Article

The spatial landscape of glial pathology and T cell response in Parkinson's disease substantia nigra

Maxwell Ma[1], Fahad Paryani[1,13], Kelly Jakubiak[1,2,13], Shengnan Xia[1], Susumu Antoku[1], Adithya Kannan[3], Jaeseung Lee[1], Nacoya Madden[1], Shailesh Senthil Kumar[1], Juncheng Li[1,2], David Chen[4,5], Gunnar Hargus[1,6], Aayushi Mahajan[3], Xena Flowers[1,7], Ashley S. Harms[2,8], David Sulzer[2,6,9,10,11], James E. Goldman[1,2,6], Peter A. Sims[4,5,12] & Osama Al-Dalahmah[1,2,6] ✉

Parkinson's Disease (PD) is an incurable neurodegenerative disease that causes movement disorders. Neurons in PD aggregate α-synuclein and are depleted from the substantia nigra (SN), which is a movement control hub. The presence of α-synuclein-reactive T cells in PD patient blood suggests a role for adaptive immunity in the pathogenesis of PD. However, the characteristics of this response within the brain are not well understood. Here, we employed single-nucleus RNAseq, spatial transcriptomics, and T cell receptor (TCR) sequencing to analyze T cell and glial cell states in post-mortem PD brain tissue. CD8 + T cells were enriched in the PD SN and characterized by clonal expansion and TCR sequences with homology to those reactive to α-synuclein. Furthermore, PD T cells were spatially correlated with CD44+ astrocytes, which increased in the PD SN. Silencing CD44 in cultured astrocytes attenuated neuroinflammatory signatures, suggesting a potential therapeutic target. These findings provide insight into the neurodegenerative niche underlying T cell-mediated neuroinflammation in PD.

Parkinson's disease (PD) is a common neurodegenerative disease, with an incidence exceeded only by Alzheimer's disease (AD)[1]. PD neuropathology is characterized by the aggregates of alpha-synuclein in neurons, known as Lewy bodies and Lewy neurites[2], that accompany loss of dopaminergic neurons in the substantia nigra (SN)[3]. While current treatments alleviate PD symptoms[4], they do not slow PD progression, and a better understanding of the

disease pathophysiology is needed to identify therapeutic strategies.

Neuropathological studies have previously identified autoimmune features associated with PD, including an increase in T cell populations in the SN of PD patients[5]. Transcriptomic studies have documented gene expression changes in neurons and glia in PD[6–10]; however, T cells have been challenging to study in the human PD brain

[1]Department of Pathology & Cell Biology, Columbia University Irving Medical Center, New York, NY, USA. [2]Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD, USA. [3]Department of Neurological Surgery, Columbia University Irving Medical Center, New York, NY, USA. [4]Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA. [5]Sulzberger Columbia Genome Center, Columbia University Irving Medical Center, New York, NY, USA. [6]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University Irving Medical Center, New York, NY, USA. [7]New York Brain Bank, New York, NY, USA. [8]Department of Neurology, University of Alabama at Birmingham, Birmingham, AL, USA. [9]Department of Psychiatry, Columbia University Irving Medical Center, New York, NY, USA. [10]Department of Pharmacology, Columbia University Irving Medical Center, New York, NY, USA. [11]Division of Molecular Therapeutics, New York State Psychiatric Institute, New York, NY, USA. [12]Department of Biochemistry and Molecular Biophysics, Columbia University Irving Medical Center, New York, NY, USA. [13]These authors contributed equally: Fahad Paryani, Kelly Jakubiak. ✉e-mail: oa2298@cumc.columbia.edu

due to the limited detection of lymphocytes. T cells in the peripheral blood specifically recognize and proliferate in response to an α-synuclein antigen challenge in PD patients[11–13], and the association between neurodegeneration and microglial activation is well-established in other neurodegenerative diseases such as AD[14]. Still, little is known about these phenomena in the PD brain. The brain microenvironment in the PD SN is considered pro-inflammatory[15], and pro-inflammatory microglia may contribute to the pathogenesis and neuronal death in PD[16]. It has also been suggested that microglia are activated in PD by exosomes secreted from neurons with α-synuclein aggregates[17] and by neuromelanin following dopaminergic neuron cell death[18]. Astrocytes have also been shown to adopt abnormal phenotypes in PD neuropathology that could be associated with antigen presentation pathways[19,20]. Thus, the interaction between brain microenvironment cells and cells of the immune system is worth further investigation.

In animal models, mice that overexpress α-synuclein exhibit dopaminergic neurodegeneration following a bout of enteric infection[21], and this is associated with a substantial entry of peripheral T cells into the brain[22–24]. The presentation of mitochondrial antigens has also been implicated in adaptive immunity in animal models of PD[25,26], and pharmacological strategies blocking T cell entry into the gut or brain are protective against dopaminergic[23] and enteric neuronal loss[21]. Finally, T cells have been shown to adopt reactive phenotypes in PD in the cerebrospinal fluid and peripheral blood and contribute to neurodegeneration alongside microglia[13,27–29].

Together, the clinical and basic data point to an important role for infiltrating T cells in the brain during PD pathogenesis. However, previous studies have mainly focused on the characterization of peripheral T cells in the blood[30–35] and cerebrospinal fluid[13,36], leaving the central role of T cells in the human SN in PD unknown. Additionally, many studies characterizing T cells of the PD brain rely on immunohistochemistry (IHC) and/or murine data[29,37–39], so questions about transcriptional profiles of T cells in the human PD brain remain unanswered. As such, there has also been little effort to compare peripheral and parenchymal CNS T cells in PD.

To address the abovementioned gaps in PD research, we created a resource for T cell and glial pathology in the human postmortem brain. Using this resource, we characterized the phenotypes of the adaptive immune response in the human PD brain, specifically from the cingulate cortex and SN. We have used multiple cutting-edge technologies paired with advanced computational techniques, including molecular analysis of one of the highest numbers of PD brain T cells that have been reported in previous study cohorts[6–8,10,40–42]. We then defined the spatial relationships between T cells and other cells in the brain microenvironment, mainly focusing on microglia and astrocytes. Astrocytes play important roles in neurodegeneration in PD as well as other neurodegenerative diseases[43], and we have previously described neuroprotective and disease-associated astrocyte states in HD[44]. In this work, we define spatial and regional differences in astrocyte states in PD.

In this study, we analyzed human postmortem cingulate cortex and substantia nigra specimens from PD and control brain donors to elucidate the roles of T cells and glia in PD pathology. We used single-nucleus RNA sequencing, TCR sequencing, and histopathology to characterize T cell phenotypes and clonality. Moreover, we characterized microglial and astrocytic phenotypic changes in both the cingulate and substantia nigra. Using spatial transcriptomics and multiplex immunohistochemistry studies, we confirmed changes in spatial colocalization between T cells, disease-associated astrocytes, and microglia in PD. Finally, we investigated the impact of silencing a key disease-associated astrocyte state protein, CD44, on gene expression and signaling pathways. Ultimately, our studies provide a detailed understanding of the neurodegenerative niche in PD, implicating T cells and astrocytes as potential therapeutic targets.

Results

T cell receptor sequencing reveals clonal expansion in the substantia nigra of Parkinson's disease subjects

Previous studies in postmortem tissue have indicated that T cells are increased in the parenchyma of the substantia nigra (SN) of subjects with PD[5,29]. Consistent with these past findings, we observed that CD8 + T cells were significantly increased in the PD SN in our cohort of PD and controls (Fig. 1a, b). Also, the number of T cells has been reported to increase in the frontal cortex of Diffuse Lewy Body Disease (DLBD) or PD with dementia (PDD)[37]. Since the cingulate cortex is typically also severely involved in DLBD, we quantified CD3 + T cells in the cingulate cortex parenchyma and the subcortical white matter (Fig. S1A, C). The results did not reveal significant differences in the density of CD3 + T cells in either region in DLBD (Fig. S1B, D), but the densities of CD3+ cells were higher overall in the white matter compared to the cortical parenchyma (Fig. S1E), where they were concentrated mainly around penetrating vessels. We did not count T cells that were located within vessel lumina.

While immunohistochemistry allows us to localize and quantify T cells, it does not tell us whether these cells are clonally expanded or allow us to investigate T cell heterogeneity. To investigate in more detail the significance of the T cell population in PD and to examine whether T cells in the PD/DLBD brain are clonally expanded rather than non-specifically polyclonal, we compared the T cell receptor (TCR) α-chain repertoires in PD/DLBD and control samples using TCR sequencing in 44 brain donors (cingulate:$n = 11$ PD and 6 controls, SN: $n = 13$ PD and 15 controls - see Supplementary Data 1 for demographic data). We chose the cingulate cortex in PDD/DLBD because it exhibits neurodegeneration but does not display increased T cell infiltration (Fig. S1). Our cohort did not have paired SN and cingulate samples, and some controls had other common neuropathological changes that did not include synucleinopathy. For simplicity, we refer to cortical PDD/DLBD as PD henceforward. As a quality control step, we determined that all libraries displayed adequate saturation (Fig. S2A and Supplementary Data 2).

As a first survey of TCR repertoires in the SN and the cingulate cortex, we compared the number of unique TCR α-chain clonotypes in PD and control SN and cingulate (Fig. 1c) using a linear model that included age, sex, and common Alzheimer's type neuropathologic changes (Braak stage) - see methods. There was a significant increase in the number of unique clonotypes in PD compared to controls in the SN, but not in the cingulate (Fig. 1d). This data is consistent with our quantitative immunohistochemistry identifying more T cells in the SN (Fig. 1a, b) and not the cingulate cortex (Fig S1). To gain more insight into the characteristics of TCRs in PD, we performed repertoire analysis and specifically examined Hill's diversity, Hill's evenness, and Shannon's entropy indices. These metrics allow us to explore repertoire diversity and infer clonal expansion (Fig. 1d). The results showed that Hill's diversity of α-chain repertoires was increased in PD in the SN, but the result was not evident in the cingulate cortex. Shannon's entropy showed a similar trend (Fig. 1d).

Based on the above results, we focused on the SN and performed additional TCR repertoire analysis to include the β-chain. To achieve this, we used a unique molecular identifier (UMI)-based method (Repseq+ – see methods), which allows us to simultaneously sequence both the α-chain and the β-chain from the same RNA sample and count RNA molecules using UMIs. Our cohort included 12 PD and 13 control SN samples– see Fig. S2B and Supplementary Data 2 for saturation metrics. As expected, the number of clonotypes discovered using this method was lower than read-based methods (Fig. 1e). Comparing the numbers of unique α-chain and β-chain clonotypes did not show significant differences between PD and control SN (Fig. 1e). Nonetheless, repertoire analysis of the α-chain and β-chain clonotypes showed that in the β-chain clonotypes of PD, there was significantly increased diversity (Hill's diversity) and reduced evenness (Fig. 1f), and the
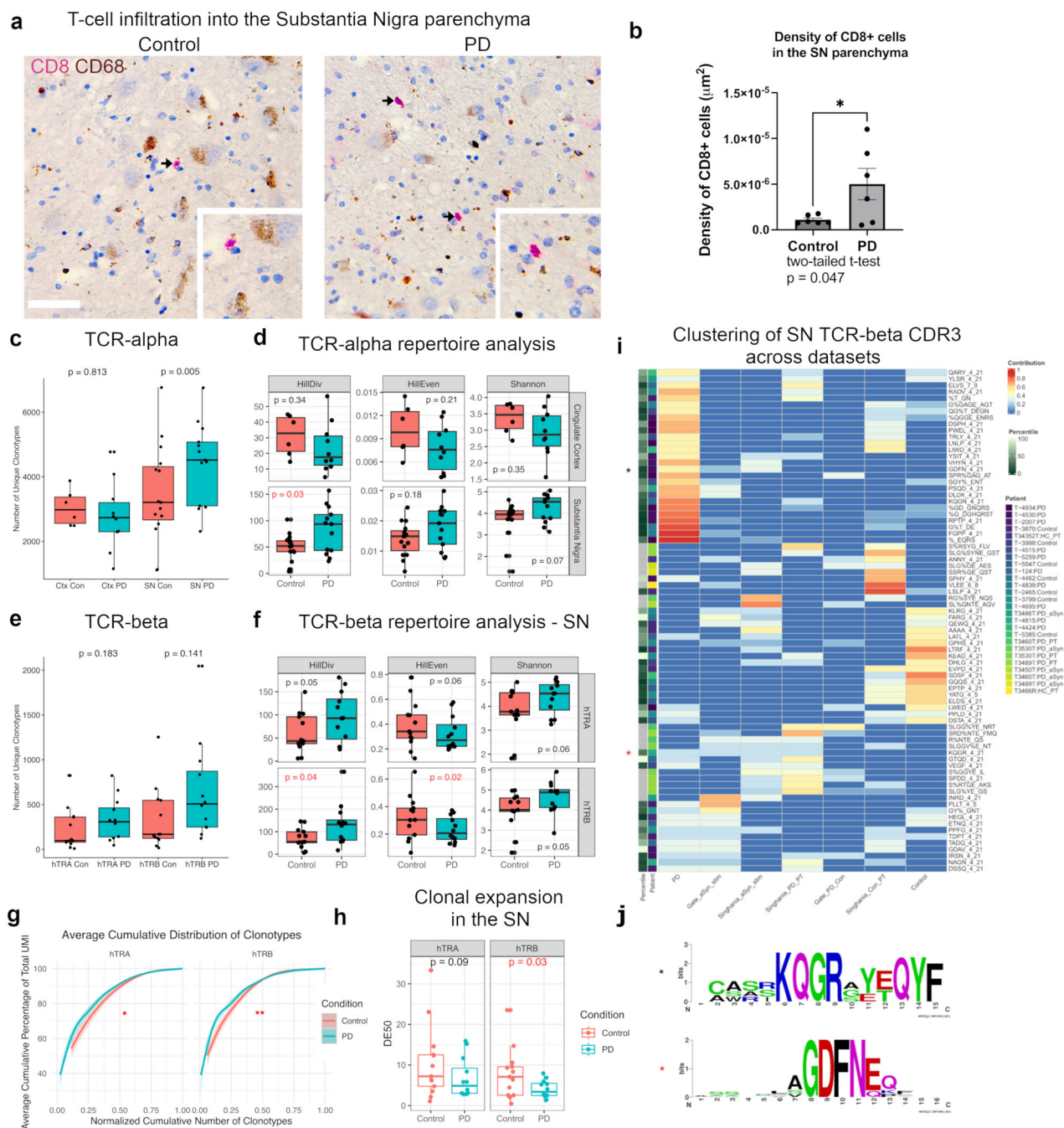
Fig. 1 | T cell receptor sequencing identifies clonal expansion in T cells of the substantia nigra in Parkinson's disease. a Representative immunohistochemistry for CD8 (red) and CD68 (brown) in the substantia nigra (SN) of control (con) and PD donors. Black arrows indicate CD8+ cells, enlarged in the insets. Scale bar: 50 μm. **b** Quantification of CD8 + T cell density in the SN parenchyma. $n = 6$/group. Two-tailed t-test p-value is indicated. Data is shown as mean +/- SEM. **c** Boxplots depicting the number of unique alpha-chain clonotypes in the cingulate cortex (ctx) and SN. Ctx-con $n = 6$, ctx-PD $n = 11$, SN-con $n = 14$, SN-PD $n = 13$. Two-tailed regression p-values are indicated. **d** TCR repertoire diversity metric of data in c. Two-tailed regression p-values are indicated. **e** Boxplots depicting the number of SN unique alpha-chain and beta-chain clonotypes. Con-n $= 13$, PD-n $= 12$. One-tailed regression p-values are indicated. **f** Diversity metrics for SN alpha-chains (hTRA) and beta-chains (hTRB). N's as in e. one-tailed regression p-values are indicated. **g** Line plots representing cumulative library contribution (y-axis) vs. the proportion

of clonotypes/sample (x-axis). P values of 2.12e-6 and 1.11e-6 were determined with a two-sided Kolmogorov-Smirnov test, d values = 0.02 and 0.05 for alpha-chains and beta-chains, respectively. **h** Boxplots depicting DE_{50} values for alpha-chains and beta-chain TCR repertoires. Two-tailed regression p-values and n's are indicated. **i** Heatmap showing the contribution of each condition (columns) to TCR-beta Gliph clusters. Two included CDR3s from both PD brain and PD patient blood T cells challenged with α-Syn peptides in two external datasets (asterisks). The percentile rank of the UMI count of the highest contributing PD brain donor sample clonotype to each Gliph cluster (rows) is indicated on the side-bar. Lower percentiles correspond to higher normalized clonotype counts. **j** Logoplots showing the consensus CDR3 sequences of the clusters indicated by the asterisks in i. For boxplots in **c**−**f** and **h**, the line indicates the median; the bounds represent the 1st and 3rd quartiles. The whiskers extend from the edges of the box to the smallest and largest values within 1.5 times the interquartile range.

α-chain clonotypes showed similar trends. These results indicate that while PD TCR repertoires are more diverse, which could be explained by increased T cell infiltration, PD TCR repertoires were not evenly distributed among clones, thus supporting the notion of T cell clonal expansion in the PD SN.

The above results are consistent with our immunohistochemical results, which showed increased numbers of T cells in the PD SN. They also provide evidence for clonal expansion in PD brain, given the reduced repertoire evenness. To more directly measure clonal expansion, we calculated the cumulative proportions of clonotypes in each donor that accounted for increasing proportions of the repertoire (Fig. S2C). Aggregating these cumulative clonotype-repertoire distributions by condition showed a significant shift of the PD cumulative distributions to the left, in both the α-chain and β-chain clonotypes (Fig. 1g). These results indicate that fewer clonotypes in PD account for larger fractions of the total UMIs per repertoire, providing independent confirmation of T cell clonal expansion in the PD brain. To confirm this conclusion, we measured the fraction of clonotypes that accounted for 50% of the total reads across all clonotypes (Diversity-equality 50, or DE50[45]). We found that for β-chain but not the α-chain clonotypes, the DE50 was significantly lower in PD vs control, which further indicates the presence of clonal expansion (Fig. 1h). The TCRseq count matrices are provided in Supplementary Data 3. Altogether, our findings show that in the PD SN, there is an increased infiltration of clonally expanded T cells.

