 2F). The number of *RIT2*⁺*NM*⁺ (including *TH*⁺ and *TH*[−]) and *RIT2*⁺*NM*[−] neurons are both decreased in the SN of PD (Fig. 2F), supporting the observation from the snRNA-seq analysis that c9, likely representing *RIT2*⁺*TH*[−] neurons, degenerate in PD (Fig. 1E). We have also examined additional marker genes of c9 through RNAscope in situ hybridization assay (fig. S2).

Table 1. Sequencing cohort demographic information. Data were represented by means ± SD for age and postmortem interval (PMI). NA, not applicable.		
	Control (n = 9)	PD (n = 23)
Age (years)	83.8 ± 8.3	78.8 ± 7.6
Sex: Male (%)	4 (44)	17 (73.9)
PMI (hours)	17.3 ± 6.3	19.1 ± 8.5
Braak	NA	II:4; III:3; IV:5; VI:8; NA:3

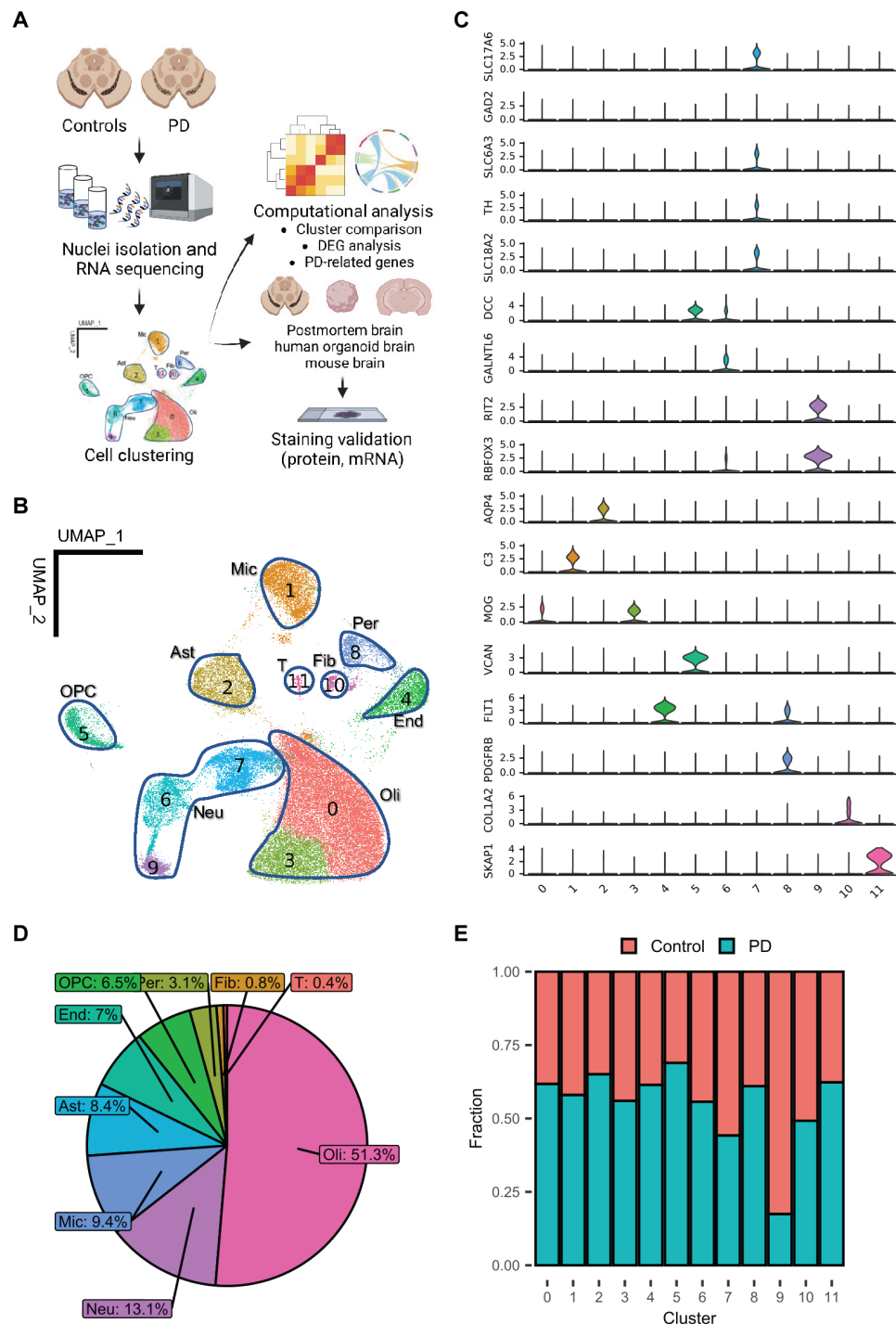


Fig. 1. Cellular diversity in the SN from patients with PD and control samples. (A) Flow chart of the experimental procedure and data processing. Barcoded single-nucleus suspension was prepared using frozen SN samples from PD and control subjects followed by RNA sequencing. Sequencing data were quality controlled and classified into cell clusters, which were annotated with known cell-type markers. Downstream analyses include cell composition changes, immunofluorescence (IF) staining, DEG identification, and cell communication alterations. (B) UMAP plot showing cell clusters. Ast, astrocytes; Mic, microglia; Oli, oligodendrocytes; OPC, oligodendrocyte progenitor cells; End, endothelial cells; Fib, fibroblast-like cells; Per, pericytes; T, T cell. (C) Expression pattern of known brain cell-type marker genes in the control cells. (D) Pie-chart for the fractions of major cell types in the human SN. (E) Stacked cell distribution represented by disease status in each cluster.

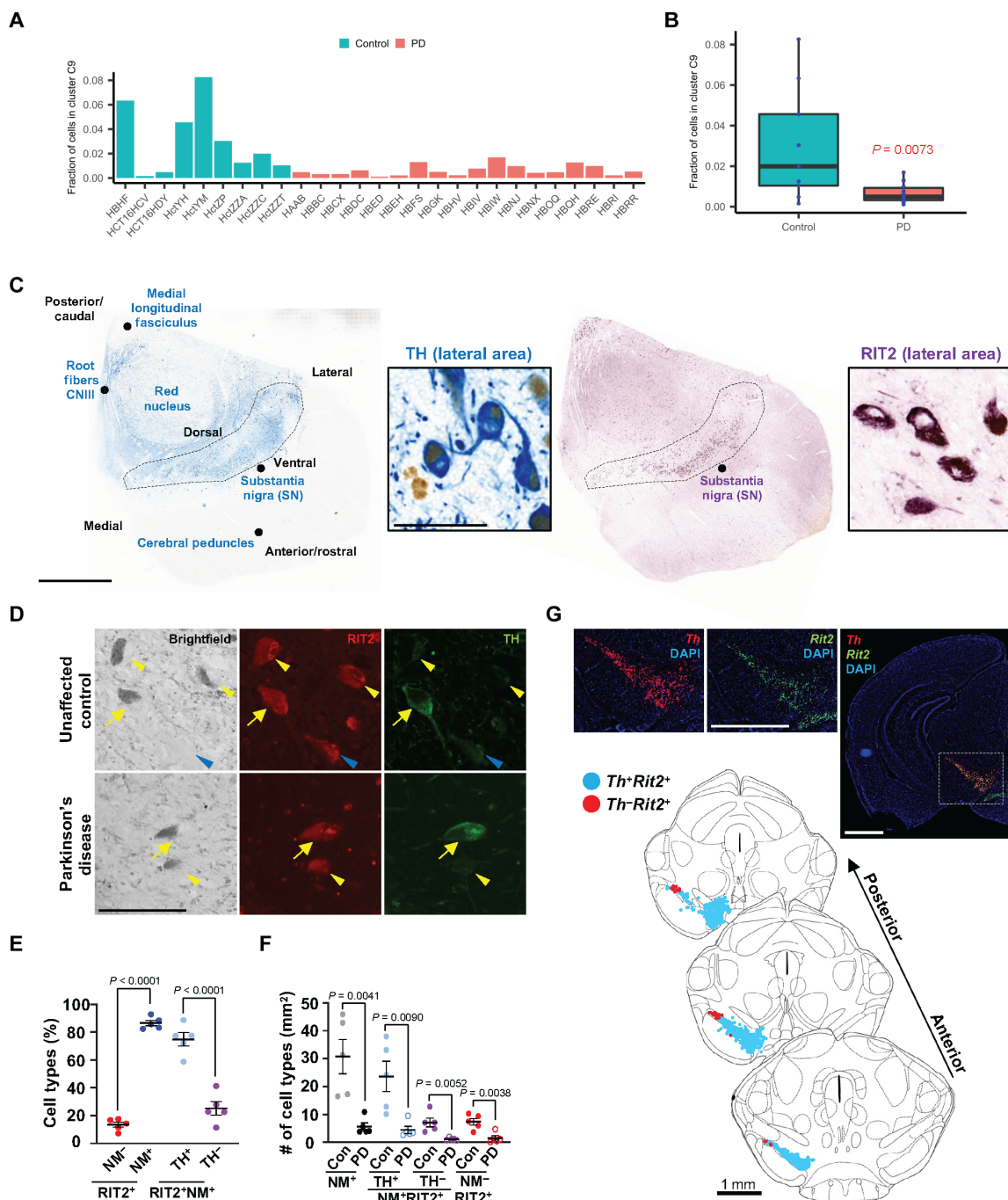


Fig. 2. Evidence of RIT2⁺ neuronal populations in health and PD brain. (A) Fraction of c9 neurons in one cohort containing PD and the controls. (B) Distribution of the fraction of c9 neurons in PD and the controls shown in (A). *P* value was computed by one-tailed Wilcoxon rank sum test. (C) IHC staining of the postmortem tissue of the human SN with anti-TH (blue, left) and -RIT2 (purple, right) antibodies. Enclosed areas are SN pars compacta (SNpc). Scale bars, 4 mm and 100 μ m in magnified images. (D) IF staining of the postmortem tissue of the human SN with anti-RIT2 (red) and anti-TH (green) antibodies. Yellow arrows, NM⁺RIT2⁺TH⁺neurons; yellow arrowheads, NM⁺RIT2⁺TH⁻neurons; blue arrowheads, NM⁻RIT2⁺TH⁻neurons. DA neurons contain neuromelanin (bright-field images). Scale bar, 100 μ m. (E) Quantification of the fractions of RIT2⁺ among NM⁻ neurons and RIT2⁺TH⁺ and RIT2⁺TH⁻ among NM⁺ neurons from five unaffected controls. (F) Quantification of the number of RIT2⁺ among NM⁻ neurons and RIT2⁺TH⁺ and RIT2⁺TH⁻ among NM⁺ neurons in the SNpc of the control (*n* = 5) and PD (*n* = 5). *P* values were calculated by unpaired two-tailed Student's *t* test. (G) RNAscope in situ hybridization assay in mouse brain. *Rit2*^{+/Th+} and *Rit2*^{+/Th-} cells were mapped on the schematic images of mouse brain. Scale bars, 1 mm.

For example, *CADPS2* is another marker gene of c9. We found the presence of *CADPS2*⁺*RIT2*⁺ cells, which are *TH*⁺ or *TH*[−] (fig. S4A), consistent with the cell cluster analysis.

Furthermore, we validated *Rit2*⁺*Th*[−] and *Rit2*⁺*Th*⁺ neurons in the SN of mouse brain via RNAscope analysis. *Rit2*⁺ cells were highly enriched in the SN and largely overlap with *Th*⁺ cells (top panels, Fig. 2G). *Rit2*⁺*Th*[−] cells are sparse in ventral SN pars compacta and clustered at lateral SN (bottom panels, Fig. 2G). Furthermore, we observed a similar distribution between *Rit2*⁺*Th*[−] and *Cadps2*⁺*Th*[−] cells at lateral SN (fig. S4B).

Together, our data suggest the presence of a transcriptomically distinct, *RIT2*-enriched neuron population (c9) in the SN, which degenerates in PD. Our observation also suggests that *RIT2*⁺ neuron populations are heterogeneous (e.g., *RIT2*⁺*TH*[−] and *RIT2*⁺*TH*⁺). The validations with IHC and RNAscope analysis showed a different proportion of *RIT2*⁺*TH*[−] versus *RIT2*⁺*TH*⁺ neurons from the snRNA-seq data, which has limitations due to the current technology.

The presence of *RIT2*-enriched neurons in the human SN and midbrain organoids

We next compared our results to the published snRNA-seq datasets of human SN from two independent cohorts, which contain only non-PD samples (13, 14). We reprocessed their original datasets using our analytic pipeline to identify subtypes of the neuron clusters (figs. S5, A to E, and S6E). In the dataset developed by Agarwal *et al.* (13), we found a *RIT2*-enriched neuron subcluster (Ac6_0), which overlaps significantly with c9 from our study and is distinguished from subcluster Ac6_1 enriched for typical DA markers such as *TH*, *SLC18A2*, and *SLC6A3* (fig. S5, B and D). Note that the *RIT2*-enriched neuron subtype Ac6_0 from Agarwal *et al.* (13) express reduced levels of *TH* and *SLC18A2* compared to Ac6_1 (fig. S5D). In the second dataset from Welch *et al.* (14), a *RIT2*-enriched neuron subcluster Wc4_5 overlaps significantly with c9 in our study and is separated from Wc4_3, which is enriched for DA markers *TH*, *SLC18A2*, and *SLC6A3* (fig. S6D). These results confirmed the presence of *RIT2*⁺ populations either *TH*^{Low} [Agarwal *et al.* (13)] or *TH*[−] [Welch *et al.* (14)], distinct from a typical DA neuron.

