Single-nucleus RNA-sequencing reveals oligodendrocytes and their progenitors as vulnerable cell types in prefrontal cortex and anterior cingulate of brains with Parkinson's disease

Mohammad Dehestani[1,2], Cornelis Blauwendraat[3], Thomas Gasser[1,2*], Vikas Bansal[2*]

1. Department of Neurodegeneration, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tuebingen, Germany
2. German Center for Neurodegenerative Diseases (DZNE), Tuebingen, Germany
3. Laboratory for Neurogenetics, National Institute of Health NIH, Bethesda, Maryland, USA

[*]**Correspondence**: thomas.gasser@uni-tuebingen.de, vikas.bansal@dzne.de

Abstract

Several prior studies have proposed the involvement of various brain regions and cell types in Parkinson's disease (PD) pathology. Here, we performed snRNA-seq on the prefrontal cortex and anterior cingulate regions from post-mortem control and PD brain tissue. We found a significant association of oligodendrocytes and oligodendrocyte precursor cells with PD-linked risk loci and report several dysregulated genes and pathways, including regulation of tau-protein kinase activity, regulation of inclusion body assembly and protein processing involved in protein targeting to mitochondria.

Keywords: Parkinson's disease, oligodendrocytes, oligodendrocyte precursor cells, snRNA-seq, prefrontal cortex, anterior cingulate cortex

Parkinson's disease (PD) is the second most common neurodegenerative disorder. It is characterized by the pathologic aggregation of alpha-synuclein and its pathology is known to progress in a predictable spatiotemporal manner. The spread of the disease pathology commences in the olfactory bulb and the lower brainstem and moves through the substantia nigra pars compacta (SN) in the midbrain, eventually reaching the meso- and neocortical areas[1,2]. PD arises from a complex interplay of various factors, such as aging, genetic predisposition, and environmental factors. While the cause of PD is unknown in most cases, specific mutations in genes such as *LRRK2* and *GBA1* have been identified to significantly increase the risk of developing the disease, although likely by affecting different molecular pathways. This may be one reason why the impact of *LRRK2* and *GBA1* mutations on clinical presentations may differ[3]. In recent decades, transcriptome profiling has emerged as a preeminent methodology for exploring human pathologies at the molecular and cellular level. In PD, alpha-synuclein pathology has been shown to be associated with the transcriptional programs of various brain cell types, including neurons and glial cells[4]. Furthermore, although the degeneration of dopaminergic neurons in PD primarily occurs in the SN region, Lewy bodies can form in other brain regions, such as the limbic system and the prefrontal cortex[5]. Moreover, there has been no investigation into the relative impacts of *LRRK2* and *GBA1* risk genotypes on transcriptional programs across several brain regions of PD patients.

In this study, we performed single-nucleus RNA-sequencing (snRNA-seq) on the prefrontal cortex (PFC) and anterior cingulate (ACC) brain regions from the same individuals (2 *LRRK2* PD, 2 *GBA1* PD and 2 Healthy Controls) to identify neuronal and non-neuronal cell-type differences (Figure 1A and Table S1). After data cleaning and quality control (see Methods), 88,876 high-quality single nuclei were retained. The clustering of these high-quality nuclei identified 13 clusters covering major broad cell types in the brain, i.e. excitatory neurons (ExN), inhibitory neurons (InN), oligodendrocytes (ODCs), oligodendrocyte precursor cells (OPCs), microglia (MG), astrocytes (Astro) and vascular cells (Vas) (Figure 1B). While ~50% of the total nuclei were annotated as ODCs, only 4.3% and 0.5% of nuclei accounted for OPCs and Vas cells, respectively (Figure 1C). On average, we obtained 7406 high-quality nuclei per sample ranging from 3.6% to 16.5% of the total nuclei (Figure 1C and Table S2). None of the clusters was significantly enriched for a particular brain region or mutation group (Figure 1D and Table S2). Although the percentage of each cell type varied between the samples, no consistent cell-type enrichment could be identified based on an individual sample, except for the ExN5 cluster, which was enriched in Sample 2 (Figure 1E). The clusters were annotated based on the expression of well-known cell-type markers (Figure 1F and Table S3). The top 5 markers of each cell type cluster are shown in Figure 1G.

Integration of snRNA-seq data with PD genome-wide association studies (GWAS)[6] showed the strongest association of ODCs and OPCs with PD-linked risk loci (Figure 2A and TableS4). Next, we performed differential expression gene analysis to identify dysregulated genes between PD and controls in ODCs and OPCs. While 1040 and 543 genes were differentially expressed between *GBA1* and control samples in ODCs and OPCs, respectively, only 278 genes in ODCs and 108 genes in OPCs were differentially expressed between *LRRK2* and control samples (Table S5). The fold change (FC) of these differentially expressed genes (DEGs) was significantly correlated with the bulk RNA-seq data (Figure 2B and Table S6). The correlation for *GBA1* DEGs was better than *LRRK2* DEGs owing to the higher number of genes. The FC of the top 10 up-regulated and top 10 down-regulated genes are shown in Figure 2C, along with the FC in bulk RNA-seq data. While 5 genes were up-regulated in all four comparisons, 11 genes were down-regulated in all four comparisons (Figure 2D). Earlier studies have found some of these genes to be dysregulated or at least associated with PD. For instance, recently *MAPT* was found to be up-regulated in oligodendrocytes of the substantia nigra in PD patients[7], and *SUMF1* has been associated with lysosomal storage disorders[8]. Enriched biological processes in up-regulated DEGs (Figure 2E and Table S7) exhibited negative regulation of inclusion body assembly (in *GBA1*/Control ODCs), regulation of microtubule nucleation (in *LRRK2*/Control ODCs), positive regulation of tau-protein kinase activity (in *GBA1*/Control OPCs) and regulation of protein polymerization (in *LRRK2*/Control OPCs). On the other hand, enriched biological processes in down-regulated DEGs exhibited protein processing involved in protein targeting to mitochondria (in *GBA1*/Control ODCs), regulation of potassium ion transport (in LRRK2/Control ODCs), modulation of chemical synaptic transmission (in *GBA1*/Control OPCs) and cholesterol biosynthetic process (in *LRRK2*/Control OPCs). Intriguingly, the pseudotime analysis indicated a significantly higher pseudotime of *GBA1* and *LRRK2* ODCs compared to controls (Figure S1A-B). Enriched biological processes of the pseudotime-associated genes (n=287) exhibited membrane depolarization during action potential and chemical synaptic transmission terms (Figure S1C-D and Table S8). Additionally, the overlap between pseudotime-associated genes and *GBA1* DEGs in ODCs suggested an enrichment of biological processes like vocal learning and retinal ganglion cell axon guidance (Figure S1E and Table S8). Interestingly, co-expression network analysis in ODCs revealed 7 co-expression modules, out of which module 2 and 6 were enriched in *GBA1* samples (Figure S1F-G and Table S9). Enriched biological processes of module 2 exhibited negative regulation of inclusion body assembly, while module 6 exhibited a cellular response to histamine and neurotransmitter secretion (Table S10).