T cells that recognize alpha-synuclein in the periphery have a broad diversity of TCRs[11]. However, as blood samples were not acquired from the subjects prior to death, we could not directly determine whether peripheral TCR sequences had been expanded in the brain. Therefore, we compared our repertoires to those described in the literature from the peripheral blood of PD donors. To do so, we employed GLIPH2[46] analysis, which provides insights into the potential antigen-specificity of T cells in the SN of PD. We compared PD-derived CDR3 amino acid sequences (TCRb) to those described by Singhania et al.[11] and Gate et al.[13], where peripheral blood-derived T cells were challenged by α-synuclein peptides, control (pertussis toxin Singhania et al.), or remained unchallenged (Gate et al.[13]), and clustered the CDR3 amino acid sequences using GLIPH2. We identified 556 clusters characterized by public, across-donor similarities in the CDR3 structure or local, within-donor sequence similarities (Fig. S3A). Of all clusters, 78 had members that spanned multiple treatment groups and conditions (Fig. 1i). Interestingly, there were several clusters whose members were solely from PD patient brains and peripheral blood. Of these, there were two clusters that only included CDR3s from PD and from PD patient blood T cells challenged with α-synuclein peptides across two external datasets (Fig. 1I asterisks and Fig. 1J). Several PD SN clonotypes that contributed to public motif clusters, including those marked with asterisks, had high UMI counts. These clonotypes also ranked highly in their relative contribution to the overall clonotypes within their respective samples. The presence of these clones suggests that T cells in the PD SN may recognize peptides related to α-synuclein. Furthermore, these SN-resident clones are possibly clonally expanded.

We next calculated the probability of finding shared CDR3 motifs using Fisher's exact test between PD SN, control SN, α-synuclein reactive peripheral blood T cells, and control peripheral blood T cells (Fig. S3B). The results showed that the overlap among groups is greater than one would expect by chance; however, there was no association between condition (PD vs. control) and α-synuclein reactivity in the blood (Chi-squared test: $X^2 = 0.077711$, p-value = 0.7804). Thus, it is not clear whether the overlap we find between the α-synuclein reactive peripheral blood T cells and the PD SN is related to disease status. Further experiments are needed to assess the functional significance of the TCRs shared between the PD SN and α-synuclein reactive peripheral blood T cells.
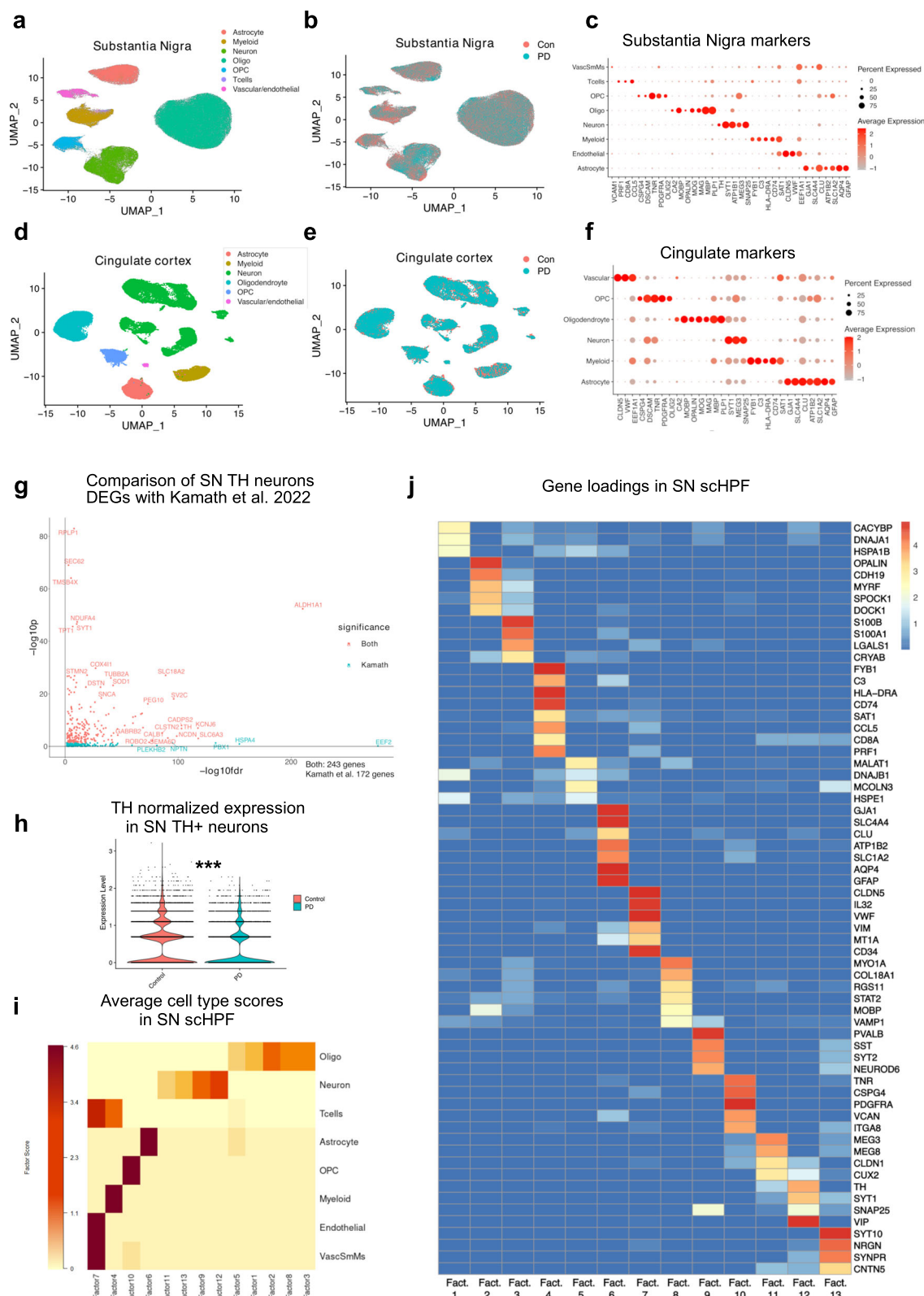
Single-nucleus RNA sequencing reveals cell-type specific DEGs in PD

To determine how PD affects the transcriptional profile of cells in the SN and cingulate cortex, we generated single-nucleus RNAseq (snRNAseq) datasets from the cingulate cortex ($n = 10$ PD and $n = 8$ controls) and the SN ($n = 13$ PD and $n = 15$ controls - Supplementary Data 1) to examine T cell gene signatures at the single cell level. The SN dataset includes 207,859 nuclei, with 96,244 derived from PD subjects, including 831 SN T cells, of which 535 were from PD donors. The cingulate dataset comprises 57,425 nuclei, 32,442 from PD subjects, of which no T cells were retrieved. We projected these nuclei in UMAP space, and, using cluster analysis (see methods), assigned cell types/lineages and disease status in the SN (Fig. 2a, b) and cingulate (Fig. 2d, e). This data was also projected in UMAP space and color-coded by donor and sex in the SN (Fig. S4A-B) and cingulate (Fig. S4C-D). The expression of select canonical marker genes per lineage for the SN and cingulate is shown in Fig. 2c, f, respectively. The cluster markers are reported in Supplementary Data 4-5 for the SN and cingulate. Metrics on cell numbers and the proportion of each cell type per sample in both regions are provided in Supplementary Data 1 and Fig. S4E–H. Consistent with the neuropathologic literature, differential abundance analysis demonstrated that dopaminergic neurons were depleted in the SN based on compositional analysis (Fig. S4I). We asked if T cells were increased in the PD SN, as we have seen in the immunohistochemistry results presented in Fig. 1a, b. Differential abundance analysis also showed that T cells were enriched in the PD SN (Fig. S4I), confirming the results from histopathologic analysis.

We determined the differentially expressed genes (DEGs) between PD and control in each of the broad lineages in the SN and cingulate (Supplementary Data 4-5). We performed this analysis while correcting for relevant biological variables – age, sex, and Alzheimer's type changes – as outlined in the methods section. We found that the largest number of altered DEGs in PD were in neurons and oligodendrocytes in both SN and the cingulate (Fig. S4J and Fig. S4K). There were also DEGs in astrocytes and myeloid cells in both regions, as well as in T cells in the SN. Notably, the number of DEGs does not necessarily reflect only how perturbed a cell type is in disease but can be affected by the abundance of a cell type and sampling, among other factors[47].

Prior studies have detailed neuronal and glial pathology in PD at the single-nucleus level[6,7]. With the goal of identifying glial and T cell pathology in PD, we first confirmed the detection of previously described transcriptomic changes in the PD SN. Thus, we subclustered SN and cingulate neuronal nuclei and identified populations similar to those we previously described in the SN of control and frontotemporal dementia with Parkinsonism donors[48] (Fig. S5A, B). Cingulate cortex neuronal clusters were also similar to the ones we described in control and Huntington disease cingulate cortices[44] (Fig. S5C, D). The distribution of neuronal subtypes was relatively consistent across donors (Fig. S5E, F). Examining the overlap between DEGs revealed that most were region-specific, with more DEGs shared between cortical neurons, including CUX2 and NRG1 glutamatergic neurons, which exhibited the largest numbers of DEGs (Fig. S5G, H).

We then focused on TH+ dopaminergic neurons, which have been investigated previously[7], and compared our DEGs in TH+ dopaminergic neurons to DEGs described by Kamath et al.[7]. Of note, the reported DEGs in Kamath et al.[7] have coefficients quantifying the effect of disease on gene expression with associated P values. These coefficients are not directly comparable to log-fold changes. Thus, we compared the adjusted P values from the genes dysregulated in dopaminergic neurons reported in Kamath et al. to the adjusted P values associated with our DEGs. This is a measure of concordance of confidence in the dysregulation of genes rather than a comparison of effect size. All of the DEGs we discovered were significantly dysregulated in Kamath et al.[7], including *ALDH1A1*, *TH*, *KCNJ6*, and *SNCA* (Fig. 2g). This result is

also consistent with published datasets[8]. As expected, *TH* was decreased in PD dopaminergic neurons (Fig. 2h). These data show the conservation of DEGs in dopaminergic neurons across different cohorts.

We next compared the KEGG pathways enriched in cortical CUX2 and NRG1 glutamatergic neurons, nigral dopaminergic neurons, and SEMA3E_TSHZ2 glutamatergic neurons, which we showed to be depleted in HD[44]. As expected, there was overlap in the enriched

pathways related to neurodegeneration, PD, and oxidative phosphorylation. However, several amino-acid metabolic pathways, such as tyrosine and beta-alanine metabolism, were only enriched in dopaminergic nigral neurons. In contrast, ubiquitin-mediated proteolysis was enriched in cortical glutamatergic neurons only (Fig. S5I). These findings point to regional and cell-type specific pathologies in PD.

We analyzed our dataset using a recently developed approach called single-cell Hierarchical Poisson Factorization[49] (scHPF; see

Fig. 2 | Single-nucleus RNA sequencing reports differences in gene expression patterns of PD lineages and identifies a PD T cell signature. a Uniform manifold approximation and projection (UMAP) graphs showing nuclei from the substantia nigra color-coded by cell type/lineage. **b** Same as a but color-coded by condition. **c** Dot plot of select gene (x-axis) marker expression in major lineages in the substantia nigra (y-axis). Size indicates percentage expression, and color indicates normalized expression levels. **d** Same as a but for the cingulate cortex. **e** Same as b but for the cingulate cortex. **f** same as c but for the cingulate cortex. **g** Comparison between the differentially expressed genes (DEGs) in PD vs. control dopaminergic neurons in our dataset and those in Kamath et al. [7]. The scatterplot shows DEGs with

the Kamath et al. reported -log10 adjusted p-value on the x-axis and our -log10 adjusted p-value on the y-axis. The color of each DEG indicates if it is significant in both datasets vs only in Kamath et al. All of the DEGs we discovered were reported in Kamath et al. P values are two-sided. **h** Violin plot showing the gene expression of *TH* in control (orange) and PD (blue) in dopaminergic neurons (logFC PD vs. control: −0.374, two-sided *p*-value 1.66e-7, as determined via differential gene analysis). **i** Heatmap showing the scores of single cell hierarchical Poisson factorization (scHPF) gene factors (columns) projected on lineages (rows). **j** Heatmap showing the gene loadings for select cell select genes in each of the single-cell Hierarchical Poisson Factorization (scHPF) factors.

methods). This method derives factors, or gene sets, that capture the sources of gene expression variability in the dataset, which could be lineage-related, disease-related, or related to other factors. When we applied scHPF to the SN snRNAseq dataset, we retrieved factors that corresponded to cell types (Fig. 2i and Supplementary Data 6). Example gene scores in each of the scHPF factors are shown in Fig. 2j, where cell-type specific genes for TH+ neurons, astrocytes, T cells, and other cell types are shown, underscoring the power and validity of the technique. Given that scHPF can identify cellular lineages and subtypes without the need for clustering, when used on a given lineage, it has the potential to discover cellular states or even disease-associated states.

Single-nucleus RNA sequencing defines a T cell PD disease signature and CD8+ resident memory phenotype

We next turned our attention to immune cells, in particular, T cells. Projecting immune cells in their own UMAP space showed that T cells are distinct from myeloid cells in our dataset (Fig. 3a). Consistent with previous reports[38,50], including our own (Fig. 1a, b), most T cells were CD8+ (Fig. 3b). We then used scHPF to extract PD-related factors by performing scHPF on T cells. T cell factor gene loadings, patient contribution to factors, and cell scores are provided in Supplementary Data 6. The factor scores of representative immune-related gene members of the T cell factors are shown in Fig. 3c.

Interestingly, the results show that several TRM and memory genes had high scores in T cell Factor 5, including *IL7R*[51], *EOMES*[52], *TBX21*[53], and *ZNF683* (Hobit)[54] (Fig. 3c). This factor also includes other genes associated with activation, such as *PRDM1 (BLIMP-1)*[55], IL-2 signaling (*IL2RB, SOS1, CD2, JAK3, LCK*), which is associated with activation[56], and memory T cell survival (*BCL2*)[57]. There are also genes involved in interleukin signaling (*IL7R, IL17RA, IL16*) and trafficking (*CCR5* and *CXCR6*)[58] (Fig. 3c and Supplementary Data 6). Together, this signature associates Factor 5 with a TRM state and T cell activation.

T cell Factor 1 included genes associated with effector T cell function with expression of granzyme B (*GZMB*), T cell activation such as *IL32*[59], *HLA-DRA* and *HLA-DRB1*, *CXCR3*[60], and proliferation (*CD37*)[61] (Fig. 3c). It is noteworthy that other memory and activation genes like *CD38* and *CD44*[62] exhibited higher scores in T cell Factor 4, suggesting that there are different programs associated with T cell activation. When measured on the population level, T cell Factor 1 was higher in control T cells (Wilcoxon rank sum test, W = 40779, P value < 2.2e-16), while PD T cells had higher scores in T cell Factor 5 (Wilcoxon rank sum test, W = 102191, P value = 3.366e-06), nominating the latter as a "disease factor" (Fig. 3d). We demonstrate several T cell Factor 5 genes in a dot plot (Fig. 3e). Interestingly, differential gene expression analysis comparing PD to control T cells at the donor level (pseudobulk) showed that T cell resident memory genes *FOS*, *RGS1*, *RGS2*, and *GADD45B*, were nominally significantly increased in PD (Fig. S3C). We then measured the donor contributions to each scHPF factor, adding another measure of correlation to disease status at the patient level. PD donors had significantly lower contributions to T cell Factor 1, which is associated with effector function, while there was variability in other
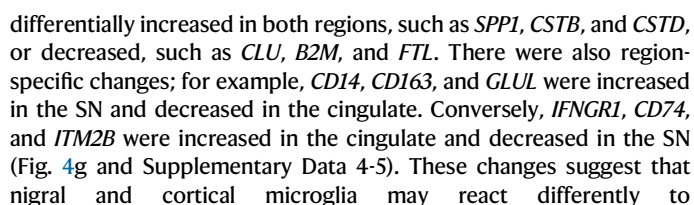
factors, including T cell Factor 5 (Fig. 3f). Additional covariate analysis showed that T cell factor 3 was lower in males, and T cell factors 1, 3 and 4 were associated with increased age. None of these covariates were significant in T cell factor 5 (Fig. S3D).

To further examine Factor 5 genes, we used preranked GSEA analysis to measure the enrichment of a T cell resident memory (TRM) gene set[50,63–70] (Supplementary Data 7; Fig. 3g) in T cell Factor 5. We found that the TRM gene set was, in fact, enriched in the T cell Factor 5, and both TRM and Factor 5 genes were enriched in T cell genes, ranked by fold change compared to controls (Fig. 3g), suggesting that PD T cells are more TRM-like.

To validate the memory phenotype of CD8 + T cells in PD SN, we used an antibody against CD103, which is expressed in TRMs[71]. The density of CD103+ cells was higher in PD compared to controls in the SN (Fig. 3h–j), and CD103 colocalized with CD3 (Fig. 3l), supporting the transcriptomic results indicating that T cells in the PD SN adopt a memory resident phenotype. We performed additional validation studies and found that both CD69 and PD1—the latter being a marker of exhaustion—were expressed in the PD SN (Fig. S6A, B). Taken together, these results indicate that PD SN T cells demonstrate a more prominent memory phenotype, which we interpret as being more antigen-experienced, although we note that it is possible that some of these T cells exhibit exhaustion.

