A single-nuclei paired multiomic analysis of the human midbrain reveals age- and Parkinson's disease–associated glial changes

Levi Adams[1,2,5], **Min Kyung Song**[1,3,5], **Samantha Yuen**[4], **Yoshiaki Tanaka**[4,✉], **Yoon-Seong Kim**[1,✉]

[1]RWJMS Institute for Neurological Therapeutics, Rutgers-Robert Wood Johnson Medical School, Piscataway, NJ, USA.

[2]Department of Biology, Bates College, Lewiston, ME, USA.

[3]College of Nursing Science, Kyung Hee University, Seoul, Republic of Korea.

[4]Department of Medicine, Maisonneuve-Rosemont Hospital Research Center (CRHMR), University of Montreal, Quebec, QC, Canada.

[5]These authors contributed equally: Levi Adams, Min Kyung Song.

Abstract

Age is the primary risk factor for Parkinson's disease (PD), but how aging changes the expression and regulatory landscape of the brain remains unclear. Here we present a single-nuclei multiomic study profiling shared gene expression and chromatin accessibility of young, aged and PD postmortem midbrain samples. Combined multiomic analysis along a pseudopathogenesis trajectory reveals that all glial cell types are affected by age, but microglia and oligodendrocytes are further altered in PD. We present evidence for a disease-associated oligodendrocyte subtype and identify genes lost over the aging and disease process, including CARNS1, that may predispose healthy cells to develop a disease-associated phenotype. Surprisingly, we found that chromatin accessibility changed little over aging or PD within the same cell types. Peak–gene association patterns, however, are substantially altered during aging and PD, identifying cell-type-

specific chromosomal loci that contain PD-associated single-nucleotide polymorphisms. Our study suggests a previously undescribed role for oligodendrocytes in aging and PD.

Parkinson's disease (PD) is the second most common neurodegenerative disease and is estimated to affect more than 10 million people globally[1]. PD prevalence increases from 41 cases per 100,000 individuals in the fourth decade of life to 425 per 100,000 individuals in the sixth decade of life and to 1,908 per 100,000 individuals in the eighth decade of life[2]. During PD pathogenesis, dopaminergic neurons in the substantia nigra degenerate, resulting in neurological, cognitive and motor symptoms[3]. Although many genetic components and environmental risk factors have been identified, the primary risk factor for PD is aging[4].

Despite the connection between aging and PD, relatively little attention has been given to changes in the aging midbrain and how these changes may predispose individuals to the development of PD. Transcriptomic studies of mouse or fly brains have provided some insight into aging but have not focused on the midbrain[5,6]. Glaab and Schneider[7] used microarray datasets to investigate shared pathways and network alterations in aging and PD, but there remains a lack of information on how alterations in gene expression during aging affect the different cell types in the midbrain and how they contribute to PD pathogenesis.

The field of single-cell transcriptomics and epigenomics has facilitated important advances in understanding how gene expression and chromatin accessibility contribute to neurodegeneration[8–10]. Recent single-cell expression studies of the midbrain have highlighted a potential role of oligodendrocytes in the midbrain in PD pathogenesis, supported by strong genome-wide association studies (GWASs) and transcriptomic data[11–14]. This is a surprising finding because vulnerable dopaminergic neurons that are lost during PD pathogenesis are sparsely myelinated[15]. Despite these crucial insights, there remains a lack of cell-type-specific high-resolution data on what distinguishes 'healthy' aging from neurodegeneration and an incomplete picture of how the epigenetic landscape changes during aging and PD.

To address this gap, we isolated nuclei from the substantia nigra of postmortem midbrains of young and aged donors with no neurological disease as well as patients with PD and compared the single-nuclei transcriptome and genome-wide chromatin accessibility simultaneously from each nucleus along aging and PD trajectories. This approach allows us to directly infer *cis*-acting elements that contribute to gene expression in the same cells and provides a more complete picture of how the aging process affects gene regulation and expression in distinct cell types in the midbrain. This multiomic analysis reveals a disease-associated oligodendrocyte subset that may contribute to PD pathogenesis.

Results

Combined single-nuclei RNA sequencing and assay for transposase accessible chromatin sequencing of the midbrain

We obtained frozen, postmortem midbrain samples of young (mean, 24 years old) and aged (mean, 75 years old) neurologically healthy donors and patients diagnosed with PD (mean, 81 years old) (Supplementary Table 1). We isolated nuclei from the substantia nigra and

performed paired single-nuclei RNA sequencing (snRNA-seq) and single-nuclei assay for transposase accessible chromatin sequencing (snATAC-seq) on each nucleus (10x Genomics Single Cell Multiome ATAC + Gene Expression Kit) (Fig. 1a and Supplementary Fig. 1a–c). After filtering to remove low-quality reads or potential multiplets (Methods), we retained 69,289 high-quality nuclei from 31 individuals (nine young donors, eight aged donors and 14 patients with PD), showing low rates of doublets based on gene expression and robust ATAC data quality (Supplementary Fig. 1d–g). Using these nuclei, we performed batch correction, variable gene and principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) dimensional reduction with Seurat version 4 (ref. 16) (Methods). We identified 23 separate clusters of nuclei in our snATAC-seq and snRNA-seq datasets (Extended Data Fig. 1a). Interrogation of gene expression patterns of known cell type markers[10,17,18] allowed us to classify our nuclei into seven major cell types (Fig. 1b,d and Extended Data Fig. 1b): neurons (Ns), oligodendrocytes (ODCs), astrocytes (ASs), microglia (MG), oligodendrocyte precursor cells (OPCs), endothelial cells (ECs) and peripheral immune cells/T cells (Ts). As expected, clustering based on ATAC profiling generated the same major clustering patterns for these cells with distinctive chromatin accessibility profiles at key loci (Fig. 1c,e and Extended Data Fig. 1b). We found that the vast majority of cells in the midbrain are ODCs, followed by MG, OPCs, ASs and Ns (Fig. 1f and Supplementary Fig. 1h). Similar to recently published data[14], we did note a statistical difference in MG and ODC populations between groups ($P < 0.01$ for ODCs and MG between young and aged and $P < 0.05$ for ODCs and MG between aged and PD; Supplementary Table 2).

Genes with cell-type-specific expression patterns also showed differential ATAC peaks nearest to the transcription start site (TSS), which implies unique promoter accessibility for each cell type (Fig. 1e) and distinctive enrichment of cell-type-specific transcription factor binding motifs (Extended Data Fig. 1c). Within each cell type, we analyzed differentially expressed genes (DEGs) between young and aged groups as well as between aged and PD groups (Supplementary Table 3). Several genes that have been previously associated with neurological disorders, such as *NEAT1*, *FKBP5* and *SLC38A2*, were differentially expressed in multiple cell types (Supplementary Fig. 2a,b and Methods). We identified cell-type-specific changes in gene expression across groups (Supplementary Fig. 2c,d,f) that showed potential alterations in neurological function and metabolic pathways (Supplementary Fig. 2e,g).

Multiomic analysis of peak–gene associations

Despite distinct ATAC profiles among cell types (Fig. 1e), we found surprisingly few differences among groups (young, aged and PD) within a given cell type (Supplementary Fig. 3a–c). Our finding is similar to previously published snATAC-seq data comparing healthy and AD brain samples (Supplementary Fig. 3d,e)[19]. Gene expression patterns of cell-type-specific genes were highly consistent with the chromatin accessibility of their promoters (Supplementary Fig. 4a). However, differential gene expression within cell types during aging or PD pathogenesis was not correlated with the chromatin accessibility of the promoters (Supplementary Fig 4b–e). These results suggest that changes over aging or PD are not strongly correlated with promoter activation but may rely, instead, on more nuanced

interactions between distal DNA elements. The paired single-nuclei gene expression and ATAC-seq data from the same nuclei provide a unique advantage because it allows us to infer the relationship between chromatin accessibility and gene expression in *cis*. We used the analytical framework developed by Ma et al.[8] to analyze paired snRNA-seq and snATAC-seq data to generate peak–gene associations (Fig. 2a). Correlated associations were calculated in the chromatin regions within ±500 kilobases (kb) of the TSS of annotated genes where there is covariation between chromatin accessibility and gene expression. In each cell type, an average of 193,732 significant peak–gene associations (5.3 peaks per gene) were identified. Most shared peak–gene associations (50.9–56.6%) were near the TSS (within 5 kb), and unique peaks were linked in more distal regions, suggesting that expression changes between groups are likely driven by more distal regions, such as enhancers (Supplementary Fig. 5a,b). In agreement with previous data[8], we also observed that not all peaks were connected to a gene (91,177 out of 210,609 peaks were connected to at least one gene) and saw an average of 6,812 genes connecting to more than 10 peaks (Supplementary Fig. 5c,d). Comparing peak–gene associations in ODCs between groups identified a subset of genes whose correlated associations were significantly changed during aging and PD progression (Fig. 2b). Increased peak–gene associations in both aged and PD groups compared to young (Fig. 2b, red dots) may indicate aging-specific changes, including the aging marker *NEAT1*. Conversely, some associations were unchanged between young and aged but decreased only in the PD cohort, including genes previously linked with longevity, such as *RASGRF1* (ref. 20) and *MSRA*[21] (Fig. 2b, blue dots). Peak–gene associations for *NEAT1* were largely shared in older (aged and PD) groups compared to young samples (Fig. 2c and Extended Data Fig. 2a), whereas, for *RASGRF1*, similar patterns were observed in normal control donors (young and aged) with much higher associations compared to PD (Fig. 2d and Extended Data Fig. 2b). Analysis of *cis*-regulatory motifs of the 43 correlated peaks for *NEAT1* specific to the older groups (aged and PD) showed enrichment of binding motifs for transcription factors related to aging, such as *EGR1*/2 (ref. 22) (Extended Data Fig. 2a,c). For *RASGRF1*, we found 61 significant shared peaks in the control groups (young and aged) and enriched motifs for *NRF2* and *ASCL1*, which are related to brain health and neurogenesis[23,24]. These motifs were not used in PD samples (Extended Data Fig. 2b,d). We also compared our peak–gene association data with previously published H3K27ac HiChIP data prepared from the human midbrain to validate the biological relevance of our findings[25] and found that our peak–gene associations of ODCs significantly overlapped with reported HiChIP loops ($P < 1.0 \times 10^{-334}$ by hypergeometric test; Extended Data Fig. 2e).

Next, we applied the ATAC peak patterns and peak–gene associations to previously published PD GWAS databases to identify cell-type-specific single-nucleotide polymorphisms (SNPs) that are related to PD. We performed linkage disequilibrium (LD) score regression analysis of GWAS SNPs to cell-type-specific ATAC peaks as described in Corces et al.[25,26]. Similarly, our midbrain ATAC peaks that were exclusive to one cell type were significantly associated with SNPs related to Alzheimer's disease (AD) in MG but displayed no significant enrichment of SNP heritability of PD (Fig. 3a,b). However, we noted that several peaks containing PD-associated SNPs were detected in more than one type of cell (Fig. 3c), so we expanded our focus to include these peaks that were shared. Inclusion

of the highly significant shared peaks that were observed in more than 10% of cells of each cell type and found that these ATAC peaks are significantly associated with SNPs related to neurodegenerative diseases (PD and AD) but less enriched with other neuronal and non-neuronal diseases (Fig. 3d). Around 8–16% of known PD-associated SNPs were located on the midbrain ATAC peaks with cell-type-specific variations (Fig. 3c and Extended Data Fig. 3c). For example, the peak on chromosome 1 containing the PD-associated gene *PINK1* has seven PD SNPs and is significantly enriched only in Ns but not in other glial cells (Fig. 3d and Extended Data Fig. 3c). The region downstream of the *SNCA* gene containing *rs356174* is accessible in ODCs and OPCs but not in Ns, ASs or MG (Fig. 3e). Similarly, we found SNPs enriched only in ASs (Fig. 3f) and MG (Fig. 3g). Although this combined analysis of snATAC-seq data with PD GWAS SNPs could provide a cell-type-specific regulatory landscape governed by a particular SNP, it is still difficult to infer which genes are regulated by each SNP. With our paired dataset, we also were able to associate gene expression with specific SNP-containing regions (Fig. 3h). For example, the *MAPT* locus on chromosome 17 has nine distinct accessible peaks containing 23 individual SNPs (Fig. 3i), and they display cell-type-specific peak–gene association patterns (Fig. 3h), implying that the same SNP could be associated with different gene expression patterns in a cell-type-specific manner. Interestingly, we also found evidence of a differential peak–gene association among young, aged and PD donors. The *BST1* locus containing three PD-associated SNPs was accessible in ASs (Fig. 3f), but comparing peak–gene associations between cohorts shows that this peak is not associated with gene expression in young donors. However, although it is significantly associated with the expression of *CD38* and *CC2D2A* in PD, it is associated only with *CC2D2A* in aged control (Extended Data Fig. 3d).

Defining pseudopathogenesis trajectory in ODC

The initial clustering analysis yielded several subclusters for each cell type (Extended Data Fig. 1a). We examined if they could be further defined based on the gene expression signatures of previously reported functional subtypes[27–33] (Methods). For ODCs, we identified clusters corresponding to newly formed, myelin-forming and mature ODCs, which show changes in expression of representative genes (Supplementary Fig. 6a,d). We also observed clusters with high expression of genes relating to neuronal and synaptic support (such as *NRXN3* and *NFASC*) and clusters with increased expression of ODC–N adhesion markers (such as *STMN1* and *HAPLN2*) (Extended Data Fig. 4a and Supplementary Fig. 6a,d). As previously reported, these subtypes have relatively distinct transcriptional programs and do not fully overlap[29,32]. MG signatures included homeostatic MG, aging MG and a small population of stage1 (*TREM2*-independent) disease-associated MG, as shown by expression of representative marker genes, such as *P2RY12*, *APOE* and *DOCK5* (Extended Data Fig. 4b and Supplementary Fig. 6b,e)[27,30]. ASs exhibited gene expression signatures for disease-associated ASs, reactive ASs and glial fibrillary acidic protein (GFAP)-low ASs, as represented by expression of genes such as *GFAP*, *FKBP5* and *LUZP2* (Extended Data Fig. 4c and Supplementary Fig. 6c,f)[28,31,34]. This cell-type-specific functional clustering led us to explore potential changes in gene expression and epigenetic dynamics during the aging and disease process and analyze each major cell type individually. To highlight heterogeneity and subtle changes within each cell type, we reclustered each type of cell individually (Methods), starting with ODCs (Fig. 4a and

Supplementary Fig. 7a). To identify potential transitions in gene expression and chromatin accessibility between clusters, we performed pseudotime analysis using Monocle3 (ref. 35). Pseudotime analysis generated a striking trajectory in both snRNA and snATAC clusters moving from young to aged and then to PD (Fig. 4a). Taking advantage of the paired multiomic nature of our data, we combined the snRNA and snATAC pseudotime and corrected for error to establish a single combined score based on both expression and epigenetic datasets that we termed the combined pseudopathogenesis (cPP) trajectory (Fig. 4b). Comparison of the cPP scores between groups showed significant increases from young to aged and from aged to PD (Fig. 4c). We found 299 genes that increased along this trajectory and 474 that decreased (Fig. 4d, Extended Data Fig. 4d and Supplementary Table 4). Differentially expressed ATAC peaks were also noticed along the cPP (Supplementary Fig. 7b). Notably, aging or PD-relevant genes that we identified in peak–gene association analysis, such as *NEAT1* and *RASGRF1* (Fig. 2c,d), were also correlated with cPP (Extended Data Fig. 4e). Gene Ontology (GO) analysis showed that pathways related to Response to Unfolded Protein, Chaperone-Mediated Autophagy (CMA) and Negative Regulation of Cell Death were increased, whereas Myelination, Receptor Clustering and Regulation of Membrane Potential were decreased with advancing pseudopathogenesis (Fig. 4d). Of note, we observed an increase in *RPS* (ribosomal protein small) family genes that were previously reported to be increased during aging[6] (Supplementary Table 4). Manually curating genes correlated with cPP score into modules based on their annotated function (Methods) showed that genes related to normal ODC functions, such as myelination and synapse formation, decrease at high cPP, whereas modules related to the unfolded protein response and chaperone-mediated autophagy increase (Fig. 4e). We also noticed a distinct polarization in the expression of several genes, including *OPALIN* and *RBFOX1*, as previously reported[14] (Extended Data Fig. 4f). Notably, both *RBFOX1* and *OPALIN* expression levels peaked at lower cPP scores and then were reduced in high cPP and were lower in PD compared to young donors (Extended Data Fig. 4g,h).