We expanded our investigation to midbrain organoids derived from human pluripotent stem cells (hPSCs) (24) for the evidence of *RIT2*⁺ neurons (c9) subtype of human SN. We performed single-cell RNA sequencing (scRNA-seq) analysis of the organoids and identified at least 17 distinct cell clusters (Fig. 3A and fig. S7). We observed that organoid cluster 0 (Oc0), 1 (Oc1), and 3 (Oc3) were enriched in *TH* expression, while Oc0 and Oc2 were enriched in *RIT2* production (Fig. 3, B and C, and table S3). By comparing cell clusters between human SN and midbrain organoids, we observed that the marker genes of c9 from human SN showed the greatest similarity to that of Oc1 (adjusted *P* value = 3.7×10^{-13} by hypergeometric test), followed by Oc4, Oc14, and Oc2 (Fig. 3D and table S4). We found that c6 of human SN shared a similar expression profile with organoid clusters Oc1 and Oc2 and a minor cluster Oc8 (adjusted *P* value ranging from 2.2×10^{-4} to 1.4×10^{-16}) (Fig. 3D and table S4). In contrast, c7 of human SN showed overlapping gene signatures with the organoid cluster Oc0 and two minor clusters (Oc10 and Oc14) (adjusted *P* value ranging from 5.5×10^{-4} to 1.9×10^{-4}) (Fig. 3D and table S4). The strong similarity of gene signatures of c6 and c9 with Oc1 and Oc2 may suggest that both c6 and c9 neurons emerge from a similar neuronal type during the developmental stage but diverge into two distinct populations with aging. Consistent with

the sequencing data, IHC analysis of the organoids showed many cells coexpressing *TH* and *RIT2* and cell populations producing *RIT2* but not *TH* (Fig. 3, E and F).

Identification of subtypes of DA neurons associated with vulnerability in PD

Next, to define the neuron clusters c6 and c7, we performed further clustering analysis and identified multiple subclusters in c6 (c6_0 to c6_5, while c6_4 and c6_5 were excluded from the further downstream analysis due to their limited cell numbers) (Fig. 4, A to C) and c7 (c7_0 to c7_3) (Fig. 4, D to F). The subclusters c6_2 and c7_3 were enriched for the expression of DA neuron marker genes such as *TH*, *SLC18A2*, and *SLC6A3*, indicative of typical DA neurons. Similar to c9 (Fig. 1E), the fractions in c6_2 and c7_3 appear disproportionately distributed between the control and PD according to the sequenced samples. For example, 49.7% of the total sequenced of 1654 nuclei in c6_2 were from the control (overall OR = 3.7, Wilcoxon *P* = 0.00096) (Fig. 4C), and 60.8% of the total sequenced of 4262 nuclei in c7_3 were from the control (overall OR = 5.9, Wilcoxon *P* = 0.0041) (Fig. 4F), suggesting the reduction of typical DA neuronal types (c6_2 and c7_3) in PD. We further showed distinct marker genes for c6_2 and c7_3 DA neuron subtypes (Fig. 4G and fig. S8A). Similar classifications of DA neurons were observed by the alignment of the DA neuron subtypes from a published mouse midbrain cell atlas using snRNA-seq (fig. S9) (10).

Multiple studies have demonstrated the classification of DA neurons in the mouse midbrain based on molecular profiling associated with the expression of distinct transcription factors (TFs) confined in specific subtypes of DA neurons (25). The TFs are critical for DA neuron specification or differentiation during development (26). We found that, however, most TFs were expressed in small fractions from all cell clusters (including the c6_2, c7_3, and c9) in the aged human SN (both the control and PD). *MYT1L* was an exception—it was expressed in large fractions of the main neuron clusters (fig. S10). c6_2 subtype of DA neurons showed much higher *SOX6* cell fractions (22.25%) than that of c7_3 (6.22%) or c9 (5%) subtypes.

To validate the DA neuron subtypes, we compared c6_2 and c7_3 with the clusters we identified in midbrain organoids (Fig. 4H, fig. S7E, and table S4). We found that the markers of c6_2 are very similar to those of Oc1 (adjusted *P* value = 1.7×10^{-16}) (Oc3 to a lesser degree, adjusted *P* value = 5.1×10^{-3}), in which *TH* was highly expressed (Fig. 3B). In contrast, the markers of c7_3 are similar to those of Oc0 (adjusted *P* value = 7.9×10^{-5}), which showed high *TH* expression (Fig. 3B), and other minor clusters from organoids (Fig. 4H, fig. S7E, and table S4). Thus, our analysis of midbrain organoids supports the finding of the molecularly distinct DA neuron subtypes in c6_2 and c7_3 from the human SN.

Altered landscape of cell type-specific transcriptomics in the SN of PD

Aside from the identification of diverse neuron types associated with PD vulnerability, we next sought to survey the global transcriptomic changes by examining differentially expressed genes (DEGs) between PD and the controls in each cell cluster. c9 (*RIT2*⁺*TH*[−]) and c7_2 (glutamatergic) presented with the highest numbers of DEGs, followed by c4 (endothelial cells) and c8 (pericytes) (Fig. 5A and table S5). Functional enrichment analyses revealed up-regulation of ribosomal genes and protein translation-related pathways in nearly all cell types (Fig. 5B and table S6). A broad increase of

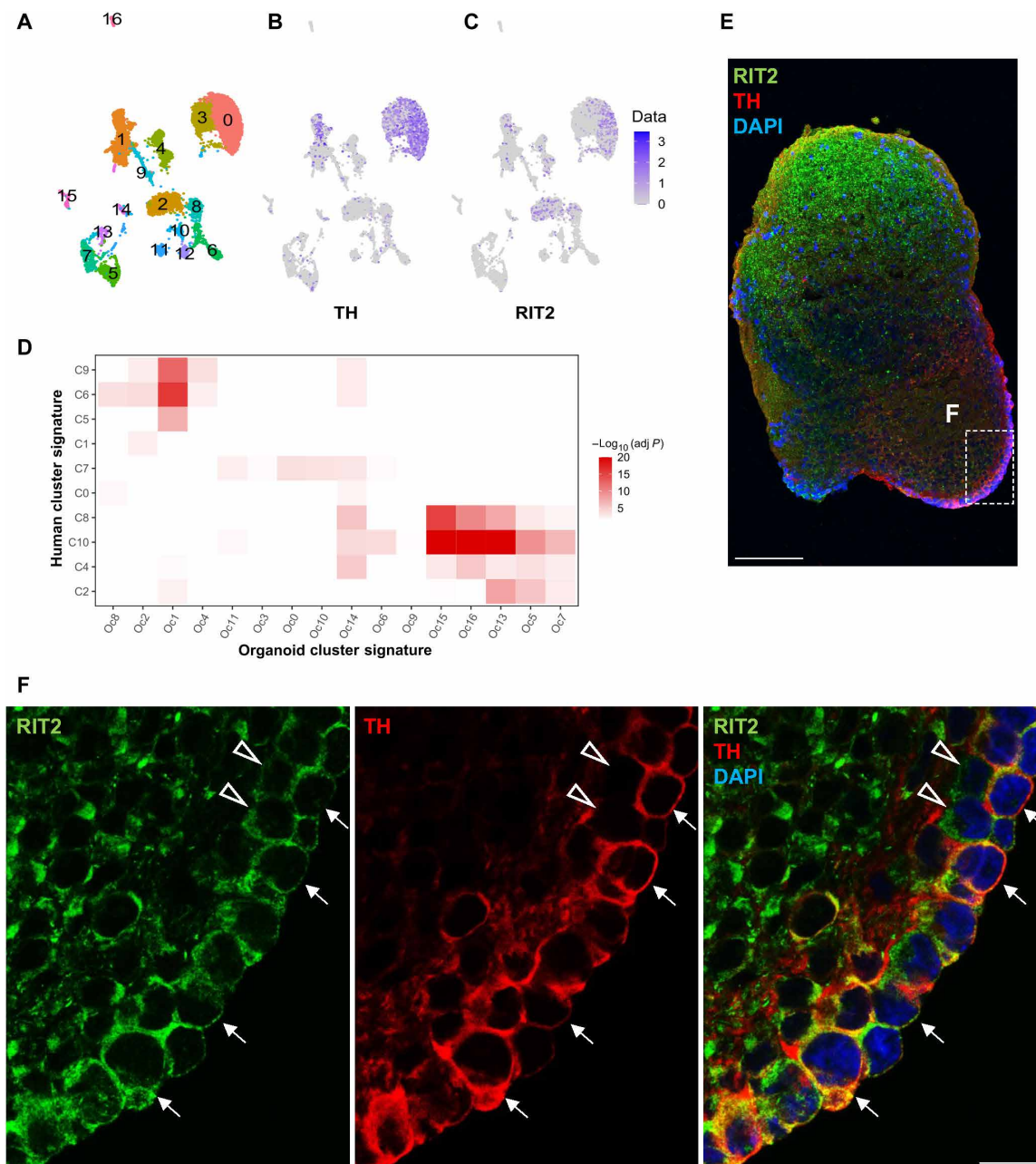


Fig. 3. scRNA-seq and immunostaining analysis of human midbrain organoids. (A) UMAP visualization of the single-cell clustering from the hiPSC-derived midbrain organoid (day 40). (B and C) UMAP visualization of the *TH* and *RIT2* gene expression. (D) Comparison between the cluster markers between the human organoids and the SN of the control samples. The minimum adjusted *P* value was set at 1×10^{-20} for visualization purpose. (E and F) IF staining of midbrain organoid with anti-*RIT2* and anti-*TH* antibodies. White arrows indicate a $RIT2^{+}TH^{+}$ cell population, and white arrowheads indicate a $RIT2^{+}TH^{-}$ cell population. Scale bars, 100 μ m (E) and 10 μ m (F).

metallothionein family genes, such as *MT2A*, *MT1E*, and *MT3*, was also found in neuronal and non-neuronal clusters in the SN of PD (Fig. 5C). The metallothionein proteins are cysteine-rich and of low molecular weight. They bind heavy metals, and some appear to play a role in detoxification and cytoprotection (27). In addition, up-regulation of several heat shock protein family members (e.g., *HSPB1*, *HSPH1*, *HSPA1*, and *HSP90AA1*) and *CRYAB* was observed in many cell types (including neuronal and non-neuronal clusters) in PD (Fig. 5C). *CRYAB*, encoding the alpha B subunit of cystallin

and a small chaperone protein associated with α -synuclein inclusion formation, was reported previously up-regulated in the SN of PD (28, 29). In contrast, vesicle trafficking, synaptic transmission, and synapse-related genes were significantly down-regulated in the neuronal clusters of PD brains (Fig. 5, B and C).