Myeloid cells in the PD SN show increased activation and enrichment of neuroinflammatory pathways

Given the role of antigen-presenting cells or glial cells in T cell interactions, we analyzed subclustered immune cells to discover different states and phenotypes in isolation of T cells and other brain cells in the SN and cingulate cortex (Fig. 4a–f). Notably, snRNAseq has limitations in recapitulating myeloid states compared to scRNAseq[72]. The heterogeneity of microglial states can be remarkable in neurodegeneration, and these may represent a continuum of states[73]. Here, we opted for a broad clustering system based on select marker genes. The results revealed three myeloid states: quiescent/homeostatic microglia, activated microglia, and monocyte-like myeloid cells better known as border-associated macrophages (BAMs) that were distributed between PD and control nuclei in the cingulate (Fig. 4a, b) and the SN (Fig. 4d–e). There was variability in the capture rate for the different classes (Fig. S7A-B). Select markers of each nigral and cortical subcluster are shown in Figs. 4c, f, respectively (Supplementary Data 4-5), including *CD74*, *C3*, and *SPP1* for disease-associated or activated microglia and *MRC1* for BAMs. We note that the cortical BAM cluster is likely heterogeneous, and other myeloid states may be included in this cluster. The quiescent clusters exhibited lower expression levels of a number of activated microglial genes (Supplementary Data 4-5), and there are several differences between activated microglia in the cingulate and the SN. For example, *CD163* was expressed in activated microglia in the cortex but not in the SN, where it was expressed in BAMs. We focused on activated microglia and compared gene expression changes between PD and control. Examining the log-fold changes for the DEGs in the SN and the cingulate showed several notable patterns (Fig. 4G). Several genes were concordantly

a Myeloid / T-cell UMAP

b General T-cell markers — CD2, CD3E, CD8A

c T-cell scHPF gene loadings

d T-cell scHPF cell loadings

e dot plot Control vs PD

f T-cell scHPF sample contribution

Factor_1 Wilcoxon, p = 0.036; Factor_2 Wilcoxon, p = 0.57; Factor_3 Wilcoxon, p = 0.6; Factor_4 Wilcoxon, p = 0.94; Factor_5 Wilcoxon, p = 0.92

g

CD8 TRM geneset in T-cell Factor 5 — NES = 1.84, p value 1.4e-06

CD8 TRM geneset min PD T-cells genes — NES = 2.01, p value = 5.9e-08

T-cell Factor 5 in PD T-cell genes — NES = 2.42, p value = 1.4e-28

h Control / PD — CD103

j Density of CD103+ cells in the SN parenchyma; One-tailed t-test p = 0.043

i CD103 / CD3 / DAPI / Merge

differentially increased in both regions, such as *SPP1*, *CSTB*, and *CSTD*, or decreased, such as *CLU*, *B2M*, and *FTL*. There were also region-specific changes; for example, *CD14*, *CD163*, and *GLUL* were increased in the SN and decreased in the cingulate. Conversely, *IFNGR1*, *CD74*, and *ITM2B* were increased in the cingulate and decreased in the SN (Fig. 4g and Supplementary Data 4-5). These changes suggest that nigral and cortical microglia may react differently to

neurodegeneration, which may be a function of the severity of neurodegeneration and regional differences in microglial responses.

To gain a deeper understanding of the gene expression changes in activated microglia in the SN and cingulate, we performed KEGG pathways enrichment analysis (Fig. 4h). Several KEGG pathways were enriched in activated microglia DEGs in the SN and cingulate, including pathways of neurodegeneration and PD. However, the enrichment of

Fig. 3 | Validation of T cell phenotypes in the SN. a Uniform manifold approximation and projection (UMAP) graphs showing substantia nigra myeloid and T cell nuclei. **b** Feature plots showing the normalized expression of select T cell genes in the UMAP space. **c** Loadings (scores) of select T cell-related genes from scHPF performed on SN T cells. **d** Heatmap of average cell score of PD and control nuclei in each T cell scHPF factor. Columns represent factors; rows represent conditions. Blue indicates a low score; red indicates a high enrichment score. **e** Dotplots showing the normalized expression of select Factor 5 genes in PD and control T cells. **f** Boxplots of the donor contribution to each T cell factor. Wilcoxon test, two-tailed, p-values are indicated. Control-n = 12, PD-n = 9. The line indicates the median, and the bounds represent the 1st and 3rd quartile. The whiskers extend from the edges of the box to the smallest and largest values within 1.5 times the interquartile range. **g** Pre-ranked gene set enrichment analysis of the $CD8 + $ T cell resident memory (TRM) gene set in T cell Factor 5 genes ranked by the gene loadings (left), $CD8 + $ TRM gene set in T cell genes ranked by logFC in PD vs control (middle), and T cell Factor genes in T cell genes ranked by logFC in PD vs control (right). Normalized enrichment scores (NES) and adjusted two-sided p-values are indicated. **h** Immunohistochemical stains for CD103 (brown) in the SN. **i** Multiplex immunofluorescence of a PD SN showing colocalization between CD103 (red) and CD3 (green). Nuclei (DAPI) are shown in blue. Scale bar = 25 μm. Representative of colocalization from two donors (biological replicates). **j** Quantification of the density of CD103-positive cells per unit area in the SN. Unpaired one-tailed t-test with n = 4 for control and $n = 5$ for PD donors. P-value is indicated. Data is shown as mean +/- SEM.

pathways involved in cytosolic DNA sensing was significant only in SN-activated microglia Fig. 4h. While the enrichment scores for antigen presentation, Th17 cell differentiation, Th1 and Th2 differentiation, T cell receptor signaling were enriched in both regions, some pathways were only enriched in cingulate cortex activated microglia such as Fc gamma R-mediated phagocytosis, MAPK signaling, and JAK/STAT pathways. Overall, the KEGG pathway enrichment highlighted heightened activation of microglia in PD in a more pronounced way in the SN. To validate microglial activation in the PD SN, we quantified the density of CD68+ microglia in the SN. As expected, CD68+ microglia were more abundant in the PD SN than in controls (Fig. 4i, j). Altogether, our data implicate increased microglial activation in the pathology of PD.

Due to the relative rarity of SN BAMs (207 nuclei in PD and 332 nuclei in control - Supplementary Data 1 and Fig. S7) and the number of covariates we account for in the DEG pipeline (age, sex, and AD neuropathologic change), none of the DEGs survived P value adjustment (Supplementary Data 4). To discover changes in BAMs in PD, we took a more supervised approach and measured the enrichment of known BAM signatures in a murine PD model[24] in ranked BAM genes in PD vs. control nuclei. The results showed that several murine BAM signatures enriched in a PD mouse model were enriched in genes that were higher in PD-derived BAMs (Fig. 4k). These included the markers of proliferative and activated BAMs. These data indicate that BAMs in PD SN resemble those activated by a more acute murine model of PD driven by α-synuclein overexpression – implicating PD BAMs in the pathogenesis of the disease.

Differential regional dysregulation of astrocytes in PD

Astrocytes play several roles in PD[43] and are central to neuroinflammation. We have shown that astrocytes can be distinguished by $CD44$ expression into fibrous-like and protoplasmic states[44]. Using snRNAseq, we discovered 31,508 astrocytic nuclei (SN: $n = 10,118$ PD and $n = 14,714$ control nuclei; cingulate: PD n = 3,772, control n = 2,904) and clustered them into protoplasmic and fibrous-like astrocytes in the cingulate cortex (Fig. 5a, b) and SN (Fig. 5d, e). Protoplasmic astrocytes expressed higher levels of the glutamate transporter $SLC1A2$, the sodium bicarbonate cotransporter $SLC4A4$, and the glutamine synthetase gene $GLUL$, while fibrous-like astrocytes expressed higher levels of $GFAP$, $S100B$, and $CD44$ (Fig. 5c, f).

For downstream analysis, we focused on protoplasmic astrocytes, as they represent the majority of nuclei (Fig. S8A, B). Differential gene expression analysis revealed several patterns of astrocytic dysregulation in PD. As expected, the expression of $GFAP$ was increased in protoplasmic astrocytes in both the cingulate cortex and SN (Fig. 5g). When we quantified the proportion of GFAP-high astrocytes by immunofluorescent staining, we found that more were GFAP-high in the cingulate, while there was a reduction in the proportion of GFAP-high astrocytes in the SN (Fig. S9E, F, G). This is compatible with previous reports[74,75] that found unchanged or reduced GFAP protein levels in the PD SN. Conversely, metallothionein gene expression (e.g., $MT3$ and $MT2A$) showed differences between the two regions, where they
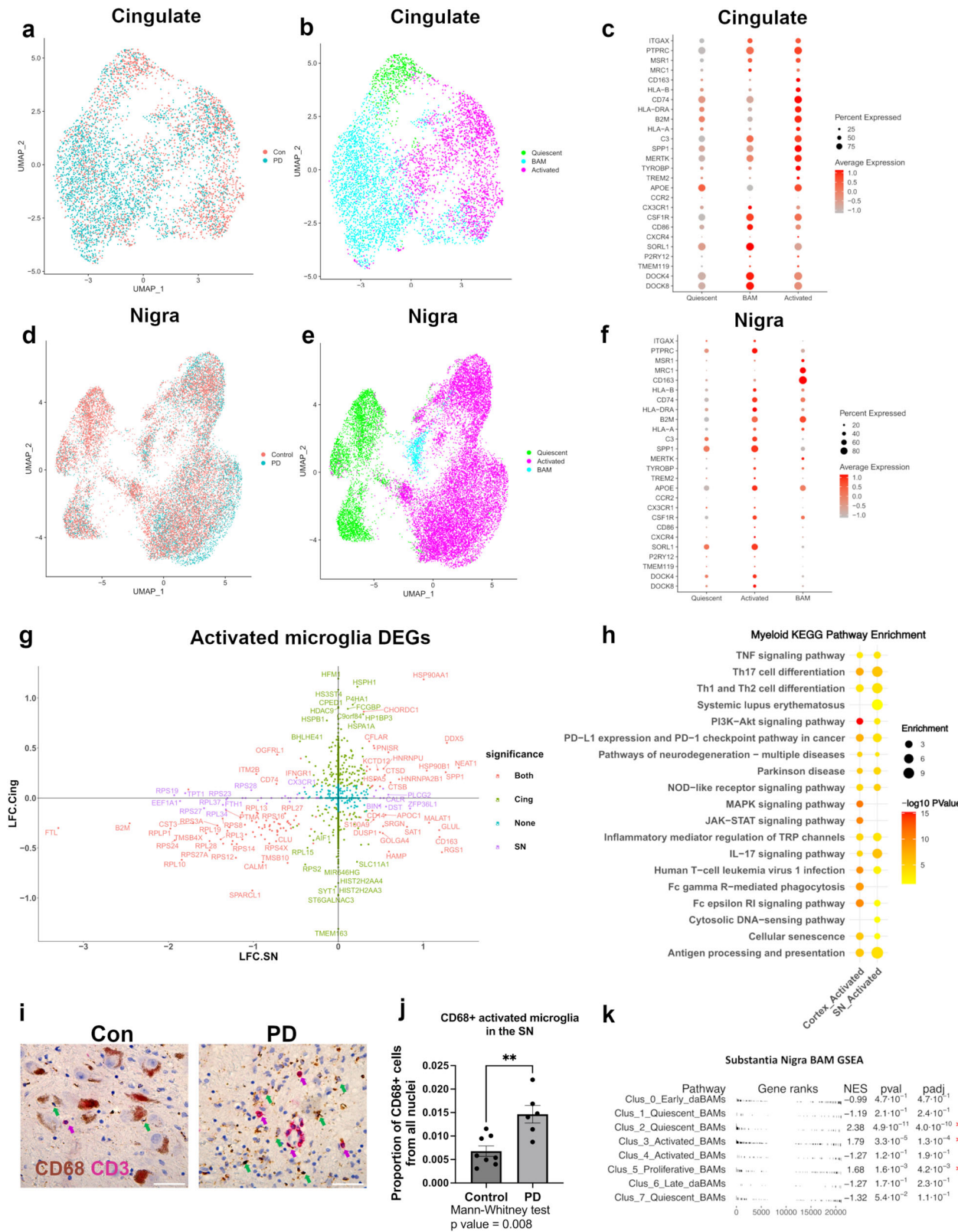
were increased in the protoplasmic astrocytes in the cingulate cortex but decreased in the SN. We confirmed this finding using multiplex immunofluorescence for MT3 (Fig. S9A–D, G). This result is reminiscent of changes seen in Huntington's disease (HD), where we also saw an increase in metallothioneins, including MT3, in the relatively less severely affected cingulate cortex but not the severely degenerated caudate nucleus[44].

We compared our SN astrocytic DEGs to those described in midbrain astrocytes in Smajic et al.[10]. Several of the DEGs we discovered overlap with those described in Smajic et al.[10], including $CD44$, CP, $CHI3L1$, and $NEAT1$ (Fig. S8C). Interestingly, while the metallothioneins $MT2A$ and $MT1G$ were reduced in the SN in our cohort, these genes were increased in the midbrains Smajic et al. analyzed (Fig. S8C). Also, $GFAP$ and $SERPINA3$ were increased only in nigral astrocytes in our dataset (Fig. S8C). This could be because our analysis was that of neuroanatomically dissected SN, while in Smajic et al, sections of the midbrain were analyzed, which included anatomic regions other than the SN.

Next, we compared the DEGs in the cingulate cortex to those described by Zhu et al.[42] in the prefrontal cortex of PD (Fig. S8D). In both datasets, protoplasmic genes like $SLC1A3$ and $ATP1B2$ were significantly reduced. Other protoplasmic genes like $SLC1A2$ and $FGFR3$ were decreased only in the cingulate cortex, which may be due to regional cortical differences, and because astrocytes were not subclustered into protoplasmic and fibrous-like in Zhu et al. Several metallothionein genes, including $MT3$, $MT2A$, and $MT1G$, and chaperone genes involved in protein folding like $CRYAB$, $HSPA1A$, and $HSPA1B$, were increased in both datasets (Fig. S8D). Previously, we have shown that MT3-high astrocytes were neuroprotective in vitro[44]. The upregulation of metallothioneins in frontal and cingulate astrocytes, but not in nigral astrocytes, may be a clue to why neurodegeneration is more severe in the SN than in the cortex; however, future experiments are needed to test this hypothesis.

Interestingly, one of the genes significantly increased in SN protoplasmic astrocytes was $CD44$ (Fig. 5G). We previously reported that CD44 increases in protoplasmic astrocytes in the severely degenerated caudate in HD but not the less severely affected cingulate cortex[44]. Thus, CD44 may represent a biomarker of disease-associated astrocytes, since we also showed it increases in astrocytes in other neurologic insults such as ischemia and seizures[76] and in glioblastoma-associated astrocytes[77]. We validated this finding by performing multiplex immunofluorescence for GFAP and CD44 in PD and control SN sections, where we identified a significant increase in the proportion of CD44 + /GFAP+ astrocytes in the PD SN (Fig. 5H, I). Together, these findings support that in the PD SN, like in the HD caudate and other pathologies, astrocytes adopt a CD44 + , fibrous-like state.

To further understand the general pathways dysregulated in PD astrocytes, we examined KEGG pathway enrichment in the DEGs (both increased and decreased) in the SN and cingulate protoplasmic astrocytes (Fig. 5j; Supplementary Data 4-5). In both brain regions, we found enrichment of pathways related to neurodegeneration, PD disease, and immune activation, including IL-17 and MAPK kinase

signaling in PD astrocyte DEGs. In the SN, further enrichment of other inflammatory pathways, such as NFκβ, was evident, while pathways related to sphingolipid signaling were enriched in cingulate astrocytes (Fig. 5j). Other shared pathways include cellular senescence and Ras signaling (Fig. 5j and Supplementary Data 4-5).

Given the detection of these neuroinflammatory phenotypes, we specifically asked if the JAK/STAT signaling pathway, an established

neuroinflammatory pathway in astrocytes[78], was enriched in PD astrocytes. Using pre-ranked GSEA for the JAK/STAT KEGG pathway in genes ranked by their relative expression in SN PD versus control, we found that JAK/STAT signaling was enriched in nigral PD astrocyte genes (Fig. 5k).

Together, these results implicate astrocytes in the neuroinflammatory network in PD. We conceive of astrocytes as present on a

Fig. 4 | Patterns of dysregulation of myeloid cells in the substantia nigra and cingulate cortex. UMAP plots of cingulate cortex myeloid cells grouped by condition (**a**) and lineage/subtype (**b**). **c** Dot plot of select gene (y-axis) marker expression in cingulate myeloid lineages (x-axis). Size indicates percentage expression, and color indicates normalized expression levels. **d** same as a but for the substantia nigra (SN). **e** Same as b but for the SN. **f** Same as c but for the SN. **g** Differentially expressed genes (DEGs) of activated microglia in PD vs. control are shown by their log2 fold change (LFC) in the SN on the x-axis and the cingulate on the y-axis. The color indicates if the genes are significantly differentially expressed in the cingulate (Cing), SN, both, or none. DEGs were considered based on two-sided adjusted *p*-values. **h** Dot plot showing KEGG pathway enrichment scores and adjusted two-sided *p*-values of select pathways of activated microglia in the SN and

cingulate cortex. The size of each dot represents its fold enrichment value, and the color represents its −log10 two-sided *p*-value, with yellow denoting lower significance and red indicating higher significance. Only statistically significant terms are shown. **i** Representative immunohistochemistry showing CD68+ cells (brown chromogen – green arrows) and CD8 + T cells (red chromogen - pink arrows) in the SN from a control and a PD donor. Scale bar: 50 μm. **j** Quantification of CD68+ cells as a proportion of all SN cells PD and controls. *N* = 8 for control and *n* = 6 for PD. Two-tailed Mann-Whitney test *p*-value is indicated. Data is shown as mean +/- SEM. **k** Preranked gene set enrichment analysis of previously described PD-murine BAM cluster markers in PD BAM genes ranked by their fold change from control. Normalized enrichment scores and adjusted two-sided *p*-values are indicated.

---

spectrum of states, with one state being a homeostatic protoplasmic-astrocyte-like state and another being a GFAP-high, reactive, fibrous-astrocyte-like state. There are likely other potential states, including compensatory ones, e.g., metallothionein-high[44], and disease-associated ones, for example, C3-high[79] and CD44-high[44,76,80].

