Title

Human microglia maturation is underpinned by specific gene regulatory networks.

Journal

Immunity, 56(9)

Authors

Han, Claudia
Li, Rick
Hansen, Emily
et al.

Publication Date

2023-09-12

DOI

10.1016/j.immuni.2023.07.016

Published in final edited form as:

Immunity. 2023 September 12; 56(9): 2152–2171.e13. doi:10.1016/j.immuni.2023.07.016.

Human microglia maturation is underpinned by specific gene regulatory networks

Claudia Z. Han[1,14], **Rick Z. Li**[1,14], **Emily Hansen**[2,3], **Samantha Trescott**[2,3], **Bethany R. Fixsen**[1], **Celina T. Nguyen**[1,2,3], **Cristina M. Mora**[2,3], **Nathanael J. Spann**[1], **Hunter R. Bennett**[1], **Olivier Poirion**[1,4], **Justin Buchanan**[1,4], **Anna S. Warden**[1,2,3], **Bing Xia**[2,3], **Johannes C.M. Schlachetzki**[1], **Martina P. Pasillas**[1], **Sebastian Preissl**[1,4,§], **Allen Wang**[1,4], **Carolyn O'Connor**[5], **Shreya Shriram**[2,3], **Roy Kim**[2,3], **Danielle Schafer**[2,3], **Gabriela Ramirez**[2,3], **Jean Challacombe**[1], **Samuel A. Anavim**[2,3], **Avalon Johnson**[2,3], **Mihir Gupta**[7], **Ian A. Glass**[8], **Birth Defects Research Laboratory**[8], **Michael L. Levy**[6], **Sharona Ben Haim**[7], **David D. Gonda**[6], **Louise Laurent**[9], **Jennifer F. Hughes**[10], **David C. Page**[10,11,12], **Mathew Blurton-Jones**[13], **Christopher K. Glass**[1,*], **Nicole G. Coufal**[2,3,5,*]

[1]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA

[2]Department of Pediatrics, University of California, San Diego, La Jolla, CA 92093, USA

[3]Sanford Consortium for Regenerative Medicine, La Jolla, CA 92037, USA

[4]Center for Epigenomics, University of California, San Diego, La Jolla, CA 92093, USA

[5]Salk Institute for Biological Studies, La Jolla, CA 92037, USA

[6]Department of Neurosurgery, University of California, San Diego-Rady Children's Hospital, San Diego, CA 92123, USA

[7]Department of Neurosurgery, University of California, San Diego, La Jolla, CA 92037, USA

[8]Department of Pediatrics, University of Washington and Seattle Children's Research Institute

[9]Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Diego, La Jolla, CA 92093, USA

---

[*]Correspondence: ckg@ucsd.edu or ncoufal@health.ucsd.edu.

[§]present address: Institute of Experimental and Clinical Pharmacology and Toxicology, Faculty of Medicine, University of Freiburg, Freiburg, Germany.

Lead contact: Nicole Coufal (ncoufal@health.ucsd.edu)

Author Contributions

Conceptualization: C.Z.H., R.Z.L, C.K.G., N.G.C. Patient identification, consent and tissue acquisition: N.G.C., L.L., M.G., S.B., D.D.G., M.L.L. Xenotransplantation: C.Z.H., C.N., A.W., B.X., M.B.J. Cell and nuclei isolation: C.Z.H., B.R.F, a N.G.C. Sorting: C.C, M.P.P., C.ZH. Sequencing libraries preparation: C.Z.H., B.R.F., J.B. Analysis: C.Z.H., R.Z.L., H.B., O.P., J.B., J.C., S.B., A.W. in vitro experiments and immunohistochemistry: E.H., A.S.W., S.S., R.K., C.N., D.S., G.R., S.A., A.J. Grant support acquisition: J.F.H., D.P., C.K.G., N.G.C. Supervision: C.K.G. and N.G.C. All authors contributed to editing and review of the manuscript.

Declaration of interests

M.B.J. is a co-inventor of patent application WO/2018/160496, related to the differentiation of pluripotent stem cells into microglia and co-founder of NovoGlia Inc. All other authors declare no competing interests.

[10]Whitehead Institute, Cambridge MA 02142, USA

[11]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

[12]Howard Hughes Medical Institute, Whitehead Institute, Cambridge, MA 02142

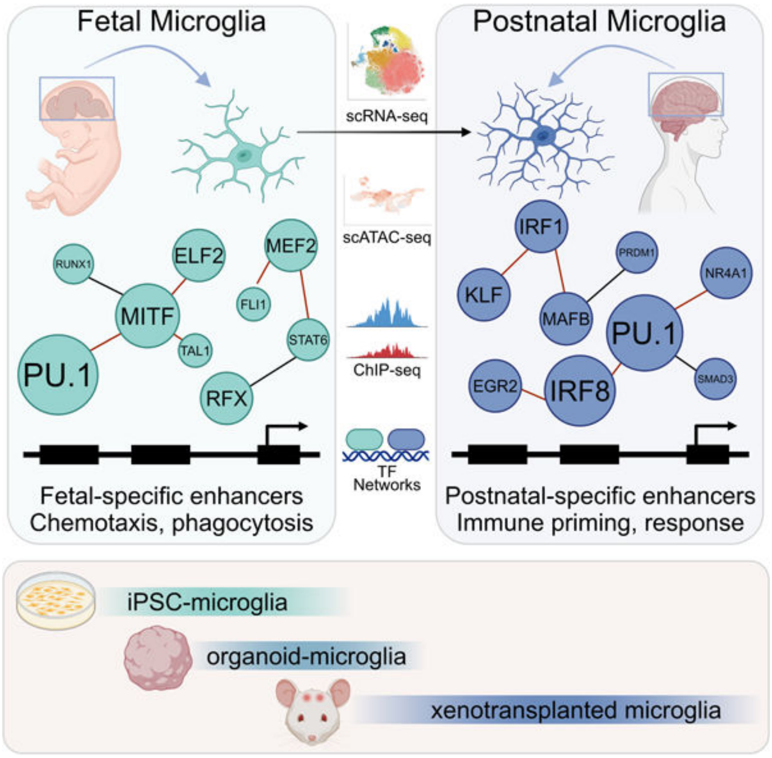
[13]Department of Neurobiology and Behavior, University of California, Irvine, Irvine CA 92696, USA

[14]These authors contributed equally to this work

Summary

Microglia phenotypes are highly regulated by the brain environment, but the transcriptional networks that specify maturation of human microglia are poorly understood. Here, we characterized stage-specific transcriptomes and epigenetic landscapes of fetal and postnatal human microglia and acquired corresponding data in induced pluripotent stem cell (iPSC) derived microglia, in cerebral organoids, and following engraftment into humanized mice. Parallel development of computational approaches that considered transcription factor co-occurrence and enhancer activity allowed prediction of shared and state-specific gene regulatory networks associated with fetal and postnatal microglia. Additionally, many features of the human fetal to postnatal transition were recapitulated in a time-dependent manner following the engraftment of iPSC cells into humanized mice. These data and accompanying computational approaches will facilitate further efforts to elucidate mechanisms by which human microglia acquire stage- and disease-specific phenotypes.

Graphical Abstract



In brief

Molecular mechanisms that drive human microglia diversity remain largely unknown. Han and Li et al., characterize transitions in transcriptomes and epigenomes of human fetal and postnatal microglia, develop computational tools that predict underlying gene regulatory networks, and demonstrate that core features of these networks can be captured by iPSC-derived microglia.

Keywords

Microglia; human brain development; transcription factors; transcriptomics; epigenomics

Introduction

In the human brain, microglia infiltration appears to precede waves of neurogenesis, astrogliogenesis and oligodendrogenesis[1], with microglia as the primary glia in the first and early second trimester[2,3] suggesting crosstalk between these cell types in establishing the brain environment. Indeed, depletion of microglia during embryonic mouse development results in abnormal positioning of subsets of interneurons[4] and altered numbers of neurons and astrocytes[5,6]. Brain development continues postnatally in both humans and mice with microglia continuing to modify neural circuitry. However, the molecular mechanisms driving microglia states during human development is unknown. Furthermore, although several hundred genes have been identified in monogenic neurodevelopmental disorders (NDDs)[7–9], the cell types involved and the cell-type specific functions perpetuating NDDs are unclear, especially with respect to the immune system[10,11], underscoring the importance of elucidating the contribution of microglial pathology to NDDs.

Cell-type specific gene expression is regulated by differential transcription factor (TF) binding and enhancer activation as a response to environmental cues[12,13]. While the transcriptional heterogeneity of human fetal cells, including microglia, have been characterized[14,15], how these genes are regulated, specifically the TF networks and the corresponding enhancers that they activate, has thus far remained elusive. Furthermore, the extent to which experimental systems based on induced pluripotent stem cells (iPSC) technology can be used to model stage-specific microglia phenotypes has not been systematically investigated.

Here, we present a thorough dissection of human microglia development by characterizing transcriptomic and epigenomic profiles of human fetal and postnatal microglia with corresponding analyses of iPSC-derived microglia in monoculture, in cerebral organoids and over time after engraftment into the mouse brain to discern the development stages best captured by iPSC-based models. In parallel, we develop two innovative computational methods to advance the interpretation of these data sets. **T**ranscription Factor **I**nteraction Inference from **M**otif co-**O**ccurrence **N**etworks (TIMON) considers the probability of TFs co-binding within active regulatory regions as the basis for the construction of state-specific regulatory networks. **T**ranscription **F**actor **Act**ivity analysis (TFAct) alternatively combines single cell RNA-seq (scRNA-seq) data and active *cis*-regulatory regions to calculate activity scores for individual TFs as drivers of specific patterns of gene expression. These studies provide substantial computational and data resources for understanding transcriptional mechanisms underlying the acquisition of stage-specific microglia phenotypes and how alterations in these mechanisms may contribute to NDDs.

Results

To investigate mechanisms regulating human microglia maturation, we performed RNA-seq on flow cytometry-sorted live microglia (Figure S1A,B) isolated from early to mid-gestation fetuses (9W to 17W gestational age, 17 subjects), referred to as fetal microglia, and from cortical tissue derived from epileptic resections of patients (18 subjects, Supplementary Table 1), referred to as postnatal microglia. Additionally, we performed RNA-seq on a portion of bulk fetal cortex (Figure S1C). Weighted gene correlation network analysis (WGCNA)[16] identified 16 clusters of highly co-expressed gene modules with respect to development and tissue type (Figure 1A). As expected, modules enriched for genes involved in axonogenesis and synapse organization were associated with fetal and postnatal cortical tissue. Modules associated with fetal, but not postnatal, microglia were enriched for genes with functional annotations for cell cycle and phagocytosis.

Comparison of fetal and postnatal microglia transcripts identified 2,291 differentially expressed genes (DEGs, FC >2, p-adj <0.05, Figure 1B) that distinguished the two developmental stages. Gene ontology (GO) analyses of genes with higher expression in fetal microglia relative to postnatal microglia yielded enrichment for cell cycle (Figure S1D) and ATP metabolic processes (Figure 1C). Postnatal microglia were enriched in genes associated with cytokine signaling and antigen processing and presentation (Figure 1A, C). We previously identified 881 transcripts[13] that were expressed >10-fold higher in human postnatal microglia than whole cortex at a false discovery rate of <0.05. Corresponding analysis of fetal microglia resulted in the identification of 992 genes that were expressed >10-fold higher in fetal microglia than in fetal cortex, of which 510 were shared with the postnatal gene signature ($p < 2.2e^{-16}$) (Figure S1E). Shared genes included genes that specify microglia development and function, while genes unique to the fetal microglial gene signature are involved in maintaining microglia survival, proliferation, adhesion and motility. Genes unique to the postnatal microglia gene signature were associated with immune response and myeloid activation (Figure S1E).

We also detected gene expression differences between first and second-trimester microglia, with second-trimester microglia upregulating genes associated with immune regulation, including *AHR, NR4A2,* and phagocytosis, such as *TIMD4* and pattern-recognition receptor *CLEC7A* (Figure 1D). In young postnatal mice, *CLEC7A* expression is associated with proliferative-region-associated microglia found in developing white matter[17]. We also found increased expression of disease-associated microglia-genes, including *LAG3*, a receptor associated with microglial reactivity in α-Synucleinopathy[18], and *CD48*, a receptor that is associated with proinflammatory responses in immune cells[19] (Figure 1D).

Mouse postnatal microglia have been reported to exhibit sex differences transcriptomically[20]. Additionally, numerous neuropathologies have a sex-bias, including autism spectrum disorder[21] and Alzheimer’s Disease[22]. We found 71 differentially expressed autosomal genes in the first trimester with expression of immune-related genes (e.g., *TNF, CCL3)* higher in female than male microglia (Figure 1E). Furthermore, analysis of the differential expression of genes associated with NDDs that have genetic variants, including monogenic and exome-sequencing mutations, in microglia and whole cortex, (Figure S2A,B, Supplemental Table 2) show that numerous genes associated with intellectual disability, autism, lysosomal storage diseases, and schizophrenia are preferentially expressed in microglia compared to whole cortex (Figure 1F, S2A,B). Taken together, our results suggest fetal microglia may play an influential role in a variety of NDDs.

Development-stage specific brain environment signals shape microglia phenotypes

To identify potential environmental ligands derived from cortical cells and downstream signaling networks involved in microglia gene expression, we utilized NicheNet, which computationally predicts ligand-receptor interactions by combining gene expression with existing knowledge of signaling pathways and gene regulatory networks[23,24]. *VCAM1,* an adhesion molecule involved in leukocyte migration, and Notch-ligand *DLL1* were two ligands predicted to preferentially influence the fetal microglia transcriptome (Figure 2A, B). Using scRNA-seq data of the developing human cortex[25], we found that radial glial cells highly express *VCAM1*, while astrocytes and intermediate progenitor cells of the medial ganglionic eminence express *DLL1* (Figure S3). Microglia interact with these cell types during embryonic development [26,27], but the consequences on fetal microglial phenotypes mediated by these ligand-receptor interactions remain to be ascertained. Manual curation of additional ligand-receptor pairs identified *CSF1* and *JAG1*, a Notch ligand, as preferentially expressed in fetal microglia, suggesting an autocrine fetal role (Figure 2C).

In the postnatal human brain, NicheNet analysis predicted TGFβ superfamily members as cortex-derived ligands influencing microglia (Figure 2A,B). Using the Allen Brain Map "Multiple Cortical Areas – SMART-seq (2019) dataset (https://portal.brain-map.org/atlases-and-data/rnaseq)[28,29], we found GABAergic and glutamatergic neurons expressed *TGFB3* and *BMP7* (Figure S3). Postnatal microglial ligands predicted by NicheNet included cytokines *IL6, CCR3 and TNF*, paralleling the immune-related pathways identified by GO analysis (Figure 2A,B). Alternatively, *APOE*, a gene involved in lipid metabolism and implicated in Alzheimer's Disease[30], was produced by both microglia and other brain cell types, including astrocytes and oligodendrocyte progenitor cells (Figure 2A,B, S3), and predicted to signal through *SORL1* in postnatal microglia. Using multiplex RNA *in situ* hybridization, we validated the expression of *APOE* and *CCL3* in both microglia and non-microglial cells in the postnatal brain and expression of *SORL1* and *CCR5* in microglia (Figure 2D, E). Taken together, our ligand-receptor analysis suggests that microglia are putatively educated by the brain environment and microglia itself.

Fetal and postnatal microglia employ differential enhancer activation

Environmental signals lead to activation of lineage determining (LDTFs) and signal-dependent transcription factors (SDTFs) which then bind to and activate distal regulatory regions, resulting in initiation of cell-specific gene expression programs[31]. To infer transcription factors (TFs) that may regulate microglia phenotypes, we defined regions of open and active chromatin that may be putative enhancers in human fetal and postnatal microglia by performing Assay for Transposase-accessible Chromatin (ATAC-seq) and chromatin immunoprecipitation (ChIP-seq) for Histone H3 Lysine 27-acetylation (H3K27ac), a mark of active promoters and enhancers[32]. We identified over 11,000 differentially active enhancers (FC >2, p-adj. <0.05, Figure 3A) between fetal and postnatal microglia. *De novo* motif analysis of enhancers preferentially active in human fetal microglia revealed enrichment for motifs of MiTF-TFE, MAF, and MEF2 TFs (Figure 3A, top right), with higher expression of *MITF* and *MAF* in fetal as opposed to postnatal microglia (Figure 3B), suggesting that these factors may activate fetal microglia-specific enhancers. Sites gaining H3K27ac in postnatal microglia were enriched for *de novo* motifs assigned to IRF, STAT2, and KLF TFs (Figure 3A, bottom right), with upregulation of *KLF4, IRF1*, and *STAT2* expression in postnatal versus fetal microglia, reflecting the immune signaling identified by NicheNet (Figure 2A).

We then performed single cell ATAC-seq (scATAC-seq) on fetal and postnatal microglia (Figure 3C) to investigate epigenomic heterogeneity during development. We identified six cell clusters with unique open chromatin landscapes. Cluster 0 contained contribution from all samples, while clusters 1, 2, and 3 were biased towards fetal samples and clusters 4–5 towards postnatal samples (Figure 3C, right). Motif analysis of scATAC-seq using chromVAR[33] =confirmed enrichment for MAF family members and the MiTF-TFE family in clusters 1 and 3, while motifs for IRF, AP-1 and EGR factors were preferentially enriched in clusters 4 and 5 (Figure 3D, E, F). We also found enrichment for the SMAD4 motif in clusters 0, 4 and 5, reflecting the NicheNet predicted TGFβ signaling in human postnatal microglia (Figure 3E, F). The motif for MEF factors was more strongly enriched in scATAC postnatal biased clusters 4 and 5, in contrast to its preferential enrichment in fetal ATAC peaks that were marked by H3K27ac. This observation suggests that fetal enhancers occupied by MEF2 transcription factors become less active following the fetal to postnatal transition. Taken together, our data suggest that differential enrichment for TF motifs in open and active chromatin results from alterations in the brain environment between fetal and postnatal states.

Transcription factor co-occurrence networks

Traditional “motif scan” methods utilize a sequential base-by-base motif scanning method, resulting in multiple TF candidates which unable to bind the DNA simultaneously due to steric hindrances. Furthermore, the binding motifs for each TF used in these methods have been hampered by motif redundancy. These challenges can cause motif scanning to suffer from high false-positive discoveries, making it unsuitable to uncover collaborative TF binding dynamics. Thus, we developed an advanced computational framework that identifies non-overlapping high confidence motifs from open chromatin regions overlaid with active enhancers to build a motif co-occurrence network predicted to specify enhancer selection. We refer to this algorithm as **T**ranscription Factor **I**nteraction Inference from **M**otif co-**O**ccurrence **N**etworks (TIMON).

Briefly, for each base location, we identified all putative TF motifs, but removed TFs that overlap with a higher scoring motif (Figure 4A, top). Therefore, a single iteration of the non-overlapping motif search will select non-overlapping motifs and produce gaps (Figure 4A, middle). This motif search is repeated in the gaps until no motif can fit (Figure 4A, bottom). Next, we built a matrix describing the frequency of TF co-occurrence in every sample. The significance of the co-occurring TF pair was tested against a background co-occurrence frequency, derived from multiple datasets of diverse cell types (Figure 4A), resulting in co-occurring TF networks. We tested TIMON on enhancer data associated with “alternative activation” of macrophages exposed to interleukin-4 (IL-4)[34]. TFs STAT6[35,36], PPARγ[37,38], EGR2[39], in addition to PU.1 and C/EBPβ collaborate at enhancers to drive the IL-4 response in macrophages[34]. We applied TIMON to the H3K27ac ChIP-seq profile of control and IL-4 treated macrophages, resulting in the accurate identification of STAT6, PPARγ, EGR2 and C/EBPβ in the TF co-occurrence network unique to IL-4 treated macrophages (Figure S4A,B). We then tested the predictive abilities of TIMON to capture biologically relevant TF connections. First, we found that detection of a co-occurring motif in the same peak significantly improved the chances of actual TF binding in the peak (Figure S4C). Second, using TF ChIP-seqs for PPARγ, STAT6 and C/EBPβ, we found that in IL-4 specific enhancers, TFs PPARγ and C/EBPβ or STAT6 and C/EBPβ were found to co-bind significantly more often than by chance (Figure S4D). Lastly, we determined the frequency of predicted TF co-binding resulting in actual co-binding (Figure S4E). TIMON predicted 27 peaks for PPARγ and C/EBPβ co-binding with experimental validation of 20 of those peaks; meanwhile for STAT6 and C/EBPβ, 38 out of 46 predicted peaks were validated (Figure S4E), as exemplified by *Tarm1* and *Slc7a2* (Figure S4F). Collectively, these data confirmed that TIMON captures the most functionally relevant interacting TF modules.

To discern the general co-occurrence network in microglia, we applied TIMON to the merged fetal and postnatal microglia datasets. The resulting network diagram revealed known microglia LDTFs, SDTFs, and their significant interactions (Figure 4B). As expected, PU.1 being the macrophage LDTF was the dominant TF. In fetal microglia, MITF, a member of the MiTF-TFE family that regulates lysosomal biogenesis[40], is next most prominent TF with the greatest number of unique partners (Figure 4C). Importantly, MITF was not overrepresented in the postnatal microglia-specific TF co-occurrence network (Figure 4D). Conversely, PRDM1 and STAT2 were highly interconnected factors in the postnatal network (Figure 4D), consistent with the enrichment of these motifs in enhancers exhibiting preferential activity in postnatal microglia (Figure 3A). We also identified SMAD2, a downstream mediator of TGFβ signaling[41], for the postnatal stage, further corroborating our NicheNet results.