To summarize, in this short report, we found a significant enrichment of ODC and OPC expressed genes within PD GWAS risk loci and revealed several dysregulated genes

and pathways. Although two decades ago, Wakabyashi et al., observed an abnormal accumulation of alpha-synuclein in the oligodendrocytes within the substantia nigra of PD patients[9], the role of oligodendrocytes in PD has been mostly neglected. Reynolds and colleagues found a significant enrichment of PD heritability in glial cell types like oligodendrocytes and astrocytes[10]. By integrating GWAS results with single-cell transcriptomic data, Bryois et al., and Agarwal et al., observed oligodendrocytes and oligodendrocyte precursor cells to be significantly associated with PD[11,12]. Moreover, an increasing amount of evidence points out that oligodendrocytes also provide support to neurons via mechanisms beyond the insulating function of myelin[13]. Therefore, it is tempting to speculate that the abnormally regulated pathways, which extend beyond myelination in ODCs, may contribute to the pathology of PD. In line with this, our snRNA-seq data from the PFC and ACC regions of post-mortem control and PD brain tissue showed a dramatic association of ODCs and OPCs with PD-linked risk loci, and revealed dysregulated pathways including vocal learning. Curiously, changes in voice have been proposed as a potential early diagnostic biomarker for PD[14]. However, we acknowledge that the respective sample sizes in this study are relatively small. In the future, it will be essential to integrate and replicate the results in a larger cohort. Nevertheless, we anticipate that our pilot study will serve as a valuable resource and prompt further research into the involvement of oligodendrocytes in the pathology of PD.

Methods

Samples used in this study

The research was conducted using fresh-frozen postmortem brain tissues obtained from four PD patients, two of which had the *LRRK2* p.G2019S mutation and the other two had the *GBA1* mutation (one with p.L444P and one with p.E326K), along with two healthy controls. All donors were males, aged between 65 and 80, and PD patients had Lewy body Braak stages of 5-6. Two brain regions, namely the prefrontal cortex and anterior cingulate cortex, were investigated for each donor. All tissues were procured from the Netherlands brain bank except one *LRRK2* brain from UCL Queen Square Brain Bank for Neurological Disorders, following the policies and regulations of the institutional ethics board at the University Hospital Tuebingen in Germany. For transcriptome analysis, single nuclei and RNA were extracted from all samples, and snRNA-seq and bulk RNA-seq were performed. The clinical characteristics of the donors have been elaborated in Table S1.

Generation of single nuclei from postmortem human brains

The process of isolating nuclei entails the utilization of a detergent lysis technique, where a detergent is employed to break down the cellular membranes, followed by the centrifugal separation of the nuclei. In brief, 300 mg of post-mortem brain tissue was

dounce-homogenized in 2 ml of Nuclei EZ Prep Lysis Buffer (Sigma Aldrich, MA, USA) spiked with 0.2 U µl−1 RNase inhibitor (Sigma Aldrich, MA, USA) , 3.3 µl DTT (Thermo Fisher Scientific, MA, USA) and 33 µl of 10% Triton X100 which were added before incubating on ice for 5 minutes in a final volume of 10 ml. Homogenized tissue was washed with 3-4 ml washing buffer which was fresh PBSB 1% and then was filtered through a 70-µm cell strainer (BD Bioscience, NY, USA). Then using a long tube, 10 ml of 1.8 M ice-cold sucrose cushion solution is added to each sample i.e the roughly 3 ml lysate. After carefully and completely discarding the supernatant and the sucrose cushion layer containing debris and myelin, 1 ml PBS buffer added to resuspend the nuclei and 4 ml nuclei suspension buffer (1% BSA-PBS solution). It was finally centrifuged on 500 g for 5 minutes. At the end, 2 µl DAPI (Sigma Aldrich, MA, USA) with the concentration of 1:100 was added to stain the nuclei. Final centrifugation step on 500 g for 5 minutes was preceded by incubating the DAPI added suspension for 15-20 minutes in the cold & dark room on the rotation wheel at 4 C. Sorting buffer which consists of 99 µl PBS and 1 µl RNase inhibitor (Sigma Aldrich, MA, USA) was added to resuspend the nuclei and make them ready for quality/quantity inspection and then to run on a 10x genomic chromium controller. Quality assessment was performed using fluorescence-activated cell sorting (FACS) to detect all DAPI-positive events, i.e. individual nuclei comprising more than 95 % of all events.

Droplet-based snRNA-seq using 10x Genomics

Single-nuclei suspension concentration was determined by automatic cell counting (DeNovix CellDrop, DE, USA) using an AO/PI viability assay (DeNovix, DE, USA) and counting nuclei as dead cells. Single-nucleus gene expression libraries were generated using the 10x Chromium Next gel beads-in-emulsion (GEM) Single Cell 3’ Reagent Kit v3.1 (10x Genomics, CA, USA) according to manufacturer’s instructions. In brief, cells were loaded on the Chromium Next GEM Chip G, which was subsequently run on the Chromium Controller (10x Genomics, CA, USA) to partition cells into GEMs. Cell lysis and reverse transcription of poly-adenylated mRNA occurred within the GEMs and resulted in cDNA with GEM-specific barcodes and transcript-specific unique molecular identifiers (UMIs). After breaking the emulsion, cDNA was amplified by PCR, enzymatically fragmented, end-repaired, extended with 3′ A-overhangs, and ligated to adapters. P5 and P7 sequences, as well as sample indices (Chromium i7 Multiplex kit, 10x Genomics, CA, USA), were added during the final PCR amplification step. The fragment size of the final libraries was determined using the Bioanalyzer High-Sensitivity DNA Kit (Agilent, CA, USA). Library concentration was determined using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, MA, USA). snRNA libraries were pooled and paired-end-sequenced on the Illumina NovaSeq 6000 platform (Illumina, CA, USA).