To gain a deeper understanding of astrocyte-state transitions, which we have described in HD, hypoxia, and seizures[44,80], we used gene trajectory analysis, which orders genes along trajectories of gene expression and indicates potential dynamics of gene expression changes as cells transition from one state to another[81]. We extracted gene programs that vary along an axis of variation in the cingulate (Fig. S10A, B) and the SN (Fig. S10D, E). This analysis allowed us to deconvolve gene expression programs at different points along a trajectory of state transitions. We identified several gene expression programs that included several astrocyte DEGs in the cortex (Fig. S10B) and the SN (Fig. S10E), including *SLC4A4* (decreased in the PD SN – gene program 2) and *CD44* (increased in PD - gene program 9), *WIF1* (decreased in cingulate PD – gene program 1) and *MT1G* (increased in cingulate PD – gene program 6). Of particular interest were programs 8-10 in the SN and cingulate, which showed a significant preponderance of fibrous-like astrocytes and PD-derived astrocytes (Fig. S10C, F). In the SN but not the cingulate, PD astrocytes had higher median scores in these gene programs (Fig. S10C, F). Together, these results show that PD astrocytes are more polarized towards a fibrous-like state with decreased metallothionein gene expression and elevated *CD44* expression in the SN, as we show in our validation studies (Fig. 5h–l).
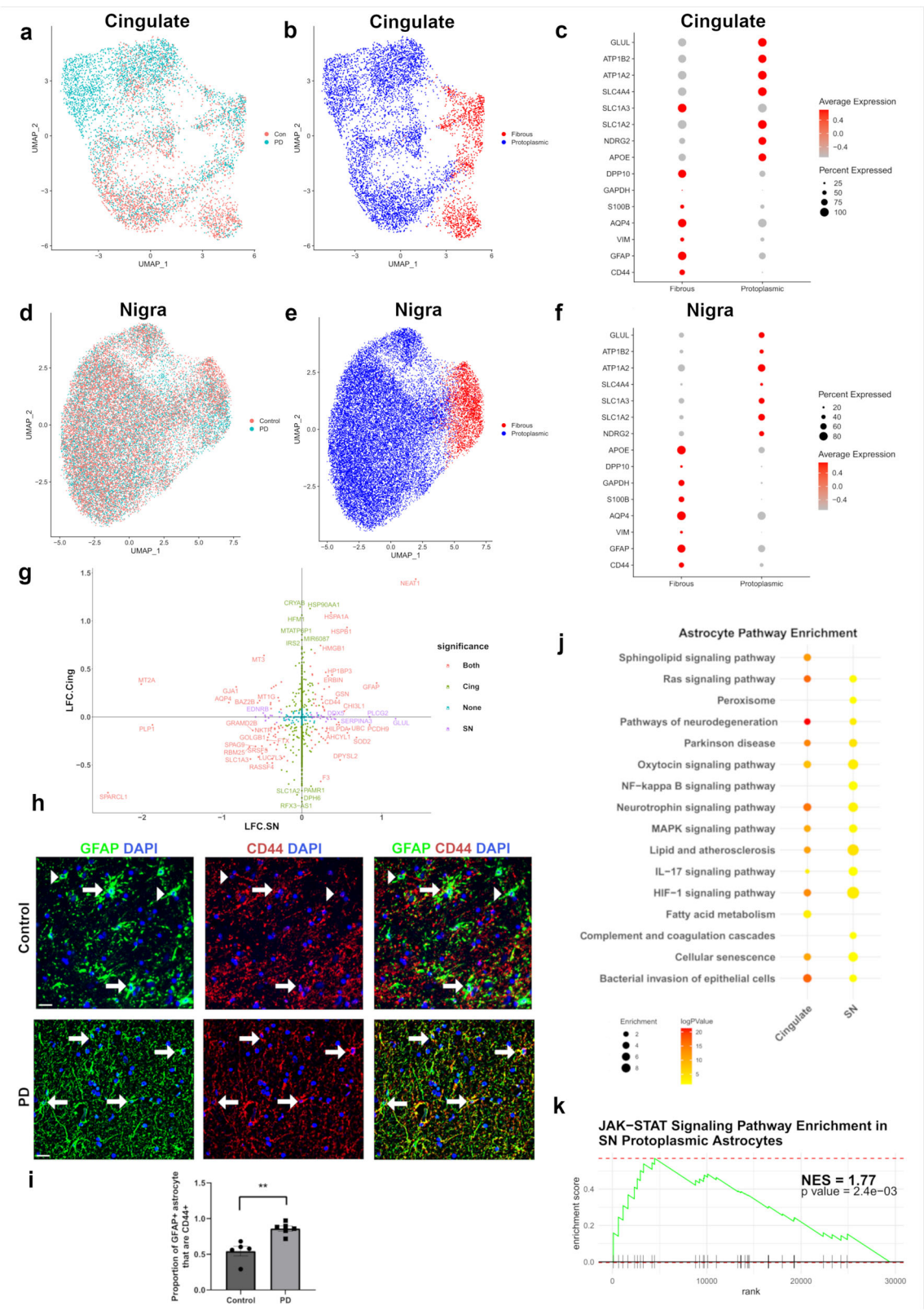
Spatial transcriptomics analysis reveals spatially diverse patterns of pathology in PD

To spatially map PD signatures within the diverse brain microenvironments, we conducted spatial transcriptomics on a subset of our SN tissue samples (*n* = 5 PD and *n* = 5 control; Fig. S11A–J). First, to evaluate cell-type-specific gene signatures in our spatial transcriptomics data, we employed Robust Cell Type Deconvolution (RCTD) to quantify the relative proportion of each cell-type/transcriptional state in each locale. Next, we measured the spot-level enrichment values for our T cell Factors 1 and 5 and identified spatial clusters/transcriptional niches defined using BayesSpace[82] (Fig. 6a; Fig. S11A–J, Fig. S12A–J, and Supplementary Results). Through comparison of deconvolved dopaminergic neuron cell-type proportions (Supplementary Data 1), *TH* expression values, and neuropathological evaluation of corresponding H&E images, we annotated the BayesSpace clusters as either SN pars compacta (SNpc) or two white matter clusters which were combined as "surrounding tissue" (Fig. 6b and Fig. S12A–J,Supplementary Results, and Supplementary Data 8). The SNpc region is defined by *TH* expression in addition to high proportions of dopaminergic neurons, as defined by deconvolution (see methods and Supplementary Data 8 for cluster markers). As expected, *TH* expression was higher in the SNpc compared to the surrounding tissue (Supplementary Data 8).

To specifically examine the interactions between T cells and astrocytes, we measured enrichment of the scores in T cell Factor 5 (TRM factor), T cell factor 1 (effector T cell), and genes increased in PD astrocytes in the SNpc and surrounding tissue regions. We found a significant increase in PD in T cell Factor 5 scores in the surrounding tissue regions, and astrocyte DEGs exhibited higher enrichment scores in both the SNpc and the surrounding tissue (Fig. 6c). The scores of T cell Factors 1 and 5 data are projected in spatial feature plots in Fig. S12A–J, which shows high expression of T cell Factor 1 in the SNpc and the surrounding tissue and relative depletion of T cell Factor 5 from the SNpc (linear mixed effect coefficients for comparing the surrounding tissue to the SNpc for Factor 5 and Factor 1: 0.0845 and 0.0109, respectively; *P* values: <2e-16 and 9.45e-16, respectively). Together, these data add a spatial dimension to the changes we observed at the single-cell level and indicate that astrocyte changes are pervasive in the SNpc and the surrounding tissue, and TRM signatures are enriched in the surrounding tissue, where large penetrating vessels are more abundant and harbor more T cells.

We next extracted spatial DEGs using our same pipeline that accounts for age, sex, and Alzheimer's changes to compare PD vs. controls in the SNpc and surrounding tissue (Supplementary Data 9 and Fig. 6d). As expected, *TH* was most decreased in PD in the SNpc compared to the surrounding tissue (Fig. 6d and Supplementary Data 9). *ALDH1A1*, one of the most significantly downregulated genes in dopaminergic neurons, which we found conserved in other datasets, as shown in Fig. 2g, was only decreased in the SNpc. Several genes were decreased in both regions, which included astrocytic genes like *MT3* and *APOE*, and increased in both regions, such as *GFAP* and *GLUL*, consistent with our snRNAseq results (Fig. 5g and Supplementary Data 4). Also consistent with snRNAseq was the reduction of *CD74*, a gene downregulated in activated microglia in the SN (Fig. 4g). Interestingly, *CD44* was increased only in the surrounding tissue (Fig. 6d). This is consistent with the localization of CD44+ astrocytes around large vessels in the white matter. Together, the results from ST add a spatial dimension to our snRNAseq results.

Finally, we used an unbiased approach to analyze ST niche-specific DEGs and measured the enrichment of KEGG pathways in the SN and surrounding white matter (Fig. 6e and Supplementary Data 9). Enriched in DEGs of both regions were pathways related to neurodegeneration, including PD, prion disease, and oxidative phosphorylation. Pathways related to inflammation, including antigen processing and presentation, IL-17 signaling, and Th17 cell differentiation, were also enriched in both regions. Interestingly, pathways related to the pentose phosphate pathway and fructose metabolism were enriched in the SNpc only, suggesting that different metabolic pathways are affected in the SNpc, which is consistent with our data on dopaminergic neurons (Fig. S5). Conversely, pathways related to TNF signaling, apoptosis, and MAPK signaling were enriched in the surrounding tissue only. Together, these results suggest that there are distinct, cell-type-specific, spatially defined pathologic signatures in PD.

Spatially defined cell-cell cohabitation patterns in the PD Substantia Nigra

To determine the spatial relationships of cell types and T cell factor 5, we performed spatial cross-correlation (SCC) analysis on the spot-level cell-type proportion values in ST datasets (Fig. 7a). SCC allows us to quantify how the cell types are correlated (SCC coefficients), assign statistical significance to the coefficients using a permutation-based method, and retrieve sample-level and disease condition-level statistics. If cell types are spatially correlated, then positive SCC values will be retrieved. If they are negatively spatially correlated, negative SCC coefficients will be retrieved. Since determining SCC across thousands of ST data points is computationally intensive and slow, we developed an approach to parallelize the computation, accelerating it by ~1200-fold (see methods and Fig. 7B and Supplementary Data 10).

Fig. 5 | Substantia nigra astrocytes in PD upregulate neuroinflammatory pathways. UMAP plots of cingulate cortex astrocytes grouped by condition (**a**) and lineage/subtype (**b**). **c** Dot plot of select gene (y-axis) marker expression in cingulate astrocytes lineages (x-axis). Size indicates percentage expression, and color indicates normalized expression levels. **d** same as a but for the substantia nigra (SN). **e** Same as b but for the SN. **f** Same as c but for the SN. **g** Differentially expressed genes (DEGs) of protoplasmic astrocytes in PD vs. control are shown by their log2 fold change (LFC) in the SN on the x-axis and the cingulate on the y-axis. The color indicates if the genes are significantly differentially expressed in the cingulate (Cing), SN, both, or none. DEGs were considered based on two-sided adjusted *p*-values. **h** Representative multiplex immunofluorescence showing CD44 (red), GFAP (green), and DAPI+ nuclei in the SN from a control and a PD donor. Scale bar: 20 μm. **i** Quantification of the CD44+ astrocytes as a proportion of all SN GFAP+ astrocytes in PD and controls. *N* = 5 for control and *n* = 6 for PD. Two-tailed Mann-Whitney test p-value is 0.0043. Data is shown as mean +/- SEM. **j** Dot plot showing KEGG pathway enrichment scores and adjusted two-sided p-values of select pathways of protoplasmic astrocytes in the SN and cingulate cortex. The size of each dot represents its fold enrichment value, and the color represents its –log10 p-value, with yellow denoting lower significance and red indicating higher significance. Only statistically significant terms are shown. **k** Preranked gene set enrichment analysis of JAK/STAT KEGG pathway in PD protoplasmic astrocyte genes ranked by their fold change from control. Normalized enrichment scores and adjusted two-sided *p*-values are indicated.

We used this approach to calculate changes in SCC between different cell types/states in PD compared to control (Fig. 7C). There were several changes in cell-cell colocalization patterns, including changes in glial and immune cells – the cell types being investigated herein. First, the spatial relationship between T cells and the T cell Factor 5 was significant in PD but not in controls. Second, we identified several cell type combinations with increased SCC, including a) activated microglia and endothelial cells, b) fibrous-like CD44+ astrocytes and T cells, c) BAMs and CD44+ fibrous-like astrocytes, d) T cells and endothelial cells, and e) T cells and oligodendrocytes. A more detailed interpretation of SCC is provided in the supplementary results section. Taken together with our DEG results, these findings highlight statistically significant patterns of increased spatial correlation and proximity between immune cells, T cells, and other glial cells in the human postmortem PD SN, which establishes a foundation to investigate neuroinflammation in PD.

We performed several validation studies to confirm the changes in SCC in PD. First, to confirm the increase in SCC between T cells and activated microglia in PD, we quantified the spatial relationships between CD68+ activated microglia and CD8 + T cells using dual IHC (Fig. 7d). We measured the minimum distance between CD68+ microglia and T cells in each section, compared the median distance between the T cell/microglia nearest neighbors using a Wilcoxon test, and found the minimum distance between T cells and activated microglia was reduced in PD (Fig. 7e). Second, to confirm the increase in SCC between T cells and fibrous-like CD44+ astrocytes in PD, we performed multiplex immunofluorescence for GFAP, CD44, and CD3 (Fig. 7f). As with microglia, the minimum distance between T cells and the closest CD44+ astrocyte was reduced in PD (Fig. 7g). Of note, we found that CD3 + T cells expressed CD44 as well, which has been reported to be a marker of memory status[62,83]. Finally, to validate the increase in SCC between T cells and oligodendrocytes, we quantified the density of CD8 + T cells in the oligodendrocyte-rich white matter of the cerebral peduncle. We found significantly higher densities of T cells in the white matter in PD (Fig. S13A, B). These findings validate the spatial changes we identified in PD between glia and immune cells using spatial transcriptomics.

CD44 loss of function reduces neuroinflammatory phenotypes in cultured astrocytes

As shown above, CD44+ astrocytes gain spatial correlation to T cells in the PD SN, CD44 was significantly increased in PD astrocytes, and PD astrocytes showed an increase in pathways and genes involved in neuroinflammation such as the JAK/STAT pathway (Fig. 5k). The intercellular domain of CD44 can be cleaved off and binds transcription factors in the nucleus, including STAT3[84]. Thus, we hypothesized that CD44 is necessary to mediate the signaling of pathways involved in neuroinflammation, such as the JAK/STAT pathway. CD44 is expressed in murine astrocytes in vitro (Fig. 8a, b), so we performed knockdown studies to abolish its expression in this system.

We tested four shRNA constructs and identified two that abolished CD44 expression in astrocytes, which we confirmed using multiplex immunofluorescence and Western blots (Fig. 8a, b). Next, we measured transcriptomic changes in control versus CD44 knockdown using RNAseq. We then interrogated the gene expression changes using differential gene expression analysis (Supplementary Data 11 and Fig. 8c) and pathway enrichment analysis (Fig. 8d, e). We compared the log-fold changes of DEGs after CD44 knockdown to the SN protoplasmic astrocyte DEGs (Fig. 8c). *CD44* and *CP*, which were increased in SN astrocytes, were reduced by CD44 knockdown, and GFAP was no longer significantly differentially expressed in CD44 knockdown. Other genes that were decreased in SN protoplasmic astrocytes were increased by CD44 knockdown, including *SYNE1* and *SON*. Since JAK/STAT and MAPK pathways were enriched in SN protoplasmic DEGs, we investigated alterations to these pathways with our CD44 knockdown model. Using decoupleR to measure pathway activation scores, we found that CD44 knockdown significantly dampened JAK/STAT pathway activation (Fig. 8d). Taking a more unbiased approach; we examined KEGG pathway enrichment in the CD44 knockdown DEGs (Fig. 8e). The results showed that several KEGG pathways were enriched in CD44 knockdown DEGs, including MAPK, sphingolipid signaling, and JAK/STAT (Fig. 8e). This is particularly interesting because JAK/STAT signaling was significantly enriched in PD astrocytes (Fig. 5k). CD44 knockdown decreased many of the genes involved in JAK/STAT signaling (Fig. 8e and Supplementary Data 11). These data suggest that CD44 upregulation in astrocytes is necessary for JAK/STAT activation, which is a neuroinflammatory pathway that is increased in PD, thus nominating CD44 as a potential target to block neuroinflammation in PD.

Discussion

The major findings in this work reveal that CD8 + T cells infiltrate the SNpc, demonstrate a resident-memory phenotype, and are clonally expanded. Importantly, we found similarities in the TCRs of PD patients and peripheral blood TCRs isolated from different patients, which react to α-synuclein. In contrast to the SN, the cingulate cortex did not show increased T cell infiltration. Through spatial transcriptomics, snRNAseq, and validation protein labeling studies, we found that the T cells reside in an inflammatory niche within a microenvironment consisting of activated microglia which upregulate SPP1, a proinflammatory molecule we previously showed to be upregulated in frontotemporal dementia SN[48], and reactive astrocytes that upregulate CD44, which our past studies have demonstrated to be representative of a disease-associated astrocyte state[44,76]. Astrocytes in the SNpc upregulated several neuroinflammatory pathways like JAK/STAT and MAPK, and functional studies revealed that the knockdown of CD44 decreases the activation of these pathways in vitro and rescues several gene expression changes we see in SN astrocytes. These results nominate CD44+ astrocytes as a potential target to abort the vicious neuroinflammatory cycle in PD.