We then extend our analysis into development stage-specific co-occurrence networks by identifying specific interactions between TF pairs (Figure 4E) or cliques (i.e., >2 interconnected nodes). For instance, in fetal microglia, one TF clique is MITF-PU.1-ELF2 (Figure 4F). Gene ontology analysis of genes associated with the enhancers containing MITF-PU.1-ELF2 motifs showed an enrichment of genes associated with chemotaxis, regulation of cytokine production and cell-substrate adhesion (Figure 4F). An example gene is *LGALS3*, which is highly expressed in fetal microglia and is involved in cell-extracellular matrix interactions and coordination of repairing of damaged lysosomes[42]. Enhancers associated with *LGALS3* with putative binding of MITF, PU.1 and ELF2 showed greater H3K27ac signal in fetal as opposed to postnatal microglia (Figure 4F). In postnatal microglia, we identified a clique with NR2C2-IRF2-MAFB-PRDM1 (Figure 4G). MHC class II gene *CD74* (Figure 4G, bottom)[43,44], is highly expressed in postnatal microglia with several intronic enhancers containing combinations of NR2C2-IRF2-MAFB-PRDM1 motifs and marked with high H3K27ac deposition. Genes linked to enhancers containing these TF motifs were associated with activation of the innate immune response, defense response, paralleling our bulk transcriptome and NicheNet analysis. Using ChIP-seq with locus-specific amplification, we validated the binding of *LGALS3*- and *CD74*-enhancer associated factors (Figure S4G, H), denoted with asterisks, in primary human fetal and postnatal microglia, respectively. Of note, TF co-occurrence predictions made by TIMON are based on best-matching motifs, but members of the same TF family (e.g., MITF-TFE and IRFs) often share highly similar motifs. Consistent with this, MITF and TFEB recognize the same motif and were found to co-occupy the *LGALS3* enhancer with PU.1 in fetal microglia (Figure S4G). In postnatal microglia, *IRF7* was more highly expressed than *IRF2* and therefore used in locus-specific amplification ChIP (Figure S4H). Overall, all tested factors showed statistically significant enrichment over background in at least one sample. In summary, TIMON goes beyond conventional motif analysis to identify combinations of TFs at the level of individual enhancers that are predicted to drive subsets of genes linked to specific biological processes.

Shared and unique microglia substates in human development

We then extended our TF analysis to elucidate the gene regulatory networks that may drive microglia state heterogeneity during development. We performed single-cell RNA-seq (scRNA-seq) on 5 fetal microglia and 3 postnatal microglia samples, yielding a total of 86,257 cells after quality control (Figure 5A). Clustering with Harmony correction[45] identified 10 clusters, with each cluster containing cellular contribution from every sample (Figure 5B). RNA velocity[46] analysis identified intra-cluster movement within cluster 0 with a transition of fetal to postnatal microglial gene expression pattern (Figure 5A, S5A), marked by decreased expression of endocytosis-phagocytosis genes and increased expression of immune priming genes (Figure S5A), similar to a previous report[15]. Expression of canonical microglia genes were relatively ubiquitous (Figure 5C, Figure S5B) with some genes, such as *P2RY12* (Figure 4C), and ligands such as *IGF1* and *CCL3* (Figure S5C) having a developmental stage bias.

Cluster 1, with a predominant fetal sample contribution, expressed genes associated with lysosomal functions (e.g., *LGALS3, CTSD*) (Figure 5C, S5D). Meanwhile, cluster 3, classified as immune modulatory, had high expression of *CD83, CH25H*, and *IL1B*, and was composed of mostly postnatal cells (Figure 5C, S5E). Both fetal and postnatal microglia had DNA replication (cluster 2) and cell cycle (cluster 4) substates (Figure 5C, S5F) with proliferation events were confirmed by immunofluorescence (Figure S5G). We also detected a small interferon responsive state (cluster 6) expressing *IFIT1, IFIT3,* and *MX1* (Figure 5C, S5H). Cluster 9 composed of cells co-expressing *CSF1R* and genes associated with neural progenitor cells (NPCs) (e.g., *DCX, SOX11*), most likely representing microglia that have phagocytosed NPCs (Figure S5I); immunohistochemistry identified rare events of likely engulfment (Figure S5J).

Lastly, we detected a monocytic population (cluster 8), derived mostly from postnatal samples, and a border-associated macrophage (BAM) cluster (cluster 5), qualified by mouse BAM markers, *CD206*[47–49] and *Lyve1*[50] (Figure 5C, S5K,L). Immunostaining of fetal brain revealed LYVE1+ positive cells only near tissue borders, marked by laminin staining (Figure S5M,N,O). Other developmental features of human microglia have also been represented in scRNA-seq analysis of developmental mouse microglia[51]. For instance, embryonic mouse microglia have a proliferative state and postnatal mouse microglia possess an immune modulatory signature. However, the human interferon-responsive microglia state is detected only in the context of white matter focal injury in mouse[51]. Collectively, our results indicate that at rest, human fetal and postnatal microglia in general have similar degrees of heterogeneity with a higher percentage of fetal microglia exhibiting a lysosomal signature and higher percentages of postnatal microglia exhibiting immune modulatory and interferon signatures.

Differential scRNA-seq cluster-based enhancer-TF activity

We then applied TIMON to our scRNA-seq datasets to decipher the TF co-occurrence network on a cluster- and development stage- basis. For the lysosomal cluster-1, which was comprised mostly of fetal microglia, the fetal-specific network identified MITF as a central partner of multiple TFs (Figure 5D). In the postnatal-specific network for this cluster, NFIC is a TF node (Figure S6A). While little is known about the role of NFIC in macrophages, NFIX, a co-family member, is involved in macrophage acquisition of an anti-inflammatory phenotype post-efferocytosis[52].

While TIMON predicts how multiple TFs can collaborate on enhancers to modulate gene expression, we sought to infer the activity of individual TFs in different microglia substates through the integration of epigenomic and scRNA-seq data. We adapted a common approach of modeling the activities of TFs as a function of the expression of their target genes[53,54], developing **T**ranscription **F**actor **Act**ivity analysis (TFAct). Briefly, TFAct identified active enhancer sequences and linked enhancers to target genes using the nearest expressed gene. Then TFAct identified nonoverlapping TF motifs within the enhancer regions and computed a TF activity score for each enhancer-associated TF using scRNA-seq data, building a TF-gene network for each development age (Figure S6B). To validate the predictive abilities of TFAct, we applied TFAct to a published single-cell dataset (Figure S6C) with accompanying epigenetic profiling of liver macrophages in control condition and nonalcoholic fatty liver disease (NASH)[55]. TFAct identified nuclear receptor family (e.g., RXR), AP-1 family factors (e.g., ATF2), and EGR (e.g., EGR2) as having statistically significant enhanced activity in liver macrophages during NASH as opposed to control conditions (Figure S6C, D). *ATF3* and *EGR2* expression are both upregulated during NASH [55], consistent with their driving the increased TF activity scores corresponding to the ATF and EGR motifs identified by TFAct (Figure S6D). LXR, a nuclear receptor, maintains Kupffer cell identity in homeostatic conditions[56], thus explaining the detected activity in the control liver macrophage cluster. Moreover, 80% of LXR ChIP-seq peaks gained in liver macrophages during NASH overlapped with ATF3 ChIP-seq peaks, suggesting that ATF3 and LXR co-bind to drive gene expression[55] as a consequence of the NASH diet. Motifs corresponding to the EGR family were enriched in LXR peaks gained in liver macrophages during NASH as well[55]. Thus, TFAct captures the “activity” of major TFs regulating cellular phenotypes on a single-cell basis.

Application of TFAct to our human microglia single-cell clusters (Figure 5F) revealed that TFE3, another MiTF-TFE family member, showed strongest activity for the lysosomal cluster-1 (Figure 5E). We also detected high TF activity for MEF2C, which when mapped to a tSNE projection, showed strong alliance with clusters with predominant contribution by fetal microglial cells (Figure 5F, left). Microglia with loss of MEF2C have exacerbated immune responses post-inflammatory challenge[57], while heterozygous loss of MEF2C in microglia leads to behavioral deficits associated with autism[8], suggesting that microglial-MEF2C plays a role in mediating microglial functions in shaping early brain development. Additionally, ATF4, a member of the AP-1 family, that has been linked to inflammation[58], showed strong activity for cluster 3, the immune modulatory cluster (Figure 5E,F). A complete list of TFs and their activity score on a cluster-based level is in Supplemental Table 3.

iPSC-microglia in organoids and mouse brain capture distinct *in vivo* phenotypes

It has become increasingly appreciated that microglia contribute to the pathogenesis of numerous neurodegenerative and neuropsychiatric diseases and techniques to study human microglia are limited. To elucidate the temporal congruence of expression of NDD-associated genes, we examined the overlap in NDD-associated gene expression between human fetal and postnatal microglia across orthologues, finding little overlap. For example, in intellectual disability, only a few genes, such as *NRAS, EZH2,* and *NFIA* showed a similar pattern in preferential expression between fetal mouse E10.5 and postnatal adult mouse microglia (Figure S7A-B).

The recent ability to differentiate induced pluripotent stem cells (iPSCs) to primitive hematopoietic progenitor cells (iHPCs) and subsequently to microglia-like cells (iMGs) and their integration into either cerebral organoids (oMGs) (Figure 6A,B) or xenotransplanted into the brains of neonatal humanized immunodeficient mice (xMGs)[59–61] (Figure 6A,B) allows for functional studies. These humanized immunodeficient mice are depleted of resident mouse microglia through the deletion of a FIRE intronic enhancer in the *Csf1r* locus[62] and overexpress the human *CSF1* ligand, allowing for extensive engraftment and survival of transplanted HPCs. We next investigated the similarity of the iPSC model systems to primary microglia to determine their applicability to experimental dissection of human microglia function. First, we examined environmental similarities and differences between oMGs and human microglia. Comparison of the ratio of gene expression between oMG and organoids revealed a high degree of correlation to that of fetal microglia to fetal cortex (r = 0.9) (Figure 6C), suggesting that the organoid environment produces signals similar to the fetal brain. Next, using Transcriptome Overlap Measure (TROM), a testing-based method for identifying transcriptomic similarities[63], we interrogated the similarity of human microglia, iMGs, oMGs, and transplanted microglia (xMG), with the latter having temporal resolution (Figure 6A). Analysis using TROM showed that iMGs and oMGs were able to capture distinct features of fetal microglia, but not of postnatal microglia (Figure 6D). xMGs at 7-days and, to a lesser degree, 12-days post-transplantation, were more similar to human fetal versus postnatal microglia. However, at 3 and 8-weeks post-transplantation, xMGs were distinctly more like human postnatal microglia on a transcriptome-level (Figure 6D). Similar to our mouse-human microglia comparisons, we also examined the gene expression patterns of NDD-associated genes comparing primary human microglia to the iPSC-derived systems. Genes belonging to multiple NDDs, including schizophrenia (e.g., *SLC9A3R1*[64]), intellectual disability (e.g., *PURA*[65]), and autism had expression profiles in the model systems that mimicked the age-dependent expression pattern in primary human microglia (Figure 6E, S6E). Taken together, utilization of our primary human microglia expression profiles in conjunction with iPSC-models could better inform on the type and timing of the model used to functionally interrogate the role of these NDD-associated genes on microglia and brain homeostasis.

To discern potential TFs that may mediate the diverse transcriptomic profiles between the different model systems, we performed ATAC-seq on iMGs, oMGs, and xMGs. PCA analysis (Figure 6F) showed a gradual reduction in PC1 variance in open chromatin between the xMGs and primary human fetal and postnatal microglia as the xMGs “matured” in the brain, reflecting the general trend that maturation of xMGs leads to gaining of qualities similar to *ex vivo* microglia. Relationships between xMGs and primary microglia defined by ATAC-seq were slightly different from our RNA-seq analysis (Figure 6D), most likely due to the inherent biology interrogated by each assay. Pairwise comparison of oMGs and xMGs 8-week post-engraftment revealed 26,307 differential peaks (Figure 6G). Motifs associated with members of AP-1, RUNX and IRF family were enriched in oMGs, while motifs highly enriched in xMG-8weeks corresponded to members of the IRF, MEF2 and SMAD family, implying alterations in TF binding and enhancer activation as the iPSC model transitions from an *in vitro* to an *in vivo* multi-cellular environment. Consistent with these findings, the expression of *MEF2C*, *IRF8*, and *SMAD4* were all significantly increased in xMG-8weeks as compared to the *in vitro* systems (Figure 6H), indicating that the mature mouse brain environment provides the necessary signals to induce the expression of these factors.

To investigate how well the *in vitro* and xenotransplantation models recapitulate the heritability of SNPs associated with complex traits of primary human microglia, we performed linkage disequilibrium score (LDSC) regression analysis[66] on these datasets. While we found no difference in the enrichment of SNPs associated with Alzheimer’s disease and multiple sclerosis between the distal regulatory elements of fetal and postnatal microglia, there was a stepwise increase in enrichment of these risk variants in these genomic regions from the *in vitro* iPSC-models to the xMGs (Figure 6I), suggesting that a multi-cellular environment is necessary to open DNA regions associated with neurodegenerative disease risk SNPs. In summary, our analysis of the transcriptome and open chromatin profile show that iPSC-derived microglia are similar to fetal microglia, but as these cells are incorporated into cerebral organoids or engrafted into mouse brains, they are educated by the changing environmental signals that drive the activation of TFs and alteration of the genomic landscape resulting in cellular profiles that best align with those of primary human postnatal microglia.

Xenotransplantation recapitulates environmental-dependent transcription factor networks

Primary microglia are highly sensitive to environmental changes as *in vitro* culturing rapidly downregulates hundreds of genes[13]. We thus used the xenotransplantation system which allows the maturation of progenitor cells to provide insight into the TF networks that may be required for maintaining microglial homeostatic gene patterns.

We analyzed the expression of genes that were differentially expressed between *ex vivo* human microglia and human microglia cultured *in vitro* for 7 days (Figure 7A) across the iPSC-model systems. Most notably, iPSC-microglia systems, from iMGs to matured xMGs, stepwise decreased expression of genes that were upregulated in primary microglia upon culturing (Figure 7B) and increased expression of genes that were downregulated in primary microglia upon culturing (Figure 7C). Some microglial genes, such as *SALL1*, a transcriptional regulator of microglia homeostasis *in vivo,* was rapidly downregulated upon microglia culturing, required 8 weeks in the mouse brain until expression was similar to primary human microglia (Figure 7C). This delay in upregulation of microglia-specific genes in the engrafted cells was markedly different from the gene expression kinetics in the murine brain, in which yolk-sac progenitors in the mouse rapidly increased expression of microglia genes, such as *Sall1*, upon entry into the brain (Figure S7D). These differences in expression dynamics between the engraftment system and mouse system could be due to cell intrinsic and/or extrinsic mechanisms.

Additionally, a subset of genes upregulated in *in vitro* as compared to *ex vivo* microglia were also upregulated in primary human fetal as compared to postnatal microglia ($p < 6.911e^{-95}$) (Figure S7C, top). GO analysis of these 464 genes revealed enrichment for terms related to ATP metabolic processes, cell cycle, and phagosome (Figure S7C, bottom). This pattern of *in vitro* culture conditions leading to a partial upregulation of a developmental program was also seen with primary mouse *ex vivo* and *in vitro* microglia[13].

We next utilized TIMON to determine the microglial TF networks that may be lost or gained in response to environmental changes. MEF2C, SMAD family members, and IRF7 and IRF9 were absent in the *in vitro* (Figure 7D) as compared to the *ex vivo* (Figure 7E) microglia TF network. Additionally, the number of unique TF connections (i.e., node size) of major microglial LDTFs PU.1 and IRF8 were reduced in the *in vitro* microglial network (Figure 7D,E). Third, the *in vitro* microglia TF network also had an enrichment of and an increase in the TF connectivity of AP-1 family factors (Figure 7D), with involvement in the TF clique JUN-C/EBPα-SREBP2 (Figure 7F). Pathways associated with the genes linked to enhancers containing these motifs include endocytosis (e.g., *CD163*), regulation of cholesterol metabolism storage and angiogenesis (Figure 7F). Expression of *SREBF2* (SREBP2), which regulates cholesterol synthesis [67], was rapidly induced in microglia upon culturing (Figure S7E). Alternatively, *EGR1* is a TF present in the *ex vivo* microglia TF network that was rapidly downregulated upon microglia culturing (Figure S7F). IRF8-EGR1-STAT2 form a clique that is found in enhancer regions associated with genes linked to regulation of leukocyte differentiation, activation (e.g., *MEF2A*), and immune effector processes (Figure 7G). Due to the expression differences of *SREFP2* (SREBP2) and *EGR1*, we then asked if there were additional TFs with environmental dependence. Indeed, we found that *IRF9* and *SMAD3*, part of the *ex vivo* TF network, were also rapidly downregulated upon culturing, while *NFKB2* and *MAFF*, were significantly higher expressed in *in vitro* microglia as compared to *ex vivo* microglia (Figure 7H). The expression of these TFs in the xMG system were nearly equivalent to postnatal human microglia by 8-weeks post-engraftment, again suggesting that the mouse brain environment provides the necessary signals for primitive hematopoietic progenitor cell differentiation into microglia (Figure S7D). Collectively, these results suggest that TIMON and TFAct are computational approaches that identify unique TF networks that are microglial-stage and subtype specific. These TF networks have variable timing and reconstitution in iPSC-model systems and xenotransplantation models, informing future disease specific studies and modeling.

Discussion

Recent single cell RNA sequencing studies have revealed the molecular heterogeneity of human fetal microglia and provided initial insights into underlying transcriptional networks[14,15,68]. However, the mechanisms by which the brain environment controls these networks during the transition from fetal to postnatal remain poorly understood. Here, we have advanced these efforts by providing data resources on the active enhancer repertoire of fetal and postnatal microglia. We then built two methods that integrate multi-omic datasets to infer roles of TFs in microglial development, TIMON and TFAct. TIMON is a technical advancement from the current motif enrichment analysis by extending consideration of TF function on enhancers from individual to groups of TFs, consistent with the well-established mechanisms that multiple TFs are required for enhancer selection and activation[12,31]. In contrast, TFAct leverages single-cell RNA-seq and epigenetic peak data to infer TF activity with single-cell resolution. Existing tools to predict TF activity, including MARA[69,70] and SCENIC[71], are limited by their focus on promoter-associated TF motifs making them insufficient to decipher cell types with extensive epigenetic controls from distal regulatory elements such as microglia[31,72].

Our application of TIMON to total human microglia enhancers, fetal and postnatal, captured the majority of TFs known to regulate microglia identity, including PU.1 and IRF8, which are well-established microglial LDTFs[73]. Notably, we did not find SALL1, a transcriptional regulator of microglial identity, in our co-occurrence network due to the lack of a SALL1 motif. However, previous studies have suggested that SALL1[74], and its family member SALL4[75], localizes to A/T rich regions. As the MEF2 motif is A/T-rich, it is possible that SALL1 is binding at these MEF2 sites or indirectly co-binding with MEF2 family members.

Of note, TIMON and TFAct uncovered MiTF-TFE family members as prominent collaborative TF specifically for fetal microglia. The MiTF-TFE family are basic helix-loop-helix leucine zipper TFs that are master regulators of lysosomal processes[40] and have been found to associate with PU.1 at promoters in phagocytes[76,77]. Furthermore, genes associated with lysosomal diseases were more highly expressed in fetal microglia than in the fetal cortex. Given that microglia shape the neural progenitor pool and the developing neural circuity through phagocytosis of apoptotic and viable progenitor cells and synapses[78,79], these results strongly suggest the importance of future investigations into the role of MiTF-

TFE family members in the microglial contribution to the neurological sequelae of pediatric lysosomal storage disorders.

A key challenge in discerning human microglia function, particularly in a disease state, is the lack of applicable human models, both *in vitro* and *in vivo*. First, there are substantial divergent transcripts between human and mice microglia and the developmental timeline of the human brain, particularly fetal development, is not replicated in its entirety by murine models. Second, microglia rapidly downregulate core microglial genes upon removal from the brain environment[13], thus rendering conclusions made through *in vitro* studies with *ex vivo* microglia confounding. Here we have shown that the iPSC-derived microglia system, cultured alone or co-cultured with cerebral organoids, are highly similar in their transcriptome and open chromatin profile to primary human fetal microglia. Our time course of engrafted microglia transcriptomes shows a gradual adoption of the human postnatal microglia signature between three and eight-weeks post-transplantation. Some core microglia genes, such as *SALL1*, did not increase in their expression after seven days post-transplantation into an intact brain environment. Potential explanations may include that the signals needed for proper microglial differentiation, beyond CSF1, are not cross-species compatible, that the necessary ligands are expressed too low in the mouse brain, or that the starting progenitor cells do not represent a true yolk-sac progenitor. Nevertheless, our studies indicate that the iPSC-based models can be utilized experimentally to dissect human microglia dynamics according to desired age and functional outcomes, with unique applicability towards a disease context as patient somatic cells can be reprogrammed and differentiated into microglia[80,81].

In concert, the present studies advance our understanding of the transcriptional networks that influence human microglia states throughout fetal and postnatal development, fulfilling a crucial unknown in the elucidation of human microglial dynamics. Additionally, our innovative computational frameworks, TIMON and TFAct, can be applied to multiple different cell types to uncover mechanisms of cell-type specific enhancer selection that establish cellular phenotypic heterogeneity in homeostasis, activation, and disease.

Limitations of Study

The degree to which sex, donor age, and genetic background affect microglial maturation and the interaction to the brain environment in xenotransplantation is currently unknown. To further the utility of the cerebral organoid system, it would be of worth to use region specific organoids (e.g., cerebellar, forebrain, hippocampal, etc.) in iMG co-cultures to determine how well these models influence microglia heterogeneity in addition to how microglia contribute to regional specificity of neuronal and glial subpopulations. For our enhancer-gene association, we used the distance-based method. Although currently there is no computational framework that outperforms distance-based methods across cell types[82], logically, distance-based methods will miss sporadic long-range interactions.

STAR Methods

LEAD CONTACT and MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nicole G. Coufal (ncoufal@health.ucsd.edu).

Materials Availability—This study did not generate new unique reagents.