Pseudopathogenesis analysis shows distinct microglial changes

We applied the same analytical framework to MG and generated a cPP score for each nucleus (Fig. 5a,b). We found significant increases in cPP among young, aged and PD samples for MG (Fig. 5c). GO analysis of the 894 increased and 254 decreased genes showed a loss of cell adhesion and chemotaxis and elevated immune activation and cytokine-mediated signaling pathway (Supplementary Table 4 and Supplementary Fig. 8a). Combined gene module analysis showed a decrease in homeostatic gene signatures and increases in aging and stage1 (TREM2-independent) disease-associated MG (Fig. 5d). We analyzed AS clusters by applying the same strategy and found a statistically significant increase in cPP score from young to aged samples but no difference between aged and PD nuclei (Extended Data Fig. 5a–c). GO analysis of trajectory-relevant gene expression showed generalized increases in apoptosis resistance and CMA pathways but a reduction in neuronal support (Extended Data Fig. 5d). Gene module analysis indicated increases in reactive AS signatures and disease-associated AS signatures near the middle of the trajectory, with a concomitant reduction of GFAP-low module signatures across cPP (Extended Data Fig. 5e). OPCs exhibited similar results to ASs (Supplementary Fig. 8b–d)

with a significant increase in cPP scores from young to aged nuclei but no difference in PD samples.

In-depth analysis of high cPP ODCs

The pseudopathogenesis trajectory provided a strong tool to analyze the incremental changes that occur with aging and PD pathology. Intriguingly, we noticed that, whereas nearly all cells from young donors had low cPP scores, there was substantial heterogeneity in aged and PD samples (Extended Data Fig. 6a). Plotting cell populations as a function of cPP score revealed three distinct peaks of cells (Fig. 6a), and we identified three populations based on kernel density estimation[36]. As low-cPP nuclei encapsulated the majority of normal cellular function for ODCs, such as myelination and neuronal support (Fig. 4d,e), we defined this population as 'healthy'. The high-cPP population was marked by a loss of canonical ODC functions and increased stress response genes (Fig. 4e), suggesting that they are 'disease-associated' ODCs, with a population of nuclei with an 'intermediate' pattern of gene activation. Healthy cells made up the vast majority of nuclei in our study (76%), and, surprisingly, 89% of nuclei from aged control donors and 48% of nuclei from patients with PD were in the 'healthy' population (Fig. 6a and Extended Data Fig. 6b), suggesting that a significant proportion of cells maintain their normal functions even in the disease state. We analyzed DEGs among healthy, intermediate and disease-associated populations (Supplementary Table 5) and noticed interesting trends when compared to DEGs among young, aged and PD cohorts. Some genes were not significantly changed over aging but were changed between healthy and intermediate or disease-associated ODCs, suggesting PD-specific changes that are unrelated to the aging process (Fig. 6b, left). The list includes reductions in multiple genes involved in the myelination pathway (*MBP* and *MOBP*) and several glial–neuron adhesion genes (*CTNNA3* and *NRXN3*) as well as increases in genes not widely explored in PD, such as *SELENOP*, *QDPR*, *SLC38A2* and *IGF1R*. We also identified a subset of genes that was differentially expressed over aging and further changed toward a disease state (Fig. 6b, right). These genes are of particular interest as they potentially represent age-related risk factors for the development of PD. They include increases in stress response genes (*HSP90AA1* and *FKBP5*) and well-known neurodegenerative markers, such as *MAPT*. We also observed an age-dependent loss of several genes that are further reduced in disease (*CARNS1* and *NKAIN2*). Of particular note, *CARNS1*, which encodes carnosine synthase 1, has been previously suggested to have a protective role in neurodegenerative conditions[37,38]. Next, we applied our cPP pipeline to analyze the age-matched control and PD midbrain dataset published by Smaji et al.[14] to see if we could find the same disease-associated population in another single-cell study. Similarly, with our datasets, kernel estimation separates three distinct groups with transcriptional signatures similar to our healthy and disease-associated populations (Extended Data Fig. 6c,d). This supports our results that identify differentially affected subpopulations of cells in PD.

We performed peak–gene analysis of healthy, intermediate and disease-associated ODCs and identified differential peak–gene associations among these groups. Interestingly, analyzing ODC-specific peaks containing PD-associated SNPs showed no appreciable peak enrichment among groups but a significant change in peak–gene association (Fig. 6c). For

example, the chromatin peak containing *rs11248060* and *rs11724804* was associated with the expression of *FGFRL1* and *PDE6B* in healthy cells, whereas it was associated with an antisense transcript of undescribed function (AC139887.4) in disease-associated ODCs. The peak containing *rs557074* showed no significant peak–gene correlation in healthy ODCs but was associated with gene expression in disease-associated ODCs (Extended Data Fig. 6e,f). As previously discussed, the *MAPT* locus contains multiple PD-associated SNPs. Five of these peaks, containing 17 individual PD-associated SNPs, have no association with *MAPT* expression in healthy cells, but all five peaks were associated with expression in disease-associated ODCs (Fig. 6d)

To confirm our bioinformatics findings, we performed RNA-FISH on formalin-fixed, paraffin-embedded (FFPE) human substantia nigra sections and performed a quantitative analysis (Supplementary Fig. 9). As expected, we observed incremental reductions in the expression of *MBP* and *CARNS1* over aging and PD (Fig. 7a,b). *RBFOX1* expression was decreased in aged controls compared to young donors (Fig. 7c,d). We also confirmed a significant decrease in *PDE1A* expression (Fig. 7e,f), whereas *SELENOP* and *QDPR* were significantly elevated in PD compared to age-matched controls (Fig. 7g–j). Of note, we saw wide variation in the levels of these genes, even in cells from the same donor, supporting the idea that subsets of cells are differentially affected during the disease process.

Discussion

In this study, we assessed paired snRNA-seq and snATAC-seq patterns in the substantia nigra region of the midbrain of young, aged controls and patients with PD to elucidate how aging processes may predispose an individual to PD. By adding data from young control donors to the current single-cell studies of the brain, our study may considerably extend understanding of the aging process and neurodegeneration at single-cell resolution. Leveraging paired analysis of RNA expression and chromatin accessibility from the same cells provides a unique opportunity to identify cell-type-specific regulation of genes and their contribution to disease processes. We found that there are marked and widespread changes in links between expressed genes and discrete chromatin regions during aging and PD. Trajectory analysis revealed that all cells change over the course of aging, and there are significant further alterations in ODC and MG populations in PD, further supporting the recently highlighted roles of these cells in PD development[11–14]. We identified a population of disease-associated ODCs and found that, even in PD, a large percentage of cells maintain a normal, healthy transcriptional profile. We highlighted genes that were specific for PD but were unchanged over aging and also identified genes that were incrementally affected by aging and PD. These aging-related and disease-associated genes may represent a source of risk that is inherent in the aging process.

Each brain region has a unique cell type composition. Our data agree with previous data[11] indicating that ODC is the major cell type in the midbrain, representing approximately 75% of all cells, which has led to increasing recognition of ODC as a potential major player in PD[11–14]. Although this finding is interesting, it should be noted that these single-cell isolation methods may lead to cell selection bias, and this warrants further research into the cellular composition of the midbrain with methods that do not rely on dissociation.

After ODCs, MG, OPCs and ASs are the next most common cell types; however, we failed to faithfully recover neuronal populations that were reported in previous studies in the midbrain[14], potentially due to their increased sensitivity to nuclear permeabilization for ATAC-seq preparation.

We adapted the peak–gene association analysis developed by Ma et al.[8] to link distal peaks to genes in *cis* based on covariation in chromatin accessibility and gene expression, which is a crucial step forward in unraveling the complex cellular heterogeneity of the brain. The recent work by Morabito et al.[19] elegantly combined separate ATAC and RNA expression data to form correlative relationships between two similar nuclei. Peak–gene association studies further advance this technique by analyzing paired data from the same nuclei. The importance of this approach is highlighted by our unexpected findings showing surprisingly little variation in ATAC peaks during aging or PD regardless of changes in gene expression. Despite this, peak–gene analysis revealed widespread alterations in gene expression–peak relationships during aging and PD pathogenesis. For example, we showed that the peak–gene association around the *MAPT* locus varies widely between cell types even though many similar peaks suggest that this region may be highly active. Overlaying known PD GWAS data shows that many nearby ATAC peaks containing PD SNPs are differentially associated with gene expression in a cell-type-specific manner. As the peak–gene association is a correlative measure, it remains to be seen how these chromatin regions may interact with the DNA and which links are upstream or downstream of gene activation. A more in-depth study of these *cis*-regulatory mechanisms underlying functional alterations in cellular subpopulations is necessary to understand the enormously complex changes that occur during aging and neurodegeneration.

Adapting pseudotime to infer pathogenic changes is a critical step toward expanding understanding of how aging may alter the midbrain and how those alterations predispose an individual to PD. Our inclusion of young donors with no neurological disease allows us to define a baseline, or 'healthy', transcriptional state. Multiomic analysis indicates that ODCs and MG in the midbrain undergo significant population shifts during aging into PD and show an increase in markers previously associated with aging and neurodegenerative disease. We did not observe a strong correlation of midbrain ASs or OPCs with PD development, although this does not preclude them from having important roles in other brain regions in PD.

Pseudopathogenesis analysis revealed that most ODCs in the midbrain maintain their normal functions during aging and PD. However, a small subset of disease-associated ODCs with distinct transcriptional signatures appears during the disease process, suggesting that a small number of highly dysregulated cells may drive the major disease process. There were sets of genes selectively changed in PD that were independent of aging, such as *QDPR* and *SELENOP*, but, more importantly, we identified genes that changed over aging that were further altered during PD. As the primary risk factor for PD is aging, these genes may help us understand the difference between so-called 'healthy' aging and aging-dependent risk factors for neurological diseases. Of particular note is *CARNS1*, which encodes carnosine synthase 1. Carnosine (β-alanine-l-histidine) is an endogenous antioxidant and neuromodulator and has been shown to confer protective effects in various neurological

conditions, including AD and PD[37,38]. We observed that both expression levels and the number of ODCs expressing *CARNS1* were incrementally decreased over aging and PD. It will be interesting to investigate if the loss of *CARNS1* expression in the subpopulation of ODCs leads to PD progression. This opens the discussion that ODCs might play a key role outside of their canonical myelination function to maintain neuronal health and that loss of this support may tip the balance of the highly vulnerable dopaminergic neurons into a degenerative state.

The associations of each brain cell type with the genetic risk of neurodegenerative diseases have been so far discussed in single-cell RNA-seq[11,12] and single-cell ATAC-seq of human brains[25]. Both studies reported that the genetic risks of AD are significantly associated with MG. In contrast, significant associations with PD were demonstrated in Ns and ODCs by single-cell RNA-seq, whereas no significant associations with PD were observed in any cell types by scATAC-seq. Similar results with this observation were also reproducible by our snATAC profiles of the human midbrain. However, we also found that the enrichment of the genetic disease risks in each cell type varies by the criteria of ATAC peaks. More inclusion of the shared ATAC peaks across cell types increased SNP heritability for AD and PD. It is imperative to investigate how these SNPs in the shared ATAC peaks play causative roles in disease pathology as a prospective study.

In future studies, a more nuanced gradient of aging samples may be required to understand the continuum of cell states and how populations shift during aging. It will also be interesting to see what findings of similar studies in different brain regions will uncover, as this is essential to building a brain-wide atlas of human aging and neurodegeneration.

Methods

Human samples

Human midbrain samples were obtained through National Institutes of Health (NIH) NeuroBioBank requests (Human Brain and Spinal Fluid Resource Center, University of Maryland Brain and Tissue Bank, Harvard Brain Tissue Resource Center and University of Miami Brain Endowment Bank). Written informed consent was obtained for the use of human postmortem tissue samples. This project used de-identified postmortem brain samples, so it is not considered to meet federal definitions for institutional review board (IRB) jurisdiction and falls outside the purview of the Rutgers IRB committee. All age, sex and pathology information are provided in Supplementary Table 1.

Nuclei isolation and sequencing

To reduce batch effects, we processed each batch containing at least one tissue block from each donor type. Two batches of four tissue blocks were processed each day until all donor samples had been processed. To capture the substantia nigra region of the midbrain precisely, we dissected approximately $3 \times 3 \times 5$-mm sections using postmortem photos provided (Supplementary Fig. 1a), yielding tissue blocks ranging from 50 mg to 75 mg. We processed the tissue in a dounce homogenizer and then centrifuged it through 1.5 M sucrose to isolate the nuclei. With this method, we were able to reliably isolate more than $1 \times$

10^6 intact nuclei from approximately 50 mg of postmortem midbrain tissue. Centrifugation through an iodixanol gradient as previously described[9] produced high-quality nuclei with low mitochondrial contamination, but the actual yield of nuclei was very low ($2.5–5 \times 10^4$ nuclei from approximately 50 mg of postmortem midbrain tissue), so we did not continue with this method, as the substantia nigra is a small region, and tissue amount is limited. We stained the nuclei with 7AAD and sorted to purify nuclei from mitochondria. After sorting, we permeabilized according to current 10x Genomics protocols (protocol CG000375, Rev A: 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween 20, 0.01% digitonin, 1% BSA, 1 mM DTT, 1 U ul^{-1} RNase inhibitor) for 2 min on ice. Nuclei were washed once before counting and proceeding with 10x Genomics protocols for transposition, nuclei isolation and barcoding and library preparation exactly as written (10x Genomics protocol Chromium Next GEM Single Cell Multiome ATAC + Gene Expression CG000338, Rev A). Libraries were sequenced at GENEWIZ using Illumina NovaSeq S4 flowcells. Our actual average sequencing depth was 1.8×10^8 reads per sample for gene expression libraries and 1.75×10^8 reads per sample for ATAC libraries. After sequencing, three samples (one each of young, aged and PD) showed strong ATAC results but poor RNA read quality, indicative of RNA degradation in the original sample. The nuclei from these three donors were excluded from downstream analysis.