In addition to the increase of clonal T cells in the PD SN, we observed an increase in CD44+ fibrous-like astrocyte genes in the substantia nigra, consistent with our gene trajectory analysis, which was coupled with an absence of upregulation of the neuroprotective
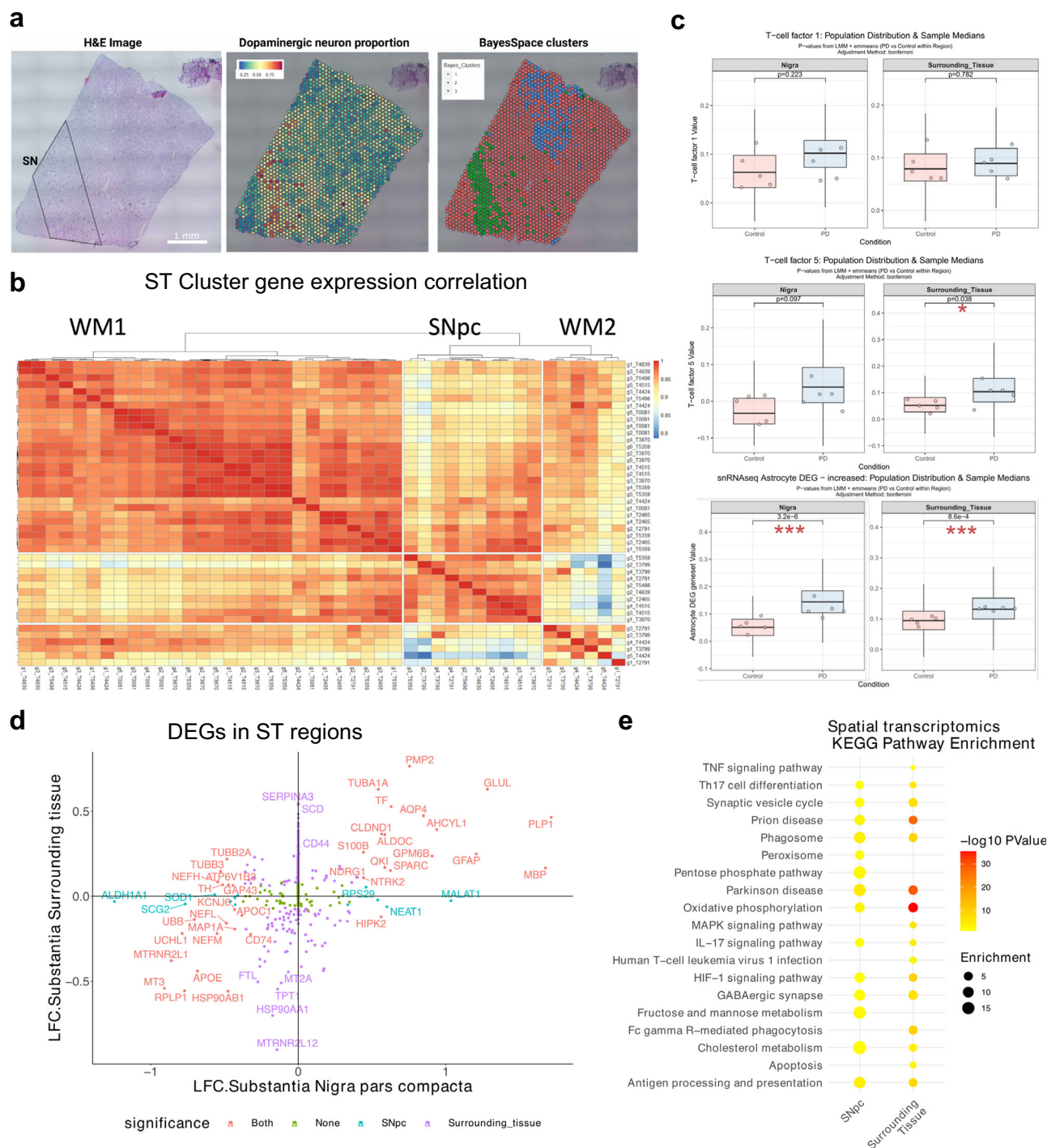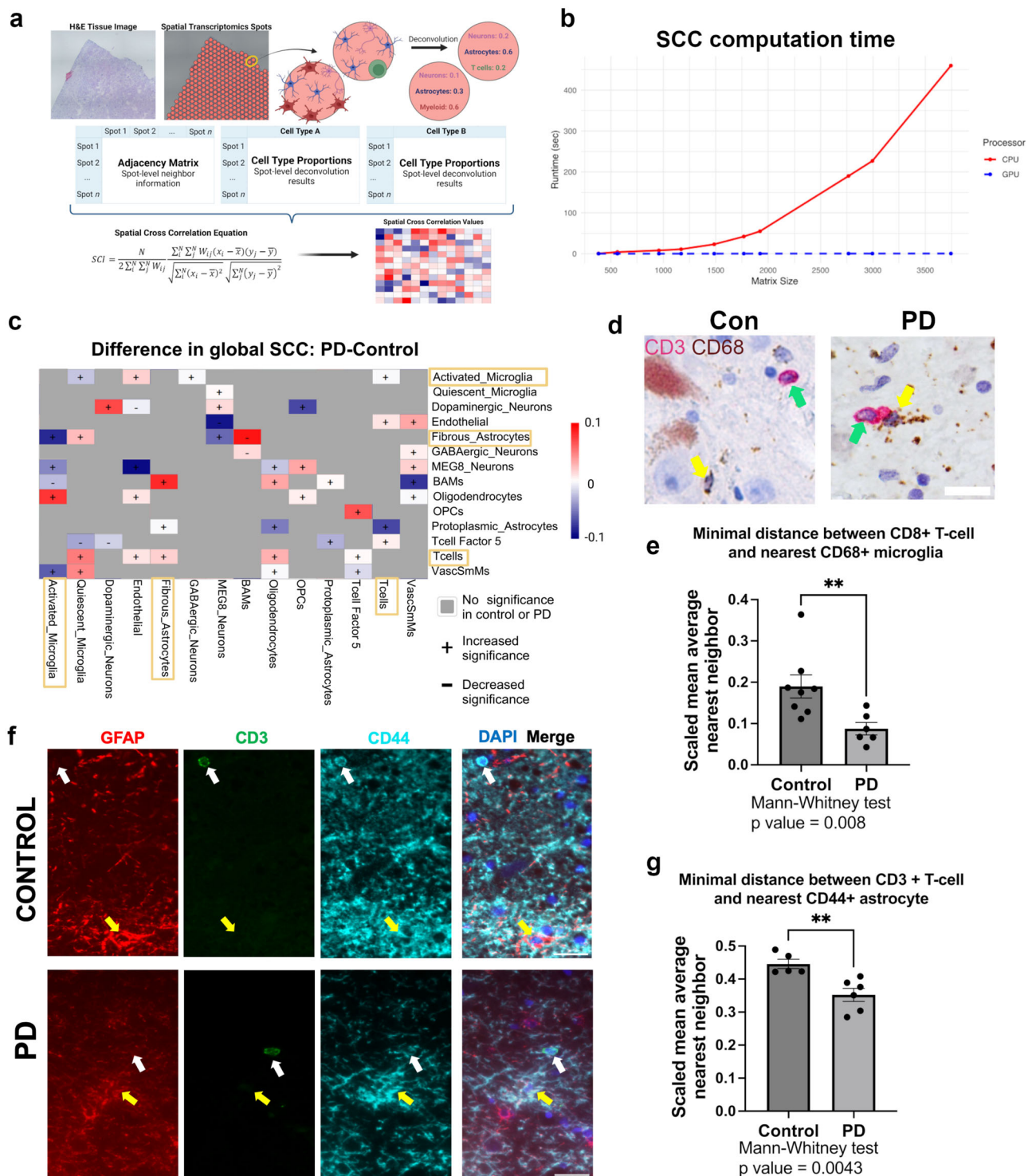
Fig. 6 | Spatial transcriptomics analyses localize cell-type specific signatures to local anatomic compartments in the PD substantia nigra. a Representative H&E image (left) from tissue used for 10X Visium ST. n = 10 donors (biological replicates). The SN is outlined and populated by neuromelanin (brown) containing neurons. Scale bar: 1 mm. Deconvolved proportions of dopaminergic neurons are shown per spot (middle). Assigned BayesSpace clusters (right) are indicated. **b** Heatmap of gene expression correlation of sample-level BayesSpace clusters. Two white matter clusters (WM1 and WM2) and a substantia nigra (SNpc) cluster are indicated. The two WM clusters were combined in the downstream analysis as "surrounding tissue." **c** Box plots of the distribution of enrichment scores of T cell Factor 1 (upper), T cell Factor 5 (middle), and snRNAseq-derived astrocyte DEGs (lower), which are enriched in the SNpc and surrounding tissue regions in control and PD ST samples. The dots indicate the median score per donor. The line of the boxplot indicates the median of the population of all spots; the bounds represent

the 1st and 3rd quartiles. The whiskers extend from the edges of the box to the smallest and largest values within 1.5 times the interquartile range. Adjusted two-sided p-values are indicated and determined with a linear mixed model – see methods and source data. Control-n = 5, PD-n = 5. **d** DEGs in Visium Spatial transcriptomics capture spots in PD vs. control are shown by their log2 fold change (LFC) in the SNpc on the x-axis and the surrounding tissue on the y-axis. The color indicates if the genes are significantly differentially expressed in the surrounding tissue, SNpc, both, or none. DEGs were considered based on two-sided adjusted p-values. **e** KEGG pathway enrichment analysis in PD vs control DEGs in SN and surrounding tissue from ST data. The size of each dot represents its fold enrichment value, and the color represents its –log10 adjusted two-sided p-value, with yellow denoting lower significance and red indicating higher significance. Only statistically significant terms are shown.

The equation shown in panel a:

$$SCI = \frac{N}{2\sum_i^N\sum_j^N W_{ij}}\frac{\sum_i^N\sum_j^N W_{ij}(x_i-\overline{x})(y_j-\overline{y})}{\sqrt{\sum_i^N(x_i-\overline{x})^2}\sqrt{\sum_j^N(y_j-\overline{y})^2}}$$

protein MT3. This is similar to our previous findings in HD and highlights a region-specific response of astrocytes to neurodegeneration, which may result from or contribute to regional vulnerability in neurodegeneration[44]. Additional studies are needed to determine the generalizability of this phenomenon and whether it is a driver of resilience to neurodegeneration, a compensatory response to neurodegeneration, or both.

The spatial analyses yielded several technical and conceptual advances. First, the TRM T cell Factor 5 was enriched in SN white matter regions, and this enrichment was more pronounced in PD than in controls. Conversely, T cell Factor 1 showed localization in both the

SNpc and the surrounding white matter and was not different between the conditions in the ST data. Astrocyte DEGs were enriched in both. To further dissect these changes, which likely represent compositional changes, we adapted our previous approach using SCC to define patterns of cellular cohabitation in infiltrating glioma[77]. Here, we have introduced a computational approach with massively parallel GPU-acceleration that yielded computation time up to 1200 times faster than CPU-based methods (see methods and Fig. 7b; Supplementary Data 10). By leveraging the spatial data from multiple ST donors to statistically measure changes in SCC, we defined a network of spatially validated neuroinflammatory niches, where

Fig. 7 | Spatial cross-correlation analysis reveals altered cellular cohabitation patterns in PD. a Schematic explaining spatial cross-correlation (SCC). Robust cell-type deconvolution (RCTD) is employed to determine the cellular composition of each spot (capture area). Spot-level neighbor information is encoded in a binary adjacency matrix, which is then combined with proportion matrices for each cell type in a previously defined SCC equation. The output is an SCC value for each cell type combination. Created in BioRender. Al Dalahmah, O. (2025) https://BioRender.com/btedio9. **b** Plot of the matrix size (number of elements) by the amount of time (seconds) taken to complete the SCC computation using our optimized algorithm, which was conducted using the CPU (red) and GPU (blue). **c** Heatmap of change in average SCC values for each cell-type combination, PD compared to control. Increased values (red) denote an increase in SCC in PD compared to control; decreased values (blue) denote a decrease. "+" symbols represent an increase in SCC significance in PD compared to control or a lower aggregated two-sided p-value. "-" symbols represent a decrease in significance. Grayed-out boxes represent

relationships that were not significant (aggregated p value > 0.05) in either PD or control or that lost significance in PD compared to control. **d** Validation of increased spatial correlation between T cells and activated microglia in PD. IHC for CD68 and CD3 in the SN of a control and a PD donor. Green arrows indicate CD3+ T-cells, and yellow arrows indicate CD68+ microglia. **e** The average min-max scaled minimum distance between a T cell and the nearest neighbor CD68+ cell was measured. Two-tailed Mann-Whitney test, n = 8 control and n = 6 PD. The p-value is indicated. Data is shown as mean +/- SEM. **f** Validation of increased spatial correlation between T cells and CD44+ astrocytes. Multiplex immunofluorescence showing GFAP (red), CD44 (cyan), CD3 (green), and DAPI nuclei (blue) in the SN of a control and a PD donor. Yellow arrows indicate CD44+ astrocytes and white arrows indicate T cells. Scale bar: 20 μm. **g** The average min-max scaled minimum distance between a T cell and the nearest neighbor CD44 + GFAP+ astrocyte was measured. Two-tailed Mann-Whitney test, n = 5 control and 6 PD. P-value is indicated. Data is shown as mean +/- SEM.

neurodegeneration in PD takes place. This network features activated microglia, T cells with TRM and activation features, and CD44+ astrocytes. We confirmed these associations between T cells and CD44+ astrocytes and T cells and activated microglia using histopathologic studies.

The neuronal pathology we document in SN and cortical neurons is consistent with prior snRNAseq reports[6,7,10]. It is interesting that many of the pathways dysregulated in vulnerable SN dopaminergic neurons were shared with cortical glutamatergic neurons. However, key metabolic KEGG pathways, including beta-alanine and tyrosine metabolism, were enriched only in the SN. These findings point to potential differences in metabolic states that may underlie neuronal vulnerability to neurodegeneration. However, further studies are needed to investigate this phenomenon.

Taken together, our work outlines a potential immune-glial signaling axis that includes fibrous-like CD44+ astrocytes, myeloid cells (activated microglia and BAMs), and T cells, which may have the net effect of eliciting a reactive, fibrous-like state in astrocytes, activation of myeloid cells, and potentially T cell clonal expansion in PD. The critical question remains to be answered: do T cells drive neuronal loss in PD, or are they only necessary for perpetuating neurodegeneration? Mouse model studies show that T cells indeed play critical roles in driving neuronal death[24,85] in the setting of overexpression of α-synuclein. Whether expanded T cells can serve as a marker for tracking the pace of neurodegeneration is an exciting avenue of research we are actively pursuing. The presence of shared motifs between PD SN TCR's and peripheral blood α-synuclein reactive T cells from two external studies is an interesting finding that merits further exploration. Thus, additional studies are needed to experimentally test if CNS TCRs in the PD SN do indeed react to α-synuclein.

Limitations
There are notable limitations of the current study. We do not have paired peripheral blood and SN samples from the same donors; therefore, we cannot draw conclusions about which brain-resident clones are also present in the periphery. We note that our studies are of neuropathology in subjects with advanced PD and low numbers of surviving SN dopaminergic neurons, and as such we cannot address the issues of whether T cells in the PD SN have increased interactions with neurons or changes in T cell characteristics at earlier disease stages when the highest rates of neuronal damage occur. As quantified by a board-certified neuropathologist (JEG), less than 1% (6 of 1151) of CD8+ cells in the PD SN were observed next to neurons, discounting statistical analyses of the phenomenon.

Methods
Human subjects and brain tissue
All study protocols were approved by the Columbia University Irving Medical Center Institutional Review Board (Protocol number
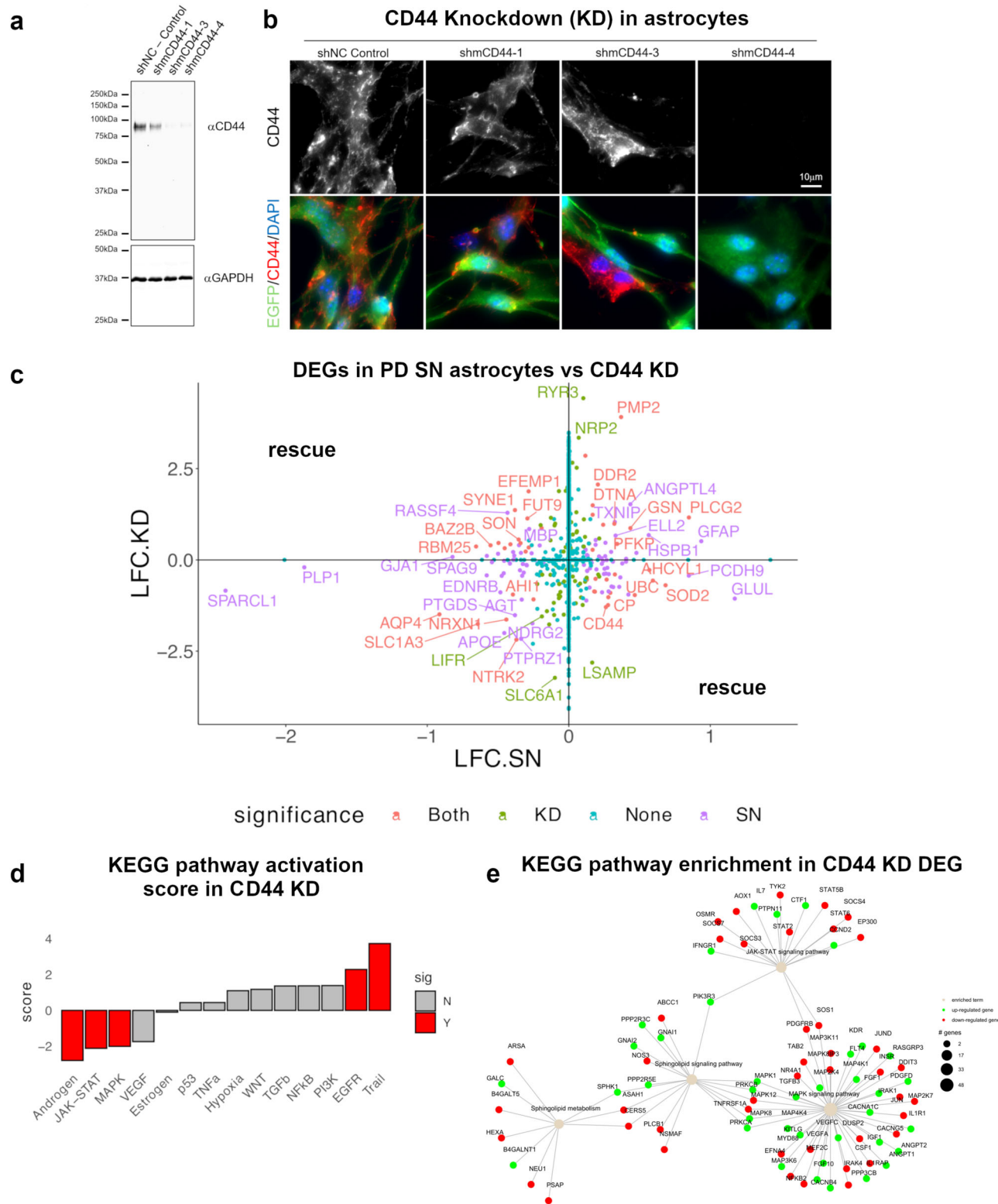
AAAU1394). Postmortem cingulate cortex and SN specimens frozen during autopsy from control (individuals whose brains did not show significant neuropathology) and PD/DLB donors were obtained from the New York Brain Bank. Consent from the next of kin was obtained for all brain donations. The tissue was dissected by a board-certified neuropathologist (OAD), or under the supervision of a board-certified neuropathologist. Forty-four cases were selected for snRNAseq and TCR sequencing, each with RNA integrity numbers >7, and ten of these were selected for spatial transcriptomics analysis. Cortical wedges, excluding subcortical white matter or SN tissue measuring ~ 5 × 4 x 3 mm, were dissected on a dry ice-cooled stage and processed immediately as described below. The demographics of the cases used are provided in Supplementary Data 1.

