IRF8 configures enhancer landscape in postnatal microglia and directs microglia specific transcriptional programs

Keita Saeki [1], Richard Pan[1,2], Eunju Lee[1], Daisuke Kurotaki[3], Keiko Ozato[1]

1: Division of Developmental Biology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, 20892

2: MD-PhD Candidate in Neurobiology and Behavior, Columbia University, School of Medicine, New York, NY

3: International Research Center for Medical Sciences, Kumamoto University, 860-011 Kumamoto City, Japan

Running title; IRF8 organizes enhancer modules in microglia

Keywords: IRF8; microglia maturation; enhancer; Alzheimer's disease

Address correspondence to Keiko Ozato (ozatok@nih.gov) or Keita Saeki (saekik2@nih.gov). Section on Molecular Genetics of Immunity, Division of Developmental Biology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, 20892

Abstract

Microglia are innate immune cells in the brain. Transcription factor IRF8 is highly expressed in microglia. However, its role in postnatal microglia development is unknown. We demonstrate that IRF8 binds stepwise to enhancer regions of postnatal microglia along with Sall1 and PU.1, reaching a maximum after day 14. IRF8 binding correlated with a stepwise increase in chromatin accessibility, which preceded the initiation of microglia-specific transcriptome. Constitutive and postnatal *Irf8* deletion led to a loss of microglia identity and gain of disease-associated microglia-like genes. Combined analysis of scRNA-seq and scATAC-seq revealed a correlation between chromatin accessibility and transcriptome at a single-cell level. IRF8 was also required for microglia-specific DNA methylation patterns. Lastly, in the 5xFAD model, constitutive and postnatal *Irf8* deletion reduced the interaction of microglia with Aβ plaques and the size of plaques, lessening neuronal loss. Together, IRF8 sets the epigenetic landscape, which is required for postnatal microglia gene expression.

Abbreviations: AD, Alzheimer's disease; DEG, differentially expressed gene; FACS, fluorescence-activated cell sorting; FC, fold change; FDR, false discovery rate; GFP, green fluorescent protein; GO, gene ontology; IRF8, interferon regulatory factor 8; TF, transcription factor; TES, transcription end sites; TSS, transcription start sites; YFP, yellow fluorescent protein; WT, wild type.

Introduction

Microglia elicit innate immunity in the brain, providing antimicrobial protection[1, 2]. Microglia also promote brain development through synaptic pruning, supporting neuronal survival, and vascular formation[3, 4, 5, 6]. Microglia continuously surveil the entire brain and respond to changes in their environment[7, 8, 9]. Dysregulation of microglia is associated with neurodegenerative diseases such as Alzheimer's disease (AD)[10, 11].

Microglia arise from the yolk sac-derived progenitors and expand in the brain in SPI1 (PU.1) and IRF8-dependent manner[12]. Transcription factors, such as Smad3 and Mafb, contribute to microglial differentiation[13]. Evidence indicates that microglia mature in the postnatal stage since they do not elicit full gene expression until later[14, 15]. Postnatal microglia maturation may be part of global change in the brain to adapt to the extra-uterine environment and develop further into adulthood[6, 14, 15]. However, postnatal microglia development and mechanisms controlling the process have not been fully understood.

IRF8 is a DNA binding transcription factor that binds to GAAAG ISRE element (ISRE) and TTCC…G/CTTT PU.1/IRF composite motif (ETS/ISRE), the latter by interacting with PU.1[16]. Outside the brain, IRF8 is expressed in developing myeloid cells in bone marrow, Ly6C^{+} macrophages, and CD8^{+} or plasmacytoid dendritic cells[17, 18]. IRF8 is essential for innate resistance against various pathogens[19]. Constitutive *Irf8* knockout (IRF8KO) mice are susceptible to infection while having excess neutrophil-like cells[20]. In the brain, IRF8 is expressed only in microglia, except for a few macrophages[21]. Besides participating in early microglia development, IRF8 responds to cytokines and

pain injury in adults[12, 22, 23]. However, the role of IRF8 in postnatal microglia maturation is largely unknown[13, 14].

CUT&RUN analysis revealed that IRF8 binds to enhancer regions of postnatal microglia gradually from P9 to P14, reaching a maximum at adult. IRF8 binding was a prerequisite for setting chromatin accessibility, histone modifications, and DNA methylation patterns unique to microglia, which preceded microglia specific gene expression. Postnatal deletion of *Irf8* (IRF8cKO) demonstrated that continuous IRF8 expression after P14 is necessary for maintaining microglia gene expression. Lastly, in the 5xFAD mouse model of AD, constitutive and postnatal deletion of *Irf8* led to reduced AD pathology due to reduced interaction of microglia with amyloidβ (Aβ) plaques. Together, IRF8 defines the epigenetic landscape in postnatal microglia, thereby directing their transcriptome programs.

Results

IRF8 binds to the microglia genome during postnatal development in a stepwise fashion

The IRF8's role in postnatal development and adult microglia remains unclear. To gain insight into this question, we sought to identify genome-wide IRF8 binding sites in postnatally developing microglia. To first verify IRF8 expression in postnatal microglia, we examined Irf8-GFP knock-in mice that express endogenous IRF8 fused to GFP[24]. FACS analyses of microglia from P9, P14, and P60 (adult) showed that IRF8 is expressed at similar levels at all times, higher than that in peritoneal macrophages (Fig. S1A, S1B).

CUT&RUN experiments were carried out for microglia isolated from Irf8-GFP mice using anti-GFP antibody. Peritoneal macrophages from the same mice were also examined as a reference. We identified 16,935 IRF8 peaks in adult microglia and 12,062 in peripheral peritoneal macrophages (Fig. 1A). About half of the IRF8 peaks were in the intergenic regions with >5 kb distal from the genes, while about 20% were at the promoter regions in both cell types (Fig. 1A). *De novo* motif analysis revealed enrichment of ETS/ISRE and ISRE motifs (Fig. 1B). DEseq2-based statistics revealed a differential IRF8 binding profile between microglia and peritoneal macrophages (Fig. 1C). Further, GREAT GO annotation indicated microgli-specific IRF8 peaks (n=9,679) to be related to cell/leukocyte activation and differentiation (Fig. 1D; upper panel). In contrast, macrophage-specific IRF8 peaks (n=2,251) were related to immune responses and inflammatory responses (Fig. 1D; lower panel). In addition, in microglia, IRF8 binding was found near microglia identity genes, such as *Sall1* and *Cx3cr1*[11, 25], but not in peritoneal macrophages (Fig. 1G).

IRF8 binding peaks on P9, P14, and adult microglia are shown in Fig. 1E. The IRF8 binding peaks on P9 microglia were very sparse. The binding increased at P14, although still less than on adult microglia, indicating a stepwise increase in IRF8 binding (Fig. 1E). Furthermore, cumulative distribution frequency graphs plotting the distance between two neighboring IRF8 binding sites (Fig. 1F, right) showed that the distance in the adult is the shortest, and the longest in P9 microglia, indicating that IRF8 binding in increases by forming arrays during postnatal development (Fig.1F left). IGV tracks corroborated the stepwise increase in IRF8 binding observed over the enhancer region of *Sall1* and *Cx3cr1* genes (Fig. 1G, left). In peritoneal

macrophages, IRF8 was bound to regulatory regions of *Clec4n and Lyz1/2* which are active in these cells (Fig. 1G, right). Thus, IRF8 binding to the microglial genome is developmentally controlled and reaches a maximum in adults when microglia specific genes become fully expressed[14, 15]. It is of note that despite a marked increase in binding activity, the amount of IRF8 was virtually the same during postnatal stages (Fig. S1B).

IRF8 directs the expression of microglia identity genes and suppresses DAM genes.

To study how IRF8 binding controls microglia's transcriptional programs, we carried out bulk RNA-seq for microglia isolated from wild-type (WT) and constitutive IRF8KO mice. FACS analyses of brain cells revealed two $CD45^+$ populations in the IRF8KO brain, one $CD11b^{high}CD45^+$ corresponding to WT microglia, and another $CD11b^{low}CD45^+$ subset, termed P3, which was absent in WT brain (Fig. S2A, S2B). The number of $CD11b^{high}CD45^+$ cells in IRF8KO brains was less than that of WT microglia. Cells termed P3, about 10 % of $CD11b^{high}CD45^+$ cells were found in all IRF8KO brains examined (Fig. S2A, S2B). The former subset expressed microglial markers, CX3CR1 and TMEM119, although its levels were lower than WT microglia. P3 cells did not express microglia markers nor border-associated macrophage marker CD206 but expressed Ly6C (Fig. S2C, S2D), indicating the presence of an additional cell population. Detailed characterization of P3 cells will be presented elsewhere. We further assessed the presence of border-associated macrophages (BAM) in our microglia preparations[21, 26]. FACS analysis revealed

that cells with CD206 and FOLR2 were less than 1% of the entire brain CD11b^{+}CD45^{+} cells prepared from WT and IRF8KO mice (Fig. S2E, S2F).

We performed bulk RNA-seq analyses for adult WT and constitutive IRF8KO microglia and found that the microglia identity genes and CLEAR lysosomal genes were drastically downregulated in IRF8KO cells (Fig.2A, B). The former include *P2ry12, Siglech,* and *Ccr5,* encoding surface proteins and a transcription factor, *Sall1,* while the latter include *Ctsb, Ctsc, Tfeb,* and *Lamp1* involved in waste clearance (Fig. 2B right)[11, 27]. Conversely, interferon-related genes and disease-associated microglia (DAM) like genes, such as *Apoe,* were strongly upregulated in IRF8KO cells[11,28, 29] (Fig. 2B).

Since IRF8 is known to activate transcription factors to promote myeloid cell development, we examined levels of transcription factors expressed in postnatal microglia[13, 18, 30]. Among them, *Batf3* and *Sall1* were markedly downregulated in IRF8KO cells (Fig. 2C, blue square). BATF3, a basic leucine zipper motif (bZIP) family of a transcription factor, is known to be critical for generating a dendritic cell subset[31]. Because the role of BATF3 in microglial transcriptome remained unelucidated, we next analyzed RNA-Seq in Batf3KO and WT microglia (Fig. 2D). As anticipated, genes downregulated in Batf3KO microglia, such as *P2ry13* and *Ctsc* were also downregulated in IRF8KO cells, indicating that IRF8 and BATF3 control an overlapping set of target genes (Fig. 2E, Table S1 for all overlapped genes). SALL1 is a transcription factor of the Spalt family, expressed selectively in microglia. Buttgereit et al. reported that SALL1 is required for the acquisition of microglia signature[25]. The correlation biplot also showed a significant overlap between SALL1 dependent genes and IRF8 dependent genes, including microglia identity genes (Fig. 2F, Table S1 for

all overlapped genes). Fixsen et al. recently identified the genome-wide Sall1 binding profile in microglia[32]. Peak overlap analysis for the IRF8 binding profile revealed that 6,998 (42%) of IRF8 co-bind to SALL1, most of which also bind to PU.1[32, 33]. Results suggest that IRF8, SALL1, and PU.1 cooperatively regulate the expression of those genes (Fig. 2G, Table S2 for all overlapped peaks and their nearest genes)

While IRF8 enhances the expression of *Batf3* and *Sall1*, they do not control IRF8 expression (Fig. 2D). These data indicate that IRF8 acts as an upstream factor in the transcription cascades and activates *Sall1* and *Batf3* expression, which in turn activates downstream target genes (Fig. 2H). We found that a complete microglia transcriptional program is realized after P14, suggesting that epigenome establishment precedes gene expression (Fig. S2G).

IRF8 configures microglia specific enhancers

To further characterize the chromatin environment in which IRF8 binds, we examined histone marks linked to active transcription, particularly those denoted for enhancers, H3K4me1 and H3K27ac[34, 35]. Of 16,935 IRF8 binding sites, about 73% and 66% of IRF8 colocalized with H3K4me1 and H3K27ac, respectively, and 64% of IRF8 binding sites carried both histone marks (Fig. 3A). However, the histone marks colocalizing with IRF8 peaks were much lower in IRF8KO cells than in WT cells, indicating that IRF8 helps set enhancer histone marks (Fig. 3B, S3A). Since the H3K27ac is a surrogate marker for super-enhancers and super-enhancers neighbor many genes critical for defining cell type specificity and function[35, 36], we examined super-enhancers

using the ROSE program[35]. As expected, many IRF8-dependent microglial identity genes were neighbored by super-enhancers, which totaled 837. Genes neighboring super-enhancers included *Sall1, Siglech,* and *Cst3* (Fig. 3C). Indeed, 735 of 837 (87.8%) super-enhancers were bound by IRF8 (Fig. 3D). While most H3K27ac^{high} regions were populated by IRF8 binding in adult microglia, this was not the case in early postnatal P9 microglia and peritoneal macrophages (Fig. 3E). In addition to super-enhancers, IRF8 was abundantly enriched in smaller enhancers found in specific intergenic and intronic regions overlapping with H3K4me1 (Fig. S3C-G). IGV tracks for the *Sall1* and *Batf3* genes showed that clusters of IRF8 binding coincide with those of H3K27ac and H3K4me1 (Fig. 3F). These results indicate that IRF8 binding guides the formation of enhancers unique to microglia during postnatal development.

IRF8 confers chromatin accessibility on postnatal microglia and sets DNA methylation patterns

To investigate genomic sites accessible for transcription and whether IRF8 influences the process, we performed an assay for transposase-accessible chromatin (ATAC) sequencing[37]. Differential accessibility analysis revealed that 10,514 peaks were lost in IRF8KO microglia while gaining 5,926 new peaks (Fig. 4A). As anticipated, gained and lost peaks were mostly localized to intergenic regions and introns (Fig. 4B). *De novo* motif analysis showed that lost peaks were enriched with the ETS/ISRE composite motifs (Fig. 4C, left). Gained peaks were enriched with motifs for PU.1, C/EBP, and Smad2/3 binding sites (Fig. 4C, right). Further, ATAC peaks lost in IRF8KO cells were enriched with WT-specific H3K4me1 and H3K27ac marks. On the other

hand, IRF8KO-specific H3K4me1 marks were found in gained ATAC peaks (Fig. S3A, 4D). Besides, IRF8KO cells had extra PU.1 binding sites absent in WT microglia (Fig. S3B). Comparison of ATAC-seq peaks and RNA-seq profiles found at lost ATAC-seq peaks in IRF8KO cells were enriched in genes downregulated in IRF8KO cells, such as microglia identity genes (Fig. 4E, in red). Whereas gained peaks were enriched in genes upregulated in IRF8KO cells (Fig. 4E, in purple). IGV tracks showed multiple ATAC peaks on and near *Sall1* and *Batf3* genes in WT cells but few in IRF8KO cells (Fig. 4I). Further, in WT cells, H3K27ac and H3K4me1 marks were abundantly present at these ATAC-seq sites but reduced in IRF8KO cells (Fig. 4I). PU.1 binding sites also revealed a correlation with ATAC-seq peaks (Fig. 4I). However, *Apoe*, expressed in IRF8KO cells but not WT cells showed multiple ATAC-seq peaks in IRF8KO cells, not evident in WT cells. Together, these data indicate that IRF8 sets chromatin accessible sites in microglia, thereby establishing a microglia specific transcriptional program.

To determine whether chromatin accessibility changes during the postnatal stage, we further analyzed ATAC-seq from P9, P14, and adult microglia. ATAC-seq peak numbers increased from P9 to P14, and P14 to adult (Fig. S4A left). Many of these peaks overlapped with IRF8 peaks (Fig S4A right). Consistent with this, *de novo* motif analysis revealed that ETS/ISRE composite elements were enriched within these peaks (Fig. S4A right). In contrast, the Batf3 motif was found during the transition from P14 to adult (Fig. S4B). These data indicate that chromatin accessibility in microglia increases during the postnatal stage in an IRF8-dependent manner. We presented a list of the other transcription factors that may co-bind to IRF8 during the postnatal

stages in Table S3. That microglia gene expression increases substantively only after adulthood (Fig S2G), while ATAC-seq peaks increase earlier, supports the view that the formation of epigenome structure precedes that of transcriptional programs[38].

