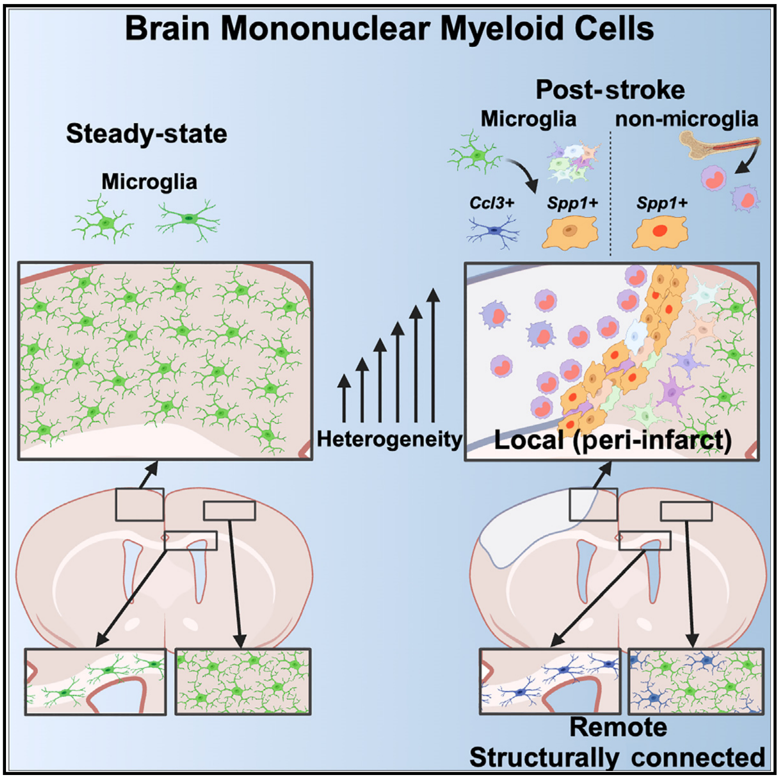
Cell Reports

Resource

Phenotypic and spatial heterogeneity of brain myeloid cells after stroke is associated with cell ontogeny, tissue damage, and brain connectivity

Graphical abstract



Highlights

- Stroke expands brain myeloid cell diversity with global spatial organization
- Brain connectivity to injury underpins locations of remote microglial reactivity
- Infarct-restricted monocyte-derived cells differentiate along dual trajectories
- Macrophages of distinct ontogeny but similar phenotype intermix around the infarct

Authors

Anirudh Patir, Jack Barrington,
Stefan Szymkowiak, ..., Karen Horsburgh,
Prakash Ramachandran, Barry W. McColl

Correspondence

barry.mccoll@ed.ac.uk

In brief

Patir et al. show the identities of resident microglia and immigrant macrophages responding to stroke in brain areas near and remote to damage. Each niche is occupied by discrete ontogeny-phenotype combinations of macrophages. The findings implicate brain connectivity in the regulation of microglia and a global brain macrophage response to focal stroke.

Patir et al., 2024, Cell Reports *43*, 114250
May 28, 2024 © 2024 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.celrep.2024.114250

Cell Reports

CellPress
OPEN ACCESS

Resource

Phenotypic and spatial heterogeneity of brain myeloid cells after stroke is associated with cell ontogeny, tissue damage, and brain connectivity

Anirudh Patir,[1,2,8] Jack Barrington,[1,2,3,8] Stefan Szymkowiak,[1,2] Gaia Brezzo,[1,2] Dana Straus,[1,2] Alessio Alfieri,[1,2] Lucas Lefevre,[1,2] Zhaoyuan Liu,[4] Florent Ginhoux,[4,5] Neil C. Henderson,[6,7] Karen Horsburgh,[2] Prakash Ramachandran,[6] and Barry W. McColl[1,2,9,*]
[1]UK Dementia Research Institute, University of Edinburgh, Edinburgh EH16 4SB, UK
[2]Centre for Discovery Brain Sciences, University of Edinburgh, Edinburgh EH16 4SB, UK
[3]Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh EH16 4SB, UK
[4]Shanghai Institute of Immunology, Department of Immunology and Microbiology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
[5]Singapore Immunology Network, Agency for Science, Technology and Research, Singapore 138648, Singapore
[6]Centre for Inflammation Research, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh EH16 4TJ, UK
[7]MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh EH4 2XU, UK
[8]These authors contributed equally
[9]Lead contact
*Correspondence: barry.mccoll@ed.ac.uk
https://doi.org/10.1016/j.celrep.2024.114250

SUMMARY

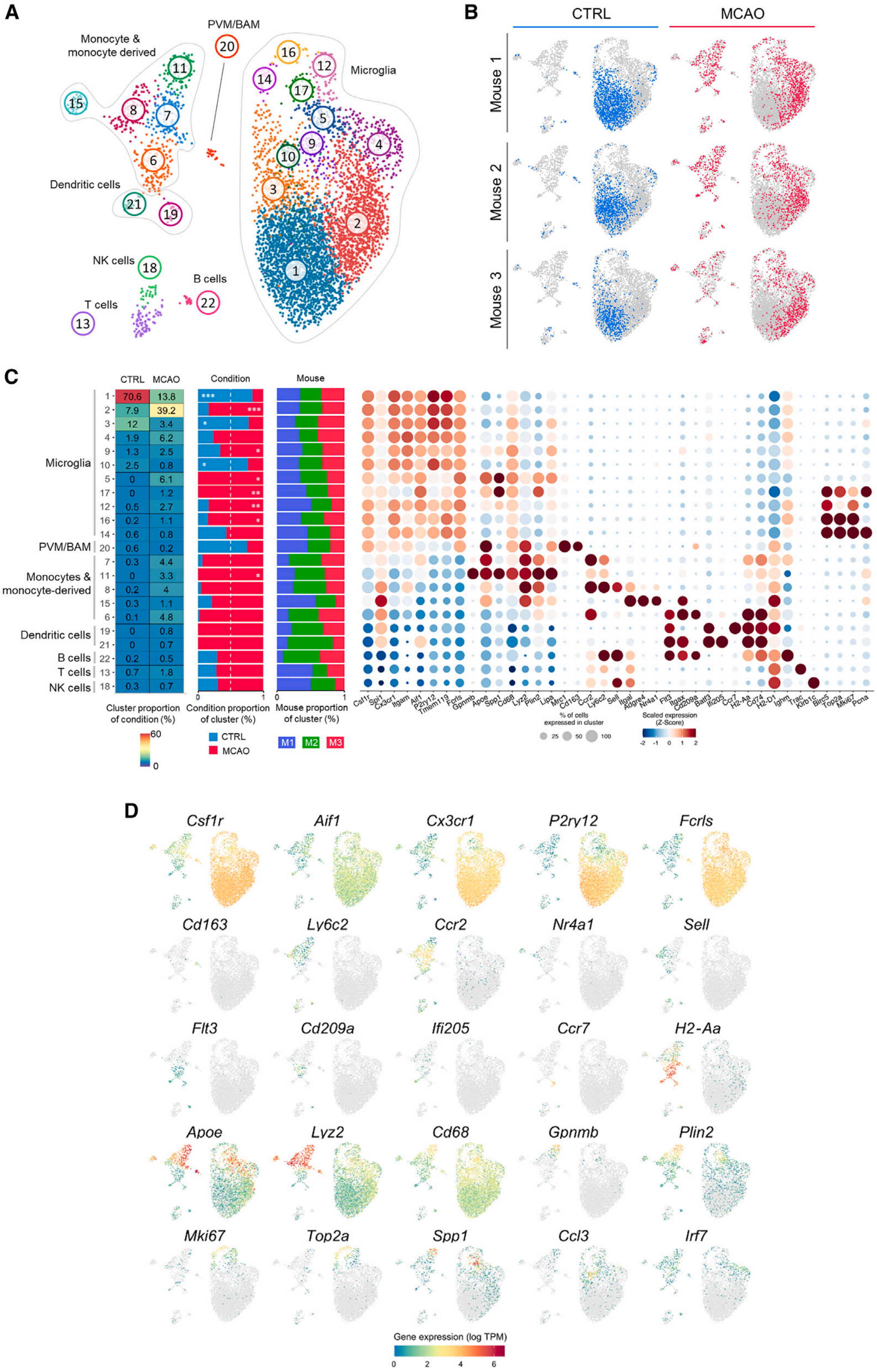
Acute stroke triggers extensive changes to myeloid immune cell populations in the brain that may be targets for limiting brain damage and enhancing repair. Immunomodulatory approaches will be most effective with precise manipulation of discrete myeloid cell phenotypes in time and space. Here, we investigate how stroke alters mononuclear myeloid cell composition and phenotypes at single-cell resolution and key spatial patterns. Our results show that multiple reactive microglial states and monocyte-derived populations contribute to an extensive myeloid cell repertoire in post-stroke brains. We identify important overlaps and distinctions among different cell types/states that involve ontogeny- and spatial-related properties. Notably, brain connectivity with infarcted tissue underpins the pattern of local and remote altered cell accumulation and reactivity. Our discoveries suggest a global but anatomically governed brain myeloid cell response to stroke that comprises diverse phenotypes arising through intrinsic cell ontogeny factors interacting with exposure to spatially organized brain damage and neuro-axonal cues.

INTRODUCTION

Acute ischemic stroke is a leading cause of death and disability worldwide, accounting for the greatest proportion of disease burden for all neurological disorders.[1] Understanding how the brain responds to initial stroke damage to contain the injury, promote brain repair, and enhance plasticity is crucial to inform the design of recovery-enhancing interventions.[2,3] Inflammatory and immune mechanisms are increasingly recognized to connect tissue damage and repair transitions throughout the body,[4] including in the central nervous system (CNS),[5–7] and thus are potential targets for treatments.

Clinical and experimental stroke triggers alterations in trafficking, accumulation, and phenotype of immune cells local to the stroke-injured brain tissue and in systemic compartments.[7–9] Innate immune cells, notably macrophages of resident microglial origin as well as those derived from infiltrating blood-derived monocytes, accumulate in and around damaged brain tissue.[10–13] While some microglial/macrophage activities have the potential to exacerbate tissue damage,[14,15] recent studies increasingly implicate these cells in brain protection and repair during subacute phases of stroke[16–20] and other CNS injuries.[21–23]

The heterogeneity of mononuclear myeloid lineage cells such as microglia, monocytes, and monocyte-derived cells (MdCs) can relate to their differing ontogenies, intra-population differentiation, and cell plasticity and give rise to diverse functional roles.[24] Previous studies used techniques such as transgene-encoded cell reporter mice, *in vivo* tracer labeling, and bone marrow chimeras to show that mixed accumulation of broad myeloid cell populations derived from resident and infiltrating sources occurs after experimental stroke.[10–13,25] Ontogeny-dependent differences in cell phenotype and function, such as phagocytic capacity, have been suggested at the broad population level.[13,15,22,26,27] Nonetheless, the above techniques are constrained for understanding deeper phenotypic heterogeneity within these cell subpopulations and the relationships among them. Spatial organization of mononuclear myeloid cell diversity



(legend on next page)

in post-stroke brain is also poorly understood, in part because most studies examining these cells have focused on infarct and peri-infarct tissue. Recent single-cell RNA sequencing (scRNA-seq)-based studies have produced valuable insight into cell diversity in stroke; however, they exclusively profiled the ipsilateral (peri-infarct) tissue, therefore precluding an understanding of important changes in more distant brain areas.[9,28–30] Remote effects of stroke on neurodegeneration, neuroplasticity, and glial reactivity, often in brain areas connected by white matter tracts to the primary area of damage, are likely critical to global brain network changes after stroke with consequences on motor and cognitive outcomes.[31–37]

Here, we used scRNA-seq and multiplexed single-molecule fluorescence *in situ* hybridization (smFISH) approaches corroborated by monocyte fate-mapping in a permanent middle cerebral artery occlusion (MCAO) stroke model. We focused analysis at 3 days post-MCAO, a critical juncture in the subacute phase, and studied cellular changes in brain areas local and distant to the focal cortical injury. We show that stroke substantially expands the diversity of mononuclear myeloid cells through spatially dependent combinations of altered cell ontogeny, differentiation, and reactivity. We identify location-dependent MdC and microglial phenotypes such as *Gpnmb*-expressing MdCs dominant in intra-/peri-infarct areas and *Ccl* chemokine-enriched microglia in remote brain areas and along connecting white matter tracts. Our findings bring further insight and clarity to existing uncertainties and prompt additional questions around the global brain cellular response to stroke. An openly accessible and interactive application, Stroke-Brain-MySeq, enables further visualization and analysis of the dataset (https://mccoll-lab-uoe.shinyapps.io/shinyapp/).

RESULTS

Flow cytometric overview of MCAO-induced immune cell changes

We acquired cytometric data during cell sorting to provide a contextual overview of major immune cell subclass abundance in response to MCAO (Figure S1 shows the overall experimental workflow). A largely homogeneous $CD45^{+}CD11b^{+}$ population consisting of $CD45^{lo}$ cells was observed in the hemisphere contralateral to MCAO (hereafter referred to as the CTRL condition) (Figure S2). Two additional $CD45^{hi}$ populations were evident in ipsilateral hemisphere samples (hereafter referred to as the MCAO condition), comprising $CD11b^{+}Ly6G^{+}$ (neutrophils) and $CD11b^{+}Ly6G^{-}$ cells. We have shown previously that CD45 surface expression levels can distinguish resident microglia from non-parenchymal myeloid cells.[38] We considered $CD45^{lo}CD11b^{+}$ $Ly6G^{-}$ cells as resident microglia and the $CD45^{hi}CD11b^{+}Ly6G^{-}$ population as blood-derived monocytes and MdCs recruited in response to MCAO (the full gating scheme for quantification is shown in Figure S3). MCAO caused a significant increase in the number of total $CD45^{+}$ cells, total $CD11b^{+}$ cells, $Ly6G^{+}$ granulocytes, microglia, and $CD11b^{+}CD45^{hi}Ly6G^{-}$ monocytes/MdCs, indicating marked innate immune cell infiltration and an increased myeloid cell abundance (Figure S2B). There were negligible $CD3^{+}$ and $CD19^{+}$ cells, indicating limited parenchymal T cell and B cell accumulation (Figures S2A and S2B). Median CD45 intensity of cells specifically within the microglial population, a measure of their broad reactive amplitude, was significantly greater (~45%) in MCAO samples, although it remained ~10-fold lower than the $CD45^{hi}$ infiltrating myeloid population (Figures S2D and S2E), as we found previously.[38] There was also a broader dispersion of individual microglial CD45 intensities within MCAO samples (Figures S2F and S2G), suggesting greater heterogeneity within the microglial population.

Altered mononuclear myeloid cell composition 3 days after MCAO

We conducted scRNA-seq on $CD45^{+}$ sorted cells negative for pan-lymphocyte (CD3 and CD19) and granulocyte (Ly6G) cell surface markers (Figure S4). Following quality control and removal of doublets marked by more than one hashtag oligonucleotide (Figure S5), downstream analysis was performed on a total of 5,213 cells.

Initial unbiased graph-based clustering identified 22 cell clusters organized into four major components when projected by uniform manifold approximation and projection (UMAP)[39] (Figure 1A). Appropriateness of clustering granularity was supported by ratio of global unshifted entropy[40] analysis (Figure S6). Overlaying mouse donors with expression clusters showed that all clusters comprised cells from each donor mouse (Figures 1B and 1C). We annotated the 22 cell clusters based on canonical marker genes for immune cell classes combined with expression of the most variable genes for each cluster (Figures 1C, 1D, and S7; Table S1). As anticipated from the cell sorting strategy, almost all clusters were of myeloid origin, and there was no identifiable granulocyte cluster. Broadly, these clusters were identified as microglia, peri-vascular-associated macrophages/border-associated macrophages (BAMs), dendritic cells (DCs), monocytes, and MdCs as guided by lineage-defining genes such as *Csf1r* combined with broad subpopulation-selective myeloid genes (e.g., *P2ry12*, *Ccr2*, *Flt3*) (Figures 1C; Table S2).

A large group of brain-resident microglia, comprising multiple subclusters (described below), was evident by the high expression of signature genes including *P2ry12*, *Tmem119*, *Gpr34*, and *Fcrls*. $Ly6C^{hi}$ (cluster 8) and $Ly6C^{lo}$ (cluster 15) monocytes were clearly distinguished by their relative expression of genes

Figure 1. Cell clustering and annotation

(A) Cells were clustered based on their transcriptomic profile and projected by UMAP.

(B) UMAPs show cells derived from the CTRL and MCAO experimental conditions and from each of the three donor mice.

(C) From left to right, the images display for each cell cluster (rows) their proportion within each condition (also indicated by the color), their distribution across condition (normalized frequencies compared between conditions; $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$, Student's t test), their distribution across three donor mice (M1, M2, and M3), and the expression of selected markers shown as a dot plot.

(D) Expression profile of selected immune cell marker genes projected by UMAP. PVM, peri-vascular macrophage; BAM, border-associated macrophage. Annotation of major cell groupings in (A) and (C) is based on annotation of individual clusters (see Table S2).