TCR sequencing
To prepare α-chain TCR libraries, we followed the iRepertoire Bulk Reagent Universal User Manual (V20200818). The starting material was 500 ng RNA per sample. We used 9 barcodes - HTAIvc kits (HTAIvc01, HTAIvc02, HTAIvc03, HTAIvc04, HTAIvc05, HTAIvc06, HTAIvc07, HTAIvc08 and HTAIvc09). We pooled one library from each of the barcoded kits together for each sequencing run. The libraries were pooled with 10% PhiX spike-in and sequenced with NextSeq High Output 300 Cycles kits (Illumina) on an Illumina NextSeq 550 (read 1: 155 cycles; read 2: 155 cycles). Five total sequencing runs were conducted.

To prepare combined α-chain and β-chain TCR libraries, we followed the iRepertoire RepSeq+ protocol; 25 RNA samples were submitted to iRepertoire, Inc. (Huntsville, AL, US). 5 μl of RNA from each sample was used as a template for iR-RepSeq + . Next-generation sequencing libraries encompassing the TCR alpha and beta chains were generated. RepSeq+ technology facilitates simultaneous amplification of TCR alpha and beta chains while mitigating common sources of errors in AIRR-seq analyses[86]. The amplification protocol employs primer pairs specific to each V-D-J combination, enabling tag extension for subsequent global amplification of TCR αβ. Unique molecular identifiers (UMIs) were incorporated during the reverse transcription step to distinguish individual RNA molecules and minimize the impact of PCR duplicates and sequencing errors.

Reverse transcription was performed using Qiagen OneStep RT-PCR mix with C-gene primer mix, followed by selection of first-strand cDNA selection and removal of remnant primers via SPRIselect bead purification (Beckman Coulter). A second round of amplification was conducted using a V-gene primer mix, followed by SPRIselect bead purification. Library amplification was performed with primers targeting communal sites engineered onto the 5' ends of the C- and V-primers. The final libraries contained Illumina dual-index sequencing adapters, a 10-nucleotide random region, and an 8-nucleotide internal barcode associated with the C-gene primer. Sequencing coverage included from within the framework 1 region to the C-region, inclusive

a

b CD44 Knockdown (KD) in astrocytes

c DEGs in PD SN astrocytes vs CD44 KD

d KEGG pathway activation score in CD44 KD

e KEGG pathway enrichment in CD44 KD DEG

of CDR1, CDR2, and CDR3. Libraries were multiplexed, pooled, and sequenced on 18% of a NextSeq 1000, P2 flow cell, 600 cycles, 36.6 million reads.

TCR sequencing analysis

Standard iRepertoire alpha-chain and RepSeq+ alpha- and beta-chain sequencing raw data were analyzed using the iRmap program[87,88]. Briefly, sequence reads were de-multiplexed according to Illumina dual indices and barcode sequences. Merged reads were mapped to germline V, D, J, and C reference sequences using an IMGT reference library. CDR3 regions were identified, extracted, and translated into amino acids. The dataset was condensed by collapsing UMIs and CDR3 sequences to correct for sequencing and amplification errors. Reads sharing identical CDR3 and UMI combinations were condensed into a single UMI count. The resulting immune receptor sequence data span from the first framework region to the beginning of the constant region, including the CDR3 hypervariable region. For read-based iRepertoire alpha-chain sequencing, consistent with previous

Fig. 8 | Astrocytic CD44 is necessary for the activation of neuroinflammatory pathways. a Western blots of murine astrocytes transduced with viruses carrying control non-targeting (shNC) versus shRNA against murine CD44 (shmCD44-1, 3, and 4). shmCD44-3 and-4 effectively knocked down CD44 expression. shmCD44-1, 3, and 4 are independent biological replicates. The lentiviruses included a GFP-tag to label infected cells. **b** Representative immunofluorescence images of murine astrocytes transduced with shRNA constructs as for panel A. CD44 (red), GFP (green), and DAPI (blue) nuclei are labeled. Scale bar: 10 μM. Note complete loss of murine CD44 in GFP+ cells in sh-mCD44 3 and 4. Uninfected GFP- cells retain CD44. For a, b, the knockdown experiment was replicated twice. **c** Differentially expressed genes (DEGs) of PD vs control substantia nigra (SN) protoplasmic astrocytes and DEGs in murine astrocytes with CD44 knockdown (KD) vs. non-targeting shRNA

control. The genes are shown by their log2 fold change (LFC) in the SN on the x-axis and CD44-KD on the y-axis. The color indicates if the genes are significantly differentially expressed in PD SN, CD44 KD, both, or none. DEGs were considered based on two-sided adjusted p-values. **d** Bar plot showing activation scores of select pathways in differentially expressed genes in CD44 KD astrocytes. The scores were calculated using Omnipath and decoupleR – see methods. Significantly repressed and activated pathways are shown in red. **e** Gene-term plot showing significantly enriched KEGG pathways derived from PathfindR analysis of CD44 DEGs, where the pathways are shown in tan nodes, the size of which corresponds to the number of genes driving the enrichment of that pathway and the genes contributing to the pathway enrichment are shown as red node (decreased DEGs) or green nodes (increased DEGs).

descriptions[89], the reads were mapped to the T cell receptor loci (TRA) and associated with V- and J-cassettes and were extracted and translated in silico in all three readings. Reading frames containing a C...FGXG amino acid motif were identified as productive CDR3 amino acid sequences. For each demultiplexed sample, all V- and J-cassettes were then reference-corrected, and the number of reads identified with each unique combination of V- and J-cassettes encoding a CDR3 amino acid sequence was counted.

Saturation levels for all sample libraries were assessed using the estimate_saturation function from the RNAseQC[90] package in R, at a depth of 200 and using 10 reps, and saturation curves were plotted for each sample (Fig. S2A, B). Read depths and unique CDRs in each dataset are provided in Supplementary Data 2.

Repertoire diversity metrics and clonal expansion
Repertoire diversity metrics were calculated using the chemodiv package[91] in R with default parameters. Shannon's entropy, Hill's diversity, and Hill's evenness were calculated based on clonotypes that included the CDR3 amino acid sequence and the V(D)J recombination. The diversity-equality 50 (DE50[45]) metric was calculated by measuring the fraction of clonotypes that accounted for 50% of the total reads across all clonotypes.

CDR3 sequence clustering
We combined our SN β-chain TCR UMI repertoires with β-chain TCR UMI counts from Gate et al.[13] and read counts from Singhania et al.[11]. Next, we used GLIPH2[92] to extract conserved motifs across CDR3 sequences using default parameters and the following arguments: accept_sequences_with_C_F_start_end = F, sim_depth = 10000, kmer_mindepth = 3. Cluster visualization was performed in turboGliph using the plot_network function with default parameters. For Fig. S3B, we used the SuperExactTest package[93] to examine conserved CDR3 motif sequences shared between blood and brain samples from control and Parkinson's disease (PD) patients. Conserved sequences from each group were supplied as list inputs to the supertest function. The background population size was set to $n = 10,000$, representing an approximate upper bound on the number of conserved sequences. P-values were computed for each multi-set intersection and adjusted for multiple comparisons using the Bonferroni correction. Logoplots in Fig. 1j were generated using weblogo.berkeley.edu with default parameters.

Extraction of nuclei and snRNAseq procedure
Nuclei were isolated from frozen postmortem brain slices in accordance with established protocols[44,48,77,94]. Libraries were prepared using Chromium Next GEM Single Cell 3′ Reagent Kit v3.1 (PN 120237), with Chromium Single Cell A Chip Kit, 48 runs (PN 120236). Target cell recovery was 10,000 cells per sample for cingulate samples and 20,000 cells for SN samples. The final number of nuclei was calculated from the average of three counts on Countess II or III (ThermoFisher©) using DAPI as a nuclear marker. The index plate used was a 10X Dual Index Kit TT Set A (PN 1000215). Chromium Next GEM Single Cell 3′

Reagent Kit v3.1 user guide (CG000315 Rev C) was followed. We used 10X Chromium v3 chemistry.

Sequencing and raw data analysis
Sequencing of the resultant libraries was performed on an Illumina NOVAseq 6000 platform V4, 150 bp paired end reads, 150 cycles. Read alignment was performed using the CellRanger pipeline (v6.1.2 – 10X genomics) to reference GRCh38.p12 (refdata-cellranger-GRCh38-1.2.0 file provided by 10x Genomics). Count matrices were generated from BAM files using the default parameters of the CellRanger pipeline. Filtering and QC were performed using DecontX[95], with default parameters, for the cingulate dataset and CellBender[96] for the SN dataset. CellBender (version 0.2.0) was run to remove ambient RNA with the addition of the '-cuda' flag to expedite the processing. Parameters were set with an expected cell count of 10,000, total droplets included at 30,000, FPR (false-positive rate) at 0.01, and a learning rate of 0.0001, utilizing 150 epochs. The total runtime for each sample ranged from 30 min to 1 h, with acceleration achieved through the use of the NVIDIA A5000 GPU. Decontamination of background was not necessary in cingulate samples. Nuclei with percent read aligning to mitochondrial genes >14% were excluded. Genes were filtered by keeping features with >500 counts per row in at least 100 cells. Doublets were identified using scDblFinder[97] and then removed.

Pre-clustering and clustering and classification of nuclei
Preclustering of nuclei was performed using Seurat's shared nearest neighbor smart local moving algorithm[98]. First, data was normalized using SCTransform[99], regressing out percent mitochondrial genes and donors. Data integration across donors was achieved using the Harmony[100] package which effectively regressed out donor effects. Harmony embeddings were used in the FindNeighbors step. Elbow plots based on PCA for each data set were used to determine an optimal number of principal components, and Clustree[101] was used to determine optimal resolution values for the FindClusters() step. Seurat's FindAllMarkers() function was used to determine basic cluster markers, which were then used to assign broad lineage identities to each cluster (astrocyte, neuron, oligodendrocyte, OPC, myeloid, endothelial, vascular, T cell). To assist with cell type sublineage assignment, we employed EnrichR[102], enabling us to garner information from multiple databases based on our representative genes. Cingulate cortex neurons were assigned in line with Paryani et al.[44]. Nuclei that did not conform to cell types were presumed to be doublets or artifactual noise and removed. The entire process was iteratively repeated for each lineage to remove aberrant cells and to assign subclusters, or sublineages/subtypes, within each lineage/cell type, including cortical myeloid cells, which underwent additional filtering after subclustering.

Differential abundance analysis
Beta-binomial generalized linear regression analysis in DCATS[103] in R was used to measure the influence of condition on the composition of dopaminergic neurons and T cells in the SN snRNAseq data. The input

to the algorithm was a matrix of all subtype compositions by donor provided in supplementary dataset 1, as well as a similarity matrix extracted from Seurat's KNN matrix. The coefficients for the cell types of interest and FDR-adjusted p-values were extracted and shown in Fig. S4I.

Differential gene expression analysis

To compare differences in gene expression between PD and control for each cell type, we used limma[104] within each lineage cluster. We controlled for donor, binned age using R's cut function, and sex in the model formula. Thresholds for most lineages were counts greater than 4 in at least 6 cells, and, for lineages containing less than 1000 cells, the threshold was lowered to counts greater than 2 in at least 4 cells. Our dataset did not include any separate batches. Only genes with p-values < 0.05 were carried through to downstream analyses. For Fig. S3C, donor-level pseudobulk counts were compared using a two-tailed Wilcoxon test.

Gene set enrichment analysis and gene ontology analyses

Packages fgsea[105] and PathfindR[106] were used to determine gene sets enriched within our differentially expressed genes for each cell type. All differentially expressed genes, along with their logFC and adjusted P values, were used as input in the run_pathfindR function, using the KEGG genesets. The parameters specified were 0.05 as the adjusted p-value threshold (using the adjusted p-value output from limma DGE analysis), minimum gene set size 5, and maximum gene set size 500. The cluster_enriched_terms function was run, with default parameters, to find representative pathways and filter out irrelevant/uninformative pathways. This analysis was used for panels 4h, 5j and 6e.

Using the fgsea package, we compared our T cell lineage to a CD8 + TRM gene set (Supplementary Data 6 – the table includes references for each gene). All genes in the T cell object were assigned a logFC value through Seurat's FindMarkers function, using PD as ident.1 and Control as ident.2, with parameters logfc.threshold, min.pct, and min.diff.pct set to -Inf to prevent filtering/removal of any genes. These genes, ranked by logFC, were input into the fgsea function with default parameters. Normalized enrichment scores and p-values were reported. This analysis was used for panels 3g, 4k, 5k,

To construct upset plots (Fig. S5G, H), we used the UpSetR[107] package. All myeloid DEG data frames from limma voom were separated into increased (logFC>0) and decreased (logFC<0), and lists of increased and decreased DEGs were input separately into the fromList function before running the upset function with default parameters.

Hierarchical poisson factorization

We used the scHPF package[49] in Python to determine interpretable factors within our SN snRNAseq dataset. The scHPF command line workflow comprised three fundamental stages: "scHPF prep," "scHPF train," and "scHPF score." In the "scHPF prep" phase, the molecular count matrix was utilized to generate a matrix market file and a gene list text file. The parameter "-m" was set to 10, filtering genes to include only those present in 10 or more cells. In the "scHPF train" stage, our SN dataset was aligned with each cell type, employing a candidate parameter range from K = 7 to 17 with a step of 2. Subsequently, for the extraction of disease factors within each cell type, the training was conducted with K values of 3, 5, 7, 9, and 11. Finally, in the "scHPF score" phase, the trained models for each K value were employed to assign gene scores to individual factors, resulting in the generation of ranked gene lists. We then selected K to prevent significant overlap in gene signatures among factors. This was mainly done by observing the factors expressed by each cell type, and evaluating which K value lent itself to the most interpretable factors (gene sets following canonical gene expression patterns).

Gene trajectory analysis

The GeneTrajectory R package was used to deconvolve specific trajectories in the snRNAseq data[108]. First, Seurat's FindVariableFeatures was used to extract highly variable genes, which were then further subset by selecting genes expressed by 1% to 50% of cells among the top 2000 variable genes. Next, we computed the diffusion map cell-cell kNN graph using the RunDM, GetGraphDistance, and CoarseGrain functions with K nearest neighbor parameter = 5 and N = 500 metacells for computational efficiency. Gene-gene distances were computed using the Wasserstein distance with the Python Optimal Transport package by implementing reticulate in R and calling the cal_ot_mat_from_numpy function with all default parameters. Finally, we extracted the gene embeddings and respective trajectories with the GetGeneEmbedding and ExtractGeneTrajectory functions with the number of trajectories N = 2, K nearest neighbor = 5, and the time step hyperparameter of the trajectory set to 7. Visualizing the extracted gene programs required us to run adaptively-thresholded low rank approximation (ALRA)[109] to smooth expression values and generate gene program bins. Seurat's RunALRA function was run with all default parameters. Next, we projected the gene trajectory values onto the UMAP with the AddGeneBinScore function where we specified the use of "alra" assay in our projection.

Spatial transcriptomics

Following 10x Visium Spatial Protocols – Tissue Preparation Guide (CG000240), OCT embedded tissue was scored to the size of the capture area targeting the SN. One 10 μm section was mounted on each capture area of the Visium slide. Tissues on the slides were fixed using a methanol-containing buffer as per the 10X Visium manual, stained with H&E or antibodies NeuN, GFAP, and DAPI as per the 10X protocol for Immunofluorescence Staining & Imaging for Visium Spatial Protocols (CG000312), and then imaged. Imaging of whole slides was done at 20X magnification on a Leica DMI8 Thunder microscope. After imaging, the slides were de-cover-slipped, and the tissue was permeabilized for 11 min (which was empirically determined to yield the best results based on the Visium Spatial Tissue Optimization Slide & Reagent Kit (PN-1000193), as detailed in the protocol provided in document CG000238 Rev D available in 10X demonstrated protocols). The remaining steps were conducted according to the manufacturer's protocol to prepare the libraries. Briefly, libraries were prepared using Visium Spatial Gene Expression Slide & Reagent Kit, 16 reactions (PN-1000184). Visium Spatial Gene Expression Reagent Kits user guide (CG000239 Rev G) was followed. The libraries were sequenced on NOVAseq (paired-end dual-indexed sequencing), targeting a minimum of 50,000 reads per spot.

The spatial transcriptomics (ST) samples were prepared using 10X Genomics Space Ranger (version 2.1.0) count commands, accompanied by Hematoxylin & Eosin (H&E) images in TIF format and manually-aligned JSON files generated from Loupe Browser (v7.0) with raw TIF images of the tissue. The loupe alignment JSON file was inputted into the loupe-alignment argument in Space Ranger along with its respective TIF image file, FASTQ reads, and slide numbers. The reference genome used for alignment was built using the Space Ranger function spaceranger mkgtf with GRCh38 as the assembly and Ensemble 91 for the transcript annotations. All other parameters were used with default settings.

