

Transformation of brain myeloid cell populations by SIV in rhesus macaques revealed by multiomics

Howard Fox

hfox@unmc.edu

University of Nebraska Medical Center <https://orcid.org/0000-0003-2032-374X>

Xiaoke Xu

University of Nebraska Medical Center

Meng Niu

University of Nebraska Medical Center <https://orcid.org/0000-0003-1055-3728>

Benjamin Lamberty

University of Nebraska Medical Center

Katy Emanuel

University of Nebraska Medical Center

Moses Apostol

University of Nebraska Medical Center

Article

Keywords: single cell, HAND, Microglia, Macrophage, snRNA-seq, snATAC-seq

Posted Date: September 18th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4916594/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

1 **Transformation of brain myeloid cell populations by SIV in rhesus macaques**
2 **revealed by multiomics**

3

4 Xiaoke Xu¹, Meng Niu¹, Benjamin G. Lamberty¹, Katy Emanuel¹, Moses Jedd Facun
5 Apostol¹, and Howard S. Fox^{1*}

6

7 ¹Department of Neurological Sciences, University of Nebraska Medical Center,
8 Omaha, Nebraska, USA

9

10 *Corresponding author

11 E-mail: hfox@unmc.edu (HSF)

12

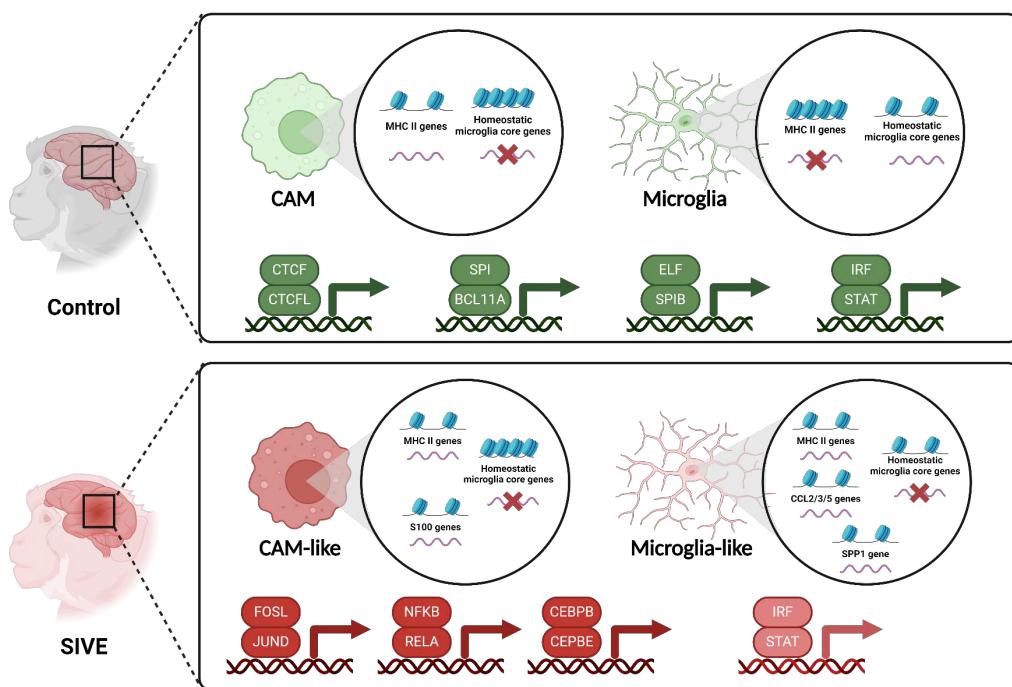
13 **Running title:** Transformation of brain myeloid cell populations by SIV

14 **ABSTRACT**

15 The primary immune constituents in the brain, microglia and macrophages, are the target for
16 HIV in people and simian immunodeficiency virus (SIV) in nonhuman primates. This infection
17 can lead to neurological dysfunction, known as HIV-associated neurocognitive disorder
18 (HAND). Given the gaps in our knowledge on how these cells respond *in vivo* to CNS
19 infection, we performed single-cell multiomic sequencing, including gene expression and
20 ATAC-seq, on myeloid cells from the brains of rhesus macaques with SIV-induced
21 encephalitis (SIVE) as well as uninfected controls. We found that the myeloid cell
22 populations were significantly changed by SIVE. In SIVE microglia-like cells express high
23 levels of chemoattractants capable of recruiting highly activated CAM-like cells to the site of
24 infection/inflammation. A unique population of microglia-like cells was found in which the
25 chromatin accessibility of genes diverged from their RNA expression. Additionally, we
26 observed a dramatic shift of upstream gene regulators and their targets in brain myeloid cells
27 during SIVE. In summary, this study further uncovers the transcriptome, gene regulatory
28 events and potential roles of different brain myeloid phenotypes in SIVE.