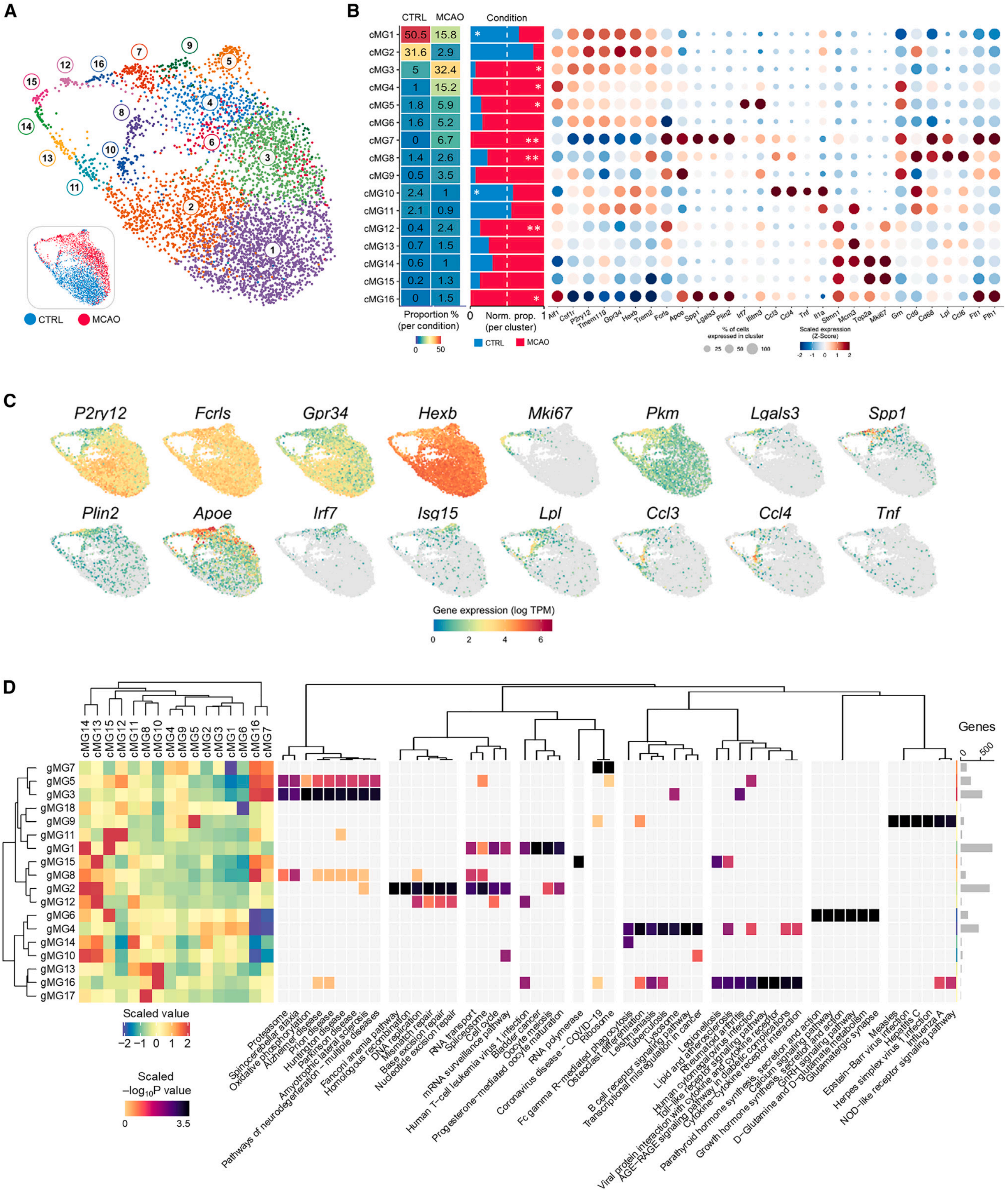


Figure 2. Multiple microglial reactive states associated with MCAO

(A) Microglia were subclustered and based on their transcriptomic profile projected by UMAP. The inset UMAP shows the MCAO and CTRL conditions cells were derived from.

(legend continued on next page)

such as *Ccr2*, *Ly6C*, *Nr4a1*, and *Adgre4*. Clusters adjacent to Ly6C^{hi} monocytes in the UMAP implicated these as differentiating MdCs progressing along macrophage (clusters 7 and 11) and DC (cluster 6) trajectories based on both canonical gene expression (e.g., *Apoe* and *Lyz2* for macrophages and *Flt3* and *Itgax* for DCs) and expression gradients indicative of their maturation from monocytes (e.g., declining but detectable *Ccr2* and *Sell* expression). Two clusters of DCs were identifiable as conventional DCs (cDCs; e.g., high expression of *Batf3*, *Irf8*, and *Ifi205*) and migratory DCs (MigDCs; e.g., high expression of *Ccr7* and *Anxa3*). The relationships among these monocyte-derived and DC clusters are examined in further detail below. An additional cluster of *P2ry12*^{lo}*Fcrls*$^{+}$ cells distinct from microglia expressed high levels of *Mrc1*, *Cd163*, and *Pf4*, which is characteristic of peri-vascular/border macrophages (cluster 20). Several clusters of cells expressed high levels of cell-cycle-associated genes (e.g., *Mki67*, *Top2a*), primarily restricted to those of microglial identity.

Heterogeneity of reactive microglia 3 days after MCAO

We explored MCAO-associated microglial heterogeneity more deeply by subclustering on the population of *P2ry12*$^{+}$*Fcrls*$^{+}$ cells, thus excluding non-parenchymal border macrophages (*P2ry12*$^{-}$*Fcrls*$^{+}$) and non-resident immune cells (*P2ry12*$^{-}$*Fcrls*$^{-}$). 16 microglial cell clusters (cMGs) were identified (Figures 2A–2C; Table S3). Unbiased weighted gene co-expression network analysis (WGCNA) on the 3,890 genes that were differentially expressed across microglia (Figure S8; Table S4) identified gene co-expression modules (gMGs) associated with these cMGs (Figures 2D and S9; Table S5), and pathway enrichment analysis explored biological pathway representation (Figure 2D; Table S5).

Four cMGs (cMG1–3, 6) expressed high levels of established homeostatic microglial genes, such as *Hexb*, *Tmem119*, *Gpr34*, and *P2ry12* contained within gMG4, in the absence of overt activation or cell-cycle-related gene modules (see below). gMG4 was enriched for terms including “P2Y receptors” (adjusted p value [*padj*] = 1.9×10^{-2}) and “transforming growth factor (TGF)-beta receptor signaling” (*padj* = 7.4×10^{-3}) consistent with a homeostatic state. cMG3/6, derived largely from MCAO samples, showed mildly reduced expression of gMG4 homeostatic genes and elevation of gMG7 lysosomal/ribosomal genes compared to cMG1/2, perhaps reflective of a transitional state (Figure 2D).

The remaining cMGs expressed relatively lower levels of the homeostatic microglial gene set (gMG4). cMG7 and cMG16, both derived exclusively from MCAO samples, expressed the lowest levels of the homeostatic gene set and shared marked induction of several large gene modules (gMG3, -5, and -7) collectively comprising over 800 differentially expressed genes, thus indicating the large-scale reactivity of these cells. High expression of gMG3 genes such as *Cst7*, *Spp1*, *Apoe*, and *Lgals3* and various cathepsin genes (Figures 2B–2D) indicated these cell clusters, particularly cMG7, as most resembling the disease-associated microglia (DAM)/activated-response microglia (ARM) state first described in chronic cerebral proteinopathy models. cMG7 and cMG16 also expressed high levels of other gMG3 gene sets involved in “lipid metabolism process” (e.g., *Apoe*, *Acadl*, *Spp1*), cell “oxidation-reduction process” (e.g., *Cybb*, *Hmox1*, *Sod2*), and glycolytic and oxidative “ATP biosynthetic process” (e.g., *Pkm* and mitochondrial complex genes) (Figures 2B–2D), implicating these microglial clusters as extremely metabolically active. The major difference distinguishing cMG7 and cMG16 was high expression in cMG16 of cell-cycle-enriched gene modules (gMG8 and -11) (Figure 2B), perhaps indicating the positioning of cMG16 microglia at the transition between cell cycle and certain differentiated reactive states. Genes defining these clusters (e.g., *Apoe*, *Spp1*) were also significantly elevated in a pseudo-bulk comparison between all MCAO- and CTRL-derived microglia (Table S6). cMG5 microglia were distinctive due to their high expression of interferon (IFN) pathway genes, including *Irf7*, *Isg15*, and *Stat1* and multiple genes from the *Ifi, Ifit*, and *Ifitm* families (all contained within the gMG9 module enriched for “IFN signaling” (*padj* = 2.4×10^{-3}) (Figures 2B–2D). cMG10 microglia were distinguished by their selectively high expression of gMG16 containing chemokine and cytokine genes (gMG16: *Ccl3* and *Ccl4*, *Egr1*, *Il1b*, *Tnf*) (Figures 2B–2D). In support of this, gMG16 was enriched for “chemokine receptors bind chemokines” (*padj* = 2.5×10^{-2}) (Figure 2D). Interestingly, cMG10 was comprised of a significantly greater number of CTRL cells, perhaps indicating an involvement of this state in microglial reactivity remote from the primary infarction, as explored further below. Consistent with our enzyme-free cell isolation methods, there was negligible/zero expression of *ex vivo* artefact-related microglial activation genes (e.g., *Dusp1*, *Fos*, *Hspa1a*)[41] in all clusters (Figure S10).

Microglial cell cycle trajectory induced by MCAO

Several microglial clusters (cMG11–16) expressed high levels of canonical cell cycle marker genes (e.g., *Pcna*, *Mki67*, *Top2a*, *Cdk1*), and WGCNA showed gene modules (e.g., gMG1, -2, and -12) enriched for cell cycle processes in these clusters (Figures 3A–3D; Tables S3–S5), consistent with our previous findings of proliferative microglia early after experimental stroke.[11] We estimated the dynamics of this cell cycle response more thoroughly, first classifying all subclustered microglia according to cell cycle phase (G1, S, G2/M) using Seurat (Figure 3A) and visualizing the expression pattern of the Seurat cell cycle gene set[42]

(B) From left to right, for each cell cluster (rows), images show the proportion of all cells contributing to each of the MCAO and CTRL conditions (also indicated by the color), the proportion of cells within the cell cluster derived from each condition (normalized frequencies compared between conditions; *$p < 0.05$ and **$p < 0.01$, Student’s t test), and a dot plot of the expression of marker genes.

(C) Expression of selected marker genes projected by UMAP.

(D) Using gene co-expression analysis (WGCNA), gene modules (rows) were identified that varied in their expression across the cell clusters (columns). Colors here signify the scaled average expression of gene modules across the various cell clusters. Gene modules were assessed for pathway enrichment (top six KEGG terms shown, *padj* < 0.05). The color captures the scaled −log10 adj. *P* value for the analysis. The rightmost image shows the number of genes within each gene module.

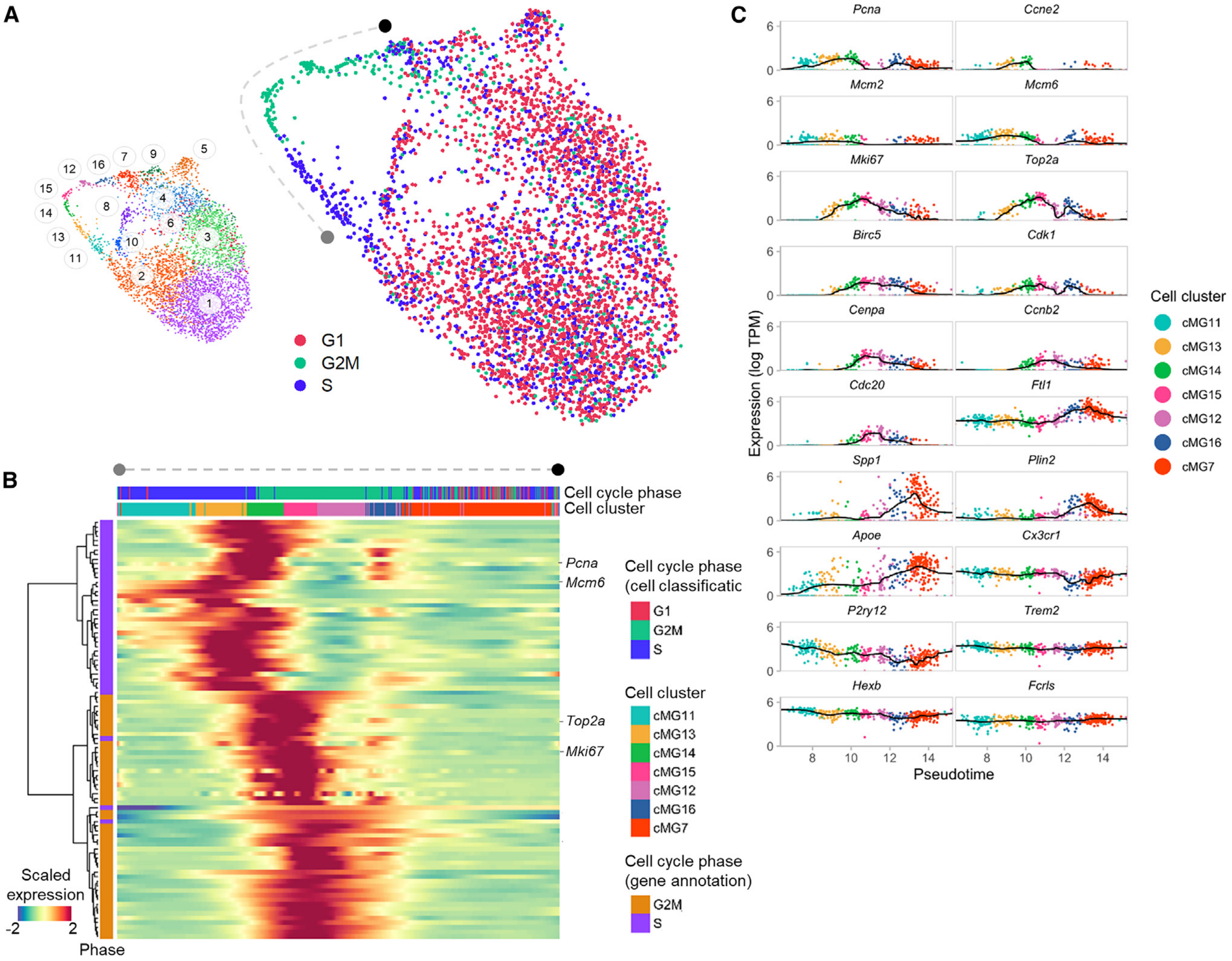


Figure 3. Microglial cell cycle analysis

(A) Cell cycle phase assignment (Seurat algorithm) of subclustered microglia. The transcriptome-based clustered UMAP from Figure 2 is shown for reference (inset). Cells progressing through S to G2M forming an arc-like arrangement are highlighted, with endpoints marked in black and gray.

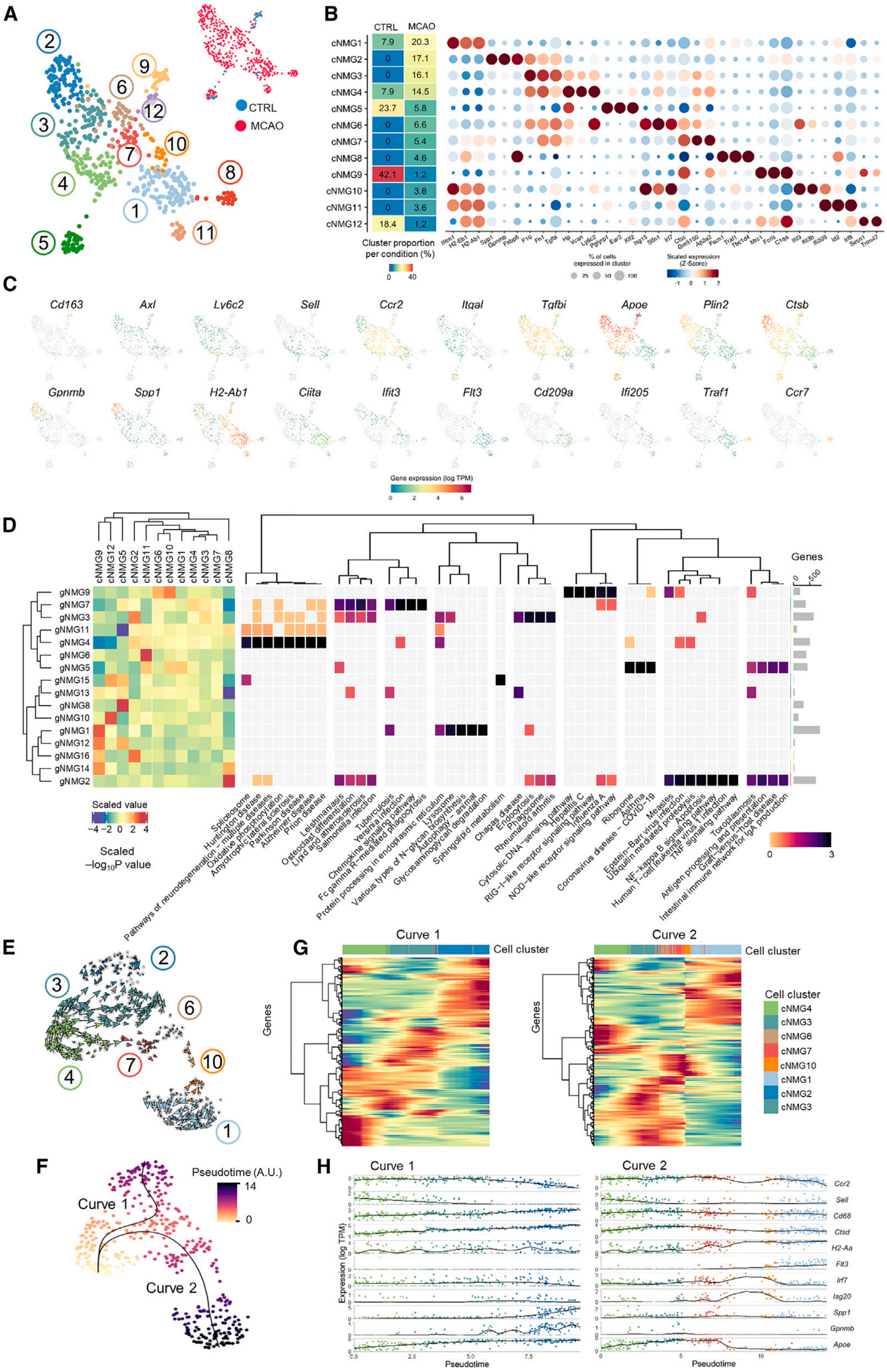
(B) Heatmap of genes significantly (*padj* < 0.05) changing along the arc of cycling cells. Genes (rows) are colored by cell cycle phase annotation, and cells (columns) are colored by cell cycle phase and transcriptome-based cell cluster.