Data and code availability—Code for WGCNA, scRNA-seq velocity analysis, TIMON, TFAct, and TROM analysis are available here: https://github.com/rzzli/FetalMicroglia. Code for scATAC-seq BAM/BED files processing, sparse matrix creation, TSS enrichment computation, matrix clustering and visualization are available here: https://gitlab.com/Grouumf/ATACdemultiplex. Description of the custom set of the 870 non-redundant motifs used as input for chromVAR analysis is described here: https://github.com/GreenleafLab/chromVARmotifs. The UCSC browser session (hg38) containing the processed RNA-seq, ATAC-seq, and ChIP-seq datasets for human fetal and postnatal microglia, iMGs, oMGs and xMGs is available at: https://genome.ucsc.edu/s/czh002/hg38_huMicroglia_Dataset. All primary human data are available through NCBI dbGaP, accession number phs001373; all other data is available through GEO: GSE226708.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human tissue—Microglia were isolated from postnatal brain tissue (in excess of that needed for pathological diagnosis) as previously described. All postnatal patients were diagnosed with refractory epilepsy and had epileptogenic focus resections at either Rady Children’s Hospital or through the UC San Diego Medical System (Jacobs Medical Center or UC San Diego Hospital). Brain tissue was obtained with informed consent from adult patients, or by informed parental consent and assent when applicable from pediatric patients under a protocol approved by the UC San Diego and Rady Children’s Hospital Institutional Review Board (IRB 160531, IRB 171361). Resected brain tissue was immediately placed on ice and transferred to the laboratory for microglia isolation or post fixation for histology within three hours after resection. Charts were reviewed for final pathological diagnosis, epilepsy medications, demographics, and timing of stereoelectroencephalography (SEEG) prior to surgery. Fetal brain samples were collected under a protocol approved by the UC San Diego Institutional Review Board (IRB 171379). Fetal brain samples were obtained within 1 hour of the pregnancy termination procedure after informed consent and transported in saline then were immediately either utilized for microglial isolation or were postfixed for histology. Frozen fetal tissue were obtained from the Birth Defects Research Laboratory at the University of Washington with ethics board approval and maternal written consent. This study was performed in accordance with ethical and legal guidelines of the University of California institutional review board. Postmortem time to freezing had a PMI of <8 hours. The reported data sets are from sequential samples for which cell viability and sequencing libraries met technical quality standards. No other criteria were used to include or exclude samples. All relevant ethical regulations were complied with.

Mice—Mice used in this study were bred and maintained at the Sanford Consortium for Regenerative Medicine. Humanized immunodeficient mice were sourced from Jackson Laboratory (Strain #017708) and were bred to Csf1r(ΔFIRE/ΔFIRE) that were generated by Drs. Clare Pridans and David Hume[62]. All animals were maintained, and procedures performed in accordance with University of California San Diego research guidelines for the care and use of laboratory animals.

Human microglia isolation—Human brain tissues were manually dissected into small 2–3 mm pieces and immersed in homogenization buffer (HBSS (Life Technologies, 14175–095), 1% bovine serum albumin (Sigma-Aldrich, A3059), 1 mM EDTA)) for mechanical dissociation using a 2 ml polytetrafluoroethylene pestle (Wheaton, 358026). Postnatal human microglia were isolated using an approach that combined Percoll enrichment and flow cytometry purification. Brain homogenate was pelleted, filtered through 40μμm filter, re-suspended in 37% isotonic Percoll (Sigma, P4937) and centrifuged at 600xg for 30 min at 16–18°C with minimal acceleration and no deceleration. Percoll gradients were utilized for all postnatal samples and only for fetal samples > 500mg. Following Percoll gradient centrifugation, pelleted cells were collected and washed twice with homogenization buffer, filtered with a 40 μm strainer (BD Falcon 352350) and incubated with Fc-receptor blocking antibody (Human TruStain FcX, BioLegend 422302) in homogenization buffer for 20 minutes on ice. Then cells were stained with the following cell surface marker antibodies for 30 min on ice (1:100 dilution, all from BioLegend): CD11b-PE (301306, clone ICRF44,), CD45-APC/Cy7 (304014, clone HI30), CD64-APC (305014, clone 10.1), CX3CR1-PerCP/Cy5.5 (341614, clone 2A9–1), CD14-AF 488 (301811, clone M5E2), HLA-DR-PE/Cy7 (307616, clone L243), and CD192-BV510 (357217, clone K036C2). CD14 and HLA-DR were included to further characterize immune cells but did not further discriminate subsets of microglia. Zombie Violet (Biolegend, 423113) or DAPI was added to the samples for viability discrimination immediately prior to sorting (1 μg/ml final concentration). Microglia were purified with either a BD Influx (100-μm nozzle, 22 PSI, 2-drop purity mode, sample chilling) or BD FACS AriaFusion (100-μm nozzle, 20 PSI, Purity mode (a 1–2 drop sort mode), sample chilling) and defined as live/DAPI-/Zombie violet-CD11b$^+$CD45^{Low}CD64$^+$CX3CR1^{High} single cells (Figure S1A). Flow cytometry data were also analyzed using FlowJo software (Tree Star).

Human nuclei isolation—Nuclei was isolated from frozen fetal brain tissue as previously described[72]. Briefly, frozen brain tissue was homogenized in 1% formaldehyde for 10 minutes and then quenched with 2.625M Glycine. Homogenate was then lysed in NF1 buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 5 mM $MgCl_2$, 0.1 M sucrose, and 0.5% Triton X-100) for 30 minutes and dounced with a 7mL glass dounce using both the loose and tight pestle. Then the homogenate was strained through a 70 μm strainer and then a 1.4M sucrose cushion was underlayed the resulting solution. Following centrifugation, the interphase, containing the nuclei was removed and stained overnight with anti-mouse/human NeuN-AF488 (MilliporeSigma, MAB377X), anti-mouse/human OLIG2-A647 (Abcam, ab225100), anti-mouse/human PU.1-PE (Cell Signaling, 81886S). Nuclei were then sorted the following day and flash-frozen for ChIP-seq studies.

Human pluripotent stem cell culture—All studies were conducted according to the human stem cell (hESCRO) protocol approved by the Embryonic Stem Cell Research Oversight (ESCRO) Committee at University of California, San Diego (IRB 171379). Human embryonic stem cell (ESC) line H1 (WiCell Research Institute, Madison, WI)[83] and induced pluripotent stem cell (iPSC) line EC11, derived from primary human umbilical vein endothelial cells (Lonza, Bioscience)(Firth et al., 2014), and an additional iPSC cell line derived from a control human aging cohort (UKERfG3G-X-001)(Mertens et al., 2015), were cultured utilizing standard techniques. In brief, cells were cultured in StemMacs iPS-Brew media (Miltenyi Biotech, Auburn, CA) or mTESR plus (Stem Cell Technologies, Vancouver, BC) and routinely passaged utilizing Gentle Cell Dissociation Reagent (Stemcell Technologies, Vancouver, BC) onto Matrigel-coated (1 mg ml^{-1}) plates. Normal karyotype was established by standard commercial karyotyping (WiCell Research Institute, Madison, WI).

Hematopoietic progenitor cell (iHPC) differentiation from iPSC/ESCs—Microglia were generated as previously described with minor modification[84]. Briefly, ESC/iPSCs were plated in iPS-Brew with 10 μM ROCK inhibitor (StemCell Technologies, Vancouver, BC) onto Matrigel-coated (1 mg ml^{-1}) 6-well plates using ReLeSR (StemCell Technologies). Cells were differentiated to CD43+ hematopoietic progenitors using the StemCell Technologies Hematopoietic Kit (Cat #05310). On day 1, cells were changed to basal media with supplement A (1:200), supplemented with an additional 1 ml/well on day 3, and changed to basal media with supplement B (1:200) on day 3–4 depending on cellular morphology. Cells received an additional 1 ml/well of medium B on days 5, 7, and 10. For xenotransplantation studies, iHPCs were collected on day 10 and frozen and the thawed the day prior to xenotransplantation.

Microglia differentiation from iHPC—Nonadherent hematopoietic cells were collected between days 11–14 depending on the differentiation. Cells were then replated onto Matrigel-coated plates (1 mg ml^{-1}) at a density of 300,000 cells/well in microglia media. Microglia media consisted of DMEM/F12 (ThermoFisher Scientific Scientific), 2x insulin-transferrin-selenite (Gibco), 2x B27 (Lifetech), 0.5x N2 (Lifetech), 1x GlutaMAX (Gibco), 2x non-essential amino acids (Gibco), 400 μM monothioglycerol, and 5 μg ml^{-1} insulin (Sigma). Microglia media was supplemented with 100 ng ml^{-1} IL-34 (Proteintech), 50 ng ml^{-1} TGFβ1 (Proteintech) and 25 ng ml^{-1} M-CSF (Proteintech). Cells were supplemented with microglia media with IL-34, TGFβ1 and M-CSF every other day. 25 days after initiation with microglia media, cells were resuspended in microglia media with IL-34, MCSF and TGFβ1 with the addition of CD200 100 ng ml^{-1} (Novoprotein) and CX3CL1 100 ng ml^{-1} (Peprotech). Cells were collected on Day 28 for experiments.

Cerebral organoid differentiation—Cerebral organoids were generated as previously described with modifications[85]. Briefly, iPSCS were grown on Matrigel, then washed with DMEM/F12 (Sigma) and were dissociated with collagenase at a concentration of 1.5mg/mL for 1 hour until colonies floated from the plate. Colonies were washed with DMEM/F12 several times and allowed to settle by gravity then were plated in embryoid body formation media (DMEM/F12 supplemented with 2mM GlutaMAX, 1% non-essential amino acids (ThermoFisher Scientific), 50nM 2-mercaptoethanol (Gibco), 20% knockout serum replacement, 2 μm dorsomorphin and 2 μm A83) with the addition of ROCK inhibitor Y27632 (40 μM) and 50ng mL^{-1} of bFGF-2 in an ultralow attachment plate (Day 1). On days 3 and 5 cells were fed with EB formation inhibitor without ROCK inhibitor or FGF. On day 5 organoids were transitioned to media for neural induction, consisting of DMEM/F12 with 2mM GlutaMAX, 1x N2 supplement (Gibco), 1% NEAA, 10ug mL^{-1} Heparin (Sigma), 1mM CHIR99201 and 1mM SB431542. On day 7 embryoid bodies were manually embedded in 18μL droplet of Matrigel for 30min at 37°C. Organoids were transitioned from neural induction to long term differentiation media on Day 12–14, consisting of DMEM/F12 with 2mM GlutaMAX, 1x N2, 1x B27 (ThermoFisher Scientific), 2.5ug ml^{-1} insulin, 55 μM 2-mercaptoethanol, 1% NEAA, 1% pen/strep. At this time organoids were moved to an orbital shaker for the remainder of the culture time. Excess Matrigel was manually removed on day 20, and cerebral organoids were utilized at 8–12 weeks of age for delineated studies.

Immunofluorescence staining and analysis of cerebral organoids—Organoids were fixed in 4% paraformaldehyde in 0.1 M Phosphate buffer saline for 45 – 60 minutes at 4°C and washed three times in PBS, then cryoprotected in 30% sucrose and embedded in tissue freezing medium (General Data) for cryo-sectioning. Twenty-μm sections were cut on a cryostat, mounted on Superfrost plus slides (Thermo Fisher Scientific, Menzel-Glaser), and stored at −80°C until staining. For immunofluorescence, sections were rehydrated, rinsed in 0.1 M TBS, then permeabilized and blocked for non-specific binding in blocking buffer containing 3% normal horse serum and 0.25% Triton X-100 (Sigma X100) in a humidified chamber for 1hr at room temperature. Slides were then incubated with the appropriate primary antibodies diluted in blocking buffer at 4°C overnight. The next day, sections were washed twice (fifteen minutes each) in 0.1M TBS, washed with blocking buffer (once for 30 minutes), and incubated with fluorophore-conjugated secondary antibodies diluted in blocking solution at RT for 2 hrs. After the two-hour incubation, sections were counter stained with DAPI for 10 minutes, rinsed three times in 0.1M TBS (15 minutes each), rinsed with 0.1 M PO_4, and mounted with Shandon Immuno-Mount (Thermo Fisher Scientific, 9990412). Imaging was performed on a Leica TCS SPE confocal microscope or a Nikon Eclipse Ti2-E with laser scanning confocal A1R HD.

RNA Scope—RNAscope probes for *P2RY12, SORL1, APOE, CCL3*, and *CCR5* were obtained from Advanced Cell Diagnostics (ACD). Slides were taken from −80 °C and dried at −20 °C for 2 hours, and RNA Scope Multiplex Fluorescent Reagent Kit V2 User Manual was followed. In summary, sections were washed with PBS for 5 minutes and baked for 30 mins at 60 °C. Sections were fixed in 4% PFA for 15 minutes at 4°C and dehydrated in an alcohol gradient, 50%, 70%, and 100% for 5 minutes each time. Sections were then incubated with ACD H2O2 for 10 mins and washed in water before treatment with target retrieval buffer (ACD) for 5 minutes at 98–100 °C. Sections were washed with water, treated with 100% alcohol, and baked at 60 °C for 5 minutes. Sections were then treated with Protease III for 30 minutes at 40 °C in the RNA scope oven (HybEZ) and washed in water. Hybridization of the probes was added (~ 2 drops/ section) and incubated for 2 hrs at 40 °C followed by a rinse in ACD wash buffer. Sections were placed in 5x saline sodium citrate overnight at RT. Slides were rinsed with wash buffer and followed by the addition of Amp 1 for 30 mins at 40°C, and rinsed in wash buffer twice for 2 minutes each. The same process was used for the addition of Amp2, and Amp 3, however, Amp 3 was only incubated for 15 minutes. To allow for signal detection, HRP-C1 was added for 15 mins at 40 °C, followed by two 2 minute washes in wash buffer, followed by C1 probe incubation (1:750) (TSA Vivid Fluorophore) for 30 mins at 40 °C. Slides were subsequently washed twice for 2 minutes, each time using wash buffer, and blocked using RNAscope Multiplex FL v2 HRP blocker for 15 mins at 40 °C followed by two, 2 minute washes in wash buffer. The same steps were used to develop C2 and C3 signal. DAPI (ACD) was then incubated for 30 seconds at RT followed by 2 drops of ProLong Gold Antifade Mountant on slides and coverslipped. Sections were left to dry overnight at RT in the dark and placed at 4 °C for storage or until imaging. For sections stained with IBA1 for immunochemistry, C3 probe step was skipped, followed by three 10 minute washes in PBS with 1% Triton. Sections were blocked for 1 hour using 10% BSA, and IBA1 antibody was incubated overnight at 4 °C (1:200). Sections were washed with PBS for 10 minutes, three times, and incubated in secondary Cy3 (1:500) for 2 hrs at RT, followed by two 10 minute PBS washes. DAPI (1:1000) was incubated for 10 minutes in PBS, and then slides were mounted using immunomount and let to dry overnight at RT in the dark. Images were acquired on a Leica TCS SP8 using a 40x (NA 1.15) objective. Images were then processed on Adobe Photoshop (CS7), where adjustments to image brightness and contrasts were made.

Isolation of iMGs and oMGs—For iMGs, cells were carefully manually scraped from the Matrigel coated plate and were concentrated by a 300 rcf x 5 minutes. For oMGs, organoids were carefully collected, allowed to settle by gravity, and then dissociated mechanically in staining buffer (HBSS 1x with 1mM EDTA and 1% BSA) using a 2 ml polytetrafluoroethylene pestle (Wheaton, 358026) in a fashion identical to fetal brain tissue. No percoll gradient was utilized. Both iMGs and oMGs were resuspended in staining buffer and were blocked with Fc receptor blocking antibody (Human TruStain FcX, BioLegend 422302) for 10 minutes. Both iMG and oMGs were stained with the following 6 antibodies, all at 1:30 dilution and all from Biolegend: CD64-APC (305014), CX3CR1 PCP-Cy5.5 (341614), CD14–488 (325610), CD11b-PE (301309), HLADR PE-Cy7 (307616), CD45-APC-CY7 (368516) for one hour. Cells were then washed and incubated in Zombie Violet (1:1000, Biolegend) for live/dead discrimination. Controls consisted of cells incubated with a combination of appropriate isotypes for each antibody (Biolegend). Microglia were purified on a BD InFlux Cytometer (Becton-Dickinson).

Xenotransplantation and isolation of xMGs—All xMGs were performed using the progeny from humanized immunodeficient mice bred to the Csf1rΔFIRE/ΔFIRE or using the humanized immunodeficient mice treated with CSF1R inhibitor BLZ945 (Selleckchem, S7725) once a day i.p. for 2 days prior to iHPC injection. Neonatal pups were genotyped at P0 and pups between P1 – P3 were used for engraftment. Engraftment was performed according to Hasselmann et al.[59], with approximately 200 to 250K cells injected per mouse. xMGs were isolated and sorted at indicated time points following engraftment as according to human microglia isolation.

Tissue processing for immunostaining—For fixation, fetal and postnatal tissue was fixed in 4% formaldehyde in phosphate-buffered saline (PBS) overnight at 4°C then transferred to 30% sucrose. The tissue was sectioned in 20-μm sections using a cryostat. Sections were stored at −80°C until staining.

Immunostaining—Sections were washed in 0.1M TBS and then non-specific binding site blocking and cell permeabilization was performed with blocking buffer containing 3% normal horse serum and 0.25% Triton X-100 (Sigma X100). Sections were incubated with primary antibody (see list below) in blocking buffer at 4°C overnight. After washing in 0.1M TBS, sections were incubated with secondary antibodies (see list below) for two hours at room temperature (1:250, Jackson Laboratories). Sections were washed with TBS before nuclear counterstaining (DAPI 1:1000, Thermo Fisher Scientific) and then mounted with Shandon Immu-Mount (Thermo Fisher Scientific, 9990412). Imaging was performed on a Leica TCS SPE confocal microscope or a Nikon Eclipse Ti2-E with laser scanning confocal A1R HD.

Cell cultures were fixed with 4% paraformaldehyde solution for 30 minutes at room temperature. Antigen blocking and cell permeabilization were performed using blocking buffer consisting of 3% horse serum and 0.25% Triton X-100 (Sigma) in TBS for 1 hour at room temperature. Primary antibodies were incubated in 3% horse serum overnight at 4°C, and secondary antibodies (1:250, Jackson Laboratories) were incubated in the same solution for 1 hour at room temperature. The cells were counterstained with DAPI for nuclei detection and mounted with Shandon Immuno-Mount (Thermo Fisher Scientific, 9990412).

The following primary antibodies were used: goat Iba1 (1:200, Abcam, ab5076), rabbit Iba1 (1:500, Wako, 019–19741), mouse Nestin (1:500, EMD Millipore, ABD69), chicken MAP2 (1:500, Abcam, ab5392), rabbit PU1 (1:250, Cell Signaling, 2266S), rabbit P2RY12 (1:200, HPA014518), rat CX3CR1 (1:100, Biolegend, 341602), rabbit Sox2 (1:200, Cell Signaling, 2748S), rabbit MafB (1:200, Abcam, ab223744), mouse Ki67 (1:1000, Leica, ACK02), rabbit Lyve1 (1:200, Abcam, ab36993), mouse CD163 (1:100, BIORAD, MCA1853), rabbit Laminin (1:100, EMD Millipore, MABE622), mouse Laminin (1:100, Telios/Gibco BRL).

The following secondary antibodies were used (all at 1:250): Donkey Cy3 anti Goat (Jackson Laboratories, 705-165-147), Donkey Alexa Fluor 488 anti-Goat (Jackson Laboratories, 705-545-147), Donkey Alexa Fluor 647 anti-Goat (Jackson Laboratories, 705-175-147), Donkey Cy3 anti Rabbit (Jackson Laboratories, 711-165-152), Donkey Alexa Fluor 488 anti-Rabbit (Jackson Laboratories, 711-545-152), Donkey Alexa Fluor 647 anti-Rabbit (Jackson Laboratories, 711-175-152), Donkey Cy3 anti Mouse (Jackson Laboratories, 715-165-151), Goat Alexa Fluor 488 anti-Mouse (Jackson Laboratories,715-545-151), Donkey Alexa Fluor 647 anti-Mouse (Jackson Laboratories, 715-545-151), Donkey Alexa Fluor 488 anti-Chicken (Jackson Laboratories, 703-545-155), Donkey Alexa Fluor 647 anti-Rat (Jackson Laboratories, 712-605-153).

Assay for Transposase-Accessible Chromatin-sequencing (ATAC-seq)—30,000–50,000 isolated human microglia, iMGs or oMGs were lysed in 50 μl lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL, CA-630, in water). Resulting nuclei were centrifuged at 500 rcf for 10 minutes. Pelleted nuclei were resuspended in 50 µl transposase reaction mix (1x Tagment DNA buffer (Illumina 15027866), 2.5 µl Tagment DNA enzyme I (Illumina 15027865), and incubated at 37°C for 30 min on a heat block. For isolations resulting in under 30,000 microglia, microglia directly placed in 50 µl transposase reaction mix, as indicated above and incubated for 37°C for 30 min. DNA was purified with Zymo ChIP DNA concentrator columns (Zymo Research D5205), eluted with 11 µl of elution buffer, and amplified using NEBNext High-Fidelity 2x PCR MasterMix (New England BioLabs M0541) with the Nextera primer Ad1 (1.25 µM) and a unique Ad2.n barcoding primer (1.25 µM) for 8–12 cycles. Libraries were size-selected through gel excision for fragments that were 175–255 bp and single-end sequenced for 51 cycles on a HiSeq 4000 or NextSeq 500.

Chromatin immunoprecipitation-sequencing (ChIP-seq)—FACS-purified microglia were spun down at 300 rcf and resuspended in 1% PFA. After rocking for 10 minutes at room temperature, PFA was quenched with 1:20 2.625M glycine for 10 minutes at room temperature. Fixed cells were washed twice and centrifuged at 800–1000 rcf for 5 minutes and snap frozen in liquid nitrogen. Snap-frozen microglia (750,000) were thawed on ice, resuspended in 130 µl of LB3 buffer (10 mM TrisHCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-Lauroylsarcosine, 1x protease inhibitors), and transferred to microtubes with an AFA Fiber (Covaris, MA). Samples were sonicated using a Covaris E220 focused-ultrasonicator (Covaris, MA) for 12 cycles of 60 secs (Duty: 5, PIP: 140, Cycles: 200, AMP/Vel/Dwell: 0.0). The sonicated sample was transferred to an Eppendorf tube, to which Triton X-100 was added to achieve a final concentration of 1%. Supernatant was spun at 21,000 rcf and the pellet discarded. 1% of the total volume was saved as DNA input control and stored at −20°C until library preparation. For the immunoprecipitation, 25 µl of Protein A DynaBeads (Thermo Fisher Scientific 10002D) and 1 µl of H3K27ac antibody (Active Motif, 39085) were added to the supernatant and rotated at 4oC overnight. Dynabeads were washed 3 times with Wash Buffer 1 (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), three times with Wash Buffer 3 (10 mM Tris-HCl pH, 250 mM LiCl, 1 mM EDTA, 1% Triton X100, 0.7% Na-Deoxycholate), three times with TET (10 mM Tris-HCl pH 8, 1 mM EDTA, 0.2% Tween20), once with TE-NaCl (10 mM Tris-HCl pH 8, 1 mM EDTA, 50 mM NaCl) and resuspended in 25 µl TT (10 mM Tris-HCl pH 8, 0.05% Tween20). Input samples were adjusted to 25 µl with TT. Sample and input libraries were prepared using NEBNext Ultra II DNA Library Prep kit (New England BioLabs E7645) according to manufacturer’s instructions. Samples and inputs were de-crosslinked (RNase A, Proteinase K, and 4.5 µl of 5M NaCl) and incubated overnight at 65°C. Libraries were PCR-amplified using NEBNext High Fidelity 2X PCR MasterMix (New England BioLabs M0541) for 14 cycles. Libraries were size-selected through gel excision for fragments that were 225 to 500 bp and single-end sequenced for 51 cycles on a HiSeq 4000 or NextSeq 500.

