

A single-cell atlas to map sex-specific gene-expression changes in blood upon neurodegeneration

Received: 17 June 2024

Accepted: 29 January 2025

Published online: 25 February 2025

Check for updates

Friederike Grandke¹, Tobias Fehlmann¹, Fabian Kern^{1,2}, David M. Gate^{3,4,5,6}, Tobias William Wolff¹, Olivia Leventhal³, Divya Channappa³, Pascal Hirsch¹, Edward N. Wilson³, Eckart Meese⁷, Chuanyu Liu⁸, Quan Shi⁸, Matthias Flotho¹, Yongping Li^{1,8}, Cynthia Chen⁸, Yeya Yu⁸, Jiangshan Xu⁸, Michael Junkin⁸, Zhifeng Wang⁸, Tao Wu⁸, Longqi Liu⁸, Yong Hou⁸, Katrin I. Andreasson^{3,5,9}, Jenny S. Gansen¹, Elvira Mass¹⁰, Kathleen Poston³, Tony Wyss-Coray^{3,4,5,6,11,13} ✉ & Andreas Keller^{1,2,3,12,13} ✉

The clinical course and treatment of neurodegenerative disease are complicated by immune-system interference and chronic inflammatory processes, which remain incompletely understood. Mapping immune signatures in larger human cohorts through single-cell gene expression profiling supports our understanding of observed peripheral changes in neurodegeneration. Here, we employ single-cell gene expression profiling of over 909k peripheral blood mononuclear cells (PBMCs) from 121 healthy individuals, 48 patients with mild cognitive impairment (MCI), 46 with Parkinson's disease (PD), 27 with Alzheimer's disease (AD), and 15 with both PD and MCI. The dataset is interactively accessible through a freely available website (<https://www.ccb.uni-saarland.de/adrcsc>). In this work, we identify disease-associated changes in blood cell type composition and the gene expression in a sex-specific manner, offering insights into peripheral and solid tissue signatures in AD and PD.

Neurodegenerative disorders such as Alzheimer's disease (AD) and Parkinson's disease (PD) show increasing rates of prevalence in the global population and severely affect the quality of life of the elderly^{1,2}. Common risk factors include the genetic background, personal lifestyle, environmental exposure and age³. Although the clinical manifestation, diagnosis and course of progression vary substantially between AD and PD, underlying common and converging cellular

pathways have been proposed⁴. Thereby, both diseases can arise from distinct precursor stages of varying duration and co-occur with certain shared symptoms, including (mild) cognitive decline^{5,6}. Notably, the interplay of the adaptive immune system and brain inflammation in neurodegenerative diseases is increasingly appreciated^{7,8}. In this context, substantial efforts have carved out the primary affected central nervous system (CNS) cell types and key risk genes driving the

¹Clinical Bioinformatics, Saarland University, 66123 Saarbrücken, Germany. ²Helmholtz Institute for Pharmaceutical Research Saarland (HIPS)–Helmholtz Centre for Infection Research (HZI), Saarland University Campus, Saarbrücken, Germany. ³Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA 94305, USA. ⁴Veterans Administration Palo Alto Healthcare System, Palo Alto, CA, USA. ⁵Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA. ⁶Chemistry, Engineering, and Medicine for Human Health, Stanford University, Stanford, CA, USA. ⁷Department of Human Genetics, Saarland University, 66421 Homburg/Saar, Germany. ⁸MGI Group, San Jose, CA, USA. ⁹Program in Immunology, Stanford University, Stanford, CA, USA. ¹⁰Life and Medical Sciences Institute, Developmental Biology of the Immune System, University of Bonn, Bonn, Germany. ¹¹The Phil and Penny Knight Initiative for Brain Resilience, Stanford University, Stanford, CA, USA. ¹²PharmaScienceHub, Saarland University Campus, Saarbrücken, Germany. ¹³These authors jointly supervised this work: Tony Wyss-Coray, Andreas Keller. ✉ e-mail: twc@stanford.edu; andreas.keller@ccb.uni-saarland.de

underlying pathology, e.g., *APOE* and *TREM2* in microglia for AD, or alpha-synuclein (*SNCA*) contained in Lewy-bodies in dopaminergic neurons for PD^{9–13}. In a similar direction, single cell technologies facilitate to uncover the role of the human brain vasculature for the development of neurodegenerative disorders^{14,15}.

While the brain regions selectively affected in these diseases are subject to intense but predominantly small-scale research due to the apparent difficulties to obtain fresh-frozen human tissue samples and associated costs, research has focused on model organisms e.g., genetically modified mice, in vitro cell systems, or on the in human better accessible peripheral organic system, for example blood or cerebrospinal fluid (CSF) and their connecting interfaces such as the blood-brain barrier (BBB)^{16–18}. Indeed, accumulating evidence points to a partial breakdown of the BBB with age and during age-related diseases, suggesting a still underestimated role of the peripheral immune system in neurodegenerative conditions, potentially through information exchange between these otherwise restrictively isolated physiological environments^{19–22}. These findings raise questions about how and at which point during the disease progression the peripheral immune system is implicated, especially in the context of driving systemic inflammation^{23–26}. One key factor which is increasingly considered in the pathology of AD and PD is the patients' biological sex. Single-cell studies for AD for instance suggest that female cells are overrepresented in disease-associated subpopulations and that transcriptional responses are different between sexes in oligodendrocytes and other cell types²¹. Sex-dependent changes in brain and the blood of AD patients have been repeatedly reported (Supplementary Table 1). In contrast, sex-dependent changes in PD patients in the blood have so far only been reported in monocytes²⁷.

Broadly accessible and low-invasively collectable plasma samples offer great promise for the detection of circulating biomarkers in neurodegeneration through a screening of peripheral blood mononuclear cells (PBMC), especially in large-scale settings^{28–32}. Substantial efforts on clinical cohorts showed that plasma markers measured with analytical methods can be used to partially predict or track neurodegenerative disease development while bearing also its own challenges^{33–37}. Thus, a next logical step is to comprehensively map molecular pathways correlating with, or even causative for these disease biomarker signatures using high-throughput techniques at high resolution. Single-cell RNA sequencing (scRNA-seq) holds the promise of unraveling cell type-specific homeostatic conditions and their alterations in human disease, as recently demonstrated for parenchymal brain tissue, brain vasculature, or PBMCs in neurodegenerative diseases^{11,12,14,15,38–44}. Reaching numbers of up to one million cellular gene expression profiles per publication, the amount of transcriptomic data generated ranges on exponential scales and is expected to further increase^{45–48}.

In this work, we sought to characterize the peripheral response through the transcriptional landscape⁴⁹ in neurodegeneration by profiling over 909k PBMCs of 290 blood samples including 155 samples of individuals with clinical neurodegeneration compared to 135 samples of healthy controls (Fig. 1a, Supplementary Data 1). We thus surpass all other currently published datasets of RNA-sequencing data of PBMCs in Alzheimer's and Parkinson's disease in both the number of patients and the number of cells (Supplementary Fig. 1, Supplementary Table 2). Our dataset offers insights into the general sensing of strongly CNS-associated diseases in the periphery, for the first time evaluating the potential of low-invasive cellular biomarkers for these, both at single-cell resolution and at scale. By including time-series follow-ups for a subset of the patients, as well as additional gray and white matter volume measurements and protein-levels of known AD-disease markers in the CSF, we combine a variety of patient-centric data. A companion web browser is available at <https://www.ccb.uni-saarland.de/adrcsc> and enables researchers to easily access the pre-processed dataset without requiring technical expertise and simplifies the testing of data-driven hypotheses.

Results

A scRNA-seq dataset of PBMCs in neurodegenerative diseases

From the Stanford Alzheimer's Disease Research Center (ADRC) we selected patients diagnosed with mild cognitive impairment (individual patients $n = 48$, total samples from the patients $n_t = 55$; MCI), Alzheimer's disease ($n = 27$, $n_t = 34$; AD), Parkinson's disease ($n = 46$, $n_t = 48$; PD), PD with MCI ($n = 15$, $n_t = 18$; PD-MCI), and healthy individuals ($n = 121$, $n_t = 135$; HC) (Fig. 1b, Table 1). We ensured a well-balanced ratio of both biological sex and age between the subgroups (Supplementary Fig. 2a). For each sample, we performed droplet-based single-cell RNA-sequencing⁴⁹ on collected PBMCs, totaling 1,374,714 cellular expression profiles. We then kept the top 909,322 single cells of highest quality (66%) that passed stringent quality control filtering (Supplementary Fig. 2b). Following a cluster-based cell type annotation and manual marker curation, we found 13 major PBMC types that were further divided into 33 fine-granular sub-cell types (Fig. 1c, d, Supplementary Fig. 3). The downstream analysis described in this paper was performed using only the baseline sample of each patient (unless stated otherwise).

Sex-specific changes in cell type proportions

As visualizing those large cell numbers in a 2-dimensional embedding becomes unclear - a phenomenon commonly denoted as overplotting - we estimated the density of cells to highlight the distribution of cells for each diagnosis group (Supplementary Fig. 4a). Interpreting the projected densities, we found slight differences in the distribution of cells, suggesting a disease-specific transcriptomics shift in the overall abundance of different cell type populations. In light of previous reports indicating that changes in cell type proportions in PBMC samples may differ between males and females⁵⁰, we separated the data by sex to independently examine disease-related changes. In general, we observed distinct changes in the cell type proportions in men and women (Supplementary Fig. 4b). A density analysis depicts these effects in the integrated data space (Fig. 2a).

