Supplementary Guide

Pyramidal neuron subtype diversity governs microglia states in the neocortex

*Jeffrey A. Stogsdill[1,2], *Kwanho Kim[1,2,3], Loïc Binan[3,4], Samouil L. Farhi[3,4], Joshua Z. Levin[2,3], Paola Arlotta[1,2]

*Authors contributed equally to the manuscript

[1]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

[2]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA

[3]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA

[4]Optical Profiling Platform, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Table of Contents

Supplementary Table Information

Table 1: Table of genes obtained from Marsh et al.[14] and Jordao et al.[13] to manually assign clusters by Mg or BAM identity, or to identify *ex vivo* activated Mg.

Table 2: Table of differentially expressed (DE) Mg genes by cluster for the P14 *Cx3cr1*^{GFP} scRNA-seq dataset. Table lists the genes with an average $\log_2$ fold-change > 0.25 and $\log_2$ fold-change as determined by MAST. Clusters are split by page.

Table 3: Table of differentially expressed (DE) Mg genes by cluster for the P60 *Cx3cr1*^{GFP} scRNA-seq dataset. Table lists the genes with an average $\log_2$ fold-change > 0.25 and ranked by $\log_2$ fold-change as determined by MAST. Clusters are split by page.

Table 4: Tables of DE genes between Homeostatic1 and Homeostatic2 Mg from the P14 and P60 *Cx3cr1*^{GFP} scRNA-seq datasets. Tables list the genes with an average $\log_2$ fold-change > 0.15 and ranked by $\log_2$ fold-change as determined by MAST.

Table 5: Table of signature genes for the identification of Mg states used through the manuscript. Gene signatures were used to generate heatmaps and for state scoring.

Table 6: Table of the genes for each latent variable (LV) retrieved by PLIER, used to estimate the variance of Homeostatic1 and Homeostatic2 Mg.

Table 7: Table of genes comprising the target probe library for MERFISH. The genes pertaining to each cell type or Mg state are listed. A selection of genes uncovered from the L-R analysis of Fig. 5 was included as part of the target probe library.

Table 8: Table of DE Mg genes by cluster for the P14 *Fezf2* Control and KO scRNA-seq dataset. Table lists the genes with an average $\log_2$ fold-change > 0.15 and ranked by $\log_2$ fold-change as determined by MAST. Clusters are split by page.

Supplementary Notes

Quality control analysis of scRNA-seq data from *Cx3cr1*^{GFP} micro-dissected layers

We purified Mg for scRNA-seq from micro-dissected S1 cortices (L1-4, L5, and L6) from *Cx3cr1*^{GFP} transgenic mice, a line which labels macrophages, including Mg[12] (Fig. 2a-b). We isolated GFP^{+} cells from juvenile (P14) and adult (P60) mice by homogenization and fluorescence activated cell sorting (FACS)[5] under conditions designed to minimize *ex vivo* activation[14]. Initial profiling of the sequenced data identified 13,191 cells from the P14 timepont and 13,589 cells from the P60 timepoint (Extended Data Fig. 2a). Quality control analyses (Extended Data Fig. 2b-e) indicated that 98.1% of cells were of Mg identity[14], with the rest identified as BAMs[13] (Extended Data Fig. 2f-g, Supplementary Table 1). Given that *ex vivo* activation of Mg can be a confounding factor in transcriptional analyses, we scored each cell for a signature of known markers of *ex vivo* activation[14] (Supplementary Table 1). Cells scoring positive for this signature were not found in the P14 dataset and were rare (< 2%) in the P60 dataset (Extended Data Fig. 2h), possibly from the more extensive dissociation necessary at this later age. Following the removal of BAMs, *ex vivo* activated cells, and low-quality cells (see Methods), the datasets comprised 23,156 high-quality Mg (Extended Data Fig. 2i-n).

Exploration of non-homeostatic Mg clusters identified in the scRNA-seq *Cx3cr1*^{GFP} micro-dissected layers