Locus-specific ChIP (ChIP-qPCR)—100K sorted PU.1-positive nuclei from fetal brain and 400K sorted PU.1-positive nuclei from postnatal brains were used for ChIP-qPCR. Briefly, ChIP for PU.1, MITF, TFEB, IRF7, NR2C2, MAFB, and PRDM1 were performed as mentioned above. Resulting library DNA was amplified for 40 cycles. Locus-specific primers used for amplification are as below: LGALS3: *ACATCTTCTGGGTTGGCTGA* (F), *GTTTCCTCACGAGGGTTCCT* (R), CD74: GAGCCAACCAGAGCTAGCAG (F), *ATGTGTGCTTAGGCCATTCC* (R). Negative control regions were detected by the following primers for fetal microglia (Chr8): *ACCAGGGGTGCAAAGTAGTG* (F), *TGAACAGTGCTTCCTTCACG* (R); negative control region primers (*VIL1*) for postnatal microglia: *CACCTTTGGAAGCTTCTTCG* (F), *GTGGGTTTTGCTCTCTGCTC* (R). Relative enrichment for factors was normalized to input and to negative indicated negative control regions.

mRNA isolation—Snap-frozen human fetal and postnatal cortical tissue were placed in TRIzol LS (Life Technologies) and homogenized using Powergen 125 homogenizer (Thermo Fisher Scientific). FACS-sorted cells were stored in TRIzol LS. Total RNA was extracted from homogenates and cells using the Direct-zol RNA MicroPrep Kit (Zymo Research R2052) and stored at −80°C until RNA-seq cDNA libraries preparation.

polyA RNA-seq—RNA-seq libraries were prepared as previously described[13]. Briefly, mRNAs were incubated with Oligo d(T) Magnetic Beads (New England BioLabs S1419), then fragmented in 2x Superscript III first-strand buffer (ThermoFisher Scientific 18080051) with 10mM DTT (ThermoFisher Scientific 18080044) at 94°C for 9 minutes. Fragment mRNA was then incubated with 0.5 μl of Random primers (3 mg/mL) (ThermoFisher Scientific 48190011), 0.5 μl of 50mM Oligo dT primer, (ThermoFisher Scientific, 18418020), 0.5 μl of SUPERase-In (ThermoFisher Scientific AM2694), 1 μl of dNTPs (10 mM) at 50°C for one minute. Then, 1 μl of 10mM DTT, 6 μl of H_2O+0.02%Tween-20 (Sigma), 0.1 μl Actinomycin D (2 mg/mL), and 0.5 μl of Superscript III (ThermoFisher Scientific) were added to the mixture. cDNA was then generated by incubating mixture in a PCR machine at the following conditions: 25°C for 10 minutes, 50°C for 50 minutes, and a 4°C hold. Product was purified using RNAClean XP beads (Beckman Coulter A63987) according to manufacturer’s instructions and eluted with 10 μl of nuclease-free H_2O. Eluate was then incubated with 1.5 μl of Blue Buffer (Enzymatics P7050L), 1.1 μl of dUTP mix (10 mM dATP, dCTP, dGTP and 20 mM dUTP), 0.2 mL of RNase H (5 U/mL Y9220L), 1.2 μl of H_2O+0.02%Tween-20, and 1 μl of DNA polymerase I (Enzymatics P7050L) at 16°C overnight. DNA was then purified using 3 μl of SpeedBeads (Thermo Fisher Scientific 651520505025) resuspended in 28 μl of 20% PEG8000/2.5M NaCl to final of 13% PEG. DNA eluted with 40 mL nuclease free H_2O+0.02%Tween-20 and underwent end repair by blunting, A-tailing and adaptor ligation as previously described[86] using barcoded adapters. Libraries were PCR-amplified for 12–15 cycles, size selected by gel extraction for 200–500 bp, and sequenced on a HiSeq 4000 (Illumina) or a NextSeq 500 (Illumina) for 51 cycles.

Single-cell RNA-seq—Sorted microglia were centrifuged for 5 minutes at 300 rcf and the supernatant was carefully aspirated leaving approximately 25 μl behind. Cells were resuspended in up to 40 μl reaction buffer (0.1% BSA (Sigma SRE0036–250ML) and 1 U/μl RNasin inhibitor (Promega PAN21110) in PBS (Corning 21-040-CV)). An aliquot of the cell suspension was mixed with Trypan Blue to count and check viability using a hemocytometer. 12,000 cells (viability 65–100%) were loaded onto a Chromium Controller (10x Genomics). Libraries were generated according to manufacturer specifications

(Chromium Single Cell 3’ Library and Gel Bead Kit v3, 1000075; Chromium Single Cell 3’ Library Construction Kit v3, 100078; Chromium Chip B Single Cell Kit, 1000153; Single Index Kit T Set A, 1000213). cDNA was amplified for 12 PCR cycles. SPRISelect reagent (Beckman Coulter B23318) was used for size selection and clean-up steps. Final library concentration was assessed by Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific Q32851) and fragment size was checked using the High Sensitivity D1000 ScreenTape assay on a Tapestation 4200 (Agilent) to ensure that fragment sizes were distributed normally about 500 bp. Libraries were sequenced using a NextSeq 500 or NovaSeq 6000 (Illumina) using these read lengths: Read 1: 28 cycles, Read 2: 91 cycles, Index 1: 8 cycles.

Single-cell ATAC-seq—Sorted microglia were centrifuged for 5 minutes at 300 rcf and the supernatant was carefully aspirated. Cells were permeabilized using lysis buffer (10 mM Tris-HCl pH 7.4 (ThermoFischer 15567027), 10 mM NaCl (Invitrogen AM9759), 3 mM $MgCl_2$ (MP Biomedicals Inc. 194698), 0.1% Tween 20 (Sigma P7949), 0.1% IGEPAL-CA630 (Sigma I8896), 0.01% Digitonin (Promega G9441), and 1% BSA in nuclease free water and incubated for 5 minutes on ice. Permeabilized nuclei were washed using washing buffer (lysis buffer without IGEPAL-CA630 and Digitonin), centrifuged for 5 minutes at 500 rcf and resuspended in Nuclei buffer (10x Genomics). An aliquot was mixed with Trypan Blue and counted using a hemocytometer. Up to 15,300 nuclei were tagmented before loading onto a Chromium Controller; libraries were generated according to manufacturer specifications (Chromium Next GEM Single Cell ATAC Library and Gel Bead Kit v1.1, 1000175; Chromium Next GEM Chip H Single Cell Kit, 1000162; Single Index Kit N Set A, 1000212, 10x Genomics). Libraries were amplified for 10 PCR cycles. SPRISelect reagent (Beckman Coulter B23318) was used for size selection and clean-up steps. Final library concentration was assessed by Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific Q32851) and fragment size was inspected using the High Sensitivity D1000 ScreenTape assay on a Tapestation 4200 (Agilent). Libraries were sequenced using a NextSeq 500 or NovaSeq 6000 (Illumina) using these read lengths: Read 1: 50 cycles, Read 2: 50 cycles, Index 1: 8 cycles, Index 2: 16 cycles.

Data analysis

Bulk analysis

Data preprocessing: Raw reads were obtained from Illumina Studio pipeline. ATAC libraries were trimmed to 30bp. RNA-seq data was mapped to hg38/mm10 genome using STAR[87] with default parameters. H3K27acetylation ChIP-seq and ATAC-seq data were mapped to hg38 genome using bowtie2[88] with default parameters. Finally, HOMER[86] tag directories were created for mapped samples.

External RNA-seq data: Data for mouse microglia development[89] and whole organoids[90] were obtained via the sequence read archive (SRA)[91] and were preprocessed with the same pipeline as described above.

RNA-seq data analysis: HOMER’s analyzeRepeats was used to calculate gene expression raw counts with the option “- condenseGenes -count exons -noadj” and transcript per kilobase million (TPM) with the option “- count exons -tpm”. Genes shorter than 200bp were removed. TPMs were matched to raw counts using accession numbers. Differentially expressed genes comparisons using human microglia, human cortical, iMGs, oMGs data were assessed with DESeq2 at an adjusted p-value < 0.05 and fold-change > 2 where indicated. Multiple testing correction was adjusted using the procedure of Benjamini and Hochberg under DESeq2 [92] framework. Mapping of orthologous genes between mouse and human was done using the one-to-one orthologs from the Ensembl (version 84) Compara database.

Weighted Correlation Network Analysis—WGCNA (Weighted Correlation Network Analysis) was performed using TPMs calculated using defined by HOMER's analyzeRepeats.pl function. Samples' tissue source (microglia or whole cortex), developmental age (fetal or postnatal) and gender were used as input traits. Soft threshold was set to 10 to reach 0.90 of scale free topology index when constructing the gene co-expression matrix. Minimum module size was set to 100 and maximum dissimilarity was set to 0.01 when merging modules. Module-trait correlations were computed using Pearson's correlation, modules with correlation coefficients p>0.7 or p<-0.7 and p-value <0.05, for at least one trait, were considered as significant and subjected to functional annotation. Gender trait was not used in further analysis as no gene module is significantly correlated with gender. Next, we computed the Kleinberg's hub centrality scores for genes in each significant module. We performed Gene Ontology (GO) enrichment analysis[93] on genes in the significant modules for functional annotation. Bioconductor package *topGO* was used to identify significantly enriched GO terms associated with genes in each module, the p-value and q-value cutoffs were set to 0.01 and 0.05, respectively. The enriched GO terms were used to annotate each gene module.

NicheNet Analysis—NicheNet models the influence of ligands expressed by a sender cell on a set of target genes in a receiver cell using a model of intracellular signaling linking receptors to target genes (available at https://github.com/saeyslab/nichenetr)[24]. NicheNet assesses whether a given ligand could drive transcription of a set of target genes relative to all expressed genes within a cell or tissue. In this study, we used NicheNet to identify ligands expressed by microglia (autocrine) and bulk cortex tissue (paracrine) that could drive fetal-specific or postnatal specific microglial gene expression. Genes that were differentially expressed at an p-adj. < 0.05 and fold change > 2 between conditions were used to establish the target set of genes for indicated states compared (e.g., the fetal and postnatal state), with all other genes serving as the background set. The NicheNet ligand-target model was filtered to only include ligands and receptors expressed above >5 TPM for both autocrine and paracrine analysis. Additionally, the NicheNet model was filtered to only include ligand-receptor interactions annotated by curated databases and exclude ligand-receptor interactions annotated by protein-protein interaction databases. Ligands that were expressed at least 8-fold higher in microglia compared to the corresponding developmental age-matched cortex were considered as likely autocrine signals, whereas ligands expressed at less than a than 2-fold increase in respective cortex versus microglia comparisons were considered to act on microglia in both an autocrine and paracrine fashion. Assignment of upstream receptors to downstream ligands in Figure 2A was performed manually using prior literature to support connections.

Neurodevelopmental disorders-associated genes—Genes associated with neurodevelopmental disorders were curated from the following: attention-deficit hyperactivity disorder[94,95], autism spectrum disorders and developmental delays[96], developmental epileptic encephalopathy[97], dyslexia[98–101], intellectual disability[102–105], schizophrenia[106–111], mitochondria disorders[112], lysosomal disorders[113,114].

ATAC-seq and ChIP-seq analysis—Peaks were called using HOMER's findPeaks command with the following parameters: "-style factor -size 200 - minDist 200" for ATAC-seq experiments and "-style histone -size 500 -minDist 1000 -region" for ChIP-seq experiments. Peaks were merged with HOMER's mergePeaks and annotated using HOMER's annotatePeaks.pl using all tag directories. For ChIP-seq experiments, peaks were annotated around ATAC-seq peaks with the parameter "-size −500,500 -pc". Subsequently, DESeq2[92] was used to identify the differentially chromatin accessible distal sites (1000bp away from known TSS) or proximal sites (<500bp away from known transcript) with p-adj < 0.05 and fold change > 2.

Motif analysis—*De novo* motif analysis was performed using HOMER's findMotifsGenome.pl with either all peaks or random genome sequences as background peaks. Motif enrichment scoring was performed using binomial distribution under HOMER's framework.

Linkage disequilibrium score regression analysis—Linkage disequilibrium score regression (LDSC) analysis[66] was applied to determine the degree of enrichment of variants associated with neurological traits in different groups of microglia epigenetic features. We obtained GWAS summary statistics for neurodevelopmental disorders [attention-deficit hyperactivity disorder[115], autism spectrum disorder[116], epilepsy[117]], neuropsychiatric [anorexia nervosa[118], anxiety disorders[119], bipolar disorder[120], major depressive disorder (MDD)[121], schizophrenia[110,122]], traits [general cognitive function[123], intelligence[124], neuroticism[125]], brain volume [brainstem, grey matter, right hippocampal, right putamen, white matter][126], neurodegenerative disorders [Alzheimer's Disease[127,128], Amyotrophic Lateral Sclerosis (ALS)[129], Parkinson's disease (PD)[130], Multiple Sclerosis (MS)[131]]. Enrichment p-values were computed using the LDSC default parameters and adjusted with the number of datasets and annotations.

Single cell RNA-seq processing—Raw sequencing data was demultiplexed and preprocessed using the Cell Ranger software package v3.0.2 (10x Genomics). Raw sequencing files were first converted from Illumina BCL files to FASTQ files using cellranger mkfastq for scRNA-seq. Demultiplexed FASTQs were aligned to the GRCh38 reference genome (10x Genomics) and reads for exonic reads mapping to protein coding genes, long non-coding RNA, antisense RNA, and pseudogenes were used to generate a counts matrix using cellranger count; expect-cells parameter was set to 5,000. Fastq files were processed using Cell Ranger 3.0.2 software to process barcodes and demultiplex reads using default parameters. The resulting count matrices were filtered and analyzed using Seurat package (version 3.2.2)[132,133]. Genes expressed in fewer than 3 cells and cells expressing fewer than 500 genes were removed from further analysis. Normalization was performed using the Seurat default parameters. We used Harmony (https://github.com/immunogenomics/harmony)[45] to integrate samples. Louvain algorithm with resolution=0.3 from Seurat package was applied to cluster cells. The result was projected on t-distributed stochastic neighbor embedding (tSNE) for visualization.

RNA velocity analysis—RNA velocity analysis was performed using the Velocyto package[46] (available at velocyto.org). Fastq files were trimmed and mapped using STAR to hg38 reference genome. The intronic and exonic information was extracted; and the ratios of the spliced and unspliced variants of each gene in every cell were computed; changes in these ratios for a gene in a cell allow for inferences in cell state changes among the population. Arrows point to the position of the future state and to infer cell trajectories.

TIMON—We developed an innovative computational framework to identify significantly enriched co-occurring motifs as a method to decipher the regulatory dynamics of microglia development. First, we developed a method to identify non-overlapping motifs from open chromatin/enhancer regions. For each open peak region (open chromatin or enhancers) at base location b, we compared the motif scores for all motifs for $\{TF1, TF2, TF3\ldots\}$. For motif TF_x to be selected as the putative motif for that given location, it needs to 1. $Score_b^{TFx}$ needs to be the highest score among all motifs at location b. 2. $Score_b^{TFx}$ needs to be higher than all motif scores n_{TFx} base pairs before location b, that is $Score_b^{TF_x} > \left\{Score_{b-1}^{\{TF1, TF2, TF3\ldots\}} \ldots Score_{b-n_{TFx}+1}^{\{TF1, TF2, TF3\ldots\}}\right\}$, where n_{TFx} represents the motif length of TF_x. We thus select TF_x as the motif for location $[b - n_{TF_x} + 1]$.

In the case TF_y has the highest score at a later location c, and $c - n_{TFy} > b$, if TF_y fulfills condition 1. and condition 2., then we then select TF_y as the motif for location $[c - n_{TFy} + 1]$ and remove TF_x as the motif for location $[b - n_{TFx} + 1]$. For example, in Figure 4A top panel, TF5 has higher score than TF4, thus TF5 is selected as the motif to occupy the region while the overlapping TF4 is removed.

We notice single iteration of the process above will leave gaps where can still fit suitable TF motifs (Figure 4A middle panel). Thus, we repeat the process in the remaining gaps of the genomic regions until a. all gaps are smaller than the smallest motif size or b. the iteration stops assigning TF motifs. In the schematic shown in Figure 4A, TF3 was reassigned to the gap because TF3 is the highest scoring motif and does not overlap with any existing selected motifs. After the non-overlapping motif discovery, we computed the co-occurrence of the motifs using the dot-product of the motif-peak matrix with its transpose. To identify significantly enriched co-occurring TF pairs, we first computed the co-occurrence of the TFs in the open chromatin regions in 35 cell types obtained from ENCODE project (https://www.encodeproject.org/). Next, we used the aggregated peaks from ENCODE projects as the background null distribution and used Monte Carlo procedures to randomly select size-matching peaks from background population 10,000 times. Finally, we compared the sample (e.g., microglia) co-occurrence of TF pairs with the background re-sampled background co-occurrences. Only TF pairs with percentile rank higher than 99.5 (equivalent of p-value 0.001) when compared to the background are selected as the significantly highly co-occurring TF pairs.

Determining effect of co-occurring motifs on TF binding—To evaluate the effects of having co-occurring motif on TF binding, we obtained PPARγ, STAT6, and C/EBPβ ChIP-seq data from IL-4 stimulated BMDMs in addition to H3K27ac-ChIP-seq data[34]. We used Monte Carlo resampling procedure to test if having the C/EBPβ motif would increase the probability of PPARγ/STAT6 bind to an active enhancer that has the PPARγ/STAT6motif. Significant difference is defined as the 95% confidence interval (CI) of the single motif only group (e.g., PPARγ only or STAT6 only) does not cover the mean of co-occurring group (PPARγ + C/EBPβ or Stat6 + C/EBPβ).

TFAct

Enhancer calling: Active enhancers were selected using the following criteria: 1) open chromatin region marked by ATAC-seq peak, 2) located outside of transcription start site (-2kb to +100bp), and 3) have at least 16 tag counts in the H3K27ac peak overlapping with the ATAC-seq peak. Peak calling and annotation were performed using Homer findPeaks and annotatePeaks.pl functions, as previously described. Fetal and postnatal specific enhancers are defined as enhancers that are not overlapped between fetal and postnatal groups.

Transcription Factor Motif: Instances of transcription factor motifs in enhancers were defined using the non-overlapping motif discovery method described in the section above. The reference motif library for motif identification was downloaded from HOCOMOCO motif library (https://hocomoco11.autosome.org/). We used the distance method (nearest gene) to assign enhancers to target genes. Motifs identified in enhancers that were assigned target genes were then used to construct transcription factor – target gene networks (TF-gene networks). Only transcription factors with mean expression at least 10 TPM from either bulk fetal RNA-seq or postnatal RNA-seq dataset were considered. Separate TF-gene networks were constructed for fetal and postnatal groups.

TF module activity score calculation: We used the fetal and postnatal microglia TF-gene networks and scRNA-seq expression data, normalized using Seurat, to compute activity score for each TF for each developmental age group. Activity score for each TF is computed using VISION[134] calcSignatureScores function with TF - gene network and scRNA-seq expression data as inputs.

Single-cell ATAC-Seq preprocessing and clustering—Raw sequencing files were first converted from Illumina BCL files to FASTQ files using cellranger-atac mkfastq. Demultiplexed FASTQs were aligned to the GRCh38 reference genome (10x Genomics) to identify chromatin accessible peaks using cellranger-atac count. Sequencing reads of the four donors were demultiplexed and processed using the Cell Ranger software package cellranger-atac v1.1.0 (10x Genomics). Reads were aligned to the human reference hg38 (Cell Ranger software package cellranger-atac-1.1.0/bwa/v0.7.17). The fragment files generated by Cell Ranger were then tagged by read and donor and combined into a unique fragment file. We then computed a Transcription Start Site (TSS) enrichment score for each cell using +/− 2kbp from the TSS as reference promoter regions. We used a flank size of 100pb at the beginning and end of the promoters, a smoothing window of 10bp, and a TSS region of 50bp to infer the maximum TSS enrichment for each cell. We called peaks on the merged fragment file using MACS2[135] with the following options: *-nomodel -keepdup-all -q 0.01 --shift 37 --extsize 73*. We then filtered the top 5% of the peaks, merged the overlapping peaks, and transformed each peak using its center with +/1000bp as boundaries. We filtered cells with less than 1500 fragments and with TSS < 7 and created a binary sparse *cell x peaks* matrix. We converted the sparse matrix into a lower dimensional embedding using a Latent Semantic Analysis (LSA) approach by weighting the features with a tf-idf scheme and extracted 25 dimensions using a Singular Value Decomposition (SVD) algorithm using *RobustSVD* and *TfIdfTransformer* classes from the scikit-learn package[136]. We used the embedding and the batch ID of each cell as input for Harmony[45] and inferred the next embedding (*ncells x 25*) now corrected for the batch effect from the donor ID. We then projected these embeddings with and without Harmony correction into a 2D space using UMAP (*umap-learn* package) using *correlation* as similarity, 2.0 as repulsion strength, and 0.01 as min distance. We clustered the UMAP spaces using HDBSCAN algorithm (python *hdbscan* package https://hdbscan.readthedocs.io/en/latest/). We called peaks on each individual cluster obtained by this procedure by agglomerating the reads of the cells according to their labels and used MACS2 with the setting described above. We merged using bedtools the peak list of each cluster into a final set of 129576 peaks. We performed the same workflow described above with the 129576 peaks as input features to create a *cells x features* binary matrix, convert into a lower dimensional embedding, and correct batch effects with Harmony. We constructed a K-Nearest Neighbor (KNN) similarity graph with K=50 and *similarity=cosine* using the scikit-learn *NearestNeighbors* class. Finally, we clustered the similarity graph with the Louvain clustering algorithm[137] with different resolutions R (1.0, 1.5, and 2.0) from the *python-louvain* package (https://github.com/taynaud/python-louvain), and selected R=1.5 (6 clusters) based on the sum(-log10(fisher p-values)) of the significant cluster features.