Authors' note on nuclear isolation

This isolation method was sufficient to obtain high-quality nuclei with minimal blebbing and worked well in our initial pilot study. However, after in-depth data analysis, we noted that barcodes for nuclei with neuronal expression patterns consistently showed unusually high RNA read count, suggesting that these nuclei were clumping together and forming multiplets, and, thus, these neurons were excluded by our quality control filtering (see below). In nuclear isolations, this is commonly caused by a partial loss of membrane integrity that allows genomic DNA to leak out and cause nuclei to stick together. We think it is likely that the digitonin permeabilization step may have been too harsh for these samples and that the larger neuronal nuclei were more sensitive to digitonin. Previous publications also noted the effect of differential responses in cell types to different isolation protocols[6]. In the months since we performed the initial isolation, multiple alternative protocols for the isolation of frozen postmortem tissue have been posted and shared online in unofficial forums that omit digitonin for this reason. In future studies, we will also adjust our isolation protocols accordingly.

Quality control of snRNA + snATAC multiome

Both snRNA-seq and snATAC-seq reads were spontaneously mapped to GRCh38 human reference transcriptome and genome, respectively, by Cell Ranger ARC (version 1.0.1) with '–mempercore = 8, –localcores = 12' (https://support.10xgenomics.com/single-cell-multiome-atac-gex/software/downloads/latest). Quality control of single-nuclei transcriptome and chromatin profiles was conducted by Seurat (version 4.0.0) and Signac (version 1.1.1) in R package version 4.0.2, respectively[16,39]. In brief, we removed cells with (1) fewer than 100 or more than 7,000 detected genes, (2) fewer than 500 or more than 20,000 reads and (3) more than 5% mitochondria-derived reads (damaged/dead cells or doublet) in the RNA datasets. We also evaluated the doublet frequency by counting cells

expressing both *STMN2* and *AQP4* (Supplementary Fig. 1d), which are usually exclusively expressed in cortical neuron and astrocyte, respectively[40]. If the ratio of $STMN2^+AQP4^+$ cells to total number of $STMN2^+AQP4^-$, $STMN2^-AQP4^+$ and $STMN2^+AQP4^+$ cells was more than 3%, the libraries were excluded from subsequent analyses. For quality control of ATAC datasets, we calculated four measurements: (1) ratio of mononucleosomal fragments to nucleosome-free fragments, (2) enrichment score in TSS, (3) total number of ATAC-seq reads per cell and (4) ratio of reads mapped to 'blacklist regions' that are genomic regions inducing aberrant mapping and artificial signals[41]. We filtered out cells with (1) more than 4 mononucleosome:nucleosome-free ratio, (2) less than 2 TSS enrichment score, (3) fewer than 1,000 and more than 60,000 ATAC reads, (4) more than 2% of reads in the blacklist regions and (5) fewer than 1,000 or more than 25,000 ATAC peaks. Our initial isolation yielded 82,735 nuclei. After quality control, we retained 69,289 high-quality nuclei that met the above criteria of RNA and ATAC for subsequent analyses (Supplementary Fig. 1d–g).

Integration of single-cell gene expression profiles

Gene expression profiles were harmonized using Seurat (version 4.0.0) as we have done previously[17,18,40,42]. In each snRNA-seq library, raw unique molecular identifier (UMI) count was normalized to total UMI count. Highly variable genes were then identified by variance-stabilizing transformation with 0.3 LOESS span and automatic setting of clip.max value. The top 2,000 variable genes were used to identify cell pairs anchoring different snRNA-seq libraries using 20 dimensions of canonical correlation analysis (CCA). After scaling gene expression values across all integrated cells, we performed dimensional reduction using PCA. For the visualization, we further projected single cells into two-dimensional UMAP space from 1st and 20th principal components (PCs). Graph-based clustering was then implemented with the shared nearest neighbor method from 1st and 20th PCs and 0.8 resolution value. DEGs in each cluster were identified with more than 1.25-fold change and $P < 0.05$ by two-sided unpaired *t*-test. Note that due to the inherent imbalance in the sex of donor samples, differential gene analysis between donor cohorts includes sex-specific genes, such as *XIST* and *UTY*, which are omitted from downstream analysis. GO analysis was performed on DEGs by the GOstats Bioconductor package (version 2.56.0)[43]. False discovery rate (FDR) was adjusted by the p.adjust function in R with the 'method = 'BH'' option.

Cluster annotation

Cell types were assigned in each cluster 'island' with uniquely expressed genes[18,42]. First, we assigned neuronal clusters with expression of *STMN2* and *TBR1*. ODC and AS clusters were classified by myelination markers (*MBP* and *MOG*) and AS-specific proteins (*GFAP* and *AQP4*) expression, respectively. $OLIG1^+OLIG2^+$ clusters without any neuronal and AS markers were defined as OPCs. ECs were annotated by substantial expression of *FLT1*, *VWF* and *PDGFRB*. We annotated clusters with *GPR34*, *TREM2* and *C1QC* without any OPC and AS markers. Because T cell infiltration into the substantia nigra is a hallmark of PD pathology, we also defined T cell clusters with their specific marker expression (*CD3E* and *CD8A*).

Integration of single-cell chromatin profiles

Integrative analysis of single-cell chromatin profiles was performed by Seurat (version 4.0.0), Signac (version 1.1.1), GenomicRanges (version 1.42.0) and Harmony (version 1.0) R packages[39,44–46]. First, ATAC peaks from all ATAC libraries were merged by the 'reduce' function in GenomicRanges. Huge (>10,000 base pairs (bp)) and tiny (<20 bp) combined peaks were removed from subsequent analyses. ATAC reads were recounted in the combined peaks by using the FeatureMatrix function in Signac. After merging the peak × cell matrices from all ATAC libraries, the ATAC read counts were normalized by term frequency–inverse document frequency. Latent semantic indexing (LSI) was then computed from the merged count matrix by the singular value decomposition method. To minimize the batch and technical difference across libraries, we ran the Harmony algorithm using LSI embeddings with the 'project.dim = F' option. For data visualization, all cells were embedded into two-dimensional UMAP space from the 2nd and 30th Harmony-corrected LSI. The 1st index of LSI may represent technical variations, such as sequencing depth, and was not used for UMAP projection.

Analysis of differential gene expression and chromatin-accessible regions

Gene expression and chromatin accessibility profiles were compared between PD and healthy aged groups, between healthy aged and young groups or across cell types. The DEGs and differential open chromatin regions were identified in each cluster by 1.25-fold change and $P < 0.05$ unpaired t-test. Cellular events and functions related to the DEGs and differential open chromatin regions were analyzed by the enrichment of GO using GOstats (version 2.56.0) as described above[43]. Genomic distribution of all ATAC peaks and differential open chromatin regions was identified by the annnotatePeak.pl script in the HOMER suite (version 4.11.1) with default parameters[47]. Analysis of previous association of DEGs was done manually[6,48–71].

Analysis of the relationship between gene expression and ATAC peaks

To investigate the relationship between regulatory sites and differential gene expression, the Spearman correlation was calculated using the normalized gene expression values and ATAC read counts in each peak–gene pair[8]. We chose all ATAC peaks within ±500 kbp of TSSs of genes for the correlation analysis. In each peak, we generated 100 permutated background peaks with the same accessibility. Spearman correlation was also calculated to the background peaks to estimate the background correlation distribution. We assumed that the background correlation follows Gaussian distribution. Thus, we determined the P value of the observed Spearman correlation using the 'pnorm' function in R with mean and standard deviation of the background correlation distribution. We defined significant peak–gene association with $P < 0.05$.

Chromatin accessibility across cell types and disease states

snATAC-seq profiles for patients with AD and age-matched controls were downloaded from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRP319543)[19]. Raw sequence reads were aligned to GRCh38 human reference genome by Cell Ranger ATAC (version 1.2.0) with '–mempercore = 8, –localcores =

12' options (https://support.10xgenomics.com/single-cell-atac/software/overview/welcome).
Data pre-processing and integration were performed by Seurat (version 4.0.0), Signac
(version 1.1.1), GenomicRanges (version 1.42.0) and Harmony (version 1.0)[39,44–46] as
described above. Chromatin accessibilities across cell types and disease states were
compared by average values of normalized ATAC read counts in all ATAC peaks.

Comparative analysis with HiChIP

Peak–gene associations were compared with H3K27ac HiChIP enhancer connectome data
from human substantia nigra (GSM4441831; 91 years old, low pathology)[25]. First, the
interaction loops of H3K27ac were obtained from the NCBI Gene Expression Omnibus.
The connections with $q < 0.05$ were defined as significant HiChIP loops. Because ODCs
are major components of human substantia nigra, the HiChIP loops were compared with
peak–gene associations in aged ODCs. We counted the number of peak–gene associations
matched with the HiChIP loops and statistically evaluated it with a hypergeometric test[8].

Enrichment analysis of GWAS SNPs in highly accessible open chromatin regions in each cell type

In each cell type, we defined highly accessible open chromatin regions, where ATAC reads
are detected in more than 10% of cells. Subsequently, the enrichment of disease-associated
GWAS SNPs in these open chromatin regions were assessed by LD score regression[26]. In
brief, we downloaded the plink format of genotype data in Phase 3 1000 Genomes Project,
HapMap3 SNPs, regression weight LD score files, baseline frequency files and GWAS
summary statistics from https://alkesgroup.broadinstitute.org/LDSCORE/. BED format files
of ATAC peaks were converted into ANNOT format files by the make_annot.py script
with the BIM format file of 1000 Genome Project genotype data. Using the ANNOT
file ('–annot-file'), LD scores were computed by ldsc.py script by inputting plink format
files of 1000 Genomes Project ('–bfile') and HapMap3 SNP file ('–print-snps') with '–l2–
ld-wind-cm 1–thin-annot' parameters. Partition heritability was then estimated by ldsc.py
script by inputting GWASfile ('–h2'), the computed LD score files ('–ref-ld-chr'), regression
weight files ('–w-ld-chr') and frequency files ('–frqfile-chr') with '–overlap-annot' options.
Finally, P value for heritability enrichment was used to represent the enrichment of disease-
associated SNPs within the ATAC peaks.

Subgrouping and pseudopathogenesis analysis in each cell type

In each cell type, individual cells were re-scaled in both RNA and ATAC profiles by
SCTransform (version 0.3.2) and LSI, respectively. After the re-scaling, individual cells
were projected into UMAP space. Subsequently, we performed cell trajectory analysis
using Monocle3 (version 0.2.3.0) to estimate pathological stages of individual cells[35]. In
brief, Seurat object was converted into Monocle3 'cell_data_set' format by SeuratWrapper
(version 0.3.0). Clustering was performed by Leiden community detection. Principal graph
of cell trajectory was constructed using the learn_graph function with default parameter
and used for pseudotime calculation by choosing cells in young group-specific clusters as
root cells. In each cell i, two pathological measurements, RNA-derived and ATAC-derived

pseudotime ($PT_{RNA,i}$ and $PT_{ATAC,i}$), were combined by dividing their variances of all cells (σ^2_{RNA} and σ^2_{ATAC}) as follows:

$$cPP_i = \frac{\dfrac{PT_{RNA,i}}{\sigma^2_{RNA}} + \dfrac{PT_{ATAC,i}}{\sigma^2_{ATAC}}}{\dfrac{1}{\sigma^2_{RNA}} + \dfrac{1}{\sigma^2_{ATAC}}}$$

Here, we called the combined values with corrected errors as cPP_i.

To identify differential genes and peaks along pseudopathogenesis, we calculated Spearman correlation with the scaled gene expression and peak intensity values. We chose genes and peaks with >0.1 or <−0.1 for subsequent analyses. Peaks associated with at least one pseudopathogenesis-correlated gene were used for subsequent motif enrichment analysis.

Functional subtype gene modules

To identify potential subsets of nuclei with functional differences, we used literature-derived gene sets that were previously repor ted[27,28,30–32,34]. For each set of genes, we used the AddModuleScore function of Seurat to establish a combined expression score for each nucleus for all the genes in the list and added it as metadata. The gene sets for each ODC module are as follows: newly formed ODC (*TCF7L2, CASR, CEMIP2* and *ITPR2*); myelin-forming ODC (*MAL, MOG, PLP1, OPALIN, SERINC5* and *CTPS1*); mature ODC (*KLK6, APOD, SLC5A11* and *PDE1A*); Synaptic Support ODC (*NFASC, NRXN3, CNTNAP2* and *ANK3*); and ODC–Neuron Adhesion Markers (*HAPLN2, STMN1, MAP1B, SEMA5A, EPHB2, S100B* and *PRKCA*). After cPP analysis, two additional modules were identified based on manual annotation of cPP-relevant genes, which we cross-referenced with GO analysis: Protein Folding (*HSP90AA1, HSPA1A, HSPA1B, CRYAB* and *FKBP5*) and Chaperone-Mediated Autophagy (*LAMP2, ST13, DNAJB1, STIP1, HSPA8* and *BAG3*). The gene sets for MG modules are as follows: Homeostatic MG (*HEXB, CST3, CX3CR1, CTSD, CSF1R, CTSS, SPARC, TMSB4X, P2RY12, C1QA* and *C1QB*); Stage 1 TREM2-independent DAM MG (*TYROBP, CTSB, APOE, B2M* and *FTH1*); Stage 2 TREM2-dependent DAM MG (*TREM2, AXL, CST7, CTSL, LPL, CD9, CSF1, ITGAX, CLEC7A, LILRB4* and *TIMP2*); and Aging MG (*IL15, CLEC2B* and *DOCK5*). The gene sets for AS modules are as follows: Disease-Associated AS (*GFAP, CSTB, VIM, OSMR, GSN* and *GGTA1P*); Widespread or PD Reactive AS (*GFAP, VIM, CHI3L1, MT1A, S100B* and *NFAT*); and GFAP-low AS (*LUZP2, SLC7A10* and *MFGE8*).

Cell classification by cPP

To infer multimodality of the cPP histogram, we performed kernel density estimation by the 'multimode' R package[33]. In brief, we applied the 'locmodes' function to cPP scores of ODC with 'mod0 = 3' and estimated the location of anti-modes (valleys) of multimodal distribution. We used anti-modes between the first and second modes and between the second and third modes as thresholds of 'intermediate' and 'disease' group, respectively. Comparison of previously published single-cell RNA-seq data[14] was performed using the same analysis pipeline previously described for ODCs. In brief, the ODC subset was

processed using SCTransform with the previously defined parameters. To identify if there was also a clear trajectory from healthy to disease states, we used the top 10 DEGs between healthy and disease to establish a gene expression module. Using the clusters with the highest expression of the healthy and disease gene modules as the root cells, we constructed a principal graph of cell trajectory using the learn_graph function with default parameters. We noted the appearance of a single large mode peak at the highest pseudotime value and denoted these nuclei as 'disease' for comparative analysis. We also applied the same analysis technique to other glial cell types from our samples, but the smaller cell populations tend to generate multiple small histogram peaks, and we elected not to include this analysis to prevent erroneous conclusions from being drawn through over-interpretation or 'force-fitting' of data.