snRNA-seq quality control

Samples were demultiplexed using Illumina's bcl2fastq conversion tool and the 10x Genomics pipeline Cell Ranger count v6.0.1 to perform alignment against the 10x Genomics pre-built Cell Ranger reference GRCh38-2020-A (introns included), filtering, barcode counting, and UMI counting. As a default, a cut-off value of 200 unique molecular identifiers expressed in at least 3 cells was used to select nuclei of sufficient complexity for further analysis. Each sample's count was normalized by the SCTransform method in Seurat v4.1.0[15] with mitochondrial reads regressed out. Two approaches were combined for quality control: (1) Doublets and multiplets were filtered out using DoubletFinder v2.0.3[16] for each individual sample; (2) outliers with a high ratio of mitochondrial and ribosomal counts (each >10%) and cells with low a number of genes (N < 1000) were removed. The core statistical parameters of DoubletFinder used to build artificial doublets for true doublet classification were determined automatically using recommended settings. After applying these filtering steps on 105,781 input nuclei, the dataset contained 88,876 high-quality single nuclei that were eligible for further analysis.

Cell annotations and differential expression

After combining all samples into a single Seurat object, genes were projected into principal component space using the principal component analysis (RunPCA). Harmony R package[17] was used for integration as well as for removing unwanted effects across subjects. The first 12 PC dimensions of data processed with Harmony were used as inputs into the FindNeighbours, FindClusters (at 0.1 resolution obtained out of a range of tested resolutions (0.1, 0.2, 0.5, 1.0)) and RunUMAP functions of Seurat. In brief, a shared-nearest-neighbor graph was constructed on the basis of the Euclidean distance metric in principal component space, and cells were clustered using the Louvain algorithm. The RunUMAP function with default settings was used to calculate 2D UMAP coordinates and search for distinct cell populations. Cluster markers and differential expression testing was performed on Seurat "RNA" assay containing seurat log-normalized counts using default Wilcoxon method implemented in Seurat v4.1.0. Differential gene expression test between cases and controls was performed for each cell type using the Wilcoxon ranked sum method implemented within the FindMarkers function. Gene ontology enrichment analysis for biological processes was performed using EnrichR[18]. In addition, hierarchical clustering of enriched GO terms was performed for a set of paired comparisons, including *LRRK2* vs. HC and *GBA1* vs. HC brains, in which differentially expressed genes showed significant enrichment (Adjusted P-Value < 0.05).

Bulk RNA-seq, data quality control and differential gene expression analysis

Total RNA was extracted from brain tissue homogenate for all samples using RNeasy plus mini kit (Qiagen, Hilden, Germany) with on-column DNase treatment according to manufacturer's protocol. Final elution was made in 65 µl of dH_2O. RNA concentration was measured using the Qubit Fluorometric Quantitation and RNA Broad-Range Assay (Thermo Fisher Scientific, MA, USA). RNA Integrity Number (RIN) was determined using the Fragment Analyzer 5300 and the Fragment Analyzer RNA kit (Agilent, CA, USA) and presented a moderate degradation pattern with RIN value between 3.4 and 7. For library preparation, ribosomal RNA fraction was depleted using the QIAseq FastSelect rRNA removal (Qiagen, Hilden, Germany) from 100 ng of total RNA. Next, libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB) according to the manufacturer's instructions. Library molarity was determined by measuring the library size (mean 340 bp) using the Fragment Analyzer 5300 and the Fragment Analyzer DNA HS NGS fragment kit (Agilent, CA, USA) and the library concentration (mean 30 ng/µl) using Qubit Fluorometric Quantitation and dsDNA high sensitivity assay (Thermo Fisher Scientific, MA, USA). The libraries were denatured according to the manufacturer's instructions, diluted to 270 pM and sequenced as paired-end 100 bp reads on an Illumina NovaSeq 6000 (Illumina, CA, USA) with a mean sequencing depth of 55 million clusters per sample. Read quality of RNA-seq data in fastq files was assessed using ngs-bits to identify sequencing cycles with low average quality, adaptor contamination, or repetitive sequences from PCR amplification. Reads were aligned using STAR v2.7.9a[19] to the GRCh37. Both unsupervised principal component analysis on normalized counts as well as differential gene expression analysis were performed using the DESeq2 R package v1.22.2 with default parameters[20]. For each gene, gene expression fold changes ($\log_2$ fold change) were computed and FDR test was performed to assess the significance.

Cell-type association with genetic risk of PD

Association analysis of cell type-specific expressed genes with genetic risk of PD was performed using Multi-marker Analysis of GenoMic Annotation (MAGMA) v1.0.0, in order to identify disease-relevant cell types in the data[21]. MAGMA, as a gene set enrichment analysis method, tests the joint association of all SNPs in a gene with the phenotype, while accounting for LD structure between SNPs. Competitive gene set analysis was performed on SNP p-values from the latest PD GWAS summary statistics including 23andMe data and the publicly available European subset of 1000 Genomes Phase 3 was used as a reference panel to estimate LD between SNPs. SNPs were mapped to genes using NCBI GRCh37 build (annotation release 105). Gene boundaries were defined as the transcribed region of each gene. An extended window of 35 kb upstream and 10 kb downstream of each gene was added to the gene boundaries.

Gene sets used in MAGMA were filtered: FDR-corrected p-values < 0.05, percentage of cells of the cluster where the expression was detected > 0.5, and logFC > 0.25.

Trajectory analysis

Slingshot v2.2.0 was used to perform the trajectory analysis[22]. ODCs and OPCs as cell types expressing most genetic risk loci identified by cell type association analysis, were extracted together and were put through the slingshot for the pseudotime inference. According to the knowledge from biology, OPCs were set as root cells during the analysis. UMAP embeddings generated from Seurat were converted to a cell_data_set object using SeuratWrappers v0.2.0 and then used as input to perform graph learning and pseudotime measurement through reversed graph embedding with slingshot. To identify genes deriving the pseudotime axis among highly variable genes (n=2000) calculated after log-normalization in Seurat, we used a generalized additive model (GAM) implemented in tradeSeq R package to firstly predict the pseudotime-associated genes and then test their differential expressions within one lineage across conditions[23].