Single-cell motif enrichment analysis—We used chromVAR (https://greenleaflab.github.io/chromVAR/)[33] to compute motif enrichment at the single-cell level. We used the merged list of cluster peaks (center +-250pb), and a list of 870 of non-redundant reference motifs as input for chromVAR workflow. We identified differentially enriched motifs for each cluster using the following strategy: for each cluster and each motif, we computed a Rank Sum test between the chromVAR Z-score distributions within and outside the cluster. We applied a Benjamini-Hochberg FDR correction on the p-values.

Multi-modal correlation matrices—Using the 6 clusters C obtained with R=1.5 and the 129576 peaks P we transformed the *cells x peaks* sparse binary matrix into a *clusters x peaks* matrix by taking the average number of cells having a peak p accessible within a cluster c for each p from P and c from C. Each column and row are then scaled to have a norm equal to 1. Using the same strategy, we created a *donors x peaks* using the donor ID (four donors) of each cell as label. Finally, we created a *bulks x peaks* matrix by counting the number of reads overlapping each peak p from P for the 30 bulk ATAC-Seq datasets. We then scaled the matrices (each column has mean=0 and std=1) and computed the *clusters x bulk* and *donors x bulk* kernel similarity matrices using the Pearson correlation as metric. Finally, we plotted the similarity matrices using the *clustermap* function from the *seaborn* python library (https://seaborn.pydata.org/api.html).

Statistical analyses—Gene expression differences were calculated with DESeq2 with Benjamini-Hochberg multiple testing correction. Genes are considered differentially expressed at >2FC, p.adj < 0.05.

Data Visualization—Heatmaps were generated with the pheatmap packages in R and other plots were made with ggplots2 in R with colors reflecting the scores/expression values, including z-scores, as noted in each figure. Circo plots showing genes related to neurodevelopmental diseases and for NicheNet ligand-receptor analysis were generated using the R package Circlize[138] package in R. Violin plots for scRNA-seq data was produced using Vlnplot function from Seurat package. Bar charts were generated using Prism 9.0 and presented as mean ± SEM; statistics were calculated using one-way ANOVA. Browser images were generated from the UCSC Genome Browser. Network plots were generated using python package *networkx* and *matplotlib*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the UC San Diego Center for Perinatal Discovery for providing fetal brain samples for this study. We thank Clare Pridans and David Hume for their generous gift of the Csf1rΔFIRE/ΔFIRE mice. We thank David Hume, Alexi Nott and Marten Hoekesema for critical reading of the manuscript. We thank Uri Manor for imaging assistance. C.Z.H. is supported by the Cancer Research Institute Irvington Postdoctoral Fellowship Program and NIH K99 MH129983. N.G.C. and C.K.G. are supported by NIH R01 NS096170 and the Chan Zuckerberg Initiative. C.K.G. is supported by the JPB foundation grant KR29574. J.F.H., D.P., N.G.C., and C.K.G. are supported by the Simons Foundation Autism Research Initiative (SFARI). N.G.C. is supported by NIH grants K08 NS109200, R01 NS124637, NS126452, the Doris Duke foundation and the Hartwell Foundation. M.B.J is supported by NIH RF1-DA048813; M.B.J. and C.K.G. are supported by the Cure Alzheimer's Fund, and Alzheimer's Association ADSF-21-829655-C. E.H. and A.J. were supported by the California Institute for Regenerative Medicine EDUC2–08388. The Flow Cytometry Core Facility of the Salk Institute is partly supported by NIH-NCI CCSG: P30 014195 and SIG S10-OD023689. Work at the Center for Epigenomics was supported in part by the UC San Diego School of Medicine. This publication includes data generated at the UC San Diego IGM Genomics Center Illumina NovaSeq 6000 purchased with NIH SIG grant (S10 OD026929). I.A.G. was supported by NICHD grant 5R24HD000836. All figure schematics were made with BioRender.

REFERENCES

1. Harry GJ (2013). Microglia during development and aging. Pharmacol Ther 139, 313–326. 10.1016/j.pharmthera.2013.04.013. [PubMed: 23644076]
2. Menassa DA, and Gomez-Nicola D (2018). Microglial Dynamics During Human Brain Development. Front Immunol 9, 1014. 10.3389/fimmu.2018.01014. [PubMed: 29881376]
3. Menassa DA, Muntslag TAO, Martin-Estebane M, Barry-Carroll L, Chapman MA, Adorjan I, Tyler T, Turnbull B, Rose-Zerilli MJJ, Nicoll JAR, et al. (2022). The spatiotemporal dynamics of microglia across the human lifespan. Dev Cell 57, 2127–2139 e2126. 10.1016/j.devcel.2022.07.015. [PubMed: 35977545]
4. Squarzoni P, Oller G, Hoeffel G, Pont-Lezica L, Rostaing P, Low D, Bessis A, Ginhoux F, and Garel S (2014). Microglia modulate wiring of the embryonic forebrain. Cell Rep 8, 1271–1279. 10.1016/j.celrep.2014.07.042. [PubMed: 25159150]

5. Dai XM, Ryan GR, Hapel AJ, Dominguez MG, Russell RG, Kapp S, Sylvestre V, and Stanley ER (2002). Targeted disruption of the mouse colony-stimulating factor 1 receptor gene results in osteopetrosis, mononuclear phagocyte deficiency, increased primitive progenitor cell frequencies, and reproductive defects. Blood 99, 111–120. 10.1182/blood.v99.1.111. [PubMed: 11756160]
6. Erblich B, Zhu L, Etgen AM, Dobrenis K, and Pollard JW (2011). Absence of colony stimulation factor-1 receptor results in loss of microglia, disrupted brain development and olfactory deficits. PLoS One 6, e26317. 10.1371/journal.pone.0026317. [PubMed: 22046273]
7. De S, Van Deren D, Peden E, Hockin M, Boulet A, Titen S, and Capecchi MR (2018). Two distinct ontogenies confer heterogeneity to mouse brain microglia. Development 145. 10.1242/dev.152306.
8. Harrington AJ, Bridges CM, Berto S, Blankenship K, Cho JY, Assali A, Siemsen BM, Moore HW, Tsvetkov E, Thielking A, et al. (2020). MEF2C Hypofunction in Neuronal and Neuroimmune Populations Produces MEF2C Haploinsufficiency Syndrome-like Behaviors in Mice. Biol Psychiatry 88, 488–499. 10.1016/j.biopsych.2020.03.011. [PubMed: 32418612]
9. Pinto B, Morelli G, Rastogi M, Savardi A, Fumagalli A, Petretto A, Bartolucci M, Varea E, Catelani T, Contestabile A, et al. (2020). Rescuing Over-activated Microglia Restores Cognitive Performance in Juvenile Animals of the Dp(16) Mouse Model of Down Syndrome. Neuron 108, 887–904 e812. 10.1016/j.neuron.2020.09.010. [PubMed: 33027640]
10. Sahin M, and Sur M (2015). Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. Science 350. 10.1126/science.aab3897.
11. Green KN, Crapser JD, and Hohsfield LA (2020). To Kill a Microglia: A Case for CSF1R Inhibitors. Trends Immunol 41, 771–784. 10.1016/j.it.2020.07.001. [PubMed: 32792173]
12. Gosselin D, Link VM, Romanoski CE, Fonseca GJ, Eichenfield DZ, Spann NJ, Stender JD, Chun HB, Garner H, Geissmann F, and Glass CK (2014). Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. Cell 159, 1327–1340. 10.1016/j.cell.2014.11.023. [PubMed: 25480297]
13. Gosselin D, Skola D, Coufal NG, Holtman IR, Schlachetzki JCM, Sajti E, Jaeger BN, O'Connor C, Fitzpatrick C, Pasillas MP, et al. (2017). An environment-dependent transcriptional network specifies human microglia identity. Science 356. 10.1126/science.aal3222.
14. Cao J, O'Day DR, Pliner HA, Kingsley PD, Deng M, Daza RM, Zager MA, Aldinger KA, Blecher-Gonen R, Zhang F, et al. (2020). A human cell atlas of fetal gene expression. Science 370. 10.1126/science.aba7721.
15. Kracht L, Borggrewe M, Eskandar S, Brouwer N, Chuva de Sousa Lopes SM, Laman JD, Scherjon SA, Prins JR, Kooistra SM, and Eggen BJL (2020). Human fetal microglia acquire homeostatic immune-sensing properties early in development. Science 369, 530–537. 10.1126/science.aba5906. [PubMed: 32732419]
16. Langfelder P, and Horvath S (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559. 10.1186/1471-2105-9-559. [PubMed: 19114008]
17. Li Q, Cheng Z, Zhou L, Darmanis S, Neff NF, Okamoto J, Gulati G, Bennett ML, Sun LO, Clarke LE, et al. (2019). Developmental Heterogeneity of Microglia and Brain Myeloid Cells Revealed by Deep Single-Cell RNA Sequencing. Neuron 101, 207–223 e210. 10.1016/j.neuron.2018.12.006. [PubMed: 30606613]
18. Gu H, Yang X, Mao X, Xu E, Qi C, Wang H, Brahmachari S, York B, Sriparna M, Li A, et al. (2021). Lymphocyte Activation Gene 3 (Lag3) Contributes to alpha-Synucleinopathy in alpha-Synuclein Transgenic Mice. Front Cell Neurosci 15, 656426. 10.3389/fncel.2021.656426. [PubMed: 33776654]
19. Vincenti JE, Murphy L, Grabert K, McColl BW, Cancellotti E, Freeman TC, and Manson JC (2015). Defining the Microglia Response during the Time Course of Chronic Neurodegeneration. J Virol 90, 3003–3017. 10.1128/JVI.02613-15. [PubMed: 26719249]
20. Thion MS, Low D, Silvin A, Chen J, Grisel P, Schulte-Schrepping J, Blecher R, Ulas T, Squarzoni P, Hoeffel G, et al. (2018). Microbiome Influences Prenatal and Adult Microglia in a Sex-Specific Manner. Cell 172, 500–516 e516. 10.1016/j.cell.2017.11.042. [PubMed: 29275859]
21. Werling DM, and Geschwind DH (2013). Sex differences in autism spectrum disorders. Curr Opin Neurol 26, 146–153. 10.1097/WCO.0b013e32835ee548. [PubMed: 23406909]

22. Mazure CM, and Swendsen J (2016). Sex differences in Alzheimer's disease and other dementias. Lancet Neurol 15, 451–452. 10.1016/S1474-4422(16)00067-3. [PubMed: 26987699]
23. Bonnardel J, T'Jonck W, Gaublomme D, Browaeys R, Scott CL, Martens L, Vanneste B, De Prijck S, Nedospasov SA, Kremer A, et al. (2019). Stellate Cells, Hepatocytes, and Endothelial Cells Imprint the Kupffer Cell Identity on Monocytes Colonizing the Liver Macrophage Niche. Immunity 51, 638–654 e639. 10.1016/j.immuni.2019.08.017. [PubMed: 31561945]
24. Browaeys R, Saelens W, and Saeys Y (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods 17, 159–162. 10.1038/s41592-019-0667-5. [PubMed: 31819264]
25. Nowakowski TJ, Bhaduri A, Pollen AA, Alvarado B, Mostajo-Radji MA, Di Lullo E, Haeussler M, Sandoval-Espinosa C, Liu SJ, Velmeshev D, et al. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science 358, 1318–1323. 10.1126/science.aap8809. [PubMed: 29217575]
26. Reemst K, Noctor SC, Lucassen PJ, and Hol EM (2016). The Indispensable Roles of Microglia and Astrocytes during Brain Development. Front Hum Neurosci 10, 566. 10.3389/fnhum.2016.00566. [PubMed: 27877121]
27. Rosin JM, Marsters CM, Malik F, Far R, Adnani L, Schuurmans C, Pittman QJ, and Kurrasch DM (2021). Embryonic Microglia Interact with Hypothalamic Radial Glia during Development and Upregulate the TAM Receptors MERTK and AXL following an Insult. Cell Rep 34, 108587. 10.1016/j.celrep.2020.108587. [PubMed: 33406432]
28. Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, Close JL, Long B, Johansen N, Penn O, et al. (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature 573, 61–68. 10.1038/s41586-019-1506-7. [PubMed: 31435019]
29. Tasic B, Yao Z, Graybuck LT, Smith KA, Nguyen TN, Bertagnolli D, Goldy J, Garren E, Economo MN, Viswanathan S, et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature 563, 72–78. 10.1038/s41586-018-0654-5. [PubMed: 30382198]
30. Yamazaki Y, Zhao N, Caulfield TR, Liu CC, and Bu G (2019). Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. Nat Rev Neurol 15, 501–518. 10.1038/s41582-019-0228-7. [PubMed: 31367008]
31. Heinz S, Romanoski CE, Benner C, and Glass CK (2015). The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol 16, 144–154. 10.1038/nrm3949. [PubMed: 25650801]
32. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A 107, 21931–21936. 10.1073/pnas.1016071107. [PubMed: 21106759]
33. Schep AN, Wu B, Buenrostro JD, and Greenleaf WJ (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods 14, 975–978. 10.1038/nmeth.4401. [PubMed: 28825706]
34. Hoeksema MA, Shen Z, Holtman IR, Zheng A, Spann NJ, Cobo I, Gymrek M, and Glass CK (2021). Mechanisms underlying divergent responses of genetically distinct macrophages to IL-4. Sci Adv 7. 10.1126/sciadv.abf9808.
35. Goenka S, and Kaplan MH (2011). Transcriptional regulation by STAT6. Immunol Res 50, 87–96. 10.1007/s12026-011-8205-2. [PubMed: 21442426]
36. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, Curina A, Prosperini E, Ghisletti S, and Natoli G (2013). Latent enhancers activated by stimulation in differentiated cells. Cell 152, 157–171. 10.1016/j.cell.2012.12.018. [PubMed: 23332752]
37. Huang JT, Welch JS, Ricote M, Binder CJ, Willson TM, Kelly C, Witztum JL, Funk CD, Conrad D, and Glass CK (1999). Interleukin-4-dependent production of PPAR-gamma ligands in macrophages by 12/15-lipoxygenase. Nature 400, 378–382. 10.1038/22572. [PubMed: 10432118]
38. Daniel B, Nagy G, Czimmerer Z, Horvath A, Hammers DW, Cuaranta-Monroy I, Poliska S, Tzerpos P, Kolostyak Z, Hays TT, et al. (2018). The Nuclear Receptor PPARgamma Controls Progressive Macrophage Polarization as a Ligand-Insensitive Epigenomic Ratchet of Transcriptional Memory. Immunity 49, 615–626 e616. 10.1016/j.immuni.2018.09.005. [PubMed: 30332629]

39. Daniel B, Czimmerer Z, Halasz L, Boto P, Kolostyak Z, Poliska S, Berger WK, Tzerpos P, Nagy G, Horvath A, et al. (2020). The transcription factor EGR2 is the molecular linchpin connecting STAT6 activation to the late, stable epigenomic program of alternative macrophage polarization. Genes Dev 34, 1474–1492. 10.1101/gad.343038.120. [PubMed: 33060136]

40. Perera RM, Di Malta C, and Ballabio A (2019). MiT/TFE Family of Transcription Factors, Lysosomes, and Cancer. Annu Rev Cancer Biol 3, 203–222. 10.1146/annurev-cancerbio-030518-055835. [PubMed: 31650096]

41. Nakao A, Imamura T, Souchelnytskyi S, Kawabata M, Ishisaki A, Oeda E, Tamaki K, Hanai J, Heldin CH, Miyazono K, and ten Dijke P (1997). TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. EMBO J 16, 5353–5362. 10.1093/emboj/16.17.5353. [PubMed: 9311995]

42. Jia J, Claude-Taupin A, Gu Y, Choi SW, Peters R, Bissa B, Mudd MH, Allers L, Pallikkuth S, Lidke KA, et al. (2020). Galectin-3 Coordinates a Cellular System for Lysosomal Repair and Removal. Dev Cell 52, 69–87 e68. 10.1016/j.devcel.2019.10.025. [PubMed: 31813797]

43. Leng L, Metz CN, Fang Y, Xu J, Donnelly S, Baugh J, Delohery T, Chen Y, Mitchell RA, and Bucala R (2003). MIF signal transduction initiated by binding to CD74. J Exp Med 197, 1467–1476. 10.1084/jem.20030286. [PubMed: 12782713]

44. Faure-Andre G, Vargas P, Yuseff MI, Heuze M, Diaz J, Lankar D, Steri V, Manry J, Hugues S, Vascotto F, et al. (2008). Regulation of dendritic cell migration by CD74, the MHC class II-associated invariant chain. Science 322, 1705–1710. 10.1126/science.1159894. [PubMed: 19074353]

45. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, and Raychaudhuri S (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296. 10.1038/s41592-019-0619-0. [PubMed: 31740819]

46. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lonnerberg P, Furlan A, et al. (2018). RNA velocity of single cells. Nature 560, 494–498. 10.1038/s41586-018-0414-6. [PubMed: 30089906]

47. Jordao MJC, Sankowski R, Brendecke SM, Sagar Locatelli G., Tai YH, Tay TL, Schramm E, Armbruster S, Hagemeyer N, et al. (2019). Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. Science 363. 10.1126/science.aat7554.

48. Mrdjen D, Pavlovic A, Hartmann FJ, Schreiner B, Utz SG, Leung BP, Lelios I, Heppner FL, Kipnis J, Merkler D, et al. (2018). High-Dimensional Single-Cell Mapping of Central Nervous System Immune Cells Reveals Distinct Myeloid Subsets in Health, Aging, and Disease. Immunity 48, 380–395 e386. 10.1016/j.immuni.2018.01.011. [PubMed: 29426702]

49. Utz SG, See P, Mildenberger W, Thion MS, Silvin A, Lutz M, Ingelfinger F, Rayan NA, Lelios I, Buttgereit A, et al. (2020). Early Fate Defines Microglia and Non-parenchymal Brain Macrophage Development. Cell 181, 557–573 e518. 10.1016/j.cell.2020.03.021. [PubMed: 32259484]

50. Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, Marques S, Munguba H, He L, Betsholtz C, et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347, 1138–1142. 10.1126/science.aaa1934. [PubMed: 25700174]

51. Hammond TR, Dufort C, Dissing-Olesen L, Giera S, Young A, Wysoker A, Walker AJ, 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 e256. 10.1016/j.immuni.2018.11.004. [PubMed: 30471926]

52. Saclier M, Lapi M, Bonfanti C, Rossi G, Antonini S, and Messina G (2020). The Transcription Factor Nfix Requires RhoA-ROCK1 Dependent Phagocytosis to Mediate Macrophage Skewing during Skeletal Muscle Regeneration. Cells 9. 10.3390/cells9030708.