DNA methylation influences the epigenetic landscape in various somatic cells[39]. In some cells, DNA methylation is linked to gene silencing[39]. However, DNA methylation status has not been deciphered fully for microglia. To determine global DNA methylation sites in microglia and the role of IRF8 in affecting DNA methylation, we performed whole genome bisulfite sequencing. We identified over 20,000 methylated loci (DML) that were differentially regulated in WT and IRF8KO microglia (Fig. 4F)[40, 41]. In addition, 1,808 DNA methylated regions (DMR) and 1,581 demethylated regions were differentially regulated in WT and IRF8KO cells, indicating that IRF8 takes part in DNA methylation in microglia (Fig. 4F). Both methylated and demethylated regions were in intergenic regions and introns, away from the proximal promoter and TSS (Fig. 4G). We found that IRF8 binding sites were located mostly in demethylated regions in WT cells (Fig. 4H, top). Further, methylated regions were enriched with the H3K4me1 and H3K27ac marks present in IRF8KO cells, while demethylated regions were enriched with those marks in WT cells (Fig. 4H). Correlation biplot analysis found that demethylated regions in WT cells correlated with ATAC-seq peaks present in WT cells, while methylated regions correlated with ATAC-seq peaks found in IRF8KO cells (Fig. S4C). These data indicate that demethylated regions were accessible to transcription factors, while methylated sites were not. Consistent with this idea, GO analysis revealed that demethylated regions were enriched with terms such as regulation of

immune responses and leukocyte activity (Fig. S4D). In IGV tracks, individual methylated and demethylated sites are shown for *Sall1*, *Batf3*, and *Apoe* (Fig. 4I, bottom two lanes). Sall1 and Batf3 were enriched with demethylated sites in WT cells, but not IRF8KO cells. Conversely, Apoe exhibited many methylated sites in WT cells but demethylated sites in IRF8KO cells. Together, our results indicate that the regulation of DNA methylation is an integral mechanism for setting epigenomic landscape in microglia, which IRF8 plays a critical role.

Single-cell RNA-seq analysis reveals different clusters for WT and IRF9KO microglia

Next, we performed single-cell RNA-Seq (scRNA-Seq) analysis to assess transcriptome diversity. After filtration, 3,504 WT and 5,303 IRF8KO cells expressing a total of 15,376 genes were selected. Leiden clustering denoted three WT clusters, seven IRF8KO clusters, and a "Common" cluster consisting of WT and IRF8KO cells (Fig. 5A and B). The SigEMD pipeline identified 6,149 DEGs between WT and IRF8KO cells, which covered 62% (3016/4848) of DEGs found in bulk RNA-seq analysis, lending credence to our data (Fig. S5A). Microglia identity gene expression was higher in all WT clusters than in all IRF8KO clusters (Fig. 5C, left). Population heterogeneity was evident within IRF8KO cells, in that ISGs were exclusively found in Cluster 4 (Fig. 5C, middle). In line with this, GO analysis of this cluster ranked interferon-related terms as the top category (Fig. S5B). DAM-like genes were enriched in IRF8KO Cluster 1, although not exclusive (Fig 5C, middle). However, CLEAR lysosomal gene expression was uniformly decreased in all IRF8KO cells (Fig. 5C, right). Interestingly, IRF8KO Cluster 6 did not express microglia identity

genes, nor other genes expressed in IRF8KO cells. Instead it expressed *Ly6c2*, encoding Ly6C (Fig. 5C, S5C). Thus, this cluster likely represents the P3 population in IRF8KO brain (Fig. S2C). The Common cluster exhibited unusual features, in that most microglial genes as well as *Ly6c2* were absent in this population (Fig. 5D). This raised the possibility that Common cluster represents a population prior to microglia maturation. We thus applied the PAGA program, which provided a trajectory network (Fig. 5E)[42]. This network illustrated a tight connection with the Common cluster to all WT and IRF8KO clusters except IRF8KO Cluster 6. Another unusual feature of Common cluster was the expression of cell cycle genes such as *Cenpa*, *Ccnb2*, and *Mki67*, a marker for cell proliferation (Fig. 5F). Lastly, Common cluster did not express genes downstream of IRF8, such as Sall1 and Batf3 (Fig.5G). Together, these results may support the idea that Common cluster represents a predecessor of microglia, possessing the capability to develop into functioning microglia and replenish microglia continually for life.

Postnatal IRF8 expression is required for transcription program in adult microglia

To assess whether IRF8 is required for postnatal microglia, we constructed Irf8$^{flox/flox}$Cx3cr1cre^{ERT2} mice to allow for conditional *Irf8* deletion. The mice also carried a Rosa26-loxP-stop-loxP-EYFP reporter, with which we could identify and isolate Cre-expressing microglia[43] (Fig. 6A). To our dismay, the efficiency of *Irf8* deletion was very poor when Tamoxifen was administered on 4-week-old mice five consecutive times (Fig. S6A). To improve deletion efficiency, we decided to delete Irf8 in younger mice between P12 and P16 and

found a considerable improvement in *Irf8* deletion efficiency (Fig. S6A). The basis of poor deletion efficiency is currently unclear, but it may be due to a shift in Cre accessibility. We first performed bulk RNA-seq for IRF8cKO and WT microglia. There were 334 genes downregulated in IRF8cKO cells, while 546 genes were upregulated in IRF8cKO cells (Fig. S6B). Many up and down-regulated genes were included in those found in constitutive IRF8KO cells (Fig S6C). Namely, many microglia identity genes and CLEAR genes were down-regulated, while DAM-like genes and Interferon-related genes were upregulated in IRF8cKO cells, indicating that the same gene sets were affected by constitutive and conditional IRF8 deletion (Fig S6D). The results suggested that IRF8 needs to be expressed in postnatal microglia to maintain the distinctive transcriptome program. Considering that *Irf8* deletion was partial, leaving uncertainty of YFP-based selection, we then performed single-cell multiome in the 10x Genomics platform, which provides scRNA-seq and scATAC-seq data for the same set of cells, enabling us to assess the relation between transcriptome profile and epigenetic landscape. We analyzed 14,319 nuclei after filtering low-quality data. For scRNA-seq data, RNA-guided clustering showed 3 WT and 3 IRF8cKO clusters (Fig. 6B, 6C). There were four additional small clusters of WT and IRF8cKO cells, termed Mixed clusters (Fig. 6B, 6C). As anticipated of incomplete Irf8 gene deletion, some cells from IRF8cKO mice were found within WT clusters, estimating that the deletion efficiency was 85% (Fig.6B, right). Microglia identity and CLEAR genes were downregulated in IRF8cKO clusters, revealing slight heterogeneity (Fig. 6D). DAM-like and Interferon-related genes upregulated in IRF8cKO cells also displayed cluster heterogeneity (Fig 6D). Notably, an equivalent of Common cluster observed in

IRF8KO microglia was not evident in the IRF8cKO model. Although the basis of the difference is uncertain, it may be due to the difference in the sample preparation steps, such as permeabilization. Incidentally, the identity of minor Mixed clusters in this analysis is currently unclear.

Next, we analyzed the ATAC-seq data to assess chromatin accessibility for the same set of cells. The clustering without reference to transcriptome exhibited six WT clusters, four IRF8cKO clusters, and three mixed clusters containing WT and IRF8cKO cells (Fig. 6E, 6F). Like scRNA-seq analysis, IRF8cKO cells contained WT ATAC clusters (Fig. 6E, right), indicating that IRF8cKO cells that did not delete *Irf8* maintained WT chromatin accessibility. Those scATAC-seq clusters showed expected correspondence with those of RNA-seq clusters, e.g., RNA-seq Cluster 1 in WT cells corresponded to WT ATAC-seq Cluster 1 as evidenced by more than 60% match (Fig 6G). Likewise, IRF8cKO RNA-seq Cluster 1 matched IRF8cKO ATAC-seq Cluster 1 by more than 50%. Nevertheless, considerable intermixing was evident, e.g., IRF8cKO RNA-seq Cluster 2 consisted of roughly 45% ATAC-seq Cluster 2, 40% ATAC-seq Cluster 1, and 15% ATAC-seq Cluster 3 cells (Fig 6G). These results suggest that each RNA-seq cluster carries a corresponding epigenome structure, although there is certain interchangeability.

There were a total of 19,899 lost peaks and 20,410 gained peaks in IRF8cKO cells (Fig. 6H). *De novo* motif analysis showed that lost peaks are enriched with ETS/ISRE composite elements, whereas gained peaks were enriched with PU.1 and C/EBP elements, consistent with bulk ATAC-seq data in IRF8KO cells (Fig. 6I). Lost peaks found in all three IRF8cKO clusters included the majority of microglia identity and CLEAR genes, although some

heterogeneity was evident (Fig, 6G, right). Conversely, DAM-like and Interferon-related genes were shown to be in the gain of ATAC peaks in IRF8cKO clusters (Fig 6H, right). This correlation supports a direct mechanistic link between chromatin accessibility and transcription operating at a single-cell level. Additionally, scATAC-seq clusters 2 and 3 in IRF8cKO correlated with the upregulation of DAM-like and Interferon-related genes that dictate heterogeneity of IRF8cKO RNA-seq clusters (Fig. 6J).

IGV profiles for scATAC-seq peaks for *Sall1* and *Batf3* regions show prominent peaks in all WT clusters but not in IRF8cKO clusters (Fig 6K, top). On the other hand, *Apoe,* one of the upregulated genes in IRF8cKO clusters 2 and 3, exhibited a gain of ATAC peaks in IRF8cKO clusters 2 and 3 but not WT clusters (Fig 6K, bottom). Together, maintaining transcriptional program and chromatin accessibility in adult microglia requires continuous IRF8 expression.

Irf8 deletion ameliorates AD pathology in the mouse model

IRF8 has been shown to promote neuroinflammation and is implicated in AD pathogenesis in mice[19, 23,44, 45]. To further address the role of IRF8 in AD, we first compared microglia transcriptome profiles of WT and 5xFAD mice at 9 months of age[46]. We found 699 genes upregulated in 5xFAD microglia over WT cells, which we regarded as the 5xFAD-associated gene set. This gene set contained many DAM-like genes, such as *Apoe*, *Tyrobp*, and *Cst7* (Fig. S7A)[10,29]. Interestingly, many DAM-like genes were upregulated in IRF8KO microglia without the 5xFAD background (Fig. 7A)[28, 45, 47]. Further, microglia from 5xFAD/IRF8KO mice showed essentially the same expression pattern as IRF8KO microglia (Fig. 7A). Grubman et al. reported that 5xFAD microglia *in*

1 *vivo* have two distinct populations, one with engulfed Aβ, and the other without,
2 which could be distinguished by staining with a brain-permeable Aβ probe,
3 MethoxyX04 (X04)[48]. X04$^+$ cells, but not X04$^-$ cells, expressed DAM-like genes
4 along with interferon-related genes[48]. However, mechanisms of setting
5 transcriptome programs in X04$^+$ and X04$^-$ microglia have remained unclear.
6 Thus, we examined the chromatin accessibility of X04$^+$ and X04$^-$ cells isolated
7 from 9-month-old 5xFAD brains (Fig. S7B). While 1-2% of microglia in
8 5xFAD/WT brains were X04$^+$, X04$^+$ cells in 5xFAD/IRF8KO brains were much
9 less (<0.1%, Fig. S7C), too few for ATAC-seq experiments. The differential
10 analysis for X04$^+$ microglia revealed that 1,650 gained peaks, while 26,882 lost
11 peaks in X04$^+$ cells relative to X04$^-$ cells (Fig 7B). Motif analysis indicated that
12 lost peaks were enriched with PU.1, and ETS/ISRE composite elements,
13 whereas gained peaks were elements for Osr1 and Rfx3 (Fig. 7C). Gained peak
14 motifs in X04$^+$ cells differed from those in IRF8KO microglia. Nearest gene
15 analysis exhibited no association between a set of DAM and Interferon-related
16 genes and gain of ATAC peaks, in contrast to IRF8KO and IRF8cKO microglia
17 (Fig. 7D). These results suggested that 5xFAD microglia assume chromatin
18 accessibility patterns different from that in IRF8KO and IRF8cKO microglia. This
19 was evident in IGV profiles of *Apoe* and *Sall1* genes where ATAC-seq peaks in
20 IRF8KO and X04$^+$ cells differed (Fig. 7E) and the heatmaps showing scATAC-
21 seq and bulk ATAC-seq signals near DAM genes (Fig. S7D). These data
22 indicate that 5xFAD microglia have epigenome structures unrelated to those in
23 IRF8KO and IRF8cKO cells.
24 We next examined how the loss of IRF8 influences AD pathology.
25 d'Errico et al. recently reported that *Irf8* depletion leads to reduced microglia

motility, resulting in reduced spreading of Aβ plaques[44]. However, whether postnatal *Irf8* deletion affects AD pathology and mechanisms by which Aβ plaque spreads within the brain, and the role of IRF8 in the process remained elusive. We performed histological analyses on four-month and one-year-old 5xFAD/WT and 5xFAD/IRF8KO brains for localization of microglia and Aβ in the cortical region (Fig. 7F). The microglia count per plaque was much fewer in 5xFAD/IRF8KO brain compared to 5xFAD/WT brain, both at four months and one year, the extent being greater at one year (Fig. 7F, 7G). These data are in line with the FACS data that 5xFAD/IRF8KO brain has fewer X04^{+} microglia (Fig. S7C). These results indicate that IRF8KO microglia are unable to recognize Aβ plaques[44]. Significantly, IRF8cKO/5xFAD brain microglia were also defective in interacting with Aβ (Fig. S7E). Our data show that postnatal *Irf8* deletion is sufficient to impair microglia's ability to interact with Aβ.

Additionally, we conducted a study to track the growth of Aβ plaques from 4 months to one year using thioflavin S (ThioS), which detects polymerized Aβ[49]. We found that plaque count and size increased in both 5xFAD/WT and 5xFAD/IRF8KO brains during the period. However, at one year, the size of plagues was much smaller in 5xFAD/IRF8KO brains than in 5xFAD/WT brains. In contrast, the sizes were comparable between 5xFAD/WT and 5xFAD/IRF8KO brains at four months old (Fig. S7F). These data also indicated a defective Aβ plaque growth without IRF8.

Next, we assessed overall neuronal health in 5xFAD/WT and 5xFAD/IRF8KO brains using the NeuN antibody, which detects intact neurons[54]. The NeuN immunoreactivity in 5xFAD/IRF8KO brains was higher than that in the 5xFAD/WT brains at one year, while it was similar at four months (Fig. 7H).

These results indicate that the deletion of IRF8 lessens neuronal damage in the 5xFAD brain.

Discussion

This study aimed to elucidate postnatal microglia development and the role of IRF8 in the process, a question that remained unanswered. We show that IRF8 binds to microglia enhancers in a stepwise fashion during P9-P14 to the adult stage. This coincided with a stepwise increase in chromatin accessibility. However, microglia specific transcriptome expression took place later in adults. Thus, IRF8 binding and setting of open chromatin preceded microglia-specific gene expression, a process consistent with other developmental models[16, 38].

It was striking that a stepwise increase in IRF8 binding was not due to corresponding increase in IRF8 expression, since IRF8 levels were constant during this period and through adulthood, suggesting an epigenome-based mechanism, yet to be deciphered. Many IRF8 peaks were in microglia specific enhancers neighboring microglia identity and CLEAR lysosomal genes, the former include transcription factors, Sall1 and Batf3, that enhance the expression of these gene sets.