29 **Keywords:** single cell, HAND, Microglia, Macrophage, snRNA-seq, snATAC-seq

30



31 **INTRODUCTION**

32 The human immunodeficiency virus (HIV) is an enveloped retrovirus that contains
33 two copies of a single-stranded RNA genome, which can cause acquired immunodeficiency
34 syndrome (AIDS) by significantly impairing the immune system. HIV remains a global health
35 challenge with profound implications for individuals, communities, and societies. The
36 estimated number of people with HIV (PWH) is 39 million worldwide as of 2022.¹ According
37 to the latest epidemiology study in 2021, PWH comprise approximately 1.2 million people in
38 the United States.² HIV infection can lead to a spectrum of neurological complications, which
39 are referred to as HIV-associated neurocognitive disorder (HAND).³ Clinically, HAND
40 manifests with a range of neurological symptoms and can be classified as asymptomatic
41 neurocognitive impairment (ANI), mild neurocognitive disorder (MND), and HIV-associated
42 dementia (HAD).⁴ Despite advancements in antiretroviral therapy (ART), which have
43 dramatically reduced the prevalence of HAD, there are still 20%-50% of the PWH with milder
44 forms of HIV-associated neurocognitive disorders (HAND).^{5,6} Like HIV in genomic, structural,
45 and virologic perspectives, the simian immunodeficiency virus (SIV) also belongs to the
46 primate retrovirus family. Infection of rhesus macaque by SIV mimics many critical aspects of
47 HIV infection in humans, including immunodeficiency, opportunistic infections, and CNS
48 infection, which can be associated with neurological impairment.^{7,8}

49 The pathogenesis of HAND remains under intense study. One neuropathological link
50 has been HIV encephalitis (HIVE), which is characterized by inflammation of the brain tissue
51 resulting from the direct infection of HIV, as well as secondary effects of viral proteins and
52 immune activation. The exact mechanism by which HIV/SIV infection leads to HIVE/SIVE is
53 still unclear,⁹ but the infiltrating monocytes/macrophages and activated microglia in the brain
54 are thought to be the main contributors.¹⁰⁻¹³ HIV/SIV infection in the CNS is initiated by the
55 entry of infected myeloid and lymphoid cells from the periphery. Once the virus seeds into
56 the brain, the resident myeloid cells (microglia and CNS-associated macrophages (CAM))
57 could be infected and activated to have enhanced ability in secreting chemokines that can

58 further recruit more activated leukocytes to amplify the infection and inflammation. Several
59 studies have implicated monocyte chemoattractant protein-1 (MCP-1, CCL2) as playing a
60 role in disseminating the virus to the brain through its chemotactic properties on myeloid
61 cells, and elevated levels of this molecule have shown a strong association with HAD.¹⁴⁻¹⁶ A
62 mutation in the CCL2 gene (encoding MCP-1) that leads to the enhanced infiltration of blood
63 myeloid cells has been reported to be a risk factor for causing HAD.¹⁷ In addition to CCL2,
64 other chemokines might also be involved in neuropathogenesis caused by HIV/SIV. The
65 activation of resident microglia and macrophages by viral proteins (e.g. gp120)¹⁸⁻²⁰ or factors
66 (e.g. cytokines and chemokines)^{10,13} released from infiltrating cells can further damage the
67 neurons and other brain cell types. Additionally, infected microglia and macrophages make
68 up a viral reservoir in the brain under suppressive ART treatment, complicating efforts for an
69 HIV/SIV cure.²¹

70 In our previous studies, single-cell transcriptomic analyses of brain myeloid cells
71 during different phases of SIV infection uncovered dysregulated genes and cellular
72 phenotypes in acute and chronic infection, the latter in the presence and absence of
73 antiretroviral treatment, as well as end-stage encephalitis.²²⁻²⁴ However, using the
74 information on RNA transcripts alone might not be enough to characterize the
75 heterogeneous myeloid populations and comprehensively understand complicated gene
76 regulatory events. Therefore, in this study, we used the single-cell multiomics technique, in
77 which paired single-nucleus RNA sequencing (snRNA-seq) with single-nucleus ATAC
78 sequencing (snRNA-seq). In this manner, we have characterized the myeloid cell
79 phenotypes in the brain to understand the cellular events in those cells during severe CNS
80 infection of SIV (SIVE).