These sex-dependent changes were confirmed using scCODA, which did not show significant changes in the cell-type distributions ($\text{abs log2FC} > 0.35$, significant according to scCODA) when leaving out sex as a covariate (Supplementary Fig. 4c). Split by male and female patients, we were able to observe significant changes (Fig. 2b), some of which were different dependent on the sex. In Parkinson's disease, CD8 + T cells and Plasma cells show a positive fold-change in males but a negative fold-change in females. Similarly, B cells are more abundant in females with PD-MCI but less abundant in males with PD-MCI. Although a sex-dependent change in B and NK cells of AD patients was indicated in the density embedding, scCODA did not show a significant difference. In the finer cell-annotation, we observed significant changes in only 3 cases in males and 9 cases in females using the raw cell type proportions (Fig. 2c). The significant changes in AD, MCI and PD match the ones found in the female patients (Supplementary Fig. 4d). Using scCODA on the finer Annotation, we found a lower proportion of T_{H} cells in AD, as previously reported (Fig. 2d). When comparing these findings with previously reported changes, some were previously described by others, such as the increased proportion of Monocytes in AD, and changes in B cells in PD (Fig. 2e).

Several but not all of the changes we report in this study have been previously described by others, such as the increased proportion of B cells in AD, and changes in CD8 + T cells in PD. Overall, our results suggest a significant shift in the cellular composition of male and female patients, with disease-specific differences. An important question to consider is whether these changes are driven by small, distinct sets of genes, or broader transcriptomic perturbations.

Differences in the sex-specific transcriptome

To gain further insight into the respective molecular changes among male and female patients with MCI, AD, or PD, we conducted a DE analysis (Supplementary Data 2). As the sample sizes of PD-MCI is too

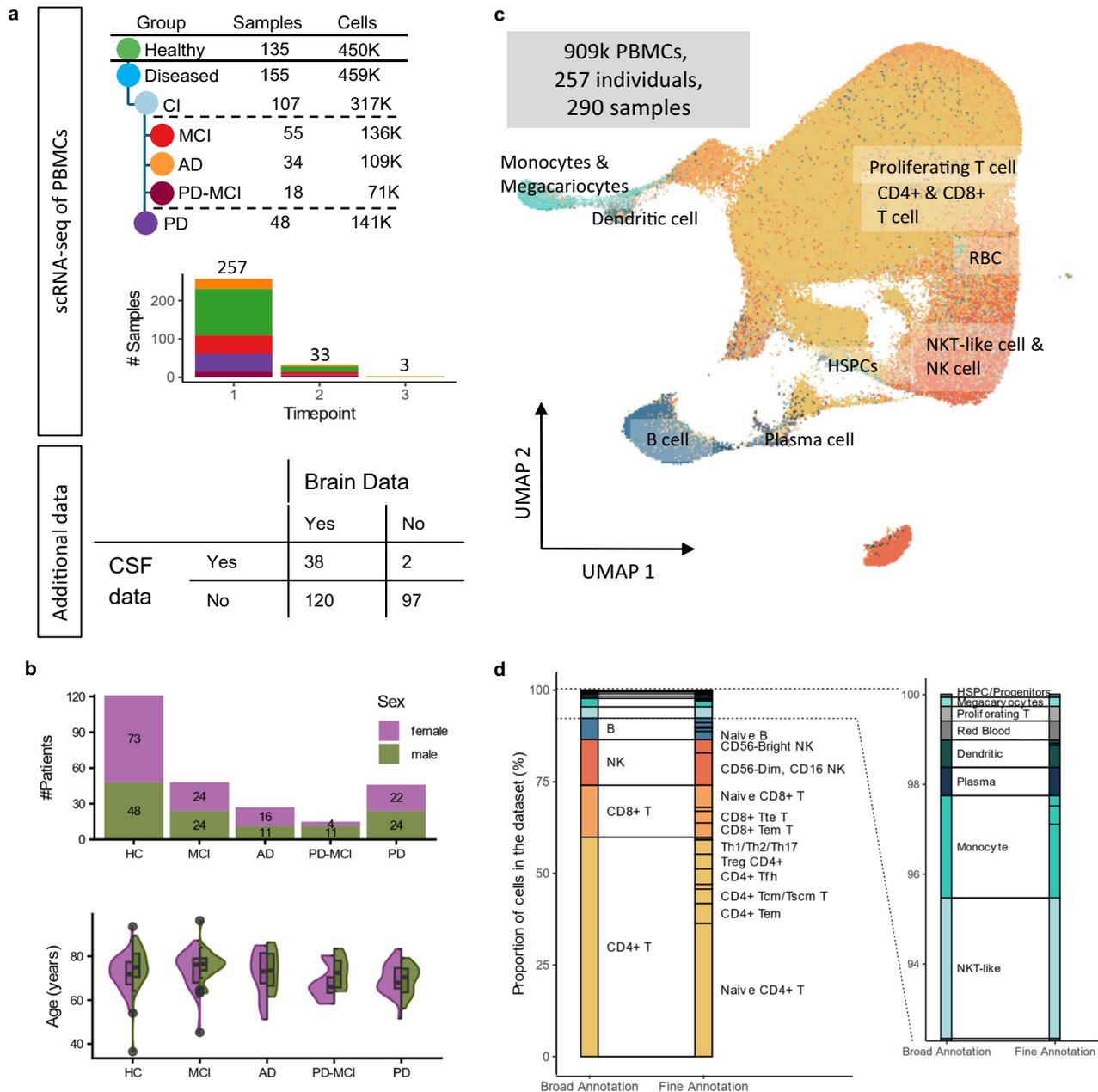


Fig. 1 | Cohort and single-cell data characterization. **a** Overview on the dataset. PBMC samples were collected from healthy patients (HC), patients with Parkinson's Disease (PD), and patients with Cognitive Impairment (CI), more detailed with Mild Cognitive Impairment (MCI), Alzheimer's Disease (AD), and Parkinson's Disease with Mild Cognitive Impairment (PD-MCI). For a subset of patients, blood samples were drawn at multiple time-points. The samples were then profiled using single-cell RNA sequencing. For a subset of patients, additional data was collected on brain volumes (158 patients) and known Alzheimer biomarkers were measured in the CSF (40 patients). **b** Distribution of biological sex and age per patient group (divided by

sex). The violin plots show the shape of the distribution, while the boxes encompass the first through third quartiles, with the central line marking the median. The whiskers mark the minimum or maximum values or, if outliers are present, the values within 1.5 times the interquartile range of the first or third quartile. Outliers are shown as dots. **c** Two-dimensional representation of the $n = 909k$ cells using Uniform Manifold Approximation and Projection (UMAP). Cells are colored by assigned cell type. **d** Cells are annotated on different levels. This allows the analysis on a broad level with 13 bigger clusters or on a finer level with 33 clusters.

small for most analysis we excluded it from further in-depth consideration. Our analysis aims to identify both, sex- and disease-related differences in gene expression. We revealed a stark contrast between sexes in the number of DEGs, particularly in Alzheimer's patients, which has been previously noted by Mathys et al.¹¹ (Fig. 3a). Specifically, we detected de-regulated genes in only 12 cell types in male patients, whereas in female patients 20 cell types exhibited de-regulated genes. A similar, albeit less sex-pronounced trend was observed in the comparison between PD and HC, as well as MCI vs. HC patients. We then compared the fold-changes in male and female subgroups and found a positive correlation for MCI vs.

HC (overall correlation of 0.35, p -value $< 2.2 \times 10^{-16}$) and AD vs. HC (overall correlation of 0.32, p -value $< 2.2 \times 10^{-16}$), particularly in cell types with a high number of significantly de-regulated genes (Fig. 3b). Conversely, cell types in PD vs. HC exhibited lower or no correlation at all that are often negative (overall correlation = -0.0064 , p -value = 0.098). Overall, our analysis revealed that most of the cell types in the PD vs. HC comparison were either negatively (9 cell types) correlated or not significantly correlated while most cell types exhibited a positive correlation in both AD (22 cell types) and MCI (22 cell types) (Fig. 3c). In total, we identified only 84 genes that were de-regulated independent of the sex.

Table 1 | Demographic breakdown of the patients included in this study

		Age in females (mean, sd and number of patients)			Age in males (mean, sd and number of patients)			Adj. P-value (m vs. f)		
Diagnosis	HC	72 ± 7 (n = 82)			75 ± 9 (n = 53)			0.318		
	AD	71 ± 10 (n = 21)			73 ± 8 (n = 13)			1		
	PD	69 ± 7 (n = 23)			69 ± 6 (n = 25)			1		
	MCI	73 ± 8 (n = 29)			76 ± 6 (n = 26)			1		
	PD-MCI	68 ± 7 (n = 6)			73 ± 6 (n = 12)			1		
Race	White (w)	71 ± 8 (n = 144)			74 ± 7 (n = 116)			0.101		
	Asian (a)	72 ± 7 (n = 12)			72 ± 14 (n = 9)			1		
	Other (o)	73 ± 7 (n = 5)			74 ± 7 (n = 4)			1		
	Race	w	a	o	w	a	o	w	a	o
Diagnosis	HC	72 ± 7 (n = 72)	71 ± 6 (n = 6)	71 ± 6 (n = 4)	76 ± 7 (n = 49)	64 ± 20 (n = 4)	- (n = 0)	0.148	1	-
	AD	70 ± 10 (n = 18)	75 ± 8 (n = 3)	- (n = 0)	73 ± 8 (n = 13)	- (n = 0)	- (n = 0)	1	-	-
	PD	69 ± 7 (n = 22)	67 - (n = 1)	- (n = 0)	68 ± 6 (n = 23)	73 ± 4 (n = 2)	- (n = 0)	1	-	-
	MCI	73 ± 8 (n = 26)	69 ± 11 (n = 2)	83 - (n = 1)	75 ± 6 (n = 20)	83 ± 1 (n = 2)	74 ± 7 (n = 4)	1	-	-
	PD-MCI	68 ± 7 (n = 6)	- (n = 0)	- (n = 0)	72 ± 7 (n = 11)	78 - (n = 1)	- (n = 0)	1	-	-

More detailed information can be found in Supplementary Data 1.