(C) Expression profile of selected genes representing cell cycle phases, microglial reactivity, and homeostasis in cells from indicated clusters arranged along pseudo-time. Each cell is colored based on its transcriptome-based cluster membership (see inset in A).

across cycling microglia and adjacent clusters (Figure 3B). Cells in S and G2/M phases showed a progression across cell clusters 11–16 (Figures 3A and 3B). Pseudo-time analysis was conducted on this arc of proliferative cells, thereby modeling a trajectory from cMG11 > cMG13 > cMG14 > cMG15 > cMG12 > cMG16. We noted with interest the rising expression in cMG16 of many genes that reached maximal expression across pseudo-time in cMG7 reactive microglia (e.g., *Apoe*, *Flt1*, *Plin2*, *Spp1*) and that this coincided with the lowest expression across all clusters of homeostatic genes (e.g., *Cx3cr1*, *P2ry12*, *Hexb*) (Figure 3C). Coupled with the observation that cMG7 microglia and cMG16 also clustered hierarchically by unbiased WGCNA (Figure 2D), these results suggest potential close relationships between the cell cycle and the emergence of selected forms of microglial reactivity.

Comparative analysis of microglial states in stroke and other neurological conditions

We quantitatively assessed the correspondence between our stroke-associated microglial clusters and those in selected mouse models of neurodegenerative disease, aging, or inflammation (Figure S11; Table S7).[43–50] Our stroke model cMG7 expression profile showed a substantial overlap with microglial clusters observed in most other conditions, including aging, amyloidopathy, de/remyelination, and CNS trauma—this cluster comprises the state commonly termed DAM/ARM characterized by genes such as *Lgals3*, *Plin2*, and *Spp1*. The high odds ratio between our stroke model cMG5 and microglial clusters observed in amyloidopathy (Sala-Frigerio, interferon-response microglia [IRM]),[48] aging (Hammond, OA3),[49] neurodegeneration (Mathys, C6),[44] and de/remyelination (Nugent, C8)[46] was



(legend on next page)

consistent with their signature of type 1 IFN pathway genes. These data indicate that IFN-enriched (e.g., IRM) and phagometabolic-enriched (e.g., DAM) microglial expression states are phenotypes common to multiple forms of CNS pathology, both acute and chronic. Contrasting with the above clusters showing widespread disease model intersection, the overlap between our stroke model cMG10 (enriched for chemokine and inflammatory genes such as *Ccl3*, *Ccl4*, and *Il1b*) and other models was more restricted, showing marked similarity only with aging (Hammond, OA2)[49] and CNS trauma (Milich, inflammatory microglia).[45] This may suggest that this microglial expression state is associated with more specific types of injury-/aging-associated tissue damage or distress signals. There was marked overlap of our stroke model cMG16 cell cycle cluster with the transiting-response microglia (TRMs)[48] (which are likely transitioning to ARMs) in CNS amyloidopathy. As described above (Figures 2 and S9), cMG16 co-expresses certain gene modules with the ARM/DAM-like cMG7 after MCAO; their relatedness was also highlighted by individual gene trajectories across pseudo-time (Figure 3). In our cross-model comparison, cMG7 additionally overlaid markedly with the dividing microglial cluster in spinal cord trauma.[45] These cross-model data further implicate direct relationships between certain reactive microglial states and the cell cycle that appear relevant to multiple CNS pathologies.

MCAO induces accumulation and differentiation of monocytes along two major trajectories

We more deeply examined myeloid cells of non-microglial (NMG) identity by re-clustering all NMG cells and used differential gene expression (Figure S12; Table S8) and WGCNA (Figure S13; Table S9) to define cluster characteristics (Figures 4A–4D; Table S10). Two clusters were identified as CNS BAMs. cNMG9 expressed high levels of the gNMG1 gene module genes including *Cd163*, *Clec10a*, and *Vcam1*, whereas cNMG12 expressed relatively greater *Itgb5*, *Aif1*, and major histocompatibility complex (MHC) class II-encoding genes, consistent with recent descriptions of peri-vascular/border macrophages segregating according to MHC class II expression[51] (Figures 4A–4D). cNMG5 cells expressed selectively high levels of *Nr4a1*, *Itgal*, and *Adgre4* (part of the gNMG8 gene module), and negligible *Ly6c2* and *Ccr2*, marking these as the patrolling/marginating Ly6C$^{lo/-}$ monocytes. Conversely, cNMG4 cells expressed very high levels of *Ly6c2*, *Sell*, and *Ccr2*, indicating these as stroke-induced Ly6C^{hi} monocyte immigrants given their almost exclusive derivation from MCAO samples (Figure 4B). cNMG7 retained *Ccr2* with diminished expression of gNMG7 module genes (e.g., *Ly6c2* and *Sell*), suggesting an early monocyte differentiation phenotype. Similarly, the gNMG7 gene set was diminished in cNMG3 cells but alongside the induction of genes (mainly within the gNMG3 module) associated with cell-matrix adhesion and migration (e.g., *Ecm1*, *Fn1*, *Tgfbi*). gNMG3 gene module genes were highly induced in cNMG2 cells that were additionally defined by the high expression of genes involved in macrophage lipid storage (*Plin2*), metabolism (*Lipa*, *Gpnmb*), lipoprotein assembly (*Apoe*, *Apoc1*, *Apobec1*), and cholesterol efflux (*Abca1*, *Npc2*). cNMG2 cells appeared in a highly transcriptionally active state given the extensive set of genes induced (e.g., gNMG3 module contains >500 genes; Figures 4D and S13). Among several enriched functional gene classes, those involved in lysosomal acidification (*Atp6v* family) and enzymatic activity (including proteases, glycosidases, lipases, and nucleases) were the most prominent, substantiated by KEGG analysis revealing “lysosome” and “phagosome” as enriched pathways (Figure 4D). cNMG6 and cNMG10 cells shared high expression of a distinctive IFN response-enriched gNMG9 module (e.g., *Irf7*, *Ifit3*, *Isg15*, and *Oasl2*), with the additional expression of MHC class II antigen-presenting genes (e.g., *H2-Aa* and *H2-Eb1*) contained within the gNMG5 gene module in cNMG10 cells. This MHC class II peptide processing and presentation gene set was also highly expressed by cNMG1 cells but in the absence of the IFN module. We also noted the MHC class II master transcriptional regulator *Ciita* as one of the most highly induced genes in cNMG1, substantiating their antigen-presenting cell phenotype, and with high expression of *Flt3* and *CD209a* but not *Zbtp46*, this is typical of a MoDC identity distinct from the cDC lineage.[52] *Zbtb46*, in contrast, was highly expressed with other canonical genes (e.g., *Ifi205*) as part of the gNMG6 gene module indicating cNMG11 cells as cDCs.[53,54] cNMG8 cells were readily identifiable as MigDCs by their high expression of genes such as *Ccr7* and *Traf1*.[53]

The observations above implicated MCAO-induced differentiation of blood-derived *Ccr2*^{hi}*Sell*$^{+}$*Ly6c2*$^{+}$ monocytes (cNMG4) toward highly differentiated monocyte-derived macrophages (MDMs) (cNMG2; enriched for lipid metabolism, lysosomal activity, matrix interactions) or MoDC (cNMG1; enriched for MHC class II antigen presentation) fates. Trajectory inference using RNA velocity (Velocyto)[55] (Figure 4E) revealed a trajectory of transcriptional change progressing from cNMG4 (monocytes)

Figure 4. Heterogeneity and differentiation trajectories of non-microglial mononuclear myeloid cells associated with MCAO

(A) Non-microglial cells were subclustered and projected by UMAP based on transcriptomic profile. The inset UMAP shows the MCAO and CTRL conditions from which subclustered cells were derived.

(B) From left to right, for each cell cluster (rows), images show the proportion of cells from that cluster of all cells within the MCAO or CTRL conditions (also shown by the color) and a dot plot showing expression of marker genes.

(C) Expression of selected marker genes projected on the cell UMAP.

(D) Using gene co-expression analysis (WGCNA), gene modules (rows) were identified that varied in their expression across the cell clusters (columns). Colors signify the scaled average expression of gene modules across the various cell clusters. Gene modules were assessed for pathway enrichment (top six KEGG terms shown, *padj* < 0.05). The rightmost image shows the number of genes within each gene module.

(E) RNA velocity analysis projected by UMAP of cells from clusters annotated as Ly6C^{hi} monocytes and monocyte-derived cells (Figures 4A and 4B; Table S10).

(F and G) Pseudo-time analysis of the above Ly6C^{hi} monocytes and monocyte-derived cells moving from light to dark with time (F). Heatmaps for each curve showing genes (rows) altered significantly (*padj* < 0.05) across cells ordered along the trajectories and organized by unsupervised hierarchical clustering (G). The color scale represents scaled expression, going from low expression levels (blue) to high (red).

(H) Expression level of selected genes in cells (dots, colored by cell cluster) along pseudo-time trajectories.

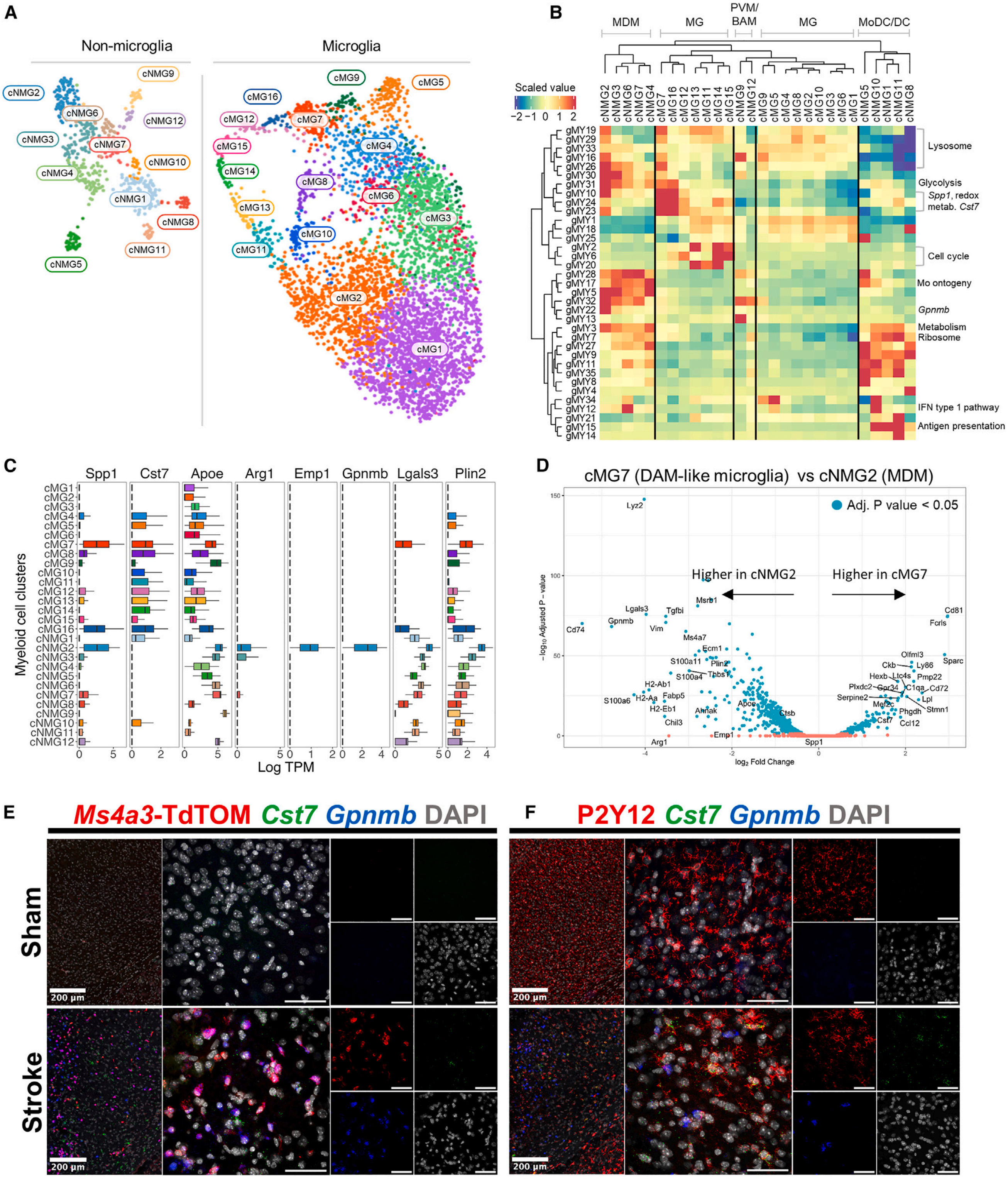


Figure 5. Comparative analysis of reactive microglia and non-microglial mononuclear myeloid cells associated with MCAO

(A) Composite transcriptome-based UMAP projection of subclustered microglia (right; from Figure 2A) and non-microglial myeloid cells (left; from Figure 4A).

(B) Using gene co-expression analysis (WGCNA), gene modules (rows) were identified that varied in their expression across all the microglial and non-microglial myeloid cell clusters (columns). Gene modules (rows) and cell clusters (columns) are arranged by unsupervised hierarchical clustering. Major groupings of cell clusters and selected gene modules with representative genes were manually annotated for orientation (see Table S12 for full pathway analysis).

(legend continued on next page)

toward cNMG2 (MDMs) via cNMG3 and toward cNMG1 (MoDCs) via cNMG7/6/10. Pseudo-time analysis[56] (Figure 4F) aligned cells along two curves representing a two-branched trajectory originating from cNMG4 (monocytes). We determined the genes statistically highly variable across pseudo-time and grouped them according to their hierarchical clustering pattern (Figure 4G). This demonstrated the progressive gradients of expression for gene sets both induced and repressed along pseudo-time with exemplar genes such as *Ccr2*, *Sell*, *Cd68*, and *H2-Aa* (Figure 4H), representative of the continuum in differentiation from monocyte to MDM and MoDC states. The transient induction of the IFN-response gene module (Figures 4G and 4H) further suggested an important role during monocyte-to-MdC transition.

Comparative analysis of reactive microglia and MdCs

We were particularly interested in exploring more formally and at a cluster-specific level how the profiles of reactive microglia and the differentiated MdCs related to each other, given their concurrent appearance and previous challenges in distinguishing these cells by bulk analyses. We integrated the subclustered microglial and NMG datasets (Figure 5A) and conducted WGCNA, thus generating an expression matrix of both shared and distinct gene modules across all mononuclear myeloid cell subclusters (Figures 5B and S14; Tables S11 and S12). Myeloid gene modules (gMYs) were annotated according to KEGG pathway analysis and manual inspection (Figure 5B; Table S12). Cell clusters formed three major top-level hierarchical groups comprising (1) all microglial states and BAMs, (2) Ly6C^{hi} monocytes and MDMs, and (3) cells of DC phenotype (including MoDCs, cDCs, and MigDCs) and Ly6C^{lo} monocytes (Figure 5B). The DC grouping was most distinct among all cell clusters and was characterized as expected by the high expression of antigen processing and presentation genes (gMY4, gMY15). Microglia did not express this antigen-presenting gene module whether of CTRL or MCAO sample origin, or of homeostatic or reactive transcriptional phenotype and more broadly, they shared few gene modules with the DC grouping. In contrast, several microglial clusters co-expressed gene modules (gMY25 and gMY33) with the Ly6C^{hi} monocyte/MDM grouping. gMY33 was enriched for lysosomal function and expressed by most microglial clusters, including those with moderate/high homeostatic gene expression (e.g., cMG1). The most striking similarity was evident between reactive MCAO-enriched cMG7 (DAM/ARM-like) microglia with severely repressed homeostatic gene set expression (gMY1: *Csf1r*, *P2ry12*, and *Trem2*) and the highly differentiated MDM cluster (cNMG2). This relatedness was driven by co-expression of the above lysosomal gene modules (e.g., gMY29 containing *Cd68*, gMY16 containing *Ctsb*) with additional modules (gMY10 and -23) that were notable for the presence of genes commonly considered part of the DAM microglial signature (e.g., *Spp1*, *Apoe*) and genes involved in glycolysis (e.g., gMG31 containing *Pkm*) (Figures 5B and 5C). A noteworthy exception was *Cst7*, which was expressed highly in reactive microglia but not MDMs. The broader transcriptional basis for the overarching distinction of these cell clusters was driven by gene modules reflecting the monocytic origin of MDMs (e.g., gMY17 containing *Ccr2* and *Sell*) and differences in oxidative metabolism and translational activity (e.g., gMY3: enriched for electron transport chain complex and mitochondrial-ribosome genes). The gMY22 expression module comprised some genes expressed selectively in MDMs compared to cMG7 microglia, including *Arg1*, *Emp1*, and *Gpnmb* (Figures 5B and 5C). We also noted several genes (e.g., *Lgals3*, *Plin2*) that, although expressed by both cMG7 microglia and cNMG2 MDMs, had levels that were markedly greater and more consistently expressed across cells within the MDM cluster (Figure 5C). Differential gene expression analysis between these two specific clusters (cMG7 vs. cNMG2) showed the individual genes that most differed (Figure 5D; Table S13), within which we noted *Gpnmb* (MDMs) and *Cst7* (DAM-like microglia) as among the most distinguishing. To corroborate the interpretations of ontogeny from scRNA-seq profiles, we assessed the localization of *Gpnmb* and *Cst7* transcripts in *Ms4a3*^{Cre}*Rosa*$^{LSL\text{-}TdT}$ fate-mapping mice, which express TdTomato (TdT) only in MdCs originating in bone marrow.[57] We observed abundant TdT^{+} cells in the brain after MCAO located in the peri-infarct area (Figure 5E and S15), consistent with scRNA-seq data showing clusters annotated as Mo-related almost exclusively derived from the MCAO (ipsilateral) hemisphere (Figures 1C and 4A). smFISH showed *Gpnmb* transcript puncta restricted to TdT^{+} cells and expression of *Cst7* only localized to TdT^{-} cells (Figure 5F). Accordingly, *Cst7* puncta were localized to cells expressing the microglial marker P2Y12 and *Gpnmb* to P2Y12^{-} cells in the same peri-infarct region, thus validating the ontogeny-phenotype relationships from scRNA-seq profiling (Figures 5F and S15).