We noticed distinct patterns of the DEGs among the two DA neuron subtypes and $RIT2^{+}TH^{-}$ neurons (c9) in PD. c9 had the greatest number of DEGs (1203), far more than c7_3 (200) and c6_2 (66), though c9 (4200) had fewer cells than c7_3 (4262). The most

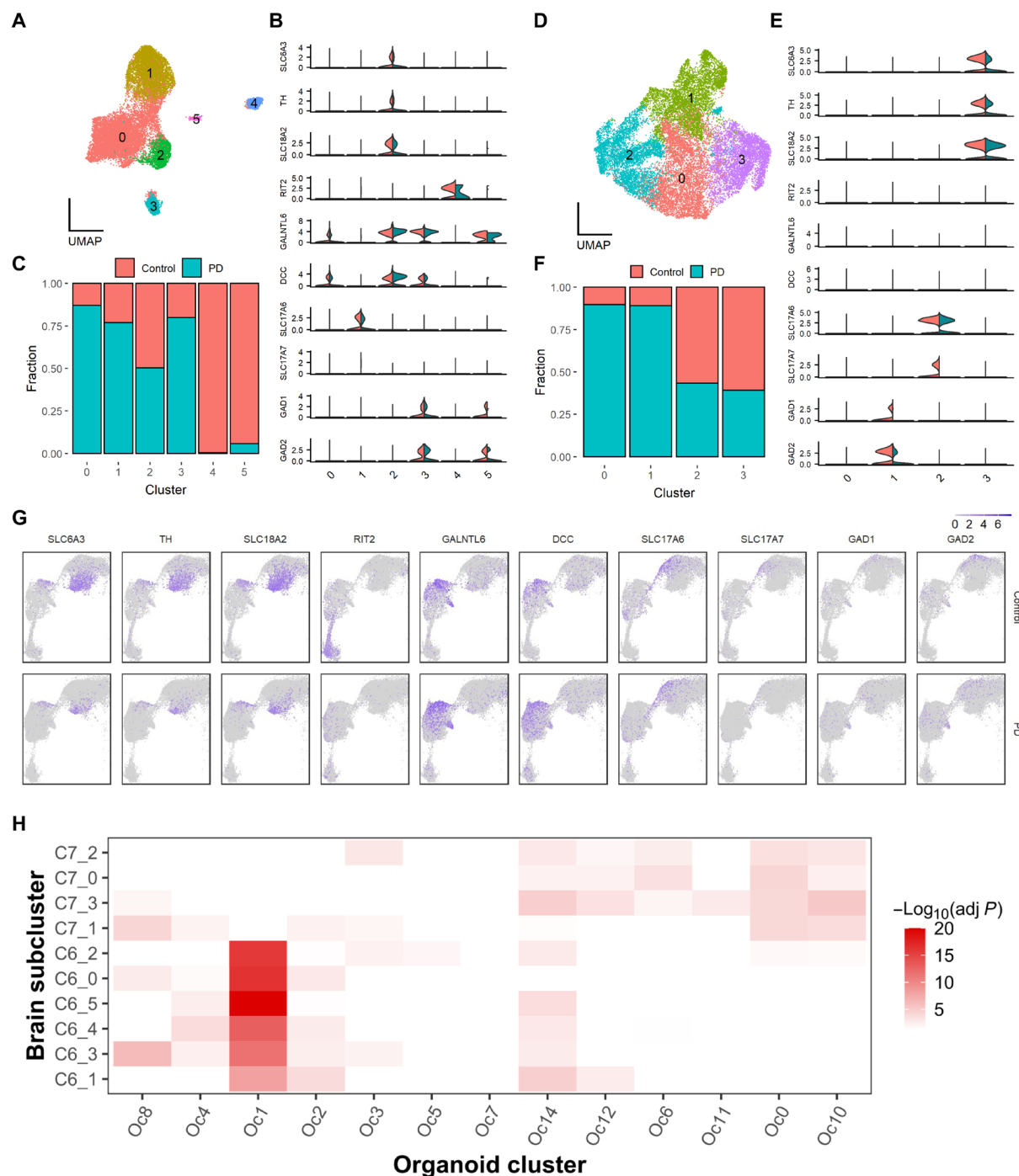


Fig. 4. Subclustering analysis of clusters c6 and c7 and identification of subtypes of DA neurons. (A to F) Subclustering analysis of clusters c6 [(A) to (C)] and c7 [(D) to (F)]. [(A) and (D)] UMAP plot for subclusters. [(B) and (E)] Expression pattern of known brain cell-type marker genes. [(C) and (F)] Cell fraction distribution represented by disease status in each subcluster. (G) UMAP plots for the expression of selected marker genes in the neuron clusters c6, c7, and c9. (H) Comparison of cluster markers between the organoids and DA neuron subclusters of human SN (control). The minimum adjusted P value was set at 1×10^{-20} for visualization purpose.

significant DEGs in c9 were associated with the down-regulation of synaptic protein interactions. In contrast, the DEGs of c7_3 showed the up-regulation of translation elongation and ribosomal proteins (Fig. 5B). Moreover, c7_3 had decreased expression of *SLC18A2*, *ALDH1A1*, *SLC6A3*, and *TH*, which are important for the regulation of dopamine release and DA neurogenesis (Fig. 5B). A few genes

sharing similar directions of changes (either up or down) among c9, c7_3, and c6_2 were also identified, which were involved in various cellular functions (fig. S8, B and C). Genes involved in single-stranded DNA sensing process such as *SSBP3* were commonly down-regulated, while *PLCL1* and *MTRNR2L1* were enhanced in all three neuron clusters (fig. S8, B and C).

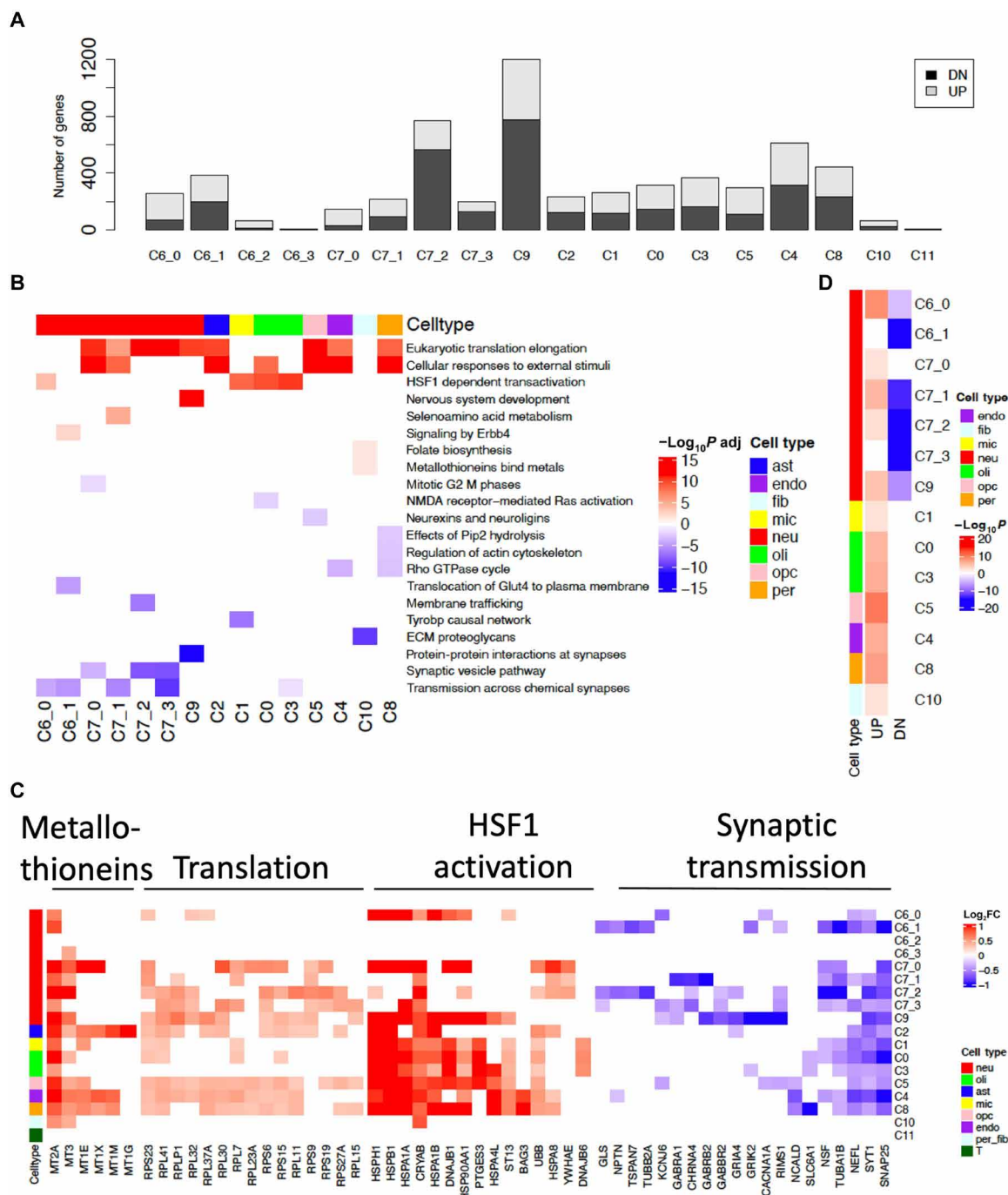


Fig. 5. Cell type-specific DEGs between PD and the control. (A) Number count of up- (UP) and down-regulated (DN) DEGs in each cluster. **(B)** Heatmap of top canonical pathways enriched for up- and down-regulated genes in each cluster. **(C)** Top DEGs involved in the indicated pathways in each cluster. **(D)** Comparison of DEGs identified in bulk tissue-based meta-analysis and snRNA-seq. FC, fold change.

Last, we observed an impaired *TYROBP*-centered causal network in the microglia (c1), indicating the immunosuppression in PD brains (Fig. 5B). These cluster-specific PD DEGs were largely consistent with those identified in our previous bulk tissue-based meta-analysis (30), with the down-regulated bulk tissue-based DEGs enriched in neuronal clusters, while others were found ubiquitously up-regulated across various cell types (Fig. 5D).

To identify potential temporal changes of gene expression during disease progression, we separated the cells into three groups according to the Braak Staging of the donors: control (Braak = 0), early stage (Braak 1 to 3), and late stage (Braak 4 to 6) (table S7). The genes were grouped into three major categories based on their expression patterns (details described in Materials and Methods): (i) Early and sustained responding genes (ESRGs); (ii) “U-shaped”

responding genes (URGs), whose early response diminished over time; (iii) late responding genes (LRGs). Each category was further divided into positive and negative responders. The positive ESRGs were primarily involved in HSF1 activation in non-neuron clusters (c0 to c3, c5, and c8) and a small set of neuron clusters (c6_0 and c7_0), while in DA neuron subclusters c7_3 and c9, DA neurogenesis and synaptic homeostasis, respectively, were disrupted at the early stage and then suppressed over the disease course (fig. S11A and table S8). Most positive URGs were enriched in metal ion regulation and metabolism in non-neuronal clusters, while a transient inactivation of *N*-methyl-D-aspartate-AMPK signaling and membrane trafficking was found in oligodendrocyte (c0) and microglia (c1) clusters (fig. S11B). Various cellular pathways were activated in all cell clusters as LRGs, including the two vulnerable neuron subtypes (c7_3 and c9), where translation elongation, chaperone activity, and GTPase cycle were altered (fig. S11C). Moreover, *TYROBP*-centered causal network and microglial pathogen phagocytosis pathway were disrupted at the late stage (fig. S11C). These results demonstrate divergent cellular stress responses in different cell types during disease progression.

Cell type-specific gene expression enrichment and deregulation of PD-associated genes in the human SN

We next examined the expression of PD-linked genes and GWAS risk alleles using our dataset. We were able to detect the expression of 22 PD-linked genes (4, 31), of which half (11 of 22) were enriched in neuron clusters, such as GABAergic (c7_1, adjusted $P = 0.019$, OR = 17.9), glutamatergic (c7_2, adjusted $P = 1.5 \times 10^{-04}$, OR = 23.0), and DA neurons (c7_3, adjusted $P = 9.1 \times 10^{-05}$, OR = 25.5). Numerous genes—such as *UCHL1* (PARK5), *SNCA* (PARK1/4), *ATP13A2* (PARK9), *VPS35* (PARK17), *SYNJ1* (PARK 20), *CHCHD2* (PARK 22), and *TMEM230*—showed strong expressions in both glutamatergic (c7_2) and DA (c7_3) subcluster neurons, while *PINK1* (PARK6), *EIF4G1* (PARK18), and *GBA* were particularly enriched in DA neuron subcluster (c7_3) (Fig. 6A, left). Unlike neuron-enriched PD genes, *LRRK2* (PARK8) was highly expressed in microglia (c1), endothelial cells (c4), and OPCs (c5). *PRKN* (PARK 2) was enriched in astrocytes (c2), microglia (c1), oligodendrocytes (c3), and OPCs (c5). Furthermore, *VPS13C* (PARK 23) and *DNAJC13* (PARK 21) were enriched in microglia (c1) (Fig. 6A, left).

The differential regulation of PD genes was heterogeneous across cell types. Nearly 30% of the PD genes (6 of 22) were down-regulated in neuronal clusters in PD. For examples, we found that *DNAJC6*, *CHCHD2*, and *SNCA* were down-regulated in DA neuron subclusters c7_3 or c9. In contrast, *DNAJC6*, *UCHL1*, and *PRKN* were up-regulated particularly in DA subcluster c9, whereas *SNCA* was up-regulated in microglia and oligodendrocytes in PD. In contrast, *PRKN* was down-regulated in pericytes (c8) and endothelial cells (c4) in PD (Fig. 6A, right).

We next extended our studies to the genes mapped to the known PD GWAS loci curated by GWAS catalog (www.ebi.ac.uk/gwas/; see Materials and Methods). Among 278 PD GWAS loci alleles, 90 genes showed cluster-specific expression as evidenced by the significant enrichment in neuron cluster (c9, adjusted $P = 0.037$, OR = 3.8), while DA neuron cluster (c7_3) was enriched with genes linked to the genetic forms of PD. Fifty-two GWAS-associated genes were differentially expressed between PD and control (Fig. 6B). By categorizing these genes by cell type, we found that many PD GWAS

loci genes were preferentially expressed in neurons (Fig. 6C) and more frequently deregulated in neurons than in other cell types (Fig. 6D). For example, *SV2C* was highly expressed in both c6_2 and c7_3 DA neuron clusters and down-regulated in c7_3 in PD (Fig. 6, C and D). *SV2C* encodes a synaptic vesicle glycoprotein 2C, which plays a role in the control of regulated secretion in neurons (32). *KTN1*, encoding an integral membrane protein belonging to the kinectin family, was highly expressed in endothelial cells (c4) (Fig. 6C) but up-regulated in c9 neurons (Fig. 6D). We found a significant down-regulation of the *RIT2* gene in c9 (Fig. 6D). The above results demonstrate the heterogeneity of expression enrichment for PD-associated genes in different cell types of the SN, suggesting the complexity of pathogenic mechanisms of PD.