Motif enrichment analysis

Transcription factor binding motifs in ATAC peaks were identified using the HOMER (version 4.11) suite. In brief, we created BED format files for specific sets of ATAC peaks and ran the findMotifsGenome.pl script with the '-size given' option to evaluate the statistical significance of the motif enrichment. $-\log_{10}(\text{FDR})$ value was used as a score of motif enrichment in the set of ATAC peaks.

SNP-containing peaks

Curated PD-related SNPs were obtained from the DisGeNET database[72]. PD-related SNP loci within ATAC peaks were detected by BEDTools[73]. To evaluate statistical significance, we calculated the expected frequency of PD-related SNPs by subsampling genomic regions, whose number and size are the same with ATAC peaks. We repeated the subsampling at 10 times and estimated Gaussian probability distribution. Then, *P* values of the observed count of PD-related SNPs within ATAC peaks were calculated from the estimated probability.

Single-molecule RNA-FISH

Multiplex RNA-FISH in 5-μm FFPE human substantia nigra brain tissue sections was performed using an RNAscope Multiplex Fluorescent v2 Assay Kit (Advanced Cell Diagnostics) according to the manufacturer's instructions. Probes were designed and manufactured by Advanced Cell Diagnostics: Hs-CARNS1-C1 (cat. no. 1171001), Hs-PDE1A-C1 (cat. no. 446121), Hs-SELENOP-C1 (cat. no. 512831), Hs-QDPR-C2 (cat. no. 560001) and Hs-MBP-C3 (cat. no. 411051). Sections were imaged using an Andor Dragonfly 2000 confocal microscope with Nikon ×40 and ×100 objectives. Quantification of RNAscope signal was performed using Imaris (version 9.9) imaging suite 'Dots' functionality (Supplementary Fig. 8). Cells with nuclear *MBP* signals were considered to be ODCs. We counted all positive RNAscope puncta over 0.75 μm in diameter within a 100-pixel (18 μm × 18 μm) square centered around the ODC nuclei, for the full depth of the tissue slice (5 μm). We used three donors from each cohort and selected 3–5 fields (based on tissue size and condition) from each tissue slide for analysis.

Statistics and reproducibility

No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications[11,14,19]. All statistical analysis methods are indicated in the figure legends where appropriate. Statistical analyses and data visualizations were performed in GraphPad Prism (version 9.0.2) or R (version 4.0.2). Data distribution was assumed to be normal, but this was not formally tested. Three samples were excluded after initial analysis indicated high ATAC signal quality with poor RNA-seq library quality, indicative of RNA degradation in the original sample. Investigators were not blinded to allocation during experiments and outcome assessment. No randomization method was used.

Extended Data
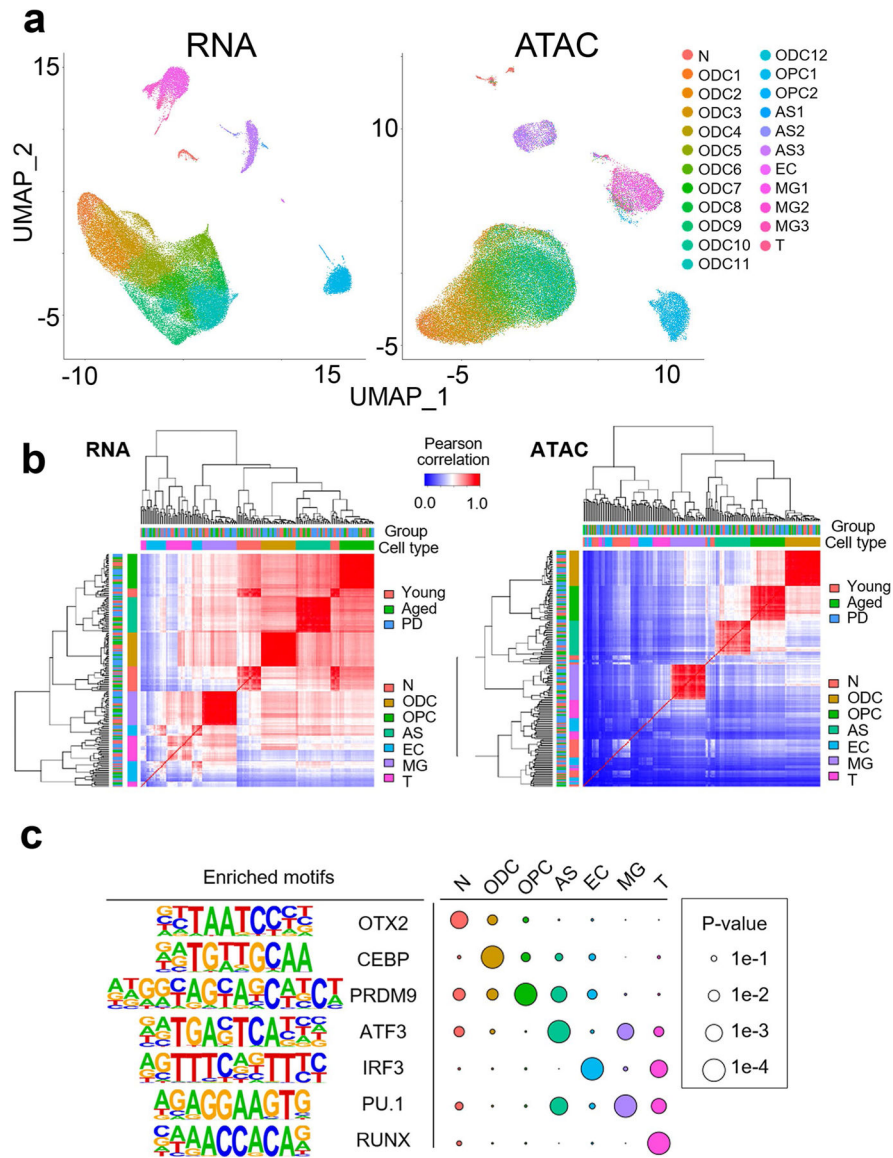
Extended Data Fig. 1 |. Multiomic analysis of human midbrain.

a, UMAP visualization of single nuclei by RNA (left) and ATAC (right) profiles. Nuclei are colored by 23 joint clusters. **b**, Heatmap showing Spearman correlation of average RNA expression (left) and ATAC peaks profiles (right) by cell types for each individual. Top and second color bars represent groups of donors and cell types, respectively. **c**, Enrichment of motifs in each annotated cell type.
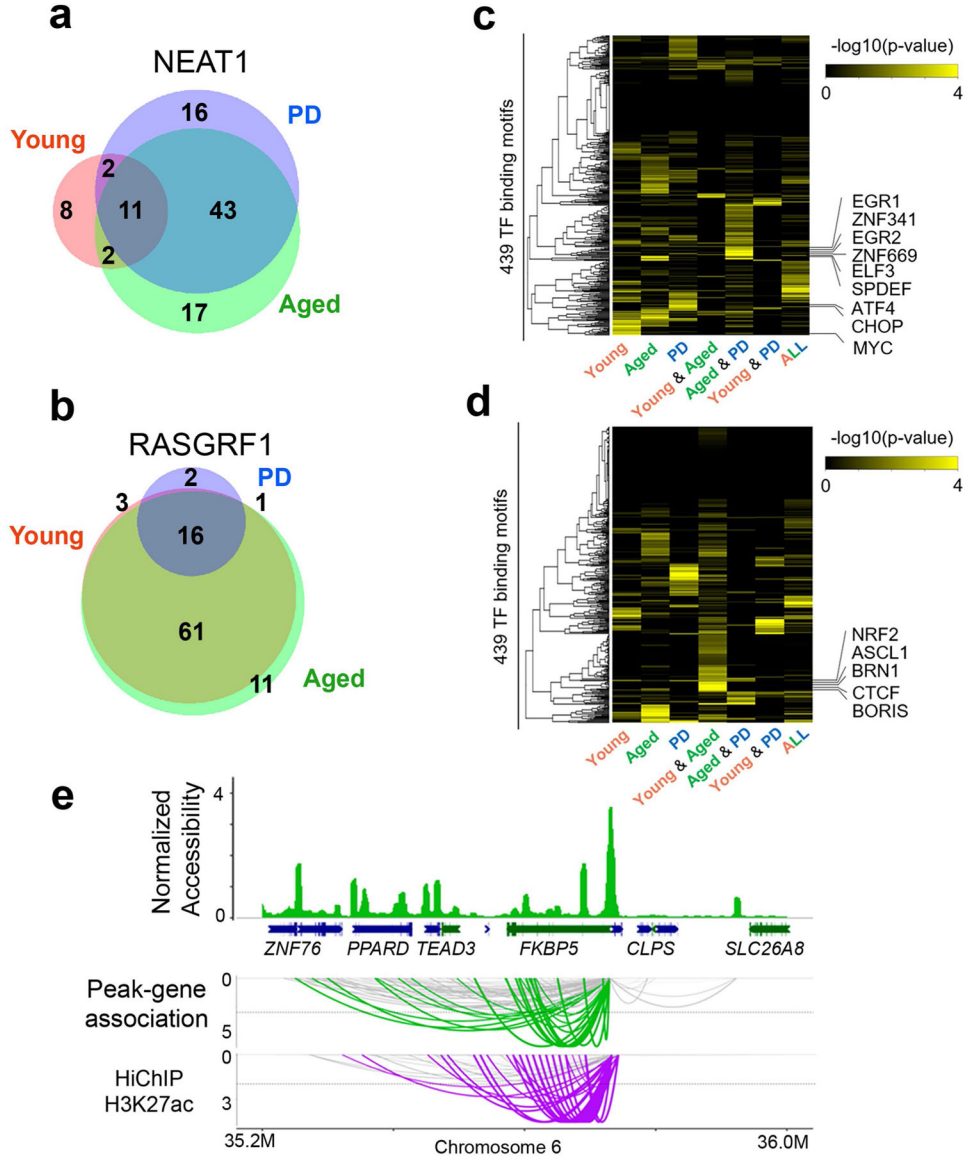

Extended Data Fig. 2 |. Changes in peak-gene connections in young, aged, and PD for NEAT1 and RASGRF1.

a,b, Venn diagram of the number of associated peaks with NEAT1 **(a)** and RASGRF1 **(b)** for young, aged, and PD midbrain. **c,d**, Heatmap showing enrichment of TF binding motifs in associated peaks with NEAT1 **(c)** and RASGRF1 **(d)**. **e**, Comparison of peak-gene association detected in our samples with H3K27ac HiChIP data from the human midbrain (Morabito et al.[19]) near the FKBP5 locus.
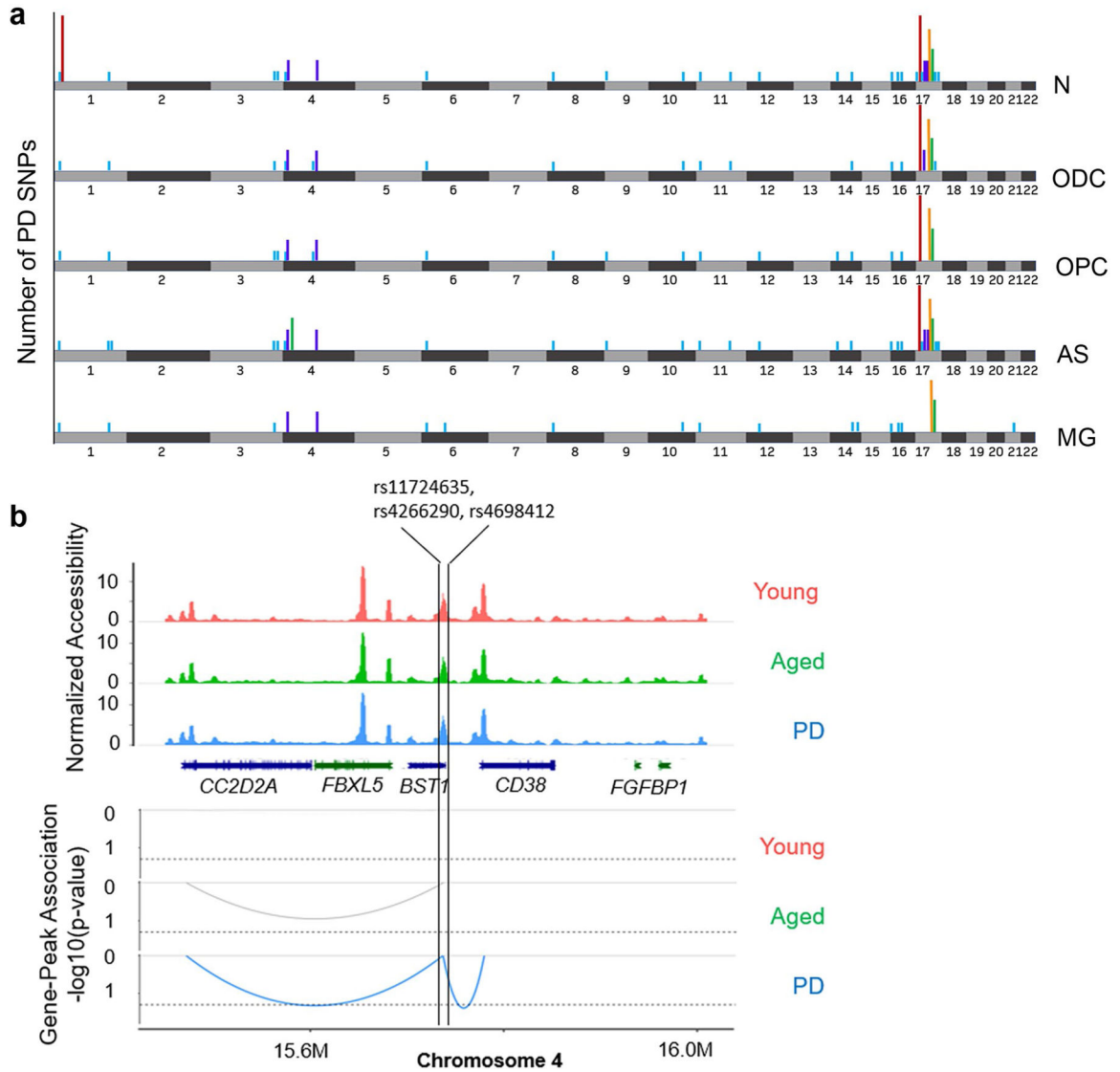
Extended Data Fig. 3 |. Different distribution of PD-associated SNPs within ATAC peaks across cell types.
a, Number of PD-related SNPs is shown on the ideogram in each cell type. (blue, 1; purple, 2; green, 3; orange, 5; red, 7 SNPs) **b**, Differential peak-gene associations in astrocytes across PD patients and healthy young and aged donors. Whereas the ATAC peak on the BST1 locus is commonly detected in PD patients and healthy donors, the peak-gene associations are different between PD patients and healthy donors. This ATAC peak contains three PD-related SNPs.
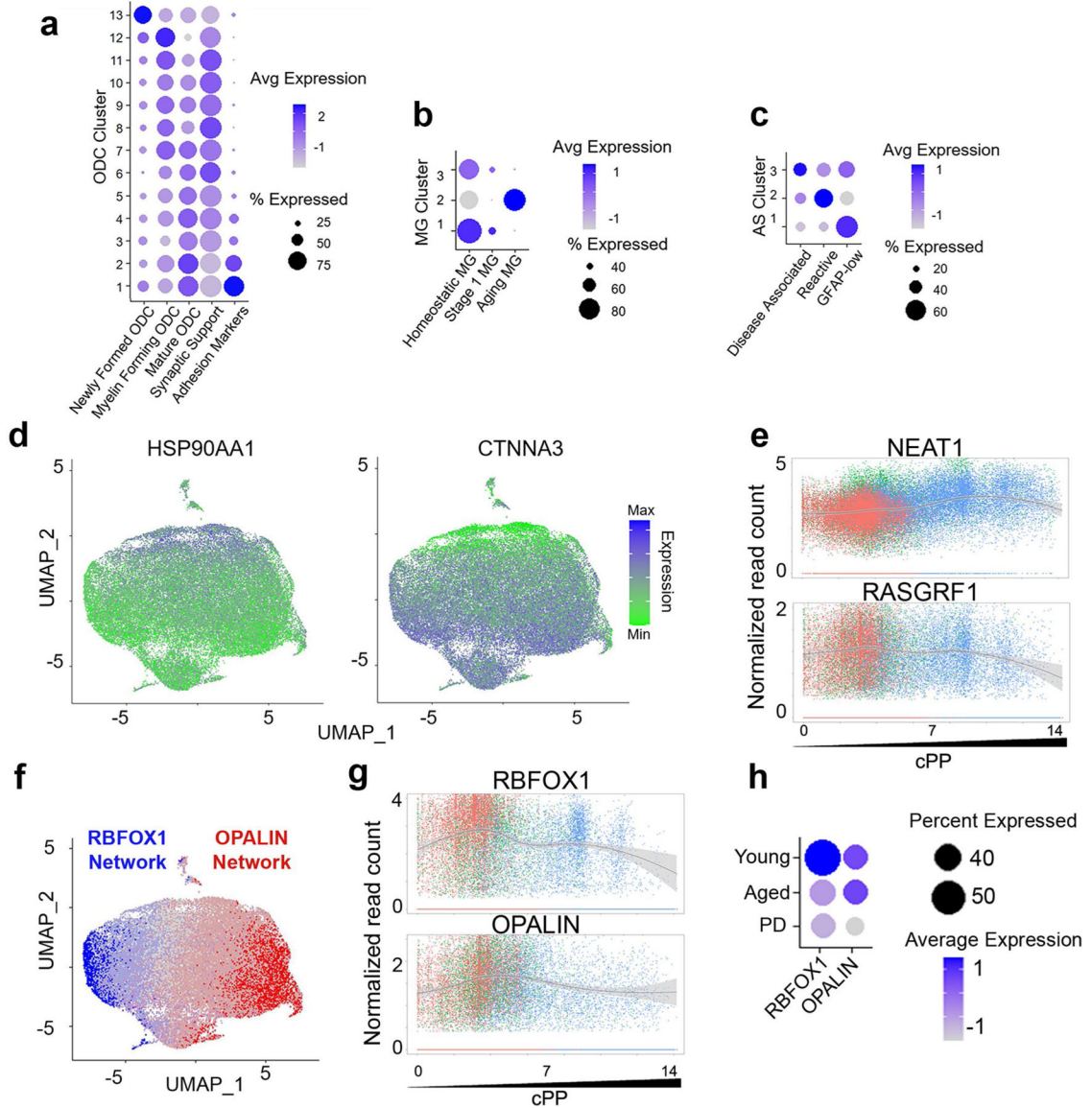
**Extended Data Fig. 4 |. cPP analysis of ODC.**