IRF8 deletion led to the loss of these enhancers, resulting in a marked reduction in the expression of these gene sets. IRF8 loss led to the formation of alternative enhancers that were near DAM-like genes (e.g., *Apoe*) and interferon-related genes that were upregulated in IRF8KO and IRF8cKO cells. We further show, by genome-wide bisulfite sequencing, that IRF8 sets DNA

methylation patterns specific for microglia, which correlated with ATAC-seq patterns.

scRNA-seq analysis revealed distinct clusters in WT and constitutive IRF8KO microglia, indicating considerable heterogeneity in gene expression patterns, most notable being interferon-related gene expression restricted to one cluster in IRF8KO microglia.

10x Multiome analysis for IRF8cKO microglia was highly informative since we could compare transcriptome profiles and chromatin accessibility patterns for the same set of cells at the single-cell level. Data revealed distinct clusters of IRF8cKO and WT cells. We found marked downregulation of microglia identity gene and CLEAR genes in distinct scRNA-seq clusters, while DAM-like and interferon-related genes were upregulated in the other distinct clusters. scATAC-seq analysis also identified distinct clusters for WT and IRF8cKO cells. A salient finding was that scATAC-seq clusters exhibited good correspondence with scRNA-seq clusters, although some interchangeability was evident. Results suggest that chromatin accessibility defines transcriptome profiles at single-cell level. Certain levels of interchangeability were interesting, as they suggest additional complexity in chromatin accessibility and gene expression. Multiome analysis also enabled us to estimate deletion efficiency in IRF8cKO accurately, which was about 85%. That way, we could remove the contribution of contaminating undeleted cells from our data. We conclude that postnatal IRF8 expression is necessary for microglia to maintain epigenome structure and transcriptional programs for adult microglia.

It has been shown that the microglia from constitutive IRF8KO mice do not spread AD plaques to other regions in the brain, presumably reducing AD

pathology[44]. To gain a mechanistic understanding of the role of IRF8 in AD, We examined (a) 5xFAD/IRF8cKO microglia and (b) development of Aβ plaques from early to late stages of the disease, questions that remained unanswered. Remarkably, microglia's ability to interact with Aβ plaques was lost by constitutive IRF8KO and postnatal *Irf8* deletion, even though deletion was incomplete. Thus, loss of IRF8 impairs microglia's ability to recognize Aβ without delay. Further, we show that Aβ plaques develop during four months to one year in both number and size and that the absence of IRF8 affects mostly the growth of plaque size. These observations add to our understanding of the role of IRF8 in AD pathology further. Lastly, we addressed the relationship between DAM-like gene expression in IRF8KO and 5xFAD brains. ATAC-seq analysis of X04^{+} 5xFAD and IRF8KO microglia displayed distinct patterns of chromatin accessibility unrelated to each other. Thus, DAM-like gene expression found upon *Irf8* deletion does not seem to have functional relevance to AD pathology in the 5xFAD model.

In summary, this study demonstrates the essential requirement of IRF8 in setting the epigenome structure of postnatal microglia and the role of AD progression.

Online Methods

Mice

B6(Cg)-*Irf8*$^{tm1.2Hm}$/J (IRF8KO), B6.Cg-*Irf8*$^{tm2.1Hm}$/J (Irf8-GFP), B6.Cg-*Irf8*$^{tm1.1Hm}$/J (Irf8flox) and control wild-type C57BL/6J mice were maintained in the NICHD animal facility[23]. B6.129S(C)-*Batf3*^{tm1Kmm}/J (Batf3KO),

B6.129P2(Cg)-*Cx3cr1*^{tm1Litt}/J (Cx3cr1^{GFP}), B6.129P2(C)-Cx3cr1$^{tm2.1(cre/ERT2)Jung}$/J (Cx3cr1cre^{ERT2}), and B6SJL-Tg$^{(APPSwFlLon, PSEN1*M146L*L286V)6799Vas}$/Mmjax (5xFAD) mice were purchased from Jackson Laboratory. Cx3cr1^{GFP} mice were crossed with IRF8KO mice. 5xFAD mice were backcrossed to C57BL/6J strain for at least seven generations and maintained hemizygote. To achieve microglia specific Irf8 depletion (IRF8cKO), we crossed Irf8flox mice with Cx3cr1cre^{ERT2} mice. Mice were given 300μg of Tamoxifen (Sigma) for five consecutive days at P12 or 2mg x 5 days at P28, and the Cre recombinase expression in microglia was confirmed using Rosa26$^{LSL-EYFP}$ reporter mice. For *in vivo* amyloid staining, 2mg/kg Methoxy-X04 (TOCRIS, #4920) in a half DMSO/saline (50%DMSO, 0.9%NaCl, pH12.0) was intraperitoneally injected 3hr before microglia isolation. Female mice were used in this study unless specified, and all animal experiments were performed according to the animal study (ASP#17-044, 20-044, and 23-044) approved by the Animal Care and Use Committees of NICHD, NIH.

Microglia preparation

Mice were euthanized by asphyxiating with CO_2 and transcardially perfused with 10 mL of ice-cold PBS. The brain was harvested, chopped with a single-edge razor blade, and ground in ice-cold Hanks-balanced salt solution using a 7ml Dounce homogenizer, then filtered through a 70μm cell strainer. The cells were resuspended in 30% isotonic Percoll and spun down 800xg for 30 min. The floating myelin layer was removed, and the cell pellet was further sorted by FACS AriaIIIμ (BD bioscience) after antibody staining. All reagents and instruments were kept in our 4°C cold room, and all procedures were

performed at less than 4°C to avoid extra activation. For FACS cell sorting, the following antibodies were used: CD11b-Brilliant Violet 421 (CD11b-BV421, clone: M1/70, Biolegend), CD45-APC (clone: 30-F11, Invitrogen). The dead cells were gated out with staining of 7-Aminoactinomycin D (7-AAD; BD pharmingen, 1:100). The live cells were counted with a hemocytometer, which was regarded as the 7-AAD negative subset in FACS for estimating the count of each cell population. To efficiently isolate cells, we took a backgating strategy in which all gates for cell sorting were forced to be set where microglia most abundantly existed. The microglial gate was determined with a 10s pilot run before an actual cell sorting. We used CX3CR1-PE (FAB5825P, R&D), Ly6C-APCCy7 (clone: HK1.4, Biolegend), F4/80-PECy7 (clone: BM8, Biolegend), CCR2-BV510 (clone: SA203G11, Biolegend), 1A8-FITC (BD Pharmingen), Tmem119-Alexa Fluor 488 (ab225497, Abcam), CD206-PE/Dazzle594 (clone: C068C2, Biolegend), FRβ-APC (clone:10/FR2, Biolegend), CD3e-PE (clone:145-2C11, Biolegend), NK1.1-PE (clone: PK136, eBioscience), B220-PE (clone: RA3-6B2, Biolegend) antibodies for the expression analysis and reanalyzed the data with FlowJo v.7.6.5 or 10.10.0 software.

Immunohistochemistry

Mice were anesthetized with 125mg/kg Ketamine and 10mg/kg Xylazine cocktail intraperitoneal injection. The brain was fixed by the transcardial perfusion of 4% PFA/PBS, followed by overnight incubation in an additional 4% PFA/PBS. After removing the fixative, the brain was cryoprotected with a 10-30% sucrose solution. The section was made with 30μm slices. Blocking was achieved with 4% BSA in PBS and 0.3% Triton X-100. Primary antibodies and working dilution

are the following: Mouse anti-β-Amyloid antibody (1:500, clone: 6E10, Biolegend), Rabbit anti-NeuN monoclonal antibody (clone: A60, Millipore), and Rabbit anti-Iba1 antibody (Wako). The tissue was then incubated with the following fluorescence-labeled secondary antibodies: Alexa Fluor 488 Goat anti-mouse IgG (H+L) (Invitrogen), Alexa Fluor 568 Goat anti-mouse IgG (H+L) (Invitrogen), Alexa Fluor 633 Goat anti-mouse IgG (H+L) (Invitrogen), Alexa Fluor 633 Goat anti-rabbit IgG (H+L) (Invitrogen). The core-dense amyloid plaques were stained with 0.5% Thioflavin S (Sigma, T1892). The nuclei were stained with Hoechst33342 dye. For each brain section, z-stack images were taken on a Zeiss AiryScan 880 confocal microscope using either a 10x or a 40x air objective, together with the Airyscan system and ZEN lite software (Zeiss). Alexa Fluor 488/Thioflavin S, Alexa Fluor 633 and Alexa Fluor 568, and Hoechst33342 dyes were excited using the 488nm, 633nm, 561nm and 405nm laser lines. The emitted fluorescence was detected using the corresponding channels on the Airyscan detectors. Raw data were further analyzed with ImageJ/Fiji[50]. Maximum intensity projection for the z-stacked images was presented, and the signal from each channel was pseudocolored as specified in the figures. The amyloid burden, plaque size, and count were quantified with five 500 μm² fields randomly chosen from each section. To quantify microglia gathered with a plaque, we randomly picked ten plaques per mouse that were positive for 6E10. We then counted the number of Iba1-positive microglia colocalized with each plaque[44]. The nuclei were stained with Hoechst33342 dye to distinguish Iba1-positive microglia as individual cells. The average microglia count per plaque was then calculated and presented for each mouse[44]. For NeuN-positive neuron count, five 500μm x 500μm fields

per section were randomly chosen in the cortical layer V and analyzed. The average value of 5 fields was presented for each mouse.

Bulk RNA-Seq

Fifteen thousand cells were sorted out into TRIzol and stored at -70°C until library preparation. Total RNA was purified by phase extraction and processed to cDNA using SMART-Seq v4 Ultra Low Input RNA Kit (Takara bio USA, CA) with 15-cycle amplification. Adaptors were added to 1ng cDNA using the Nextera XT library prep kit (Illumina, CA). The libraries were sequenced on Nextseq500 with 55bp paired end read. For sequence data analysis, adapter sequences were trimmed with Trimmomatic v0.33 after the initial quality check, and then data was aligned with STAR v2.5.2b. Sequence duplicates were removed using Picard v2.17.11. Aligned tags were counted with Subread featureCounts, and differentially expressed genes were estimated using DESeq2[51] with FDR<0.01. Batch effects were normalized at this point if needed. Heatmaps were depicted with Morpheus (https://software.broadinstitute.org/morpheus).

Single-cell RNA-seq for constitutive Irf8 knockout microglia

Ten thousand cells were sorted out from 3 mice pools and immediately emulsified and subjected to the libraries using a 10x Genomics Chromium kit v2 as the manufacturer's protocol. The libraries were sequenced 26bp for R1 and 98bp for R2 on HiSeq2000. The fastq files were first processed with Cellranger v3.1 pipeline to make gene expression matrices, and the following analyses were performed with Scanpy[52]. Cells with less than 500 or greater than 6,000

detected genes were filtered from analysis. Genes detected in less than three cells were also removed from the data. The feature counts were normalized to 10,000 reads per cell and logarithmized. The data were regressed out based on the highly variable genes marked with *min_mean=0.0125, max_mean=3, min_disp=0.5* options, and the cells exhibiting high standard deviation (SD>15) were excluded from the analysis. UMAP dimension reduction and Leiden clustering were employed to visualize the data with *n_neighbors=80, n_pcs=15,* and *min_dist=1.0* parameters. Each cluster's connectivity was computed using PAGA[42] with the default parameters.

Single-cell multiome for conditional Irf8 knockout microglia

Forty thousand male microglia were isolated from each mouse and immediately emulsified and subjected to the libraries using a Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent as the manufacturer's protocol. The ATAC libraries were sequenced 50bp for R1, 8bp for the i7 index, 24bp for the i5 index, and 49bp for R2 on NovaSeq6000. The RNA libraries were sequenced 28bp for R1, 10bp for i7, 10bp for i5, and 90bp for R2. The fastq files were processed with Cellranger-arc pipeline to create a single sequence aggregate, and the following analyses were performed with Muon[53].

Cells with low-quality sequences were filtered from the analysis. To that end, the following criteria were used: less than 200 detected transcripts, less than 500 unique molecular identifiers (UMIs), or >1.0% mitochondrial gene contamination for scRNA-seq, peaks within less than 10 cells, cells with less than 1000 or greater than 80,000 UMIs, cells with less than 100 or greater than

30,000 peak regions detected, for scATAC-seq. The feature counts were log-normalised, the same as the above scRNA-seq analysis. The data were regressed out based on the highly variable genes marked with *min_mean=0.02, max_mean=4, and min_disp=0.5* options. ATAC highly variable counts were filtered with the Cellranger flavor option. For 14,391 cells, UMAP dimension reduction and Leiden clustering were employed to visualize the data with *n_neighbors=100, n_pcs=50,* and *min_dist=1.0* parameters for scRNA-seq, *n_neighbors=100, n_pcs=30,* and *min_dist=1.0* for scATAC-seq. The transcriptome and chromatin accessibility data with identical barcodes were then integrated. The differentially accessible peaks were identified by the student's t-test combined with Benjamini-Hochberg correction, and peaks with adjusted p-value < 0.01 were considered statistically significant. HOMER was used for motif analysis and peak overlap identification. We utilized ArchR[54] to generate pseudobulk bigwigs, which were normalized based on the number of fragments.

CUT&RUN

CUT&RUN methodology[55] was used with some modifications. Protein A-bound micrococcal nuclease (pA-MNase) was kindly provided by Steven Henikoff. For histone marks, live microglial cells were sorted out and washed once with HEPES-NaCl buffer (HEN buffer; 20mM HEPES pH7.5, 150mM NaCl, 500μM Spermidine, Protease inhibitors cocktail), then counted. Ten thousand cells were immobilized to activated BioMag®Plus Concanavalin A (ConA)-beads (Bangs Laboratories, IN) in the presence of Mn^{2+} and Ca^{2+} ions (0.1mM $MnCl_2$, 0.1mM $CaCl_2$) at room temperature. The beads complex was then

1 permeabilized using 0.005% Digitonin (Merck)-containing HEN buffer (Digitonin
2 buffer) and incubated for 12hrs at 4°C with 0.25μg of the following antibodies in
3 the presence of 2mM EDTA: rabbit anti-H3K27ac (ab4729, Abcam), rabbit anti-
4 H3K4me1 (ab8895, Abcam). After antibody binding, beads were washed twice
5 with Digitonin buffer and bound pA-MNase. After the washout of extra pA-
6 MNase, beads-immunocomplex was activated in the presence of 3mM $CaCl_2$ at
7 0°C for 30 min. The digestion was stopped by adding 10x Stop buffer (1.70M
8 NaCl, 100mM EDTA, 20mM EGTA, 0.25mg/ml Glycogen, 0,25μg/μl RNaseA)
9 and 15pg MNase-digested yeast DNA spike-in control. The nuclei were digested
10 at 56°C for five hours after adding 40μg of Proteinase K and SDS at the final
11 concentration of 0.1%. The immunoprecipitated DNA was purified by phase
12 separation and removed undigested DNA with a half volume of AMPure beads.
13 For library preparation, purified DNA was processed using SMARTer®
14 ThruPLEX DNA-Seq Kit (TAKARA Bio USA, CA) as the manufacturer's protocol
15 except for PCR cycles to avoid amplifying long fragments (13 cycles, 10s for
16 extension). For transcription factors, the antibody-labeled brain homogenate
17 was fixed with 1% Paraformaldehyde-PBS at room temperature for 10 minutes
18 right before cell sorting. Sorted cells were immediately lysed with nuclear
19 extraction buffer (20mM HEPES pH7.5, 10mM KCl, 500μM Spermidine, 0.1%
20 Triton X-100, 20% Glycerol, Protease Inhibitor cocktail) and counted. Thirty-
21 thousand nuclei were immobilized with ConA-beads in the same nuclear
22 extraction buffer and blocked using Wash buffer (20mM HEPES pH7.5, 150mM
23 NaCl, 500μM Spermidine, 0.1% Bovine Serum Albumin, 1% Triton X-100,
24 0.05% SDS, Protease inhibitor cocktail) containing 4mM EDTA. The nuclei were
25 incubated with a rabbit anti-GFP antibody (0.5μg/reaction, ab290, Abcam) or a

rabbit anti-mouse PU.1 (EPR22624-20, Abcam) for 12hrs at 4°C. After binding to pA-MNase in Wash buffer, DNA was cleaved by adding 3mM $CaCl_2$ for 30min at 0°C. MNase was inactivated by adding chelators mixture at the final concentration of 170mM NaCl, 10mM EDTA, 2mM EGTA, 0.005% NP-40, 25μg/ml Glycogen, and 15pg yeast spike-in control. Total DNA was extracted, and de-crosslinked by incubating overnight at 65°C in the high-salt buffer containing 0.1% SDS, 0.5mg/ml Proteinase K and 0.1mg/mL RNase A. The undigested DNA was removed as above, and the small fragments for sequencing libraries were harvested. The final libraries were purified by AMPure beads and sequenced with a 30bp paired end read. For data analysis, the adapter and low-quality sequences were removed from raw fastq data using TrimGalore. Then the data was aligned on mm10 mouse genome using Bowtie2 with *--local --very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700 --no-overlap --no-dovetail* options. Peak calling was performed by MACS2 with *--broad --keep-dup all -p 0.01 --min-length 1000* option for a broad histone mark, *--keep-dup all -q 0.01 –max-gap 640* option for a narrow histone mark and IRF8[56]. For PU.1, we employed *-q 0.05* instead because it is more consistent with the previous dataset (GSM1533906)[33]. The peak co-occurrence, genomic position, and motif analysis were performed using HOMER[34] with *-d given* option. To determine highly concordant peaks across biological replicates, we employed “the Majority rule,” which picks up only peaks found within more than 50% of replicates[57]. The differential peaks were obtained using Diffbind-DESeq2[58] with FDR<0.05 cut-off. Then all bam files were merged and used for IGV visualization and depicting average plots and heatmaps with Deeptools[59].