ST object preprocessing and quality control

The number of counts per spot per ST sample is shown in Fig. S11A–J. The plots of ST experiments shown in Fig. 6a, S11, and S12 were generated using Seurat's SpatialFeaturePlot and SpatialDimPlot functions. A total of 10 samples were analyzed (Supplementary Data 1). Finally, any spots with zero counts were removed, and spot-level gene expression was normalized using SCTransform in Seurat.

Cell type deconvolution

Deconvolution using RCTD[110] was used to determine the proportion of each defined cell type in each ST spot from our data. We used the normalized counts matrix and nUMI from our SN snRNAseq object with annotated cell lineages and sublineages as a reference. Queries for RCTD were generated using coordinates from the "image" and "row" columns in the Seurat object, normalized counts, and nUMI for each sample. The function run.RCTD was run with parameter doublet_mode = "full". Otherwise, default parameters were used.

Spatial cross-correlation

We implemented spatial cross-correlation analyses to determine how different cell types were correlated with one another on a spatial plane[77,111]. For these analyses, we first created adjacency matrices for each sample using the getSpatialNeighbors function from the MERINGUE package[112] to denote which spots were first-order neighbors.

To avoid false neighbor assignment of nearby cells that were not true neighbors (e.g., separated by a break in the tissue), adjacency matrices were first created using all spots, whether in tissue or not, as listed in the Space Ranger "tissue_positions" csv output file. Next, all spots assigned as "in_tissue" were kept for downstream processing, and the rest were removed. This way, spots not directly next to each other would not be labeled as first-order neighbors.

RCTD cell-type enrichment values per spot, along with each sample's corresponding adjacency matrix, were combined to create spatial cross-correlation metrics by matrix multiplication. We used the same principles employed by MERINGUE's[112] spatial cross-correlation function; however, due to the large sizes of our input matrices, spatial cross-correlation was implemented by matrix multiplication in Tensorflow to expedite the processing time. Specifically, the local measurement of spatial cross-correlation involves multiplying two large matrices and obtaining the diagonal elements of the resulting matrix. The speed was further enhanced by utilizing the Einsum function in the TensorFlow package, which allows for element-wise computation. The code is available at: https://github.com/dalhoomist/T-cell_and_glial_pathology_in_PD.

Spatial transcriptomics clustering

To assign spatial clusters, we employed the R package BayesSpace[82]. We first processed our data with the spatialPreprocess function, using 7 principal components and 2000 highly variable genes for PCA, with log.normalize set to TRUE. We used the qTune function, evaluating q values between 2 and 10, and assessed the subsequent qPlot to determine the optimal number of clusters, q, defined by the elbow plot inflection point (Supplementary Data 1). We then used the spatialCluster function on the SCT counts for each sample, using the top 7 principal components, error model t, and 1000 MCMC iterations with 100 MCMC iterations excluded as a burn-in period. All other parameters were used in their defaults.

Additionally, we sought to neuroanatomically classify each spatial cluster using gene expression data. We first merged all ST objects using scCustomize and normalized them using scTransform, as described above in our snRNAseq processing methods. We extracted all 2000 variable features from the SCT assay. We then returned to our original objects (split by sample) and derived the spot-level gene expression values for each of the previously defined variable features. We then ran a correlation test using the "cor" function of the stats package. We generated a heatmap using the pheatmap package in R (version 1.0.12), with Manhattan distances and ward.D clustering (Fig. 6b). This resulted in three metaclusters, from which we derived cluster markers using Seurat's FindAllMarkers function with default parameters. Evaluation of the cluster markers (Supplementary Data 8) showed one cluster with high SN neuronal genes (SNpc cluster) and two white matter clusters (WM1 and WM2), which were merged as "Surrounding tissue".

T cell factor scores in spatial transcriptomics

To compare T cell factor scores and snRNAseq-derived Astrocyte DEGs between PD and control across the two major spatial regions, Nigra and Surrounding Tissue (Fig. 6c), while accounting for variability between individual samples, a Linear Mixed-Effects Model (LMM) approach was utilized via the lme4 and lmerTest packages in R. Separate models were fit for astrocyte DEG (increased in PD only), T_cell_factor1, and T_cell_factor5 as response variables with condition, region, and their interaction (Condition * Region) as fixed effects. The sample variable was included as a random intercept effect (1 | Sample) to address the non-independence of measurements from the same individual. P-values for the overall fixed effects (main effects and interaction) were obtained using the Kenward-Roger approximation for degrees of freedom provided by lmerTest. To get p-values for the specific comparisons of interest (PD vs. control within each region), post-hoc tests were conducted on the Estimated Marginal Means (EMMs) derived from the LMMs using the emmeans package, by calling the pairs() function on the EMMs from condition by region (specs = ~ Condition | Region). The p-values reported for these specific pairwise contrasts were generated using the Kenward-Roger method for degrees of freedom and adjusted using the Bonferroni method.

Immunohistochemistry, multiplex immunofluorescence, and histology

Paraffin-embedded formalin-fixed tissue sections of 5–7 micron thickness of the midbrain at the level of the red nucleus or of the anterior cingulate cortex were used. All immunostains were conducted on a Leica© Bond RXm automated stainer. For a description of antibodies, see Supplementary Data 1. For chromogenic DAB stains, a generic IHC protocol was employed per manufacturer protocols. Standard deparaffinization and rehydration steps preceded antigen retrieval in Leica ER2 (Cat. No. AR9640) antigen retrieval buffer for 10–20 min according to manufacturer protocols. Then, a peroxide block was applied for 10 minutes followed by three wash steps using bond wash solution (Cat. No. AR9590). A one-hour incubation in a blocking buffer in 10% donkey serum containing PBS-based buffer preceded primary antibody (the source and dilutions are found Supplementary Dataset 1) labeling for 1 h at ambient temperature. This was followed by three wash steps, after which the Post Primary was dispersed for 8 min, followed by three wash steps prior to the Polymer being dispersed for 8 minutes, followed by another three wash steps. The slides were then treated with deionized water for one minute prior to incubating in Mixed DAB refined for 10 min, followed by three washes of deionized water. Slides were stained with Hematoxylin for five minutes, followed by a wash with deionized water, then Bond wash solution, and lastly, a deionized water wash. For multiplexing immunostains using antibodies raised in non-overlapping hosts, we used a generic immunofluorescence protocol. Slides were baked in a 65 °C oven for a minimum of 2 h. The following protocol was then used: After a dewaxing step, incubation in BOND Epitope Retrieval Solution 2 (Cat. No. AR9640) for 20 min was used for heat-induced epitope retrieval. Next, the slides were washed in 1X PBS before washing twice in Bond Wash Solution (Cat. No. AR9590) – 10 min/wash. Next, they were incubated in 10% blocking donkey serum for 60 min, followed by the primary antibody diluted in a blocking buffer for 60 min. After three washes, the slides were incubated in the secondary antibody containing buffer for 60 min. After three washes, a DAPI containing mounting solution (Everbright TrueBlack Hardset Mounting Medium with DAPI, Cat. No. 23018) was used to label nuclei and quench autofluorescence prior to cover-slipping. A volume of 150 ml/slide was used for all steps. All steps were conducted at ambient temperature.

Brightfield images were acquired with a Leica Aperio LSM™ slide scanner under a 20X objective. All immunofluorescent images were acquired on the Leica Thunder imager DMi8. Images were acquired

using 20X or 40X oil immersion objectives on a Leica K5 camera. Leica Biosystems LAS X software was used for image capture. Tiles covering the cingulate and SN were taken and stitched. Leica Thunder instant computational clearing was used to remove out-of-focus light.

Image analysis and quantification

All image analysis was performed in QuPath 0.42[113]. All observers were blinded to experimental conditions. Annotations detailing the cingulate, peduncle, or SN were manually drawn. To detect cells, we used the "cell detection" function under the analysis menu, with DAPI as the Detection Channel. We modified the background threshold per image to eliminate non-specific detections. We then trained an object classifier to classify the detections for the different channels. Training data was created from each image to delineate cells that were positive for the specific antigens in question. One classifier per channel was trained by calling the "train object classifier" function with the following parameters: type = Random Trees, measurements = Cell: Channel X standard deviation, mean, max, and min measurements for the channel in question. To increase the accuracy of the classifier, additional training annotations were created on the image in question until the classification results matched the impression of the observer. Once a classifier was trained for each channel, "create composite classifier" was called to create a classifier consisting of multiple individual classifiers, one for each channel on the image. Classifiers were trained for each image separately. For the DAB stains, positive cell detection was used by detecting optical density sum to detect nuclei for CD8+ cells. An object classifier was again trained by using the "train object classifier" function, with the following parameters: type=Random Trees, measurements = all measurements, and selected classes= CD8+ and CD8-. The numbers of cells identified as CD8+ were then normalized by dividing by the area of the annotation in which the analysis was done. The same classifier-based approach was used to quantify CD68+ cell density in IHC stained sections, CD44 + GFAP+ or MT3 + GFAP+ astrocytes, and CD3+ cells in multiplex immunofluorescence sections. The numbers of CD103+ cells in the SN were counted manually by two board-certified neuropathologists (OAD, JEG). These counts were then divided by the area of the respective region.

To calculate the distance between astrocytes and CD3 + T cells, and CD68+ activated microglia and CD8 + T cells (Fig. 7), the x-y coordinates of the centroids of the classified objects were extracted, and the Euclidian distances between every T cell objects and all astrocyte objects or all microglia object centroids was measured. For each T cell, the minimum distance was recorded. The distances were then min-max normalized using values from all samples. The median minimal distances were compared between PD and control donors using an unpaired two-tailed Mann-Whitney test.

Cloning

pLV-EF1a-EGFP-N4-IRES-puro was derived from inserting EGFP sequence of pEGFP-C4 (Clontech) into pLV-EF1a-IRES-puro vector (Addgene, Cat#85132). pLKO2-EGFP-T2A-puro hH1 was derived from TRC2 pLKO-puro vector by introducing EGFP-T2A sequence between human PGK promotor and puromycin resistance cDNA and replacing human U6 promotor with human H1 promoter from pSUPER-puro (Oligoengine). The shRNA sequence for the control non-targeting shRNA (NC), sh mCD44-1, sh mCD44-2, sh mCD44-3, and sh mCD44-4 were obtained from Sigma©: SHC002, 5'- caacaagatgaagagcaccaa-3'; TRCN0000262948, 5'- ccaaccacacaggagtatata-3'; TRCN0000065357, 5'- ccgaattagctggacactcaa-3'; TRCN0000262945, 5'-gtgtagtgcctacgc-cattaa-3'; and TRCN0000262946, 5'- gataccttcatgtccatattt-3', respectively. These were inserted in pLKO2-EGFP-T2A-puro hH1 vector with AgeI and EcoRI restriction sites. All constructs were confirmed by DNA sequencing.

Lentiviral production and transduction

293 T cells (ATCC, CRL-3216) were maintained with Dulbecco's Modified Eagle's Medium (DMEM) with 10% (v/v) fetal bovine serum (Gemini Bio., 900-108-500) and 1% (v/v) Penicillin-Streptomycin (ThermoFisher Scientific, Cat#15070063) at 37 °C and 5% CO_2 incubator. 293 T cells were transfected with lentiviral vectors and the second-generation packaging plasmids (psPAX2, Addgene, Cat#12260, and pMD2.G, Addgene, Cat#12259) by Ca2PO4 transfection method. Medium containing the produced virus was harvested 24, 28, and 32 h after transfection. The 0.45 mm filtered virus medium was then concentrated by using ultracentrifugation with SW28 rotor (Beckman Coulter) at 113,000 x g for 1.5 h. The virus was added to mouse astrocyte culture in the presence of 2 mg/ml polybrene and incubated. On the next day, the cells were replated in the presence of 1 mg/ml puromycin to allow for selection. Knockdown of Cd44 was confirmed using immunofluorescence and western blotting.

Western blotting

For western blotting, cells were lysed with a lysis buffer (25 mM Tris-HCl pH = 7.4, 150 mM NaCl, 5 mM EDTA pH = 8.0, 1% (v/v) Triton X-100, 10 mM sodium pyrophosphate, 10 mM b-glycerophosphate, 10 mM NaF, 1 mM sodium orthovanadate, 10% (v/v) glycerol, 1 mM dithiothreitol, DTT, and 1 mM phenylmethylsulfonyl fluoride, PMSF). The lysate was incubated on ice for 30 minutes and spun at 29,800 x g for 15 min. The soluble fraction of lysate was mixed with 5X SDS sample buffer (312.5 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 8% (w/v) DTT, and 0.02% (w/v) bromophenol blue), and boiled. The proteins in the sample were resolved by SDS-PAGE and transferred to the nitrocellulose membrane (Cytiva, Cat#10600011). The proteins were probed with primary anti-CD44 (E7K2Y) XP antibody (Cell Signaling, Cat#37259) followed by fluorophore-conjugated secondary antibody donkey anti-rabbit IgG, AF 647 (Li-COR Inc, Cat# A32795). The signal was detected using Odyssey Imager (Li-COR Inc.).

Immunocytofluorescence and imaging

Cells were fixed with 4% (v/v) paraformaldehyde (Cat#15714, Electron Microscopy Sciences) in phosphate-buffered saline (PBS) for 20 min, and permeabilized and blocked with PBS containing 0.1 % (v/v) Triton-X and 1% (w/v) bovine serum albumin (BSA) (Sigma-Aldrich, Cat#7906) for 30 min. The cells were labeled first with primary CD44 monoclonal antibody (IM7) (Invitrogen, Cat#14-0441-82) and then fluorescently labeled secondary donkey anti-rat IgG, AF568 (Invitrogen, Cat#A78946) and DAPI. Images were acquired with 60× Apo TIRF Oil objective, NA 1.49, (Nikon) and Hamamatsu ORCA-Fusion BT camera on Nikon Eclipse Ti2-E microscope with CrestOptics X-Light V3 spinning disk confocal module controlled by Nikon's NIS-Elements software.

Bulk RNAseq of mouse astrocytes

Total RNA was extracted for sequencing in quantities ranging from 200 ng to 1000 ng, with RNA Integrity Numbers (RIN) of 9.9 to 10, as determined using a Bioanalyzer. Sequencing was performed on the AVITI platform, utilizing the Standard RNA-seq pipeline (STRDPOLYA). Poly-A pull-down was employed to enrich mRNAs from the total RNA samples, followed by library construction using Illumina TruSeq chemistry. The final PCR step was modified with KAPA HiFi HotStart Ready Mix. Libraries were sequenced on the Element AVITI system at the Columbia Genome Center. Samples were multiplexed in each lane, yielding 30 million reads of paired-end 75 bp reads per sample. Bases2fastq version 1.5.0.962525890 was used to convert BCL to fastq format and to trim the adaptor. Pseudoalignment to a kallisto[114] index created from Ensembl v96 transcriptomes (Human: GRCh38.p12; Mouse: GRCm38.p6) was performed using kallisto (version 0.44.0). Differential gene expression analysis was carried out using DESeq2[115]. Pathway activation scores (Fig. 8d) were measured using[116] the

run_mlm function in decoupleR on the list of differentially expressed genes with default parameters. Gene term network plots were generated in PathfindR[106] using the gene_term_graph function.

Statistical testing

All statistical analyses were conducted in GraphPad® Prism 10 or R version 4.2.2. One-tailed or two-tailed unpaired *t*-tests were used to compare PD vs control, as indicated in the figure legends. A one-tailed *t*-test was used when we had a prior hypothesis informed by the transcriptomic data. For comparing two groups, two-tailed *t*-tests were used, as seen in Fig. 1b, Fig. S1B, S1D, Fig. S9G, and S13B. A paired two-tailed *t*-test was used to compare T cell densities in the cingulate cortex between the white matter and the cortex with paired measurements from the same samples (Fig. S1E). In cases involving proportions, two-tailed Mann-Whitney tests were utilized, as in Figs. 4j, 5j. One-tailed *t*-tests were used in Fig. 3j, when a directional hypothesis was being tested informed by the literature or our transcriptional data. Linear models were used in Fig. 1c–f, h to account for confounding variables like age, sex, and Alzheimer's changes. For replicate data on the same samples (Fig. 1e, f), one-tailed *p*-values were retrieved. Linear mixed models were used in Fig. 6C to assess T cell Factor scores in ST data, with the sample being the random effect (see above). Wilcoxon tests were used in Fig. 3f to compare donor contributions to T cell Factors, and two-tailed *p*-values were reported. Kolmogorov-Smirnov tests were used to assess cumulative library contributions in Fig. 1g, and Mann-Whitney tests were used to compare gene program distributions in Fig. S10C and S10F. Gene set enrichment analysis (GSEA) was done using pre-ranked gene lists, as shown in Figs. 3g, 4k, and 5k (see methods section on GSEA). DEG analysis (Figs. 3e, 2g, h, 4g, 5g, 6d, and 8c) was done using limma or DESeq2 as indicated above in the respective sections on snRNAseq DEG, ST DEG, and bulk RNAseq DEG.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The files, both raw and processed, used for the analysis of human TCR sequencing, snRNAseq, and spatial transcriptomics datasets generated in this study have been deposited in the GEO database under accession codes: GSE253462 (snRNAseq), GSE253975 (ST), GSE253981 (TCR-seq). Processed snRNAseq and ST R objects are available embedded in our GitHub repository: https://github.com/dalhoomist/T-cell_and_glial_pathology_in_PD. All other datasets are provided in the Supplementary Data. Raw data are available upon request. Source data are provided with this paper.

Code availability

Code required for spatial cross-correlation analysis is provided here: https://github.com/dalhoomist/T-cell_and_glial_pathology_in_PD.