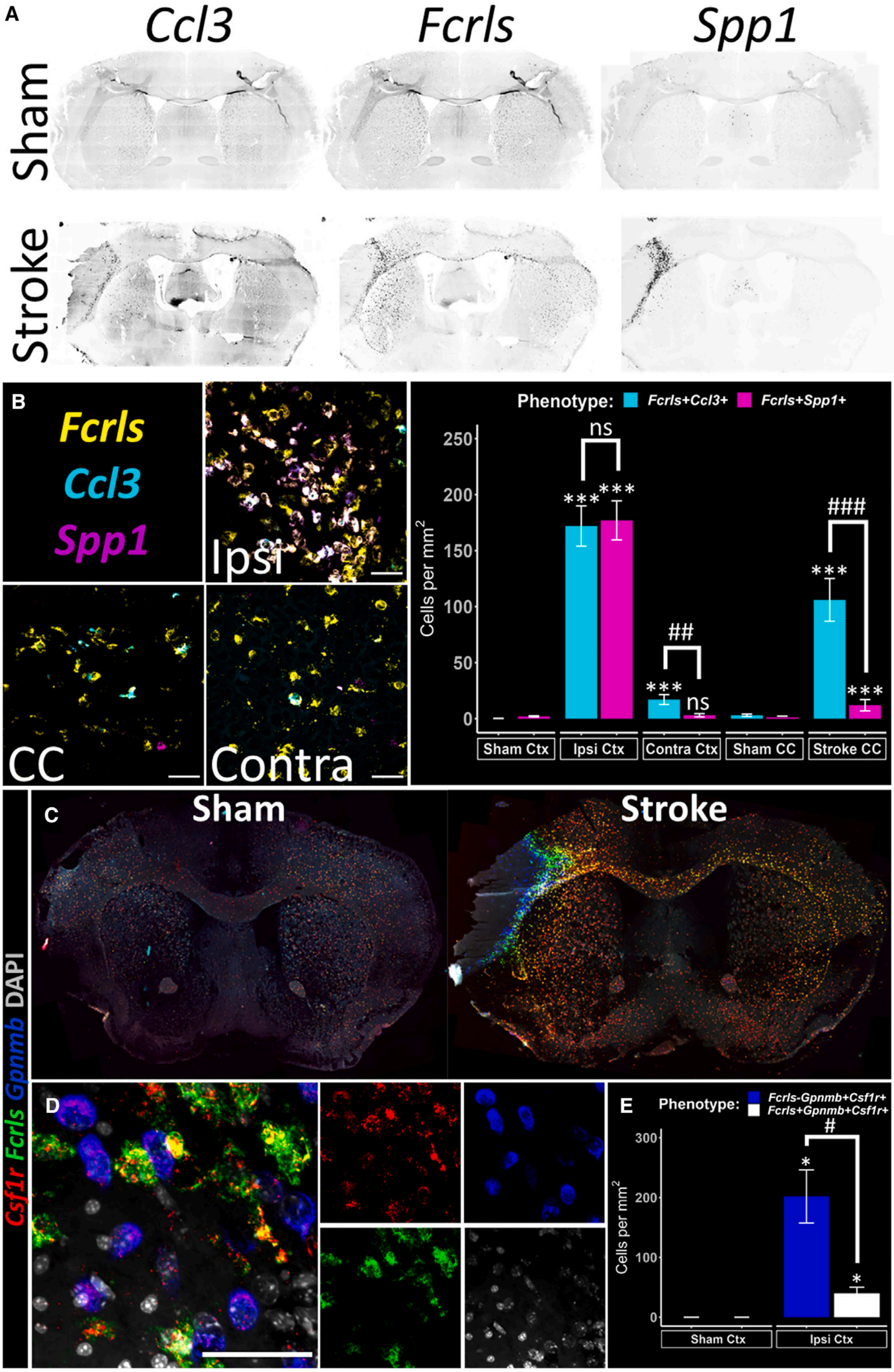
In situ analysis reveals spatial organization of myeloid cell responses associated with structural brain connectivity after MCAO

By sampling cells from hemispheres ipsilateral and contralateral to arterial occlusion, our scRNA-seq expression patterns implicated spatially organized cell types and states. We used multiplexed smFISH to orthogonally validate this spatial heterogeneity, comparing the regional distribution of cMG7 (DAM-like) with cMG10 (chemokine-enriched) using cluster-defining genes *Spp1* and *Ccl3* alongside the microglial marker *Fcrls* (Figures 6A and 6B). *Ccl3* was not expressed at detectable levels in brain tissue in sham-operated controls (Figure 6A). High-magnification confocal imaging of combined *Fcrls*, *Ccl3*, and *Spp1* revealed that *Fcrls* did not co-localize with either cluster-defining marker in sham control brains, indicating that both phenotypes were

(C) Expression distribution of representative genes across each cell cluster.

(D) Analysis of differential gene expression between cMG7 (DAML) and cNMG2 (MDM). Differentially expressed genes (*padj* < 0.05) are colored blue.

(E and F) Images of *Cst7* and *Gpnmb* transcripts in peri-infarct brain area 3 days after stroke (or equivalent in sham) in combination with immunofluorescence for (E) TdTomato (TdTOM) reporter from *Ms4a3*$^{CreERT2/+}$*Rosa*$^{TdT/+}$ mice or (F) P2Y12. Images show low and high magnification (composite and individual channels) from a representative animal from sham and stroke groups. Images from all animals are shown in Figure S15. MDM, monocyte-derived macrophage; MG, microglia; PVM, peri-vascular macrophage; BAM, border-associated macrophage; MoDC, monocyte-derived dendritic cell; DC, dendritic cell.



(legend on next page)

reactive to stroke induction and were not features of a homeostatic or surgically induced microglial landscape (Figures 6B and S16). In contrast, labeling of all three genes was increased in the brain 3 days post-MCAO, albeit with differing spatial distributions (Figure 6A). The increased labeling of both *Fcrls* and *Ccl3* was observed in peri-infarct regions and along brain regions structurally connected to the cortical infarct/peri-infarct, such as transcallosal and cortico-striatal fibers. On the contrary, MCAO-specific *Spp1* labeling was highly enriched to infarct and peri-infarct brain tissue (Figures 6A and 6B), consistent with our scRNA-seq finding that cMG7 was specific to the ipsilateral hemisphere (Figure 6B). Quantification confirmed the presence of *Fcrls*$^+$*Ccl3*$^+$ and *Fcrls*$^+$*Spp1*$^+$ cells around the infarct that were absent in sham controls. *Fcrls*$^+$*Ccl3*$^+$ cells were, however, ~10-fold more abundant than *Fcrls*$^+$*Spp1*$^+$ cells along transcallosal fibers in the corpus callosum ($padj = 2.4 \times 10^{-6}$; Figure 6B) and ~6-fold in the contralateral cortex (Figure 6B). Thus, *Ccl3*$^+$ microglia (i.e., reflective of cMG10) appear both in/around primary infarcted tissue and in perturbed communicating fibers in remote, structurally connected brain regions. *Spp1*$^+$ microglia (i.e., reflective of cMG7), in contrast, are mostly restricted to infarct/peri-infarct tissue 3 days post-MCAO. However, a small but significant increase in cell number above sham tissue in the corpus callosum (Figure 6B; $padj = 0.001$) may indicate that this reactive state also appears in connected regions. In support of the structural brain connectivity largely explaining the spatial pattern of microglial reactive states, areas of remote microglial reactivity revealed by smFISH mapped to brain areas directly connected to the site of primary stroke injury when viewed on matched coronal planes from the viral tracing-based Allen Brain Connectivity Atlas (https://connectivity.brain-map.org/) (Figures S17A and S17B). We also observed morphological signs of reactive P2Y12$^+$ microglia by immunostaining only in local and remote (connected), but not remote (unconnected), brain areas (Figures S17C and S17D).

Spp1 was highly enriched in both cMG7 (DAM-like) and cNMG2 (MDM) cells in our scRNA-seq analyses (Figures 2 and 4) and thus likely labels cells of both ontogenies *in situ*. To better separate these two cell classes, we performed smFISH on the pan-mononuclear phagocyte marker *Csf1r*, the cNMG2-specific marker *Gpnmb* (validated by *Ms4a3*-based fate-mapping above), and microglial-enriched *Fcrls* (Figures 6C–6E). *Gpnmb* was notably absent in sham control brain tissue, whereas *Csf1r* and *Fcrls* co-localized across the entire brain section (Figure 6C). In contrast, *Gpnmb* was strongly expressed within and around the infarct of MCAO brain tissue. High-magnification confocal imaging revealed that *Gpnmb*$^+$ cells expressed low levels of *Csf1r* compared to neighboring *Fcrls*$^+$ cells (Figure 6D), consistent with the finding that *Csf1r* is more highly expressed by microglia and BAMs over cells of bone marrow origin (Figures 1C and 1D). These *Gpnmb*$^+$*Fcrls*$^-$*Csf1r*$^{+(lo)}$ (cNMG2, MDM) cells were ~5-fold more abundant in peri-infarct tissue than *Gpnmb*$^+$*Fcrls*$^+$*Csf1r*$^{+(hi)}$ cells (Figure 6E; $padj = 0.04$). In accordance with our fate-mapping approaches (Figures 5E, 5F, and S15), these data indicate that cNMG2 MDMs recruited to the brain following MCAO are confined to the infarct and peri-infarct tissue, in close proximity to cMG7 microglia that share transcriptional features.

Our scRNA-seq analysis indicated that there were two potential routes to increasing the brain macrophage pool following stroke, i.e., the proliferation of resident microglia and the recruitment of immature monocytes and subsequent maturation into MDMs (Figures 1, 2, and 4). To identify the spatial correlates for these parallel routes, we combined *Fcrls* with the monocyte marker *Ccr2* and the cell cycle marker *Mki67* (Figures 7A–7D). Labeling of *Ccr2* was notably absent in sham-operated controls, and *Mki67* labeling in shams was consistent with known anatomical proliferative cell niches in steady state (periventricular zones and hippocampus) (Figure 7A). *Mki67* rarely co-localized with *Fcrls* in control brains (Figure 7B). 3 days post-MCAO, *Ccr2* was present within and around the infarct, and *Mki67* was found in peri-infarct regions and in structurally connected regions, such as the thalamus (Figure 7A). High-magnification confocal microscopy revealed that ~15% of the *Fcrls*$^+$ cells (~50 of a total ~350 *Fcrls*$^+$ cell/mm^2 pool) in peri-infarct tissue were positive for the cell cycle marker *Mki67*, whereas none of the *Ccr2*$^+$ cells were *Mki67*$^+$, despite the close proximity of *Fcrls*$^+$ and *Ccr2*$^+$ cells (Figures 7B and 7C). No *Ccr2*$^+$ cells were observed in the corpus callosum or contralateral cortex of MCAO brains, although some *Mki67*$^+$*Fcrls*$^+$ cells could be seen in these regions structurally connected to the primary infarct (Figure 7D), consistent with the respective bilateral recovery of these cells in our scRNA-seq experiment (Figure 1). We next looked to compare *in situ* labeling of the *Mki67* gene with immunolabeling for the encoded protein Ki-67 (Figures 7E–7G). In line with transcript-level data, Ki-67$^+$P2Y12$^+$ microglia were rare in sham-operated control cortical brain tissue (Figure 7E). However, a large amount of Ki-67$^+$P2Y12$^+$ microglia were evident in peri-infarct tissue 3 days post-MCAO. Most of these Ki-67$^+$ microglia were more ameboid and expressed lower levels of P2Y12 than neighboring Ki-67$^-$

Figure 6. Spatially resolving microglial and non-microglial transcriptional phenotypes

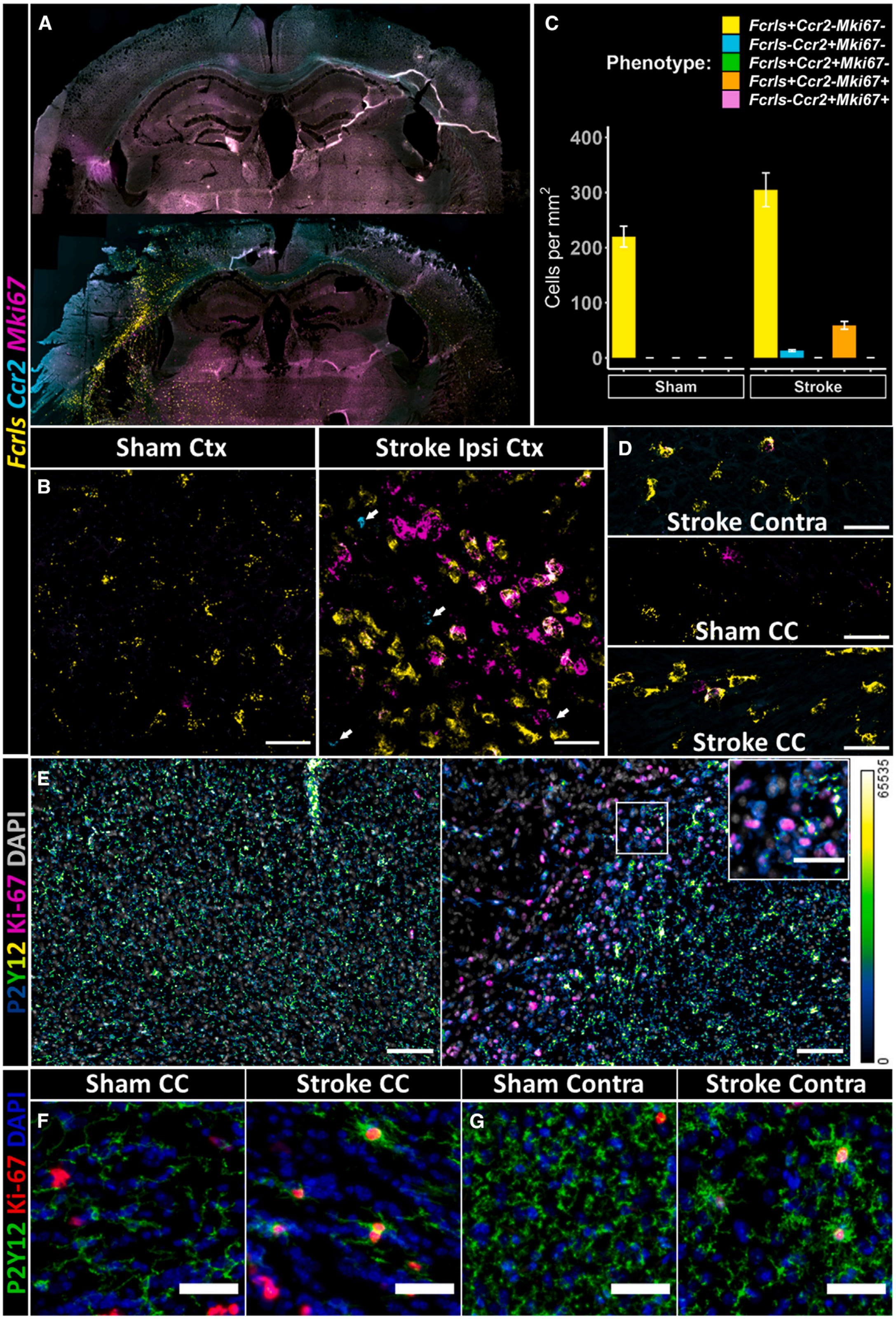
(A) Slide-scanned images of *Ccl3*, *Fcrls*, and *Spp*1 transcripts in (top) sham-operated control brain sections and (bottom) 3 days following stroke. Individual channels are separated and presented in grayscale.

(B) Maximum intensity projections from high-power confocal stacks in ipsilateral peri-infarct (Ipsi), corpus callosum (CC), and contralateral cortex (Contra) of stroke brains (scale bar, 50 μm). *Fcrls*$^+$*Ccl3*$^+$ (cyan) and *Fcrls*$^+$*Spp1*$^+$ (magenta) phenotypes were manually quantified from 3–6 non-overlapping regions in each sham and stroke brain and abundance compared across brain regions (*significantly different from sham, $^\#$significantly different from other transcriptional phenotype; linear mixed model with Holm-Sidak post hoc pairwise comparison; *$padj < 0.05$, **$padj < 0.01$, and ***$padj < 0.001$; $n = 4$ biological replicates [animals] per group). Data are expressed as mean ± SEM.

(C) Slide-scanned images of *Csf1r*, *Fcrls*, and *Gpnmb* transcripts in sham and stroke.

(D) Maximum intensity projection of transcript expression in the peri-infarct region with (left) combined and (right) individual channels shown (scale bar, 50 μm).