1 H3K27ac^{high} large regions were identified as “super-enhancer regions” of the
2 ROSE’s readouts with default parameters[35].
3
4 *Bulk ATAC-Seq*
5 ATAC-Seq libraries were constructed as described before[37]. Three
6 thousand (X04 microglia) to fifty thousand isolated microglia were lysed in lysis
7 buffer (10mM Tris-HCl pH7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% NP-40). Following
8 lysis, nuclei were tagmented using the Nextera DNA library prep kit (Illumina) for
9 30 minutes. Then, DNA was purified and amplified for 9-14 cycles based on the
10 SYBR-qPCR determination. The libraries were sequenced with a 50bp paired
11 end read. After removing the adapter and low-quality sequences, the data were
12 aligned to the mm10 genome with Bowtie2. Peaks were called on aligned
13 sequences using MACS2 with *-p 0.01 --nomodel --shift -37 --extsize 73 -B --*
14 *SPMR --keep-dup* all options. Peaks were then evaluated with less than 0.01
15 irreversible differential rates. Differentially accessible regions were identified
16 using Diffbind-EdgeR with FDR<0.05[58]. Deeptools and HOMER were employed
17 for further analyses, including motif enrichment analysis with the indicated
18 options.
19
20 *Whole-genome bisulfite sequence*
21 Genomic DNA was extracted from 40,000 male microglial cells using
22 Quick-DNA Microprep plus kit (Zymo Research) as the manufacturer’s protocol.
23 Then, 50ng DNA was bisulfite-converted and processed to the library using Pico
24 Methyl-Seq Library Prep Kit (Zymo Research), with 8-cycle PCR amplification.
25 The libraries were sequenced with 75bp paired end read on HiSeq3000 in the

NHLBI sequencing core (9.2x coverage on average). Based on the Bismark M-bias plot, the adapters, low-quality sequence, and an additional 10bp of 5'-end and 2bp of 3'-end were removed from the sequencing data using TrimGalore. The data were first aligned using Bismark[40] with *-q --score-min L,0,-0.4 --ignore-quals --no-mixed --no-discordant --dovetail --maxins* options, and then unmapped sequences were aligned additionally with single-end mode. After deduplication and methylation calls, both files were merged into a single coverage file. Differential analyses were performed using R/DSS package[41] using FDR<0.05 for the differentially methylated loci (DMLs) and p<0.001 for the regions (DMRs).

Data availability: All high-throughput sequence datasets generated in this paper are available in GSE231406 and GSE266424. Some gene sets were obtained from publications as denoted individually.

Acknowledgments: We thank Steven Coon and James Iben, NICHD DIR MGC, for their guidance for scRNAseq and multiome, Yuesheng Li in NHLBI sequencing core for whole-genome bisulfite sequencing, and Fayaz Seifuddin for data analysis. We also thank Avinash Bhandola for providing Rosa26-EYFP mice, Lisa William-Simons for cell sorting, and the NICHD animal facility staff for caring for our mice and for other technical assistance. We gratefully acknowledge colleagues in the NIH and elsewhere for valuable advice and discussions. This work was supported by the NICHD Intramural programs ZIA HD008015-13.

Disclosure: There is no conflict of interest.

Author Contributions: K.S and K.O conceived the framework of this work; K.S generated IRF8cKO microglia, devised whole genome analyses, and executed experiments along with R.P; E.L performed immunohistology; D.K. provided advice on experiments; K.S and K.O wrote the manuscript.

1 Figure Legends

2 Figure 1 Microglia specific IRF8 occupancy increases postnatally

3 A. Genome wide distribution of IRF8 peaks in adult microglia (top) and
4 peritoneal macrophages (bottom). The genomic position of each peak
5 was determined with HOMER.

6 B. HOMER *de novo* motif analysis for IRF8 binding sites in microglia. Motifs
7 with a match score >0.85 are shown. A motif "Sfpi" is equivalent to the
8 PU.1 motif.

9 C. An M-A plot visualizing the difference between microglia IRF8 peak
10 intensity and that of peritoneal macrophages (n=3). The differential peaks
11 with FDR<0.05 are indicated in pink.

12 D. Top 10 Gene ontology terms for microglia specific IRF8 peaks (top) and
13 macrophage specific IRF8 peaks (bottom), analyzed with GREAT.
14 Sequences covering 500kb distal from the TSS to 1kb from TES were set
15 as an associating region. The terms of the GO biological process and the
16 corresponding binomial FDR values were shown.

17 E. The number of IRF8 peaks identified in P9, P14, and adult microglia.

18 F. Schematic illustration of the distance between two neighboring IRF8
19 peaks; the shortening peak intervals represent an array-like appearance
20 of peaks (left). The CDF plot shows that the frequency of short interval
21 peaks increases along with the postnatal microglia development and
22 more than peritoneal macrophages. (right; $p=2.2x10^{-16}$, Kolmogorov-
23 Smirnov two-sided Test).

G. IGV examples of IRF8 peaks in developing and adult microglia and peritoneal macrophages at microglia specific genes (*Sall1* and *Cx3cr1*; left) or macrophage specific genes (*Clec4n*, *Lyz1*, and *Lyz2*; right).

Figure 2 Identification of IRF8-driven transcription factor cascades in shaping microglia-specific transcriptional programs

A. Volcano plot for RNA-seq comparing WT and IRF8KO microglia. Genes upregulated in IRF8KO microglia are indicated in pink (n=2,356), and genes downregulated in IRF8KO cells are in blue (n=2,682, FDR<0.01). Representative genes are shown with a label.

B. Heatmaps for indicated gene sets. "Microglia identity" and "Interferon-related" gene sets were referred to as "cluster 6: Microglia" and "Interferon-Related" in the previous publication (ref.11). The DAM-like (ref.29) and CLEAR lysosome gene set were obtained from elsewhere (ref.27).

C. Transcription factors downregulated in IRF8KO microglia (FDR<0.01 cut-off) are shown. A list of transcription factor genes was obtained from the previous publication (ref.13). Genes that show more than four log-scale reductions in IRF8KO cells are boxed, and Sall1 and Batf3 in the blue box were further investigated.

D. Volcano plot for RNA-Seq data comparing WT and Batf3KO microglia. The genes upregulated in Batf3KO cells are shown in pink (n=172) and downregulated genes in blue (n=274, FDR<0.01). The representative genes are indicated in the plot.

E. Correlation biplot showing that Batf3 dependent genes correlate with those dependent on IRF8. The DEGs from Batf3KO microglia RNA-Seq were analyzed with the Kendall rank correlation test (r=0.19, p=2.4x10^{-9}). Microglia identity genes shared between Batf3KO DEGs and IRF8KO DEGs are labeled red.

F. Correlation biplot showing that Sall1 dependent and IRF8 dependent genes overlap. All up and down DEGs from Sall1KO microglia RNA-Seq obtained from the previous publication (ref.25). were analyzed with the Kendall rank correlation test (r=0.23, p=5.6x10^{-39}). The microglia identity genes shared between Sall1KO DEGs and IRF8KO DEGs are labeled red, and those of interferon-related genes are blue.

G. IRF8 peaks overlapped with SALL1 and PU.1 binding sites. The SALL1 and PU.1 binding profiles were obtained from GSE226090 (ref.32) and GSM1533906 (ref.33), respectively.

H. A model illustrating IRF8 dependent transcription cascade. IRF8 induces expression of downstream TFs, BATF3 and SALL1. These TFs, along with IRF8, then activate indicated target genes. This dual cascade would cover all four gene sets, microglia identity, interferon related, DAM-like, and CLEAR lysosomal genes.

Figure 3 Adult IRF8 is an active enhancer that establishes an enhancer landscape in adult microglia

A. A Venn diagram showing the overlap among IRF8 peaks, H3K4me1, and H3K27ac marks.

B. Heatmaps showing H3K27ac (left) and H3K4me1 (right) signal intensity centered on IRF8 peaks ± 3.0kb in adult WT and IRF8KO microglia. Shown above heatmaps are average peak intensities.

C. H3K27ac ranked super enhancers in adult WT microglia. IRF8 dependent genes neighboring super enhancers are indicated in red.

D. Venn Diagram depicting the overlap of H3K27ac super enhancers and IRF8 peaks in adult WT microglia.

E. IRF8 peak counts of adult microglia, P9 microglia, and peritoneal macrophages in the H3K27ac^{high} regions.

F. Representative IGVs visualizing the distribution of IRF8 peaks, H3K4me1, and H3K27ac marks near/at Sall1 (top) and Batf3 genes (bottom). An H3K27ac^{high} region is marked by a red bar. IRF8 is bound densely to the Batf3 regulatory region (orange bar), distinct from super enhancers. See also Fig.S3C-G for further description.

Figure 4 IRF8 confers chromatin accessibility on postnatal microglia and sets DNA methylation patterns

A. MA plot showing ATAC-seq signals comparing WT and IRF8KO microglia (n=2). The differentially accessible regions (FDR<0.05) are colored in pink. The number of gained or lost accessible regions in IRF8KO microglia was shown at the corresponding corner of the area.

B. The genome-wide distribution of gained and lost ATAC-seq peaks as determined by HOMER.

C. HOMER *de novo* motif analysis for gained and lost ATAC-seq peaks. The top 4 motifs are shown.

D. The number of IRF8, H3K27ac, and H3K4me1 peaks (WT or IRF8KO specific) overlapped with lost or gained ATAC-seq peaks. See Fig.S3A for the WT or IRF8KO specific peak identification.

E. Volcano plot showing a correlation between ATAC-seq peaks and RNA-seq data. Microglia identity genes (representative "down" genes) and interferon-related genes (representative "up" genes) are labeled red and blue, respectively. The nearest genes for gained or lost ATAC-seq peaks correlate with genes up or down-regulated in IRF8KO microglia.

F. The number of differentially methylated single nucleotide loci (DML; FDR<0.05) and the differential regions (DMR; p<0.001). Numbers in the graph are differentially methylated and demethylated loci/regions in WT cells.

G. Genome-wide distribution of the differentially demethylated (top) and methylated (bottom) regions identified with HOMER.

H. Numbers of IRF8, H3K27ac, and H3K4me1 (WT or IRF8KO specific) peaks in differentially methylated or demethylated regions.

I. Representative IGVs visualizing IRF8. H3K4me1, H3K27ac, ATAC-seq, PU.1 peaks and methylated DNA (CpG) status on Sall1 (top), Batf3 (middle), and Apoe (bottom) in WT and IRF8KO cells. Methylated and demethylated regions are indicated by black bars.

Figure 5 scRNA-seq analysis reveals transcriptome heterogeneity and the presence of a progenitor like population

A. UMAP derived from WT and IRF8KO microglia aggregates. Leiden clusters were colored and labeled as indicated. The data were obtained from the pool of three individual WT and IRF8KO microglia preparations.
B. Percentages of WT and IRF8KO genotypes in indicated clusters.
C. UMAP presentation of indicated gene sets (compare with heatmaps in Fig. 2B).
D. Gene expression heatmap showing the expression of microglia identity gene set in each WT and IRF8KO cluster. Note that expression of all identity genes is absent in IRF8KO Cluster 6.
E. PAGA network depicted with the default parameters. Each node indicates the corresponding Leiden cluster.
F. Rank plot aligning the top 20 genes by scoring expression level. The genes in the Common cluster were compared to those of all the other clusters.
G. Heatmap displaying the correlation coefficients between Common cluster and other clusters for IRF8, BATF3, or SALL1 dependent gene sets.

Figure 6 Postnatal deletion of *Irf8* causes gene expression and chromatin accessibility changes analogous to constitutive IRF8KO microglia

A. The strategy for conditional Irf8 deletion in postnatal microglia. P12 $Irf8^{flox/flox}Cx3cr1Cre^{ERT2/+}Rosa26^{YFP}$ mice were injected with Tamoxifen 5 times, followed by eight weeks of rest. Microglia in adults were isolated and used for bulk RNA-seq and single-cell multiome analysis.

B. RNA-guided Leiden clusters on UMAP displaying single-cell multiome sequencing data obtained from WT and IRF8cKO microglia (left). The cells from WT mice were colored yellow in the right panel.

C. Frequency of WT and IRF8cKO cells in indicated RNA-guided clusters.

D. Heatmap presentation of indicated gene sets (compare with heatmaps in Fig. 2B).

E. ATAC-guided Leiden clusters of WT and IRF8cKO microglia, unsupervised with scRNA-seq data (left). The cells from WT mice were colored yellow in the right panel.

F. Frequency of WT and IRF8cKO cells in indicated ATAC-guided clusters.

G. Percentages occupied by the indicated ATAC-guided cluster cells in WT and IRF8cKO-specific RNA-guided clusters.

H. Volcano plot displaying differentially accessible regions. The more accessible regions in WT-specific clusters (WT_atacCL1-4) were colored in pink, whereas those in IRF8cKO-specific clusters (IRF8cKO_atacCL1-4) were in blue. The set of genes presented in Fig.2B was labeled in the four right panels as the nearest genes from peaks, respectively.

I. *De novo* motif analysis for differentially accessible regions presented in Fig. 6H

J. Volcano plot displaying differentially accessible regions comparing IRF8cKO ATAC cluster 1 and IRF8 ATAC cluster 2+3. DAM-like (left) and Interferon-related genes (right) presented in Fig. 2B were labeled, respectively.

K. Representative IGVs of *Sall1*, *Batf3,* and *Apoe* regions. The differential accessible regions were bracketed. IRF8cKO microglia clusters exhibit

heterogeneity in the expression of DAM-like and interferon-related genes, which is evident in the *Apoe* bracketed region.