Altered cell-cell communication networks in the SN of PD

We next predicted altered cell-cell communications by using an R package, CellChat, where cell communications are characterized by ligand-receptor (LR) interactions between the source and target cells (details in Materials and Methods) (33). First, by comparing the differential number and strength of LR interactions among major cell types between control and PD, we observed a global decrease of cell communications for neuronal cells but increased communications for microglia, pericytes, endothelial cells, and fibroblasts (Fig. 7A). Second, we assessed how individual cell clusters were affected by PD, according to the total outgoing and incoming signals. A number of neuron and oligodendrocyte-related cell clusters showed a loss of incoming and/or outgoing cell communications in PD, including c9 (Neu), c5 (OPC), c2 (Ast), c3 (Oli), c7 (Neu), c0 (Oli), and c6 (Neu), while the other cell types showed no obvious change in incoming or outgoing interaction strength (Fig. 7B).

By aggregating the LR pairs into major signaling pathways, we found that GRN, EPHB, GAS, PERIOSTIN, EDN, OCLN, interleukin-16, major histocompatibility complex-II (MHC-II), WNT, PACAP, and CXCL pathways were inactivated in PD (Fig. 7C and table S9). The β -amyloid precursor protein pathway was enhanced, while the Angiopoietin-like protein (ANGPTL) pathway was activated in PD. Multiple signaling pathways also showed cell cluster-specific regulation in PD. For example, two neuronal clusters c7 and c9, both containing subtypes of DA neurons, lost Cadherin (CDH) signaling in PD (Fig. 7, C and D). Disruption of the CDH pathways was also seen in pericytes (c8). *CDH2*, which encodes N-cadherin, is the primary member of CDH. Because N-cadherin exerted a neuroprotective effect on DA neurons (34), the loss of CDH interactions in DA neurons may disrupt the functions of DA neurons. Ephrins and Eph receptors are multifunctional in various biological conditions, including axon guidance and regeneration (35). Aside from the loss of EPHB signaling in PD, we noticed that the Ephrin type-A receptors (EPHA) input from neuron clusters c6 and c9 into neuron cluster c7, oligodendrocyte cluster c0, and pericytes cluster c8 were lost in PD (Fig. 7D). Together, our analysis suggests that multiple pathways are altered in PD. Specifically, we found extensive disruptions and rearrangements of CDH and EPHA/EPHB signaling pathways in the SN of PD.

DISCUSSION

Emerging scRNA-seq approaches have become instrumental in deciphering the intricate heterogeneity and composition of cell types of complex human diseases including Alzheimer's disease (36). In

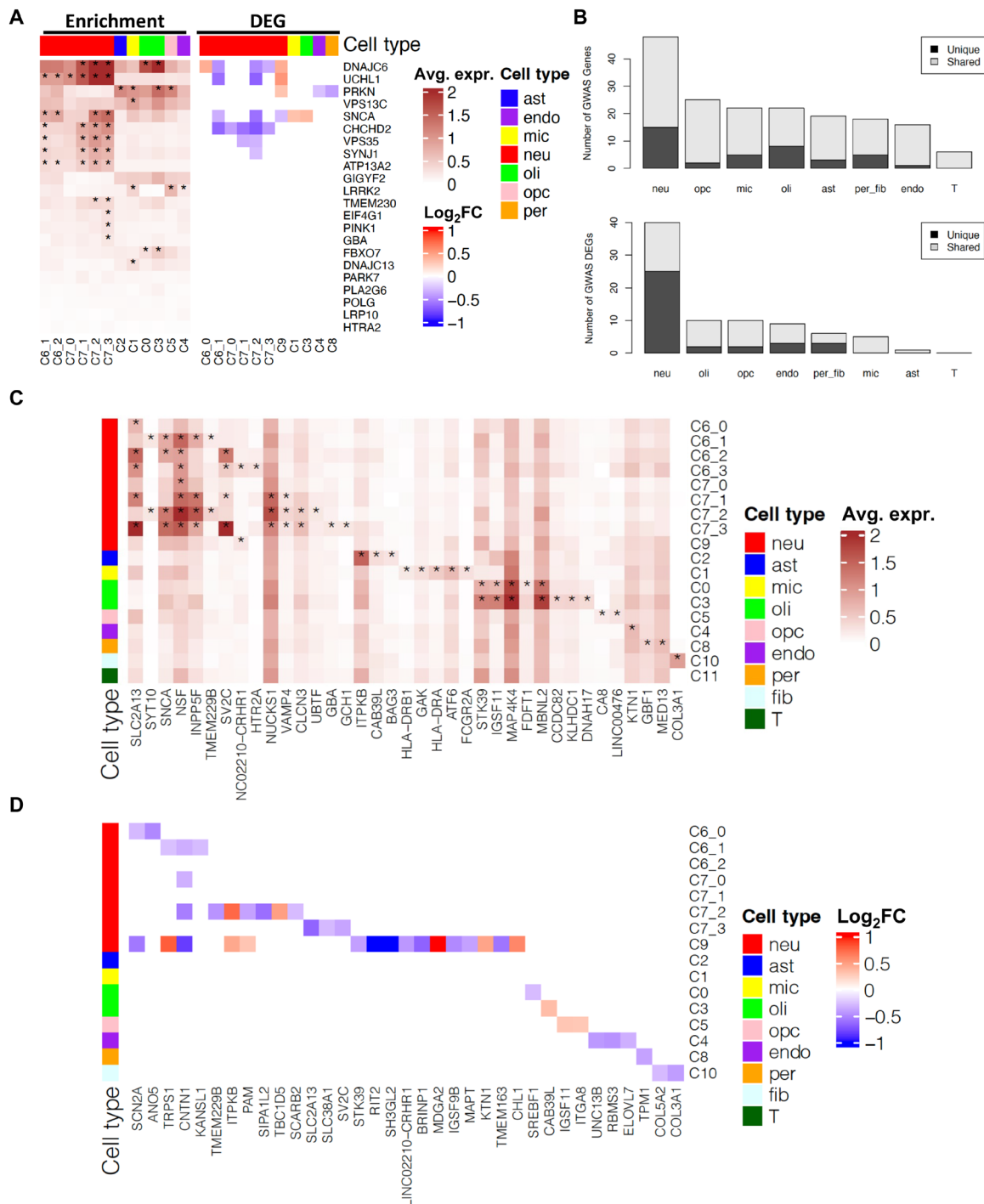


Fig. 6. Cell type-specific expression and DEG of PD-associated genes and risk loci. (A) Heatmap for cell type-specific expression (left) and DEG (right) of *PARK* family genes in PD. (B) Number of GWAS loci related genes enriched (top) and differentially expressed (bottom) in each cell type. Dark gray indicated genes unique to each cell type, and light gray indicated genes shared among cell clusters. (C) Expression patterns of selected PD GWAS loci-related genes in different cell clusters. (D) Heatmap for the log₂ fold change of differential expression of selected PD GWAS loci genes in each cluster.

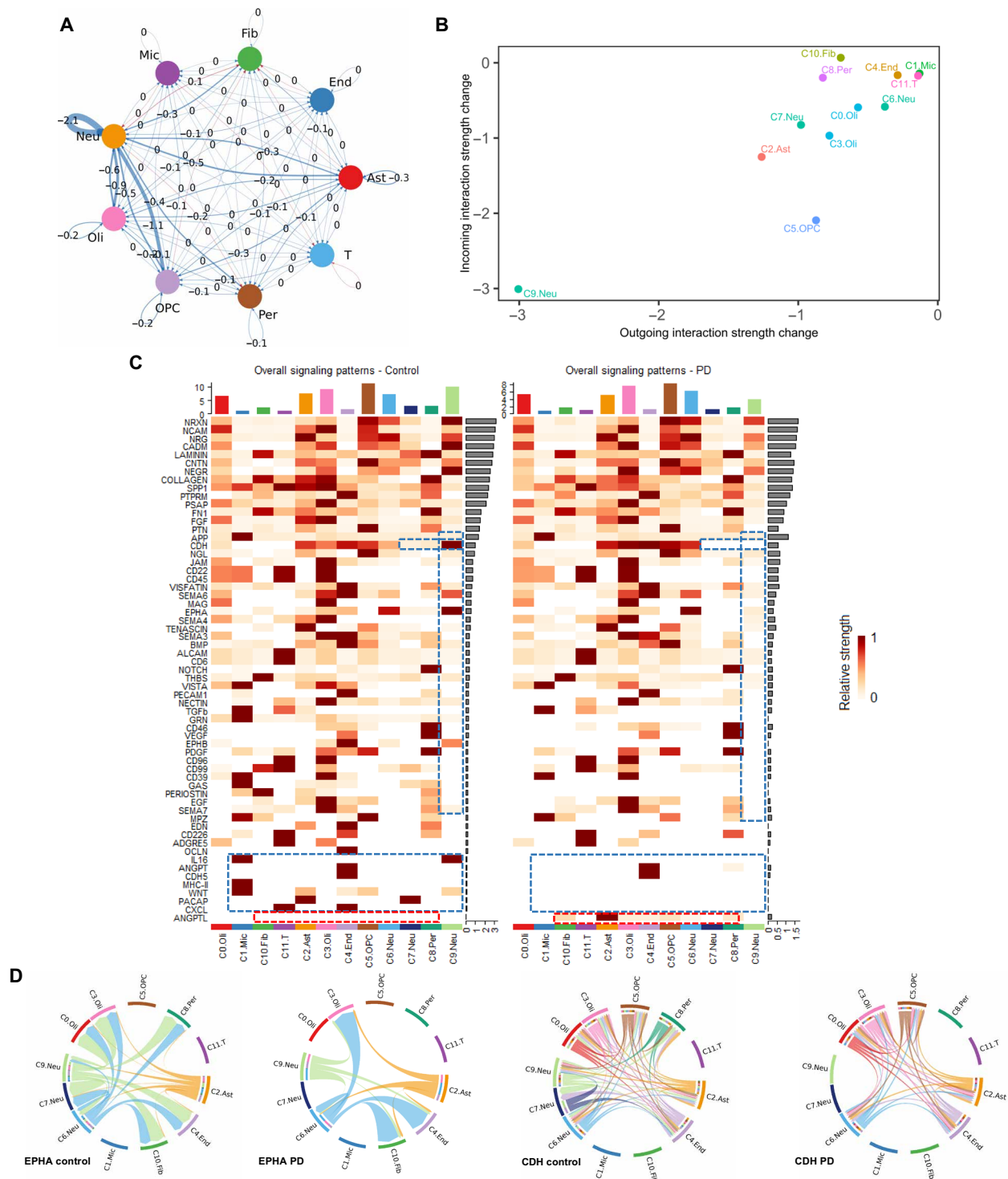


Fig. 7. Altered cell communication networks in PD. (A) Differential LR interactions between PD and the controls. Blue lines and red lines indicated decreased and increased interactions, respectively. The line width was proportional to the difference. (B) Distribution of cell cluster based on their relative changes in the incoming and outgoing signaling strengths between PD and control. (C) Information flow changes of major signaling pathways between PD and control in each cell cluster. Dashed boxes highlighted gain (red) or loss (blue) of signaling in specific cell clusters. (D) Chord diagram of EPHA and CDH signaling in control and PD.

this study, we profiled a large number of nuclei in the human SN from both PD and controls with an average age of 81. We developed a comprehensive single-cell transcriptomic atlas of the human SN. Our data revealed cell type-specific molecular alterations and disruption of cell-cell communication networks in the SN of PD. Our results highlight the cell heterogeneity and molecular basis for the complexity underlying disease mechanism.

The profiling of human SN has led to the identification of a *RIT2*-enriched neuron subpopulation, a fraction of which expressed little *TH* (*RIT2*⁺*TH*[−]) and showed a distinct transcriptomic signature from other neuron types including DA neurons (*TH*⁺*SLC18A2*⁺*SLC6A3*⁺). *RIT2*⁺ neurons have never been previously characterized. The *RIT2*⁺ neurons (including *RIT2*⁺*TH*[−] and *RIT2*⁺/*TH*⁺) showed remarkable vulnerability in PD. We validated the presence of *RIT2*⁺*TH*[−] neurons in two previously published datasets from the human SN, human organoids, and the mouse SN, although a potentially different distribution of the specific neuron population was noted between mouse (Fig. 2G) and human (Fig. 2C). Our study also indicated that NM are found in *RIT2*⁺*TH*[−] and *RIT2*⁺*TH*⁺ subtypes. A previous study found differential rate of degeneration between NM⁺ and *TH*⁺ neurons in the SN of PD, suggesting the heterogeneity of NM⁺ neurons—not all NM⁺ neurons express *TH* (18). It remains possible that some *RIT2*⁺*TH*[−] cells account for those NM⁺ neurons.