Weighted gene co-expression network analysis

Utilizing the hdWGCNA v0.2.16 software package[24], we conducted a single-cell level weighted gene coexpression network analysis to identify gene modules exhibiting comparable expression patterns in all ODCs. In consideration of their biological relevance and markedly substantial cell count, we performed the analysis only on ODCs. The initial phase involved the creation of metacells, i.e. aggregates of homogenous cells stemming from identical biological groups. The k-nearest neighbor (KNN) algorithm was used to identify groups of similar cells for aggregation, and then the average expression of these cells was calculated, resulting in a metacell gene expression matrix. A correlation adjacency matrix was created, which is essential for inferring co-expression relationships between genes. After calculating the appropriate soft power threshold, co-expression networks were then constructed using the ConstructNetwork function. Module eigengenes (MEs), a commonly used metric to summarize the gene expression profile of an entire co-expression module, were calculated by performing principal component analysis (PCA) on the subset of the gene expression matrix that comprises each module. Next, a function OverlapModulesDEGs was used to explore the overlap between modules and cell-type marker genes that were distinctly expressed in each mutation group, here called as marker genes, and were calculated using the Seurat embedded function of FindAllMarkers using only ODCs.

Ethics approval and consent to participate

Our study was approved by the Ethics Committee of the University Hospital and Medical Faculty Tuebingen.

Availability of data and materials

A Seurat object of the processed snRNA-seq data has been uploaded to Zenodo (https://doi.org/10.5281/zenodo.7886802).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by the Charitable Hertie Foundation and in part by the Intramural Research Programs of the National Institute on Aging (NIA). V.B. is supported by a Career Development Fellowship at DZNE Tuebingen.

Authors' contributions

M.D., T.G. and V.B. conceived the project. T.G. and V.B. supervised the study. M.D. performed all the computational analysis and nuclei isolation. C.B. performed the MAGMA analysis. M.D. and V.B. wrote the manuscript. All authors contributed to and reviewed the manuscript.

Acknowledgments

We thank the NGS Competence Center Tübingen (NCCT) core facility for performing 10x Genomics library preparations and sequencing. We also thank the Jonas Neher research group at DZNE Tübingen for providing the single nucleus isolation protocol and FACS facility. We would like to express our gratitude to the brain banks and the donors.

Main figure legends

Figure 1: Overview of snRNA-seq profiling in the human post-mortem brain tissues. (A) Schematic overview of the experimental plan. **(B)** Uniform manifold approximation and projection (UMAP) visualization of the snRNA-seq clusters from 88,876 high quality nuclei. **(C)** Percentage of nuclei for each cell type across samples, mutation group and brain regions. **(D)** UMAP embeddings of nuclei colored by mutation group and brain regions. **(E)** Barplot displaying the distribution of cell-type percentage sample-wise. **(F)** Violin plot illustrating the expression distribution of known gene markers. **(G)** Genes most up-regulated in identified cell types: excitatory neurons (ExN), inhibitory neurons (InN), oligodendrocytes (ODCs), oligodendrocyte precursor cells (OPCs), microglia (MG), astrocytes (Astro) and vascular cells (Vas).

Figure 2: Association of PD susceptibility with ODCs and OPCs. (A) Multi-marker analysis of genomic annotation (MAGMA) gene set enrichment based on the snRNA-seq data showed significant associations with oligodendrocytes (ODCs) and oligodendrocyte precursor cells (OPCs). **(B)** Spearman's correlation between the ODCs and OPCs DEGs (GBA1 or LRRK2 vs controls), and the bulk RNA-seq data, were calculated based on the logarithmic fold-change. **(C)** Top ten up- and down-regulated genes ranked by logarithmic fold-change in snRNA-seq. **(D)** Venn diagrams displaying the overlap between up-regulated (left) or down-regulated (right) genes identified in ODCs and OPCs. **(E)** Gene ontology enrichment analysis of up-regulated (left) or down-regulated (right) genes. Top five biological process terms for each gene list are indicated.

Supplementary figure and table legends

Figure S1: Pseudotime and weighted gene co-expression network analysis. (A) Trajectory and pseudotime inference by Slingshot. **(B)** Boxplots showing the distribution of pseudotime in ODCs and OPCs separated by mutation groups. The p-values are calculated using the Kolmogorov-Smirnov test. **(C)** Pseudotime-associated genes (n=287) were identified using a generalized additive model (GAM) implemented in tradeSeq R package. **(D)** Gene ontology enrichment analysis of pseudotime-associated genes. The top ten biological process terms are indicated. **(E)** Venn diagrams displaying the overlap between pseudotime-associated genes and ODCs and OPCs DEGs. **(F)** WGCNA dendrogram showing the different co-expression modules resulting from the network analysis in ODCs. Each leaf on the dendrogram represents a single gene, and the color at the bottom indicates the co-expression module assignment. **(G)** Results from the OverlapBarPlot function implemented in the hdWGCNA R package demonstrating the overlap between ODCs positive markers in each mutation group and the modules.

Table S1: Sample overview with meta-data.

Table S2: Number of nuclei for cell types across samples, mutation group and brain regions.

Table S3: Known and identified cluster markers.

Table S4: MAGMA snRNA-seq cell-type enrichment results for PD GWAS.

Table S5: Differentially expressed genes in snRNA-seq for ODCs and OPCs.

Table S6: Differentially expressed genes from bulk RNA-seq.

Table S7: Enriched gene ontology terms for DEGs in ODCs and OPCs.

Table S8: Pseudotime-associated genes and enriched gene ontology terms.

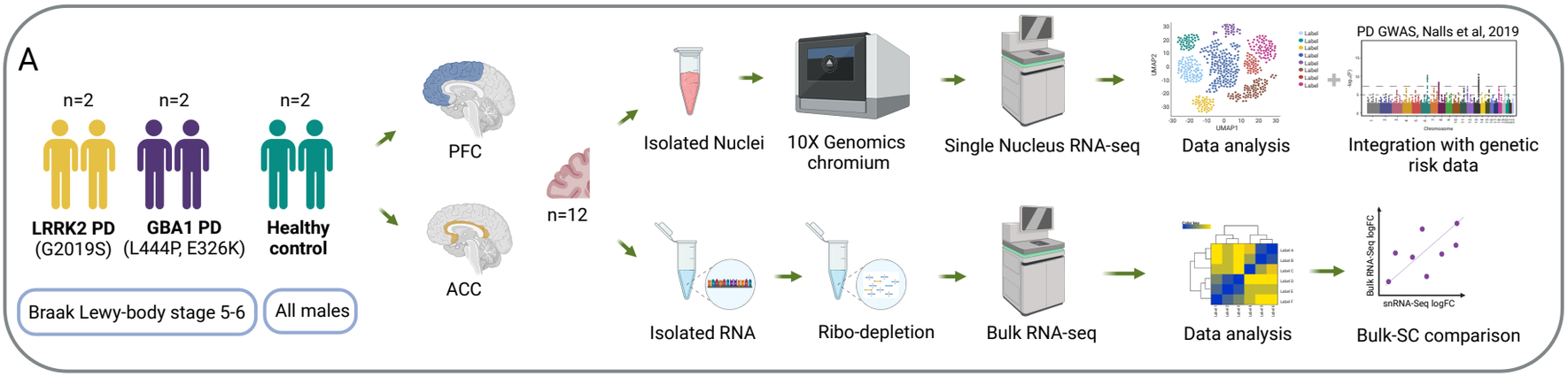
Table S9: Module genes from weighted gene co-expression network analysis.