a-c, Dot plot displaying enrichment of gene expression modules for functionally distinct subpopulations for ODC **(a)**, MG **(b)**, and AS **(c)** clusters in the human midbrain. **d**, Example genes that have a correlated increase (HSP90AA1) and decrease (CTNNA3) of expression with cPP trajectory. **e**, Plot of expression of NEAT1 and RASGRF1 across cPP trajectory. NEAT1 expression is correlated with increasing cPP score; RASGRF1 expression is inversely correlated with cPP score (NEAT1: Spearman correlation = 0.314, p < 2.2e-16, RASGRF1: Spearman correlation = −0.110, p < 2.2e-16). Black line indicates loess-smoothed curve, and the gray outline represents 95% CI. **f**, UMAP of RBFOX1 and OPALIN and their coexpressed genes are mutually exclusive in ODC. **g**, Expression plot of RBFOX1 and OPALIN across cPP trajectory. n = 15,192 (Young), 11,973 (Aged), and 18,415 (PD). **h**, Dot plot of gene expression for RBFOX1 and OPALIN in each donor cohort.
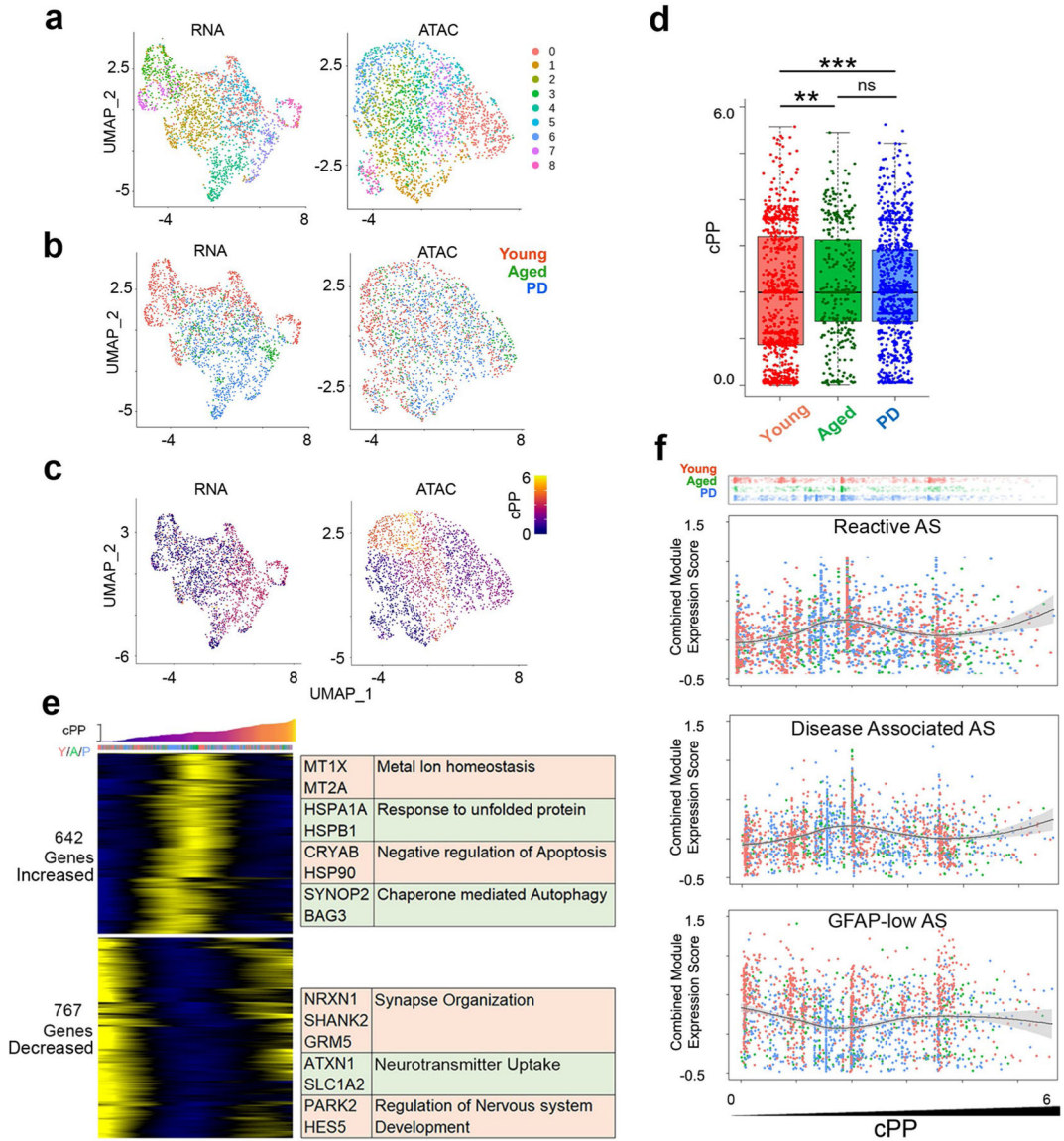
Extended Data Fig. 5 |. Establishment of pseudopathogenesis trajectory in AS.
a, UMAP plot of AS nuclei colored by subclusters **b**, UMAP plot of AS nuclei colored
by young, aged, and PD donor. **c**, UMAP plot of AS nuclei colored by cPP. **d**, cPP scores
of individual AS nuclei from young, aged, and PD midbrain are significantly changed over
aging but not a disease state. (One-way ANOVA with Tukey's post-hoc analysis, p-value for
Y/A = 0.002, p-value for Y/P = 5.28e-4, p-value for A/P = 0.24). p-values are represented as
** $p < 0.01$ and *** $p < 0.001$. **e**, Heatmap showing AS genes correlated with cPP trajectory.
X-axis represents individual cells sorted by cPP. Y-axis of heatmap represents positively
(upper)- and negatively (bottom)-correlated genes. Representative genes and significant GO
terms are shown in the right panel (Spearman correlation > 0.1 or < −0.1).n = 999 (Young),
397 (Aged), and 1,032 (PD). The bottom, center, and top of the box represent 25, 50,
and 75 percentile. Whiskers represent $1.5 \times$ IQR. **f**, Gene expression modules across AS
cPP trajectory. Top panel shows individual nuclei AS along with cPP scores and donor

group. X-axis shows the cPP score. Y-axis is the combined expression level for all genes in the expression module. Black line indicates loess-smoothed curve, and the gray outline represents 95% CI.
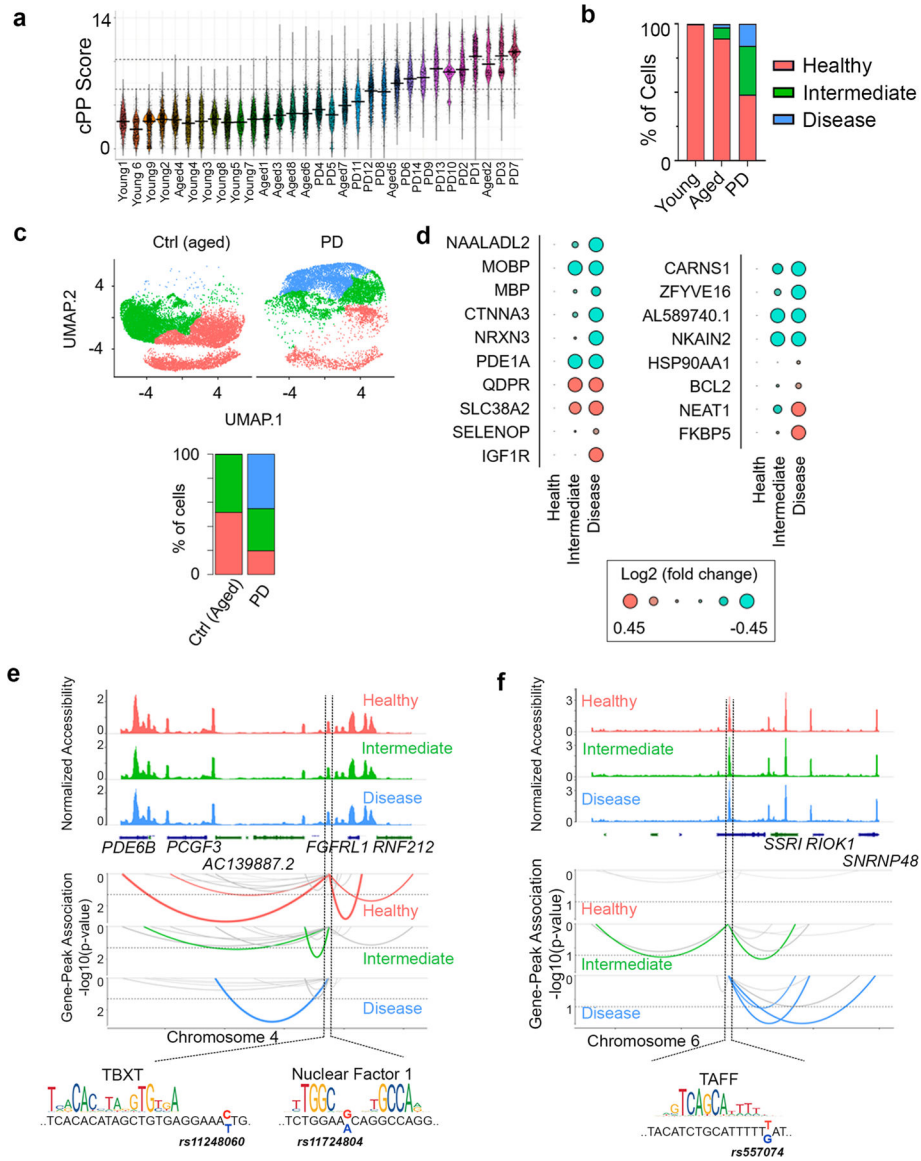
Extended Data Fig. 6 |. Aging- and disease-specific analysis.
a, Violin plot of cPP score of every individual ODC nuclei by donor. **b**, Bar graph showing the percentage of nuclei in donor cohort from healthy, intermediate, and disease groups. **c**, UMAP showing healthy, intermediate, and disease subsets from publicly available snRNA-seq data (Smaji et al, 2022) from the human PD and aged control midbrain after pseudopathogenesis analysis. **d**, Dot plots of genes from the same dataset (Smaji et al, 2022) showing similar expression patterns among healthy, intermediate, and disease subsets as our multiomic dataset. Circle size represents relative gene expression to healthy subsets. **e,f**, Representative peak-gene connection plots for peaks containing PD-associated

SNPs that have decreased (**e**) or increased (**f**) gene connections in disease-associated ODC compared to healthy ones. Motif information was obtained from the JASPAR Transcription Factors track in the UCSC genome browser. SNPs associated with each peak are shown below.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data availability

Processed data for the samples presented in this study are available in the Gene Expression Omnibus database under accession number GSE193688. Raw data are available through the dbGaP portal under accession number phs002819.v1.p1. snRNA-seq and ATAC-seq reads were mapped to GRCh38 human reference (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.26/).