While the emergence and significance of *RIT2*⁺*TH*[−] neurons remain to be elucidated, *RIT2* was previously identified as a PD risk gene (3, 22). We found *RIT2* expression was down-regulated in the c9 in PD with advanced Braak stages (0.8-fold, adjusted *P* value = 0.047), consistent with a previous report of reduced expression of *RIT2* in the SN of PD brain (37). The *RIT2* protein belongs to the RAS superfamily of small GTPases, which interact with and regulate DAT levels in a sex-dependent manner in mice (38). One question is whether the loss of *RIT2*⁺*TH*[−] neurons is linked to any symptom of PD. Our study suggests *RIT2* as a potential molecular marker for age-related pathological alterations in PD due to the loss of *RIT2*⁺ (*TH*⁺ and *TH*[−]) neurons in the SN and the link of *RIT2* variants to PD risk.

The PD samples used in our snRNA-seq study are from advanced stage of PD (~80 years), by which point most of DA neurons are lost. Therefore, our data are unlikely to provide an insight into the causes of neurodegeneration. However, the study of the remaining DA neurons at advanced stages may reveal a clue for how they are adapted and survive, while others are lost. Kordower *et al.* (18) reported a rapid decline of DA neuron number in the SN of PD at an early stage of the symptoms, while a subpopulation of DA neurons remained unchanged in number for at least a decade and at advanced stages of the disease. At present, it remains to be further determined what could render the remaining DA neuron resistant to the death at late PD. By comparing transcriptomic profiles of the DA neurons between PD and control cohorts, our study may provide an opportunity to decipher potential mechanisms for their resilience.

Our study gains an insight into the cell type-specific gene expression for the PD-linked genes and GWAS risk alleles. Our results demonstrated the cell heterogeneity of expression enrichment for PD-associated genes in the SN. The DEG analysis suggested common and distinct cellular pathways that are affected in various cell types of the SN of PD. Our result revealed that most of the DEGs in PD are associated with neurons. It is worth noting that *LRRK2* was produced primarily in microglia and OPCs, but little change was observed in any cell type in PD. Furthermore, *SNCA* expression

was reduced in subpopulations of DA and glutamatergic neurons but was enhanced in microglia and oligodendrocytes in PD. The transcriptional elevation of *SNCA* has not been previously reported in glia in the SN of PD, and the significance of this observation should be investigated in the future. The above observations highlight the diversity of the molecular mechanisms underlying DA neuron degeneration.

Many studies have indicated activation of microglia or astrocytes in PD. Several groups reported the appearance of amoeboid microglia producing MHC class II, intercellular adhesion molecule-1 (ICAM-1), and LFA-1, the markers for activated microglia, and reactive astrocytes expressing ICAM-1 in the SN and the putamen of PD brain (39–43). Previous studies also detected increased binding of a radiotracer, 11C-(R)-PK11195, in PD brains compared to controls. This tracer is known to bind to the 18-kDa translocator protein (TSPO) expressed mainly by microglia (44). However, other groups failed to observe activated microglia or astrocytes in human PD brains (45, 46). In our study, we found little evidence supporting the extensive activation of inflammation-related molecules or disease-associated microglia signature in glial clusters, although we detected the up-regulation of multiple genes, such as *AKT-PIP3*, *FOXO1/3*, and *ERBIN*, as well as *NGR3*, which are known to regulate macrophage/microglia activation, migration, proliferation, and inflammation (47–51). The discrepancy of the results could be due to many factors such as sample sources and analytic procedures. It remains possible that glia become less active at advanced stage of PD as shown in our study. Furthermore, accumulated evidence also suggests the increase of senescent/dystrophic microglia in human aged brains (52, 53).

Despite the large number of nuclei analyzed by snRNA-seq in our study, the sample size used in the investigation is still relatively small, considering the variations among postmortem samples such as postmortem interval (PMI) and pathological differences. The observation of altered gene expression in PD could be biased and should be rigorously validated with large sample sizes from independent cohorts. Moreover, the development and physiological function of *RIT2*-enriched neurons in the human SN has yet to be elucidated. Nonetheless, our study has established a transcriptomic atlas of the human SN at the single-cell resolution and delineated the landscape of molecular and cellular alterations in PD. Our study not only provides a valuable resource for dissecting molecular and cellular compositions and structures of the human SN but also presents an unprecedented opportunity to understand in-depth pathogenic mechanisms, identify key therapeutic targets, and develop novel clinical biomarkers for PD.

MATERIALS AND METHODS

Postmortem brain sample collection

The postmortem brain samples were requested from National Institutes of Health (NIH) Neurobiobank (www.neurobiobank.nih.gov) and fulfilled by Brain Endowment Bank at Miller School of Medicine, University of Miami. The samples were pretested for known genetic mutations linked to familial PD, including *SNCA*, *LRRK2*, and *GBA*. The samples did not harbor any of the abovementioned mutations.

The information of the sample demographics was summarized in Table 1 and table S1. Specifically, frozen punches of SN were obtained and then pulverized in a liquid-nitrogen-chilled mortar and aliquoted. Approximately 50 mg of tissues was used for snRNA-seq.

Midbrain organoid differentiation

The midbrain organoids were generated from the male human embryonic stem cell line, WA01 TH #2, containing the TH-TdTomato reporter for DA neurons, as previously described (24, 54). Briefly, the hPSCs were seeded into a 125-ml disposable spinner flask (Corning, #3152) in StemFlex supplemented with 10 μ M Y-27632 ROCK inhibitor (Tocris, #1254) and Pen/Strep (Gibco, #15140122) in a total volume of 120 ml. The flask was placed on a ninth position stir plate (Dura-Mag, #CLS-4100-09) at a speed of 65 rpm in the incubator. On day 2 after seeding, half of the culture medium (60 ml) was changed. Midbrain patterning was initiated on day 4 after seeding. The base medium throughout the differentiation consisted of DMEM/F12 + GlutaMAX (Gibco, #10565018) supplemented with 1 $\times$ B27 supplement, minus vitamin A (Gibco, #12587010) and 1 $\times$ N-2 supplement (Gibco, #17502048). Half of the media (60 ml) was changed daily throughout the patterning stage with the exception of 80 ml of media changes on day 8 (D8) and D12. The D0 to D1 medium contained the transforming growth factor- β inhibitor, SB431542 (10 μ M; Stemgent, #04-0010), and the bone morphogenetic protein inhibitor, LDN193189 (100 nM; Tocris, #6053). Purmorphamine (2 μ M; STEMCELL Technologies, #72202) and 3-chloro-N-[trans-4-(methylamino)cyclohexyl]-N-[[3-(4-pyridinyl)phenyl]methyl]-benzo[b]thiophene-2-carboxamide (SAG) (1 μ M; Cayman Chemical, #11914) were added to the D2 to D3 medium. The WNT activator, CHIR99021 (1.5 μ M; Tocris, #99021), was also added to the D4 to D7 medium. The D8 to D11 medium only contained LDN193189 and CHIR99021. On D12, the medium was switched to the terminal differentiation medium consisting of brain-derived neurotrophic factor (20 ng/ml; R&D Systems, #248-BD), glial cell line-derived neurotrophic factor (20 ng/ml; R&D Systems, #212-GD), N-[2S-(3,5-difluorophenyl)acetyl]-L-alanyl-2-phenyl-glycine, 1,1-dimethylethyl ester (DAPT) (10 μ M; Cayman Chemical, #13197), ascorbic acid (0.2 mM; Fisher BioReagents, #BP351), and Dibutyryl-cAMP,N⁶,2'-O-Dibutyryl-adenosine-3',5'-cyclic monophosphate (dcAMP) (0.1 mM; BioLog, #D009). Samples were collected on D30 for immunofluorescence (IF) staining and on D40 for scRNA-seq analysis.

Nuclei isolation and sequencing

Single-nucleus gene expression sequencing was performed on the samples using the Chromium platform (10x Genomics, Pleasanton, CA) with the Next GEM Single cell 3'GEX Reagent Kit and an input of ~10,000 nuclei from a debris-free suspension. Briefly, nuclei were isolated from frozen tissue, as per 10x Genomics' recommendations, using chilled, 0.1% NP-40 lysis buffer with gentle homogenization and washed. Gel-bead in emulsions (GEMs) were generated on the sample chip in the Chromium controller. Barcoded cDNA was extracted from the GEMs by Post-GEM RT-cleanup and amplified for 12 cycles. Amplified cDNA was fragmented and subjected to end repair, polyadenylate tailing, adapter ligation, and 10x-specific sample indexing following the manufacturer's protocol. Libraries were quantified using Bioanalyzer (Agilent) and QuBit (Thermo Fisher Scientific) analysis. Libraries were sequenced using a 2 $\times$ 100PE configuration on a NovaSeq instrument (Illumina, San Diego, CA), targeting a depth of 50,000 to 100,000 reads per nucleus.

Sequencing data were aligned and quantified using the Cell Ranger Single-Cell Software Suite (version 3.1.0, 10x Genomics) against the provided GRCh38 reference genome using default parameters, including introns. Before sequencing all samples, three samples were

randomly selected for sequencing in a pilot run. Then, all 31 samples were processed and sequenced in one batch. For the three samples with replicated libraries, we observed similar data quality between the pilot and final sequencing run. Therefore, we combined libraries from both the pilot and final sequencing, resulting in a total number of 457,453 nuclei (on average, 13,455 nuclei and median 12,188 per library) before quality control (QC).

snRNA-seq data preprocessing and preclustering analysis

Starting from Cell Ranger-derived unique molecular identifier (UMI) count matrices from all sequencing libraries, we performed QC by removing low-quality nuclei with either too few genes (<200) or an excessive number (>2500) of genes detected, retaining 355,157 nuclei after filtering. Then, we removed insufficiently detected genes by keeping 30,038 genes expressed in more than one nucleus. Mitochondrial reads (mean fraction of 18.5%, median fraction of 13.4%) were discarded to avoid biases introduced during the nuclei isolation because they are not expressed inside nucleus (55–57). After QC, there were, on average, 9290 (median of 9010) nuclei per sequencing library sample. We obtained, on average, 817 (median of 686) unique genes per nucleus per individual and 1536 (median of 1080) UMI per nucleus. We performed a preclustering analysis using a well-established scRNA-seq data integration workflow based on R packages Harmony (20) and Seurat (v3) (19). Briefly, the UMI data were first normalized by sequencing depth and log-transformed using the LogNormalize method implemented in Seurat. A total of 2000 most variable gene features were identified, scaled, and centered after regression out covariates sex, age, and PMI. Next, dimensional reduction was performed using principal components analysis (PCA) based on the 2000 most variable genes. The top 30 principal components (collectively explaining more than 90% of the variance) as determined by an elbow approach were selected for integration of snRNA-seq data across all sequencing libraries with Harmony (20). Top 20 embeddings in the Harmony space were used for calculating two-dimensional (2D) reductions by *t*-distributed stochastic neighbor embedding (58) and Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) (59). The same top 20 Harmony embeddings were also used to compute the nearest neighbor graph and the subsequent cell preclusters with the Louvain algorithm implemented in Seurat (19). This initial preclustering analysis resulted in 14 preclusters at a resolution of 0.2. Two smallest preclusters c12 and c13, dominated by nuclei from one or two donors, overlapped with precluster c0 on the UMAP space (fig. S1, A and B).

Cluster stability analysis

To assess the stability and robustness of the preclusters, we performed repeated subsampling analysis by making use of software tool scclusteval (60). In each subsample, we sampled without replacement a subset of 80% of the nuclei in the full QCed dataset and then repeated the data normalization, scaling, PCA, Harmony data integration, and clustering procedure on this subset of data as above described. We repeated subsampling 100 times. For each subsample, we compared its clusters with those preclusters from the full data by Jaccard index analysis and returned a maximum Jaccard index coefficient for each of the original preclusters. We found that precluster c12 had almost close to 0 Jaccard index coefficients in all subsamples, indicating that it was an unstable cluster dissolved in the subsamples. Precluster c13 was also dissolved in 13 of the subsamples, suggesting that it was a potentially unstable cluster. For the remaining

preclusters, they all showed Jaccard index coefficients larger than 0.50, except for c9 where one subsample had a Jaccard index coefficient of less than 0.25. Next, for each cell i in precluster c12 or c13, we assessed the coclustering probability between i and all the preclusters as the mean fraction of cells in the preclusters that clustered together with i in the repeated subsamples by using equation

$$p_{c,i} = \frac{1}{m} \sum_{r=1}^m |x_c \cap x_{r,i}| / |c \cap x_r|$$

where m denotes the number of subsamples that included cell i , c denotes a precluster, x_c denotes the set of cells in precluster c , x_r denotes the set of cells in subsample r , and $x_{r,i}$ denotes the set of cell subcluster that contains cell i in x_r . The distribution of $p_{c,i}$ stratified by c is shown in fig. S1D. Cells in precluster c12 tended to cocluster with cells in precluster c0, followed by cells in c3 and c13, while cells in precluster c13 tended to cocluster with cells in precluster c12, followed by cells in c0. Together with the spatial distribution of the cells in the UMAP space, we decided to merge the two unstable preclusters c12 and c13 into their adjacent bigger neighbor precluster c0, leading to 12 clusters (c0 to c11) for further analysis.