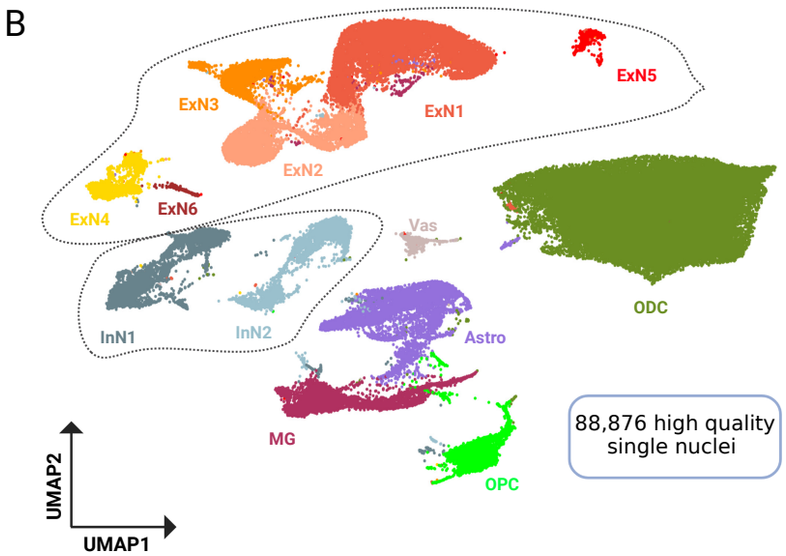
Table S10: Enriched gene ontology terms for WGCNA module genes.

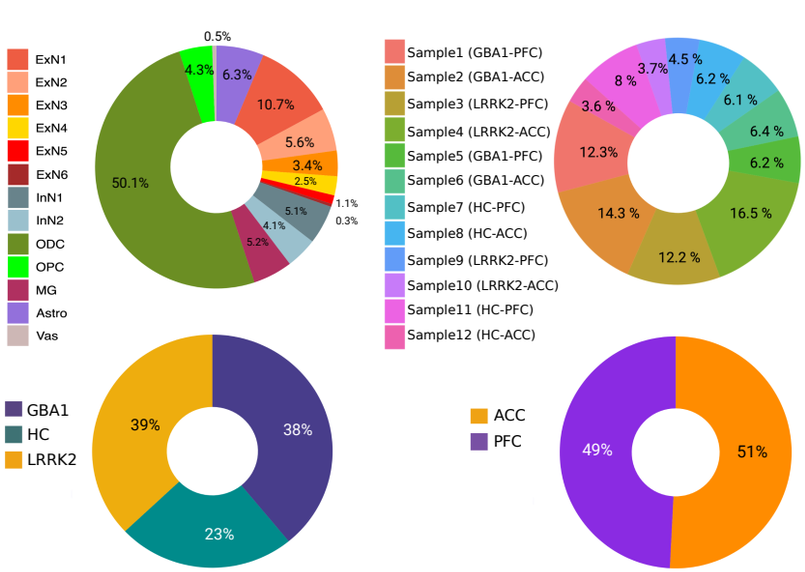
References

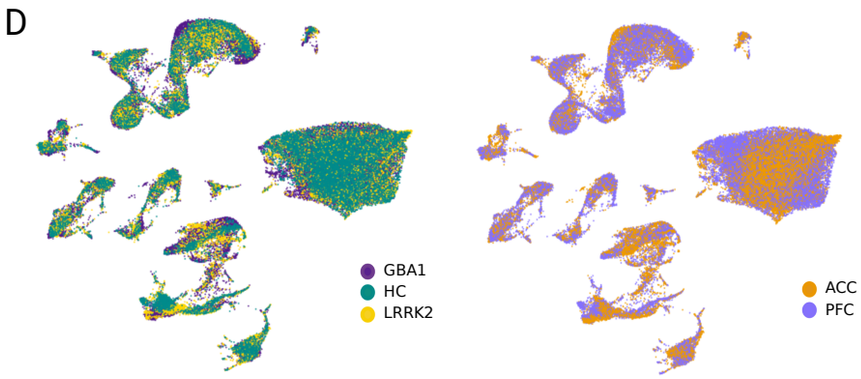
1. Braak, H. *et al.* Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **24**, 197–211 (2003).
2. Obeso, J. A. *et al.* Past, present, and future of Parkinson's disease: A special essay on the 200th Anniversary of the Shaking Palsy. *Mov. Disord.* **32**, 1264–1310 (2017).
3. Gan-Or, Z. *et al.* LRRK2 and GBA mutations differentially affect the initial presentation of Parkinson disease. *Neurogenetics* **11**, 121–125 (2010).
4. Olsen, A. L. & Feany, M. B. Glial α-synuclein promotes neurodegeneration characterized by a distinct transcriptional program in vivo. *Glia* **67**, 1933–1957 (2019).
5. Dumitriu, A. *et al.* Gene expression profiles in Parkinson disease prefrontal cortex implicate FOXO1 and genes under its transcriptional regulation. *PLoS Genet.* **8**, e1002794 (2012).
6. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
7. Lee, A. J. *et al.* Characterization of altered molecular mechanisms in Parkinson's disease through cell type-resolved multiomics analyses. *Sci Adv* **9**, eabo2467 (2023).
8. Zhao, Y.-W. *et al.* The Association Between Lysosomal Storage Disorder Genes and Parkinson's Disease: A Large Cohort Study in Chinese Mainland Population. *Front. Aging Neurosci.* **13**, 749109 (2021).
9. Wakabayashi, K., Hayashi, S., Yoshimoto, M., Kudo, H. & Takahashi, H. NACP/alpha-synuclein-positive filamentous inclusions in astrocytes and oligodendrocytes of Parkinson's disease brains. *Acta Neuropathol.* **99**, 14–20 (2000).
10. Reynolds, R. H. *et al.* Moving beyond neurons: the role of cell type-specific gene regulation in Parkinson's disease heritability. *NPJ Parkinsons Dis* **5**, 6 (2019).
11. Bryois, J. *et al.* Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat. Genet.* **52**, 482–493 (2020).
12. Agarwal, D. *et al.* A single-cell atlas of the human substantia nigra reveals cell-specific pathways associated with neurological disorders. *Nat. Commun.* **11**, 4183 (2020).
13. Pöyhönen, S., Er, S., Domanskyi, A. & Airavaara, M. Effects of Neurotrophic Factors in Glial Cells in the Central Nervous System: Expression and Properties in Neurodegeneration and Injury. *Front. Physiol.* **10**, 486 (2019).
14. Ma, A., Lau, K. K. & Thyagarajan, D. Voice changes in Parkinson's disease: What are they telling us? *J. Clin. Neurosci.* **72**, 1–7 (2020).
15. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).