81

82 **RESULTS**

83 **SIV-induced encephalitis dramatically changed the brain myeloid cell populations.**

84 We used both RNA expression and chromatin accessibility information for cell
85 clustering. In total, we identified 12 different cell clusters (**Figure 1A**). After quality control,
86 doublet removal, dimensionality reduction, and batch effect correction, the SIVE or control
87 group samples aggregated sufficiently. However, the samples from SIVE animals (21T and
88 34T) and uninfected animals (104T and 106T) were still separated (**Figure 1B**), which
89 suggests that the SIVE significantly changed the myeloid cell phenotypes in the normal brain
90 (**Figure 1C**). After screening the cell markers using snRNA-seq data (**Figures 1D and**
91 **1E**) and snATAC-seq data (**Figures 1F and 1G**), we identified microglia, CNS-associated
92 macrophages (CAMs), and a few lymphocytes as the three central cell populations in our
93 dataset.

94 We removed the lymphocyte cluster (C11) and one undefined cluster with only 28
95 cells (C1) from our dataset. The final dataset included ten myeloid cell clusters (**Figure 2A**),
96 with six clusters (C3, C4, C7, C8, C10, and C12) found predominately in the uninfected
97 brains, and four (C2, C5, C6, and C9) specifically found in SIVE animals (**Figure 2B**). The
98 myeloid cells in uninfected brains were different phenotypes of microglia and CAM. The
99 microglial cells were characterized as cells with high expression of homeostatic microglia
100 core genes (e.g. P2RY12, GPR34, CX3CR1, and SALL1) but low expression of MHC class II
101 molecules (e.g. MAMU-DRA, -DRB1, -DRB5, and CD74). Conversely, the CAM were
102 characterized as the cells with high expression of MHC class II molecules but low expression
103 of homeostatic microglial core genes.²⁵ The discrepancies between chromatin accessibility
104 (predicted gene activity) and RNA expression for myeloid cells in uninfected brains were
105 minimal (**Figures 2C and 2D**), indicating that the RNA expression of the genes largely
106 reflected their chromatin accessibility in uninfected conditions. As for four SIVE-specific
107 clusters (C2, C5, C6, and C9), we found three of them (C5, C6, and C9) that shared several
108 common findings. For example, they had minimal chromatin accessibility (predicted gene
109 activity) and RNA expression of homeostatic microglial core genes (**Figures 2C and 2D**).
110 Additionally, they were closer to each other but far from the C2 on the UMAP.