Doublet prediction analysis

After finalizing the cell clusters, we predicted doublets by making use of the scDblFinder package (61, 62). scDblFinder first simulates doublets from the provided cell clusters and then computes a doublet prediction score for each cell by combining the fraction of simulated doublets in its neighborhood with another score based on coexpression of mutually exclusive gene pairs (19). The doublet prediction scores are iteratively refined, and a classification model is trained to best characterize the putative doublets by integrating a number of discriminating metrics (61, 62). Figure S1E shows the final doublet prediction score and classification. A total of 39,290 predicted doublets were removed, resulting in 315,867 singlets of the final clusters for all downstream analyses. The number of high-quality nuclei and the number of detected genes across all nuclei of each sample are listed in table S1. Overall, there are, on average, 9290 (median of 9010) singlets per sequencing library sample.

Cluster cell-type annotation

For each major cluster or subcluster, we first interrogated the expression patterns of known gene markers to annotate clusters into major cell types: neurons (*RBFOX3*, *GAD1*, and *NRGN*), astrocytes (*AQP4* and *GFAP*), oligodendrocytes (*MOG*), microglia (*C3*, *CSF1R*, *CD74*, and *TYROBP*), oligodendrocyte progenitor cells (*VCAN*), endothelial cells (*FLT1*), and pericytes (*PDGFRB*). Next, we calculated de novo cluster signatures by comparing the cells in this cluster against the cells of the rest clusters using Wilcoxon rank sum test in Seurat. We defined cluster up-regulated genes as those up-regulated by at least 1.2-fold and with Bonferroni adjusted P value less than 0.05, from which we further defined cluster-enriched de novo signatures (i.e., marker genes) as those up-regulated by at least twofold. To assist the annotation of cell type of each cluster, we overlapped the de novo cluster signatures with a large-scale collection of cell-type markers curated from more than 1054 single-cell experiments (21), with P value significance of the overlaps computed by a hypergeometric test. Because cell type marker expression may change in PD cells, only the control cells were used for investigating the marker gene expression pattern and calculating the cluster signatures.

Subclustering analysis

To perform subclustering analysis for a given cluster, we first extracted the normalized and covariates adjusted data for the cluster. As in the preclustering analysis, we computed dimensional reduction using PCA. Because the number of cells contributed from each individual donor ranged from 18 to 2170 (cluster 6), and 19 to 2069 (cluster 7), we did not conduct Harmony analysis to avoid biased data integration due to small and uneven cell numbers. The top 10 principal components as determined by an elbow approach were selected to compute UMAP, the nearest neighbor graph, and the subsequent cell subclusters with the Louvain algorithm. Subcluster marker signatures were defined by comparing each subcluster with all other cells, including other subclusters and major clusters, using Wilcoxon rank sum test in Seurat as above.

Cell cluster proportion change

To test if there was a significant change in the proportion of a cell cluster between PD and controls, we first calculated the OR of proportion difference using formula: $OR = \left(\frac{N_{c,Control}}{N_{c,PD}} \right) / \left(\frac{N_{\bar{c},Control}}{N_{\bar{c},PD}} \right)$, where $N_{c,PD}$ ($N_{c,Control}$) denotes the total number of cells in cluster c in all PD (control) samples, while $N_{\bar{c},PD}$ ($N_{\bar{c},Control}$) denotes the total number of cells in the remaining clusters in all PD (control) samples. To compute the P value significance, we computed the odds of cells assigned to cluster c in individual i as $odds_{c,i} = f_{c,i} / f_{\bar{c},i}$, where $f_{c,i}$ denotes the fraction of cluster c cells in the i -th individual and $f_{\bar{c},i}$ denotes the fraction of cells in all the other clusters in the i -th individual. Then, a one-tailed Wilcoxon rank sum test was conducted to compare the difference in odds between PD and controls.

Replication of cell clusters in two independent non-PD SN samples

To replicate the present cell clusters, we compared our data with two published snRNA-seq datasets from non-PD SN samples. We first reprocessed the snRNA-seq data reported by Agarwal and colleagues from non-PD SN samples. By using the pipeline described in the precluster analysis section, we identified clusters of major brain cell types from the Agarwal data (13), including astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and oligodendrocyte progenitor cells (fig. S5, A and C). We further performed a subclustering analysis on the neuronal cluster c6 and identified three subclusters, c6_0 (*RIT2* and *RBFOX3* enriched), c6_1 (*TH* enriched), and c6_2 (*GAD1* and *GAD2* enriched) (fig. S5, B and D). To compare the cluster similarity between our data and the Agarwal data, we assessed the significance of intersection of cluster signatures between ours and the Agarwal dataset using hypergeometric test (fig. S5E). Similarly, we reprocessed another snRNA-seq data from non-PD SN samples reported by Welch *et al.* (14) and identified clusters of major brain cell types (fig. S6, A and C). We performed a subclustering analysis on the neuronal cluster c4 and identified seven subclusters, among which c4_3 is enriched for *TH/SLC6A3/SLC18A2* and c4_5 is uniquely enriched for both *RIT2* and *RBFOX3* (fig. S6, B and D). Then, we compared the cluster signature similarity between our data and the Welch data using hypergeometric test (fig. S6E).

Midbrain organoid processing for scRNA-seq

The midbrain organoids were dissociated as previously described (24). Briefly, up to 20 organoids were washed with 5 ml of phosphate-buffered saline (PBS)^{−/−} in a well of a six-well plate and incubated in

5 ml of papain (Worthington, #LK003176) in Hanks' balanced salt solution-H (STEMCELL Technologies, #37150) supplemented with deoxyribonuclease (DNase; Worthington, #LK003170) and 10 μ M Y-27632 ROCK inhibitor for 20 to 40 min (depending on organoid size) on a shaker in the incubator. The organoids were gently triturated using fire-polished glass pipets of decreasing tip sizes. The cell suspension was filtered through a 70- μ m strainer (Fisherbrand, #22-363-548) into a 50-ml tube containing 10 ml of fluorescence-activated cell sorting (FACS) buffer [hibernate-A medium (BrainBits #HA) supplemented with 1 $\times$ B27 (minus vitamin A), 1% bovine serum albumin solution, DNase, and 10 μ M Y-27632 ROCK inhibitor]. The cells were centrifuged at 800g for 5 min, the pellet was resuspended in 1 ml of FACS buffer, and the suspension was filtered through a 35- μ m strainer (STEMCELL Technologies, #100-0087) into a FACS tube. The samples were stored on ice for FACS. The BD FACS Aria Cell Sorter (BD Biosciences) was used to enrich for the TH-TdTomato⁺ cells for the scRNA-seq experiment. The FACS gates consisted of the P1 gate to exclude cell debris, singlets to exclude doublets, and % live to exclude dead cells. The TH-TdTomato gate was set on the basis of a negative control. The FSC Express 7 Software was used to analyze the results.

Single-cell transcriptomics analysis of midbrain organoids

scRNA-seq analysis of midbrain organoids was performed using the Chromium platform (10x Genomics, Pleasanton, CA). The raw sequencing data were first processed by the standard 10x Genomics cell ranger pipeline to derive UMI count matrix. Then, cells with less than 200 genes, more than 6000 genes, or with more than 5% mitochondrial reads were discarded, and genes detected in less than two cells were removed. After filtering, the UMI count matrix was normalized by sequencing depth and log-transformed using the LogNormalize method implemented in Seurat (19). A total of 2000 most variable gene features were identified, scaled, and centered after regression out percentage of mitochondrial reads. Next, dimensional reduction was performed using PCA based on the 2000 most variable genes. The top principal components collectively explaining more than 90% of the variance were selected for calculating 2D reductions by UMAP (59). The same top principal components were also used to compute the nearest neighbor graph and the subsequent cell preclusters with the Louvain algorithm at a clustering resolution of 0.5. This clustering analysis resulted in 18 initial clusters. Doublets were predicted using the scDblFinder package (61, 62) and removed from further analysis. The smallest cluster was dominated by doublets (90 of 93 cells) and thus removed. De novo cell cluster markers were computed for each of the remaining 17 clusters (named Oc0 to Oc16 for simplicity). To compare the cluster similarity between human SN data and the organoid data, we assessed the significance of intersection of cluster markers using hypergeometric test.

Cluster-specific differential gene expression and functional enrichment analysis

DEGs between PD and control and DEGs between different Braak stages and control in each cluster/subcluster were identified using the R package MAST (19, 63) implemented in Seurat, with correction for PMI, age, and sex. DEGs were identified at the cutoff of Bonferroni corrected P value ≤ 0.05 and fold change $\geq 20\%$. For Braak stage-dependent differential expression analysis, we separated the cells into three groups by the Braak stage of the donors:

control (Braak = 0), early stage (Braak 1 to 3), and late stage (Braak 4 to 6). Genes significantly increased/decreased in the early versus control and late versus control contrasts were considered as positive/negative early and sustained responders. Genes significantly increased in the early versus control contrast but decreased in the late versus early contrast and vice versa were defined as positive and negative U-shaped responders, respectively. Genes significantly up-regulated/down-regulated only in the late versus control or also in the late versus early contrast were late responders. Functional enrichment of DEGs with MSigDB gene annotation collections was examined by Fisher's exact test (FET) with Benjamini-Hochberg (BH) correction. Results with BH adjusted P value < 0.05 were considered statistically significant. When comparing cluster-specific DEGs among three or more clusters, R package SuperExactTest (64) was used to visualize and compute the P value significance of the overlap.

Comparisons between cluster-specific with bulk tissue-based DEGs

Bulk tissue-based DEGs were defined by a meta-analysis described in our previous work (30). The intersection between the present cell cluster-specific DEGs and the bulk tissue-based DEGs was examined by FET followed by BH correction using the common genes identified in both bulk and snRNA-seq as the background. Results with BH adjusted P value < 0.05 were considered statistically significant.

PD-associated gene expression and regulation patterns in the snRNA-seq

PD-linked genes were defined as genes whose mutations were directly linked to familial PD as reviewed in (3, 31). PD GWAS genes were downloaded from GWAS catalog (www.ebi.ac.uk/gwas/) and defined based on the mapped genes closest to the risk loci. Both PD-linked genes and PD GWAS genes were considered as PD-associated genes. We first examined whether PD-associated genes were preferentially expressed in certain cell types by overlapping them with top 10% up-regulated genes ranked by fold change in each cell cluster compared to the rest using control cells only. The overlap was tested by FET using GeneOverlap package in R and the results with Bonferroni adjusted P value < 0.05 were considered statistically significant. We then examined whether PD-associated genes were differentially regulated in PD by overlapping them with cluster-specific PD versus control DEGs. The total number of PD-associated genes that showed cell type-specific enrichment and differential regulation were summarized, and they were further separated into single cell-type enriched or shared across multiple cell types.

Cell communication analysis

Cell communication analysis and visualization were performed using the default setting in CellChat package (33). In CellChat, the differential number of LR pairs between two conditions was determined by differential gene analysis. If either or both components of a LR pair disappeared in the disease status, then such a LR pair between source and target cells was considered lost. The signaling strength, or communication probability, of a LR pair is modeled by the law of mass action. The communication probability of a signaling pathway summarizes the probabilities of its associated LR pairs. A weighted-directed network is used to demonstrate the information flow among various cell groups, where the directed edges represent the summarized communication strength of all possible signaling

pathways between source and target cells. Therefore, the incoming and outgoing signals for a given cell type can be reflected by the thickness of the inward edges and outward edges, respectively. All LR pairs in cell-cell contact, extracellular matrix receptors, and secreted signaling were included. Specifically, we focused on the gain or loss of interactions in general cell types, the shift of the outgoing and incoming interactions, and the top signaling pathways altered among cell clusters between PD and control.

IHC staining for the human brain

Formalin-fixed paraffin-embedded (FFPE) blocks were sectioned into 5- μ m slices, deparaffinized with xylene for 10 min for three times, and rehydrated with gradient EtOH (100%, 100%, 75%, 50%, and water). Antigen retrieval was performed using Antigen Unmasking Solution, Citric Acid Based (Vector Laboratories, H-3300) with a pressure cooker (Presto Electronic, WI) for 15 min. For immunofluorescence staining, after blocking with 1% goat serum (Vector laboratory, CA, S-1012-50), brain slices were incubated with RIT2 antibody (1:100; Thermo Fisher Scientific, CF501757, clone OTI3F4) and TH antibody (1:100; Millipore, AB152) followed by secondary antibodies conjugated with Alexa-fluorescein. After masking autofluorescence background with TrueBlack Lipofuscin Autofluorescence Quencher (Biotium, #23007), brain slices were mounted with ProLong Gold Antifade Mountant (Thermo Fisher Scientific, #23007). For chromogenic method, endogenous peroxidase activity was quenched with 3% hydrogen peroxide for 30 min in MeOH. After blocking, brain slices were incubated with RIT2 antibody or TH antibody and the signals were developed through ImmPRESS-AP Horse Anti-Mouse IgG Polymer Kit (alkaline phosphatase; Vector laboratory, CA, MP5402) with Vector Blue (alkaline phosphatase substrate) or ImmPRESS HRP Horse Anti-Mouse IgG PLUS Polymer Kit (peroxidase; Vector laboratory, CA, MP7452) with ImmPACT VIP (peroxidase substrate, SK-4605) according to the manufacturer's instructions. After staining, brain slices were dehydrated by 100% EtOH followed by HistoChoice Clearing Agent (H2779, Sigma-Aldrich, a substitute of xylene) and mounted with VectaMount Permanent Mounting Medium (Vector laboratory, CA, H-5000-60). Images were taken by All-in-One Fluorescence Microscope (KEYENCE, IL).