Most of these genes were found in CD4⁺ T_{cm}/T_{scm} T cells, NKT-like cells (Fig. 3d) or in CD56-Dim, CD16 NK cell. These findings lead us to evaluate if the similarities between the different diseases varies by sex. We thus determined the overlap between the significantly de-regulated genes in the three diseases to control comparisons. The distribution of shared genes between diseases in males (Fig. 3e) compared to those in females (Fig. 3f) is similar and significantly de-regulated genes are mainly shared between MCI vs. HC and PD vs. HC. In particular, 35 genes are differentially expressed in all three comparisons for male and 27 for female patients. Overall, our results thus suggest rather sex-independent expression patterns in the different diseases, while in certain cases - depending on the disease and cell type - the signatures for the sexes differ substantially. When comparing the gene-expression changes in the different cell-types, we observed that the changes in sub-cell types from the same cell type show a high similarity (Supplementary Fig. 5).

Peripheral pathways affected in neurodegeneration

In the light of the significant variability in gene expression patterns among diseases, cell types, and sexes, we sought to verify that signatures in our dataset reflect the dysregulation of known disease gene markers. To accomplish this, we specifically analyzed genes from the KEGG Alzheimer's disease and Parkinson's disease pathways that were de-regulated in our PBMCs. Astonishingly, we found that gene expression in the KEGG Alzheimer's disease pathway is largely independent of both sex and cell type. This might be due to a bias in how the genes on the pathways are identified in previous studies, largely excluding sex specific effects. While the fold-changes in some cell types differ between sexes, the direction of the de-regulation is consistent (Supplementary Fig. 6a). Conversely, we observed differences in gene expression between males and females and between different cell types for some genes in the KEGG Parkinson's disease pathway (Supplementary Fig. 6b).

This poses the question, which pathways are affected specifically in males and females. The results of a pathway analysis of the change in all cell-types and sexes (Supplementary Data 3) shows that the affected pathways in AD and in MCI are similar in both sexes (Fig. 4a). In PD the pathways are more sex specific. The overall most frequently enriched pathway in AD is "SRP-dependent cotranslational protein targeting to membrane". This is also the most frequently enriched pathway in females MCI and males with PD (Fig. 4b). In PD, the "Asthma" pathway is most affected in females and in MCI "Oxidative phosphorylation" in males.

In PD, MCI and in males in AD, the pathways are similar across all cell-types (Supplementary Fig. 7). In females with AD, B cells show a similar

cluster of pathways consisting of pathways related to junctions (adherens junction, cell-substrate junction, cell-substrate adherens junction) and pathways related to catabolic and biosynthetic processes ("amide biosynthetic process", "aromatic compound catabolic process",...). In males, these pathways appear across all groups of cell-types.

Overall, there are 15 pathways shared between PD, MCI and AD in females, but 53 in males (Fig. 4c, d). 114 pathways are shared between AD and PD in males and 24 in females.

Similar disease signatures in PBMCs and neural cells

Having determined changes that can be observed between Alzheimer's disease and healthy controls in PBMC samples, we sought to contrast peripheral cell transcriptomes to those from the brain. A large number of studies describe single cell gene expression across multiple regions in the brain with respect to different disease phenotypes and physiological conditions, complicating a joint analysis.

We thus used a database that integrates single cell gene expression data from 21 studies, allowing for precise insights into the regulation of genes across cell types in relation to age, sex, and disease (Fig. 5a)⁵¹. As there are not enough cells from the same cell-type in both brain and blood, we rely on overlaps of gene-expression patterns and pathways when looking for similar patterns in neurodegenerative diseases.

We identified genes that exhibit differential expression patterns between Alzheimer's disease and healthy individuals in male and female patients for the PBMCs and investigated whether these exhibit similar expression patterns in brain cells of Alzheimer's patients (Supplementary Data 4, Supplementary Fig. 8a). Our findings suggest that gene expression changes in peripheral cells may not necessarily reflect those in the brain, given the natural disparity of cell types between blood and brain. However, emphasizing the need for further in-depth investigation knowing the role of the immune system in Alzheimer's, the associated brain inflammation and infiltration by specifically activated immune cells.

Overall, there were 36 genes in males and 7 genes in females that were significantly de-regulated in both PBMCs and the brain in the main cell-types (Fig. 5b, c). The genes in males (adj. P-value < 0.05 and abs. log₂ Fold-change > 0.5 in both datasets) were enriched for pathways related to the regulation of the immune system and the membrane (Supplementary Fig. 8b, c). The female genes (adj. P-value < 0.05 in both datasets) were enriched for the Herpes Simplex Virus 1 pathway. Herpes simplex has previously been linked to an increased risk for the development of Alzheimer's and is discussed as a key factor in disease development^{52,53}.

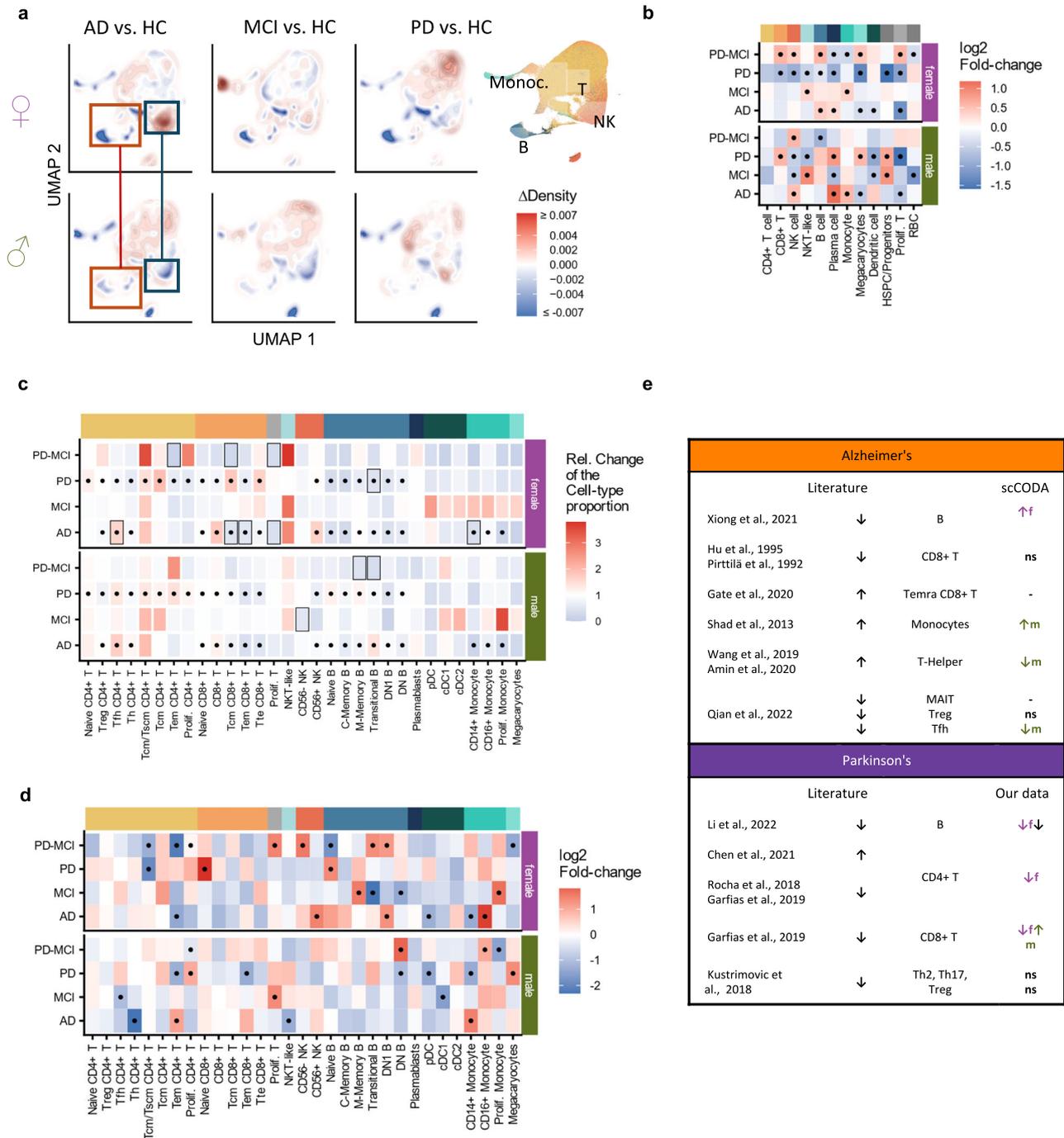


Fig. 2 | Sex-dependent changes of cell type proportions. **a** Differential density UMAP embedding for male and female patients shows differences in the cell type distribution for AD, PD and MCI. Each embedding shows the spots with the highest density of the diseased patient group (red) or the healthy control group (blue). **b** Comparison of the cell type proportions (on the broad annotation) using scCODA with Age and ApoE as covariates. Celltypes with an absolute Fold-change of at least 0.5 that were reported as significant by scCODA are marked with a dot. **c** Relative change in the cell type-proportion in % between Healthy and the diagnosis groups for the different cell types reveals sex-specific changes. Significant values (adj. p-value < 0.05, unpaired Student's t-test) are marked by a frame, previously described changes from the literature are marked with a dot. **d** Comparison of the changes of cell type proportions found in the literature and the cell type proportion changes found in our PBMC data using scCODA. Changes are only considered if they were reported as significant by scCODA between healthy and disease. If no sex is given, the test was performed without considering the sex. **e** Comparison of the cell type proportions on the fine Annotation using scCODA as described in (b). Celltypes with an absolute Fold-change of at least 1 that were reported as significant by scCODA are marked with a dot.