In the P14 and P60 *Cx3cr1*^{GFP} scRNA-seq datasets, we found clusters expressing markers that map to homeostatic or non-homeostatic Mg states as previously identified[5,15]. We identified four distinct clusters of non-homeostatic Mg, which we termed "*Apoe*^{High}", "*Ccr1*^{High}", "Innate Immune", and "Inflammatory" (Fig. 2c), which combined accounted for 27.2% and 15% of P14 and P60 Mg, respectively. *Apoe*^{High} Mg displayed strongly the upregulated expression of two genes, *Apoe* and *Lyz2* (Extended Data Fig. 2q, Supplementary Tables 2-3). *Apoe* transcript expression in Mg is dramatically elevated in Alzheimer's models[46]; however, this cluster likely represents a previously reported immature state of Mg expressing *Apoe* in healthy controls[47]. The *Ccr1*^{High} Mg upregulated genes (e.g., *Ccr1* and *Tmem176a*) associated with a previously described population implicated in the recruitment of conventional dendritic cells[48] (Extended Data Fig. 2r, Supplementary Tables 2-3). The Innate Immune Mg displayed elevated expression of interferon-induced genes[49], including *Ifi27l2a*, *Ifitm3*, *Oasl2,* and *Rtp4* (Extended Data Fig. 2s,

Supplementary Table 2-3), that map to innate immune gene ontology terms (Extended Data Fig. 2t, see Methods) as previously reported[5]. Lastly, the Inflammatory Mg upregulated the expression of cytokines and other inflammatory genes, including *Ccl3, Ccl4*, *Cd63,* and *Abcg1* (Extended Data Fig. 2u, Supplementary Table 2-3)[5], and whose differentially expressed genes map to inflammatory response and cytokine gene ontology terms (Extended Data Fig. 2v).

Quality control and annotation of cell types and states identified by MERFISH in S1 cortex

We performed spatial transcriptomics using multiplex error-robust fluorescent *in situ* hybridization (MERFISH[19]) to visualize the spatial distribution of 7 Mg states, 8 PN subtypes and all major classes of cell types in the P14 S1 cortex. We constructed a probe library (Supplementary Table 7) targeting 75 cell type- and Mg state-enriched genes, taken from published work[20] and our profiling data (Extended Data Fig. 7): 15 genes specific to major cell types, including PNs, interneurons (INs), Mg, astrocytes, oligodendrocyte-lineage cells (Oligo), BAMs, and pericytes; 28 genes that mark distinct subtypes of PNs; and 32 Mg state signature genes from our scRNA-seq data (Extended Data Fig. 5a-b). The 75 genes were imaged by MERFISH in three biological replicates of P14 *Cx3cr1*^{GFP} brains (Extended Data Fig. 5c). Automated cell segmentation (Extended Data Fig. 5d-e; see Methods) yielded a total of 20,253 cells, which showed consistent per-cell transcript and gene expression levels across samples (Extended Data Fig. 5f-h).

Unsupervised clustering of all cells produced 6 clusters representing major cell types (Fig. 3a left). The clusters were annotated via the expression of cell type-specific genes such as *Slc17a7* (PNs), *Erbb4* (INs), *Tmem119* (Mg), *Sox10* (Oligos), *Aldh1l1* (Astrocytes), and *Slc6a20a* (Pericytes; Extended Data Fig. 6a-b). BAMs were not identified within the dataset. All major cell types of the cortical parenchyma could be detected (Fig. 3b left and Extended Data Fig. 6c) and showed abundances similar to those previously reported[50].

We then sub-clustered the PNs to characterize their diversity and define the layers of the neocortex in our data. We sub-setted the MERFISH dataset to retrieve all *Slc17a7*$^{+}$/*Neurod2*$^{+}$ PNs and performed unsupervised clustering (Fig. 3a center). Mapping PN subtype-specific genes from our probe list to the clusters uncovered 8 classes of PNs (Extended Data Fig. 6d-f), which appeared at expected frequencies[21]. We annotated the clusters based on their gene expression and laminar location (Fig. 3b center and Extended Data Fig. 6d-e): L2/3 CPN (*Stard8*$^{+}$, *Pou3f2*$^{+}$), L4 Stellate (*Rorb*$^{+}$/*Kcnh5*$^{+}$), L5 Subcerebral PN (SCPN; *Tcerg1l*$^{+}$/*Fezf2*$^{+}$, *Bcl11b*$^{+}$), L5 Near Projecting (NP; *Bcl11b*$^{+}$, *Trhr*$^{+}$, *Dkk2*$^{+}$), L5 Corticostriatal PN (CStrPN; *Fezf2*$^{+}$, *Deptor*$^{+}$, *Foxo1*$^{+}$), L6 CThPN

(*Tle4*$^{+}$, *Foxp2*$^{+}$), and L6b Subplate (*Ctgf*$^{+}$, *Nxph4*$^{+}$). In addition, we retrieved a *Sulf1*$^{+}$, *Bcl11b*^{low} L6-enriched PN subtype that was *Tle4* and *Foxp2* negative, which we identified as L6 CPNs (Extended Data Fig. 5b and 6d-e). Cortical layer demarcations were calculated by normalizing the distance of each PN from the pia. Layer boundaries were defined based upon the known abundance of PN classes by layer, as follows: L1 = no PNs, L2/3 = upper-layer CPNs, L4 = L4 Stellate, L5 = SCPN + NP + CStrPN, L6 = CThPN + L6b. PN subtypes demonstrated a tight spatial sequestration corresponding to their expected layer of residence (Fig. 3d and Extended Data Fig. 6g).