(E) *Fcrls*$^-$*Gpnmb*$^+$*Csf1r*$^+$ (blue) and *Fcrls*$^+$*Gpnmb*$^+$*Csf1r*$^+$ (white) phenotypes were manually quantified in sham cortex (left) and peri-infarct of stroke (right) and cell abundance compared across phenotypes (*significantly different in one-sided t test against zero, $^\#$significantly different in paired t test; Bonferroni correction applied for three comparisons; *$padj < 0.05$; $n = 4$ biological replicates [animals] per group). Data are expressed as mean ± SEM.



(legend on next page)

ramified microglia that were a greater distance from the infarct border. Such a reduction in P2Y12 expression is consistent with our transcriptional dataset indicating that microglia downregulation of homeostatic markers is maximal in cycling microglia (Figure 3C). We were also able to visualize the appearance of stroke-specific Ki-67$^+$P2Y12$^+$ hypertrophic microglia in the corpus callosum (Figure 7F) and contralateral cortex (Figure 7G). Together, these data indicate that microglia enter the cell cycle around infarcted tissue and in structurally connected, remote brain regions. We also note with interest the close proximity of proliferating microglia, immigrant monocytes, and differentiated MDMs in peri-infarct zones, which implies shared molecular determinants of mononuclear phagocyte fate at this time point after stroke independent of ontogeny.

Overall, these smFISH results align closely with our scRNA-seq data and, by comparing sham- and MCAO-operated tissue, provide *in situ* confirmation that these reactive phenotypes, whether of proximal or remote location to the infarct, are induced by stroke. For additional corroboration at scRNA-seq level, we mined an independent MCAO dataset conducted on the same droplet-based platform (Zheng et al., GEO: GSE174574)[9] and identified sham-enriched microglia by re-clustering and annotating microglia according to the authors' original marker genes (Figures S18A–S18C; see STAR Methods for full details). This identified Zheng_c0 as a sham-enriched homeostatic microglial cluster. Comparison of normalized and scaled expression of top marker genes between sham-enriched Zheng_c0 and microglial clusters from our study showed a similarity in the expression profiles between our homeostatic microglia (cMG1, cMG2) and Zheng_c0 but a clear distinction between our reactive microglial clusters and Zheng_c0 (Figure S18D). This distinction was evident for reactive microglial clusters in our study whether they were predominant in the ipsilateral or contralateral hemisphere. This provides further independent validation of the stroke-induced specificity of reactive phenotypes we observe in hemispheres ipsilateral and contralateral to stroke, consistent with our smFISH results and emphasizing the global pattern of microglial reactivity to focal ischemic brain damage.

DISCUSSION

The present study demonstrates a globally organized and diversely comprised mononuclear myeloid cell response to experimental stroke. This increased cellular and spatial diversity presumably reflects the array of *de novo* environmental signals and complex tissue changes differing by brain location that require a greater range of cell specializations in the aftermath of stroke, particularly at this transitional time point as an initial wave of damage and danger/distress signals triggers brain-wide events involved in neurorepair, neuroplasticity, and neurodegeneration.[3]

Our scRNA-seq data showed multiple microglial clusters with marked suppression of a panel of canonical homeostatic genes differing according to their combinations of elevated metabolic, inflammatory, and phagolysosomal gene modules. The DAM/ARM-like microglial cluster (cMG7) was notable for its vast array of genes altered, many associated with bioenergetic, redox, and lipid metabolic processes marking this state as particularly reactive. Type 1 IFN-enriched (cMG5) and chemokine-enriched (cMG10) microglia, in contrast, highly elevated entirely distinct gene sets of smaller size. This might suggest the segregation of functional properties through exposure to distinct stimuli and is supported by spatially dependent expression of microglial transcripts characteristic of these discrete scRNA-seq clusters. *Spp1*$^+$*Fcrls*$^+$ microglia were most abundant in the peri-infarct area and elevated in number compared to sham-operated controls. In contrast, *Ccl3*$^+$*Fcrls*$^+$ microglia were evident in both peri-infarct areas and remotely, including in contralateral regions. Beyond this bi-hemispheric pattern, there was a non-random distribution of these contralateral *Ccl3*$^+$*Fcrls*$^+$ microglia, with most located in cortical areas connected by axonal tracts to the infarct and peri-infarct region and along the axonal tracts themselves (e.g., callosal fibers). This specific pattern, and the elevated numbers of *Ccl3*$^+$*Fcrls*$^+$ microglia in MCAO compared to sham-operated mice, also supports that these remote reactive microglia and the cMG10 cluster are a specific response to MCAO and not surgical stress or reflecting a pre-existent baseline microglial state. Given their anatomical distribution, this state of microglial reactivity may be associated with altered structural or functional neural connectivity between regions proximal and distant to the primary stroke injury. Remote neurodegeneration and neuroplasticity in connected brain areas are increasingly recognized to occur after acute stroke and are implicated in long-term motor and cognitive outcomes.[31,33,37]

Our quantitative cross-model/disease analysis showed that stroke-enriched cMG7 reactive microglia shared significant similarity with those observed in CNS amyloidopathy, chronic de/remyelination, spinal cord trauma, axotomy, and aging (Figure S11). Their common profile is defined by the high expression of marker genes such as *Lgals3*, *Spp1*, *Plin2*, and *Apoe:* multiple elevated cathepsin and metabolic genes, reflective of an expression state originally referred to as the DAM/ARM state upon their discovery in CNS amyloidopathy models.[43,48] Two recent

Figure 7. Spatially mapping cell cycle microglia at the transcript and protein levels

(A) Slide-scanned images of *Fcrls*, *Ccr2*, and *Mki67* transcripts in (top) sham-operated control brain sections and (bottom) 3 days following stroke.

(B) Maximum intensity projection from high-power confocal stacks in (left) sham cortex and (right) stroke peri-infarct (scale bar, 50 μm).

(C) *Fcrls*$^+$*Ccr2*$^-$*Mki67*$^-$ (yellow), *Fcrls*$^-$*Ccr2*$^+$*Mki67*$^-$ (cyan), *Fcrls*$^+$*Ccr2*$^+$*Mki67*$^-$ (green), *Fcrls*$^+$*Ccr2*$^-$*Mki67*$^+$ (orange), and *Fcrls*$^-$*Ccr2*$^+$*Mki67*$^+$ (violet) phenotypes were manually quantified from 3–6 non-overlapping regions of sham cortex and stroke peri-infarct in each mouse ($n = 4$ biological replicates [animals] per group). Data are expressed as mean ± SEM.

(D) Representative images of stroke Contra, sham CC, and stroke CC (scale bar, 50 μm).

(E) Protein Ki-67 immunolabeled in (left) sham cortex and (right) stroke peri-infarct alongside the homeostatic microglial marker P2Y12. P2Y12 intensity is visualized on a continuous scale with black/blue representing low, green average, and yellow/white high expression (scale right). Inset is higher resolution of the white box (scale bars: large image, 100 μm and inset, 50 μm).

(F and G) Representative images of Ki-67 and P2Y12 labeling in the CC of (left) sham and (right) stroke (F) and Contra (G) of same groups (scale bar, 25 μm).

studies using a different (transient) stroke model also observed a cluster of microglia/myeloid cells with the characteristic DAM/ARM profile,[9,28] and while some stroke-specific elements have been proposed, there appear many more similarities than differences with chronic disease DAM/ARM. The induction of a microglial state recently demonstrated after spinal cord trauma (termed "migrating microglia"; see Figure 6)[45] is largely indistinguishable from the profile we show here after MCAO, and this spinal trauma phenotype also overlapped closely with chronic disease microglia. In both stroke and spinal cord trauma, the DAM-like state is evident within days of the insult, highlighting that this expression state of microglia is not confined to chronic disease but is a conserved response evident in multiple acute and chronic pathologies. Whether cells converge on this expression state despite exposure to different signals or there is an as-yet unidentified signal common to all pathologies remains to be determined. As noted above, *Spp1*$^{+}$*Fcrls*$^{+}$ microglia were mostly located in peri-infarct areas after MCAO. The *Spp1*$^{+}$ cMG7 microglial cluster we describe here expressed high levels of lipometabolism genes such as *Apoe*, *Fabp5*, *Nceh1*, and *Plin2*. Debris and lipid processing are a feature of peri-infarct remodeling, and our data align with the recent report of lipid-laden and lipid-processing microglia/macrophages of similar transcriptional profiles in the infarcted tissue after transient MCAO.[28] The IFN or interferon-stimulated gene (ISG)-enriched reactive state (cMG5) also appears relatively conserved across pathological conditions. Recent studies have begun to indicate important disease-modifying roles for IFN-enriched microglia in chronic disease and aging.[49,58] In contrast, the CCL chemokine-expressing cMG10 profile was relatively less well represented across disease conditions, albeit evident in aging and spinal cord injury and, interestingly, in acute,[49] but not chronic,[46] demyelination models. Given that the *Ccl3*$^{+}$*Fcrls*$^{+}$ microglia after MCAO were associated spatially with white matter tracts and brain areas remote but connected to the primary infarct, this state may be driven by altered neuroaxonal activity and/or glia-axonal signaling.

Microglial clusters with a cell cycle signature collectively amounted to ∼9% of MCAO-derived microglia, and *in situ* analyses showed that ∼15% of peri-infarct microglia were Ki-67^{+}. Our data clearly show that microglia, but not monocytes/MdCs, are in cell cycle at this time point, consistent with a recent scRNA-seq study in spinal cord trauma.[45] This is despite the close spatial intermixing of microglia and MDMs in peri-infarct tissue and ostensibly similar environmental exposure, thus suggesting that injury-induced cues combine with ontogeny-determined factors to influence cell cycle entry. Inhibiting microglial proliferation after MCAO is detrimental to outcome,[59] whereas protective effects have been reported in chronic disease settings.[60–62] Our data here point to a potentially important relationship between the cell cycle and microglial reactive state development. Co-expression and pseudo-temporal analysis both suggested the G2M-phase-enriched cMG16 as an intermediary state between cell cycle and the DAM-like cMG7. WGCNA of differentially expressed genes showed several gene modules shared by cMG16 and cMG7, thus emphasizing their transcriptional relatedness. CSF1R and TREM2 are both known regulators of microglial proliferation. CSF1R inhibitors lowered, whereas TREM2 agonists augmented, the frequency of DAM/ARM-like cells within the microglial population during CNS amyloidosis.[63,64] Collectively, these data warrant that future studies more deeply explore the connections between the microglial cell cycle and the emergence of reactive states.

The influx of MdCs to the post-stroke brain is well established, and our scRNA-seq data highlight that they contribute >15% of the cells forming the more abundant and phenotypically diverse mononuclear myeloid population in MCAO. A key function of injury-recruited monocytes is to differentiate to more specialized derivatives, which, after stroke, are not well defined. Our clustering and pseudo-temporal analyses show clear differentiation along two major pathways: one producing cells with a highly specialized macrophage phenotype enriched in phagolysosomal and lipometabolism gene modules and the other generating cells of DC phenotype highly enriched for antigen presentation genes. We noted with interest that the MoDC cluster (cNMG1), exclusive to the MCAO condition, was the most abundant that demonstrated a DC phenotype. The complete segregation of MoDC and cDC clusters from all microglia by unsupervised clustering of cell clusters based on gene module co-expression (Figure 5B) reinforces the lack of any DC-like signature of microglia at this time point. A recent study showed that brain-infiltrating myeloid cells are functionally much more capable than microglia of stimulating T cell proliferation after MCAO and that CD11c expression on microglia does not correlate with any difference in their antigen-presenting capacity.[65]

In contrast to the distinctiveness of all microglial clusters and MoDCs, our unbiased WGCNA-based clustering of all myeloid cell clusters showed a striking similarity between the MDM (cNMG2) and DAM/ARM-like microglial (cMG7) clusters based on shared high expression of multiple gene modules. These contained many archetypal DAM/ARM genes and were also enriched for lysosomal, phagocytic, and lipometabolic processes. Our data therefore suggest that a brain macrophage phenotype characterized by a phagolyosomal and lipometabolic transcriptional identity, and thus resembling chronic disease DAM/ARM, arises after stroke via dual sources, i.e., from resident microglia and immigrant monocyte precursors. A recent study showed that brain macrophages expressing MMP12 and osteopontin (encoded by *Spp1*), and thus proposed to reflect key features of the DAM/ARM-like phenotype, were derived from both resident microglia and *Cxcr4*-expressing bone marrow cells.[28] The contribution of dual monocyte and resident brain sources of macrophages accumulating in chronic disease around amyloid deposits has also recently been proposed.[66] Here, our use of *Ms4a3*-based fate-mapping confirmed the resident and immigrant sources of macrophage phenotypes. *Fcrls*$^{-}$*Csf1r*$^{\mathrm{lo}}$*Gpnmb*$^{+}$ cells with a highly ameboid morphology (reflecting *Gpnmb*-expressing MDMs) were almost entirely located in the peri-infarct zone, which was concordant with the almost exclusive origin of Ly6C^{+} MdCs from the MCAO hemisphere in our scRNA-seq dataset (Figure 3). These *Ms4a3*-TdT^{+}*Gpnmb*$^{+}$ MDMs were highly intermixed with peri-infarct *Fcrls*$^{\mathrm{hi}}$ microglia and co-located within the same territory as *Fcrls*$^{\mathrm{hi}}$*Spp1*$^{+}$ microglia. Spatial signals enriched in the peri-infarct zone may shape macrophages in this area toward a conserved phenotype, despite differing origins.

Whether peri-infarct co-located stroke-induced macrophages of distinct ontogenies yet similar effector phenotypes engage in similar tasks functionally is unclear. Due to the restricted tools available, previous studies were limited to comparing functional properties of bulk resident microglia and non-resident monocyte-derived populations but, nonetheless, suggested that the MdCs are more phagocytic.[27] Future studies harnessing advanced tools will be needed to examine comparative functions at more discrete cell state resolution and of functional properties beyond phagocytosis. Moreover, the proximity of microglial and MDM highlights their potential for communication. We previously showed how MDMs can affect microglial function *ex vivo* and the impact of monocytes on long-term microglial phenotype after CNS trauma.[21]

In summary, combinatorial influences of resident and immigrant myeloid ontogeny and brain location relative to primary stroke infarct pathology and its remote connected sites create patterns of spatially organized microglial and MdC states after stroke. We propose that these states will influence global brain function and key stroke outcomes such as cognitive and motor recovery. Further studies achieving more targeted manipulation of specific myeloid states are warranted to empirically define their functional roles.

Limitations of the study

The inclusion of ipsilateral and contralateral hemispheres was a productive approach to discovering the global brain scRNA-seq profile not available to date but can cause challenges in interpreting stroke-specific effects in some cases. This is offset by our validation of scRNA-seq patterns with a published dataset containing sham-operated controls and our *in situ* tissue approaches that included sham-operated controls. This showed marked concordance between scRNA-seq and ISH patterns confirming stroke-specific induction. We studied a single time point 3 days after MCAO, which precludes assessment of how the composition of cell states and types may progress into the chronic phase of stroke important for longer-term outcomes. It is unclear whether some cell states represent transient or discrete fixed phenotypes over a more protracted time frame, and the long-term fate of cells and substates remains unknown currently. A recent study of note suggests that the DAM/ARM-like microglial state is present chronically after CNS trauma,[67] but whether this results from early conversion and then persistence or later but transient conversion is unknown, a distinction with considerable implications for therapeutic manipulation. We used only young adult male mice, and given that sex- and age-related differences in myeloid cell properties are increasingly recognized and stroke is more prevalent in older individuals, these factors will need to be systematically integrated into future studies. Profiling of human ischemic stroke tissue will be required to determine which features apply across species and inform interventional strategies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Experimental stroke model
- METHOD DETAILS
  - Cell sorting and flow cytometry
  - Single cell library preparation and sequencing
  - Single-cell RNA-Seq quality control and pre-processing
  - Cell clustering
  - Comparing cell proportions
  - Generating pseudo-bulk samples from scRNA-Seq to construct gene co-expression networks
  - Differential gene expression
  - Functional enrichment analysis of WGCNA gene modules
  - Cell cycle analysis
  - RNA velocity and pseudotime analysis
  - Cross-model microglial comparison
  - Multiplex fluorescence *in situ* hybridisation
  - Immunofluorescence
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114250.

ACKNOWLEDGMENTS

This work is supported by the UK Dementia Research Institute (award number UK DRI-4005) through UK DRI, Ltd., principally funded by the UK Medical Research Council. We acknowledge funding from the Medical Research Council (MR/R001316/1 and MR/L003384/1) (A.A., B.W.M.); the Leducq Foundation Transatlantic Network of Excellence; Stroke-IMPaCT (19CVD01) (B.W.M.); Alzheimer's Research UK (ARUK-PG2016B-6) (K.H., B.W.M.); the Alzheimer's Research UK Scotland Network (N.C.H., K.H., B.W.M.); Chief Scientist Office Scotland (CGA/18/46) (B.W.M.); and the Stroke Association (TSA PPA 2017/01) (J.B.). N.C.H. is supported by a Wellcome Trust Senior Research Fellowship in Clinical Science (ref. 219542/Z/19/Z). P.R. is supported by an MRC Senior Clinical Fellowship (MR/W015919/1). We thank Adrian Olmos-Alonso for technical contributions and staff at Edinburgh Genomics, FACS, and biological research facilities for their invaluable support to the studies. For the purpose of open access, the author has applied a creative commons attribution (CC BY) license to any author-accepted manuscript version arising from this submission.

AUTHOR CONTRIBUTIONS

Conceptualization, P.R. and B.W.M.; methodology, all authors; investigation, all authors; analysis, all authors; supervision, B.W.M.; project administration, B.W.M.; writing – original draft, A.P., J.B., and B.W.M.; writing – review & editing, all authors; funding acquisition, N.C.H., K.H., and B.W.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 18, 2023
Revised: March 21, 2024
Accepted: May 2, 2024
Published: May 18, 2024

REFERENCES

1. Feigin, V.L., Nichols, E., Alam, T., Bannick, M.S., Beghi, E., Blake, N., Culpepper, W.J., Dorsey, E.R., Elbaz, A., Ellenbogen, R.G., et al. (2019). Global,

regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. *18*, 459–480. https://doi.org/10.1016/S1474-4422(18)30499-X.