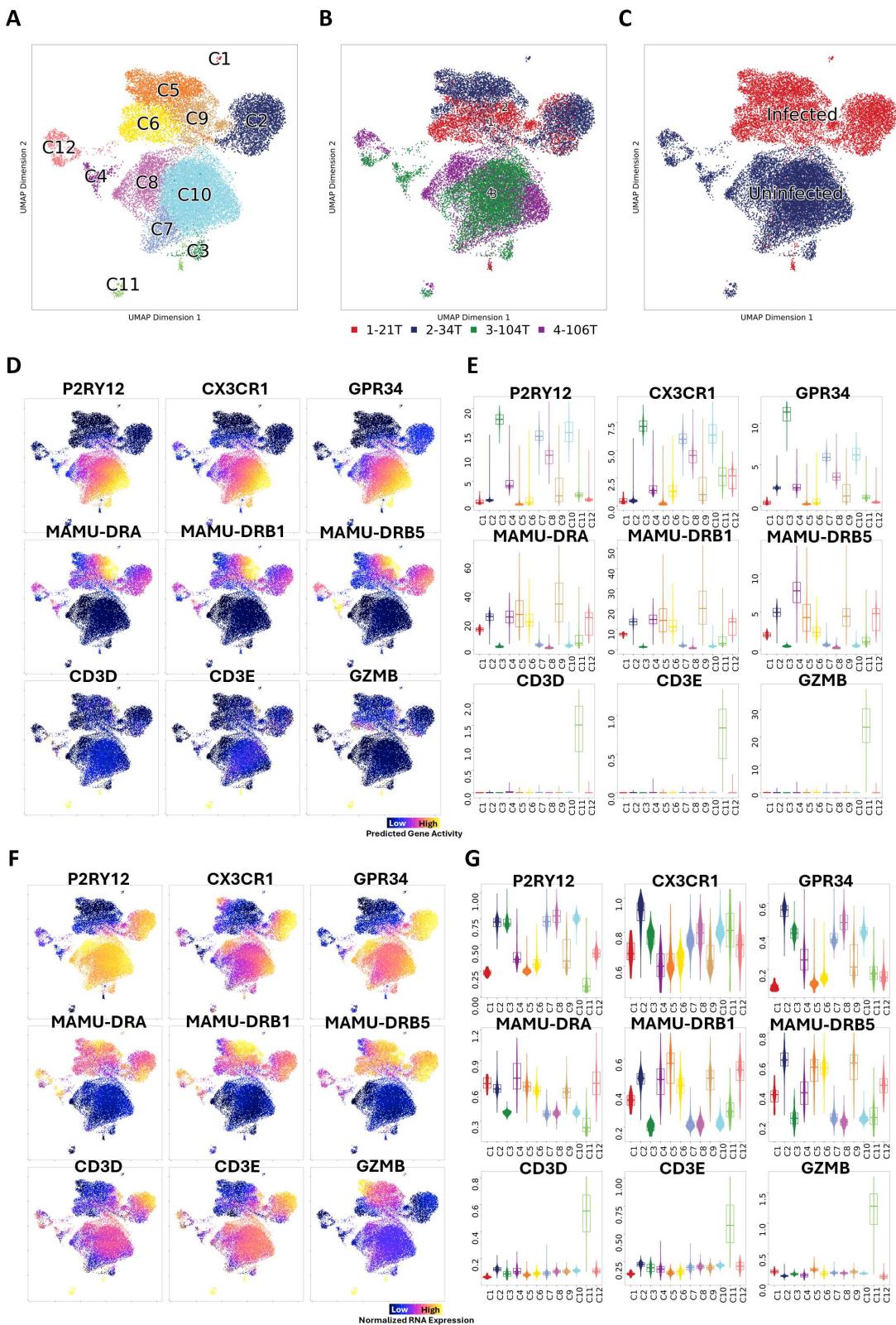


Figure 1. Characterization of different cell phenotypes in the dataset. **(A)** UMAP projection of 24,633 single-cell chromatin and RNA expression profiles, colored by graph-based clustering results **(B)** UMAP projection of single cells in the dataset. The color of each cell indicated the animals that they came from. **(C)** UMAP projection of single cells in

the dataset. The color of each cell indicated they were from infected or uninfected animals. **(D and E)** UMAP projection and violin plots of predicted gene activity of known marker genes for microglia, CNS-associated macrophages (CAMs), and lymphocytes. **(F and G)** UMAP projection and violin plots of RNA expression of known marker genes for microglia, CAMs, and lymphocytes.