With the PN subtype maps and cortical layers defined, we then characterized the heterogeneity and spatial localization of Mg states. The MERFISH probe library targeted seven Mg states identified by scRNA-seq, including: Homeostatic1 (7 genes), Homeostatic2 (10 genes), Innate Immune (2 genes), Inflammatory (5 genes), *Apoe*^{High} (2 genes), *Ccr1*^{High} (3 genes), and Proliferative (3 genes). We extracted all *Tmem119*$^{+}$/*Fcrls*$^{+}$ Mg and annotated them using the state-enriched signature genes (see Methods). We first identified each of the non-homeostatic Mg states based on the high standardized expression of key genes enriched in each state (Extended Data Fig. 6h). The remaining, unassigned Mg were categorized as either Homeostatic1 or Homeostatic2 based on their highest signature score for these two groups (Extended Data Fig. 6i). Plotting the Mg in *t*-SNE space, coded by their state signatures, did not result in separate clusters (Fig. 3a right), likely because our MERFISH library only interrogates 32 Mg genes, compared to the thousands of genes detected by scRNA-seq (Extended Data Fig. 6j). The relative frequency of each Mg state was reproducible among samples (Extended Data Fig. 6k), in line with our scRNA-seq results (Fig. 2e).

Quality control analysis of scRNA-seq data from *Fezf2 Control and KO* micro-dissected cortical layers

We performed scRNA-seq on Mg isolated from layer micro-dissected cortices from *Fezf2* Control and *Fezf2* KO mice, to ask whether Mg were affected non-cell autonomously by a change in L5-6 PN identity. Visual landmarks of cortical layers are difficult to distinguish in the *Fezf2* KO. Therefore, to define L5, we crossed the *Fezf2* KO line with the tamoxifen-inducible *Tcerg1l*-CreER^{T2}/LSL-tdTomato reporter line[25], which marks L5b PNs following administration of tamoxifen at P2 (Fig. 4a-b, Extended Data Fig. 8a). Unlike controls, P14 *Fezf2* KO mice showed few tdTomato^{+} cells in L5b, indicating a severe disruption in L5 PN identity (Fig. 4a-b). The

persisting tdTomato^{+} cells permitted a guided micro-dissection of L1-4, L5, and L6. Single-cell suspensions were immunolabelled with a triple-antibody combination to identify Mg (anti-Cd11b, -Cd45, and -*Cx3cr1*),[8] and triple-positive Mg were isolated by FACS and profiled by scRNA-seq (Fig. 4c, Extended Data Fig. 8b-f). This approach resulted in a high proportion of Mg, with almost no BAMs (Extended Data Fig. 8g). After removing low quality cells, BAMs, and *ex vivo*-activated Mg (Extended Data Fig. 8h-k), the dataset comprised 16,108 cells, which expressed markers of Mg identity (Extended Data Fig. 8l-m).

Supplemental Information References

46. Chen, Y. *et al.* Interplay Between Microglia and Alzheimer's Disease-Focus on the Most Relevant Risks: APOE Genotype, Sex and Age. *Front Aging Neurosci* **13**, 631827, doi:10.3389/fnagi.2021.631827 (2021).

47. Masuda, T. *et al.* Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature* **566**, 388-392, doi:10.1038/s41586-019-0924-x (2019).

48. Gallizioli, M. *et al.* Dendritic Cells and Microglia Have Non-redundant Functions in the Inflamed Brain with Protective Effects of Type 1 cDCs. *Cell Rep* **33**, 108291, doi:10.1016/j.celrep.2020.108291 (2020).

49. Grabert, K. *et al.* Microglial brain region-dependent diversity and selective regional sensitivities to aging. *Nat Neurosci* **19**, 504-516, doi:10.1038/nn.4222 (2016).

50. Keller, D., Ero, C. & Markram, H. Cell Densities in the Mouse Brain: A Systematic Review. *Front Neuroanat* **12**, 83, doi:10.3389/fnana.2018.00083 (2018).