References

1. Parkinson's Foundation. Statistics. https://www.parkinson.org/understanding-parkinsons/statistics

2. Pringsheim T, Jette N, Frolkis A & Steeves TD The prevalence of Parkinson's disease: a systematic review and meta-analysis. Mov. Disord 29, 1583–1590 (2014). [PubMed: 24976103]

3. Dauer W & Przedborski S Parkinson's disease: mechanisms and models. Neuron 39, 889–909 (2003). [PubMed: 12971891]

4. Hindle JV Ageing, neurodegeneration and Parkinson's disease. Age Ageing 39, 156–161 (2010). [PubMed: 20051606]

5. Davie K et al. A single-cell transcriptome atlas of the aging *Drosophila* brain. Cell 174, 982–998 (2018). [PubMed: 29909982]

6. Ximerakis M et al. Single-cell transcriptomic profiling of the aging mouse brain. Nat. Neurosci 22, 1696–1708 (2019). [PubMed: 31551601]

7. Glaab E & Schneider R Comparative pathway and network analysis of brain transcriptome changes during adult aging and in Parkinson's disease. Neurobiol. Dis 74, 1–13 (2015). [PubMed: 25447234]

8. Ma S et al. Chromatin potential identified by shared single-cell profiling of RNA and chromatin. Cell 183, 1103–1116 (2020). [PubMed: 33098772]

9. Mathys H et al. Single-cell transcriptomic analysis of Alzheimer's disease. Nature 570, 332–337 (2019). [PubMed: 31042697]

10. Nott A et al. Brain cell type-specific enhancer–promoter interactome maps and disease–risk association. Science 366, 1134–1139 (2019). [PubMed: 31727856]

11. 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). [PubMed: 32826893]

12. 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). [PubMed: 32341526]

13. Reynolds RH et al. Moving beyond neurons: the role of cell type-specific gene regulation in Parkinson's disease heritability. NPJ Parkinsons Dis 5, 6 (2019). [PubMed: 31016231]

14. Smajic S et al. Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. Brain 145, 964–978 (2021).

15. Braak H & Del Tredici K Poor and protracted myelination as a contributory factor to neurodegenerative disorders. Neurobiol. Aging 25, 19–23 (2004). [PubMed: 14675725]

16. Hao Y et al. Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587 (2021). [PubMed: 34062119]

17. Cakir B et al. Engineering of human brain organoids with a functional vascular-like system. Nat. Methods 16, 1169–1175 (2019). [PubMed: 31591580]

18. Tanaka Y, Cakir B, Xiang Y, Sullivan GJ & Park IH Synthetic analyses of single-cell transcriptomes from multiple brain organoids and fetal brain. Cell Rep 30, 1682–1689 (2020). [PubMed: 32049002]

19. Morabito S et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat. Genet 53, 1143–1155 (2021). [PubMed: 34239132]

20. Borras C et al. RasGrf1 deficiency delays aging in mice. Aging (Albany NY) 3, 262–276 (2011). [PubMed: 21422498]

21. Salmon AB et al. Effects of transgenic methionine sulfoxide reductase A (MsrA) expression on lifespan and age-dependent changes in metabolic function in mice. Redox Biol 10, 251–256 (2016). [PubMed: 27821326]

22. Duclot F & Kabbaj M The role of early growth response 1 (EGR1) in brain plasticity and neuropsychiatric disorders. Front. Behav. Neurosci 11, 35 (2017). [PubMed: 28321184]

23. Kim EJ, Ables JL, Dickel LK, Eisch AJ & Johnson JE Ascl1 (Mash1) defines cells with long-term neurogenic potential in subgranular and subventricular zones in adult mouse brain. PLoS ONE 6, e18472 (2011). [PubMed: 21483754]

24. Sandberg M, Patil J, D'Angelo B, Weber SG & Mallard C NRF2-regulation in brain health and disease: implication of cerebral inflammation. Neuropharmacology 79, 298–306 (2014). [PubMed: 24262633]

25. Corces MR et al. Single-cell epigenomic analyses implicate candidate causal variants at inherited risk loci for Alzheimer's and Parkinson's diseases. Nat. Genet 52, 1158–1168 (2020). [PubMed: 33106633]

26. Bulik-Sullivan BK et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet 47, 291–295 (2015). [PubMed: 25642630]

27. Galatro TF et al. Transcriptomic analysis of purified human cortical microglia reveals age-associated changes. Nat. Neurosci 20, 1162–1171 (2017). [PubMed: 28671693]

28. Habib N et al. Disease-associated astrocytes in Alzheimer's disease and aging. Nat. Neurosci 23, 701–706 (2020). [PubMed: 32341542]

29. Jakel S et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. Nature 566, 543–547 (2019). [PubMed: 30747918]

30. Keren-Shaul H et al. A unique microglia type associated with restricting development of Alzheimer's disease. Cell 169, 1276–1290 (2017). [PubMed: 28602351]

31. Liddelow SA et al. Neurotoxic reactive astrocytes are induced by activated microglia. Nature 541, 481–487 (2017). [PubMed: 28099414]

32. Marques S et al. Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. Science 352, 1326–1329 (2016). [PubMed: 27284195]

33. Teeple E et al. Single nuclei sequencing of human putamen oligodendrocytes reveals altered heterogeneity and disease-associated changes in Parkinson's disease and multiple system atrophy. Preprint at bioRxiv 10.1101/2021.05.06.442967 (2021).

34. Escartin C et al. Reactive astrocyte nomenclature, definitions, and future directions. Nat. Neurosci 24, 312–325 (2021). [PubMed: 33589835]

35. Cao J et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502 (2019). [PubMed: 30787437]

36. Ameijeiras-Alonso J, Crujeiras RM & Rodriguez-Casal A multimode: an R package for mode assessment. J. Stat. Softw 97, 1–32 (2021).

37. Hipkiss AR Glycotoxins: dietary and metabolic origins; possible amelioration of neurotoxicity by carnosine, with special reference to Parkinson's disease. Neurotox. Res 34, 164–172 (2018). [PubMed: 29417441]

38. Kubota M et al. Carnosine suppresses neuronal cell death and inflammation induced by 6-hydroxydopamine in an in vitro model of Parkinson's disease. PLoS ONE 15, e0240448 (2020). [PubMed: 33052927]

39. Stuart T et al. Comprehensive integration of single-cell data. Cell 177, 1888–1902 (2019). [PubMed: 31178118]

40. Xiang Y et al. Fusion of regionally specified hPSC-derived organoids models human brain development and interneuron migration. Cell Stem Cell 21, 383–398 (2017). [PubMed: 28757360]

41. Amemiya HM, Kundaje A & Boyle AP The ENCODE blacklist: identification of problematic regions of the genome. Sci. Rep 9, 9354 (2019). [PubMed: 31249361]

42. Xiang Y et al. Dysregulation of BRD4 function underlies the functional abnormalities of MeCP2 mutant neurons. Mol. Cell 79, 84–98 (2020). [PubMed: 32526163]

43. Falcon S & Gentleman R Using GOstats to test gene lists for GO term association. Bioinformatics 23, 257–258 (2007). [PubMed: 17098774]

44. Korsunsky I et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods 16, 1289–1296 (2019). [PubMed: 31740819]

45. Lawrence M et al. Software for computing and annotating genomic ranges. PLoS Comput. Biol 9, e1003118 (2013). [PubMed: 23950696]

46. Stuart T, Srivastava A, Madad S, Lareau CA & Satija R Single-cell chromatin state analysis with Signac. Nat Methods 18, 1333–1341 (2021). [PubMed: 34725479]

47. Heinz S et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589 (2010). [PubMed: 20513432]

48. Al-Dalahmah O et al. Single-nucleus RNA-seq identifies Huntington disease astrocyte states. Acta Neuropathol. Commun 8, 19 (2020). [PubMed: 32070434]

49. Baldo B et al. A screen for enhancers of clearance identifies huntingtin as a heat shock protein 90 (Hsp90) client protein. J. Biol. Chem 287, 1406–1414 (2012). [PubMed: 22123826]

50. Barry G et al. The long non-coding RNA NEAT1 is responsive to neuronal activity and is associated with hyperexcitability states. Sci. Rep 7, 40127 (2017). [PubMed: 28054653]

51. Barry G, Guennewig B, Fung S, Kaczorowski D & Weickert CS Long non-coding RNA expression during aging in the human subependymal zone. Front. Neurol 6, 45 (2015). [PubMed: 25806019]

52. Binder EB et al. Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. Nat. Genet 36, 1319–1325 (2004). [PubMed: 15565110]

53. Blair LJ et al. Accelerated neurodegeneration through chaperone-mediated oligomerization of tau. J. Clin. Invest 123, 4158–4169 (2013). [PubMed: 23999428]

54. Chang LC et al. A conserved BDNF, glutamate- and GABA-enriched gene module related to human depression identified by coexpression meta-analysis and DNA variant genome-wide association studies. PLoS ONE 9, e90980 (2014). [PubMed: 24608543]

55. Evers M, Cunningham-Rundles C & Hollander E Heat shock protein 90 antibodies in autism. Mol. Psychiatry 7, S26–S28 (2002). [PubMed: 12142940]

56. Karis K et al. Altered expression profile of IgLON family of neural cell adhesion molecules in the dorsolateral prefrontal cortex of schizophrenic patients. Front. Mol. Neurosci 11, 8 (2018). [PubMed: 29434535]

57. Katsel P et al. The expression of long noncoding RNA NEAT1 is reduced in schizophrenia and modulates oligodendrocytes transcription. NPJ Schizophr 5, 3 (2019). [PubMed: 30696826]

58. Kim JJ et al. Identification of antibodies to heat shock proteins 90 kDa and 70 kDa in patients with schizophrenia. Schizophr. Res 52, 127–135 (2001). [PubMed: 11595400]

59. Labadorf A et al. RNA sequence analysis of human Huntington disease brain reveals an extensive increase in inflammatory and developmental gene expression. PLoS ONE 10, e0143563 (2015). [PubMed: 26636579]

60. Lee CH et al. Heat shock protein 90 and its cochaperone, p23, are markedly increased in the aged gerbil hippocampus. Exp. Gerontol 46, 768–772 (2011). [PubMed: 21601628]

61. Liu Y & Lu Z Long non-coding RNA NEAT1 mediates the toxic of Parkinson's disease induced by MPTP/MPP+ via regulation of gene expression. Clin. Exp. Pharmacol. Physiol 45, 841–848 (2018). [PubMed: 29575151]

62. Maccarrone G et al. Psychiatric patient stratification using biosignatures based on cerebrospinal fluid protein expression clusters. J. Psychiatr. Res 47, 1572–1580 (2013). [PubMed: 23962679]

63. Mariani E et al. Meta-analysis of Parkinson's disease transcriptome data using TRAM software: whole substantia nigra tissue and single dopamine neuron differential gene expression. PLoS ONE 11, e0161567 (2016). [PubMed: 27611585]

64. Patel H et al. Transcriptomic analysis of probable asymptomatic and symptomatic alzheimer brains. Brain Behav. Immun 80, 644–656 (2019). [PubMed: 31063847]

65. Patel N, Crider A, Pandya CD, Ahmed AO & Pillai A Altered mRNA levels of glucocorticoid receptor, mineralocorticoid receptor, and co-chaperones (FKBP5 and PTGES3) in the middle frontal gyrus of autism spectrum disorder subjects. Mol. Neurobiol 53, 2090–2099 (2016). [PubMed: 25912394]

66. Sha L et al. Pharmacologic inhibition of Hsp90 to prevent GLT-1 degradation as an effective therapy for epilepsy. J. Exp. Med 214, 547–563 (2017). [PubMed: 28028152]

67. Sinclair D, Fillman SG, Webster MJ & Weickert CS Dysregulation of glucocorticoid receptor co-factors FKBP5, BAG1 and PTGES3 in prefrontal cortex in psychotic illness. Sci. Rep 3, 3539 (2013). [PubMed: 24345775]

68. Sunwoo JS et al. Altered expression of the long noncoding RNA NEAT1 in Huntington's disease. Mol. Neurobiol 54, 1577–1586 (2017). [PubMed: 27221610]

69. Uryu K et al. Convergence of heat shock protein 90 with ubiquitin in filamentous α-synuclein inclusions of α-synucleinopathies. Am. J. Pathol 168, 947–961 (2006). [PubMed: 16507910]

70. Woo JA et al. Enhanced tau pathology via RanBP9 and Hsp90/Hsc70 chaperone complexes. Hum. Mol. Genet 26, 3973–3988 (2017). [PubMed: 29016855]

71. Zhao MY et al. The long-non-coding RNA NEAT1 is a novel target for Alzheimer's disease progression via miR-124/BACE1 axis. Neurol. Res 41, 489–497 (2019). [PubMed: 31014193]