53. Duren Z, Chen X, Jiang R, Wang Y, and Wong WH (2017). Modeling gene regulation from paired expression and chromatin accessibility data. Proc Natl Acad Sci U S A 114, E4914–E4923. 10.1073/pnas.1704553114. [PubMed: 28576882]

54. Kamimoto K, Stringa B, Hoffmann CM, Jindal K, Solnica-Krezel L, and Morris SA (2023). Dissecting cell identity via network inference and in silico gene perturbation. Nature 614, 742–751. 10.1038/s41586-022-05688-9. [PubMed: 36755098]

55. Seidman JS, Troutman TD, Sakai M, Gola A, Spann NJ, Bennett H, Bruni CM, Ouyang Z, Li RZ, Sun X, et al. (2020). Niche-Specific Reprogramming of Epigenetic Landscapes Drives Myeloid Cell Diversity in Nonalcoholic Steatohepatitis. Immunity 52, 1057–1074 e1057. 10.1016/j.immuni.2020.04.001. [PubMed: 32362324]

56. Sakai M, Troutman TD, Seidman JS, Ouyang Z, Spann NJ, Abe Y, Ego KM, Bruni CM, Deng Z, Schlachetzki JCM, et al. (2019). Liver-Derived Signals Sequentially Reprogram Myeloid Enhancers to Initiate and Maintain Kupffer Cell Identity. Immunity 51, 655–670 e658. 10.1016/j.immuni.2019.09.002. [PubMed: 31587991]

57. Deczkowska A, Matcovitch-Natan O, Tsitsou-Kampeli A, Ben-Hamo S, Dvir-Szternfeld R, Spinrad A, Singer O, David E, Winter DR, Smith LK, et al. (2017). Mef2C restrains microglial inflammatory response and is lost in brain ageing in an IFN-I-dependent manner. Nat Commun 8, 717. 10.1038/s41467-017-00769-0. [PubMed: 28959042]

58. Pitale PM, Gorbatyuk O, and Gorbatyuk M (2017). Neurodegeneration: Keeping ATF4 on a Tight Leash. Front Cell Neurosci 11, 410. 10.3389/fncel.2017.00410. [PubMed: 29326555]

59. Hasselmann J, Coburn MA, England W, Figueroa Velez DX, Kiani Shabestari S, Tu CH, McQuade A, Kolahdouzan M, Echeverria K, Claes C, et al. (2019). Development of a Chimeric Model to Study and Manipulate Human Microglia In Vivo. Neuron 103, 1016–1033 e1010. 10.1016/j.neuron.2019.07.002. [PubMed: 31375314]

60. Mancuso R, Van Den Daele J, Fattorelli N, Wolfs L, Balusu S, Burton O, Liston A, Sierksma A, Fourne Y, Poovathingal S, et al. (2019). Stem-cell-derived human microglia transplanted in mouse brain to study human disease. Nat Neurosci 22, 2111–2116. 10.1038/s41593-019-0525-x. [PubMed: 31659342]

61. Kiani Shabestari S, Morabito S, Danhash EP, McQuade A, Sanchez JR, Miyoshi E, Chadarevian JP, Claes C, Coburn MA, Hasselmann J, et al. (2022). Absence of microglia promotes diverse pathologies and early lethality in Alzheimer's disease mice. Cell Rep 39, 110961. 10.1016/j.celrep.2022.110961. [PubMed: 35705056]

62. Rojo R, Raper A, Ozdemir DD, Lefevre L, Grabert K, Wollscheid-Lengeling E, Bradford B, Caruso M, Gazova I, Sanchez A, et al. (2019). Deletion of a Csf1r enhancer selectively impacts CSF1R expression and development of tissue macrophage populations. Nat Commun 10, 3215. 10.1038/s41467-019-11053-8. [PubMed: 31324781]

63. Li WV, Chen Y, and Li JJ (2017). TROM: A Testing-Based Method for Finding Transcriptomic Similarity of Biological Samples. Stat Biosci 9, 105–136. 10.1007/s12561-016-9163-y. [PubMed: 28781712]

64. Wilkinson B, Evgrafov OV, Zheng D, Hartel N, Knowles JA, Graham NA, Ichida JK, and Coba MP (2019). Endogenous Cell Type-Specific Disrupted in Schizophrenia 1 Interactomes Reveal Protein Networks Associated With Neurodevelopmental Disorders. Biol Psychiatry 85, 305–316. 10.1016/j.biopsych.2018.05.009. [PubMed: 29961565]

65. Lalani SR, Zhang J, Schaaf CP, Brown CW, Magoulas P, Tsai AC, El-Gharbawy A, Wierenga KJ, Bartholomew D, Fong CT, et al. (2014). Mutations in PURA cause profound neonatal hypotonia, seizures, and encephalopathy in 5q31.3 microdeletion syndrome. Am J Hum Genet 95, 579–583. 10.1016/j.ajhg.2014.09.014. [PubMed: 25439098]

66. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson N, Daly MJ, Price AL, and Neale BM (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47, 291–295. 10.1038/ng.3211. [PubMed: 25642630]

67. Madison BB (2016). Srebp2: A master regulator of sterol and fatty acid synthesis. J Lipid Res 57, 333–335. 10.1194/jlr.C066712. [PubMed: 26798145]

68. Bian Z, Gong Y, Huang T, Lee CZW, Bian L, Bai Z, Shi H, Zeng Y, Liu C, He J, et al. (2020). Deciphering human macrophage development at single-cell resolution. Nature 582, 571–576. 10.1038/s41586-020-2316-7. [PubMed: 32499656]

69. Balwierz PJ, Pachkov M, Arnold P, Gruber AJ, Zavolan M, and van Nimwegen E (2014). ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. Genome Res 24, 869–884. 10.1101/gr.169508.113. [PubMed: 24515121]

70. Consortium F, Suzuki H, Forrest AR, van Nimwegen E, Daub CO, Balwierz PJ, Irvine KM, Lassmann T, Ravasi T, Hasegawa Y, et al. (2009). The transcriptional network that controls growth

arrest and differentiation in a human myeloid leukemia cell line. Nat Genet 41, 553–562. 10.1038/ng.375. [PubMed: 19377474]

71. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods 14, 1083–1086. 10.1038/nmeth.4463. [PubMed: 28991892]
72. Nott A, Holtman IR, Coufal NG, Schlachetzki JCM, Yu M, Hu R, Han CZ, Pena M, Xiao J, Wu Y, et al. (2019). Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. Science 366, 1134–1139. 10.1126/science.aay0793. [PubMed: 31727856]
73. Kierdorf K, Erny D, Goldmann T, Sander V, Schulz C, Perdiguero EG, Wieghofer P, Heinrich A, Riemke P, Holscher C, et al. (2013). Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. Nat Neurosci 16, 273–280. 10.1038/nn.3318. [PubMed: 23334579]
74. Yamashita K, Sato A, Asashima M, Wang PC, and Nishinakamura R (2007). Mouse homolog of SALL1, a causative gene for Townes-Brocks syndrome, binds to A/T-rich sequences in pericentric heterochromatin via its C-terminal zinc finger domains. Genes Cells 12, 171–182. 10.1111/j.1365-2443.2007.01042.x. [PubMed: 17295837]
75. Kong NR, Bassal MA, Tan HK, Kurland JV, Yong KJ, Young JJ, Yang Y, Li F, Lee JD, Liu Y, et al. (2021). Zinc Finger Protein SALL4 Functions through an AT-Rich Motif to Regulate Gene Expression. Cell Rep 34, 108574. 10.1016/j.celrep.2020.108574. [PubMed: 33406418]
76. Hume DA, Summers KM, Raza S, Baillie JK, and Freeman TC (2010). Functional clustering and lineage markers: insights into cellular differentiation and gene function from large-scale microarray studies of purified primary cell populations. Genomics 95, 328–338. 10.1016/j.ygeno.2010.03.002. [PubMed: 20211243]
77. Summers KM, Bush SJ, and Hume DA (2020). Network analysis of transcriptomic diversity amongst resident tissue macrophages and dendritic cells in the mouse mononuclear phagocyte system. PLoS Biol 18, e3000859. 10.1371/journal.pbio.3000859. [PubMed: 33031383]
78. Cunningham CL, Martinez-Cerdeno V, and Noctor SC (2013). Microglia regulate the number of neural precursor cells in the developing cerebral cortex. J Neurosci 33, 4216–4233. 10.1523/JNEUROSCI.3441-12.2013. [PubMed: 23467340]
79. Michell-Robinson MA, Touil H, Healy LM, Owen DR, Durafourt BA, Bar-Or A, Antel JP, and Moore CS (2015). Roles of microglia in brain development, tissue maintenance and repair. Brain 138, 1138–1159. 10.1093/brain/awv066. [PubMed: 25823474]
80. Soldner F, and Jaenisch R (2012). Medicine. iPSC disease modeling. Science 338, 1155–1156. 10.1126/science.1227682. [PubMed: 23197518]
81. Yamanaka S (2020). Pluripotent Stem Cell-Based Cell Therapy-Promise and Challenges. Cell Stem Cell 27, 523–531. 10.1016/j.stem.2020.09.014. [PubMed: 33007237]
82. Moore JE, Pratt HE, Purcaro MJ, and Weng Z (2020). A curated benchmark of enhancer-gene interactions for evaluating enhancer-target gene prediction methods. Genome Biol 21, 17. 10.1186/s13059-019-1924-8. [PubMed: 31969180]
83. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, and Jones JM (1998). Embryonic stem cell lines derived from human blastocysts. Science 282, 1145–1147. 10.1126/science.282.5391.1145. [PubMed: 9804556]
84. McQuade A, Coburn M, Tu CH, Hasselmann J, Davtyan H, and Blurton-Jones M (2018). Development and validation of a simplified method to generate human microglia from pluripotent stem cells. Mol Neurodegener 13, 67. 10.1186/s13024-018-0297-x. [PubMed: 30577865]
85. Qian X, Nguyen HN, Song MM, Hadiono C, Ogden SC, Hammack C, Yao B, Hamersky GR, Jacob F, Zhong C, et al. (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. Cell 165, 1238–1254. 10.1016/j.cell.2016.04.032. [PubMed: 27118425]
86. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589. 10.1016/j.molcel.2010.05.004. [PubMed: 20513432]

87. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]

88. Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. 10.1038/nmeth.1923. [PubMed: 22388286]

89. Matcovitch-Natan O, Winter DR, Giladi A, Vargas Aguilar S, Spinrad A, Sarrazin S, Ben-Yehuda H, David E, Zelada Gonzalez F, Perrin P, et al. (2016). Microglia development follows a stepwise program to regulate brain homeostasis. Science 353, aad8670. 10.1126/science.aad8670. [PubMed: 27338705]

90. Bian S, Repic M, Guo Z, Kavirayani A, Burkard T, Bagley JA, Krauditsch C, and Knoblich JA (2018). Genetically engineered cerebral organoids model brain tumor formation. Nat Methods 15, 631–639. 10.1038/s41592-018-0070-7. [PubMed: 30038414]

91. Leinonen R, Sugawara H, Shumway M, and International Nucleotide Sequence Database, C. (2011). The sequence read archive. Nucleic Acids Res 39, D19–21. 10.1093/nar/gkq1019. [PubMed: 21062823]

92. Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]

93. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, and Chanda SK (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10, 1523. 10.1038/s41467-019-09234-6. [PubMed: 30944313]

94. Kniffin CL, and McKusick VA (2013). OMIM: 143465 Attention Deficit-Hyperactive Disorder; ADHD. https://www.omim.org/entry/143465.

95. Faraone SV, and Mick E (2010). Molecular genetics of attention deficit hyperactivity disorder. Psychiatr Clin North Am 33, 159–180. 10.1016/j.psc.2009.12.004. [PubMed: 20159345]

96. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L, et al. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell 180, 568–584 e523. 10.1016/j.cell.2019.12.036. [PubMed: 31981491]

97. Kniffin CL, and McKusick VA (2019). OMIM: 308350 Developmental and Epileptic Encephalopathy 1; DEE1. https://www.omim.org/entry/308350.

98. Centanni TM (2020). Neural and Genetic Mechanisms of Dyslexia In (Springer International Publishing), pp. 47–68. 10.1007/978-3-030-35687-3_4.

99. Luciano M, Gow AJ, Pattie A, Bates TC, and Deary IJ (2018). The Influence of Dyslexia Candidate Genes on Reading Skill in Old Age. Behav Genet 48, 351–360. 10.1007/s10519-018-9913-3. [PubMed: 29959602]

100. Mascheretti S, De Luca A, Trezzi V, Peruzzo D, Nordio A, Marino C, and Arrigoni F (2017). Neurogenetics of developmental dyslexia: from genes to behavior through brain neuroimaging and cognitive and sensorial mechanisms. Transl Psychiatry 7, e987. 10.1038/tp.2016.240. [PubMed: 28045463]

101. Nopola-Hemmi J, Myllyluoma B, Haltia T, Taipale M, Ollikainen V, Ahonen T, Voutilainen A, Kere J, and Widen E (2001). A dominant gene for developmental dyslexia on chromosome 3. J Med Genet 38, 658–664. 10.1136/jmg.38.10.658. [PubMed: 11584043]

102. Deciphering Developmental Disorders S (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature 519, 223–228. 10.1038/nature14135. [PubMed: 25533962]

103. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature 511, 344–347. 10.1038/nature13394. [PubMed: 24896178]

104. de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, et al. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med 367, 1921–1929. 10.1056/NEJMoa1206524. [PubMed: 23033978]

105. Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, et al. (2012). Range of genetic mutations associated with severe

non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet 380, 1674–1682. 10.1016/S0140-6736(12)61480-9. [PubMed: 23020937]

106. Farrell MS, Werge T, Sklar P, Owen MJ, Ophoff RA, O'Donovan MC, Corvin A, Cichon S, and Sullivan PF (2015). Evaluating historical candidate genes for schizophrenia. Mol Psychiatry 20, 555–562. 10.1038/mp.2015.16. [PubMed: 25754081]
107. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, Georgieva L, Rees E, Palta P, Ruderfer DM, et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature 506, 179–184. 10.1038/nature12929. [PubMed: 24463507]
108. Gulsuner S, Walsh T, Watts AC, Lee MK, Thornton AM, Casadei S, Rippey C, Shahin H, Consortium on the Genetics of, S., Group, P.S., et al. (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell 154, 518–529. 10.1016/j.cell.2013.06.049. [PubMed: 23911319]
109. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, O'Dushlaine C, Chambert K, Bergen SE, Kahler A, et al. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506, 185–190. 10.1038/nature12975. [PubMed: 24463508]
110. Schizophrenia Working Group of the Psychiatric Genomics, C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427. 10.1038/nature13595. [PubMed: 25056061]
111. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, Tooley K, Presumey J, Baum M, Van Doren V, et al. (2016). Schizophrenia risk from complex variation of complement component 4. Nature 530, 177–183. 10.1038/nature16549. [PubMed: 26814963]
112. Stenton SL, and Prokisch H (2020). Genetics of mitochondrial diseases: Identifying mutations to help diagnosis. EBioMedicine 56, 102784. 10.1016/j.ebiom.2020.102784. [PubMed: 32454403]
113. Brozzi A, Urbanelli L, Germain PL, Magini A, and Emiliani C (2013). hLGDB: a database of human lysosomal genes and their regulation. Database (Oxford) 2013, bat024. 10.1093/database/bat024. [PubMed: 23584836]
114. Greiner-Tollersrud OK, and Berg T (2000-2013). Lysosomal Storage Disorders. Madame Curie Bioscience Database Landes Bioscience.
115. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, Baldursson G, Belliveau R, Bybjerg-Grauholm J, Baekvad-Hansen M, et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 51, 63–75. 10.1038/s41588-018-0269-7. [PubMed: 30478444]
116. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, Pallesen J, Agerbo E, Andreassen OA, Anney R, et al. (2019). Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 51, 431–444. 10.1038/s41588-019-0344-8. [PubMed: 30804558]
117. International League Against Epilepsy Consortium on Complex, E. (2018). Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. Nat Commun 9, 5269. 10.1038/s41467-018-07524-z. [PubMed: 30531953]
118. Boraska V, Franklin CS, Floyd JA, Thornton LM, Huckins LM, Southam L, Rayner NW, Tachmazidou I, Klump KL, Treasure J, et al. (2014). A genome-wide association study of anorexia nervosa. Mol Psychiatry 19, 1085–1094. 10.1038/mp.2013.187. [PubMed: 24514567]
119. Otowa T, Hek K, Lee M, Byrne EM, Mirza SS, Nivard MG, Bigdeli T, Aggen SH, Adkins D, Wolen A, et al. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. Mol Psychiatry 21, 1391–1399. 10.1038/mp.2015.197. [PubMed: 26754954]
120. Psychiatric, G.C.B.D.W.G. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 43, 977–983. 10.1038/ng.943. [PubMed: 21926972]
121. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TMF, et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 50, 668–681. 10.1038/s41588-018-0090-3. [PubMed: 29700475]
122. Pardinas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, Legge SE, Bishop S, Cameron D, Hamshere ML, et al. (2018). Common schizophrenia alleles are enriched in

mutation-intolerant genes and in regions under strong background selection. Nat Genet 50, 381–389. 10.1038/s41588-018-0059-2. [PubMed: 29483656]

123. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, Hagenaars SP, Ritchie SJ, Marioni RE, Fawns-Ritchie C, et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun 9, 2098. 10.1038/s41467-018-04362-x. [PubMed: 29844566]

124. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, Nagel M, Awasthi S, Barr PB, Coleman JRI, et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet 50, 912–919. 10.1038/s41588-018-0152-6. [PubMed: 29942086]

125. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, Savage JE, Hammerschlag AR, Skene NG, Munoz-Manchado AB, et al. (2018). Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet 50, 920–927. 10.1038/s41588-018-0151-7. [PubMed: 29942085]

126. Zhao B, Luo T, Li T, Li Y, Zhang J, Shan Y, Wang X, Yang L, Zhou F, Zhu Z, et al. (2019). Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. Nat Genet 51, 1637–1644. 10.1038/s41588-019-0516-6. [PubMed: 31676860]

127. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hagg S, Athanasiu L, et al. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer’s disease risk. Nat Genet 51, 404–413. 10.1038/s41588-018-0311-9. [PubMed: 30617256]

128. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, van der Lee SJ, Amlie-Wolf A, et al. (2019). Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. Nat Genet 51, 414–430. 10.1038/s41588-019-0358-2. [PubMed: 30820047]

129. van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, van der Spek RA, Vosa U, de Jong S, Robinson MR, et al. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat Genet 48, 1043–1048. 10.1038/ng.3622. [PubMed: 27455348]

130. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, Tan M, Kia DA, Noyce AJ, Xue A, et al. (2019). Identification of novel risk loci, causal insights, and heritable risk for Parkinson’s disease: a meta-analysis of genome-wide association studies. Lancet Neurol 18, 1091–1102. 10.1016/S1474-4422(19)30320-5. [PubMed: 31701892]

131. International Multiple Sclerosis Genetics, C., Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, Cotsapas C, Shah TS, Spencer C, Booth D, et al. (2013). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 45, 1353–1360. 10.1038/ng.2770. [PubMed: 24076602]

132. Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420. 10.1038/nbt.4096. [PubMed: 29608179]

133. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. (2021). Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587 e3529. 10.1016/j.cell.2021.04.048. [PubMed: 34062119]

134. DeTomaso D, Jones MG, Subramaniam M, Ashuach T, Ye CJ, and Yosef N (2019). Functional interpretation of single cell similarity maps. Nat Commun 10, 4376. 10.1038/s41467-019-12235-0. [PubMed: 31558714]

135. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137. 10.1186/gb-2008-9-9-r137. [PubMed: 18798982]

136. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. (2011). Scikit-learn: Machine Learning in Python. J Mach Learn Res 12, 2825–2830.

137. Blondel VD, Guillaume JL, Lambiotte R, and Lefebvre E (2008). Fast unfolding of communities in large networks. J Stat Mech-Theory E. Artn P10008 10.1088/1742-5468/2008/10/P10008.

138. Gu Z, Gu L, Eils R, Schlesner M, and Brors B (2014). circlize Implements and enhances circular visualization in R. Bioinformatics 30, 2811–2812. 10.1093/bioinformatics/btu393. [PubMed: 24930139]

Highlights:

- Human microglia have unique developmental stage-dependent epigenomic landscapes.
- TIMON predicts transcription factor cooperativity in human microglia maturation.
- Transcriptomes and epigenomes inform interpretation of disease risk alleles.
- iPSC-derived microglia can model developmental stage-specific phenotypes.

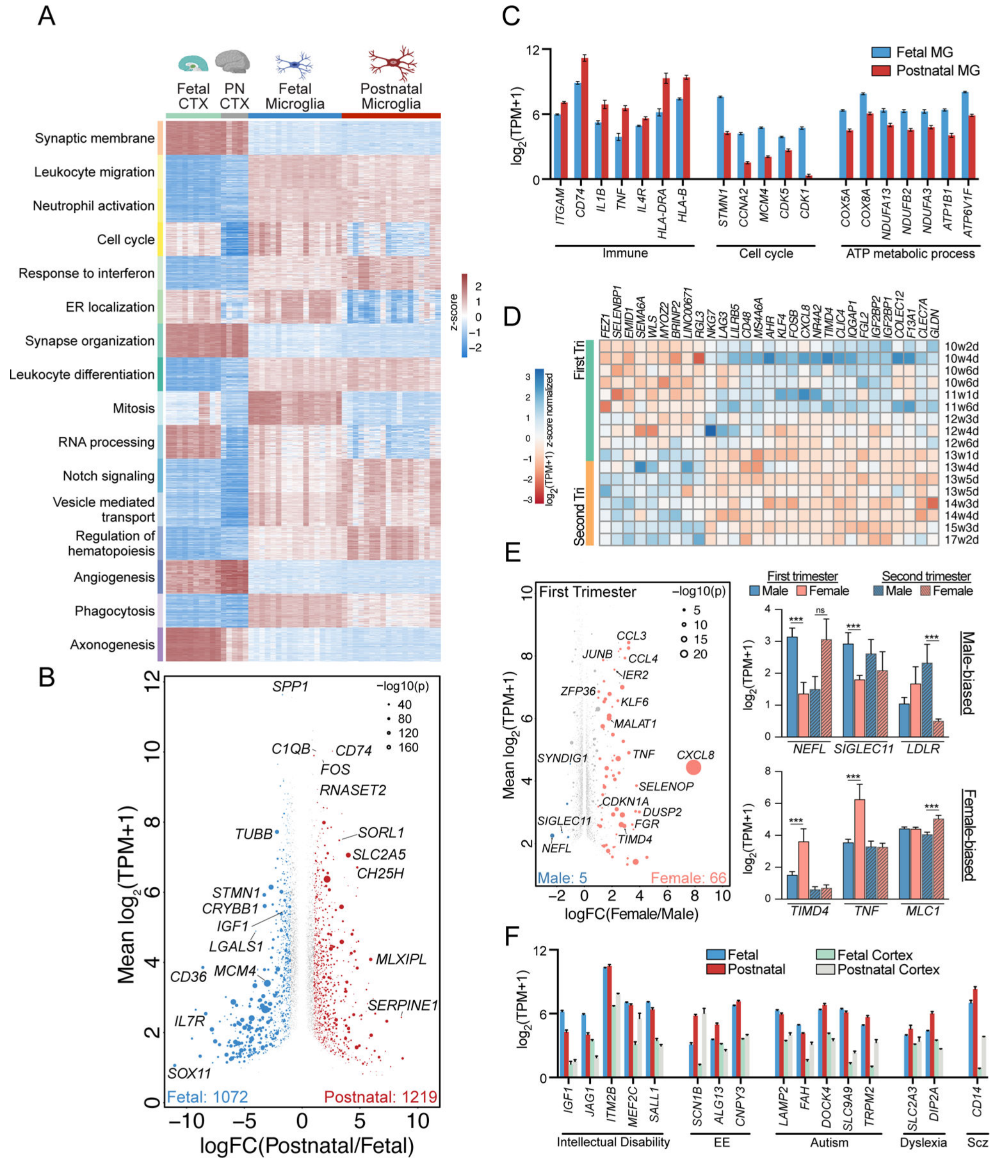


Figure 1. Changing transcriptome profile of human microglia during fetal development.

A. Heatmap expression z-scores of the top 100 genes from each significantly correlated module identified by WGCNA, ranked by Kleinberg's hub centrality scores. PN=postnatal, CTX=cortex

B. MA plot of gene expression between human fetal (DEG, blue) and postnatal microglia (DEG, red).

C. Bar charts of the expression of genes represented in GO term analysis of DEGs between fetal and postnatal microglia. Only genes with p-adj < 0.001 and FC > 2 are shown.

D. Heatmap expression z-scores of DEGs between first and second trimester fetal microglia.

E. MA plot of gene expression differences between male (DEG, blue) and female (DEG, orange) first trimester fetal microglia (left). Bar charts of expression of sex-biased autosomal genes in fetal microglia with male-bias (top right) and female bias (bottom right).

F. Bar charts of expression of NDD-associated genes in microglia and brain cortex. Genes with p-adj < 0.001 and FC > 2 are shown. EE=epileptic encephalopathy, Scz=Schizophrenia. Please also see Figure S1-2.