To further validate these findings and to see if the expression changes with different brain-regions, we tested for the general overlap of differentially expressed genes (abs. \log_2 Fold-change > 0.5 and adj. P-value < 0.05) in this PBMC dataset, the ZEBRA dataset (21 studies) and an additional ROSMAP dataset (4 studies; 6 brain regions) (Supplementary Data 5). Of note, two studies from the prefrontal cortex

samples of the ROSMAP dataset are a subset of the ZEBRA dataset and thus not independent. The other brain-regions are independent between the datasets. We found an overlap of 32 genes in the male patients and 8 genes in the female patients. 24 out of the 32 male genes and 5 out of the 8 female genes have been previously reported in the context of Alzheimer's disease, most in the context of the immune

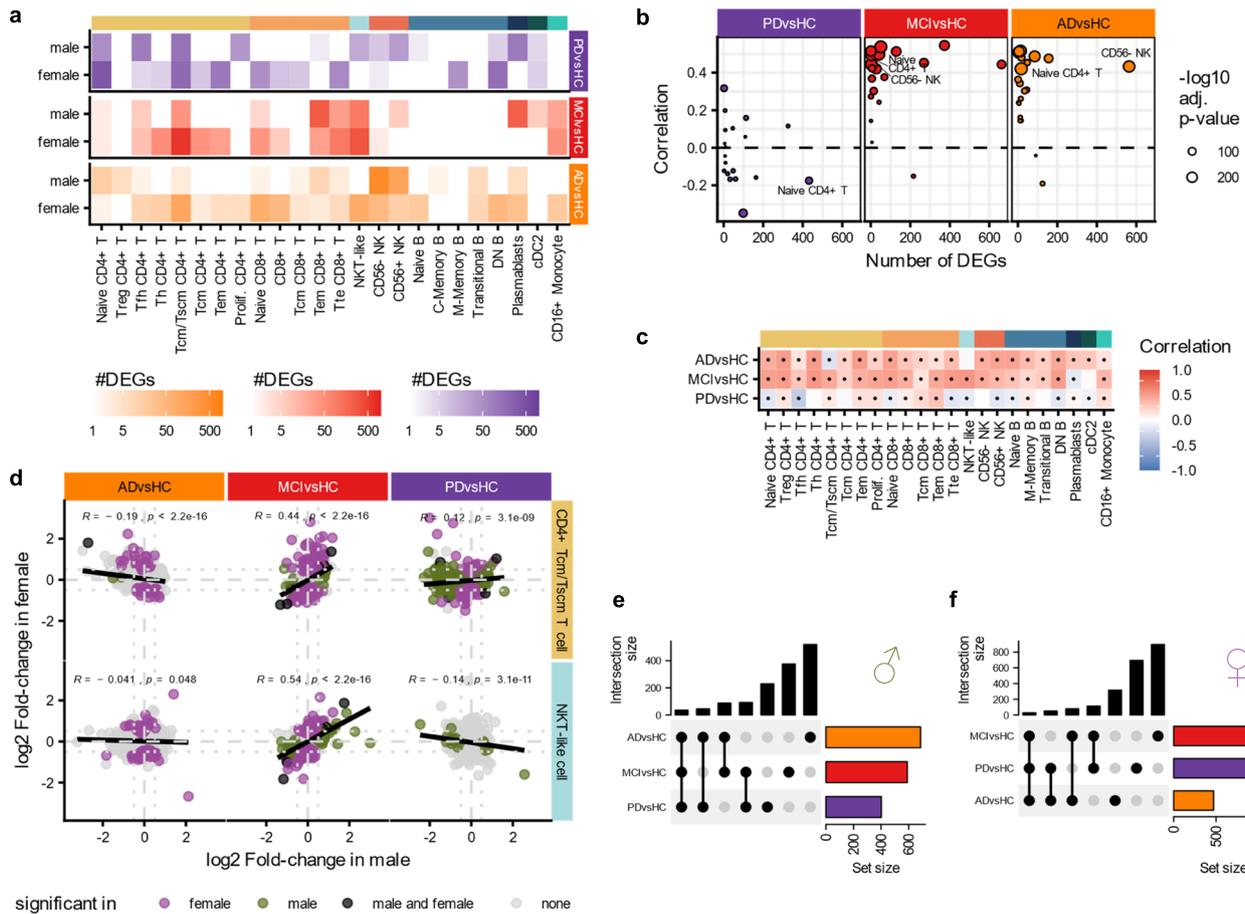


Fig. 3 | Changes in gene-expression patterns in male and female patients.
a Comparison of the number of genes that are significantly de-regulated (adj. p-value < 0.05 and absolute log₂ fold-change > 0.5, see Methods) between the healthy and diseased patients in females and in males. **b** Cell types with more significantly de-regulated genes (as in **a**) show higher correlation values in AD and MCI. The Pearson correlation coefficient was calculated between the fold-changes in male and female of all genes in the dataset. **c** Correlations of the fold-changes in male and female patients in PD is often lower and in 9 cell types negative, but positive in most cell types in MCI and AD. Significant values

(Pearson’s correlation, adj. p-value < 0.05) are marked with a dot. **d** Comparison of the log₂ Fold-changes in male and female patients in CD4+ T_{cm}/T_{scm} cells and NKT-like cells shows the number of genes that are significantly de-regulated in both or only one sex. The genes are colored by the sex in which they are significant in (adj. p-value < 0.05, see Methods). **e**, **f** Upset plot showing the number of significantly de-regulated genes that are shared between the different comparisons of healthy with the diagnosis groups for the male sub-group (**e**) and for the female sub-group (**f**) show similar patterns in the overlap between diseases in both sexes.

system, blood-brain barrier, Astrocytes and Microglia (Supplementary Table 3, Fig. 5d, e).

To gain a deeper understanding for the changes in PBMCs and brain cells and how they might interact, we compared the de-regulated genes with the CellChat signaling database⁵⁴. We found genes from the *CCL* (Supplementary Fig. 8d) and the *MHC-I* signaling pathway in the de-regulated genes in females and the *CD45* and the *CCL* signaling pathway in males. All of these genes have already been reported in the context of Alzheimer’s in the literature^{55–57}. This indicates that changes in the blood might influence the brain or vice versa, but as there are only few genes of these pathways de-regulated, further experiments would be necessary to confirm this.

Additionally, we performed a pathway analysis using both the de-regulated genes in male and female samples (Supplementary Data 6 and 7). We observed a general enrichment of membrane, ribosome, and adherens junction-related pathways in both blood and brain (Fig. 5f). Especially the “SRP-dependent cotranslational protein targeting to membrane” pathway is frequently enriched in all three datasets. This suggests that the observed symptoms might be caused by a systemic change that affects both blood and the brain.

To further investigate the relationship between changes in the gene-expression in both tissues, we compared them to an additional bulk RNA-

sequencing dataset from the ROSMAP cohort (Synapse <https://www.synapse.org/#!Synapse:syn3388564>, Supplementary Fig. 8e). Although fold-changes in the brain single-cell and bulk data showed a significant positive correlation (overall correlation: 0.14, p-value: 0.00025), we did not find a significant correlation of the PBMC single-cell with the brain bulk data (Supplementary Fig. 8f).

Access for the community

To allow easy, fast and uncomplicated data access to our large-scale findings, we developed a webserver for the dataset. The web-resource allows the visualization of gene-expression values (Supplementary Fig. 9a, b). The direct access to the list of de-regulated genes and its visualization (Supplementary Fig. 9c) allow fast hypothesis-testing.

In addition to the scRNA-seq data used for the analysis shown in this manuscript, the webserver also allows researchers to access the longitudinal data, brain volume measurements and CSF-marker values.

Our dataset contains 32 patients with more than 1 time-point (Supplementary Fig. 10a) and 3 with more than 2. All patient with 3 Visits were female and we observed differences in the age-distribution between male and female patients in this subgroup (Supplementary Fig. 10b). The average time between the first and second visit was less than a year for AD and between 2 and 3 years for the other diseases

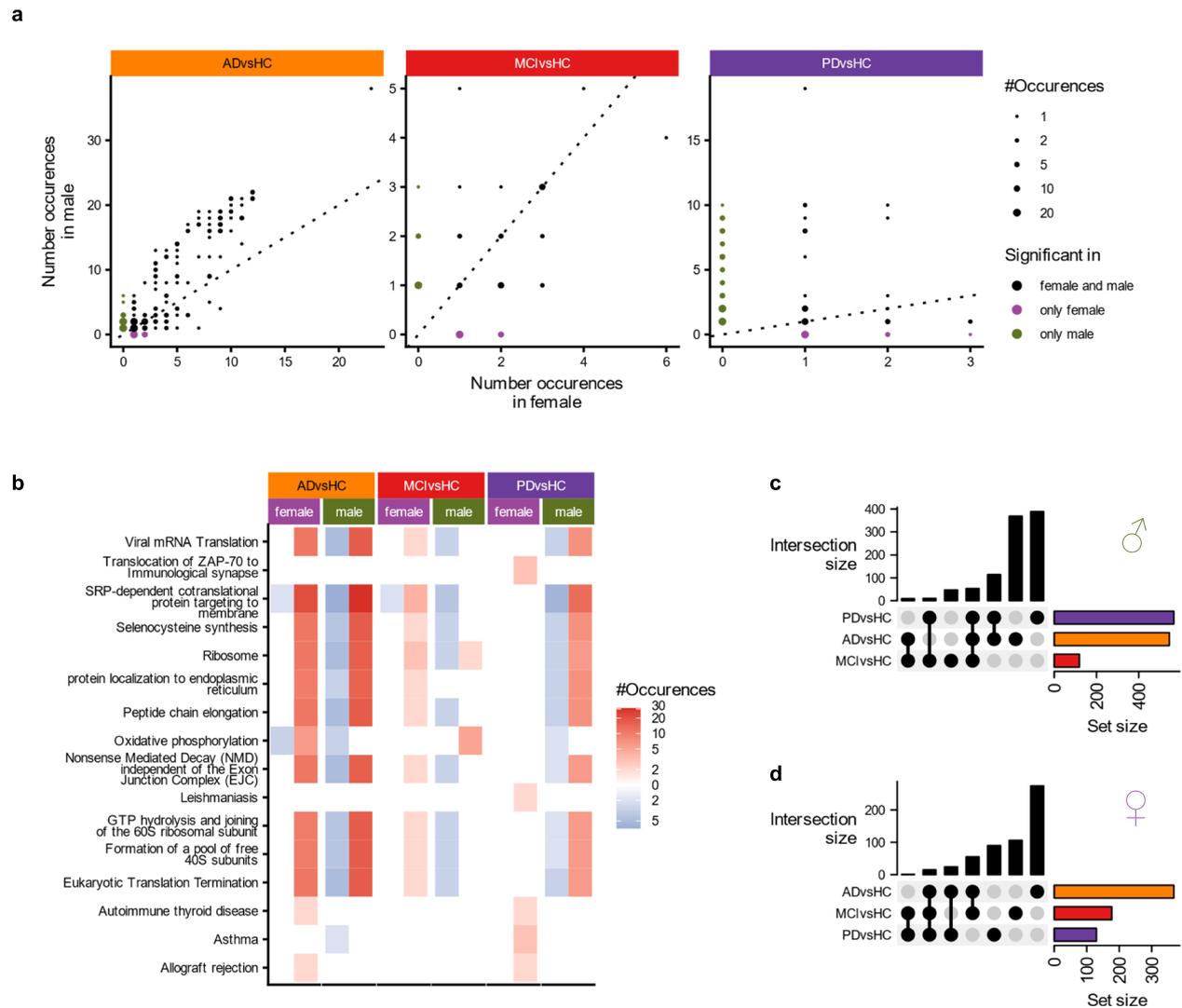


Fig. 4 | Pathways analysis shows sex-dependent changes in PD and MCI. a The number of cell types in which a pathway is enriched or depleted in males and females in the different comparisons. The dots are colored if they occur only in males or in females. **b** Number of occurrences of the top 5 most frequently enriched