2. Corbett, D., Carmichael, S.T., Murphy, T.H., Jones, T.A., Schwab, M.E., Jolkkonen, J., Clarkson, A.N., Dancause, N., Weiloch, T., Johansen-Berg, H., et al. (2017). Enhancing the Alignment of the Preclinical and Clinical Stroke Recovery Research Pipeline: Consensus-Based Core Recommendations From the Stroke Recovery and Rehabilitation Roundtable Translational Working Group. Neurorehabilitation Neural Repair *31*, 699–707. https://doi.org/10.1177/1545968317724285.
3. Carmichael, S.T. (2016). The 3 Rs of Stroke Biology: Radial, Relayed, and Regenerative. Neurotherapeutics *13*, 348–359. https://doi.org/10.1007/s13311-015-0408-0.
4. Wynn, T.A., and Vannella, K.M. (2016). Macrophages in Tissue Repair, Regeneration, and Fibrosis. Immunity *44*, 450–462. https://doi.org/10.1016/j.immuni.2016.02.015.
5. Rayasam, A., Hsu, M., Kijak, J.A., Kissel, L., Hernandez, G., Sandor, M., and Fabry, Z. (2018). Immune responses in stroke: how the immune system contributes to damage and healing after stroke and how this knowledge could be translated to better cures? Immunology *154*, 363–376. https://doi.org/10.1111/imm.12918.
6. Yong, H.Y.F., Rawji, K.S., Ghorbani, S., Xue, M., and Yong, V.W. (2019). The benefits of neuroinflammation for the repair of the injured central nervous system. Cell. Mol. Immunol. *16*, 540–546. https://doi.org/10.1038/s41423-019-0223-3.
7. Iadecola, C., Buckwalter, M.S., and Anrather, J. (2020). Immune responses to stroke: mechanisms, modulation, and therapeutic potential. J. Clin. Invest. *130*, 2777–2788. https://doi.org/10.1172/jci135530.
8. McColl, B.W., Allan, S.M., and Rothwell, N.J. (2007). Systemic inflammation and stroke: aetiology, pathology and targets for therapy. Biochem. Soc. Trans. *35*, 1163–1165. https://doi.org/10.1042/bst0351163.
9. Zheng, K., Lin, L., Jiang, W., Chen, L., Zhang, X., Zhang, Q., Ren, Y., and Hao, J. (2022). Single-cell RNA-seq reveals the transcriptional landscape in ischemic stroke. J. Cerebr. Blood Flow Metabol. *42*, 56–73. https://doi.org/10.1177/0271678x211026770.
10. Schilling, M., Besselmann, M., Leonhard, C., Mueller, M., Ringelstein, E.B., and Kiefer, R. (2003). Microglial activation precedes and predominates over macrophage infiltration in transient focal cerebral ischemia: a study in green fluorescent protein transgenic bone marrow chimeric mice. Exp. Neurol. *183*, 25–33.
11. Denes, A., Vidyasagar, R., Feng, J., Narvainen, J., McColl, B.W., Kauppinen, R.A., and Allan, S.M. (2007). Proliferating resident microglia after focal cerebral ischaemia in mice. J. Cerebr. Blood Flow Metabol. *27*, 1941–1953. https://doi.org/10.1038/sj.jcbfm.9600495.
12. Garcia-Bonilla, L., Faraco, G., Moore, J., Murphy, M., Racchumi, G., Srinivasan, J., Brea, D., Iadecola, C., and Anrather, J. (2016). Spatio-temporal profile, phenotypic diversity, and fate of recruited monocytes into the post-ischemic brain. J. Neuroinflammation *13*, 285. https://doi.org/10.1186/s12974-016-0750-0.
13. Kronenberg, G., Uhlemann, R., Richter, N., Klempin, F., Wegner, S., Staerck, L., Wolf, S., Uckert, W., Kettenmann, H., Endres, M., and Gertz, K. (2018). Distinguishing features of microglia- and monocyte-derived macrophages after stroke. Acta Neuropathol. *135*, 551–568. https://doi.org/10.1007/s00401-017-1795-6.
14. Li, T., Zhao, J., Xie, W., Yuan, W., Guo, J., Pang, S., Gan, W.B., Gómez-Nicola, D., and Zhang, S. (2021). Specific depletion of resident microglia in the early stage of stroke reduces cerebral ischemic damage. J. Neuroinflammation *18*, 81. https://doi.org/10.1186/s12974-021-02127-w.
15. Park, K.W., Ju, H., Kim, I.-d., Cave, J.W., Guo, Y., Wang, W., Wu, Z., and Cho, S. (2022). Delayed Infiltration of Peripheral Monocyte Contributes to Phagocytosis and Transneuronal Degeneration in Chronic Stroke. Stroke *53*, 2377–2388. https://doi.org/10.1161/STROKEAHA.122.038701.
16. Pedragosa, J., Miró-Mur, F., Otxoa-de-Amezaga, A., Justicia, C., Ruíz-Jaén, F., Ponsaerts, P., Pasparakis, M., and Planas, A.M. (2020). CCR2 deficiency in monocytes impairs angiogenesis and functional recovery after ischemic stroke in mice. J. Cerebr. Blood Flow Metabol. *40*, S98–S116. https://doi.org/10.1177/0271678x20909055.
17. Wattananit, S., Tornero, D., Graubardt, N., Memanishvili, T., Monni, E., Tatarishvili, J., Miskinyte, G., Ge, R., Ahlenius, H., Lindvall, O., et al. (2016). Monocyte-Derived Macrophages Contribute to Spontaneous Long-Term Functional Recovery after Stroke in Mice. J. Neurosci. *36*, 4182–4195. https://doi.org/10.1523/jneurosci.4317-15.2016.
18. Szalay, G., Martinecz, B., Lénárt, N., Környei, Z., Orsolits, B., Judák, L., Császár, E., Fekete, R., West, B.L., Katona, G., et al. (2016). Microglia protect against brain injury and their selective elimination dysregulates neuronal network activity after stroke. Nat. Commun. *7*, 11499. https://doi.org/10.1038/ncomms11499.
19. Otxoa-de-Amezaga, A., Miró-Mur, F., Pedragosa, J., Gallizioli, M., Justicia, C., Gaja-Capdevila, N., Ruíz-Jaen, F., Salas-Perdomo, A., Bosch, A., Calvo, M., et al. (2019). Microglial cell loss after ischemic stroke favors brain neutrophil accumulation. Acta Neuropathol. *137*, 321–341. https://doi.org/10.1007/s00401-018-1954-4.
20. Gliem, M., Schwaninger, M., and Jander, S. (2016). Protective features of peripheral monocytes/macrophages in stroke. Biochim. Biophys. Acta *1862*, 329–338. https://doi.org/10.1016/j.bbadis.2015.11.004.
21. Greenhalgh, A.D., Zarruk, J.G., Healy, L.M., Baskar Jesudasan, S.J., Jhelum, P., Salmon, C.K., Formanek, A., Russo, M.V., Antel, J.P., McGavern, D.B., et al. (2018). Peripherally derived macrophages modulate microglial function to reduce inflammation after CNS injury. PLoS Biol. *16*, e2005264. https://doi.org/10.1371/journal.pbio.2005264.
22. Mastorakos, P., Mihelson, N., Luby, M., Burks, S.R., Johnson, K., Hsia, A.W., Witko, J., Frank, J.A., Latour, L., and McGavern, D.B. (2021). Temporally distinct myeloid cell responses mediate damage and repair after cerebrovascular injury. Nat. Neurosci. *24*, 245–258. https://doi.org/10.1038/s41593-020-00773-6.
23. Schwartz, M. (2010). "Tissue-repairing" blood-derived macrophages are essential for healing of the injured spinal cord: From skin-activated macrophages to infiltrating blood-derived cells? Brain Behav. Immun. *24*, 1054–1057. https://doi.org/10.1016/j.bbi.2010.01.010.
24. Benakis, C., Garcia-Bonilla, L., Iadecola, C., and Anrather, J. (2014). The role of microglia and myeloid immune cells in acute cerebral ischemia. Front. Cell. Neurosci. *8*, 461. https://doi.org/10.3389/fncel.2014.00461.
25. Werner, Y., Mass, E., Ashok Kumar, P., Ulas, T., Händler, K., Horne, A., Klee, K., Lupp, A., Schütz, D., Saaber, F., et al. (2020). Cxcr4 distinguishes HSC-derived monocytes from microglia and reveals monocyte immune responses to experimental stroke. Nat. Neurosci. *23*, 351–362. https://doi.org/10.1038/s41593-020-0585-y.
26. Rajan, W.D., Wojtas, B., Gielniewski, B., Gieryng, A., Zawadzka, M., and Kaminska, B. (2019). Dissecting functional phenotypes of microglia and macrophages in the rat brain after transient cerebral ischemia. Glia *67*, 232–245. https://doi.org/10.1002/glia.23536.
27. Ritzel, R.M., Patel, A.R., Grenier, J.M., Crapser, J., Verma, R., Jellison, E.R., and McCullough, L.D. (2015). Functional differences between microglia and monocytes after ischemic stroke. J. Neuroinflammation *12*, 106. https://doi.org/10.1186/s12974-015-0329-1.
28. Beuker, C., Schafflick, D., Strecker, J.-K., Heming, M., Li, X., Wolbert, J., Schmidt-Pogoda, A., Thomas, C., Kuhlmann, T., Aranda-Pardos, I., et al. (2022). Stroke induces disease-specific myeloid cells in the brain parenchyma and pia. Nat. Commun. *13*, 945. https://doi.org/10.1038/s41467-022-28593-1.
29. Guo, K., Luo, J., Feng, D., Wu, L., Wang, X., Xia, L., Tao, K., Wu, X., Cui, W., He, Y., et al. (2021). Single-Cell RNA Sequencing With Combined Use of Bulk RNA Sequencing to Reveal Cell Heterogeneity and Molecular Changes at Acute Stage of Ischemic Stroke in Mouse Cortex Penumbra Area. Front. Cell Dev. Biol. *9*, 624711. https://doi.org/10.3389/fcell.2021.624711.

30. Li, X., Lyu, J., Li, R., Jain, V., Shen, Y., Del Águila, Á., Hoffmann, U., Sheng, H., and Yang, W. (2022). Single-cell transcriptomic analysis of the immune cell landscape in the aged mouse brain after ischemic stroke. J. Neuroinflammation *19*, 83. https://doi.org/10.1186/s12974-022-02447-5.

31. Carmichael, S.T. (2003). Plasticity of cortical projections after stroke. Neuroscientist *9*, 64–75. https://doi.org/10.1177/1073858402239592.

32. Dancause, N., Barbay, S., Frost, S.B., Plautz, E.J., Chen, D., Zoubina, E.V., Stowe, A.M., and Nudo, R.J. (2005). Extensive Cortical Rewiring after Brain Injury. J. Neurosci. *25*, 10167–10179. https://doi.org/10.1523/JNEUROSCI.3256-05.2005.

33. Hall, G.R., Kaiser, M., and Farr, T.D. (2021). Functional Connectivity Change in Response to Stroke Is Comparable Across Species: From Mouse to Man. Stroke Vasc. Interv. Neurol. *52*, 2961–2963. https://doi.org/10.1161/STROKEAHA.121.034097.

34. Dihné, M., Grommes, C., Lutzenburg, M., Witte, O.W., and Block, F. (2002). Different Mechanisms of Secondary Neuronal Damage in Thalamic Nuclei After Focal Cerebral Ischemia in Rats. Stroke *33*, 3006–3011. https://doi.org/10.1161/01.STR.0000039406.64644.CB.

35. Morioka, T., Kalehua, A.N., and Streit, W.J. (1993). Characterization of microglial reaction after middle cerebral artery occlusion in rat brain. J. Comp. Neurol. *327*, 123–132. https://doi.org/10.1002/cne.903270110.

36. Shi, K., Tian, D.C., Li, Z.G., Ducruet, A.F., Lawton, M.T., and Shi, F.D. (2019). Global brain inflammation in stroke. Lancet Neurol. *18*, 1058–1066. https://doi.org/10.1016/s1474-4422(19)30078-x.

37. Zhang, J., Zhang, Y., Xing, S., Liang, Z., and Zeng, J. (2012). Secondary Neurodegeneration in Remote Regions After Focal Cerebral Infarction. Stroke *43*, 1700–1705. https://doi.org/10.1161/STROKEAHA.111.632448.

38. Davies, C.L., Patir, A., and McColl, B.W. (2019). Myeloid Cell and Transcriptome Signatures Associated With Inflammation Resolution in a Model of Self-Limiting Acute Brain Inflammation. Front. Immunol. *10*, 1048. https://doi.org/10.3389/fimmu.2019.01048.

39. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. *37*, 38–44. https://doi.org/10.1038/nbt.4314.

40. Liu, B., Li, C., Li, Z., Wang, D., Ren, X., and Zhang, Z. (2020). An entropy-based metric for assessing the purity of single cell populations. Nat. Commun. *11*, 3155. https://doi.org/10.1038/s41467-020-16904-3.

41. Marsh, S.E., Walker, A.J., Kamath, T., Dissing-Olesen, L., Hammond, T.R., de Soysa, T.Y., Young, A.M.H., Murphy, S., Abdulraouf, A., Nadaf, N., et al. (2022). Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. Nat. Neurosci. *25*, 306–316. https://doi.org/10.1038/s41593-022-01022-8.

42. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196. https://doi.org/10.1126/science.aad0501.

43. Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T.K., David, E., Baruch, K., Lara-Astaiso, D., Toth, B., et al. (2017). A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. Cell *169*, 1276–1290.e17. https://doi.org/10.1016/j.cell.2017.05.018.

44. Mathys, H., Adaikkan, C., Gao, F., Young, J.Z., Manet, E., Hemberg, M., De Jager, P.L., Ransohoff, R.M., Regev, A., and Tsai, L.-H. (2017). Temporal Tracking of Microglia Activation in Neurodegeneration at Single-Cell Resolution. Cell Rep. *21*, 366–380. https://doi.org/10.1016/j.celrep.2017.09.039.

45. Milich, L.M., Choi, J.S., Ryan, C., Cerqueira, S.R., Benavides, S., Yahn, S.L., Tsoulfas, P., and Lee, J.K. (2021). Single-cell analysis of the cellular heterogeneity and interactions in the injured mouse spinal cord. J. Exp. Med. *218*, e20210040. https://doi.org/10.1084/jem.20210040.

46. Nugent, A.A., Lin, K., van Lengerich, B., Lianoglou, S., Przybyla, L., Davis, S.S., Llapashtica, C., Wang, J., Kim, D.J., Xia, D., et al. (2020). TREM2 Regulates Microglial Cholesterol Metabolism upon Chronic Phagocytic Challenge. Neuron *105*, 837–854.e9. https://doi.org/10.1016/j.neuron.2019.12.007.

47. Tay, T.L., Sagar, Dautzenberg, J., Grün, D., and Prinz, M. (2018). Unique microglia recovery population revealed by single-cell RNAseq following neurodegeneration. Acta Neuropathol. Commun. *6*, 87. https://doi.org/10.1186/s40478-018-0584-3.

48. Sala Frigerio, C., Wolfs, L., Fattorelli, N., Thrupp, N., Voytyuk, I., Schmidt, I., Mancuso, R., Chen, W.-T., Woodbury, M.E., Srivastava, G., et al. (2019). The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques. Cell Rep. *27*, 1293–1306.e6. https://doi.org/10.1016/j.celrep.2019.03.099.

49. Hammond, T.R., Dufort, C., Dissing-Olesen, L., Giera, S., Young, A., Wysoker, A., Walker, A.J., Gergits, F., Segel, M., Nemesh, J., et al. (2019). Single-Cell RNA Sequencing of Microglia throughout the Mouse Lifespan and in the Injured Brain Reveals Complex Cell-State Changes. Immunity *50*, 253–271.e6. https://doi.org/10.1016/j.immuni.2018.11.004.

50. Sousa, C., Golebiewska, A., Poovathingal, S.K., Kaoma, T., Pires-Afonso, Y., Martina, S., Coowar, D., Azuaje, F., Skupin, A., Balling, R., et al. (2018). Single-cell transcriptomics reveals distinct inflammation-induced microglia signatures. EMBO Rep. *19*, e46171. https://doi.org/10.15252/embr.201846171.

51. Van Hove, H., Martens, L., Scheyltjens, I., De Vlaminck, K., Pombo Antunes, A.R., De Prijck, S., Vandamme, N., De Schepper, S., Van Isterdael, G., Scott, C.L., et al. (2019). A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by ontogeny and tissue environment. Nat. Neurosci. *22*, 1021–1035. https://doi.org/10.1038/s41593-019-0393-4.

52. Menezes, S., Melandri, D., Anselmi, G., Perchet, T., Loschko, J., Dubrot, J., Patel, R., Gautier, E.L., Hugues, S., Longhi, M.P., et al. (2016). The Heterogeneity of Ly6C(hi) Monocytes Controls Their Differentiation into iNOS(+) Macrophages or Monocyte-Derived Dendritic Cells. Immunity *45*, 1205–1218. https://doi.org/10.1016/j.immuni.2016.12.001.

53. Miller, J.C., Brown, B.D., Shay, T., Gautier, E.L., Jojic, V., Cohain, A., Pandey, G., Leboeuf, M., Elpek, K.G., Helft, J., et al. (2012). Deciphering the transcriptional network of the dendritic cell lineage. Nat. Immunol. *13*, 888–899. http://www.nature.com/ni/journal/v13/n9/abs/ni.2370.html#supplementary-information.