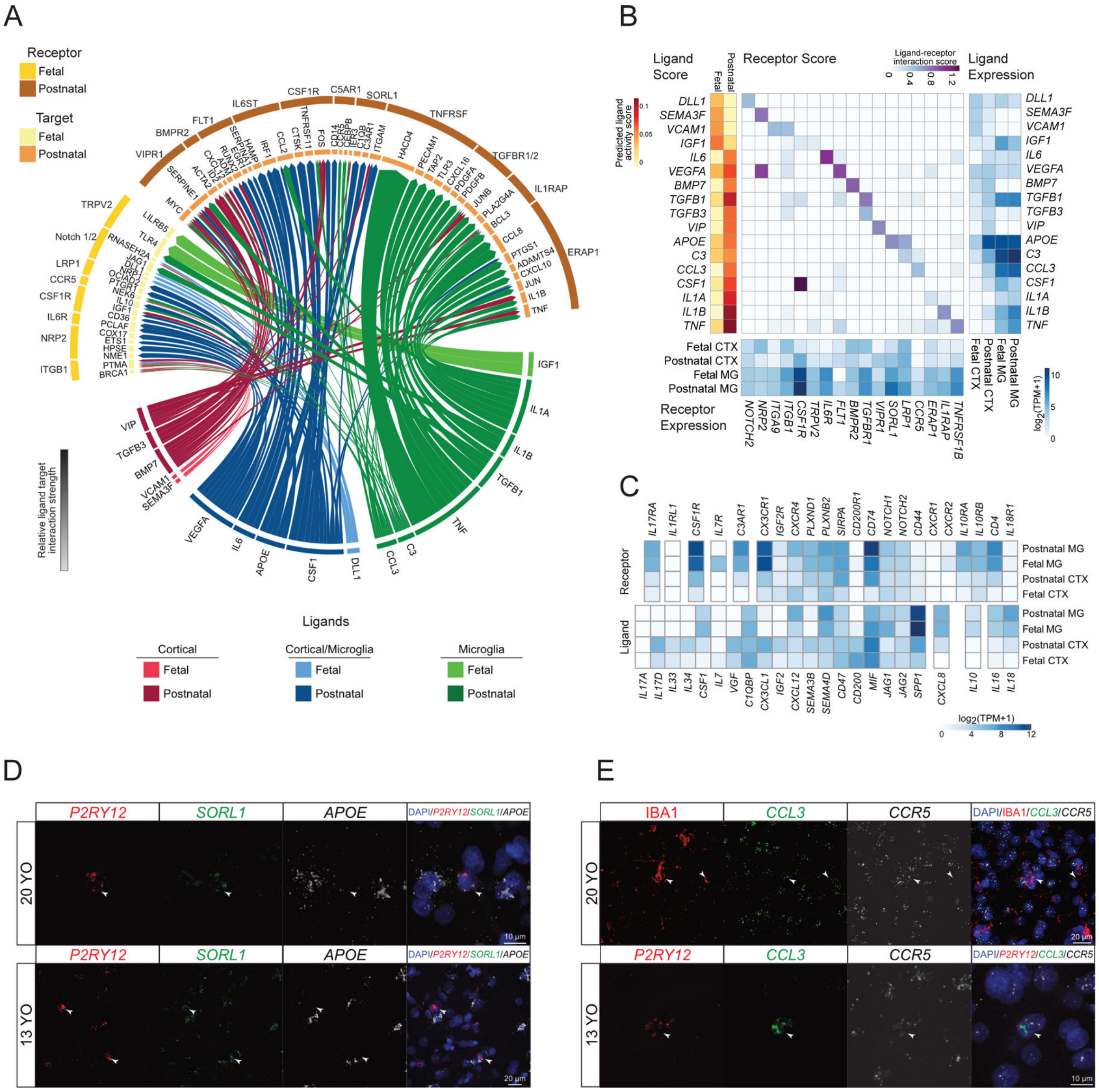


Figure 2. Putative ligand-receptor interactions influencing microglia development.

A. Circos plot indicating NicheNet prediction of ligand to target genes and predicted receptors in fetal and postnatal microglia. Please also see Figure S3.

B. Heatmaps depicting ligand activity score (left) and ligand-receptor interaction score (middle) represented in 2A and of RNA expression of ligands (right) and receptors (bottom).

C. Heatmap of gene expression of ligand (bottom)-receptor (top) pairs between human fetal bulk cortex, postnatal bulk cortex, fetal microglia and postnatal microglia.

D. Representative images of 3 independent experiments of multi-fluorescent RNAscope of human postnatal brains probed for *P2RY12* (red), *SORL1* (green), *APOE* (white); and DAPI. Arrowheads indicate microglia co-expressing *P2RY12* and *SORL1*.

E. Representative images of 3 independent experiments multi-fluorescent RNAscope of human postnatal brains probed for *P2RY12* (red) or IHC for IBA1 (red), *CCL3* (green), *CCR5* (white); and DAPI. Arrowheads indicate microglia co-expressing *P2RY12* or IBA1, *CCL3*, and *CCR5*.

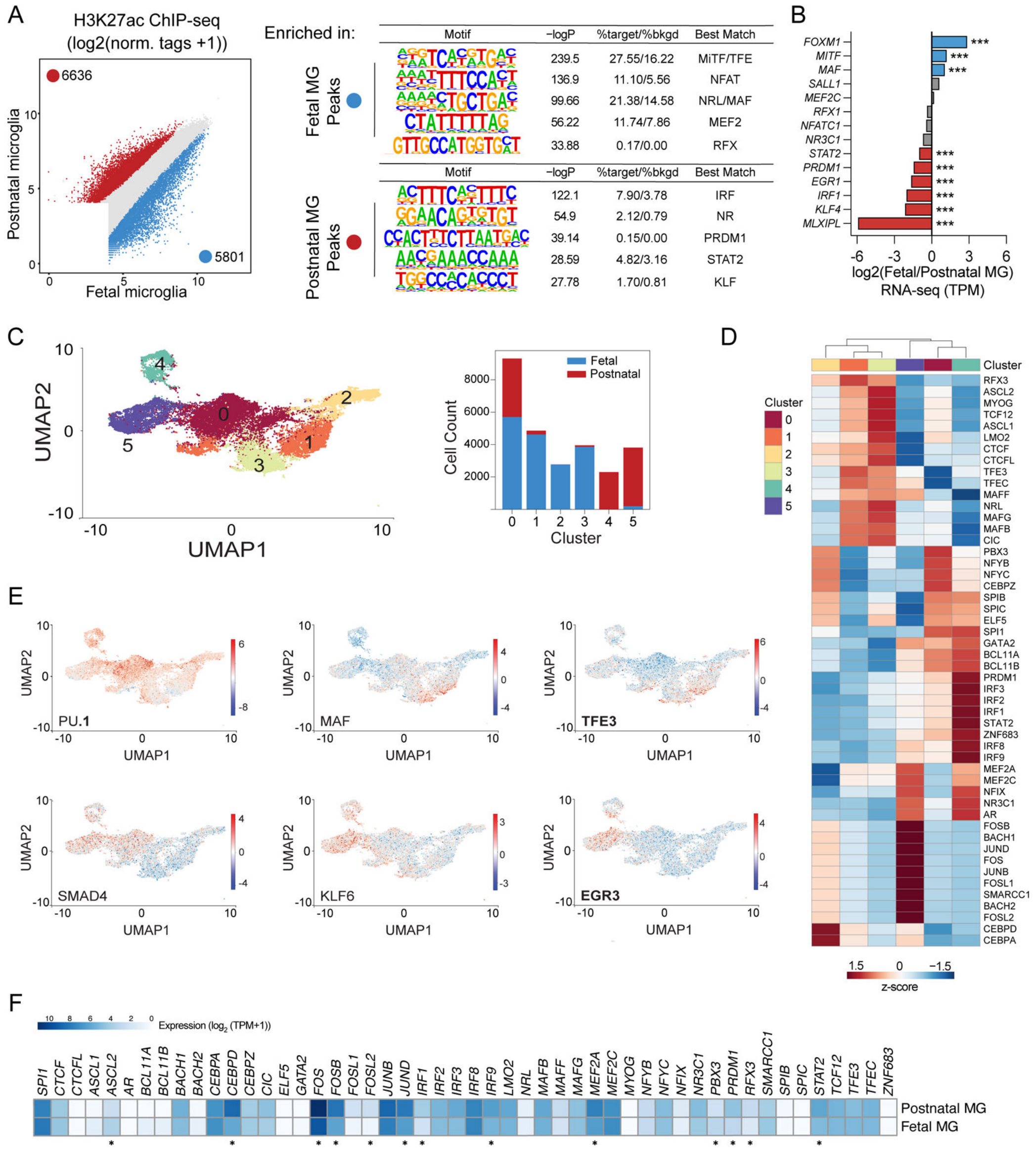


Figure 3. Maturation of human microglia remodels the active enhancer landscape.

A. (left) Scatter plot of H3K27ac ChIP-seq signal around distal (>1000bp from TSS) ATAC-seq peaks. Differentially acetylated regions (FC > 2, p-adj < 0.05) enriched in fetal (blue) and postnatal (red) microglia are colored. (right) *De novo* motifs analysis of differentially acetylated regions in fetal (top) or postnatal (bottom) microglia.

B. Bar chart showing log2FC of candidate TFs known to bind motifs identified in A, with TFs in blue higher expressed in fetal microglia and red higher expressed in postnatal microglia. *** indicates FC > 2, p-adj < 0.001. Grey, no significant differences.

C. (left) UMAP projection and color clustering of 27,041 scATAC-seq profiles of fetal and postnatal microglia. Each dot represents one cell. (right) Bar chart indicating the relative fetal and postnatal contribution to each cluster.

D. Heatmap of average chromVAR score per motif and per cluster. Scores are averaged for all cells within each cluster and z-score normalized.

E. UMAP visualization of enrichment for indicated motifs using ChromVar.

F. Heatmap of gene expression of TFs best matching motifs identified in Figure 2E in fetal and postnatal microglia. Stars indicate genes that are differentially expressed (FC > 2, p-adj <0.05) between fetal and postnatal microglia.

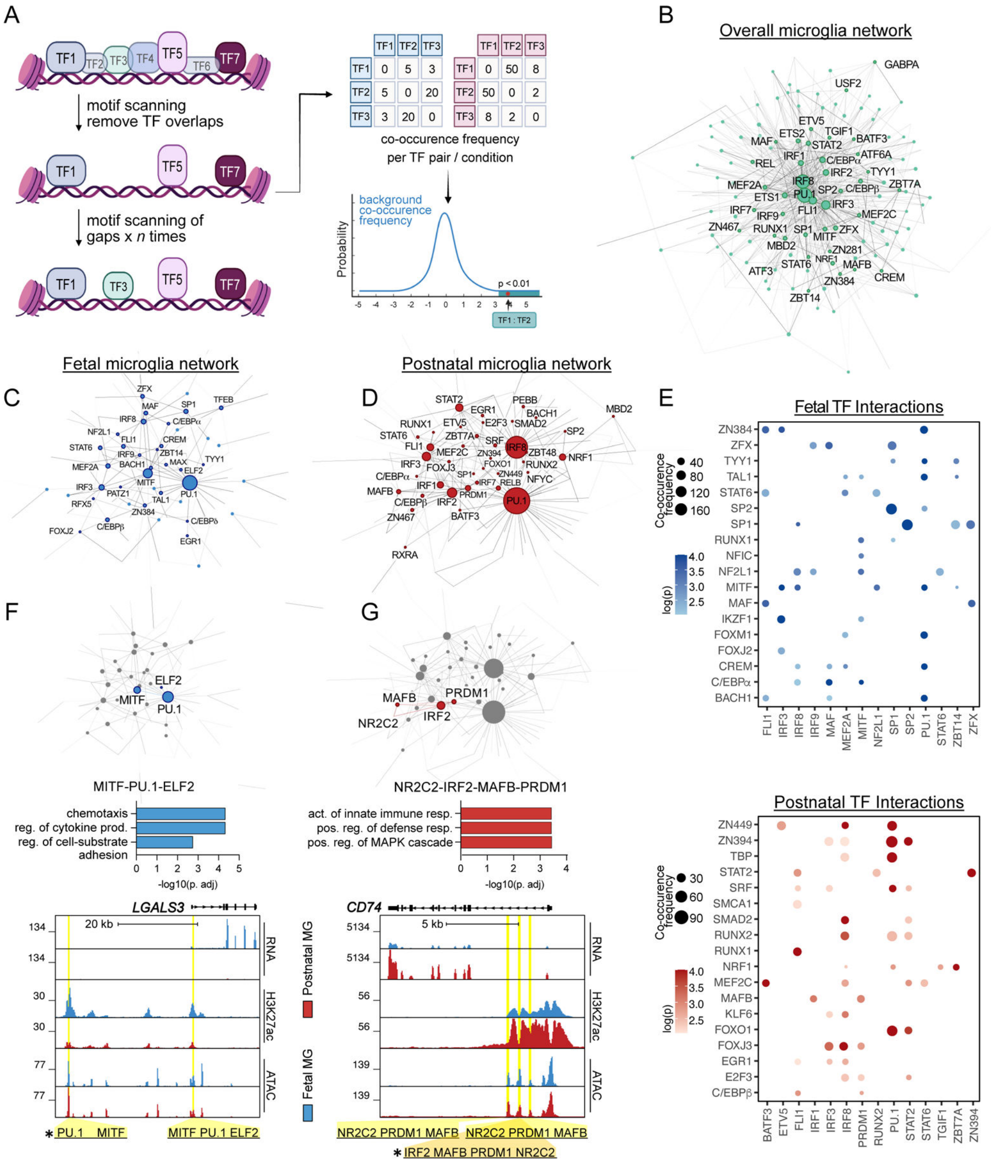


Figure 4. Human microglia have gene regulatory networks unique to development-stage.

A. Schematic of Transcription factor Interaction inference from Motif co-Occurrence Networks (TIMON). Height of TF is directly correlated with motif score.

B. TIMON analysis of all microglia (fetal and postnatal). Node represents a TF motif, edge represents significant ($p<0.001$) co-occurrence between two transcription factors. Node sizes are proportional to the node degree. Nodes with degree >10 are labeled.

C. Fetal microglia specific enhancer motif co-occurrence networks, nodes with degree >3 are labeled.

D. Postnatal microglia specific enhancer motif co-occurrence networks, nodes with degree >3 are labeled.

E. Bubble plots of significant co-occurring transcription factor pairs in fetal and postnatal microglia. Node size represents the co-occurrence frequency, opacity is proportional to the significance level.

F. MITF-PU.1-ELF2 (top) is a TF clique of the fetal microglia network. (middle) GO terms of genes linked to enhancers containing MITF-PU.1-ELF2. (bottom) UCSC browser tracks showing expression of *LGALS3* in fetal (blue) and postnatal microglia (red) and corresponding ATAC-seq and H3K27ac ChIP-seq peaks associated with *LGALS3*. Enhancer regions with MITF-PU.1-ELF2 combinatorial motifs are highlighted yellow. Star indicates the enhancer region in which binding of indicated TFs were validated in Supplemental Figure 4G.

G. NR2C2-IRF2-MAFB-PRDM1 (top) is a clique of the postnatal microglia network. (middle) GO terms of genes linked to enhancers containing NR2C2-IRF2-MAFB-PRDM1. (bottom) UCSC browser tracks showing expression of *CD74* in fetal (blue) and postnatal microglia (red) and corresponding ATAC-seq and H3K27ac ChIP-seq peaks associated with *CD163*. Combinatorial NR2C2-IRF2-MAFB-PRDM1 motifs are highlighted yellow. Star indicates the enhancer region in which binding of indicated TFs were validated in Supplemental Figure 4H.

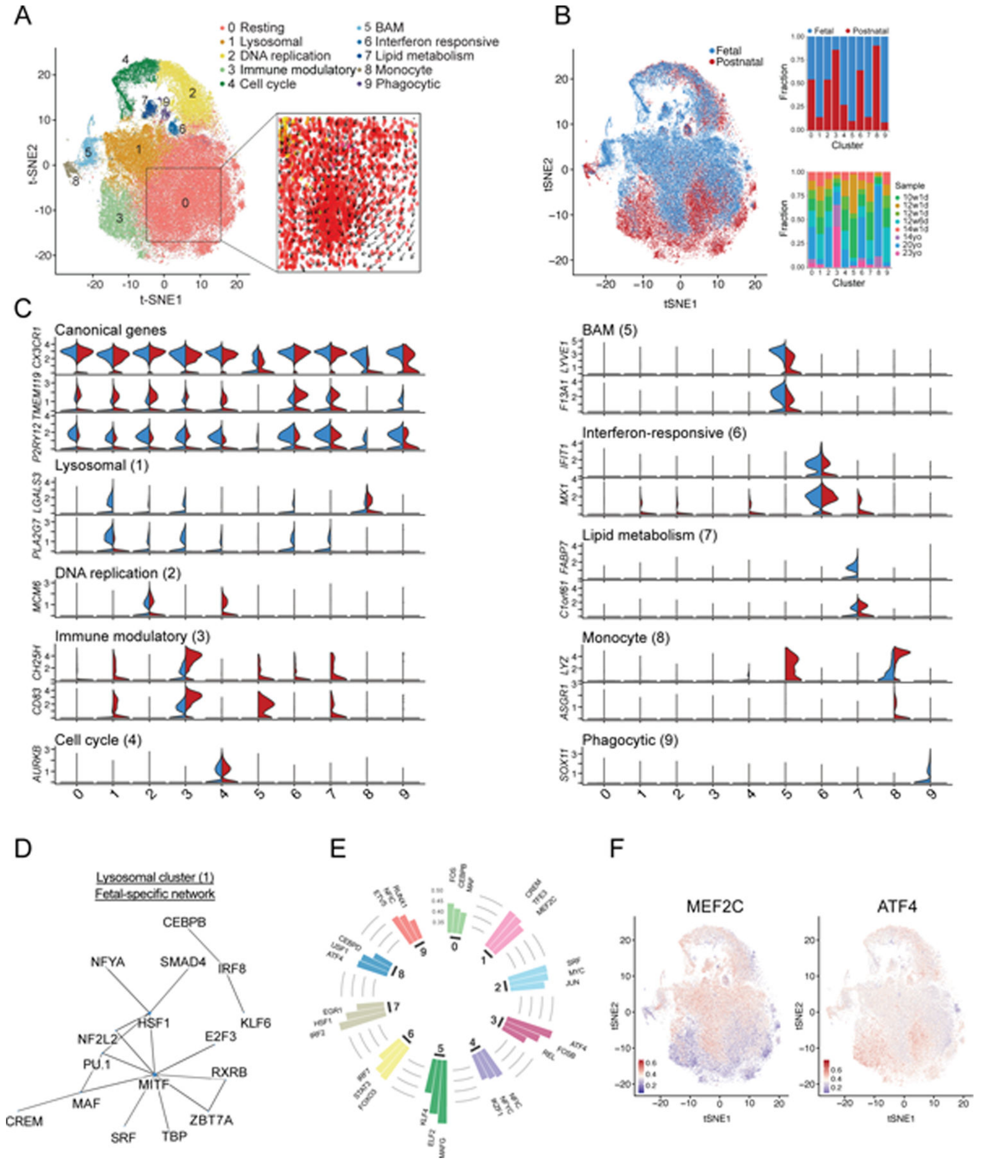


Figure 5. Transcriptional heterogeneity of primary human microglia during development.

A. Annotation of scRNA-seq clusters with inset depicting RNA velocity analysis of cluster 0.

B. (left) tSNE projection of scRNA-seq analysis of fetal and postnatal microglia. Each dot represents one cell with coloring indicating age contribution. Bar graphs illustrating age contribution (middle) and sample contribution (right) to each cluster.

C. Split violin plots showing the distribution of gene expression per cluster for fetal (blue) and PN (red) microglia.

D. Fetal microglia specific TIMON analysis for cluster 1.

E. Circle plot representing TFAct scores per cluster with height of bars indicative of TF activity scores. Please also see Figure S6.

F. tSNE projections of TF activity scores for MEF2C (left) and ATF4 (right).

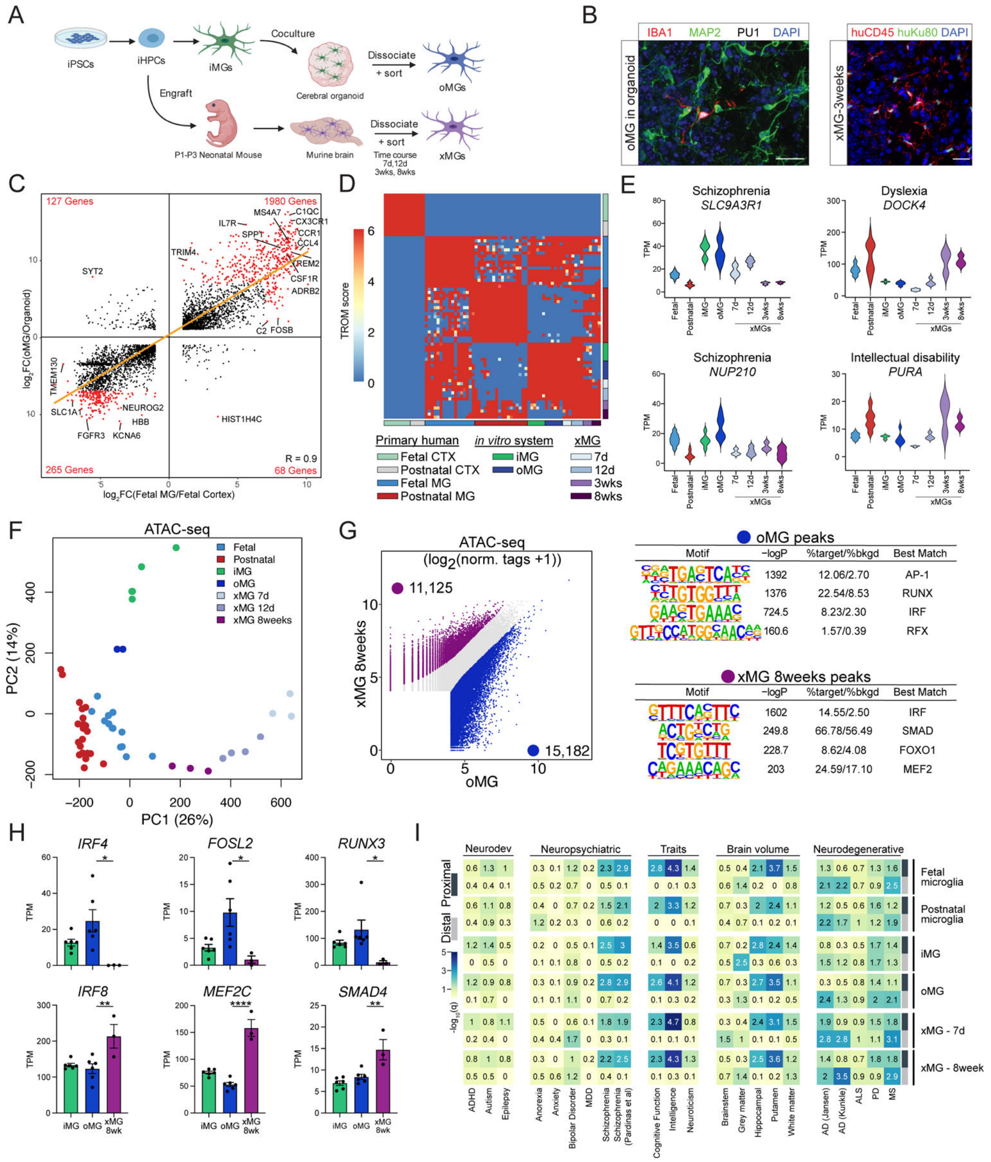


Figure 6. iPSC and xenotransplantation systems capture distinct stages of human microglia maturation.