IF staining for midbrain organoids

The midbrain organoids were washed with PBS^{-/-} and fixed in fresh 4% paraformaldehyde at room temperature for 30 to 60 min. (depending on organoid size) on a shaker. The samples were then washed with PBS^{-/-}, transferred to 30% sucrose, and stored at 4°C overnight. On the following day, the organoids were transferred to optimal cutting temperature compound and placed in cryomolds. The cryomolds were flash-frozen on dry ice and stored at -80°C until sectioning. The blocks were sectioned using a cryostat, and sections (30 μ m) were prepared and mounted on positively charged slides. IF staining was performed on the cryosections by first oven drying at 56°C for 30 min to adhere organoids to slide. Sections were washed twice with PBS for 5 min, and antigen retrieval was performed using Antigen Unmasking Solution, Citric Acid Based (Vector Laboratories, H-3300) with a pressure cooker (Presto Electronic, WI) for 15 min. Endogenous peroxidase activity was blocked using 3% H₂O₂ in methanol for 30 min. Sections were subsequently washed twice with 0.1% PBS-T for 5 min and blocked using 2.5% normal goat serum (Vector Laboratories, S-1012) for 30 min and incubated with RIT2 antibody (1:100; OriGene, CF501757) and TH

antibody (1:250; Millipore, AB152) overnight. Sections were then washed twice with 0.1% PBS-T for 5 min and incubated with secondary antibodies conjugated to Alexa fluorescein (1:500) for 1 hour followed by 0.1% PBS-T washes twice for 5 min. Autofluorescence was quenched with addition of TrueBlack Lipofuscin Autofluorescence Quencher (Biotium, 23007) and washed with PBS. Sections were allowed to dry and mounted with ProLong Diamond Antifade Mountant with DAPI (P36962, Thermo Fisher Scientific). Samples were imaged using Zeiss LSM780.

RNAscope in situ hybridization

Mouse brains were fixed overnight at 4°C in 4% paraformaldehyde. Fixed brains were stored at 4°C in a 30% sucrose solution until they sank. A series of coronal sections (15 μ m) were obtained with a cryostat (Leica, Wetzlar, Germany). Probes for *Rit2* (catalog no. 589041), *Cadps2* (catalog no. 529361), and *Th* (catalog no. 317621, Advanced Cell Diagnostics, CA) were applied and visualized according to the manual for RNAscope Multiplex Fluorescent V2 Assay (catalog no. 323100). FFPE human brain sections (7 μ m) were stained with probes for *RIT2* (catalog no. 534358), *CADPS2* (catalog no. 592718), and *TH* (catalog no. 441658) using RNAscope LS (Leica Biosystems) Multiplex assay kit (catalog no. 322800) and Leica Bond RX auto-stainer (Leica, IL). Autofluorescence signals were briefly masked with TrueBlack Lipofuscin Autofluorescence Quencher (Biotium, #23007) for 30 s.

Statistical analysis

The data analyses were performed using R/4.0.3 and GraphPad Prism 9 (GraphPad Software, CA, USA). For demographic information, the results were represented as mean $\pm$ SD for continuous and *N*(%) for discrete variables, respectively. For IHC, the results were reported as means $\pm$ SEM. The statistical significance of differences between two groups was determined using the unpaired two-tailed Student's *t* test. *P* value <0.05 was considered statistically significant.

Supplementary Materials

This PDF file includes:

Figs. S1 to S11

Legends for tables S1 to S9

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S9

REFERENCES AND NOTES

1. E. Paredes-Rodriguez, S. Vegas-Suarez, T. Morera-Herreras, P. De Deurwaerdere, C. Miguelez, The noradrenergic system in Parkinson's disease. *Front. Pharmacol.* **11**, 435 (2020).
2. N. Giguère, S. Burke Nanni, L. E. Trudeau, On cell loss and selective vulnerability of neuronal populations in Parkinson's disease. *Front. Neurol.* **9**, 455 (2018).
3. M. A. Nalls, C. Blauwendraat, C. L. Vallerga, K. Heilbron, S. Bandres-Ciga, D. Chang, M. Tan, D. A. Kia, A. J. Noyce, A. Xue, J. Bras, E. Young, R. von Coelln, J. Simón-Sánchez, C. Schulte, M. Sharma, L. Krohn, L. Pihlström, A. Siitonen, H. Iwaki, H. Leonard, F. Faghri, J. R. Gibbs, D. G. Hernandez, S. W. Scholz, J. A. Botia, M. Martinez, J.-C. Corvol, S. Lesage, J. Jankovic, L. M. Shulman, M. Sutherland, P. Tienari, K. Majamaa, M. Toft, O. A. Andreassen, T. Bangale, A. Brice, J. Yang, Z. Gan-Or, T. Gasser, P. Heutink, J. M. Shulman, N. W. Wood, D. A. Hinds, J. A. Hardy, H. R. Morris, J. Gratten, P. M. Visscher, R. R. Graham, A. B. Singleton; 23andMe Research Team; System Genomics of Parkinson's Disease Consortium; International Parkinson's Disease Genomics Consortium, Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: A meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
4. C. Blauwendraat, M. A. Nalls, A. B. Singleton, The genetic architecture of Parkinson's disease. *Lancet Neurol.* **19**, 170–178 (2020).

5. D. Chang, M. A. Nalls, I. B. Hallgrímsdóttir, J. Hunkapiller, M. van der Brug, F. Cai, G. A. Kerchner, G. Ayala, B. Bingol, M. Sheng, D. Hinds, T. W. Behrens, A. B. Singleton, T. R. Bhargava, R. R. Graham, A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* **49**, 1511–1516 (2017).
6. P. W. Hook, S. A. McClymont, G. H. Cannon, W. D. Law, A. J. Morton, L. A. Goff, A. S. McCallion, Single-cell RNA-Seq of mouse dopaminergic neurons informs candidate gene selection for sporadic Parkinson disease. *Am. J. Hum. Genet.* **102**, 427–446 (2018).
7. G. La Manno, D. Gyllborg, S. Codeluppi, K. Nishimura, C. Salto, A. Zeisel, L. E. Borm, S. R. W. Stott, E. M. Toledo, J. C. Villaescusa, P. Lönnerberg, J. Ryge, R. A. Barker, E. Arenas, S. Linnarsson, Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell* **167**, 566–580.e519 (2016).
8. J. F. Poulin, G. Caronia, C. Hofer, Q. Cui, B. Helm, C. Ramakrishnan, C. S. Chan, D. A. Dombeck, K. Deisseroth, R. Awatramani, Mapping projections of molecularly defined dopamine neuron subtypes using intersectional genetic approaches. *Nat. Neurosci.* **21**, 1260–1271 (2018).
9. J. F. Poulin, J. Zou, J. Drouin-Ouellet, K. Y. Kim, F. Cicchetti, R. B. Awatramani, Defining midbrain dopaminergic neuron diversity by single-cell gene expression profiling. *Cell Rep.* **9**, 930–943 (2014).
10. A. Saunders, E. Z. Macosko, A. Wysoker, M. Goldman, F. M. Krienen, H. de Rivera, E. Bien, M. Baum, L. Bortolin, S. Wang, A. Goeva, J. Nemesh, N. Kamitaki, S. Brumbaugh, D. Kulp, S. A. McCarroll, Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030.e1016 (2018).
11. K. Tiklova, A. K. Björklund, L. Lahti, A. Fiorenzano, S. Nolbrant, L. Gillberg, N. Volakakis, C. Yokota, M. M. Hilscher, T. Hauling, F. Holmström, E. Joodmardi, M. Nilsson, M. Parmar, T. Perlmann, Single-cell RNA sequencing reveals midbrain dopamine neuron diversity emerging during mouse brain development. *Nat. Commun.* **10**, 581 (2019).
12. J. Bryois, N. G. Skene, T. F. Hansen, L. J. A. Kogelman, H. J. Watson, Z. Liu, L. Brueggeman, G. Breen, C. M. Bulik, E. Arenas, J. Hjerling-Leffler, P. F. Sullivan, Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat. Genet.* **52**, 482–493 (2020).
13. D. Agarwal, C. Sandor, V. Volpato, T. M. Caffrey, J. Monzón-Sandoval, R. Bowden, J. Alegre-Abarrategui, R. Wade-Martins, C. Webber, A single-cell atlas of the human substantia nigra reveals cell-specific pathways associated with neurological disorders. *Nat. Commun.* **11**, 4183 (2020).
14. J. D. Welch, V. Kozareva, A. Ferreira, C. Vanderburg, C. Martin, E. Z. Macosko, Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887.e1817 (2019).
15. S. Smajic, C. A. Prada-Medina, Z. Landoulsi, J. Ghelfi, S. Delcambre, C. Dietrich, J. Jarazo, J. Henck, S. Balachandran, S. Pachchek, C. M. Morris, P. Antony, B. Timmermann, S. Sauer, S. L. Pereira, J. C. Schwamborn, P. May, A. Grunewald, M. Spielmann, Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. *Brain* **145**, 964–978 (2021).
16. T. Kamath, A. Abdullaouf, S. J. Burris, J. Langlieb, V. Gazestani, N. M. Nadaf, K. Balderrama, C. Vanderburg, E. Z. Macosko, Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. *Nat. Neurosci.* **25**, 588–595 (2022).
17. L. Zecca, R. Fariello, P. Riederer, D. Sulzer, A. Gatti, D. Tampellini, The absolute concentration of nigral neuromelanin, assayed by a new sensitive method, increases throughout the life and is dramatically decreased in Parkinson's disease. *FEBS Lett.* **510**, 216–220 (2002).
18. J. H. Kordower, C. W. Olanow, H. B. Dodiya, Y. Chu, T. G. Beach, C. H. Adler, G. M. Halliday, R. T. Bartus, Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. *Brain* **136**, 2419–2431 (2013).
19. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
20. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.-r. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
21. O. Franzén, L.-M. Gan, J. L. M. Björkengren, PanglaoDB: A web server for exploration of mouse and human single-cell RNA sequencing data. *Database* **2019**, baz046 (2019).
22. N. Pankratz, G. W. Beecham, A. L. DeStefano, T. M. Dawson, K. F. Doherty, S. A. Factor, T. H. Hamza, A. Y. Hung, B. T. Hyman, A. J. Ivins, D. Krainc, J. C. Latourelle, L. N. Clark, K. Marder, E. R. Martin, R. Mayeux, O. A. Ross, C. R. Scherzer, D. K. Simon, C. Tanner, J. M. Vance, Z. K. Wszolek, P. M. Zabetian, R. H. Myers, H. Payami, W. K. Scott, T. Foroud; PD GWAS Consortium, Meta-analysis of Parkinson's disease: Identification of a novel locus, RIT2. *Ann. Neurol.* **71**, 370–384 (2012).
23. I. P. Grigor'ev, E. G. Sukhorukova, E. A. Kolos, D. E. Korzhnevskii, Neuromelanin in neurons of substantia nigra lacking tyrosine hydroxylase. *Morfologiya* **141**, 65–67 (2012).
24. L. Sarrafha, G. M. Parfitt, R. Reyes, C. Goldman, E. Coccia, T. Kareva, T. Ahfeldt, High-throughput generation of midbrain dopaminergic neuron organoids from reporter human pluripotent stem cells. *STAR Protoc.* **2**, 100463 (2021).
25. J. F. Poulin, Z. Gaertner, O. A. Moreno-Ramos, R. Awatramani, Classification of midbrain dopamine neurons using single-cell gene expression profiling approaches. *Trends Neurosci.* **43**, 155–169 (2020).
26. L. Tian, M. Al-Nusaif, X. Chen, S. Li, W. Le, Roles of transcription factors in the development and reprogramming of the dopaminergic neurons. *Int. J. Mol. Sci.* **23**, 845 (2022).
27. Y. Okita, A. N. Rcom-H'cheo-Gauthier, M. Goulding, R. S. Chung, P. Faller, D. L. Pountney, Metallothionein, copper and alpha-synuclein in alpha-synucleinopathies. *Front. Neurosci.* **11**, 114 (2017).
28. Y. Liu, Q. Zhou, M. Tang, N. Fu, W. Shao, S. Zhang, Y. Yin, R. Zeng, X. Wang, G. Hu, J. Zhou, Upregulation of alphaB-crystallin expression in the substantia nigra of patients with Parkinson's disease. *Neurobiol. Aging* **36**, 1686–1691 (2015).
29. S. Z. Lu, Y. S. Guo, P. Z. Liang, S. Z. Zhang, S. Yin, Y. Q. Yin, X. M. Wang, F. Ding, X. S. Gu, J. W. Zhou, Suppression of astrocytic autophagy by α B-crystallin contributes to α -synuclein inclusion formation. *Transl. Neurodegener.* **8**, 3 (2019).
30. Q. Wang, Y. Zhang, M. Wang, W. M. Song, Q. Shen, A. McKenzie, I. Choi, X. Zhou, P. Y. Pan, Z. Yue, B. Zhang, The landscape of multiscale transcriptomic networks and key regulators in Parkinson's disease. *Nat. Commun.* **10**, 5234 (2019).
31. A. Lunati, S. Lesage, A. Brice, The genetic landscape of Parkinson's disease. *Rev. Neurol.* **174**, 628–643 (2018).
32. A. R. Dunn, K. A. Stout, M. Ozawa, K. M. Lohr, C. A. Hoffman, A. I. Bernstein, Y. Li, M. Wang, C. Sgobio, N. Sastry, H. Cai, W. M. Caudle, G. W. Miller, Synaptic vesicle glycoprotein 2C (SV2C) modulates dopamine release and is disrupted in Parkinson disease. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E2253–E2262 (2017).
33. S. Jin, C. F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C. H. Kuan, P. Myung, M. V. Pliukis, Q. Nie, Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
34. T. Zuo, J. Y. Qin, J. Chen, Z. Shi, M. Liu, X. Gao, D. Gao, Involvement of N-cadherin in the protective effect of glial cell line-derived neurotrophic factor on dopaminergic neuron damage. *Int. J. Mol. Med.* **31**, 561–568 (2013).
35. A. Kania, R. Klein, Mechanisms of ephrin-Eph signalling in development, physiology and disease. *Nat. Rev. Mol. Cell Biol.* **17**, 240–256 (2016).
36. M. Wang, W. M. Song, C. Ming, Q. Wang, X. Zhou, P. Xu, A. Krek, Y. Yoon, L. Ho, M. E. Orr, G. C. Yuan, B. Zhang, Guidelines for bioinformatics of single-cell sequencing data analysis in Alzheimer's disease: Review, recommendation, implementation and application. *Mol. Neurodegener.* **17**, 17 (2022).
37. K. Bossers, G. Meerhoff, R. Balesar, J. W. van Dongen, C. G. Kruse, D. F. Swaab, J. Verhaagen, Analysis of gene expression in Parkinson's disease: Possible involvement of neurotrophic support and axon guidance in dopaminergic cell death. *Brain Pathol.* **19**, 91–107 (2009).
38. C. G. Sweeney, P. J. Kearney, R. R. Fagan, L. A. Smith, N. C. Bolden, R. Zhao-Shea, I. V. Rivera, J. Kolpakova, J. Xie, G. Gao, A. R. Tapper, G. E. Martin, H. E. Melikian, Conditional, inducible gene silencing in dopamine neurons reveals a sex-specific role for Rit2 GTPase in acute cocaine response and striatal function. *Neuropsychopharmacology* **45**, 384–393 (2020).
39. K. J. Doorn, T. Moors, B. Drukarch, W. van de Berg, P. J. Lucassen, A. M. van Dam, Microglial phenotypes and toll-like receptor 2 in the substantia nigra and hippocampus of incidental Lewy body disease cases and Parkinson's disease patients. *Acta Neuropathol. Commun.* **2**, 90 (2014).
40. K. Imamura, N. Hishikawa, M. Sawada, T. Nagatsu, M. Yoshida, Y. Hashizume, Distribution of major histocompatibility complex class II-positive microglia and cytokine profile of Parkinson's disease brains. *Acta Neuropathol.* **106**, 518–526 (2003).
41. P. L. McGeer, S. Itagaki, B. E. Boyes, E. G. McGeer, Reactive microglia are positive for HLA-DR in the substantia nigra of Parkinson's and Alzheimer's disease brains. *Neurology* **38**, 1285–1291 (1988).
42. J. Miklossy, D. D. Doudet, C. Schwab, S. Yu, E. G. McGeer, P. L. McGeer, Role of ICAM-1 in persisting inflammation in Parkinson disease and MPTP monkeys. *Exp. Neurol.* **197**, 275–283 (2006).
43. Y. Ouchi, E. Yoshikawa, Y. Sekine, M. Futatsubashi, T. Kanno, T. Ogasu, T. Torizuka, Microglial activation and dopamine terminal loss in early Parkinson's disease. *Ann. Neurol.* **57**, 168–175 (2005).
44. A. Gerhard, N. Pavese, G. Hotton, F. Turkheimer, M. Es, A. Hammers, K. Eggert, W. Oertel, R. B. Banati, D. J. Brooks, In vivo imaging of microglial activation with [11 C](R)-PK11195 PET in idiopathic Parkinson's disease. *Neurobiol. Dis.* **21**, 404–412 (2006).
45. C. Knott, G. P. Wilkin, G. Stern, Astrocytes and microglia in the substantia nigra and caudate-putamen in Parkinson's disease. *Parkinsonism Relat. Disord.* **5**, 115–122 (1999).
46. B. Mirza, H. Hadberg, P. Thomsen, T. Moos, The absence of reactive astrocytosis is indicative of a unique inflammatory process in Parkinson's disease. *Neuroscience* **95**, 425–432 (2000).
47. M. Calvo, N. Zhu, C. Tsantoulas, Z. Ma, J. Grist, J. A. Loeb, D. L. Bennett, Neuregulin-ErbB signaling promotes microglial proliferation and chemotaxis contributing to microgliosis and pain after peripheral nerve injury. *J. Neurosci.* **30**, 5437–5450 (2010).
48. Y. Xiang, T. Liu, H. Yang, F. Gao, H. Xiang, A. Manyande, Y. Tian, X. Tian, NRG1-ErbB signalling promotes microglia activation contributing to incision-induced mechanical allodynia. *Eur. J. Pain* **19**, 686–694 (2015).