Figure 7 IRF8KO microglia lessens AD pathology in the 5xFAD mouse model, despite spontaneous DAM/NDG gene induction

A. Heatmap showing the 5xFAD associated gene set expression in WT, 5xFAD/WT, IRF8KO, and 5xFAD/IRF8KO microglia. Data are from RNA seq analysis of microglia from nine months old mice. Data were clustered by K-mean methodology (k=2), and representative genes were labeled on the right.
B. An M-A plot showing ATAC-seq data comparing MethoxyX04-positive and -negative microglia from 10 months old 5xFAD mouse (n=2). The differentially accessible regions (FDR<0.05) are colored in pink. The number of gained or lost accessible regions in MethoxyX04-positive microglia was shown at the corresponding corner of the area.
C. *De novo* motif analysis for differentially accessible regions presented in Fig. 7B
D. Volcano plot showing gene expression profiles comparing 5xFAD/WT and non-5xFAD control (same data as Fig.S7A with rescaled). The genes that are nearest to the gained peaks in $X04^+$ microglia were plotted in blue. Among them, significantly upregulated genes in 5xFAD microglia (referred to as 5xFAD-associated genes) were highlighted in red and labeled. The gain of ATAC in $X04^+$ microglia exhibits some connection with transcriptome in 5xFAD microglia rather than DAM-like genes.

E. Representative IGVs of *Apoe* and *Sall1* regions. The X04^{+} microglia do not show an increase of chromatin accessibility at the promoter region of *Apoe* gene (bracketed), whereas IRF8KO or IRF8cKO microglia exhibited a significant increase in that region. For the *Sall1* region, the accessibility was lost entirely in IRF8KO, while some were retained in IRF8cKO and X04^{+} microglia in a cell-type-specific manner (see bracketed regions).

F. Representative histology images depicting immunostaining of the cortex regions of 4 months (top) and 1-year-old (bottom), 5xFAD/WT or 5xFAD/IRF8KO brain with 6E10 for Aβ, Iba1 for microglia, and Hoechst33342 for nuclei, respectively, to quantify microglia colocalization with plaques. Merge images on the right illustrate a lack of extensive colocalization in the 5xFAD/IRF8KO brain. On the right are enlarged images from the bracketed regions on the left. Scale bar; 50 μm.

G. Quantification of microglia colocalized with plaques in Fig.7F. The data was statistically tested using one-way ANOVA followed by Tukey's post hoc test ($F(3,20)=2.788$, $p<0.001$). Graph bar; mean +/- SEM.

H. Representative histology images of NeuN staining in the cortex area of 4 months (top) and a year-old (bottom) 5xFAD/WT or 5xFAD/IRF8KO brain. Scale bar; 200 μm. On the right are enlarged images from the bracketed regions on the left.

I. Quantification of NeuN positive cells. Values are the average of five fields per brain. The data was statistically tested using one-way ANOVA followed by Tukey's post hoc test ($F(3,33)=2.570$, $p<0.001$). Graph bar; mean +/- SEM.

1

2 Figure 8 Model for the role of IRF8 in postnatal microglia development

3 At the early postnatal stage (P9 microglia), IRF8 binding to the microglia

4 genome is sparse (Top). After P14 through adulthood, IRF8 binding becomes

5 denser, forming arrays along with histone modifications (Middle). Full IRF8

6 binding leads to the formation of microglia-specific enhancers along with Sall1

7 and PU.1, opens relevant regions of chromatin and stably demethylates

8 enhancer regions of DNA (bottom), which, taken together, establishes

9 transcriptional programs in microglia.

10

11

12

References

1. Rivest, S. Regulation of innate immune responses in the brain. *Nat Rev Immunol* **9**, 429-439 (2009).

2. Prinz, M., Jung, S. & Priller, J. Microglia Biology: One Century of Evolving Concepts. *Cell* **179**, 292-311 (2019).

3. Ueno, M. *et al.* Layer V cortical neurons require microglial support for survival during postnatal development. *Nat Neurosci* **16**, 543-551 (2013).

4. Fourgeaud, L. *et al.* TAM receptors regulate multiple features of microglial physiology. *Nature* **532**, 240-244 (2016).

5. Fantin, A. *et al.* Tissue macrophages act as cellular chaperones for vascular anastomosis downstream of VEGF-mediated endothelial tip cell induction. *Blood* **116**, 829-840 (2010).

6. Schafer, D.P. *et al.* Microglia sculpt postnatal neural circuits in an activity and complement-dependent manner. *Neuron* **74**, 691-705 (2012).

7. Wang, X. *et al.* Requirement for Microglia for the Maintenance of Synaptic Function and Integrity in the Mature Retina. *J Neurosci* **36**, 2827-2842 (2016).

8. Reshef, R. *et al.* The role of microglia and their CX3CR1 signaling in adult neurogenesis in the olfactory bulb. *Elife* **6** (2017).

9. Nimmerjahn A, K.F., Helmchen F. . Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo. *Science* **308(5726)**, 1314-1318 (2005).

10. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* **169**, 1276-1290 e1217 (2017).

11. Friedman, B.A. *et al.* Diverse Brain Myeloid Expression Profiles Reveal Distinct Microglial Activation States and Aspects of Alzheimer's Disease Not Evident in Mouse Models. *Cell Rep* **22**, 832-847 (2018).

12. Kierdorf K, E.D., Goldmann T, Sander V, Schulz C, Perdiguero EG, Wieghofer P, Heinrich A, Riemke P, Hölscher C, Müller DN, Luckow B, Brocker T, Debowski K, Fritz G, Opdenakker G, Diefenbach A, Biber K, Heikenwalder M, Geissmann F, Rosenbauer F, Prinz M. . Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat Neurosci.* **16(3)**, 273-280 (2013).

13. Matcovitch-Natan, O. *et al.* Microglia development follows a stepwise program to regulate brain homeostasis. *Science* **353**, aad8670 (2016).

14. Bennett, M.L. *et al.* New tools for studying microglia in the mouse and human CNS. *Proc Natl Acad Sci U S A* **113**, E1738-1746 (2016).

15. Hammond, T.R. *et al.* 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 (2019).

16. Mancino, A. *et al.* A dual cis-regulatory code links IRF8 to constitutive and inducible gene expression in macrophages. *Genes Dev* **29**, 394-408 (2015).

17. Das, A., Chauhan, K.S., Kumar, H. & Tailor, P. Mutation in Irf8 Gene (Irf8(R294C)) Impairs Type I IFN-Mediated Antiviral Immune Response by Murine pDCs. *Front Immunol* **12**, 758190 (2021).

18. Kurotaki D, O.N., Nishiyama A, Yamamoto M, Ban T, Sato H, Nakabayashi J, Umehara M, Miyake N, Matsumoto N, Nakazawa M, Ozato K, Tamura T. . Essential role of the IRF8-KLF4 transcription factor cascade in murine monocyte differentiation. *Blood.* **7;121(10)**, 1839-1849. (2013).

19. Gupta, M. *et al.* IRF8 directs stress-induced autophagy in macrophages and promotes clearance of Listeria monocytogenes. *Nat Commun* **6**, 6379 (2015).

20. Holtschke T, L.J., Kanno Y, Fehr T, Giese N, Rosenbauer F, Lou J, Knobeloch KP, Gabriele L, Waring JF, Bachmann MF, Zinkernagel RM, Morse HC 3rd, Ozato K, Horak I. . Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene. *Cell.* **18;87(2)**, 307-317. (1996).

21. Masuda, T. *et al.* Specification of CNS macrophage subsets occurs postnatally in defined niches. *Nature* **604**, 740-748 (2022).

22. Masuda, T. *et al.* IRF8 is a critical transcription factor for transforming microglia into a reactive phenotype. *Cell Rep* **1**, 334-340 (2012).

23. Yoshida, Y. *et al.* The transcription factor IRF8 activates integrin-mediated TGF-beta signaling and promotes neuroinflammation. *Immunity* **40**, 187-198 (2014).

24. Wang, H. *et al.* A reporter mouse reveals lineage-specific and heterogeneous expression of IRF8 during lymphoid and myeloid cell differentiation. *J Immunol* **193**, 1766-1777 (2014).

25. Buttgereit, A. *et al.* Sall1 is a transcriptional regulator defining microglia identity and function. *Nat Immunol* **17**, 1397-1406 (2016).

26. Brioschi, S. *et al.* A Cre-deleter specific for embryo-derived brain macrophages reveals distinct features of microglia and border macrophages. *Immunity* **56**, 1027-1045 e1028 (2023).

27. Palmieri, M. *et al.* Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Hum Mol Genet* **20**, 3852-3866 (2011).

28. Prater, K.E. *et al.* Human microglia show unique transcriptional changes in Alzheimer's disease. *Nat Aging* **3**, 894-907 (2023).

29. Chen, Y. & Colonna, M. Microglia in Alzheimer's disease at single-cell level. Are there common patterns in humans and mice? *J Exp Med* **218** (2021).

30. Tamura, T. *et al.* IFN regulatory factor-4 and -8 govern dendritic cell subset development and their functional diversity. *J Immunol* **174**, 2573-2581 (2005).

31. Hildner K, E.B., Purtha WE, Diamond M, Matsushita H, Kohyama M, Calderon B, Schraml BU, Unanue ER, Diamond MS, Schreiber RD, Murphy TL, Murphy KM. . Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity. *Science* **14;322(5904):1097-100.** (2008).

32. Fixsen, B.R. *et al.* SALL1 enforces microglia-specific DNA binding and function of SMADs to establish microglia identity. *Nat Immunol* **24**, 1188-1199 (2023).

33. Gosselin, D. *et al.* Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. *Cell* **159**, 1327-1340 (2014).

34. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).

35. Whyte, W.A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319 (2013).

36. Dey, A. *et al.* BRD4 directs hematopoietic stem cell development and modulates macrophage inflammatory responses. *EMBO J* **38** (2019).

37. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. & Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218 (2013).

38. Kurotaki, D. *et al.* Chromatin structure undergoes global and local reorganization during murine dendritic cell development and activation. *Proc Natl Acad Sci U S A* **119**, e2207009119 (2022).

39. Deaton, A.M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev* **25**, 1010-1022 (2011).

40. Krueger, F. & Andrews, S.R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571-1572 (2011).

41. Feng, H., Conneely, K.N. & Wu, H. A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Res* **42**, e69 (2014).

42. Wolf, F.A. *et al.* PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol* **20**, 59 (2019).

43. Yona, S. *et al.* Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* **38**, 79-91 (2013).

44. d'Errico, P. *et al.* Microglia contribute to the propagation of Abeta into unaffected brain tissue. *Nat Neurosci* **25**, 20-25 (2022).

45. Zhou, Y. *et al.* Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med* **26**, 131-142 (2020).

46. Oakley, H. *et al.* Intraneuronal beta-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in amyloid plaque formation. *J Neurosci* **26**, 10129-10140 (2006).

47. Ulland, T.K. & Colonna, M. TREM2 - a key player in microglial biology and Alzheimer disease. *Nat Rev Neurol* **14**, 667-675 (2018).

48. Grubman, A. *et al.* Transcriptional signature in microglia associated with Abeta plaque phagocytosis. *Nat Commun* **12**, 3015 (2021).

49. Rajamohamedsait, H.B. & Sigurdsson, E.M. Histological staining of amyloid and pre-amyloid peptides and proteins in mouse tissue. *Methods Mol Biol* **849**, 411-424 (2012).

50. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).

51. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

52. Wolf, F.A., Angerer, P. & Theis, F.J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15 (2018).

53. Bredikhin, D., Kats, I. & Stegle, O. MUON: multimodal omics analysis framework. *Genome Biol* **23**, 42 (2022).

54. Granja, J.M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**, 403-411 (2021).

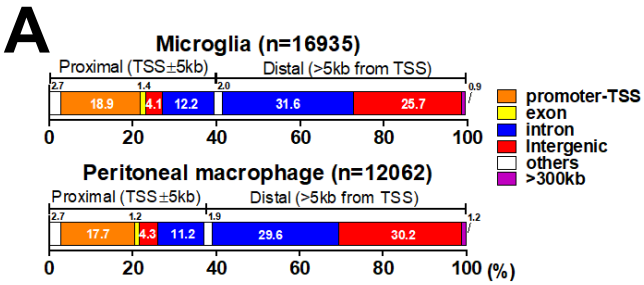
55. Skene, P.J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6** (2017).

56. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

57. Yang, Y. *et al.* Leveraging biological replicates to improve analysis in ChIP-seq experiments. *Comput Struct Biotechnol J* **9**, e201401002 (2014).

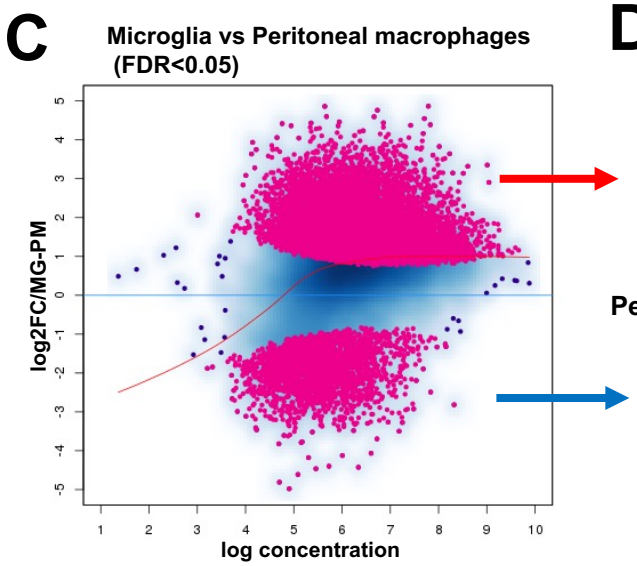
58. Ross-Innes, C.S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389-393 (2012).

59. Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-191 (2014).



B

Rank	Motif	p-value	%T/B	TF
1		1e-1765	48.06 /17.76	Sfpi1 (0.969)
2		1e-378	29.60 /16.67	IRF8 (0.881)



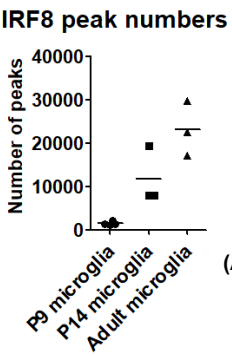
D

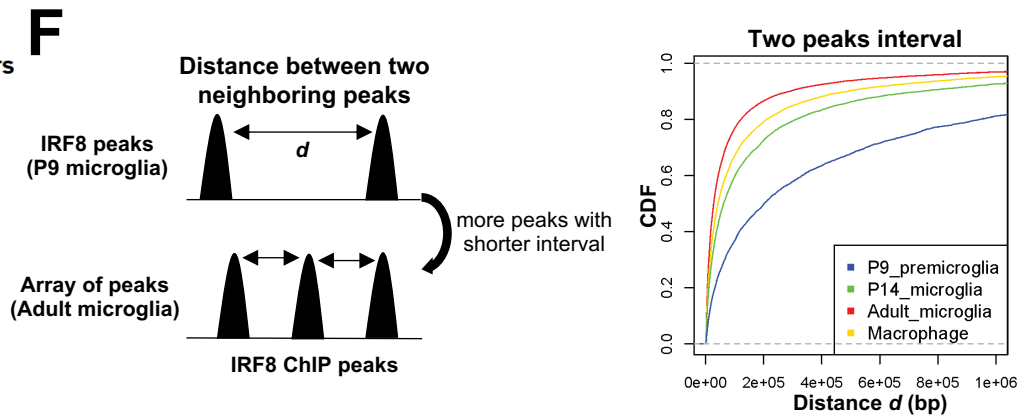
Microglia specific IRF8 (n=9679)

# Term Name	FDR
positive regulation of immune system process	1.19210e-111
cell activation	3.09162e-99
leukocyte activation	4.62254e-84
regulation of immune response	1.59865e-77
regulation of cell activation	5.56143e-77
regulation of lymphocyte activation	2.94470e-69
regulation of leukocyte activation	1.50084e-68
regulation of hemopoiesis	4.11654e-67
regulation of leukocyte differentiation	3.70227e-66
leukocyte differentiation	5.52185e-61

Peritoneal macrophage specific IRF8 (n=2251)

# Term Name	FDR
regulation of immune system process	3.99626e-40
regulation of MAPK cascade	8.16172e-29
inflammatory response	1.53254e-28
response to molecule of bacterial origin	4.08295e-28
defense response	4.10928e-28
positive regulation of immune system process	9.64933e-27
response to other organism	3.52883e-26
response to bacterium	4.63784e-26
response to external biotic stimulus	4.65288e-26
response to lipopolysaccharide	1.25573e-25





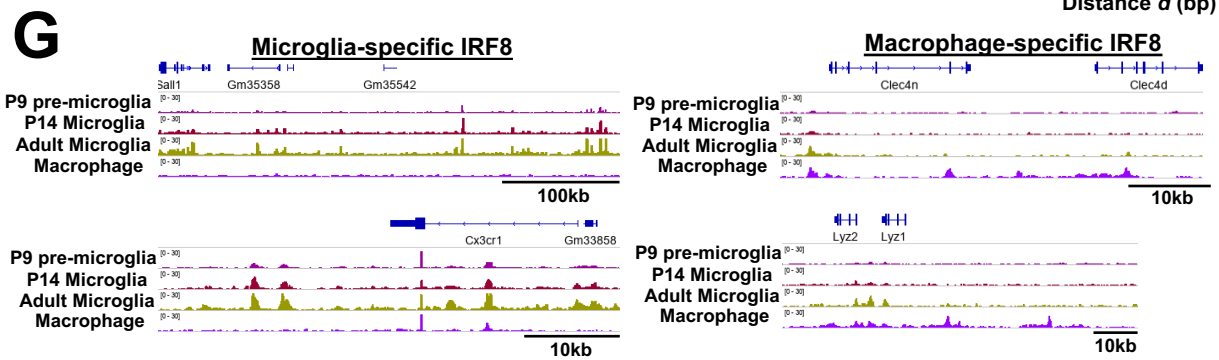


Fig.1 Saeki et al.