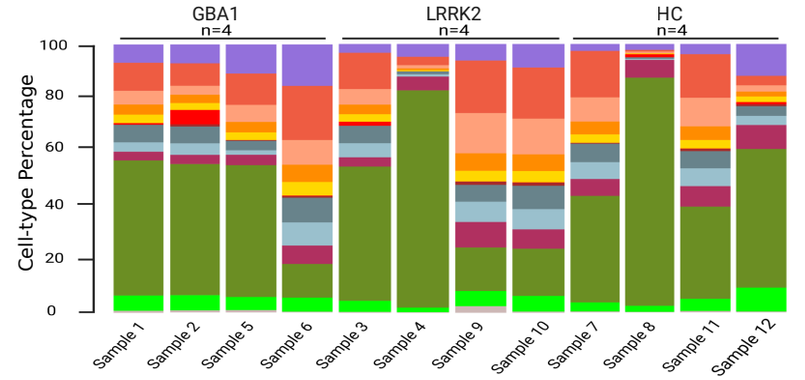
16. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329–337.e4 (2019).
17. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
18. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
19. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
20. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
21. Skene, N. G. *et al.* Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
22. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
23. Van den Berge, K. *et al.* Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).
24. Morabito, S. *et al.* Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat. Genet.* **53**, 1143–1155 (2021).

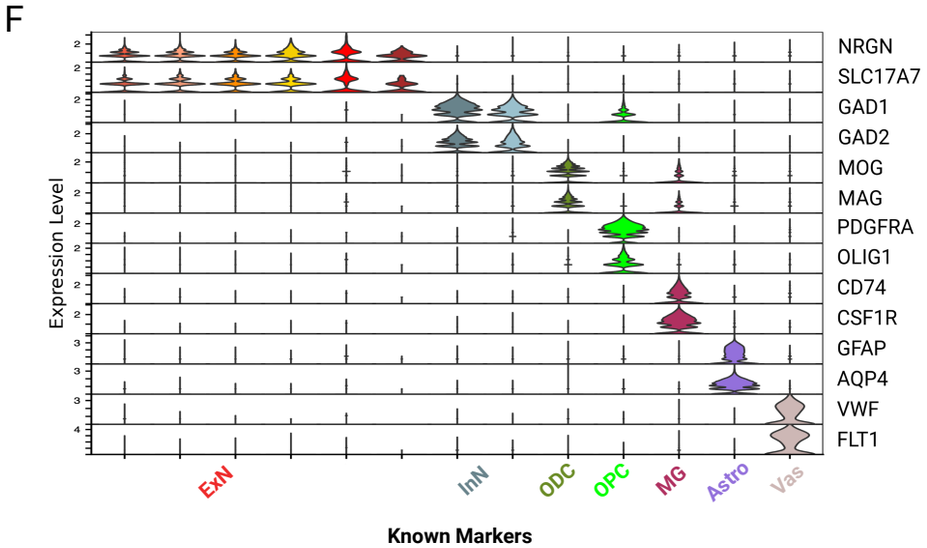

A
n=2
n=2
n=2
LRRK2 PD
(G2019S)
GBA1 PD
(L444P, E326K)
Healthy
control
Braak Lewy-body stage 5-6
All males
PFC
ACC
n=12
Isolated Nuclei
10X Genomics
chromium
Single Nucleus RNA-seq
Data analysis
PD GWAS, Nalls et al, 2019
Integration with genetic
risk data
Isolated RNA
Ribo-depletion
Bulk RNA-seq
Data analysis
Bulk-SC comparison

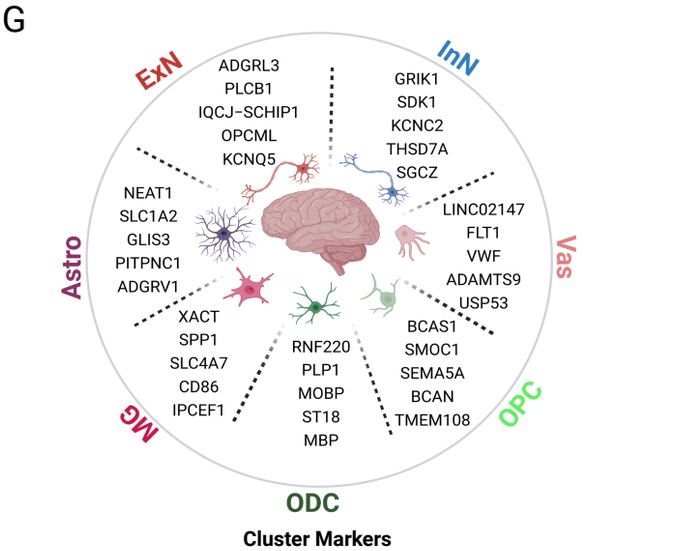

B
ExN3
ExN1
ExN5
ExN2
ExN4
ExN6
Vas
InN1
InN2
Astro
ODC
MG
OPC
88,876 high quality
single nuclei
UMAP2
UMAP1