49. Y. N. Chen, H. H. Sha, Y. W. Wang, Q. Zhou, P. Bhuiyan, N. N. Li, Y. N. Qian, H. Q. Dong, Histamine 2/3 receptor agonists alleviate perioperative neurocognitive disorders by inhibiting microglia activation through the PI3K/AKT/FoxO1 pathway in aged rats. *J. Neuroinflammation* **17**, 217 (2020).
50. L. T. Guo, S. Q. Wang, J. Su, L. X. Xu, Z. Y. Ji, R. Y. Zhang, Q. W. Zhao, Z. Q. Ma, X. Y. Deng, S. P. Ma, Baicalin ameliorates neuroinflammation-induced depressive-like behavior through inhibition of toll-like receptor 4 expression via the PI3K/AKT/FoxO1 pathway. *J. Neuroinflammation* **16**, 95 (2019).
51. Y. Zhang, Y. Xiang, X. Wang, L. Zhu, H. Li, S. Wang, X. Pan, H. Zhao, Cerebral dopamine neurotrophic factor protects microglia by combining with AKT and by regulating FoxO1/mTOR signaling during neuroinflammation. *Biomed. Pharmacother.* **109**, 2278–2284 (2019).
52. R. K. Shahidepour, R. E. Higdon, N. G. Crawford, J. H. Neltner, E. T. Ighodaro, E. Patel, D. Price, P. T. Nelson, A. D. Bachstetter, Dystrophic microglia are associated with neurodegenerative disease and not healthy aging in the human brain. *Neurobiol. Aging* **99**, 19–27 (2021).
53. W. J. Streit, N. W. Sammons, A. J. Kuhns, D. L. Sparks, Dystrophic microglia in the aging human brain. *Glia* **45**, 208–212 (2004).
54. A. Rigamonti, G. G. Repetti, C. Sun, F. D. Price, D. C. Reny, F. Rapino, K. Weisinger, C. Benkler, Q. P. Peterson, L. S. Davidow, E. M. Hansson, L. L. Rubin, Large-scale production of mature neurons from human pluripotent stem cells in a three-dimensional suspension culture system. *Stem Cell Reports* **6**, 993–1008 (2016).
55. L. Schirmer, D. Velmeshev, S. Holmqvist, M. Kaufmann, S. Werneburg, D. Jung, S. Vistnes, J. H. Stockley, A. Young, M. Steindel, B. Tung, N. Goyal, A. Bhaduri, S. Mayer, J. B. Engler, O. A. Bayraktar, R. J. M. Franklin, M. Haeussler, R. Reynolds, D. P. Schafer, M. A. Friese, L. R. Shiow, A. R. Kriegstein, D. H. Rowitch, Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).
56. H. Mathys, J. Davila-Velderrain, Z. Peng, F. Gao, S. Mohammadi, J. Z. Young, M. Menon, L. He, F. Abdurrob, X. Jiang, A. J. Martorell, R. M. Ransohoff, B. P. Hafler, D. A. Bennett, M. Kellis, L.-H. Tsai, Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
57. B. B. Lake, S. Chen, M. Hoshi, N. Plongthongkum, D. Salamon, A. Knoten, A. Vijayan, R. Venkatesh, E. H. Kim, D. Gao, J. Gaut, K. Zhang, S. Jain, A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat. Commun.* **10**, 2832 (2019).
58. L. van der Maaten, G. Hinton, Visualizing non-metric similarities in multiple maps. *Mach. Learn.* **87**, 33–55 (2012).
59. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv:1802.03426 [stat.ML] (2018).
60. M. Tang, Y. Kaymaz, B. L. Logeman, S. Eichhorn, Z. S. Liang, C. Dulac, T. B. Sackton, Evaluating single-cell cluster stability using the Jaccard similarity index. *Bioinformatics* **37**, 2212–2214 (2021).
61. P. Germain, A. Lun, W. Macnair, M. Robinson, Doublet identification in single-cell sequencing data using *scDblFinder*. *F1000Res* **10**, 979 (2021).
62. P.-L. Germain, A. Sonrel, M. D. Robinson, pipeComp, a general framework for the evaluation of computational pipelines, reveals performant single cell RNA-seq preprocessing tools. *Genome Biol.* **21**, 227 (2020).
63. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
64. M. Wang, Y. Zhao, B. Zhang, Efficient test and visualization of multi-set intersections. *Sci. Rep.* **5**, 16923 (2015).

Acknowledgments: We thank the members from Z.Y. laboratory and B.Z. laboratory for the critical reading and comments of the manuscript. **Funding:** This work was supported by grants PF-RC-936279, NIH/P20NS123220 (to Z.Y.), NIH/R21NS109895 (to Z.Y.), NIH/U01AG046170 (to B.Z.), NIH/R01AG054008 (to J.F.C.), NIH/R01NS095252 (to J.F.C.), NIH/R01NS086736 (to J.F.C.); the Rainwater Charitable Foundation and a gift from S. Katz and J. Martin (to J.F.C.); ASAP/MJFF and NIH/R01NS114239 (to J.W.B.); NIH/K01 AG078442 (to I.C.); NIH/K01 AG070326 (to K.F.); NYSTEM-C32561GG (to L.S.); NIH/T32MH087004 (to M.L.); the National Health and Medical Research Council Investigator grant (1175781) (to J.E.P.); ASAP/MJFF (to J.E.P.); and a Fok Family Fellowship (to J.E.P.). Z.Y., B.Z., J.F.C. and J.W.B. are also supported by Parkinson's Foundation (PF-RC-936279). **Author contributions:** Z.Y. and B.Z. conceived the project. Q.W. initiated the project and performed data analysis. M.W. performed data analysis. I.C. and M.L. performed IHC and IF validation experiments. K.F., C.D.S., and J.F.C. contributed to validation experiments. L.H. prepared brain samples for sequencing. K.G.B. and R.S. provided support in snRNA-seq. D.A.D. and X.S. provided human brain samples and related information. L.S. provided the midbrain organoid samples for the IF staining and the organoid scRNA-seq dataset. D.N. remapped the organoid scRNA-seq dataset for integration. T.A. and J.B. contributed to the midbrain organoid sample preparation and scRNA-seq. Z.Y., Q.W., M.W., I.C., L.S., and B.Z. wrote the manuscript. L.H., K.G.B., R.S., K.F., J.F.C., D.A.D., L.S., J.B., D.N., J.P., and X.S. provided comments and edits. All authors approved the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The raw data of our snRNA-seq are available from the Gene Expression Omnibus (GEO) under the accession code GSE184950. The raw data from Agarwal *et al.* (13) are available from GEO under the accession code GSE140231. The raw data from Welch *et al.* (14) are available from GEO under the accession code GSE126836. The code for data analysis in this study is deposited to Zenodo (<https://zenodo.org/records/10119032>).

Submitted 20 May 2023

Accepted 12 December 2023

Published 10 January 2024

10.1126/sciadv.adi8287