111

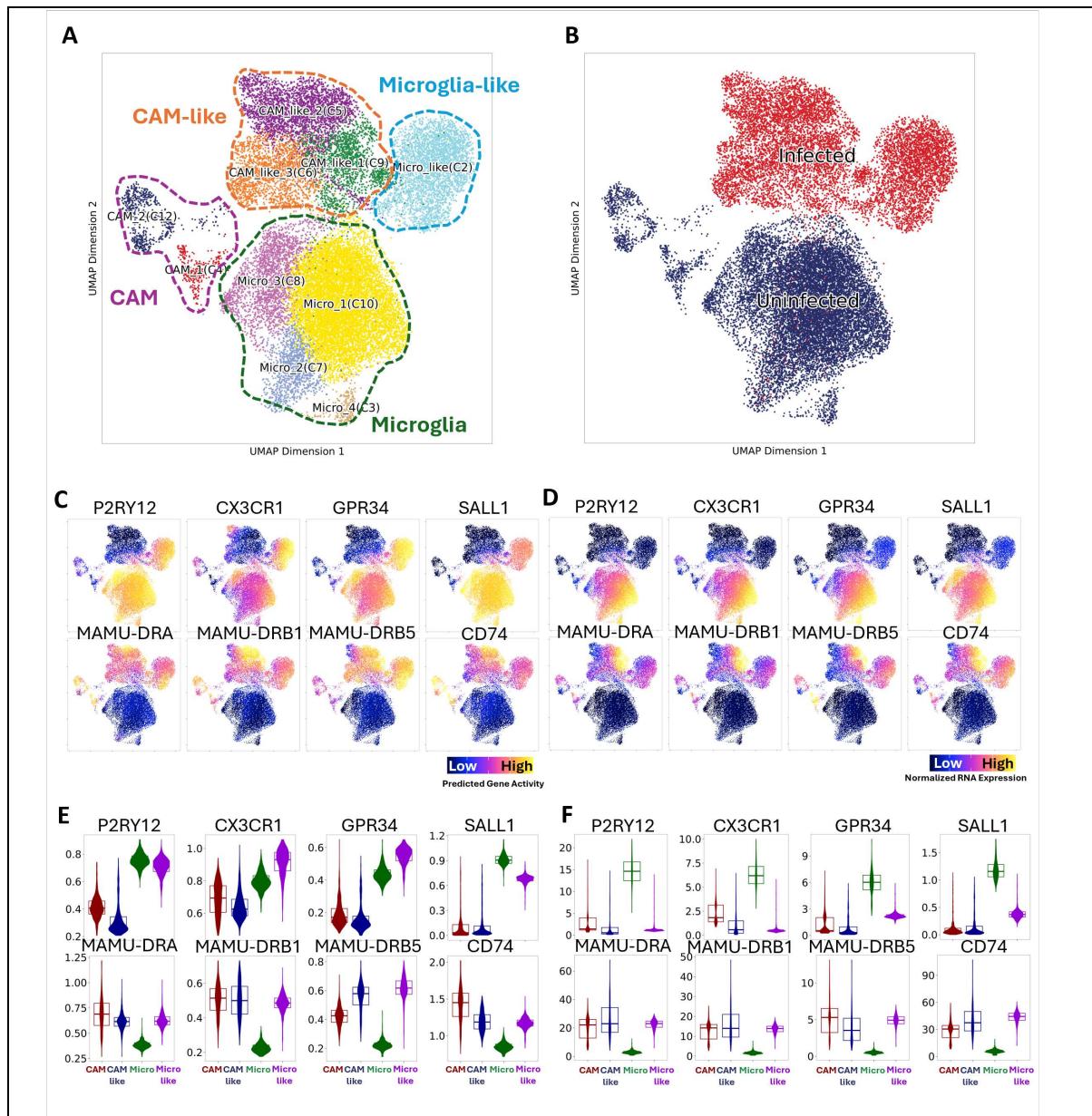


Figure 2. Multiomic characterizations of myeloid cells in the brain of uninfected and SIV-infected animals with encephalitis. (A) UMAP projection of 24,420 myeloid-cell chromatin and RNA expression profile, colored by graph-base clustering results. **(B)** UMAP projection was colored by uninfected (blue) and infected (red) status. **(C)** UMAP projection of predicted gene activity of homeostatic microglial core genes and MHC class II molecules. **(D)** UMAP projection of normalized RNA expression of homeostatic microglial core genes and MHC class II molecules. **(E)** Violin plots

of predicted gene activity of homeostatic microglial core genes and MHC class II molecules. The cells in CAM, CAM-like, Microglia (Micro), and Microglia-like clusters were aggregated for plotting. **(F)** Violin plots of normalized RNA expression of homeostatic microglial core genes and MHC class II molecules. The cells in CAM, CAM-like, Microglia (Micro), and Microglia-like clusters were aggregated for plotting.