72. Pinero J et al. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. Database (Oxford) 2015, bav028 (2015).

73. Quinlan AR & Hall IM BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010). [PubMed: 20110278]
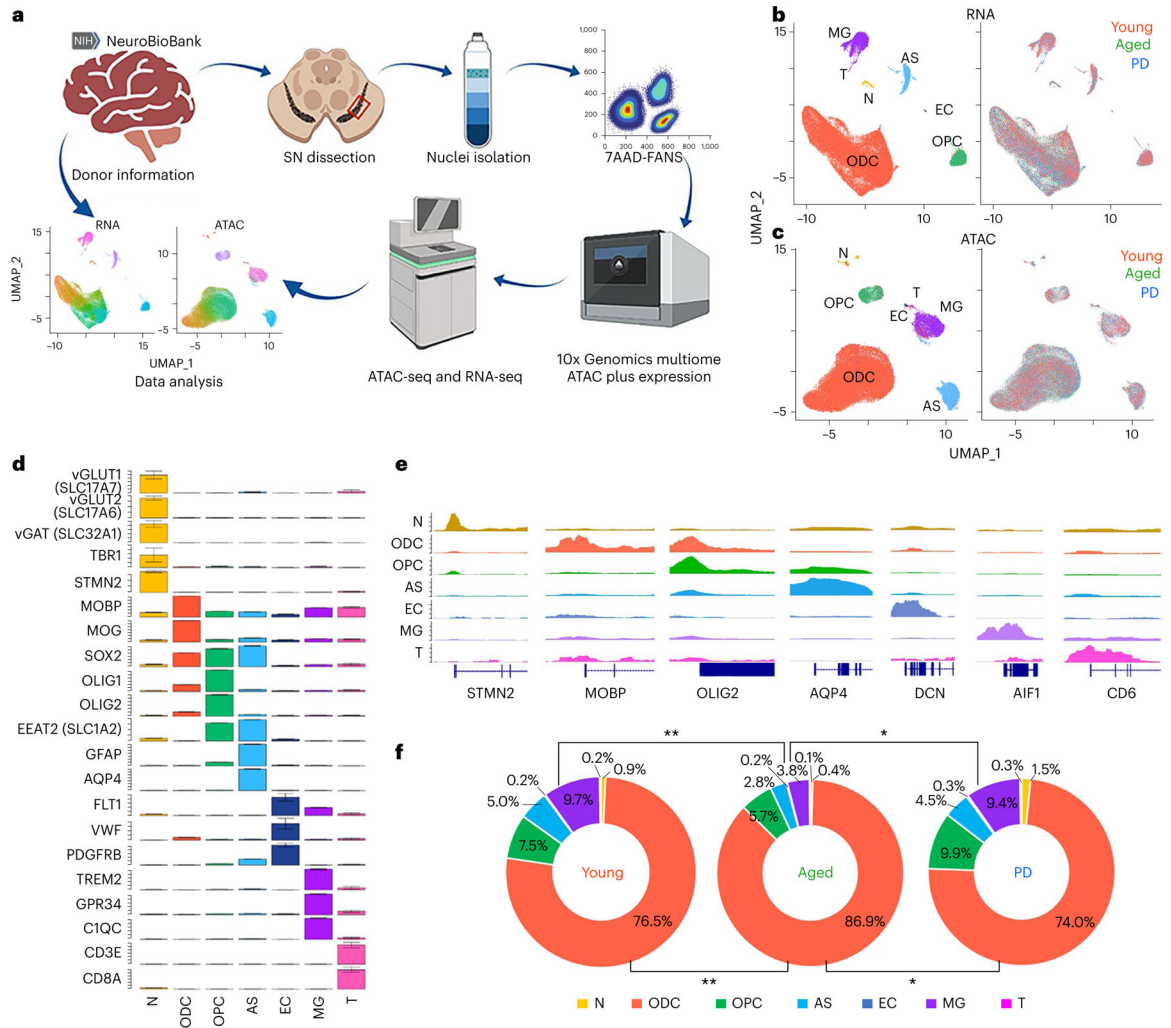
Fig. 1 |. Multiomic analysis of human midbrain.
a, Schematic of isolation and snATAC plus gene expression analysis for human postmortem midbrain (created with BioRender). **b,c**, UMAP visualization of single nuclei by RNA (**b**) and ATAC (**c**) profiles. Nuclei are colored by identified cell type (left) and donor type (right). **d**, Expression of cell-type-specific genes in each annotated cell type cluster. Error bars represent mean ± s.e.m. **e**, Chromatin accessibility at cell-type-specific genes in each annotated cell type cluster. **f**, Percentage of each cell type passing quality control criteria identified in young, aged and PD. Two-sided Student's *t*-test shows that ODCs significantly changed from young to aged ($P = 0.002$) and aged to PD ($P = 0.023$). MG significantly changed from young to aged ($P = 0.007$) and aged to PD ($P = 0.042$). *P* values are represented as *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$. Ns ($n = 489$), ODCs ($n = 45,580$), ASs ($n = 2,428$), MG ($n = 4,482$), OPCs ($n = 4,396$), ECs ($n = 125$) and Ts ($n = 110$). SN, substantia nigra.
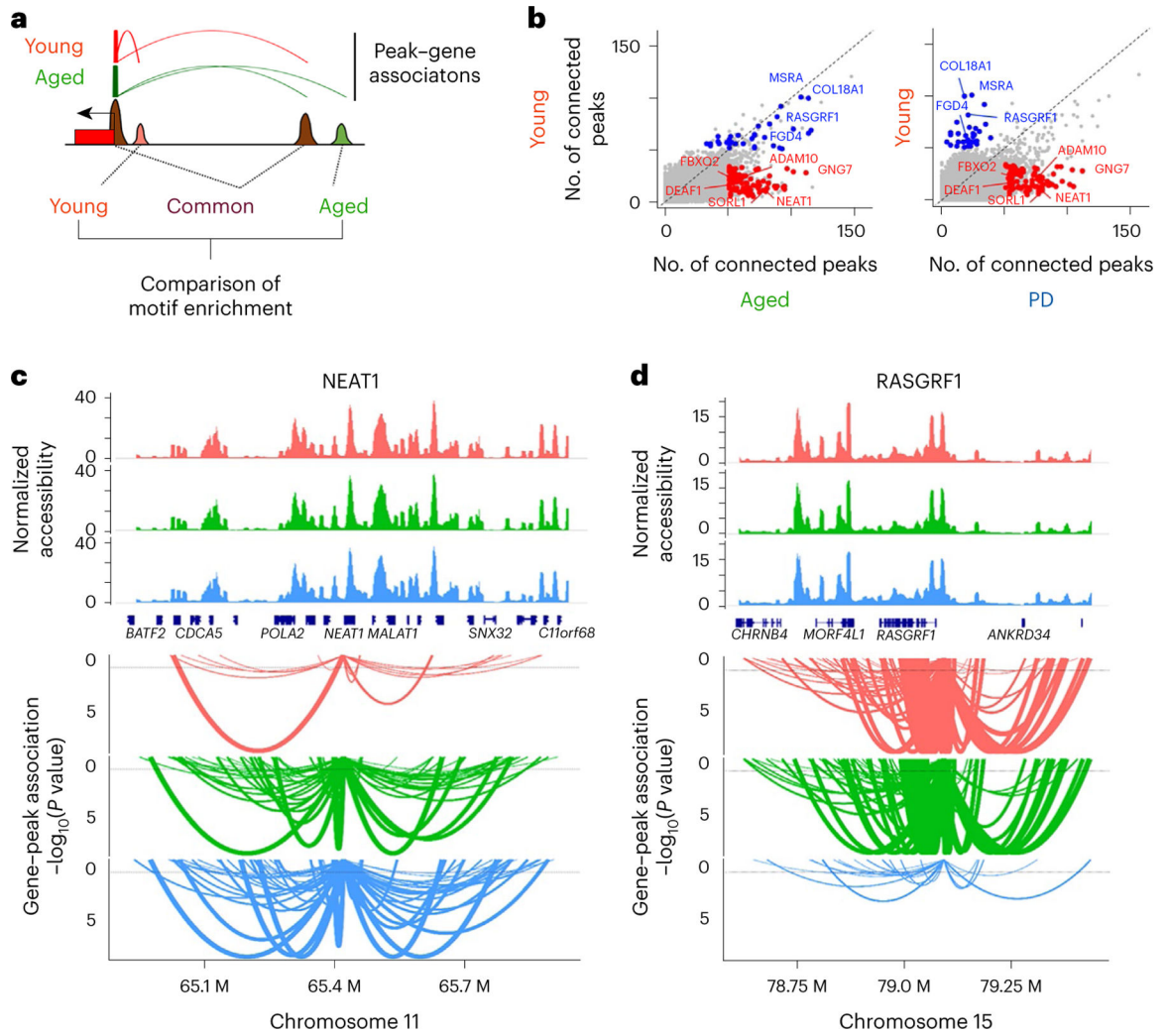
Fig. 2 |. Analysis of peak–gene connections in the human midbrain.
a, Schematic of peak–gene analysis between samples. **b**, Comparison of the number of connected peaks in each gene between young and aged (left) and young and PD (right) midbrain. Some genes show altered peak connections from young in both aged and PD (red dots). Some genes show altered peak connections only in PD (blue dots). **c,d**, Peak–gene connection plots for NEAT1 (**c**) and RASGRF1 (**d**) among young, aged and PD midbrain. Top, normalized ATAC read count distribution. Bottom, arcs show ATAC peaks significantly correlated with gene expression. Arc height represents statistical significance.
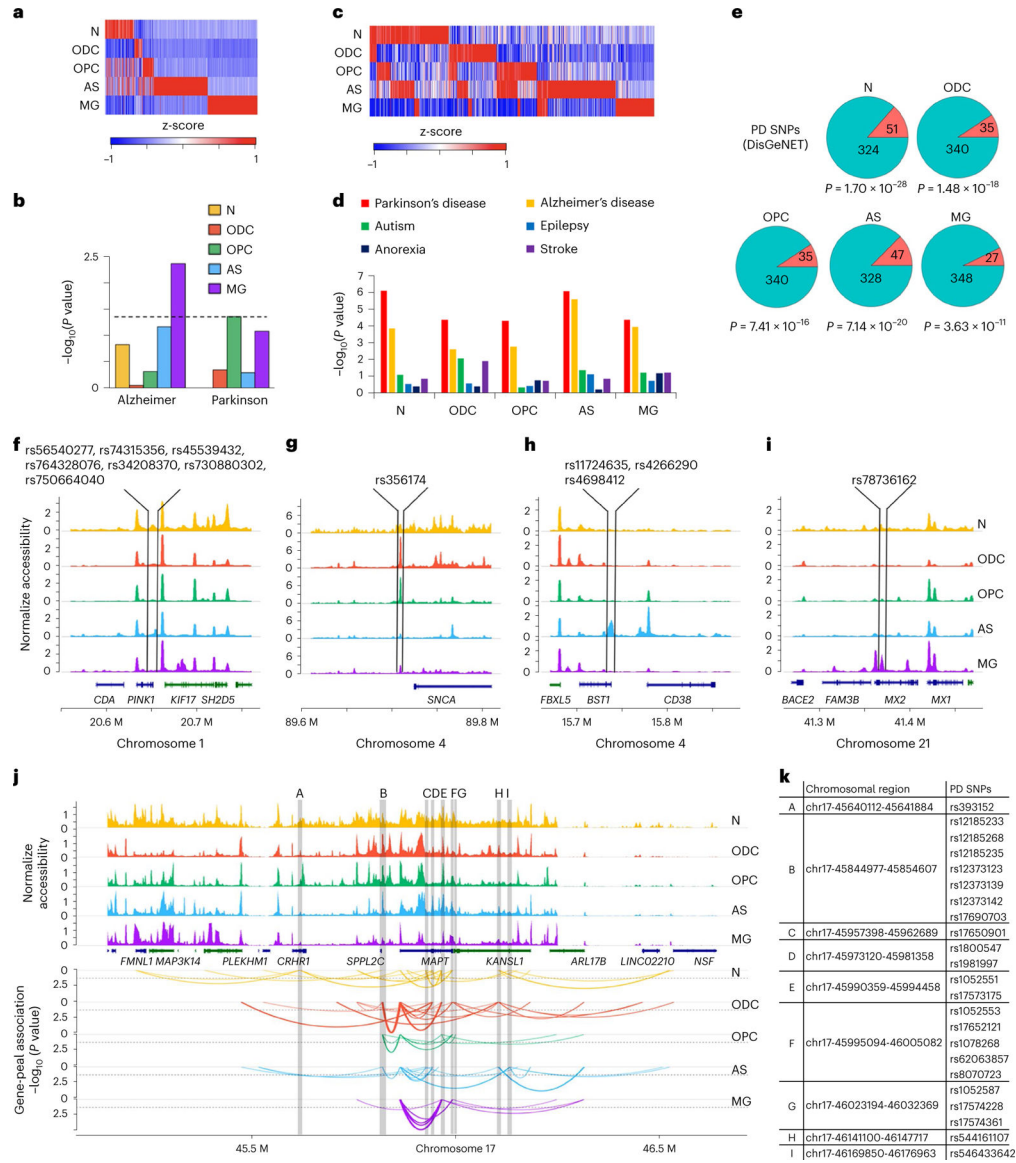
Fig. 3 |. Enrichment of PD-associated SNPs in highly accessible open chromatin regions.
a, Heat map showing the normalized chromatin accessibility of cell-type-specific ATAC peaks by Corces et al.[25] criteria. **b**, The enrichment of GWAS SNPs from AD and PD in the cell-type-specific ATAC peaks. **c**, Heat map showing the normalized chromatin accessibility of highly accessible peaks. **d**, LD score regression analysis of GWAS SNPs associated with neuronal and non-neuronal diseases. $-\log_{10}(P$ value) for the heritability enrichment is shown by a bar graph. **e**, Pie chart representing the number of PD-associated SNPs within cell-type-specific ATAC peaks. Statistical significance was calculated by hypergeometric test. **f–i**, Chromatin accessibility around cell-type-specific ATAC peaks in four different regions. These peaks are unique to Ns (**f**), ODCs/OPCs (**g**), ASs (**h**) and MG (**i**) and contain the indicated PD-associated SNPs. **j**, Differential chromatin accessibility and peak–gene associations around MAPT gene across cell types. Peak–gene associations were calculated from single-cell multiome data of patients with PD. Significant peak–gene associations

are represented by colored thick curves. **k**, Chart showing PD-associated SNPs in the highlighted regions of **j**.
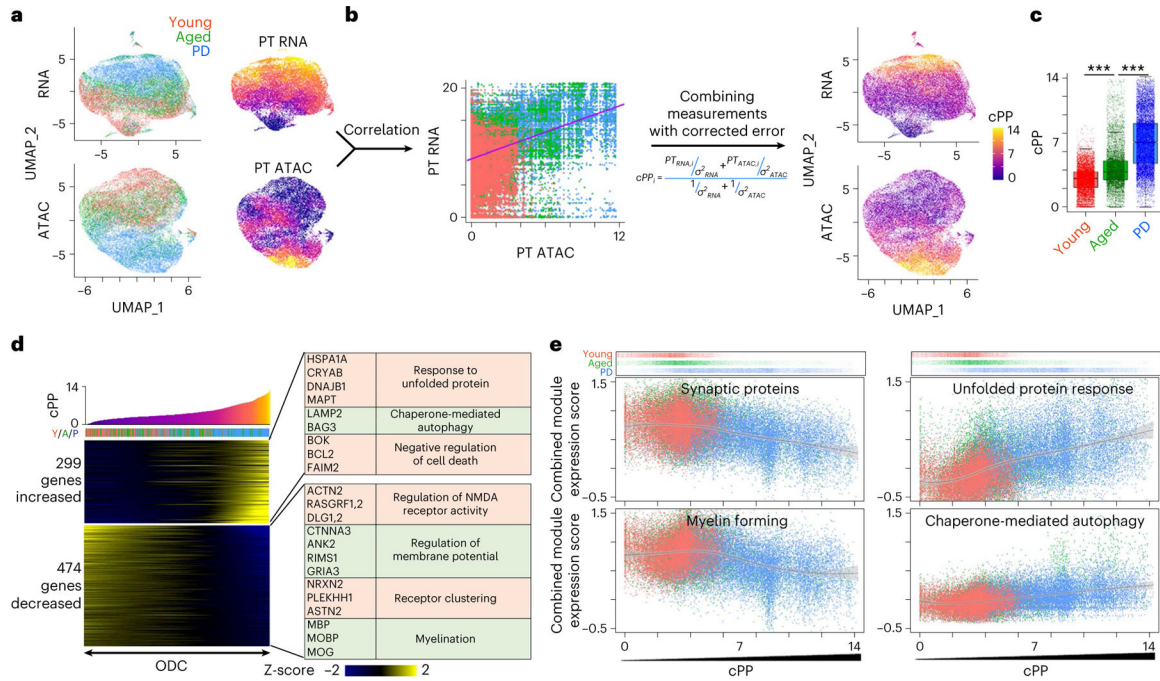
Fig. 4 |. Establishment of pseudopathogenesis trajectory in ODCs.