References

1. Dorsey, E. R. et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology* **68**, 384–386 (2007).
2. Gómez-Benito, M. et al. Modeling Parkinson's Disease With the Alpha-Synuclein Protein. *Front Pharm.* **11**, 356 (2020).
3. Pollanen, M. S., Dickson, D. W. & Bergeron, C. Pathology and Biology of the Lewy Body. *J. Neuropathol. Exp. Neurol.* **52**, 183–191 (1993).
4. Ntetsika, T., Papathoma, P.-E. & Markaki, I. Novel targeted therapies for Parkinson's disease. *Mol. Med* **27**, 17 (2021).
5. Brochard, V. et al. Infiltration of CD4+ lymphocytes into the brain contributes to neurodegeneration in a mouse model of Parkinson disease. *J. Clin. Invest.* **119**, 182–192 (2008).
6. 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).
7. Kamath, T. et al. Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. *Nat. Neurosci.* **25**, 588–595 (2022).
8. Martirosyan, A. et al. Unravelling cell type-specific responses to Parkinson's Disease at single cell resolution. *Mol. Neurodegeneration* **19**, 7 (2024).
9. Shwab, E. K. et al. Single-nucleus multi-omics of Parkinson's disease reveals a glutamatergic neuronal subtype susceptible to gene dysregulation via alteration of transcriptional networks. *Acta Neuropathol. Commun.* **12**, 111 (2024).
10. Smajic, S. et al. Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. *Brain* **145**, 964–978 (2022).
11. Singhania, A. et al. The TCR repertoire of alpha-synuclein-specific T cells in Parkinson's disease is surprisingly diverse. *Sci. Rep.* **11**, 302 (2021).
12. Frosch, M., Amann, L. & Prinz, M. CNS-associated macrophages shape the inflammatory response in a mouse model of Parkinson's disease. *Nat. Commun.* **14**, 3753 (2023).
13. Gate, D. et al. CD4(+) T cells contribute to neurodegeneration in Lewy body dementia. *Science* **374**, 868–874 (2021).
14. Hansen, D. V., Hanson, J. E. & Sheng, M. Microglia in Alzheimer's disease. *J. cell Biol.* **217**, 459–472 (2018).
15. Pajares, M., I. Rojo, A., Manda, G., Boscá, L. & Cuadrado, A. Inflammation in Parkinson's Disease: Mechanisms and Therapeutic Implications. *Cells* **9**, 1687 (2020).
16. Lazdon, E., Stolero, N. & Frenkel, D. Microglia and Parkinson's disease: footprints to pathology. *J. Neural Transm. (Vienna)* **127**, 149–158 (2020).
17. Zhang, X., Zhang, R., Nisa Awan, M. U. & Bai, J. The Mechanism and Function of Glia in Parkinson's Disease. *Front. Cell. Neurosci.* **16**, 903469 (2022).
18. Zhang, W. et al. Neuromelanin activates microglia and induces degeneration of dopaminergic neurons: implications for progression of Parkinson's disease. *Neurotox. Res* **19**, 63–72 (2011).
19. Lee, H. J. et al. Direct transfer of alpha-synuclein from neuron to astroglia causes inflammatory responses in synucleinopathies. *J. Biol. Chem.* **285**, 9262–9272 (2010).
20. Rostami, J. et al. Astrocytes have the capacity to act as antigen-presenting cells in the Parkinson's disease brain. *J. Neuroinflam.* **17**, 119 (2020).
21. Garretti, F. et al. Interaction of an α-synuclein epitope with HLA-DRB1∗15:01 triggers enteric features in mice reminiscent of prodromal Parkinson's disease. *Neuron* **111**, 3397–3413.e3395 (2023).
22. Allen Reish, H. E. & Standaert, D. G. Role of α-Synuclein in Inducing Innate and Adaptive Immunity in Parkinson Disease. *J. Parkinson's. Dis.* **5**, 1–19 (2015).
23. Williams, G. P. et al. CD4 T cells mediate brain inflammation and neurodegeneration in a mouse model of Parkinson's disease. *Brain* **144**, 2047–2059 (2021).
24. Schonhoff, A. M. et al. Border-associated macrophages mediate the neuroinflammatory response in an alpha-synuclein model of Parkinson disease. *Nat. Commun.* **14**, 3754 (2023).
25. Matheoud, D. et al. Parkinson's Disease-Related Proteins PINK1 and Parkin Repress Mitochondrial Antigen Presentation. *Cell* **166**, 314–327 (2016).
26. Williams G. P. et al. PINK1 is a target of T cell responses in Parkinson's disease. *J. Clin. Investig.* **135**, e180478 (2024).
27. Wang, P. et al. Single-cell transcriptome and TCR profiling reveal activated and expanded T cell populations in Parkinson's disease. *Cell Discov.* **7**, 52 (2021).

28. Hobson, R. et al. Clonal CD8 T cells in the leptomeninges are locally controlled and influence microglia in human neurodegeneration. *bioRxiv*, 2023.2007.2013.548931 (2023).

29. Galiano-Landeira, J., Torra, A., Vila, M. & Bove, J. CD8 T cell nigral infiltration precedes synucleinopathy in early stages of Parkinson's disease. *Brain* **143**, 3717–3733 (2020).

30. Baba, Y., Kuroiwa, A., Uitti, R. J., Wszolek, Z. K. & Yamada, T. Alterations of T-lymphocyte populations in Parkinson disease. *Parkinsonism Relat. Disord.* **11**, 493–498 (2005).

31. Bhatia, D. et al. T-cell dysregulation is associated with disease severity in Parkinson's Disease. *J. neuroinflam.* **18**, 250 (2021).

32. Dhanwani, R. et al. Transcriptional analysis of peripheral memory T cells reveals Parkinson's disease-specific gene signatures. *NPJ Parkinsons Dis.* **8**, 30 (2022).

33. Sun, C. et al. Abnormal subpopulations of peripheral blood lymphocytes are involved in Parkinson's disease. *Ann. Transl. Med.* **7**, 637 (2019).

34. Williams, G. P. et al. Unaltered T cell responses to common antigens in individuals with Parkinson's disease. *J. Neurol. Sci.* **444**, 120510 (2023).

35. Yan, Z. et al. Dysregulation of the Adaptive Immune System in Patients With Early-Stage Parkinson Disease. *Neurol. Neuroimmunol. Neuroinflam.* **8**, e1036 (2021).

36. Schröder, J. B. et al. Immune Cell Activation in the Cerebrospinal Fluid of Patients With Parkinson's Disease. *Front. Neurol.* **9**, 1081 (2018).

37. Iba, M. et al. Neuroinflammation is associated with infiltration of T cells in Lewy body disease and alpha-synuclein transgenic models. *J. Neuroinflam.* **17**, 214 (2020).

38. Smolders, J. et al. Tissue-resident memory T cells populate the human brain. *Nat. Commun.* **9**, 4593 (2018).

39. Williams, G. P. et al. T cell infiltration in both human multiple system atrophy and a novel mouse model of the disease. *Acta Neuropathologica* **139**, 855–874 (2020).

40. Chatila, Z. K. et al. RNA- and ATAC-sequencing Reveals a Unique CD83+ Microglial Population Focally Depleted in Parkinson's Disease. *bioRxiv*, (2023).

41. Welch, J. D. et al. Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* **177**, 1873–1887 e1817 (2019).

42. Zhu, B. et al. Single-cell transcriptomic and proteomic analysis of Parkinson's disease brains. *Sci. Transl. Med* **16**, eabo1997 (2024).

43. Booth, H. D. E., Hirst, W. D. & Wade-Martins, R. The Role of Astrocyte Dysfunction in Parkinson's Disease Pathogenesis. *Trends Neurosci.* **40**, 358–370 (2017).

44. Paryani, F. et al. Multi-omic analysis of Huntington's disease reveals a compensatory astrocyte state. *Nat. Commun.* **15**, 6742 (2024).

45. Hosoi, A. et al. Increased diversity with reduced "diversity evenness" of tumor infiltrating T-cells for the successful cancer immunotherapy. *Sci. Rep.* **8**, 1058 (2018).

46. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020).

47. Skinnider, M. A. et al. Cell type prioritization in single-cell data. *Nat. Biotechnol.* **39**, 30–34 (2021).

48. Al-Dalahmah, O. et al. Osteopontin drives neuroinflammation and cell loss in MAPT-N279K frontotemporal dementia patient neurons. *Cell Stem Cell.* **31**, 676–693 (2024).

49. Levitin, H. M. et al. De novo gene signature identification from single-cell RNA-seq with hierarchical Poisson factorization. *Mol. Syst. Biol.* **15**, e8557 (2019).

50. Wakim, L. M. et al. The Molecular Signature of Tissue Resident Memory CD8 T Cells Isolated from the Brain. *J. Immunol.* **189**, 3462–3471 (2012).

51. Kaech, S. M. et al. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nat. Immunol.* **4**, 1191–1198 (2003).

52. Lin, Y. H. et al. Small intestine and colon tissue-resident memory CD8(+) T cells exhibit molecular heterogeneity and differential dependence on Eomes. *Immunity* **56**, 207–223.e208 (2023).

53. Intlekofer, A. M. et al. Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin. *Nat. Immunol.* **6**, 1236–1244 (2005).

54. Li, G. et al. T-Bet and Eomes Regulate the Balance between the Effector/Central Memory T Cells versus Memory Stem Like T Cells. *PLoS one* **8**, e67401 (2013).

55. Fu, S. H., Yeh, L. T., Chu, C. C., Yen, B. L. & Sytwu, H. K. New insights into Blimp-1 in T lymphocytes: a divergent regulator of cell destiny and effector function. *J. Biomed. Sci.* **24**, 49 (2017).

56. Schenkel, J. M. et al. cell memory. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science* **346**, 98–101 (2014).

57. Grayson, J. M., Zajac, A. J., Altman, J. D. & Ahmed, R. Cutting edge: increased expression of Bcl-2 in antigen-specific memory CD8+ T cells. *J. Immunol.* **164**, 3950–3954 (2000).

58. Yang, Q. et al. Cutting Edge: Characterization of Human Tissue-Resident Memory T Cells at Different Infection Sites in Patients with Tuberculosis. *J. Immunol.* **204**, 2331–2336 (2020).

59. Gruber, T. et al. IL-32gamma potentiates tumor immunity in melanoma. *JCI Insight* **5**, e138772 (2020).

60. Pontes Ferreira, C. et al. CXCR3 chemokine receptor contributes to specific CD8+ T cell activation by pDC during infection with intracellular pathogens. *PLoS Negl. Trop. Dis.* **14**, e0008414 (2020).

61. van Spriel, A. B. et al. A regulatory role for CD37 in T cell proliferation. *J. Immunol.* **172**, 2953–2961 (2004).

62. Baaten, B. J., Li, C. R. & Bradley, L. M. Multifaceted regulation of T cells by CD44. *Commun. Integr. Biol.* **3**, 508–512 (2010).

63. Kumar, B. V. et al. Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep.* **20**, 2921–2934 (2017).

64. Topham, D. J. & Reilly, E. C. Tissue-Resident Memory CD8+ T Cells: From Phenotype to Function. *Front. Immunol.* **9**, 515 (2018).

65. Crowl, J. T. et al. Tissue-resident memory CD8+ T cells possess unique transcriptional, epigenetic and functional adaptations to different tissue environments. *Nat. Immunol.* **23**, 1121–1131 (2022).

66. Anadon, C. M. et al. Ovarian cancer immunogenicity is governed by a narrow subset of progenitor tissue-resident memory T cells. *Cancer Cell* **40**, 545–557.e513 (2022).

67. Gueguen, P. et al. Contribution of resident and circulating precursors to tumor-infiltrating CD8 + T cell populations in lung cancer. *Sci. Immunol.* **6**, eabd5778 (2021).

68. Szabo, P. A. et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, 4706 (2019).

69. Cho, B. K., Wang, C., Sugawa, S., Eisen, H. N. & Chen, J. Functional differences between memory and naive CD8 T cells. *Proc. Natl. Acad. Sci. USA* **96**, 2976–2981 (1999).

70. Berard, M. & Tough, D. F. Qualitative differences between naïve and memory T cells. *Immunology* **106**, 127–138 (2002).

71. Mackay, L. K. et al. The developmental pathway for CD103+CD8+ tissue-resident memory T cells of skin. *Nat. Immunol.* **14**, 1294–1301 (2013).

72. Thrupp, N. et al. Single-Nucleus RNA-Seq Is Not Suitable for Detection of Microglial Activation Genes in Humans. *Cell Rep.* **32**, 108189 (2020).

73. Olah, M. et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nat. Commun.* **11**, 6129 (2020).

74. Song, Y. J. C. et al. Degeneration in Different Parkinsonian Syndromes Relates to Astrocyte Type and Astrocyte Protein Expression. *J. Neuropathol. Exp. Neurol.* **68**, 1073–1083 (2009).

75. Tong, J. et al. Low levels of astroglial markers in Parkinson's disease: relationship to alpha-synuclein accumulation. *Neurobiol. Dis.* **82**, 243–253 (2015).

76. Al-Dalahmah, O. et al. The Matrix Receptor CD44 Is Present in Astrocytes throughout the Human Central Nervous System and Accumulates in Hypoxia and Seizures. *Cells* **13**, 129 (2024).

77. Al-Dalahmah, O. et al. Re-convolving the compositional landscape of primary and recurrent glioblastoma reveals prognostic and targetable tissue states. *Nat. Commun.* **14**, 2586 (2023).

78. O'Callaghan, J. P., Kelly, K. A., VanGilder, R. L., Sofroniew, M. V. & Miller, D. B. Early activation of STAT3 regulates reactive astrogliosis induced by diverse forms of neurotoxicity. *PloS one* **9**, e102003 (2014).

79. Hartmann, K. et al. Complement 3(+)-astrocytes are highly abundant in prion diseases, but their abolishment led to an accelerated disease course and early dysregulation of microglia. *Acta Neuropathol. Commun.* **7**, 83 (2019).

80. Al-Dalahmah, O. et al. Single-nucleus RNA-seq identifies Huntington disease astrocyte states. *Acta Neuropathol. Commun.* **8**, 19 (2020).

81. Van den Berge, K. et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).

82. Zhao, E. et al. Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* **39**, 1375–1384 (2021).

83. Walzer, T., Arpin, C., Beloeil, L. & Marvel, J. Differential in vivo persistence of two subsets of memory phenotype CD8 T cells defined by CD44 and CD122 expression levels. *J. Immunol.* **168**, 2704–2711 (2002).

84. Lee, J. L., Wang, M. J. & Chen, J. Y. Acetylation and activation of STAT3 mediated by nuclear translocation of CD44. *J. Cell Biol.* **185**, 949–957 (2009).

85. Harms, A. S. et al. MHCII Is Required for -Synuclein-Induced Activation of Microglia, CD4 T Cell Proliferation, and Dopaminergic Neurodegeneration. *J. Neurosci.* **33**, 9592–9600 (2013).

86. Beechem, J. M. High-Plex Spatially Resolved RNA and Protein Detection Using Digital Spatial Profiling: A Technology Designed for Immuno-oncology Biomarker Discovery and Translational Research. *Methods Mol. Biol.* **2055**, 563–583 (2020).

87. Yang, Y. et al. Distinct mechanisms define murine B cell lineage immunoglobulin heavy chain (IgH) repertoires. *eLife* **4**, e09083 (2015).

88. Wang, C. et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc. Natl. Acad. Sci. USA* **107**, 1518–1523 (2010).

89. Sims, J. S. et al. Diversity and divergence of the glioma-infiltrating T-cell receptor repertoire. *Proc. Natl. Acad. Sci. USA* **113**, E3529–E3537 (2016).

90. DeLuca, D. S. et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530–1532 (2012).

91. Petren, H., Kollner, T. G. & Junker, R. R. Quantifying chemodiversity considering biochemical and structural properties of compounds with the R package chemodiv. *N. Phytol.* **237**, 2478–2492 (2023).

92. Chiou, S. H. et al. Global analysis of shared T cell specificities in human non-small cell lung cancer enables HLA inference and antigen discovery. *Immunity* **54**, 586–602.e588 (2021).

93. Wang, M., Zhao, Y. & Zhang, B. Efficient Test and Visualization of Multi-Set Intersections. *Sci. Rep.* **5**, 16923 (2015).

94. Lim, R. G. et al. Huntington disease oligodendrocyte maturation deficits revealed by single-nucleus RNAseq are rescued by thiamine-biotin supplementation. *Nat. Commun.* **13**, 7791 (2022).

95. Yang, S. et al. Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol.* **21**, 57 (2020).

96. Fleming, S. J. et al. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* **20**, 1323–1335 (2023).

97. Germain, P. L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet identification in single-cell sequencing data using scDblFinder *F1000Res.* **10**, 979 (2021).

98. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

99. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).

100. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

101. Zappia L., Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, giy083 (2018).

102. Xie, Z. et al. Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).

103. Lin, X., Chau, C., Ma, K., Huang, Y. & Ho, J. W. K. DCATS: differential composition analysis for flexible single-cell experimental designs. *Genome Biol.* **24**, 151 (2023).

104. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

105. Korotkevich, G. et al. Fast gene set enrichment analysis. *BioRxiv*, https://doi.org/10.1101/060012 (2021).

106. Ulgen, E., Ozisik, O. & Sezerman, O. U. pathfindR: An R Package for Comprehensive Identification of Enriched Pathways in Omics Data Through Active Subnetworks.*Front Genet.* **10**, 858 (2019).

107. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).

108. Qu, R. et al. Gene trajectory inference for single-cell data by optimal transport metrics. *Nat. Biotechnol.* **43**, 258–268 (2025).

109. Linderman, G. C. et al. Zero-preserving imputation of single-cell RNA-seq data. *Nat. Commun.* **13**, 192 (2022).

110. Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).

111. Chen, Y. A New Methodology of Spatial Cross-Correlation Analysis. *PloS one* **10**, e0126158 (2015).

112. Miller, B. F., Bambah-Mukku, D., Dulac, C., Zhuang, X. & Fan, J. Characterizing spatial gene expression heterogeneity in spatially resolved single-cell transcriptomic data with nonuniform cellular densities. *Genome Res.* **31**, 1843–1855 (2021).

113. Bankhead, P. et al. QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).

114. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

115. 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).

116. Badia, I. M. P. et al. decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv.* **2**, vbac016 (2022).

Acknowledgements

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at
https://doi.org/10.1038/s41467-025-62478-3.

Correspondence and requests for materials should be addressed to Osama Al-Dalahmah.

Peer review information *Nature Communications* thanks Michael Henderson and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at
http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.