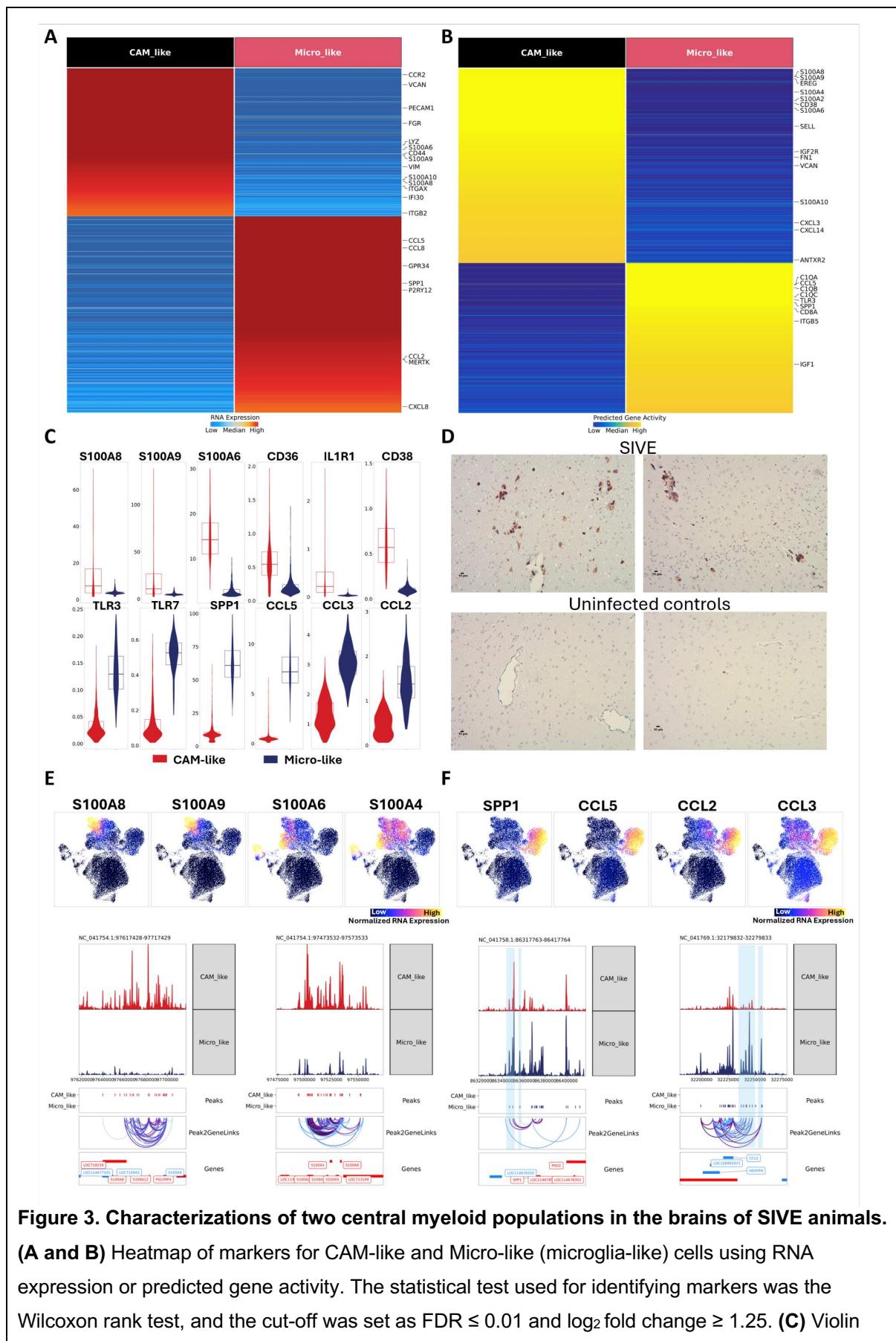
112
113 Distinct from the C5, C6, and C9 clusters, C2 cluster had low RNA expression but high levels
114 of chromatin accessibility of homeostatic microglia core genes, which shared some
115 similarities with microglia in uninfected brains. Given the resemblance between SIVE-
116 specific myeloid clusters and microglia and CAM, we annotated the cells in the C2 cluster as
117 microglia-like cells and those in the C5, C6, and C9 as CAM-like cells. In summary, open
118 chromatin was found for homeostatic microglia genes in both microglia and microglia-like
119 cells (**Figure 2E**), but there was a divergence in RNA expression (**Figure 2F**) for those
120 genes between microglia and microglia-like cells. In contrast, CAM and CAM-like cells
121 shared similarities for MHC class II molecules in chromatin accessibility and RNA expression
122 (**Figures 2E and 2F**).