and depleted pathways in each comparison and sex. **c, d** Size of the overlap of the pathways in the different comparisons, independent of cell type for males (**c**) and females (**d**).

(Supplementary Fig. 10c). When correlating the cell-type proportions with the number of the visit, we observed no significant correlations (Supplementary Fig. 10d). Generally, both the cell-type proportions of the same patient at different visits and between different patients show large fluctuations (Supplementary Fig. 10e).

Although the statistical power of the longitudinal data is limited, we included it on the website, allowing interested researchers to study this data in more detail. We additionally included the data of the brain volume measurements (Supplementary Fig. 11a) and CSF-marker values (Supplementary Fig. 11b) for direct comparison. Similarly to the time-series data, we correlated it with the cell-type proportions (Supplementary Fig. 11c). We found significant correlations (adj. P-value < 0.05) mainly in healthy controls. We still include this data as a resource for other researchers.

Our dataset showed sex-dependent differences in the changes of the cell type composition and in the gene-expression profiles of neurodegenerative diseases and Parkinson's disease specifically.

Discussion

The era of single-cell -omics has enabled studying human diseases at an unprecedented cellular resolution combined with massive

throughput, generating millions of transcriptomic profiles at ease. While large-scale single-cell analysis of fresh-frozen post-mortem brains seems to come in close reach for the upcoming years, it is of great importance to understand which systemic changes precede and co-occur neurodegeneration in the human body as to inform our current understanding of early biomarkers as well as individualized risk or progression markers. Our dataset is composed of 290 gene-expression profiles of PBMCs from MCI, AD, PD, PD-MCI and HC individuals and includes additional information about Brain volume measurements and CSF-biomarker levels of known AD biomarkers. This dataset allows us to generally study disease- and sex-specific changes in patients with neuro-degenerative diseases.

To study the exact mechanisms, a more specialized dataset with more samples (e.g. for AD)⁴⁸ and additional demographic and clinical information (e.g. medication use, other disease, lifestyle factors,...) would be necessary. Changes in the sub-cell types and potential sub-groups in the cell clusters furthermore need to be studied in more targeted experiments, as the number of cells in each of these clusters is limited.

When comparing the cell type composition, we found large differences between the changes in men and women that are often specific for certain sub-cell types. Changes in B cells and Monocytes in

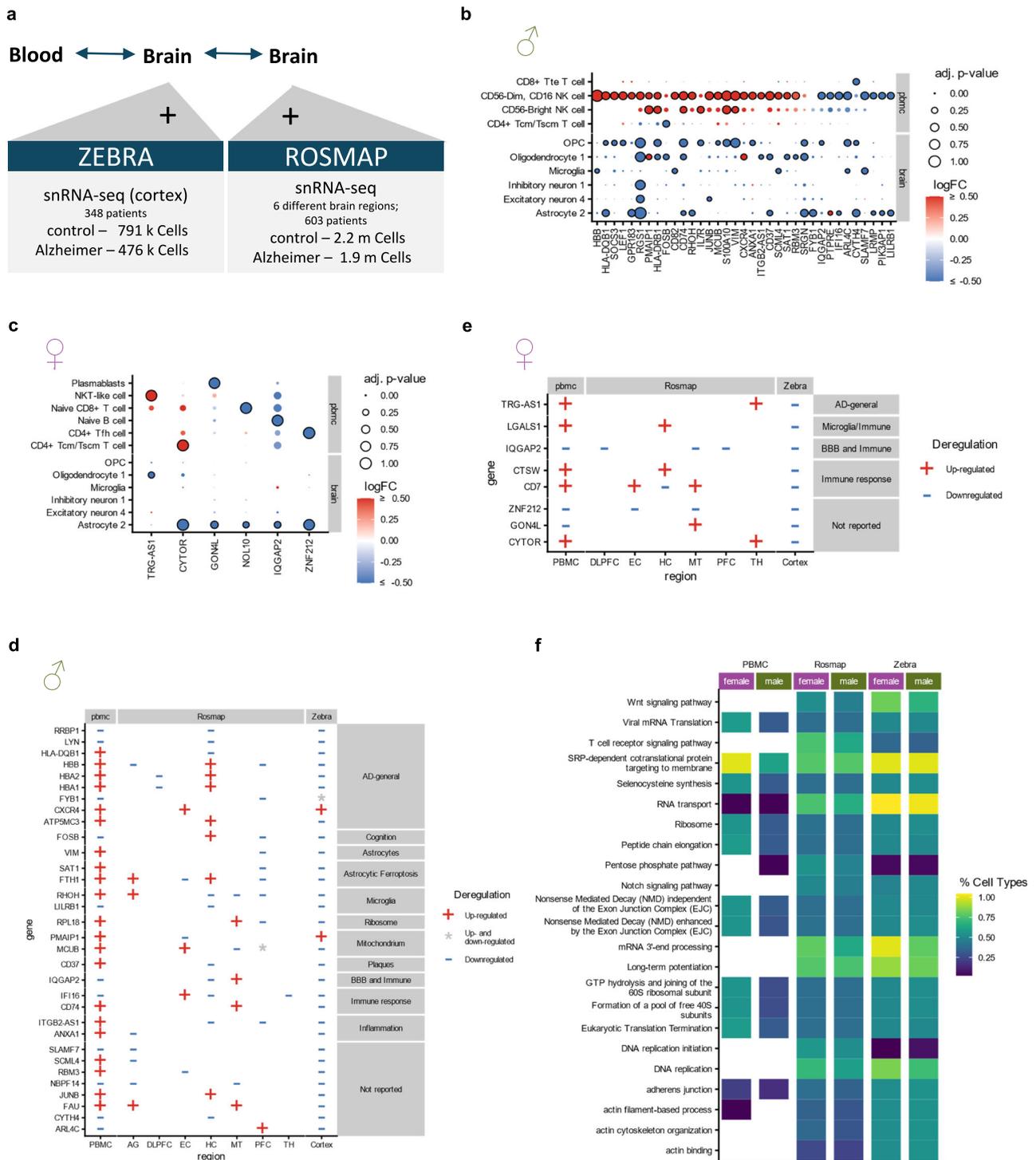


Fig. 5 | Comparison of gene-expression in brain and blood of Alzheimer's patients. **a** Single-cell RNA-sequencing data from the cortex from the ZEBRA-dataset was used for the comparison of changes in PBMC samples and brain samples. **b** Fold-changes in PBMCs and cortex of the genes with an adj. p-value < 0.05 (see Methods) in the PBMCs the brain in males. **c** De-regulation of genes in PBMCs and cortex with an adj. p-value < 0.05 in PBMCs and the brain and an abs. log2 fold-change bigger than 0.6 in the PBMCs (see Methods). **d, e** Deregulation of the

significantly de-regulated genes in all three datasets (see Methods) and the direction of de-regulation in males (**d**) and females (**e**) (+: Up-regulated in all cell-types, -: Down-regulated in all cell-types and *: Mixed signals). Genes are annotated with the context in which they have been previously reported in for Alzheimer's (see Supplementary Table 3). **f** Top 10 most frequently enriched or depleted pathways in both brain datasets and in PBMCs and the proportion of cell-types they are enriched/depleted in.

Alzheimer's disease were inconsistent across male and female patients. It has been previously shown that depletion of B cells from mice led to cognitive deficits and an increase of Aβ plaques⁵⁸. B cells also play an important role in the regulation of local immune responses and might thus indicate inflammatory states in Alzheimer's. The increased levels

of monocytes in AD can furthermore be a sign of chronic inflammation and may have a negative effect on structures like the blood brain barrier (BBB)⁵⁹. In contrast to the large sex-specific differences in the PD cell type compositions that we observed, previous studies mainly found changes in the T-cell population of Parkinson's patients. Of note,

the findings of the previous studies sometimes contradict each other. This can be partially explained by the large fluctuations in the cell type composition that can be observed even of the same patient⁶⁰ so that differences in the cell type composition can vary largely between samples. Moreover, some of our findings have not yet been documented, or only partially align with previous studies. This could be attributed to biases in the cohorts used (such as differences in ethnicity, disease severity, or other factors), technological limitations, or other biases.

To identify sex- and disease-specific gene signatures, a machine-learning model was applied to the dataset (Supplementary Notes 1, Supplementary Data 8). Testing out a random forest-based feature selection and three other machine-learning methods, we concluded that the application on an Alzheimer's disease dataset is too complex and would require a larger dataset (ideally with different studies) to reliably predict disease-specific features across patients.