54. Brown, C.C., Gudjonson, H., Pritykin, Y., Deep, D., Lavallée, V.-P., Mendoza, A., Fromme, R., Mazutis, L., Ariyan, C., Leslie, C., et al. (2019). Transcriptional Basis of Mouse and Human Dendritic Cell Heterogeneity. Cell *179*, 846–863.e24. https://doi.org/10.1016/j.cell.2019.09.035.

55. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498. https://doi.org/10.1038/s41586-018-0414-6.

56. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genom. *19*, 477. https://doi.org/10.1186/s12864-018-4772-0.

57. Liu, Z., Gu, Y., Chakarov, S., Bleriot, C., Kwok, I., Chen, X., Shin, A., Huang, W., Dress, R.J., Dutertre, C.-A., et al. (2019). Fate Mapping via Ms4a3-Expression History Traces Monocyte-Derived Cells. Cell *178*, 1509–1525.e19. https://doi.org/10.1016/j.cell.2019.08.009.

58. Roy, E.R., Chiu, G., Li, S., Propson, N.E., Kanchi, R., Wang, B., Coarfa, C., Zheng, H., and Cao, W. (2022). Concerted type I interferon signaling in microglia and neural cells promotes memory impairment associated with amyloid β plaques. Immunity *55*, 879–894.e6. https://doi.org/10.1016/j.immuni.2022.03.018.

59. Lalancette-Hébert, M., Gowing, G., Simard, A., Weng, Y.C., and Kriz, J. (2007). Selective Ablation of Proliferating Microglial Cells Exacerbates Ischemic Injury in the Brain. J. Neurosci. *27*, 2596–2605. https://doi.org/10.1523/jneurosci.5360-06.2007.
60. Olmos-Alonso, A., Schetters, S.T.T., Sri, S., Askew, K., Mancuso, R., Vargas-Caballero, M., Holscher, C., Perry, V.H., and Gomez-Nicola, D. (2016). Pharmacological targeting of CSF1R inhibits microglial proliferation and prevents the progression of Alzheimer's-like pathology. Brain *139*, 891–907. https://doi.org/10.1093/brain/awv379.
61. Gómez-Nicola, D., Fransen, N.L., Suzzi, S., and Perry, V.H. (2013). Regulation of Microglial Proliferation during Chronic Neurodegeneration. J. Neurosci. *33*, 2481–2493. https://doi.org/10.1523/jneurosci.4440-12.2013.
62. Mancuso, R., Fryatt, G., Cleal, M., Obst, J., Pipi, E., Monzón-Sandoval, J., Ribe, E., Winchester, L., Webber, C., Nevado, A., et al. (2019). CSF1R inhibitor JNJ-40346527 attenuates microglial proliferation and neurodegeneration in P301S mice. Brain *142*, 3243–3264. https://doi.org/10.1093/brain/awz241.
63. Wang, S., Mustafa, M., Yuede, C.M., Salazar, S.V., Kong, P., Long, H., Ward, M., Siddiqui, O., Paul, R., Gilfillan, S., et al. (2020). Anti-human TREM2 induces microglia proliferation and reduces pathology in an Alzheimer's disease model. J. Exp. Med. *217*, e20200785. https://doi.org/10.1084/jem.20200785.
64. Hu, Y., Fryatt, G.L., Ghorbani, M., Obst, J., Menassa, D.A., Martin-Estebane, M., Muntslag, T.A.O., Olmos-Alonso, A., Guerrero-Carrasco, M., Thomas, D., et al. (2021). Replicative senescence dictates the emergence of disease-associated microglia and contributes to Aβ pathology. Preprint at bioRxiv. https://doi.org/10.1101/2021.03.22.436174.
65. Gallizioli, M., Miró-Mur, F., Otxoa-de-Amezaga, A., Cugota, R., Salas-Perdomo, A., Justicia, C., Brait, V.H., Ruiz-Jaén, F., Arbaizar-Rovirosa, M., Pedragosa, J., et al. (2020). Dendritic Cells and Microglia Have Non-redundant Functions in the Inflamed Brain with Protective Effects of Type 1 cDCs. Cell Rep. *33*, 108291. https://doi.org/10.1016/j.celrep.2020.108291.
66. Silvin, A., Uderhardt, S., Piot, C., Da Mesquita, S., Yang, K., Geirsdottir, L., Mulder, K., Eyal, D., Liu, Z., Bridlance, C., et al. (2022). Dual ontogeny of disease-associated microglia and disease inflammatory macrophages in aging and neurodegeneration. Immunity *55*, 1448–1465.e6. https://doi.org/10.1016/j.immuni.2022.07.004.
67. Hakim, R., Zachariadis, V., Sankavaram, S.R., Han, J., Harris, R.A., Brundin, L., Enge, M., and Svensson, M. (2021). Spinal Cord Injury Induces Permanent Reprogramming of Microglia into a Disease-Associated State Which Contributes to Functional Recovery. J. Neurosci. *41*, 8441–8459. https://doi.org/10.1523/JNEUROSCI.0860-21.2021.
68. Liu, Z., Gu, Y., Chakarov, S., Bleriot, C., Chen, X., Shin, A., Huang, W., Dress, R.J., Dutertre, C.-A., Schlitzer, A., et al. (2019). Fate mapping via Ms4a3 expression history traces monocyte-derived cells. Preprint at bioRxiv. https://doi.org/10.1101/652032.
69. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. https://doi.org/10.1038/ncomms14049.
70. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.
71. Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol. *19*, 224. https://doi.org/10.1186/s13059-018-1603-1.
72. Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods *14*, 865–868. https://doi.org/10.1038/nmeth.4380.
73. McGinnis, C.S., Patterson, D.M., Winkler, J., Conrad, D.N., Hein, M.Y., Srivastava, V., Hu, J.L., Murrow, L.M., Weissman, J.S., Werb, Z., et al. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat. Methods *16*, 619–626. https://doi.org/10.1038/s41592-019-0433-8.
74. Chung, N.C., and Storey, J.D. (2015). Statistical significance of variables driving systematic variation in high-dimensional data. Bioinformatics *31*, 545–554. https://doi.org/10.1093/bioinformatics/btu674.
75. Freeman, T.C., Horsewell, S., Patir, A., Harling-Lee, J., Regan, T., Shih, B.B., Prendergast, J., Hume, D.A., and Angus, T. (2020). Graphia: A platform for the graph-based visualisation and analysis of complex data. Preprint at bioRxiv. https://doi.org/10.1101/2020.09.02.279349.
76. van Dongen, S., and Abreu-Goodger, C. (2012). Using MCL to Extract Clusters from Networks. In Bacterial Molecular Networks: Methods and Protocols, J. van Helden, A. Toussaint, and D. Thieffry, eds. (Springer New York), pp. 281–295. https://doi.org/10.1007/978-1-61779-361-5_15.
77. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. *9*, 559. https://doi.org/10.1186/1471-2105-9-559.
78. Kaufman, L., and Rousseeuw, P.J. (1986). CLUSTERING LARGE DATA SETS. In Pattern Recognition in Practice, E.S. Gelsema and L.N. Kanal, eds. (Elsevier), pp. 425–437. https://doi.org/10.1016/B978-0-444-87877-9.50039-X.
79. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.
80. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29. https://doi.org/10.1038/75556.
81. Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., et al. (2022). The reactome pathway knowledgebase 2022. Nucleic Acids Res. *50*, D687–D692. https://doi.org/10.1093/nar/gkab1028.
82. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30. https://doi.org/10.1093/nar/28.1.27.
83. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.
84. Meo, P.D., Ferrara, E., Fiumara, G., and Provetti, A. (2011). Generalized Louvain Method for Community Detection in Large Networks, pp. 88–93. https://doi.org/10.1109/ISDA.2011.6121636.
85. Schmid, K.T., Cruceanu, C., Böttcher, A., Lickert, H., Binder, E.B., Theis, F.J., and Heinig, M. (2020). Design and power analysis for multi-sample single cell genomics experiments. Preprint at bioRxiv. https://doi.org/10.1101/2020.04.01.019851.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
TotalSeq^{TM}-A0301 anti-mouse Hashtag 1 Antibody	Biolegend	Cat# 155801; RRID: AB_2750032
TotalSeq^{TM}-A0302 anti-mouse Hashtag 2 Antibody	Biolegend	Cat# 155803; RRID: AB_2750033
TotalSeq^{TM}-A0303 anti-mouse Hashtag 3 Antibody	Biolegend	Cat# 155805; RRID: AB_2750034
Monoclonal rat antibody targeting mouse CD16/32 (FC Block)	Biolegend	Cat# 101301 (also 101302); RRID: AB_312800
Anti-CD45-PE	Biolegend	Cat# 103106 (also 103105); RRID: AB_312971
Anti-CD11b-BV711	Biolegend	Cat# 101242 (also 101241); RRID: AB_2563310
Anti-Ly6G-APC	Biolegend	Cat# 127613 (also 127614); RRID: AB_1877163
anti-CD3-APC-Cy7	Biolegend	Cat# 100221 (also 100222); RRID: AB_2057374
anti-CD19-PE-Cy7	Biolegend	Cat# 115520 (also 115519); RRID: AB_313655
Goat-*anti*-tdTomato	AMS Biotechnology	Cat# AB8181 (also AB8181-200); RRID:AB_2722750
Rabbit-*anti*-mouse-P2Y12	AnaSpec	Cat# 55043A; RRID: AB_2298886)
Donkey-*anti*-Goat-AF555	ThermoFisher	Cat# A-21432 (also A21432); RRID: AB_2535853
Donkey-*anti*-Rabbit-AF555	ThermoFisher	Cat# A-31572 (also A31572); RRID: AB_162543)
Rabbit-*anti*-Ki-67 (ab15580)	Abcam	Cat# ab15580; RRID: AB_443209)
Biotinylated Goat-*anti*-Rabbit	Vector	Cat# BA-1000 (also BA-1000-1.5); RRID: AB_2313606)
Goat-*anti*-Rabbit-AlexaFluor-647	ThermoFisher	Cat# A-21244; RRID: AB_2535812)
Chemicals, peptides, and recombinant proteins		
Percoll	Merck	Cat: GE17-0891-02
Tyramide 488	Invitrogen	Cat: B40953
Tyramide 555	Invitrogen	Cat: B40955
Tyramide 647	Invitrogen	Cat: B40958
DAPI (4′,6-Diamidino-2-Phenylindole, Dihydrochloride)	Invitrogen	Cat: D1306
Critical commercial assays		
Chromium Single Cell 3' with Feature Barcoding technology for Cell Surface Protein Reagent Kit v3	10X Genomics	Cat: PN-1000128
NovaSeq 6000 SP Reagent Kit	Illumina	Cat: 20028402
RNAScope^{TM} Multiplex Fluorescent Assay v2	ACDBio	Cat: 323100
Tyramide signal amplification kit	Invitrogen	Cat: B40933
Deposited data		
Raw and analyzed data	This paper	GEO: GSE240368

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
C57BL/6J	Charles River	RRID:IMSR_JAX:000664
C57BL/6J-Ms4a3em2(cre)Fgnx/J	F. Ginhoux	RRID:IMSR_JAX:036382
B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J (Ai14)	The Jackson Laboratory	RRID:IMSR_JAX:007914
Oligonucleotides		
RNAScope Probe-Mm-Fcrls-C2	ACDBio	Cat: 441231-C2
RNAScope Probe-Mm-Ccl3	ACDBio	Cat: 319471
RNAScope Probe-Mm-Spp1-C3	ACDBio	Cat: 435191-C3
RNAScope Probe-Mm-Csf1r	ACDBio	Cat: 428191
RNAScope Probe-Mm-Gpnmb-C3	ACDBio	Cat: 489511-C3
RNAScope Probe-Mm-Ccr2-O1	ACDBio	Cat: 501681
RNAScope Probe-Mm-Mki67-C3	ACDBio	Cat: 416771-C3
RNAScope Probe-Mm-Cst7-C2	ACDBio	Cat: 498711-C2
Software and algorithms		
FCS Express 7	*De Novo* Software	FCS Express (RRID:SCR_016431)
CellRanger v3.0.2	10× Genomics	Cell Ranger (RRID:SCR_017344)
Seurat v3	Satija Lab	Seurat (RRID:SCR_016341)
CITE-seq-Count	https://hoohm.github.io/CITE-seq-Count/	CITE-seq-Count (RRID:SCR_019239)
MULTI-Seq	https://github.com/chris-mcginnis-ucsf/MULTI-seq	Multi-Seq
Graphia	Graphia Technologies	https://graphia.app/
ROGUE	https://github.com/PaulingLiu/ROGUE	ROGUE (Ratio of Global Unshifted Entropy)
WGCNA	https://cran.r-project.org/web/packages/WGCNA/index.html	Weighted Gene Co-expression Network Analysis (RRID:SCR_003302)
clusterProfiler	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html	clusterProfiler (RRID:SCR_016884)
velocyto	http://velocyto.org/	Velocyto (RRID:SCR_018167)
Slingshot	https://github.com/kstreet13/slingshot	Slingshot (RRID:SCR_017012)
GraphPad Prism	GraphPad Software	GraphPad Prism (RRID:SCR_002798)
Lme4	https://cran.r-project.org/web/packages/lme4/index.html	R package: lme4 (RRID:SCR_015654)
R	http://www.r-project.org/	R Project for Statistical Computing (RRID:SCR_001905)
Python	http://www.python.org/	Python Programming Language (RRID:SCR_008394)
Other		
Stroke Brain MySEQ interactive shiny app	This paper	https://mccoll-lab-uoe.shinyapps.io/shinyapp/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Barry McColl (barry.mccoll@ed.ac.uk).

Materials availability

No new or unique materials were generated as part of this study.

Data and code availability

- Single cell RNA sequencing (scRNAseq) data have been deposited at GEO (GSE240368) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Key processed scRNAseq data are provided in Supplemental tables. Interactive visualisation and analysis of the dataset can be carried out at our openly accessible interactive web browser application, Stroke-Brain-MySeq (https://mccoll-lab-uoe.shinyapps.io/shinyapp/). All other data supporting the findings of this study are available from the lead contact upon reasonable request.
- This paper does not report original code. Sources of standard code used are cited in the STAR Methods.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Experimental stroke model

Procedures were conducted on male 10–12 week old C57BL/6J mice (Charles River Laboratories) or *Ms4a3*$^{\mathrm{CreERT2/+}}$*Rosa*$^{\mathrm{TdT/+}}$ mice, generated by breeding *Ms4a3*$^{\mathrm{CreERT2}}$[68] and *Rosa*$^{\mathrm{LSL-TdT}}$ (Ai14) mice (JAX:007914). Mice were maintained under specific pathogen-free conditions and a standard 12 h light/dark cycle with unrestricted access to food and water. Mice were housed in individually ventilated cages in groups of up to five mice and were acclimatised for a minimum of one week prior to procedures. All procedures involving live animals were carried out under the authority of a UK Home Office Project Licence in accordance with the 'Animals (Scientific Procedures) Act 1986' and Directive 2010/63/EU and were approved by The University of Edinburgh Bioresearch & Veterinary Services Animal Welfare and Ethics Review Body (AWERB). Experimental stroke was induced by permanent distal middle cerebral artery occlusion (dMCAO) under isoflurane anesthesia (mixed with 0.2 L/min O_2 and 0.5 L/min N_2O). Core body temperature was maintained at 37 ± 0.5°C throughout the procedure with a homeothermic system (Harvard Apparatus, UK). A vertical incision between the left eye and ear was made and the main trunk and bifurcations of the middle cerebral artery exposed via a subtemporal craniectomy. The site was cooled continuously with saline application. The middle cerebral artery was electrocoagulated at the junction of its main trunk and bifurcation and cessation of blood flow confirmed visually prior to cutting through the coagulated area. The temporal muscle was repositioned, the incision sutured and topical local anesthetic (lidocaine/prilocaine, LMX4) applied. Anesthesia was discontinued and mice were administered 0.5 mL sterile saline subcutaneously. Mice were returned to home cages on a heating blanket and then transferred to pre-surgery holding areas.

METHOD DETAILS

Cell sorting and flow cytometry

Mice were perfused transcardially with 0.9% saline under terminal isoflurane anesthesia to remove circulating blood and brain samples containing the infarct/peri-infarct region or the equivalent region in the contralateral hemisphere were dissected. Samples were placed in 1X Hank's buffered saline solution (HBSS) (without Ca^{2+} and Mg^{2+}) (Gibco, Life Technologies) on ice and minced to fine pieces using a scalpel blade and then transferred to a 15 mL Dounce tissue homogeniser and manually dissociated. Samples were centrifuged at 400 g for 5 min at 4°C with no brake, supernatant was aspirated and pellets resuspended in 35% Percoll (GE healthcare) in HBSS (without Ca^{2+} & Mg^{2+}) and carefully overlaid with 1X HBSS. Samples were centrifuged at 800 g for 45 min at 4°C with no brake for density separation of cells from myelin. Supernatant and myelin layer were removed, and cells resuspended in 1X HBSS (without Ca^{2+} or Mg^{2+}). Cell suspensions were centrifuged at 400 g for 5 min at 4°C with no brake and resuspended in staining buffer (DPBS (without Ca^{2+} and Mg^{2+}) (Gibco, Life Technologies) containing 0.1% low endotoxin BSA (Sigma). Cells were incubated with anti-CD16/32 (Clone: 93, Biolegend) for 30 min at 4°C to block non-specific Fc receptor binding and then for 30 min at 4°C with the following fluorochrome-conjugated antibodies (anti-CD45-PE (clone 30-F11), anti-CD11b-BV711 (clone M1/70), anti-Ly6G-APC (clone 1A8), anti-CD3-APC-Cy7 (clone 17A2), anti-CD19-PE-Cy7 (clone 6D5), Biolegend) mixed with one of three distinct barcoded hashtag antibody-oligonucleotide (HTO) conjugates (TotalSeq-A, Biolegend) for each mouse to enable cell hashing. Cells were washed, centrifuged 400*g* for 5 min at 4°C with no brake, and resuspended in FACS buffer (0.04% BSA in DPBS). A myeloid-enriched cell population ($CD3^{-}$CD19-Ly6G-$CD45^{+}$) was sorted into FACS buffer using a FACSAria IIu (Becton Dickinson, Oxford, UK) and cell suspensions stored on ice. Unstained and single-stained controls were used to define the sorting and cytometry analysis gating positions. Approximately 30,000 cells per individual brain sample from each mouse (*n* = 3 mice) were sorted and cells ipsilateral or contralateral to MCAO pooled i.e., a pooled ipsilateral sample and a pooled contralateral sample each comprising HTO-barcoded cells from three bio-replicates (mice) was formed. Cytometry analysis was conducted on the FCS3 data files acquired during the cell sorting procedure using FCS Express 7 (*De Novo* Software).