a, UMAP plot of ODC nuclei colored by young, aged and PD donor (left) and heat map of RNA-based and ATAC-based pseudotime trajectory (right). **b**, Schematic of calculation of cPP score from transcriptome and ATAC profiles (Spearman correlation = 0.419; $P = 2.2 \times 10^{-16}$). **c**, Significant difference in cPP scores of individual ODC nuclei from young, aged and PD midbrain (one-way ANOVA with Tukey's post hoc analysis; P for Y/A = 1 $\times 10^{-199}$, P for Y/P = 1×10^{-199} and P for A/P = 1×10^{-199}). P values are represented as $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$. $n = 15{,}192$ (young), $n = 11{,}973$ (aged) and $n = 18{,}415$ (PD). The bottom, center and top of the box represent 25th, 50th and 75th percentiles, respectively. Whiskers represent 1.5× interquartile range. **d**, Heat map showing genes correlated with cPP trajectory. x axis represents individual cells sorted by cPPs. y axis of heat map represents positively (upper) and negatively (bottom) correlated genes. Representative genes and significant GO terms are shown in the right panel (Spearman correlation > 0.1 or Spearman correlation < −0.1). **e**, Change in gene expression modules across cPP trajectory in ODCs. Top panel shows individual ODC nuclei along cPP scores and donor group. x axis shows cPP score. y axis is the combined expression score for all genes in the expression module. Black line indicates LOESS-smoothed curve, and the gray outline represents 95% confidence interval. PT, pseudotime trajectory.
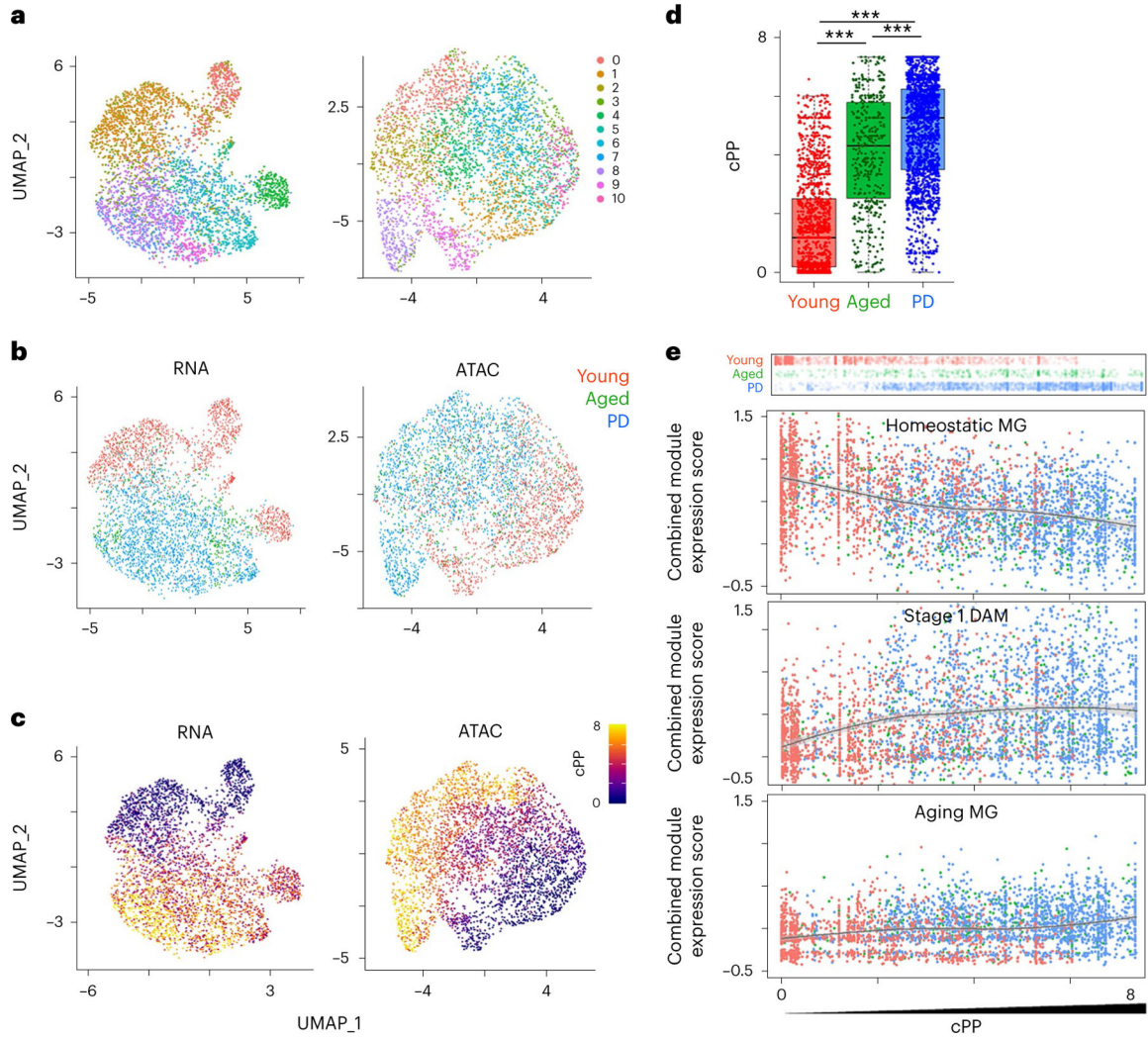
Fig. 5 |. Establishment of cPP trajectory in MG.

a, UMAP plot of MG nuclei colored by subclusters **b**, UMAP plot of MG nuclei colored by young, aged and PD donors. **c**, UMAP plot of MG nuclei colored by cPP. **d**, Significant difference in cPP scores of individual MG nuclei from young, aged and PD midbrain (one-way ANOVA with Tukey's post hoc analysis; P for Y/A = 1.71×10^{-8}, P for Y/P = 1.71×10^{-8} and P for A/P = 1.71×10^{-8}). P values are represented as $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$. $n = 1,930$ (young), $n = 460$ (aged) and $n = 2,092$ (PD). The bottom, center and top of the box represent 25th, 50th and 75th percentiles, respectively. Whiskers represent $1.5\times$ interquartile range. **e**, Gene expression modules across MG cPP trajectory. Top panel shows individual MG nuclei along with cPP scores and donor group. x axis shows cPP score. y axis is the combined expression score for all genes in the expression module. Black line indicates LOESS-smoothed curve, and the gray outline represents 95% confidence interval.
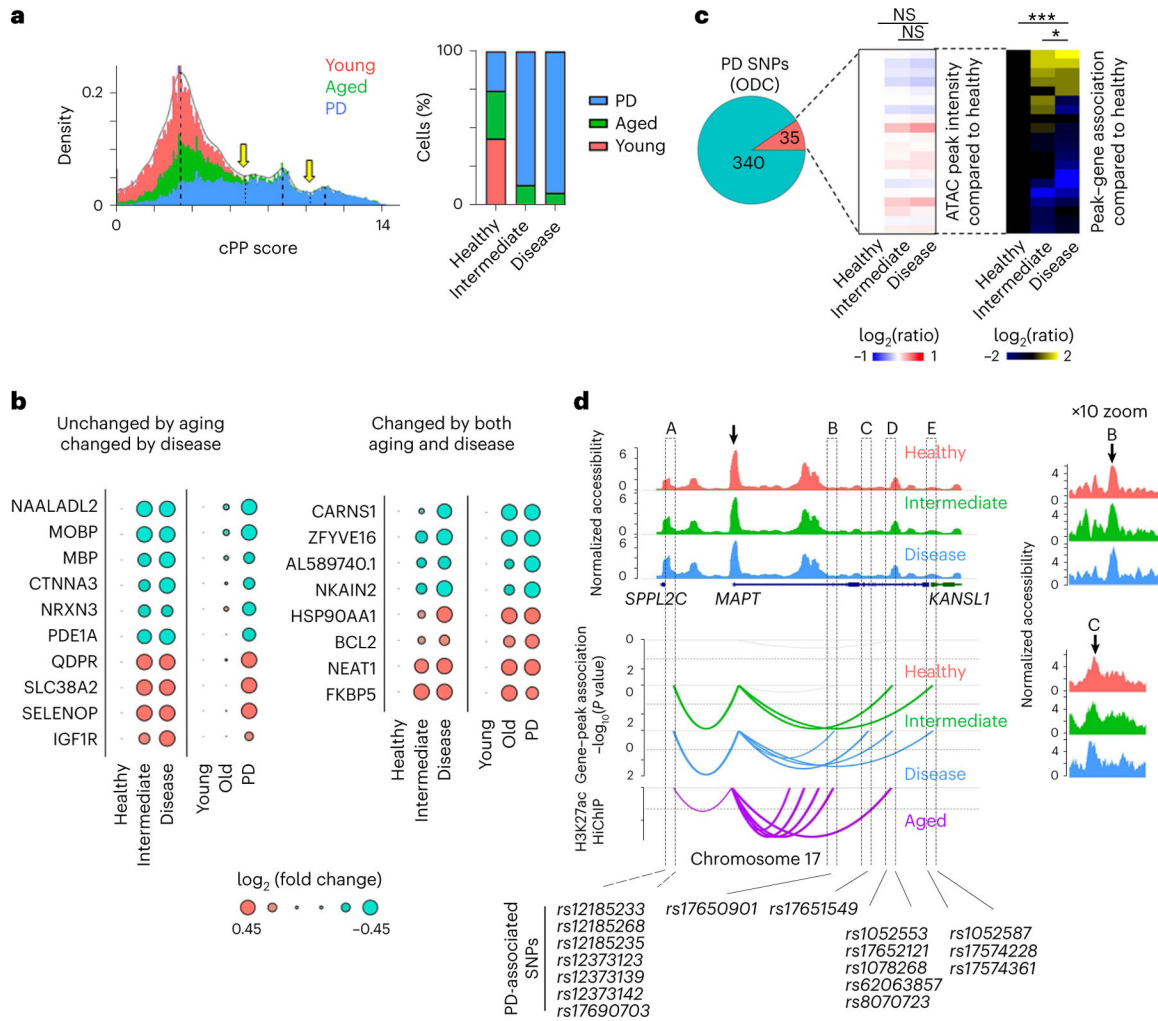
Fig. 6 |. Analysis of disease-associated ODCs.

a, Histogram and inferred multimodal distribution of ODCs across cPP. Mode valleys were identified using the 'multimode' package in R. Peaks were identified as healthy, intermediate and disease-associated ODCs. Bar graph shows the percentage of nuclei in each group from the three donor cohorts. **b**, Dot plot of selected genes with differential expression among healthy, intermediate and disease-associated cells. Left, genes with disease-specific changes with no differential expression between young and aged/PD. Right, genes differentially expressed over aging/PD. Circle size represents relative gene expression to healthy subsets or young. **c**, Heat maps represents relative peak intensity and peak–gene association to healthy group ($\log_2$(ratio)) for PD-associated SNP-containing peaks among healthy, intermediate and disease-associated ODCs. Statistical difference of the peak intensity and peak–gene association across the three groups were assessed by two-sided Student's t-test. P for H/D = 8.28×10^{-4} and P for I/D = 1.49×10^{-2}. P values are represented as $*P < 0.05$ and $***P < 0.001$. **d**, Peak–gene association plots for peaks with PD-associated SNPs that are significantly correlated with MAPT gene expression in healthy, intermediate and disease-associated ODCs. Black arrow indicates the location of MAPT promoter. SNPs associated with each peak are shown below. NS, not significant.
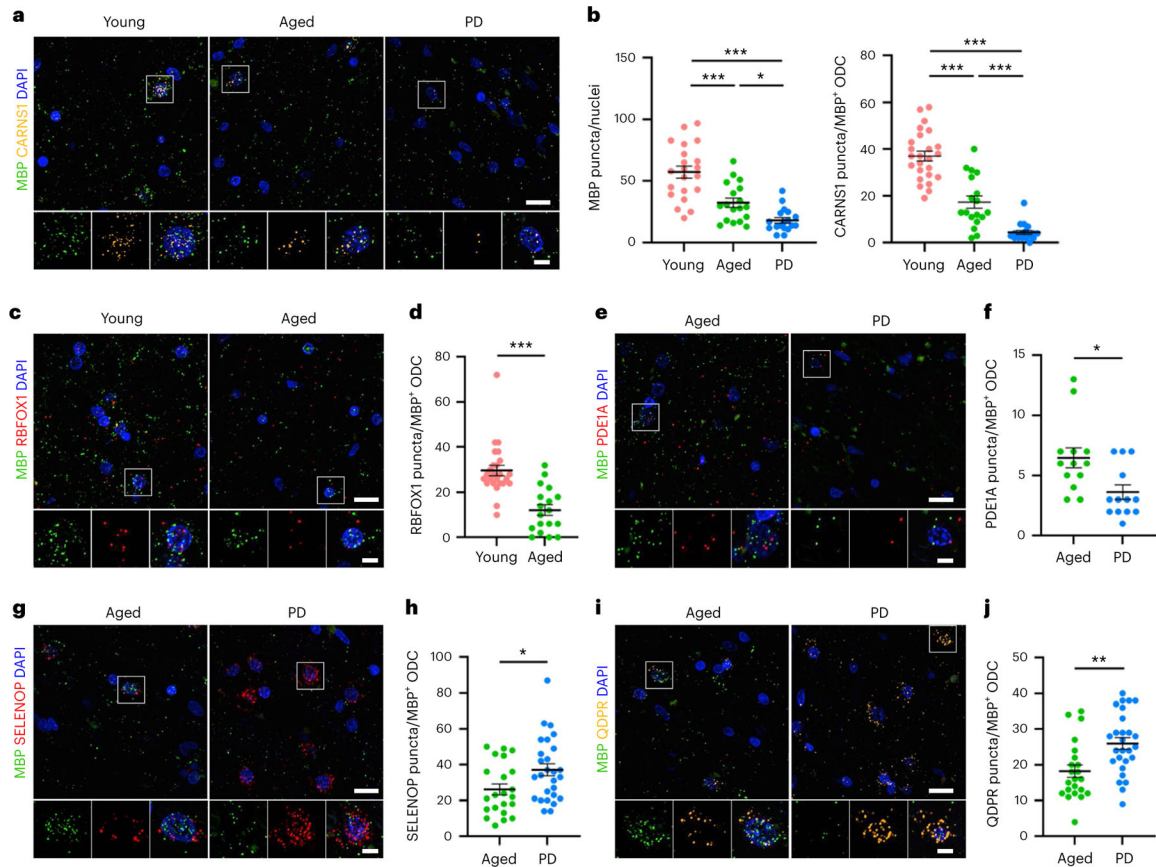
Fig. 7 |. RNA-FISH of human midbrain samples.
Confocal imaging (×40) of RNA-FISH for FFPE human midbrain sections for indicated targets (scale bar, 15 μm). Inset (×100) (scale bar, 5 μm). **a**, Representative fluorescence images of CARNS1 and MBP showing reduced expression over aging/PD. **b**, Quantification of MBP and CARNS1 puncta within a 100×100-pixel square around MBP + ODC nuclei (one-way ANOVA with Tukey's post hoc test; MBP, P for Y/A $= 8.58 \times 10^{-5}$, P for Y/P $= 9.34 \times 10^{-9}$ and P for A/P $= 0.0387$ (n: Y $= 21$, A $= 18$ and PD $= 18$); CARNS1, P for Y/A $= 1.23 \times 10^{-8}$, P for Y/P $= 1.49 \times 10^{-11}$ and P for A/P $= 2.78 \times 10^{-4}$ (n: Y $= 26$, A $= 18$, PD $= 18$)). **c**, Representative fluorescence images of RBFOX1 showing a reduction in aged control compared to young samples. **d**, Quantification of RBFOX1 puncta in MBP+ nuclei (two-sided Student's t-test, $P = 0.0001$, n: Y $= 25$ and A $= 18$). **e**, Representative fluorescence images of PDE1A showing a decrease in PD compared to aged control. **f**, Quantification of PDE1A puncta in MBP+ nuclei (two-sided Student's t-test, $P = 0.0106$, n: A $= 13$ and PD $= 13$). **g,i**, Representative fluorescence images of SELENOP (**g**) and QDPR (**i**) showing increases in PD compared to aged control. **h,j**, Quantification of SELENOP (**h**) and QDPR (**j**) puncta in MBP+ nuclei (two-sided Student's t-test. SELENOP, $P = 0.0231$ n: A $= 22$ and PD $= 27$; QDPR, $P = 0.0021$, n: A $= 22$ and PD $= 27$). For each target, we selected 3–5 fields from three unique donors from each cohort. P values are represented as *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$. Error bars represent s.e.m. A, aged; Y, young.