The data furthermore suggest sex-specific differences between gene-expression signatures of PD, more similarities, but still sex-specific differences in AD. Sex-specific differences in PD have been previously found in Monocytes²⁷ but sex-independent expression-changes were reported in T-Memory cells⁵¹. We have found an anti-correlation of gene-expression changes in male and female indicating large differences between the disease-specific changes in PD, while the sex-specific differences in AD can be mainly observed by the effect size of the gene-expression change. We have seen that we found much more DEGs in female patients than in male patients. Even though this effect can be observed across all diseases in this study, it is particularly strong in AD, which has previously been reported by Mathys et al.¹¹. In Parkinson this effect has been described previously in blood Monocytes²⁷, whereas the exact opposite case has been observed in the brain⁶².

The changes that were observed in Alzheimer's disease independent of the sex mainly focused on genes from the "SRP-dependent cotranslational protein targeting to membrane" pathway. This pathway has been previously reported in Alzheimer's disease⁶³ and Parkinson's disease^{64,65}. This indicates similar changes in male and female Alzheimer's patients and male Parkinson's patients. Changes in those genes have been showing an increased expression during microglia activation, a state that has previously been associated with Alzheimer's⁶³.

The "SRP-dependent cotranslational protein targeting to membrane" pathway has also been observed in brain samples from both the ZEBRA and the ROSMAP dataset. The genes involved thus do not only show changes in PBMC, but also the Brain. The genes of this pathway have previously been discussed in the context of periodontitis, a condition mainly caused by *Porphyromonas gingivalis*^{63,66}. This bacterium is known to secrete gingipains, which have been linked to Alzheimer's pathology in mice and humans^{67,68}.

Diving deeper into the changes that are observed in both the blood and the brain, we found 30 overlapping genes in males and 7 in females that are significantly deregulated in both blood and the brain. A majority of those genes have been previously reported in relation to Alzheimer's disease in different tissues and cell-types (Supplementary Table 3). These include genes related to changes in the Immune system and the Blood-brain barrier (especially in females) and changes related to Astrocytes and Microglia. This indicates general changes in the immune system, possibly in relation to an infection (such as herpes simplex) or age-related inflammation processes. The infection with Herpes simplex can lead to a higher risk of developing Alzheimer's in APOE4 carriers^{52,53} and is discussed to play a major role in the disease development.

As patients of the ADRC cohort are continuously monitored, two major shortcomings of our study will be addressed in the future. First, longitudinal trajectories will be important to trace down the temporal dynamics on the molecular level. Second, a patient-matched analysis of both brain and blood tissue will enable us to validate the here reported effects ruling out inter-individual variation.

In sum, our dataset indicates that there are strong sex-related differences in the neuro-degenerative diseases discussed in this study. While the exact relationship between the immune-system and the brain remains unknown, this study highlights the importance to consider possible influences of the immune system and the sex on the development of these diseases. Beyond the analyses that we provide in the present study, we are convinced that exploring the very broad dataset supports a broad range of further research topics. We thus make not only the raw data available but also implemented a convenient web interface with substantial functionality to explore the data (<https://www.ccb.uni-saarland.de/adrcsc>).

Methods

Cohort

The patients included in the study were enrolled by the Stanford Alzheimer's Disease research Center (ADRC). The group of patients includes patients diagnosed with mild cognitive impairment (individual patients $n = 51$, total samples from the patients $n_t = 63$; MCI), Alzheimer's disease ($n = 31$, $n_t = 43$; AD), Parkinson's disease ($n = 55$, $n_t = 61$; PD), PD with MCI ($n = 20$, $n_t = 24$, PD-MCI), and healthy individuals ($n = 154$, $n_t = 172$, HC). We selected patients with respect to a similar ratio of both biological sex and age between the subgroups. The Stanford University IRB approved the study and all patients provided written informed consent. Due to the strict quality control, the final dataset only contains a total of 290 samples from 257 individuals. CSF measurements were performed for NFL, UCHL1, Tau and GFAP (Quanterix4plex), Ab40, Ab42 and Tau (Quanterix3plex) and pTau181 (Quanterix). Brain Volume measurements were performed using PET and MRI imaging.

Single-cell suspension preparation

Frozen PBMCs vials were rapidly thawed in a 37°C water bath for 2 minutes, and the vials were removed when a tiny ice crystal was left. Thawed PBMCs were quenched with 13 ml 37°C pre-warmed 1X phosphate-buffered saline (PBS, Thermo Fisher Scientific, #10010031) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher Scientific, #A3160601). Cells were centrifuged at 300 x g for 10 minutes at room temperature. The supernatant was removed, and cell pellet was resuspended in 3 ml 1X PBS containing 10% FBS, passed through a 40 µm cell strainer (Falcon, #352340), then centrifuged at 300 x g for 10 minutes at room temperature. Dead cells were removed by magnetic beads purification (Miltenyi Biotech, #130-090-101) according to the manufacturer's protocol. Cells were resuspended with cell resuspension buffer at a concentration of 1000 viable cells/µl.

Single-cell RNA-seq with DNBelab C4 system

The DNBelab C Series Single-Cell Library Prep Set (MGI, #1000021082) was utilized as previously described⁴⁹. In brief, single-cell suspensions were used for droplet generation, emulsion breakage, beads collection, reverse transcription, and cDNA amplification to generate bar-coded libraries. Indexed scRNA-seq libraries were constructed according to the manufacturer's protocol. The DNA nanoballs (DNBs)-based libraries were sequenced by the ultra-high-throughput DIPSEQ T1 sequencer at China National GeneBank (CNGB). The read structure was paired-end with Read 1, covering 30 bases inclusive of 10-bp cell barcode 1, 10-bp cell barcode 2 and 10-bp unique molecular identifier (UMI), and Read 2 containing 100 bases of transcript sequence, 10-bp sample index.

Primary single-cell RNA-seq data processing

Raw sequencing reads from DIPSEQ T1 sequencer were filtered and demultiplexed using PISA (version 0.2) (<https://github.com/shiquan/PISA>). Reads were aligned to hg38 genome using STAR⁴⁹ (version 2.7.4a), and sorted by sambamba⁷⁰ (version 0.7.0). Genes were

annotated according to GENCODE⁷¹. Cell versus gene UMI count matrix was generated with PISA.

Quality control and data processing

For each count matrix, SoupX⁷² (1.4.8) was used to estimate and remove ambient RNA contamination. Only samples with at most 10% estimated contamination were kept for further analyses. Subsequently samples were filtered such that each sample contained at least 300 cells, each cell contained at most 7% mitochondrial gene counts and between 300 and 4000 genes were expressed per cell. Doublet filtering was also performed sample-wise using the R Bioconductor package scDblFinder⁷³ 1.2 with an expected doublet ratio of 1% per 1000 cells. Next, all samples were merged.

Dimensionality-reduction and clustering

Dimensionality-Reduction and clustering were performed using Seurat (version 4.0.0)⁷⁴. First the data were again normalized, scaled and the top 2000 variable genes were selected using Seurats NormalizeData, FindVariableFeatures and ScaleData functions. For the dimensionality-reduction, we used the RunPCA and the RunUMAP function with the first 10 dimensions, which was determined using Seurats ElbowPlot-Function. We performed clustering using the FindNeighbors-function with the first 10 dimensions and the FindClusters-Function with a resolution of 0.1.

Cell type annotation

This resulted in multiple clusters which could be assigned to B cells, T and NK cells, myeloid cells, Plasma cells and a cluster with mixed cell populations. To identify cell subclusters we next performed a second round of dimensionality-reduction and clustering of each of the previous clusters. The T/NK cell clusters are first processed independently for each batch for the filtering and then merged afterwards to perform cell type annotation. We then filtered out sub-clusters with more than 50% cells that originate only from one sample and repeated the dimensionality-reduction and clustering on the filtered data, as described for the complete set of cells. We repeated this step until no cells are filtered out. Cell sub-clusters are then checked for known marker-genes. Clusters that did not express known markers of PBMC cells or that are indicated to be low-quality or doublets are removed from the dataset, followed by another dimensionality-reduction and clustering step. The T/NK cells of the different batches are then merged, dimensionality-reduction and clustering were performed and the cells were split into CD4⁺, CD8⁺ and NK cells before continuing with the sub-clustering and the cell type annotation. After having filtered out the unwanted cells, the final sub-clusters were annotated based on known markers from literature.

Low dimensional embedding and density computation

An overall cell density per cell type and patient cohort was determined by first estimating the cell density per sample and cell type with the two-dimensional kernel density estimation implemented in the kde2d function of the MASS R-package (version 7.3-55). Subsequently, the density estimates were aggregated per patient cohort by averaging the obtained density estimates.

Cell-type composition analysis

Changes in the cell-type composition were evaluated using an unpaired Student's t-Test on the raw cell-type proportions and the p-values were corrected using the Benjamini-Hochberg procedure. scCODA (version 0.1.9)⁷⁵ was first used with an automatically selected reference cell type and Age and AgoE as covariates. Using the selected reference cell type (CD4⁺ T cell for the broad annotation and Treg CD4⁺ cell for the fine annotation), the analysis was repeated separately for males and females. The corrected Fold-changes were reported, and significant changes were determined using a p-value cut-off of 0.05.

Differential gene expression analysis

Differentially expressed genes on the PBMC data were determined with the pseudobulk approach implemented by the muscat⁷⁶ package (version 1.6.0). Cells were aggregated with the aggregate data function at cell type and sample level by summing up the SoupX corrected counts. The pbDS function was then used with the limma-voom⁷⁷ package to calculate the set of differentially expressed genes. The design matrix included besides the diagnosis, sample batch and sex as factors. Multiple testing adjustment was performed with the Benjamini-Hochberg procedure. Genes were subsequently filtered to be expressed in at least 5% of all cells of a cell type in at least one of the comparison patient cohorts. Genes were considered significantly down-regulated when they showed an absolute log₂ fold-change above 0.5 (i.e. a log₂ FC < -0.5 or > 0.5) and an adjusted p-value smaller than 0.05. Adjusted p-values were reported as given by the muscat package.