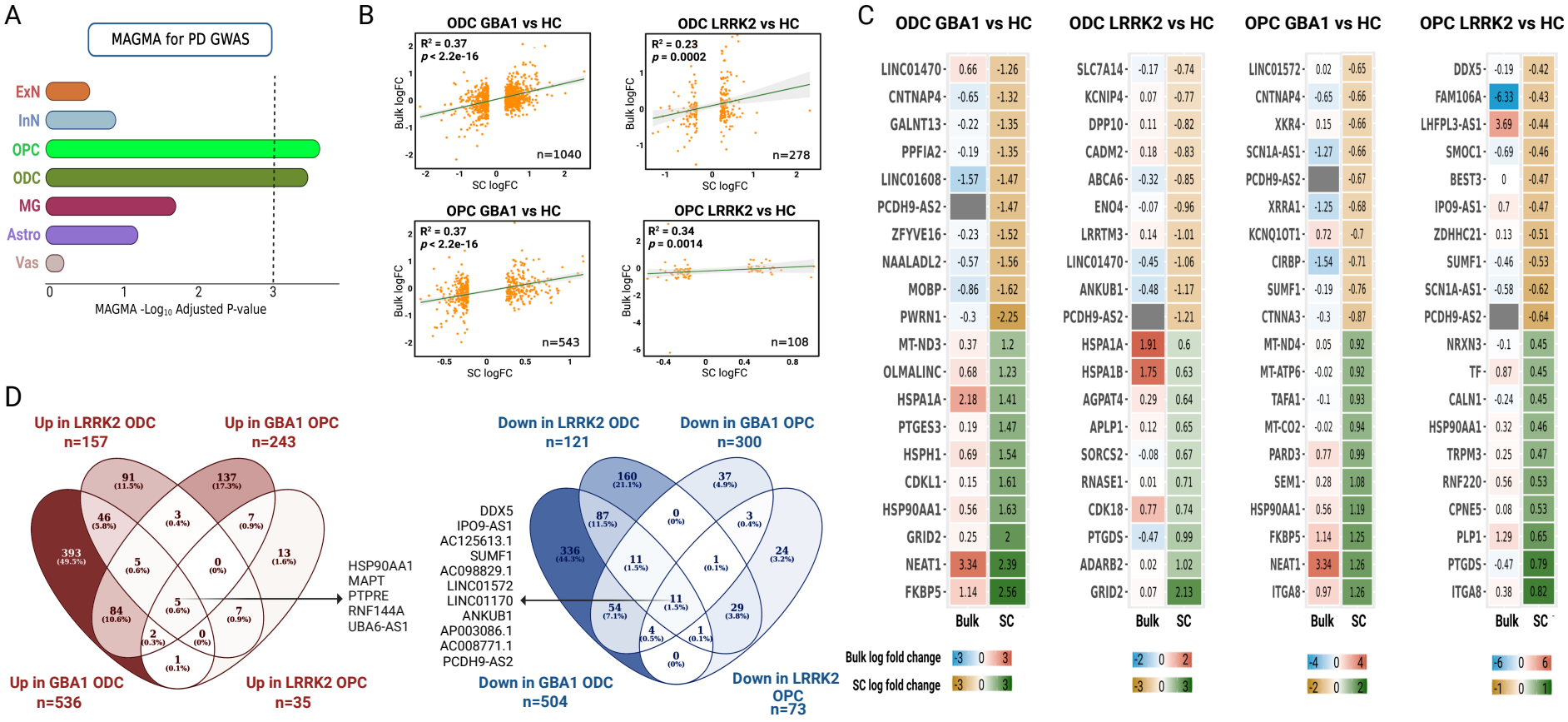

C
ExN1
ExN2
ExN3
ExN4
ExN5
ExN6
InN1
InN2
ODC
OPC
MG
Astro
Vas
0.5%
4.3%
6.3%
10.7%
5.6%
3.4%
2.5%
1.1%
0.3%
5.1%
4.1%
5.2%
50.1%
Sample1 (GBA1-PFC)
Sample2 (GBA1-ACC)
Sample3 (LRRK2-PFC)
Sample4 (LRRK2-ACC)
Sample5 (GBA1-PFC)
Sample6 (GBA1-ACC)
Sample7 (HC-PFC)
Sample8 (HC-ACC)
Sample9 (LRRK2-PFC)
Sample10 (LRRK2-ACC)
Sample11 (HC-PFC)
Sample12 (HC-ACC)
4.5 %
3.7%
8 %
6.2 %
6.1 %
3.6 %
6.4 %
12.3%
6.2 %
14.3 %
16.5 %
12.2 %
GBA1
HC
LRRK2
39%
38%
23%
ACC
PFC
49%
51%


D
GBA1
HC
LRRK2
ACC
PFC


E
GBA1
n=4
LRRK2
n=4
HC
n=4
Cell-type Percentage
100
80
60
40
20
0
Sample 1
Sample 2
Sample 5
Sample 6
Sample 3
Sample 4
Sample 9
Sample 10
Sample 7
Sample 8
Sample 11
Sample 12


F
Expression Level
NRGN
SLC17A7
GAD1
GAD2
MOG
MAG
PDGFRA
OLIG1
CD74
CSF1R
GFAP
AQP4
VWF
FLT1
ExN
InN
ODC
OPC
MG
Astro
Vas
Known Markers


G
ExN
ADGRL3
PLCB1
IQCJ-SCHIP1
OPCML
KCNQ5
InN
GRIK1
SDK1
KCNC2
THSD7A
SGCZ
Astro
NEAT1
SLC1A2
GLIS3
PITPNC1
ADGRV1
Vas
LINC02147
FLT1
VWF
ADAMTS9
USP53
MG
XACT
SPP1
SLC4A7
CD86
IPCEF1
ODC
RNF220
PLP1
MOBP
ST18
MBP
OPC
BCAS1
SMOC1
SEMA5A
BCAN
TMEM108
Cluster Markers