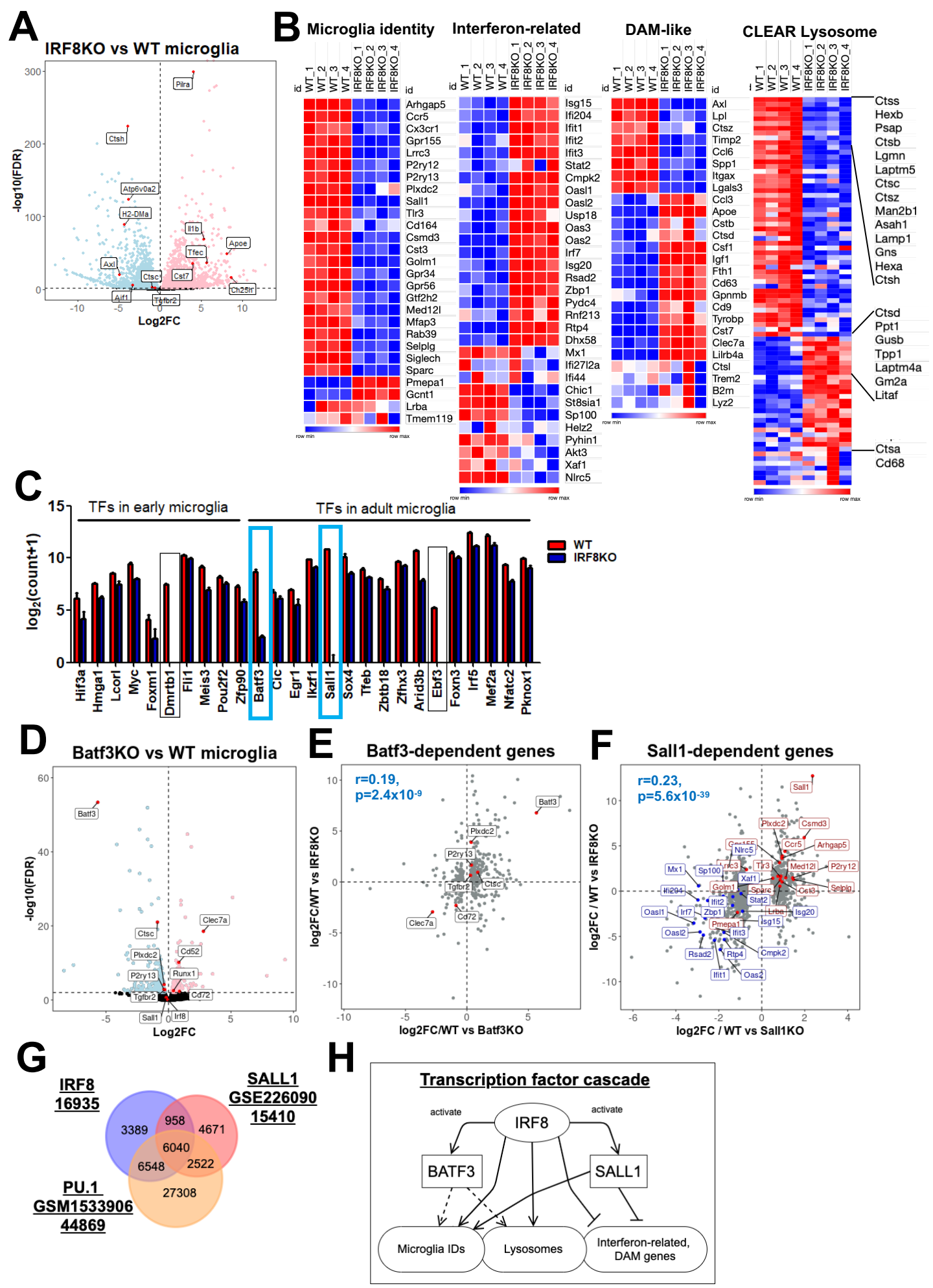


Fig.2 Saeki et al.

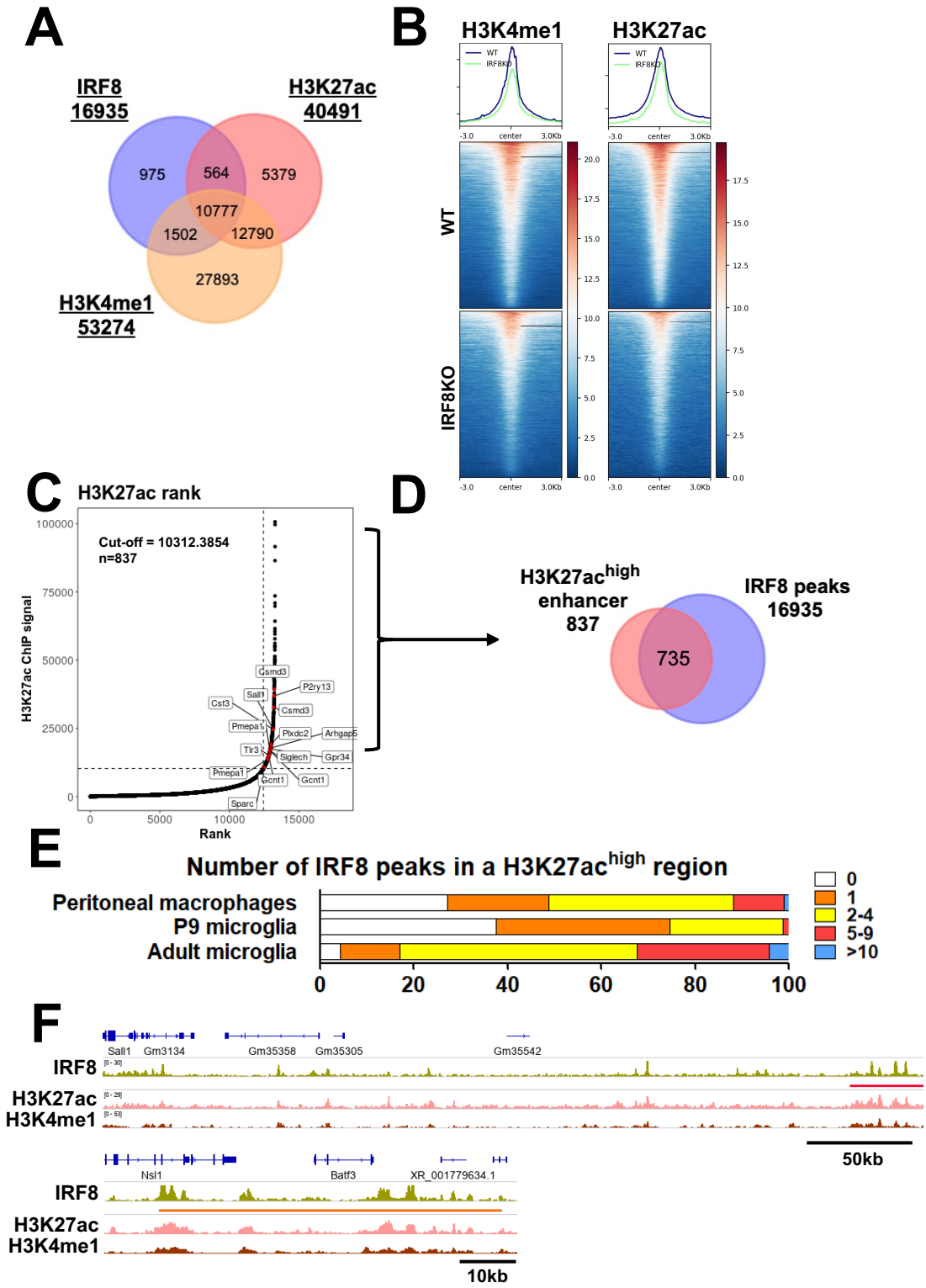
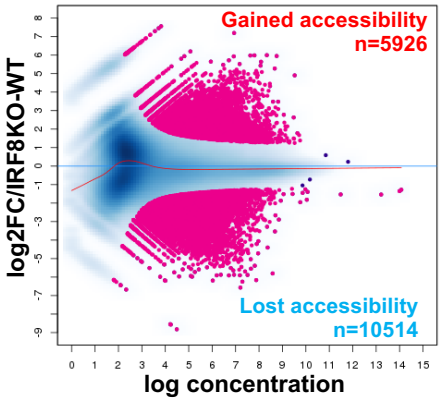


Fig.3 Saeki et al.

A

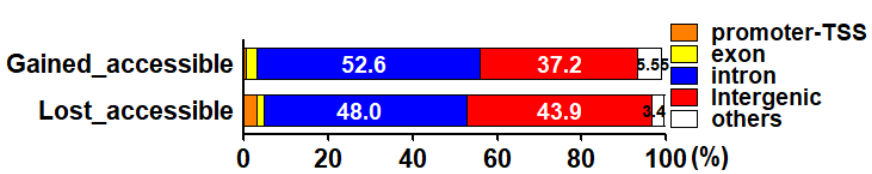


D

IRF8 peaks, H3K4me1 and H3K27ac marks in lost and gained ATAC-seq peaks

	Lost accessible regions (n=10514)	Gained accessible regions (n=5926)
IRF8 peaks (n=16935)	4103	158
WT-specific H3K4me1 (n=3882)	1489	8
KO-specific H3K4me1 (n=4952)	15	984
WT-specific H3K27ac (n=2819)	1210	7
KO-specific H3K27ac (n=1275)	1	163

B

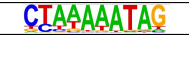


C

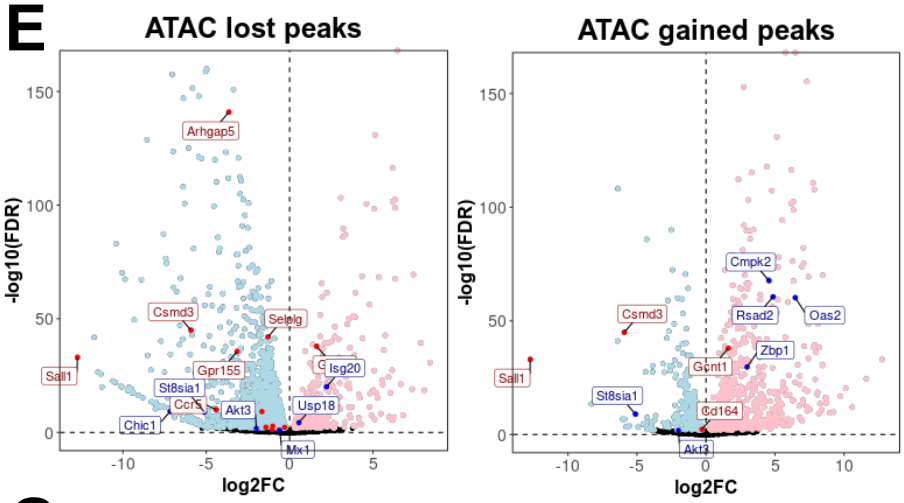
Lost accessible regions in IRF8KO microglia

Rank	Motif	P-value	T/B (%)	TF
1		1e-3245	43.01/4.14	PU.1:IRF8 (0.992)
2		1e-807	34.14/11.52	ELF3 (ETS, 0.918)

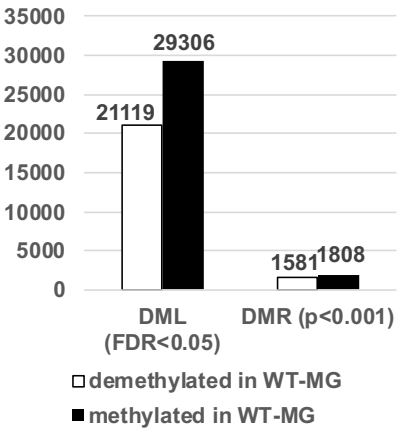
Gained accessible regions in IRF8KO microglia

Rank	Motif	P-value	T/B (%)	TF
1		1e-792	54.74/19.38	Sfpi1 (0.949)
2		1e-612	37.90/11.31	CEBPA (0.928)
3		1e-430	25.84/7.10	Smad2:Smad3 (0.981)
4		1e-236	28.70/12.58	RUNX2 (0.947)
5		1e-136	13.85/5.25	MEF2A (0.965)

E



F



G

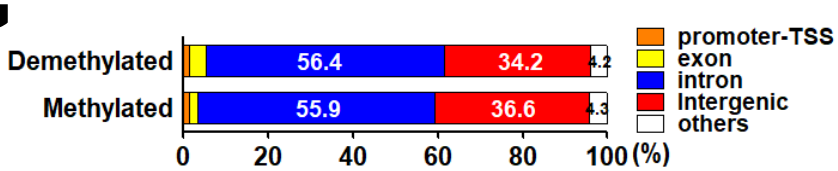


Fig.4 Saeki et al.