123 Cluster-specific markers, based on RNA expression (**Figure S1A**) and chromatin
124 accessibility (**Figure S1B**), were found for each myeloid cell cluster by setting a cut-off of
125 FDR ≤ 0.01 and $\log_2\text{FC} \geq 1.25$ (**Table S1**). The two CAM clusters identified in uninfected
126 controls comprised two different phenotypes that might be derived from two central blood
127 monocyte populations. Classical monocytes have a high capacity for phagocytosis and are
128 efficient producers of pro-inflammatory cytokines, which play a central role in the early
129 stages of the immune response, whereas non-classical monocytes exhibit a patrolling
130 behavior, continuously surveying the endothelial lining of blood vessels for signs of injury or
131 infection.^{26,27} The CAM_1 cluster had higher expression of molecules related to antigen
132 presentation and phagocytosis, whereas the CAM_2 cluster upregulated various adhesion
133 molecules. This suggests that CAM_1 might share an origin with classical monocytes and
134 CAM_2 with non-classical monocytes. The different microglial clusters identified in
135 uninfected brains, while distinct from the CAM clusters, did not show many differences with

136 each other, and they all highly expressed homeostatic microglial core genes. While the
137 CAM-like clusters showed very different transcriptomic profiles and open chromatin regions
138 than microglia-like clusters, the differences between each of the CAM-like clusters were
139 minimal. Therefore, we focused on the differences between CAM-like and microglia-like cells
140 to understand these cell populations in SIVE.

141 **The gene activity and transcriptomic profile of CAM-like myeloid cells were vastly**
142 **different from those of microglia-like cells.**

143 We first compared RNA expression (**Figure 3A**) and gene activity (**Figure 3B**) for CAM-like
144 cells and microglia-like cells (**Table S2**). There were 166 markers in CAM-like cells and 154
145 in microglia-like cells that were common between RNA expression and chromatin
146 accessibility analyses. Although microglia-like and CAM-like cells were both activated
147 myeloid cells, their upregulated genes differed (**Figure 3C**). Compared to the microglia-like
148 cells, the CAM-like cells had high RNA expression of molecules in the S100 family (**Figure**
149 **3C**), which have been reported to play critical roles in macrophage signaling. For example,
150 heterodimers formed by S100A8 and S100A9, named calprotectin, serve as an indicator for
151 inflammation.^{28,29} Immunostaining for S100A9 revealed a notable increase in S100A9+ brain
152 myeloid cells in SIVE (**Figure 3D**). Interestingly, most of the S100A9+ myeloid cells in the
153 SIVE brains are located in the perivascular space, indicating that the CAM-like cells might be
154 activated CAM cells in response to SIVE. In addition, S100A6 and S100A4, which were
155 found to be highly expressed in CAM-like cells, play a vital role in myeloid cell activation
156 through AP-1 signaling. The genomic regions for S100A4 and S100A6 genes are close to
157 each other, as are those for S100A8 and S100A9, and they were all very accessible in CAM-
158 like cells but not in microglia-like cells (**Figure 3E**). Furthermore, Ingenuity Pathway Analysis
159 (IPA), assessing the RNA expression DEGs, revealed that the S100 family signaling
160 pathway was significantly upregulated (p-value: 3.66E10⁻¹⁴) in the CAM-like cells. From the
161 gene networks for the S100 family signaling pathway, we found that most of the DEGs in the



plot for RNA expression of selected genes between CAM-like and Micro-like cells. **(D)** Representative images (scale bars, 10 μ m) of S100A9+ cells in the temporal lobe of two SIVE animals (21T and 34T). The same brain region of two uninfected animals was stained as a comparison. **(E)** UMAP projection of RNA expression of S100A8/9 and S100A4/6 proteins for all myeloid cell clusters. (upper panel) Representative browser plot visualization of pseudo-bulk ATAC-seq, peak-to-gene links (Peak2GeneLinks), and genes at the S100A8/9 and S100A4/6 loci for Micro-like cells and CAM-like cells. (lower panel) **(F)** UMAP projection of SPP1, CCL5, CCL3, and CCL2 RNA expression for all myeloid cell clusters. (upper panel) Representative browser plot visualization of pseudo-bulk ATAC-seq, peak-to-gene links (Peak2GeneLinks), and genes at the SPP1 and CCL5 loci for Micro-like and CAM-like cells (lower panel).

162

163 CAM-like cluster were enriched in S100 proteins that can induce their expression through
164 AP-1 and NF- κ B activation (**Figure S2A**). Some surface markers involved in S100 protein
165 induced NF- κ B and AP-1 activation, including CD36, IL1R1, and CD38, were also
166 upregulated in CAM-like cells (**Figure 3C**). These results suggest that NF- κ B and AP-1
167 signaling, significantly activated by S100 proteins, are critical for CAM-like cells mediated
168 neuroinflammation in SIVE animals. Interestingly the three CAM-like clusters were found to
169 upregulate different S100 proteins. For example, the CAM_like_2 cluster had higher RNA
170 expression of S100A8 and S100A9 proteins, but the CAM_like_3 cluster were found to have
171 higher RNA expression of S100A4 and S100A6 (**Figure 3E and Table S1**).