Single cell library preparation and sequencing

Single cell 3′ gene expression (mRNA) and cell surface hashtag-oligonucleotide (HTO) libraries were prepared according to the Chromium Single Cell 3' with Feature Barcoding technology for Cell Surface Protein Reagent Kit v3 protocol (10X Genomics) with modifications for incorporation of TotalSeq-A HTO antibodies (clones M1/42; 30-F11, BioLegend). Briefly, generation of gel beads in emulsion (GEM) partitioned with single cells, reverse transcription to generate cDNA and HTO-barcoded DNA, and DNA amplification were performed according to manufacturer instructions with the addition of HTO primers during DNA amplification. Amplified cDNA and HTO-derived DNA were separated by size selection, fragmented and libraries constructed. Expected fragment size distributions for cDNA and HTO libraries were confirmed by electrophoresis (Agilent Bioanalyzer). cDNA and HTO libraries were mixed at a ratio of 4:1 for the pooled sample ipsilateral to MCAO and separately for the pooled sample contralateral to MCAO. Ipsilateral and contralateral samples (each containing combined cDNA and HTO libraries) were loaded on separate lanes of an SP flow cell and paired-end sequencing (Read 1: 28 cycles; i7 index read 1: 8 cycles; Read 2: 91 cycles) was conducted on a Novaseq 6000 (Illumina) to achieve a minimum read depth of 20,000 reads per cell for cDNA (gene expression) and 5000 reads per cell for HTO (mouse ID). cDNA and HTO sequencing reads were demultiplexed according to unique i7 indexes.

Single-cell RNA-Seq quality control and pre-processing

Count matrices were generated for each library using the 10X Genomics CellRanger v3.0.2[69] pipeline. First, raw reads were demultiplexed to produce fastq files. Second, reads were aligned to the mouse reference genome (GRCm38) and quantified. Resultantly, two unique molecule identifiers (UMI) based expression matrices were generated each for the ipsilateral and contralateral conditions which were pre-processed and analyzed using the Seurat v3 package in R.[70] As the samples were prepared and sequenced together, they represented a single batch and hence analyzed together as a merged dataset. Subsequently, quality control was conducted for cells and genes, where genes expressed in at least three cells were incorporated in downstream analyses. Cells were initially filtered based on the number of genes (at least 700) expressed and the mitochondrial percent (less than 8%). Subsequent to TPM like log normalisation using a scaling factor of 10^4, cells with normalised UMI between 2,500 and 4,000 were considered for downstream analyses. To identify multiplets, HTOs[71] were quantified for quality controlled cells of each condition using CITE-seq-Count[72] (https://github.com/Hoohm/CITE-seq-Count). The hashtag oligos (HTO) for each cell were normalized using a centered log-ratio transformation. Cells were then demultiplexed using the MULTI-Seq approach[73] and those assigned as multiplets or empty droplets were filtered out. To remove donor effects, canonical correlation analysis (CCA) was adopted using default parameters from Seurat.[70] Briefly, the approach identifies similar cells or ‘‘anchors’’ between datasets (in this case donors) based on their similarity in a joint reduced space, CCA components. The differences between these anchors represents the differences between donors and is accordingly used to calculate a weighted correction vector. These vectors are then applied on cells of the original gene expression matrix to remove donor effects.

Cell clustering

The integrated scRNA-Seq dataset was then scaled and regressed for percent mitochondrial content. Principal component analysis (PCA) was used to reduce the dimensions of the dataset to the 47 most significant ($p < 0.05$) PCs based on Jackstraw permutations.[74] Clustering on the whole dataset was done using a network-based approach using Graphia.[75] A k-nearest neighbor ($k = 10$) network of cells was constructed by connecting cells with a high Pearson similarity coefficient (based on the significant PCs) $r \geq 0.5$ to accommodate all cells. The resulting network was clustered using the Markov Clustering algorithm (MCL) with a low inflation MCLi = 1.35.[76] To estimate appropriateness of clustering, we calculated ROGUE[40] scores for each cluster in each of the datasets (all cells, subclustered microglia and subclustered non-microglial myeloid cells). ROGUE scores range from 0 to 1, with under-clustering represented by lower values < 0.8 and pure clusters by a value of 0.9–1. We sought a clustering granularity that was conservative and aligning with key biological ‘‘checkpoints’’ (e.g., discrimination of cell cycle phases) to avoid over-clustering.

Comparing cell proportions

The scRNA-Seq experiment was conducted by enriching for myeloid cell types, and the considerably altered composition of mononuclear phagocytes in response to MCAO made a direct comparison between MCAO and CTRL relative abundance inaccurate. Hence, we assumed that the number of resident homeostatic or homeostatic-like microglia from cluster 1 and 2 respectively would not change significantly between conditions, therefore we normalised cell frequencies of each cluster (considering condition and donor) with those of cluster 1 and 2. The normalized cell frequencies were then compared between conditions for cell clusters using a t test.

Generating pseudo-bulk samples from scRNA-Seq to construct gene co-expression networks

All the gene co-expression networks were constructed by first constructing pseudo-bulk samples from the scRNA-Seq data and then conducting weighted gene co-expression network analysis or WGCNA.[77] This preceding step of constructing pseudo-bulk samples aimed to improve the signal-to-noise ratio of single-cell data, capture the intra- and inter-cell type variation and generate a cell type balanced correlation space i.e., each cell type is equally represented as each has equal number of representative pseudo-bulk samples. The pseudo-bulk algorithm was used to generate five pseudo-bulk samples for each cluster. The strategy for generating these pseudo-bulk samples was conducted in the following two steps. 1) Each of the clusters was further clustered into five sub-clusters using Clara[78] (an extension of k-medoid clustering) based on the cell-to-cell similarity determined by the significant PCs. In the process, the medoid cell of each sub-cluster was also identified. 2) For each of these representative medoid cells, the expression values for genes were averaged across their ten nearest neighbors (based on significant PCs) within the cluster. As a result, five sub-clusters/pseudo-bulk samples were generated for each of the original clusters. The resultant pseudo-bulk vs. gene expression matrix was used to identify co-expressing genes using WGCNA.[77] A soft threshold of $\beta = 6$ was used to construct an adjacency matrix of pairwise Pearson correlations between genes. This matrix was then transformed into a topological overlap matrix to take into account the structure of the data. Hierarchical clustering was performed with parameters set for the module size (minModuleSize = 10), the granularity of clustering (deepSplit = 4) and the number of clusters (mergeCutHeight = 0.15).

Differential gene expression

Differentially expressed genes for each cluster were identified using the negative binomial generalized linear model available in Seurat.[79] Furthermore, only those genes were tested whose log fold change was greater than 0.25 (in either direction) and expressed in a minimum of 25% of cells in the groups examined. To provide functional annotation to subclustered microglia and non-microglia cell clusters, differentially-expressed genes for each cluster were examined for their enrichment in biological processes as defined by gene ontology (GO),[80] Reactome[81] and Kyoto encyclopedia of genes and genomes (KEGG)[82] databases using clusterProfiler in R.[83]

Functional enrichment analysis of WGCNA gene modules

Gene modules generated from WGCNA were examined for their enrichment in biological processes as defined by gene ontology (GO),[80] Reactome[81] and Kyoto encyclopedia of genes and genomes (KEGG)[82] databases using clusterProfiler in R.[83] Furthermore, to compare the enrichment of KEGG terms across gene clusters, the top 6 terms for each cluster were compared across all clusters shown in Figures. In all cases, the biological processes with significant ($Padj < 0.05$) enrichment were considered with P adjusted using the Benjamini-Hochberg correction for multiple testing.

Cell cycle analysis

Cell clusters were annotated for one of the three cell cycle stages, including G1, G2 and S phase using the cell cycle scoring algorithm available in Seurat v3[70] adapted from.[42] The algorithm generates a standardised score for each cell based on their average expression of G2 and S phase gene signatures. Based on these scores, cells are assigned to one of the three phases, where G1 (non-cycling) phase presents poor scores for both G2 and S phase gene signatures.

RNA velocity and pseudotime analysis

To estimate the direction and velocity of differentiating cells, the velocyto package in python was used with parameters proposed by the authors (https://github.com/velocyto-team/velocyto-notebooks/blob/master/python/Haber_et_al.ipynb). First, the spliced and unspliced reads were quantified from the BAM files generated from the cellranger pipeline. Cells previously selected based on the Seurat QC were used in the analysis. Genes were filtered based on their minimum expression, average expression and coefficient of variation. The data was subsequently normalized, and the 3,000 most variable genes were used to calculate PCAs. The top nine PCs were then used for kNN-imputation with k = 70 and a gamma distribution was fitted to each gene. The resultant expression matrix was used to estimate the velocity and future cellular states i.e., the direction.

Highly variable genes along the differentiation trajectory were determined using pseudotime analysis using Slingshot with default parameters.[56] First, the algorithm re-clustered the data using a Gaussian mixture model. Next, the start and endpoints of differentiation previously determined by RNA velocity were used to construct a minimum spanning tree on which simultaneous principal curves were fitted. Each cell was then aligned to the curves, thus aligning them along pseudotime. Finally, genes significantly changing ($p < 0.05$) along pseudotime were determined using a general additive model.

Cross-model microglial comparison

For disease model comparison, nine studies (including this study) were selected to represent different neurological disease conditions.[43–50] Microglial clusters from within these studies were identified and included for the analysis. As a note, DEG analysis for most studies compared one microglial cluster to all others, while DEGs from Mathys et al.[44] were identified by comparing early/late response microglial subtypes (Cluster 3, 6 and 7) with homeostatic microglia (Cluster 2). To capture genes representative of a cluster, DEGs were filtered by log fold change >0.25, and Padj <0.05. In the case of Mathys et al.,[44] genes with maximum likelihood estimation >0 (alternate metric to fold change) were considered. Furthermore, of these filtered DEGs the top 500 genes (based on log fold change) were considered for each microglial cluster across publications. The resultant gene lists were compared using Fisher's exact t test which provided an odds ratio (values greater than 1 are indicative of associated/correlated gene lists) and a P. Furthermore, only significant ($p < 0.05$ and odds ratio >1) pairwise odds ratios were examined.

For comparing subclustered microglia profiles in the present study to sham-derived microglia from an independent dataset, we selected the Zheng et al. dataset (GEO:GSE174574).[9] The dataset was downloaded and processed using the Seurat pipeline. Following log normalisation the most variable genes were identified and PCA was conducted. The top 20 PCs were used to cluster the data using the Louvain algorithm.[84] Microglia were subsetted using the following markers highlighted in the original study (*Hexb, Aldoc, Plp1, Ttr, Itm2a* and *Acta2*) and clusters annotated using markers identified by the authors (Figure S18B). To compare sham-enriched cluster 0 (c0) from Zheng et al. with microglia clusters from the present study, the two datasets were normalised and expression values for each gene scaled by converting expression values of genes to Z-scores within each dataset. Normalised expression values for top cluster marker genes were then compared across Zheng_c0 and microglia clusters from the present study, visualised by heatmap (Figure S18C).

Multiplex fluorescence *in situ* hybridisation

RNAScope Multiplex Fluorescent Assay v2 (ACDBio, 323100) was performed on 20μm cryostat-sectioned fixed-frozen brain sections using slight modifications to manufacturer's protocol. Sections were allowed to equilibrate to room temperature (RT) before washing with dH_2O and baking on to slides at 60°C. Slides were post-fixed (10% neutral-buffered formalin) for 15 min at 4°C and dehydrated through alcohols. Hydrogen peroxide was applied for 10 min at RT and antigen retrieval was performed for 15 min in a pre-heated (20 min) plastic Coplin jar in a 97.5°C waterbath. Sections were left at RT overnight before placing into humidity chambers and protease III was applied for either 20 min (for combination with protein immunofluorescence) or 30 min at 37°C. The following probes were incubated for 2 h at 40°C *Fcrls* (441231-C2)*, Ccl3* (319471)*, Spp1* (435191-C3)*, Csf1r* (428191)*, Gpnmb* (489511-C3)*, Ccr2* (501681)*, Mki67* (416771-C3), *Cst7* (498711-C2). Probes were subsequently amplified and labeled with Tyramide 488 (Invitrogen, B40953), 555 (Invitrogen, B40955) or 647 (Invitrogen, B40958). When co-labelling with immunofluorescence, sections were immediately placed into an antibody dilution buffer containing either Goat-*anti*-tdTomato (10μg/mL, AB8181) or Rabbit-*anti*-P2Y12 (0.66μg/mL, AS55043A) and left to incubate at 4°C overnight. Donkey-*anti*-Goat-AF555 or Donkey-*anti*-Rabbit-AF555 were subsequently incubated for fluorescent labeling. Nuclei were labeled with 1 μg/mL DAPI. Low-power images were acquired using an Axioscan Slide Scanner and high-power confocal images using an Opera Phenix Plus high-content confocal imager. Maximum intensity projections were processed from confocal stacks and cells from 3 to 6 non-overlapping 300μm × 300μm regions of interest were manually quantified per brain region per section in each individual mouse brain (n = 4 mice per condition).

Immunofluorescence

For immunolabelling of P2Y12 and Ki-67, sections were equilibrated to RT, washed with dH_2O and baked for 30 min at 60°C. Antigen retrieval was performed using Tris-EDTA buffer (10mM Tris Base, 1 mM EDTA, pH 8.6) for 20 min in a non-preheated plastic Coplin jar in a water bath set to 97.5°C for 20 min. Rabbit-*anti*-Ki-67 (1.25μg/mL, ab15580) was incubated overnight at 4°C. Endogenous peroxidas activity quenched with 0.3% H_2O_2 RT 30 min. Biotinylated Goat-*anti*-Rabbit (ZH0615) incubated for 90 min, then strep-HRP (Tyramide amplification kit, B40933) incubated for 45 min before amplification with Tyramide 555. To allow multiplexing of antibodies raised in the same species, antigen retrieval was performed again in the same manner to elute Ki-67 antibodies (leaving Tyramide signal untouched). Rabbit-*anti*-P2Y12 (AS55043A) was incubated overnight at 4°C and fluorescently labeled using Goat-*anti*-Rabbit-AlexaFluor-647 (A21244), incubated for 90 min RT. Primary antibodies were incubated in buffer containing 1% BSA, 0.3% Triton X-100, 1X PBS, other antibodies were incubated in 0.1% BSA, 0.5% Tween, Tris-buffered Saline. 3 × 2 min washes were performed between each step following primary antibody incubation using 0.5% Tween, Tris-buffered Saline. Widefield images of P2Y12 and Ki67 were acquired using an AxioImager.D2.

To compare microglial morphology across brain regions connected and unconnected to the primary infarct, we searched the Allen brain connectivity atlas (https://connectivity.brain-map.org/) for experiments that injected tracers into brain areas matching those affected by the primary infarct based on cresyl violet staining of 20μm sections from our MCAO brains. We chose experiment 100141780 that injected a tracer into the primary motor area and extracted the projection segmentation image from section 44. This section demonstrates high connectivity between hemispheres through transcallosal fibers and the lack of connectivity of the primary motor area with medial septal complex (MSC), both supported by projection volumes calculated by the Allen brain atlas (contralateral primary motor cortex 1.48, medial septal nucleus 0.00). Images of P2Y12-labelled brain sections were acquired from the peri-infarct, contralateral cortex and MSC across three brain sections per brain from sham-operated and MCAO mice (n = 5). Three non-overlapping 150 × 150 μm boxes were drawn per image and microglial morphology quantified on a per-cell basis based on the following scoring scale: 1, Ramified; 2, De-/Re-ramifying; 3, Hypertrophic; 4, Rod-like; 5, Ameboid. Scoring was averaged for analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Our primary aim was to conduct a hypothesis-free scRNASeq profiling study. Formal sample size estimates were not conducted however cell isolation was guided by estimating the number of cells required to detect a rare cell type within each individual sample using the following parameters (frequency = 0.01, minimum number of cells = 10, detection power = 0.95) (https://satijalab.org/howmanycells/).[85] Cell hashing (see above) from three independent mouse brains provided bio-replication within the dataset. Statistical approaches for analysis of scRNA data are described above. Cytometry data were analyzed by Welch's t test using Graphpad Prism (v9). For quantification of multiplex smFISH, a linear mixed-effect model was utilised to compare the distribution of cell phenotypes within brain regions across sham and stroke brains (n = 4 mice per condition). Within-subject design with random intercepts was used and factors (phenotype, region) and interactions were treated as fixed-effects and their inclusion evaluated through log likelihood ratio. Holm-Sidak post-hoc tests were utilised for *a-priori* planned pairwise contrasts. Normality and homoscedasticity were evaluated graphically and box-cox transformations applied when necessary. One sided t-tests were utilised to compare stroke phenotypes to 0 counts seen in sham controls. A paired t test was used to compare phenotype abundance within stroke brains and all p values were adjusted using the Bonferroni method. All smFISH analyses were performed using R (Version 4.1.2).