ZEBRA Brain Atlas preparation and DEG analysis

The study collection, pre-processing, and aggregating of the brain single-nucleus data was performed as described in Flotho et al.⁵¹. The Blanchard et al.⁷⁸ was filtered and the Hardwick et al.⁷⁹ dataset was removed due to overlaps with other studies in the dataset. The marker genes have been computed using a pseudo-bulk approach with edgeR (v3.36.0). Only cells from studies which reported a particular gene were considered for computing the DEGs. If the cells for computing the DEGs originated from more than a single study the study ID was used as a latent variable in the edgeR design matrix. For significance testing the glmQLFit method was used. Finally, the p-values have been adjusted using Benjamini-Hochberg correction of p.adjust method from the R stats package (v 4.1.3).

ROSMAP Brain Atlas preparation and DEG analysis

We collected and harmonized a scRNA-seq AD/CT dataset from state of the art studies using sample from the Religious Order and Aging Project (ROSMAP)⁸⁰. In particular we used the studies syn52293433, syn18681734, syn51062116 and syn2580853.

We utilized the Synapse client for Python (synapseclient version 4.5.1) with Python 3.11.10 to acquire all necessary metadata. For realigning the counts we use the CellRanger v8.0.1 software with the 2T2 reference genome. We removed redundancies in the data by overlapping cells and shared donors across studies and removed all duplicated cells from our downstream analysis. We apply the following QC thresholds: Genes have to be present in at least 3 cells, each cell with less than 100 unique genes has been discarded. We remove putative doublets using scrublet v0.2.2⁸¹ with default parameters, and run CellBender v0.3.2⁸² as well with defaults.

We use the Allen Brain Institutes MapMyCells (RRID:SCR_024672) for cell type labeling, here only cell types with more than 100 cells have been considered for further analysis. For computing DEGs we followed the already described methodology (Methods: Differential gene expression analysis) and included all NCI and AD cases in our comparison.

Pathway enrichment analysis

Enriched pathways on the whole gene-set were determined by performing an unweighted gene set enrichment analysis using GeneTrail 3⁸³ and the "GO - Biological Process", "GO - Cellular Component", "GO - Molecular Function", "KEGG - Pathways", "Reactome - Pathways", "WikiPathways" and "Pfam - Protein families" databases. Genes were ordered by the -log₁₀ adjusted p-value and the absolute log₂ fold-change. Only genes which were used by the models were considered. Pathway analysis for pre-selected gene subsets were performed using over-representation analysis using GeneTrail 3 with the same database⁸³. Pathways with an adjusted p-value smaller than 0.05 were considered significantly enriched or depleted. P-values were reported as given by GeneTrail.

Cell-cell communication analysis

Changes in the cell-cell communication were evaluated using the CellChat package⁵⁴ (version 1.5.0).

Literature Search

We performed a manually curated literature-research using PubMed-Search by combining the key-words "Alzheimer's" and "Parkinson's" with "blood", "peripheral immunity", "peripheral blood", "pbmc" or "CSF marker" or "brain" with and without the keyword "sex" and selected papers that provide further information and previous findings on the topics of this paper. For the single-cell and bulk PBMC datasets of Alzheimer's and Parkinson's, we performed a manually curated literature-research using PubMed-Search with the terms "transcriptomics", "single-cell", "pbmc", "human", "RNA-seq", "blood", "parkinson", "alzheimer" and "neurodegeneration" and selected all single-cell and bulk RNA-seq datasets.

Correlation analysis

All correlation values were determined using the Pearson's correlation coefficient (cor.test Function). P-values were adjusted using Benjamini-Hochberg. Correlations were considered significant if they had an adjusted p-value smaller than 0.05.

Cosine similarity

The similarity between changes in the gene expression was calculated as the cosine similarity between the unfiltered gene-lists ordered by their fold-changes.

Statistics and reproducibility

All test statistics used were conducted two-tailed, if not indicated otherwise. Computational tools requiring seed generators were set to use random but fixed initialization constants. Unless stated differently, p-values were corrected for multiple hypothesis testing using Benjamini-Hochberg. Unless stated otherwise, all analyses were performed using the baseline sample for each patient, with specifically defined analyses involving also longitudinal samples.

Webserver implementation

The interactive visualization website was implemented using Django v2.2.16 for the backend and bootstrap v5.0.1 for the front end. For the visualization, plotly.js v2.0.0 was used for the graphs and datatables v1.10.25 for the differential expression table. The data was stored in the h5ad format and accessed with anndata v0.7.6.

Reporting summary

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

Data availability

Raw sequencing data is freely available from the sequence read archive (SRA) using accession ID SRP312418. An interactive webserver to explore the scRNA-seq count data is freely available at <https://www.ccb.uni-saarland.de/adrcsc>. As metadata, we publish sex and diagnosis to comply to Nature's scientific data policy. Source data are provided with this paper.

Code availability

The code used to analyze the data is available from GitHub⁵⁴. Source data are provided with this paper.

References

- Scheltens, P. et al. Alzheimer's disease. *Lancet* **397**, 1577–1590 (2021).
- Bloem, B. R., Okun, M. S. & Klein, C. Parkinson's disease. *Lancet* **397**, 2284–2303 (2021).
- Hou, Y. et al. Ageing as a risk factor for neurodegenerative disease. *Nat. Rev. Neurol.* **15**, 565–581 (2019).
- Gan, L., Cookson, M. R., Petrucelli, L. & La Spada, A. R. Converging pathways in neurodegeneration, from genetics to mechanisms. *Nat. Neurosci.* **21**, 1300–1309 (2018).
- Jessen, F. et al. The characterisation of subjective cognitive decline. *Lancet Neurol.* **19**, 271–278 (2020).
- Gauthier, S. et al. Mild cognitive impairment. *Lancet* **367**, 1262–1270 (2006).
- Lindestam Arlehamn, C. S., Garretti, F., Sulzer, D. & Sette, A. Roles for the adaptive immune system in Parkinson's and Alzheimer's diseases. *Curr. Opin. Immunol.* **59**, 115–120 (2019).
- Ransohoff, R. M. How neuroinflammation contributes to neurodegeneration. *Science* **353**, 777–783 (2016).
- Marschallinger, J. et al. Lipid-droplet-accumulating microglia represent a dysfunctional and proinflammatory state in the aging brain. *Nat. Neurosci.* **23**, 194–208 (2020).
- Wightman, D. P. et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat. Genet.* **53**, 1276–1282 (2021).
- Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
- Grubman, A. et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat. Neurosci.* **22**, 2087–2097 (2019).
- Smajic, S. et al. Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. *Brain* **145**, 964–978 (2022).
- Yang, A. C. et al. A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature* **603**, 885–892 (2022).
- Garcia, F. J. et al. Single-cell dissection of the human brain vasculature. *Nature* **603**, 893–899 (2022).
- Hampel, H. et al. Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. *Nat. Rev. Neurol.* **14**, 639–652 (2018).
- Gate, D. et al. Clonally expanded CD8 T cells patrol the cerebrospinal fluid in Alzheimer's disease. *Nature* **577**, 399–404 (2020).
- Gate, D. et al. CD4(+) T cells contribute to neurodegeneration in Lewy body dementia. *Science* **374**, 868–874 (2021).
- Banks, W. A., Reed, M. J., Logsdon, A. F., Rhea, E. M. & Erickson, M. A. Healthy aging and the blood-brain barrier. *Nat. Aging* **1**, 243–254 (2021).
- Sweeney, M. D., Kisler, K., Montagne, A., Toga, A. W. & Zlokovic, B. V. The role of brain vasculature in neurodegenerative disorders. *Nat. Neurosci.* **21**, 1318–1331 (2018).
- Sweeney, M. D., Sagare, A. P. & Zlokovic, B. V. Blood-brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nat. Rev. Neurol.* **14**, 133–150 (2018).
- Yang, A. C. et al. Physiological blood-brain transport is impaired with age by a shift in transcytosis. *Nature* **583**, 425–430 (2020).
- Amin, J. et al. Peripheral immunophenotype in dementia with Lewy bodies and Alzheimer's disease: an observational clinical study. *J. Neurol. Neurosurg. Psychiatry* **91**, 1219–1226 (2020).
- Cao, W. & Zheng, H. Peripheral immune system in aging and Alzheimer's disease. *Mol. Neurodegener.* **13**, 51 (2018).
- Ni Chasaide, C. & Lynch, M. A. The role of the immune system in driving neuroinflammation. *Brain Neurosci. Adv.* **4**, 2398212819901082 (2020).
- Prinz, M. & Priller, J. The role of peripheral immune cells in the CNS in steady state and disease. *Nat. Neurosci.* **20**, 136–144 (2017).
- Carlisle, S. M. et al. Sex-based differences in the activation of peripheral blood monocytes in early Parkinson disease. *NPJ Parkinsons Dis.* **7**, 36 (2021).