A
MAGMA for PD GWAS
ExN
InN
OPC
ODC
MG
Astro
Vas
0
1
2
3
MAGMA -Log10 Adjusted P-value
B
ODC GBA1 vs HC
R² = 0.37
p < 2.2e-16
n=1040
ODC LRRK2 vs HC
R² = 0.23
p = 0.0002
n=278
OPC GBA1 vs HC
R² = 0.37
p < 2.2e-16
n=543
OPC LRRK2 vs HC
R² = 0.34
p = 0.0014
n=108
Bulk logFC
SC logFC
C
ODC GBA1 vs HC
LINC01470 0.66 -1.26
CNTNAP4 -0.65 -1.32
GALNT13 -0.22 -1.35
PPFIA2 -0.19 -1.35
LINC01608 -1.57 -1.47
PCDH9-AS2 -1.47
ZFYVE16 -0.23 -1.52
NAALADL2 -0.57 -1.56
MOBP -0.86 -1.62
PWRN1 -0.3 -2.25
MT-ND3 0.37 1.2
OLMALINC 0.68 1.23
HSPA1A 2.18 1.41
PTGES3 0.19 1.47
HSPH1 0.69 1.54
CDKL1 0.15 1.61
HSP90AA1 0.56 1.63
GRID2 0.25 2
NEAT1 3.34 2.39
FKBP5 1.14 2.56
ODC LRRK2 vs HC
SLC7A14 -0.17 -0.74
KCNIP4 0.07 -0.77
DPP10 0.11 -0.82
CADM2 0.18 -0.83
ABCA6 -0.32 -0.85
ENO4 -0.07 -0.96
LRRTM3 0.14 -1.01
LINC01470 -0.45 -1.06
ANKUB1 -0.48 -1.17
PCDH9-AS2 -1.21
HSPA1A 1.91 0.6
HSPA1B 1.75 0.63
AGPAT4 0.29 0.64
APLP1 0.12 0.65
SORCS2 -0.08 0.67
RNASE1 0.01 0.71
CDK18 0.77 0.74
PTGDS -0.47 0.99
ADARB2 0.02 1.02
GRID2 0.07 2.13
OPC GBA1 vs HC
LINC01572 0.02 -0.65
CNTNAP4 -0.65 -0.66
XKR4 0.15 -0.66
SCN1A-AS1 -1.27 -0.66
PCDH9-AS2 -0.67
XRRA1 -1.25 -0.68
KCNQ1OT1 0.72 -0.7
CIRBP -1.54 -0.71
SUMF1 -0.19 -0.76
CTNNA3 -0.3 -0.87
MT-ND4 0.05 0.92
MT-ATP6 -0.02 0.92
TAFA1 -0.1 0.93
MT-CO2 -0.02 0.94
PARD3 0.77 0.99
SEM1 0.28 1.08
HSP90AA1 0.56 1.19
FKBP5 1.14 1.25
NEAT1 3.34 1.26
ITGA8 0.97 1.26
OPC LRRK2 vs HC
DDX5 -0.19 -0.42
FAM106A -6.33 -0.43
LHFPL3-AS1 3.69 -0.44
SMOC1 -0.69 -0.46
BEST3 0 -0.47
IPO9-AS1 0.7 -0.47
ZDHHC21 0.13 -0.51
SUMF1 -0.46 -0.53
SCN1A-AS1 -0.58 -0.62
PCDH9-AS2 -0.64
NRXN3 -0.1 0.45
TF 0.87 0.45
CALN1 -0.24 0.45
HSP90AA1 0.32 0.46
TRPM3 0.25 0.47
RNF220 0.56 0.53
CPNE5 0.08 0.53
PLP1 1.29 0.65
PTGDS -0.47 0.79
ITGA8 0.38 0.82
Bulk
SC
Bulk log fold change
SC log fold change
D
Up in LRRK2 ODC
n=157
Up in GBA1 OPC
n=243
Up in GBA1 ODC
n=536
Up in LRRK2 OPC
n=35
HSP90AA1
MAPT
PTPRE
RNF144A
UBA6-AS1
Down in LRRK2 ODC
n=121
Down in GBA1 OPC
n=300
Down in GBA1 ODC
n=504
Down in LRRK2 OPC
n=73
DDX5
IPO9-AS1
AC125613.1
SUMF1
AC098829.1
LINC01572
LINC01170
ANKUB1
AP003086.1
AC008771.1
PCDH9-AS2
E
Up in GBA1 ODC
Up in LRRK2 ODC
Up in GBA1 OPC
Up in LRRK2 OPC
chondroitin sulfate proteoglycan biosynthetic process (GO:0050650)
chondroitin sulfate metabolic process (GO:0030204)
cell morphogenesis involved in neuron differentiation (GO:0048667)
chondroitin sulfate biosynthetic process (GO:0030206)
response to unfolded protein (GO:0006986)
regulation of cellular response to stress (GO:0080135)
regulation of insulin receptor signaling pathway (GO:0046626)
cellular protein-containing complex assembly (GO:0034622)
regulation of cellular response to heat (GO:1900034)
regulation of protein polymerization (GO:0032271)
ERBB2 signaling pathway (GO:0038128)
regulation of cellular protein localization (GO:1903827)
cellular response to heat (GO:0034605)
regulation of protein modification by small protein conjugation or removal (GO:1903320)
central nervous system neuron axonogenesis (GO:0021955)
regulation of microtubule nucleation (GO:0010968)
central nervous system neuron development (GO:0021954)
positive regulation of tau-protein kinase activity (GO:1902949)
negative regulation of inclusion body assembly (GO:0090084)
Combined score (scaled)
Down in GBA1 ODC
Down in LRRK2 ODC
Down in GBA1 OPC
Down in LRRK2 OPC
regulation of mRNA splicing, via spliceosome (GO:0048024)
cholesterol metabolic process (GO:0008203)
regulation of calcium-mediated signaling (GO:0050848)
regulation of stress fiber assembly (GO:0051492)
axonogenesis (GO:0007409)
positive regulation of cell-cell adhesion (GO:0022409)
negative regulation of stress fiber assembly (GO:0051497)
secondary alcohol biosynthetic process (GO:1902653)
cholesterol biosynthetic process (GO:0006695)
sterol biosynthetic process (GO:0016126)
synapse organization (GO:0050808)
regulation of synapse assembly (GO:0051963)
modulation of chemical synaptic transmission (GO:0050804)
regulation of nervous system process (GO:0031644)
modulation of excitatory postsynaptic potential (GO:0098815)
regulation of AMPA receptor activity (GO:2000311)
nucleic acid metabolic process (GO:0090304)
establishment of protein localization to membrane (GO:0090150)
protein localization to plasma membrane (GO:0072659)
positive regulation of epithelial cell migration (GO:0010634)
regulation of potassium ion transport (GO:0043266)
regulation of cation transmembrane transport (GO:1904062)
transmembrane receptor protein tyrosine phosphatase signaling pathway (GO:0007185)
protein processing involved in protein targeting to mitochondrion (GO:0006627)
Combined score (scaled)