H IRF8 peaks, H3K4me1 and H3K27ac marks in methylated and demethylated regions

	Methylated in WT(n=1808)	Demethylated in WT(n=1581)
IRF8 peaks (n=16935)	41	721
WT-specific H3K4me1 (n=3882)	5	230
KO-specific H3K4me1 (n=4952)	254	1
WT-specific H3K27ac (n=2819)	8	222
KO-specific H3K27ac (n=1275)	33	0

I

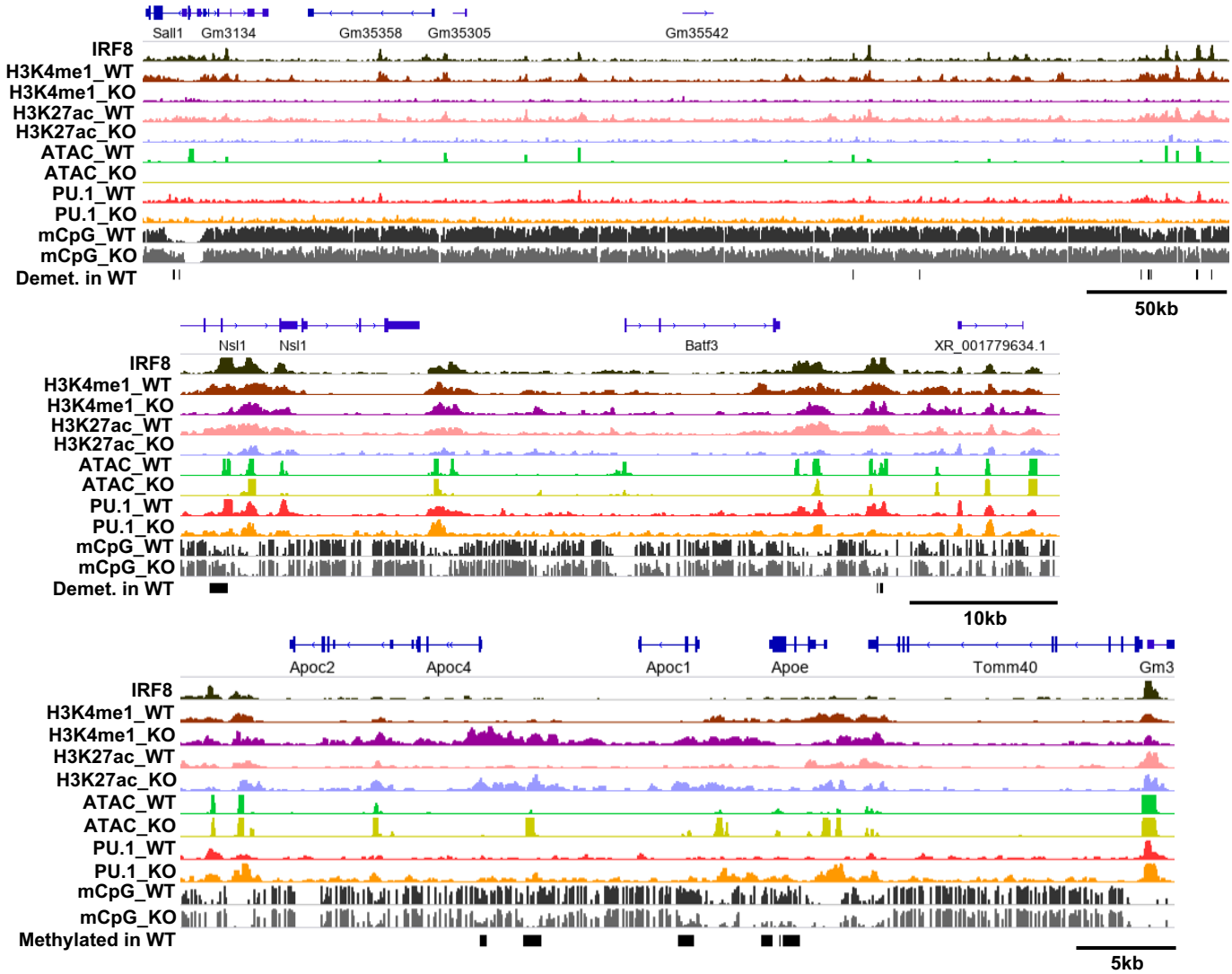


Fig.4 Saeki et al.

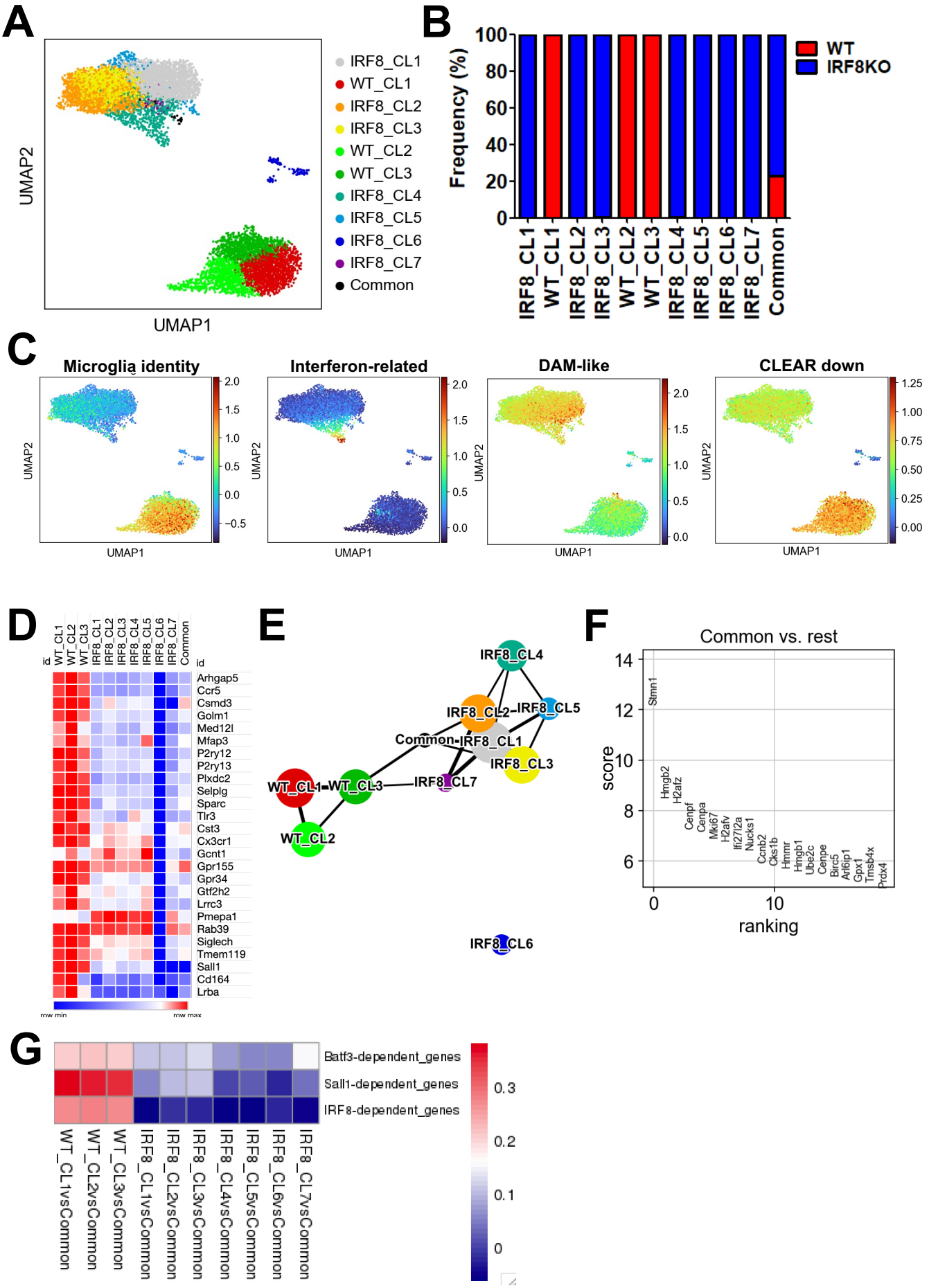


Fig.5 Saeki et al.

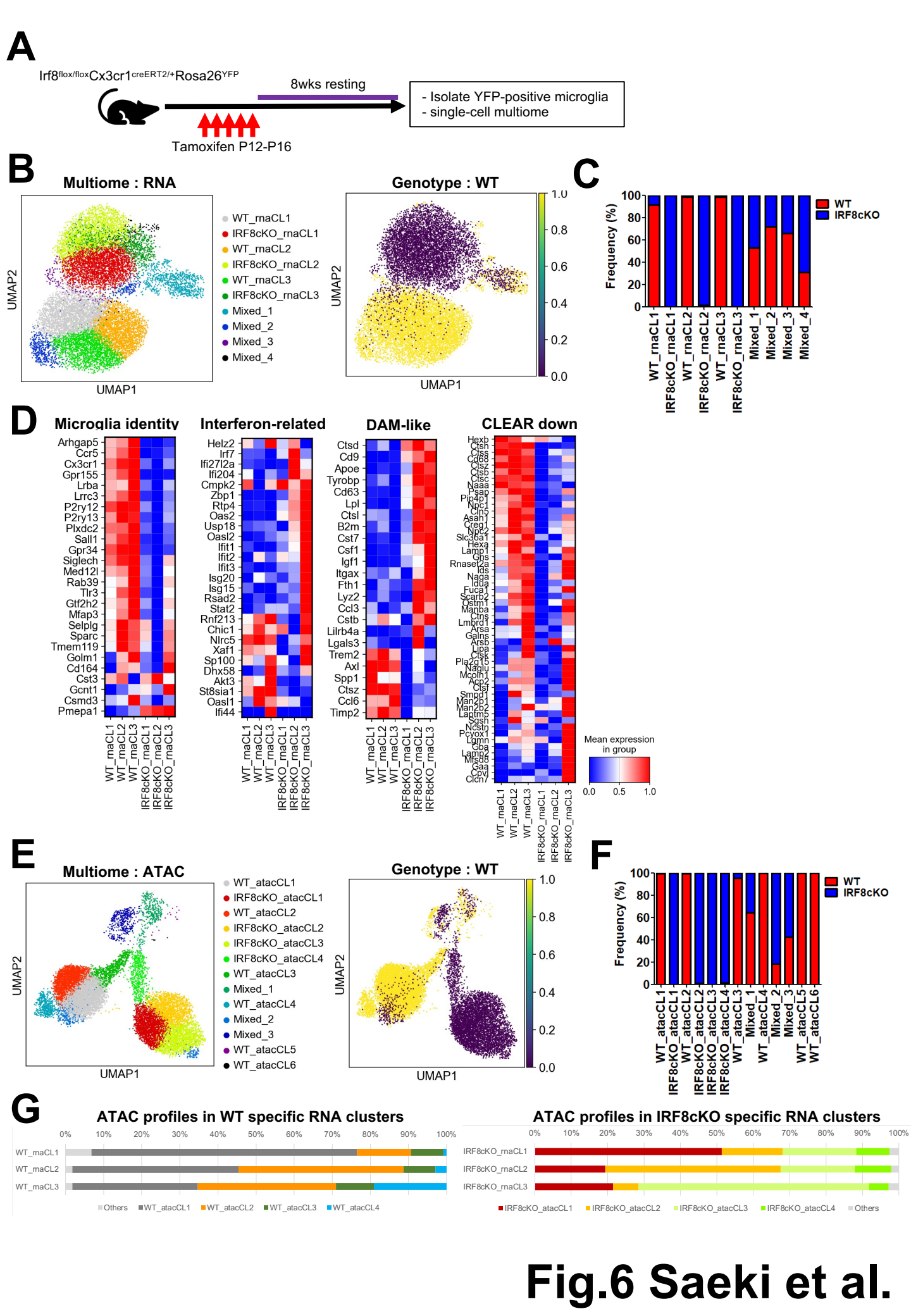
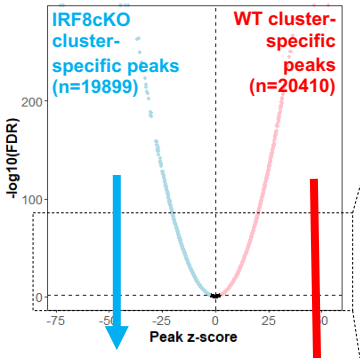
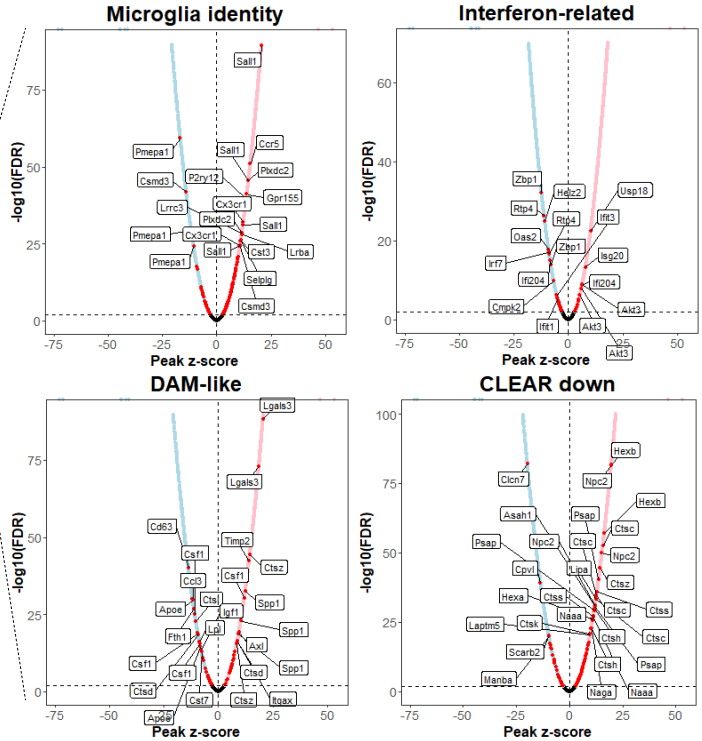


Fig.6 Saeki et al.

H WT-specific ATAC clusters vs IRF8cKO-specific ATAC clusters





I

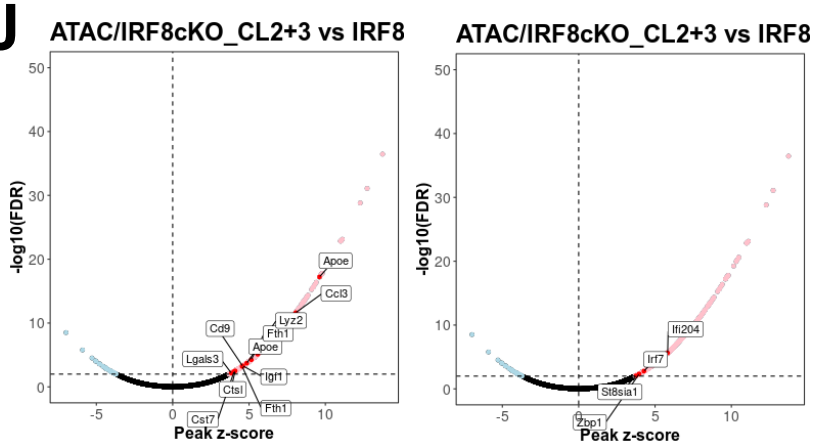
IRF8cKO cluster specific peaks

Rank	Motif	P-value	T/B (%)	TF
1		1e-1209	40.28/17.69	Sfpi1 (0.963)
2		1e-330	23.29/13.13	Cebpe (0.938)
3		1e-252	19.91/11.54	cJun (0.991)
4		1e-176	12.84/7.21	Mef2d (0.955)
5		1e-27	21.79/18.67	Znf384 (0.919)

WT cluster specific peaks

Rank	Motif	P-value	T/B (%)	TF
1		1e-2054	58.15/25.88	Sfpi1 (0.962)
2		1e-1097	32.70/13.31	PU.1:IRF8 (0.973)