28. Tumpara, S. et al. Polymerization of misfolded Z alpha-1 antitrypsin protein lowers CX3CR1 expression in human PBMCs. *Elife* **10**, <https://doi.org/10.7554/eLife.64881> (2021).
29. Shigemizu, D. et al. Identification of potential blood biomarkers for early diagnosis of Alzheimer's disease through RNA sequencing analysis. *Alzheimers Res Ther.* **12**, 87 (2020).
30. Schlachetzki, J. C. M. et al. A monocyte gene expression signature in the early clinical course of Parkinson's disease. *Sci. Rep.* **8**, 10757 (2018).
31. Konstantin Nissen, S. et al. Changes in CD163+, CD11b+, and CCR2+ peripheral monocytes relate to Parkinson's disease and cognition. *Brain Behav. Immun.* **101**, 182–193 (2022).
32. Kern, F. et al. Deep sequencing of snRNAs reveals hallmarks and regulatory modules of the transcriptome during Parkinson's disease progression. *Nat. Aging* **1**, 309–322 (2021).
33. Mila-Aloma, M. et al. Plasma p-tau231 and p-tau217 as state markers of amyloid-beta pathology in preclinical Alzheimer's disease. *Nat. Med* **28**, 1797–1801 (2022).
34. Schindler, S. E. & Karikari, T. K. Comorbidities confound Alzheimer's blood tests. *Nat. Med* **28**, 1349–1351 (2022).
35. Palmqvist, S. et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat. Med* **27**, 1034–1042 (2021).
36. Bateman, R. J., Barthelemy, N. R. & Horie, K. Another step forward in blood-based diagnostics for Alzheimer's disease. *Nat. Med* **26**, 314–316 (2020).
37. Zetterberg, H. & Schott, J. M. Biomarkers for Alzheimer's disease beyond amyloid and tau. *Nat. Med* **25**, 201–203 (2019).
38. Phongpreecha, T. et al. Single-cell peripheral immunoprofiling of Alzheimer's and Parkinson's diseases. *Sci. Adv.* **6**, <https://doi.org/10.1126/sciadv.abd5575> (2020).
39. Kapellos, T. S. et al. Human Monocyte Subsets and Phenotypes in Major Chronic Inflammatory Diseases. *Front Immunol.* **10**, 2035 (2019).
40. Gazestani, V. et al. Early Alzheimer's disease pathology in human cortex involves transient cell states. *Cell* **186**, 4438–4453 e4423 (2023).
41. Sun, N. et al. Human microglial state dynamics in Alzheimer's disease progression. *Cell* **186**, 4386–4403.e4329 (2023).
42. Mathys, H. et al. Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**, 4365–4385.e4327 (2023).
43. Gabitto, M. I. et al. Integrated multimodal cell atlas of Alzheimer's disease. *Res Sq.* <https://doi.org/10.21203/rs.3.rs-2921860/v1> (2023).
44. Sun, N. et al. Single-nucleus multiregion transcriptomic analysis of brain vasculature in Alzheimer's disease. *Nat. Neurosci.* **26**, 970–982 (2023).
45. Yang, A. C. et al. Dysregulation of brain and choroid plexus cell types in severe COVID-19. *Nature* **595**, 565–571 (2021).
46. Ren, X. et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* **184**, 1895–1913.e1819 (2021).
47. Stephenson, E. et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med* **27**, 904–916 (2021).
48. Fujita, M. et al. Cell subtype-specific effects of genetic variation in the Alzheimer's disease brain. *Nat. Genet* **56**, 605–614 (2024).
49. Liu, C. et al. A portable and cost-effective microfluidic system for massively parallel single-cell transcriptome profiling. *bioRxiv* <https://doi.org/10.1101/818450> (2019).
50. Huang, Z. et al. Effects of sex and aging on the immune cell landscape as assessed by single-cell transcriptomic analysis. *Proc. Natl Acad. Sci. USA* **118**, <https://doi.org/10.1073/pnas.2023216118> (2021).
51. Flotho, M. et al. ZEBRA: a hierarchically integrated gene expression atlas of the murine and human brain at single-cell resolution. *Nucleic Acids Res* **52**, D1089–D1096 (2024).
52. Lovheim, H. et al. Herpes simplex infection and the risk of Alzheimer's disease: A nested case-control study. *Alzheimers Dement* **11**, 587–592 (2015).
53. Linard, M. et al. Interaction between APOE4 and herpes simplex virus type 1 in Alzheimer's disease. *Alzheimers Dement* **16**, 200–208 (2020).
54. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat. Protoc.* <https://doi.org/10.1038/s41596-024-01045-4> (2024).
55. Wang, H. C., Zong, Y., Zhu, L., Wang, W. Y. & Han, Y. S. Chemokines in patients with Alzheimer's disease: A meta-analysis. *Front Aging Neurosci.* **15**, ARTN 1047810 (2023).
56. Kim, M. S. et al. Neuronal MHC-I complex is destabilized by amyloid-beta and its implications in Alzheimer's disease. *Cell Biosci.* **13**, 181 (2023).
57. Zhu, Y. et al. CD45 deficiency drives amyloid-beta peptide oligomers and neuronal loss in Alzheimer's disease mice. *J. Neurosci.* **31**, 1355–1365 (2011).
58. Xiong, L. L. et al. Single-cell RNA sequencing reveals B cell-related molecular biomarkers for Alzheimer's disease. *Exp. Mol. Med* **53**, 1888–1901 (2021).
59. Shad, K. F., Aghazadeh, Y., Ahmad, S. & Kress, B. Peripheral markers of Alzheimer's disease: surveillance of white blood cells. *Synapse* **67**, 541–543 (2013).
60. Kashima, Y. et al. Intensive single-cell analysis reveals immune-cell diversity among healthy individuals. *Life Sci Alliance* **5**, <https://doi.org/10.26508/lsa.202201398> (2022).
61. Dhanwani, R. et al. Transcriptional analysis of peripheral memory T cells reveals Parkinson's disease-specific gene signatures. *NPJ Parkinsons Dis.* **8**, 30 (2022).
62. Tranchevent, L. C., Halder, R. & Glaab, E. Systems level analysis of sex-dependent gene expression changes in Parkinson's disease. *NPJ Parkinsons Dis.* **9**, 8 (2023).
63. Patel, S. et al. Donor-Specific Transcriptomic Analysis of Alzheimer's Disease-Associated Hypometabolism Highlights a Unique Donor, Ribosomal Proteins and Microglia. *eNeuro* **7**, <https://doi.org/10.1523/ENEURO.0255-20.2020> (2020).
64. Flinkman, D. et al. Regulators of proteostasis are translationally repressed in fibroblasts from patients with sporadic and LRRK2-G2019S Parkinson's disease. *NPJ Parkinsons Dis.* **9**, 20 (2023).
65. Hernandez, S. M. et al. Unexpected Implication of SRP and AGO2 in Parkinson's Disease: Involvement in Alpha-Synuclein Biogenesis. *Cells* **10**, <https://doi.org/10.3390/cells10102792> (2021).
66. Felsky, D. et al. The Caribbean-Hispanic Alzheimer's disease brain transcriptome reveals ancestry-specific disease mechanisms. *Neurobiol. Dis.* **176**, 105938 (2023).
67. Ilievski, V. et al. Chronic oral application of a periodontal pathogen results in brain inflammation, neurodegeneration and amyloid beta production in wild type mice. *PLoS One* **13**, e0204941 (2018).
68. Dominy, S. S. et al. Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Sci. Adv.* **5**, eaau3333 (2019).
69. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
70. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
71. Frankish, A. et al. GENCODE 2021. *Nucleic Acids Res* **49**, D916–D923 (2021).
72. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, <https://doi.org/10.1093/gigascience/giaa151> (2020).
73. 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).

74. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
75. Buttner, M., Ostner, J., Muller, C. L., Theis, F. J. & Schubert, B. scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun.* **12**, 6876 (2021).
76. Crowell, H. L. et al. muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat. Commun.* **11**, 6077 (2020).
77. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
78. Blanchard, J. W. et al. APOE4 impairs myelination via cholesterol dysregulation in oligodendrocytes. *Nature* **611**, 769–779 (2022).
79. Hardwick, S. A. et al. Single-nuclei isoform RNA sequencing unlocks barcoded exon connectivity in frozen brain tissue. *Nat. Biotechnol.* **40**, 1082–1092 (2022).
80. Pérez-González, A. P. et al. The ROSMAP project: aging and neurodegenerative diseases through omic sciences. *Front Neuroinform.* **18**, 1443865 (2024).
81. Samuel L. Wolock, Romain Lopez, & Allon M. Klein. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281–291.e9 (2019).
82. 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).
83. Gerstner, N. et al. GeneTrail 3: advanced high-throughput enrichment analysis. *Nucleic Acids Res* **48**, W515–W520 (2020).
84. Grandke, F., Fehlmann, T., Flotho, M. & Wolff, T. W. A single-cell atlas to map sex-specific gene-expression changes in blood upon neurodegeneration. *GitHub: adrcsc* <https://doi.org/10.5281/zenodo.14641642> (2025).

Acknowledgements

The study has been supported by the Michael J Fox Foundation under grant agreements 14446, 17047 (subc. 125491594) and MJFF-021418. The Schaller-Nikolich-foundation provided further support for the study. Computational resources used within this study were financed through the DFG under project 466168626. The development of the software pipeline has been supported by funds of the EU (Project 101057548-EPIVIN). We thank Shao Ning Pei, I-Jane Chen, Jody Beecher, Yongwei Zhang from Complete Genomics and Zhouchun Shang from MGI-SZ for their support.

Author contributions

F.G., T.F., F.K., M.F., and T.W.W. performed computational analysis. F.G., T.F., and F.K. analyzed the data and assembled the figures. J.S.G. performed the literature research. F.G., T.F., F.K., and A.K. wrote the manuscript with input from all authors. T.F. and Q.S. performed primary data processing. E.Mass and D.M.G. assisted with celltype annotation. D.M.G., O.L., D.C. provided and organized samples. P.H. developed the webserver. E.N.W. performed the CSF biomarker measurement

experiments and K.I.A. supervised this part of the study. E.Meese, D.M.G. and K.P. provided field-expertise and guided clinical interpretation. C.L., Y.L., C.C., Y.Y., J.X., M.J., Z.W., T.W. performed the sequencing experiments. L.L. and Y.H. supervised the sequencing experiments. A.K. and T.W.C. designed and supervised the study.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

A.K. was advisor of the company Firalis, researching RNA-based biomarkers for AD, while preparing the manuscript. C.L., Q.S., Y.L., C.C., Y.Y., J.X., M.J., Z.W., T.W., L.L., and Y.H. work for the company MGI group while working on the project. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-56833-7>.

Correspondence and requests for materials should be addressed to Tony Wyss-Coray or Andreas Keller.

Peer review information *Nature Communications* thanks Jennifer Yokoyama, who co-reviewed with Jake Oxendine, Sergiu Groppa and the other, anonymous, reviewer(s) 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.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025