A. Schematic demonstrating derivations of HPCs, iMGs, oMGs, and xMGs from iPSCs.

B. Immunohistochemistry depicting oMGs (IBA1, red; PU1, white) in proximity to neurons (MAP2, green) in organoids (left) and xMGs (right) (huCD45, red; human nuclei-Ku80, green; DAPI, blue) 3 weeks after engraftment into humanized mouse brains.

C. Ratio-ratio plot of genes comparing FC in gene expression between oMGs relative to cerebral organoids versus FC in gene expression between fetal microglia relative to fetal cortex. Pearson's correlation coefficient is indicated in bottom right.

D. TROM correspondence map of the transcriptomes of fetal and postnatal microglia and cortex, iMG, oMGs, and xMGs. Values are TROM scores, 6 being best match.

E. Violin plots of gene expression (TPM) that are associated with indicated neurodevelopmental disorders.

F. PCA of distal ATAC-seq peaks (> 1000bp from TSS) of primary human microglia, *in vitro* iPSC models and *in vivo* xMGs.

G. (left) Scatterplot of distal ATAC-seq peaks showing differentially accessible regions in oMGs (blue) and xMGs-8weeks post-engraftment (purple). (right) *De novo* motif analysis of differentially accessible regions in oMGs (top) and xMGs (bottom).

H. Bar chart of expression of TFs known to bind to DNA motifs identified in (G). $*p<0.05$, $** p<0.01$, $*** p<0.001$, $**** p<0.0001$.

I. Heatmap of LDSC analysis of enrichment of genetic variants associated with listed conditions displayed as -log10(q) value for significance of enrichment for promoter-proximal (light grey) and distal elements (dark grey) in indicated cell types.

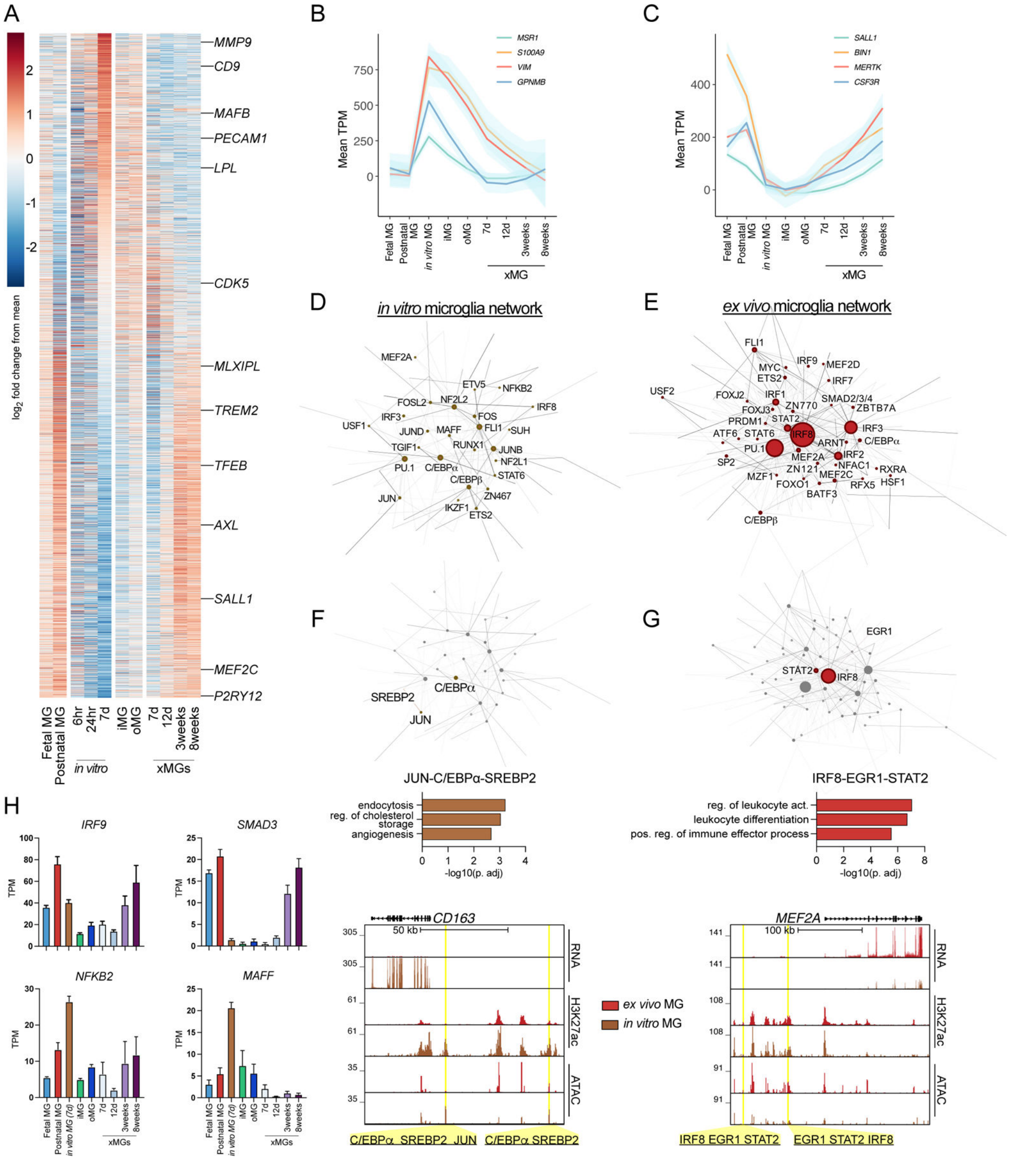


Figure 7. Transcription factor networks underlying environment-dependent genes in microglia.

A. Heatmap of gene expression changes of environmentally dependent genes across primary human microglia and iPSC-systems.

B. Mean TPM (LOESS fit) of microglia genes that are increased in *in vitro* microglia as compared to *ex vivo* microglia.

C. Mean TPM (LOESS fit) of environmentally dependent microglia genes that regain expression in xMGs.

D. Transcription factor network, derived from TIMON, of *in vitro* human microglia.

E. Transcription factor network, derived from TIMON, of *ex vivo* human microglia.

F. JUN-C/EBPα-SREBP2 (top) is a clique of the *in vitro* microglia network. (middle) GO terms of genes linked to enhancers containing JUN-C/EBPα-SREBP2. (bottom) UCSC browser tracks showing RNA expression of *CD163* in *in vitro* (brown) and *ex vivo* microglia (red) and corresponding ATAC-seq and H3K27ac ChIP-seq peaks associated with

CD163. Enhancer regions with JUN-C/EBPα-SREBP2 combination motifs are highlighted in yellow.

G. IRF8-EGR1-STAT2 (top) is a clique of the *ex vivo* microglia network. (middle) GO terms of genes linked to enhancers containing IRF8-EGR1-STAT2. (bottom) UCSC browser tracks showing RNA expression of *MEF2A* in *in vitro* (brown) and *ex vivo* microglia (red) with corresponding ATAC-seq and H3K27ac ChIP-seq peaks. Enhancer regions in which combinations of IRF8-EGR1-STAT2 motifs are detected highlighted in yellow.

H. Bar chart of expression of select brain environment regulated TFs. Genes are $p < 0.001$ by one-way ANOVA.

KEY RESOURCE TABLE

Reagent/Resource	**Source**	**IDENTIFIER**
Antibodies		
Anti-human CD11b PE (clone ICRF44)	Biolegend	301306; RRID: AB_314158
Anti-human CD45 APC-Cy7 (clone HI30)	Biolegend	304014; RRID: AB_314402
Anti-human CD45 (clone HI30)	Biolegend	304001; RRID: AB_314389
Anti-human CD64 APC (clone: 10.1)	Biolegend	305014; RRID: AB_1595428
Anti-human CX3CR1 PerCP-Cy5.5 (clone: 2A9–1)	Biolegend	341614; RRID: AB_11219203
Anti-human CD14-AF 488 (clone M5E2)	Biolegend	301811; RRID: AB_493159
Anti-human HLA-DR PE-Cy7 (clone L243)	Biolegend	307616; RRID: AB_493588
Anti-human CX3CR1 (clone 2A9–1)	Biolegend	341602; RRID: AB_1595422
Anti-human CD192-BV510 (clone K036C2)	Biolegend	357217; RRID: AB_2566504
Rabbit anti-mouse/human OLIG2 A647 (clone EPR2673)	Abcam	ab225100; RRID: AB_10861310
Mouse anti-mouse/human NeuN AF488 (clone A60)	MilliporeSigma	MAB377X; RRID: AB_2149209
Rabbit anti-mouse/human PU.1 PE (clone 9G7)	Cell Signaling	81886S; RRID:AB_2799984
Goat Anti-Iba1	Abcam	ab5076; RRID:AB_2224402
Rabbit Anti-Iba1	Wako	019–19741; RRID:AB_839504
Human Anti-Nestin	EMD Millipore	ABD69; RRID: AB_2744681
Chicken Anti-MAP2	Abcam	ab5392; RRID: AB_2138153
Rabbit Anti-PU.1	Cell Signaling Technology	2266; RRID:AB_10692379
Rabbit Anti-P2RY12	Millipore Sigma	HPA014518; RRID:AB_2669027
Goat Anti-IL-17/IL-17A	R&D	AF-317-NA; RRID:AB_354463
Mouse Anti-IL-17RA/IL-17R	R&D	MAB177; RRID:AB_2125546
Rabbit Anti-Sox2	Cell Signaling Technology	2748; RRID:AB_823640
Rabbit anti-MafB (BLR046F)	Bethyl Laboratories	A700–046; RRID: AB_2891845
Rabbit anti-TFEB	Bethyl Laboratories	A303–673A; RRID: AB_11204751
Rabbit anti-PU.1 (9G7)	Cell Signaling	2258; RRID: AB_2186909
Mouse anti-NR2C2 (OTI4B1)	Origene	TA807275
Rabbit anti-IRF7 (G-8)	Santa Cruz Antibodies	sc-74472; RRID:AB_2280489
Mouse anti-MiTF	Active Motif	91201; RRID: AB_2793801
Rabbit anti-PRDM1	Rockland	600-401-B52; RRID:AB_1961755
Mouse anti-Ki67	Leica Biosystems	ACK02; no RRID
Rabbit anti-LYVE1	Abcam	ab36993; RRID:AB_2138663
Mouse anti-human CD163 (clone EDHu-1)	Bio-Rad	MCA1853; RRID:AB_2074540
Donkey Anti-Goat IgG Cy3	Jackson ImmunoResearch	705-165-147; RRID: AB_2307351
Donkey Anti-Goat IgG AF488	Jackson ImmunoResearch	705-545-147; RRID: AB_2336933
Donkey Anti-Goat IgG AF647	Jackson ImmunoResearch	705-605-147; RRID: AB_2340437
Donkey Anti-Goat IgG H&L Cy3	Abcam	Ab6949; RRID:AB_955018
Donkey Anti-Rabbit IgG Cy3	Jackson ImmunoResearch	711-165-152; RRID: AB_2307443

Reagent/Resource	**Source**	**IDENTIFIER**
Donkey Anti-Rabbit IgG AF488	Jackson ImmunoResearch	711-545-152; RRID: AB_2313584
Donkey Anti-Rabbit IgG AF647	Jackson ImmunoResearch	711-605-152; RRID: AB_2492288
Donkey Anti-Mouse IgG Cy3	Jackson ImmunoResearch	715-165-151; RRID: AB_2315777
Donkey Anti-Mouse AF488	Jackson ImmunoResearch	715-545-151; RRID: AB_2341099
Goat anti-Mouse AF488	Jackson ImmunoResearch	115-545-062; RRID: AB_2338845
Donkey anti-Mouse AF488	Jackson ImmunoResearch	715-545-151; RRID: AB_2341099
Donkey anti-Chicken AF488	Jackson ImmunoResearch	703-545-155; RRID: AB_2340375
Donkey anti-Rat AF647	Jackson ImmunoResearch	712-605-153; RRID: AB_2340694
Mouse anti-H3K27ac (clone MABI 0309)	Active Motif	39085; RRID: AB_2793305
Chemicals, Peptides and Recombinant Proteins		
KAPA SYBR FAST qPCR Master mix (2X)	Kapa Biosystems	07959427001
Dynabeads Protein A	Thermo Fisher Scientific	10002D
SpeedBeads magnetic carboxylate modified particles	GE Healthcare	65152105050250
TRIzol LS Reagent	Thermo Fisher Scientific	10296028
Formaldehyde, 37% by weight	Thermo Fisher Scientific	F79–1
Dulbecco’s PBS (DPBS) solution	Thermo Fisher Scientific	MT21031CV
Disuccinimidyl glutarate (DSG)	ProteoChem	C1104
Dimethyl sulfoxide (DMSO)	MilliporeSigma	D2650
UltraPure DNase/RNase-free distilled water	Thermo Fisher Scientific	10977023
Glycine	MilliporeSigma	4810
1M Tris-HCl, pH 8.0	Thermo Fisher Scientific	15568025
0.5 M EDTA, pH 8.0	Thermo Fisher Scientific	15575020
1M $MgCl_2$	Thermo Fisher Scientific	AM9530G
Sucrose	Thermo Fisher Scientific	S6500
Triton X-100	MilliporeSigma	T8787
1,4-Dithiothreitol	Thermo Fisher Scientific	BP172–5
Bovine serum albumin	MilliporeSigma	A3059
4′,6-Diamidino-2-phenylindole, dilactate (DAPI)	BioLegend	422801
Oligo $d(T)_{25}$ Magnetic Beads	NEB	S1419S
DTT	Thermo Fisher Scientific	P2325
SUPERase-In	Ambion	AM2696
Oligo dT	Thermo Fisher Scientific	18418020
Random Primers	Thermo Fisher Scientific	48190011
Agencourt RNA Clean XP Beads	Beckman Coulter	A63987
10X Blue Buffer	Enzymatics	P7050L
DNA Polymerase I	Enzymatics	P7050L
SuperScript III Reverse Transcriptase	Thermo Fisher Scientific	18080044
5x First-strand Buffer	Thermo Fisher Scientific	18080044
Actinomycin D	Sigma	A1410

Reagent/Resource	**Source**	**IDENTIFIER**
DMEM / F12 (1:1) (1X)	Thermo Fisher Scientific	Cat#11330–032
GlutaMAX (100X)	Thermo Fisher Scientific	Cat#35050–061
MEM NEAA (100X)	Thermo Fisher Scientific	Cat#11140–050
Penicillin-Streptomycin	Thermo Fisher Scientific	Cat#15140–122
B-27 Supplement (50X)	Thermo Fisher Scientific	Cat#17504–044
N-2 Supplement (100X)	Thermo Fisher Scientific	Cat#17502–048
Insulin-Transferrin-Selenium (100X)	Thermo Fisher Scientific	Cat#41400–045
1-thioglyercol	Sigma-Aldrich	Cat#M1753
Insulin solution human	Sigma-Aldrich	Cat#I9278
rhM-CSF	Proteintech	Cat#HZ-1192
rhIL-34	Proteintech	Cat#HZ-1316
rhTGFβ1	Proteintech	Cat#HZ-1087
STEMdiff Hematopoietic Kit	STEMCELL Technologies	Cat#5310
2-mercaptoethanol (50mM)	Thermo Fisher Scientific	Cat#31350010
ReLeSR	STEMCELL Technologies	Cat#05872
KnockOut Serum Replacement	Thermo Fisher Scientific	Cat#10828–028
Heparin	Sigma-Aldrich	Cat#H3149
Dorsomorphin dihydrochloride	Tocris	Cat#3093
A83–01	AdooQ Bioscience	Cat#A12358–50
CHIR-99021	AdooQ Bioscience	Cat#A10199–100
SB431542	StemRD	Cat#SB-050
Y-27632	STEMCELL Technologies	Cat#72308
Collagenase IV	Thermo Fisher Scientific	Cat#17104–019
Cultrex Reduced Growth Factor Basement Membrane Extract, PathClear	R&D Systems	Cat#3433-010-01
RNAscope 3-plex Positive Control Probe-HS	Advanced Cell Diagnostics	320861
RNAscope 3-plex Negative Control Probe	Advanced Cell Diagnostics	320871
TSA Vivid Fluorophore 520 - FITC	Advanced Cell Diagnostics	323271
TSA Vivid Fluorophore 570 – Cy3	Advanced Cell Diagnostics	323272
TSA Vivid Fluorophore 650 – Cy5	Advanced Cell Diagnostics	323273
RNAscope Probe-Hs-ApoE	Advanced Cell Diagnostics	433091
RNAscope Probe-Hs-SORL1-C2	Advanced Cell Diagnostics	579461-C2
RNAscope Probe-Hs-P2RY12- C3	Advanced Cell Diagnostics	450391-C3
RNAscope Probe-Hs-CCR5	Advanced Cell Diagnostics	601501

Reagent/Resource	Source	IDENTIFIER
RNAscope Probe-Hs-CCL3-C2	Advanced Cell Diagnostics	455331-C2
Critical Commercial Assays		
Direct-zol RNA MicroPrep Kit	Zymo Research	R2062
Qubit dsDNA HS Kit	Thermo	Q32851
Nextera DNA Library Prep Kit	Illumina	FC-121-1030
NEBNext Ultra II DNA library prep kit	NEB	E7645L
ChIP DNA Clean and Concentrator Kit	Zymo Research	D5205
Chromium Single Cell 3’ Library and Gel Bead Kit v3	10X Genomics	PN-1000075
Chromium Single Cell 3’ Library Construction Kit v3	10X Genomics	PN-100078
Chromium Chip B Single Cell Kit	10X Genomics	PN-1000153
Single Index Kit T Set A	10X Genomics	PN-1000213
Chromium Next GEM Single Cell ATAC Library and Gel Bead Kit v1.1	10X Genomics	PN-1000175
Chromium Next GEM Chip H Single Cell Kit	10X Genomics	PN-1000162
Single Index Kit N Set A	10X Genomics	PN-1000212
RNAscope Multiplex Fluorescent Reagent Kit v2	Advanced Cell Diagnostics	323100
Experimental models: Cell lines		
H1 (ID: WAe001-A)	WiCell	Wa01; RRID:CVCL_9771
EC11		Firth et al PMID 26299960
UKERfG3G-X-001		Schultze et al 29986767
Experimental models: Organisms/Strains		
Mouse: Csf1rΔFIRE/ΔFIRE	David A. Hume & Clare Pridans	Rojo et al, PMID: 31324781
C;129S4-Rag2tm1.1Flv Csf1tm1(CSF1)Flv Il2rgtm1.1Flv/J	Jackson Laboratory	Strain #:017708; RRID:IMSR_JAX:01770
Data		
Human fetal microglia RNA-, ATAC-, H3K27ac-ChIP-seq		This paper
Human postnatal microglia RNA-, ATAC-, H3K27ac-ChIP-seq		This paper dbGaP: phs001373.v2.p2
Human fetal and postnatal microglia scRNA- and scATAC-seq		This paper
iMG, oMG, xMG RNA-, ATAC-seq		This paper
IL-4 treated BMDMs RNA-, ATAC-,H3K27ac-ChIP-seq, PPARγ, STAT6, C/EBPβ-ChIP-seq	GEO	GSE159630
Control and NASH scRNA-seq	GEO	GSE128334
Software and Algorithms		
Bowtie2	(Langmead and Salzberg, http://bowtie-bio.sourceforge.net/2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Cell Ranger		https://github.com/10XGenomics/cellranger
FlowJo		https://www.flowjo.com/

Reagent/Resource	**Source**	**IDENTIFIER**
HOMER		http://homer.ucsd.edu/homer/
Metascape		http://metascape.org/gp/index.html#/main/step1
R package: Seurat		https://satijalab.org/seurat/
R package: Pheatmap		https://CRAN.R-project.org/package=pheatmap
R package: Tidyverse		https://CRAN.R-project.org/package=tidyverse
R package: RColorBrewer		https://CRAN.R-project.org/package=RColorBrewer
R package: WGCNA		https://cran.r-project.org/web/packages/WGCNA/index.html
Velocyto		
R package: NicheNet		https://github.com/saeyslab/nichenetr
R package: VISION		https://github.com/YosefLab/VISION
R package: MACS2		https://github.com/macs3-project/MACS
R package: chromVar		https://bioconductor.org/packages/release/bioc/html/chromVAR.html
Python: LDSC		https://github.com/bulik/ldsc
R package: DEseq2		https://bioconductor.org/packages/release/bioc/html/DESeq2.html
R package: Velocyto		https://smorabit.github.io/tutorials/8_ velocyto/
GraphPad Prism		
Other		
BD Influx		Equipment
BD FACSARIA Fusion		Equipment
MoFlo Astrios		Equipment