J



K

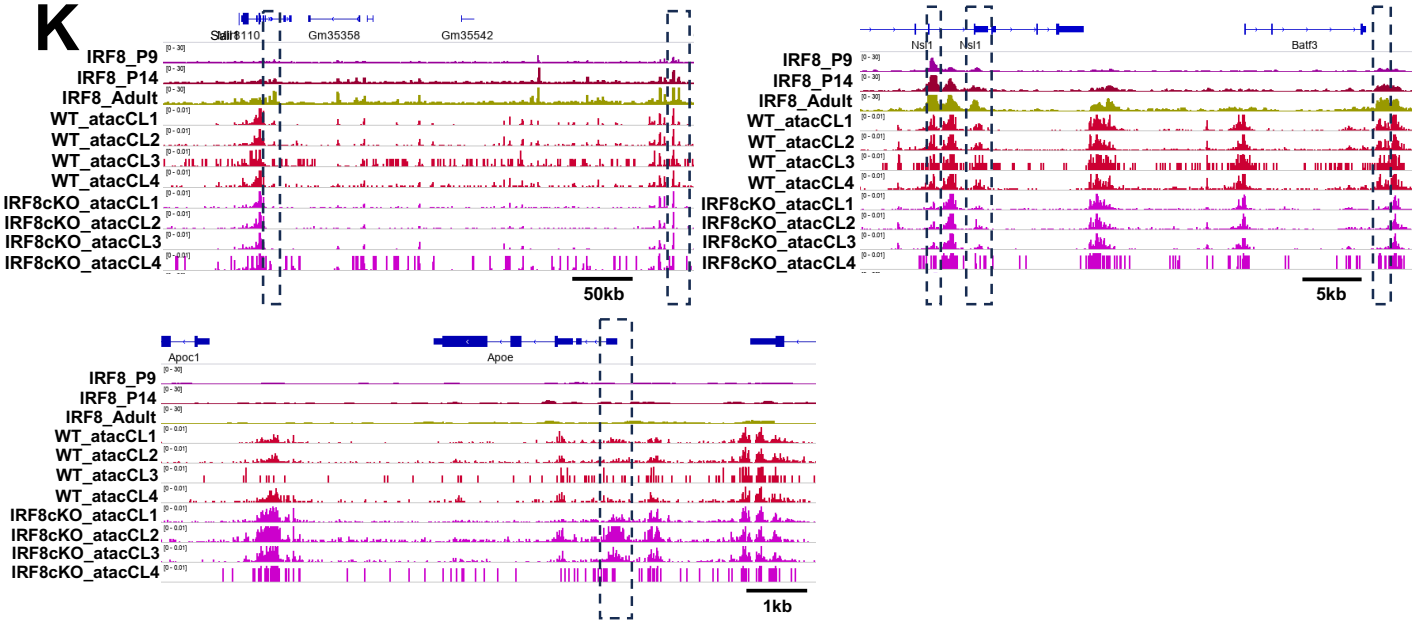
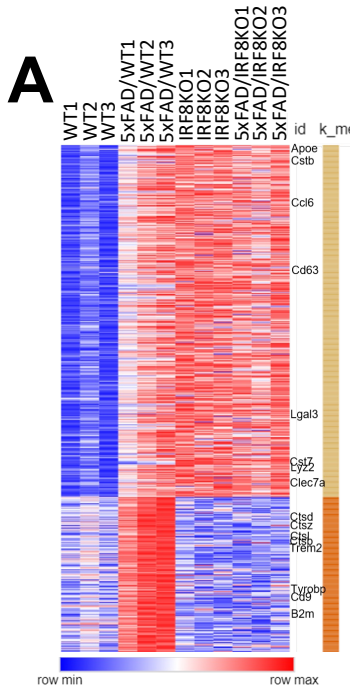
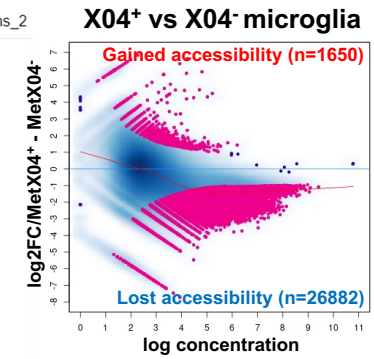


Fig.6 Saeki et al.





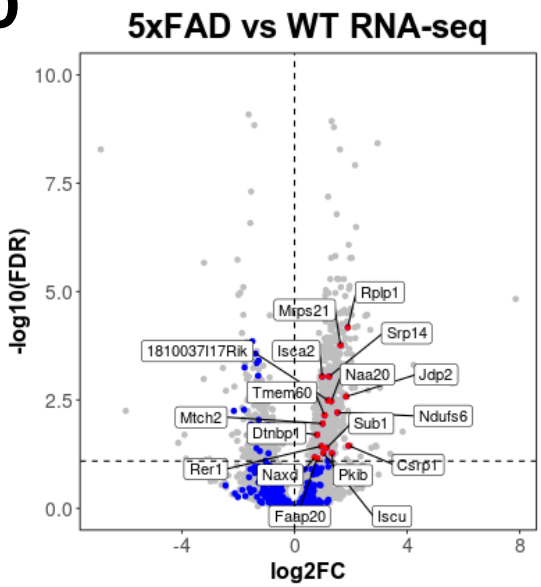
C

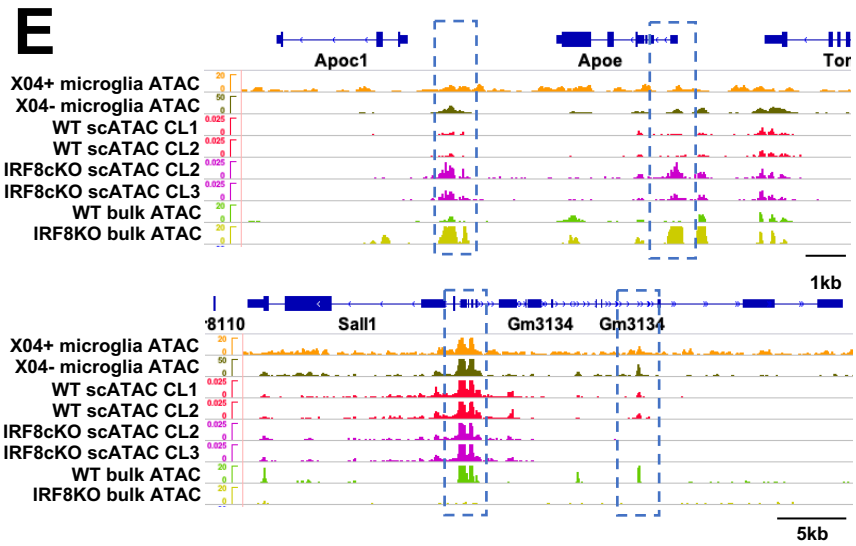
Accessibility gained in X04+ microglia

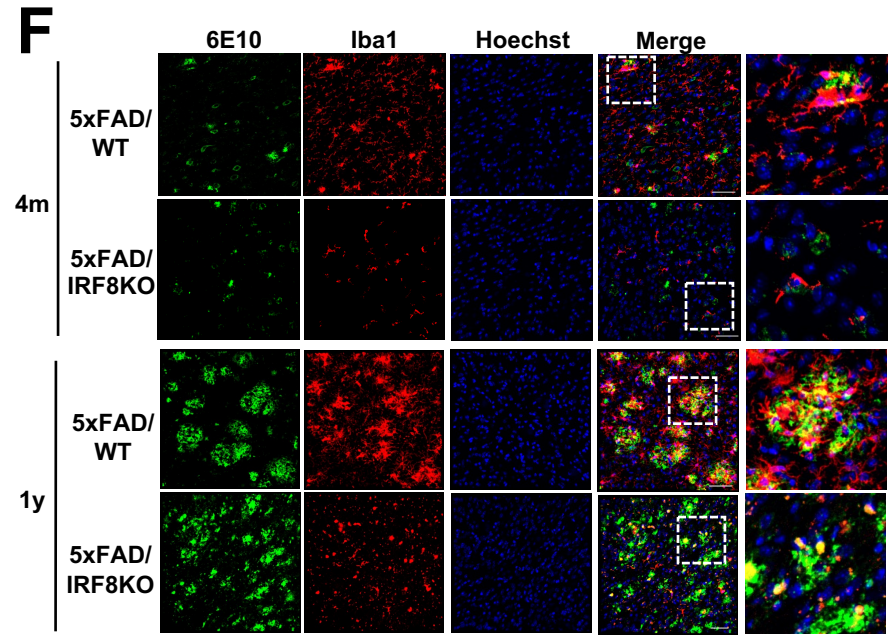
Rank	Motif	P-value	T/B (%)	TF
1	CTCCCGAGTAGC	1e-33	1.52/0.03	Osr1 (0.625)
2	TCCCATGCCAAC	1e-32	6.30/1.49	Rfx3 (0.915)
3	GTTATGTAAG	1e-25	4.97/1.20	HLF (bZIP; 0.946)
4	GCACAGTGCG	1e-22	43.39/31.86	DCE_S_II (0.697)

Accessibility reduced in X04+ microglia

Rank	Motif	P-value	T/B (%)	TF
1	ACTTCCTCTTTC	1e-4301	58.04/19.15	Sfpi1 (PU.1; 0.965)
2	GAAGTGAAAC	1e-754	12.32/3.76	PU.1-IRF (0.925)
3	GGTTGCGCAA	1e-206	17.48/11.16	CEBPE (0.870)
4	GCCCCCGCCCCC	1e-190	13.62/8.26	Sp2 (0.945)







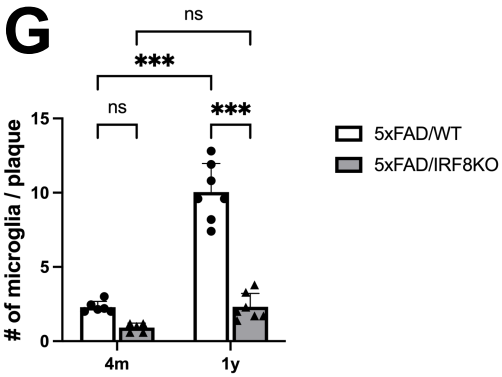


Fig.7 Saeki et al.

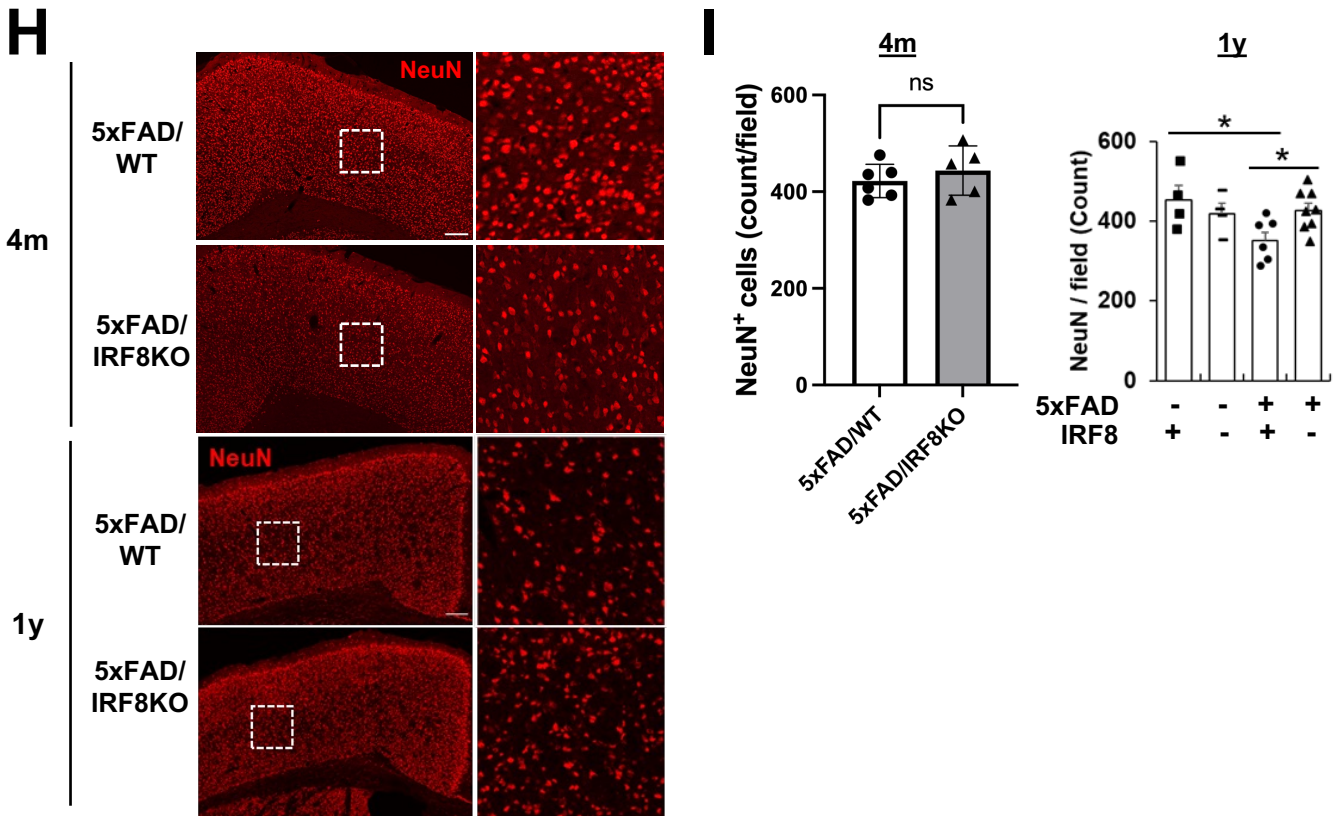


Fig.7 Saeki et al.

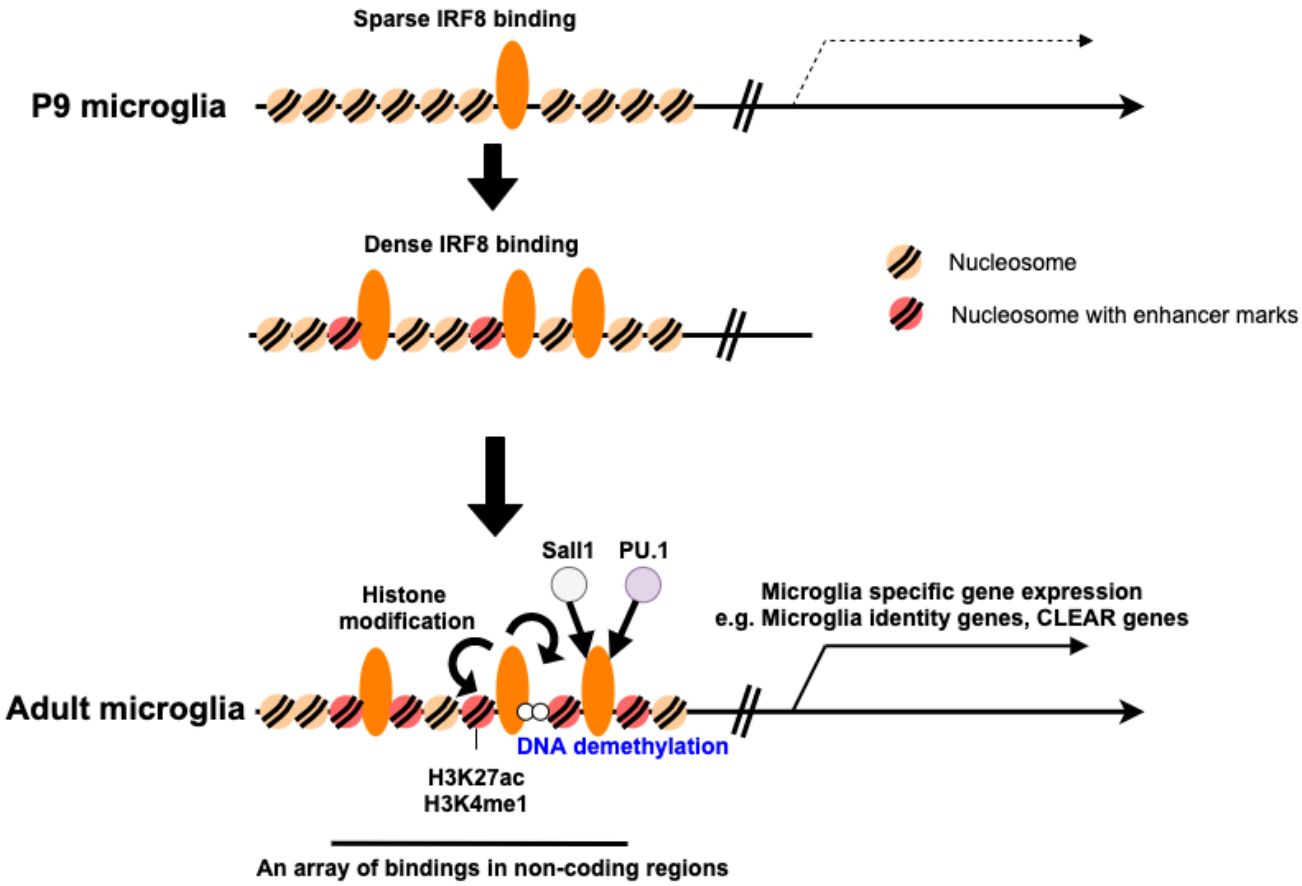


Fig.8 Saeki et al.