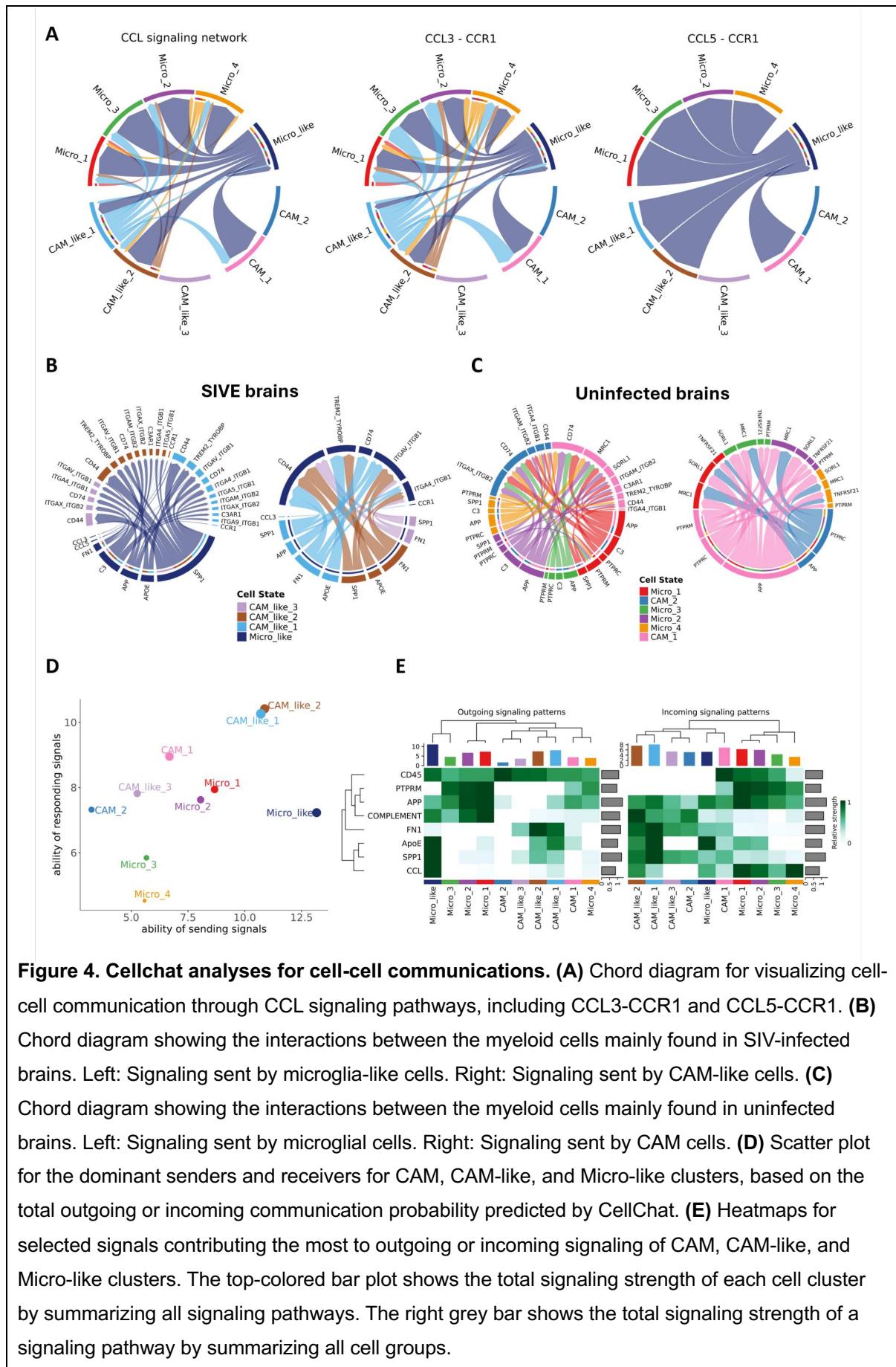
172 In contrast, the microglia-like cells were more likely to play roles in immune
173 interactions. The genes encoding complement components (e.g. C1QA, C1QB, and C1QC)
174 and Toll-like receptors (e.g. TLR3 and TLR7) were upregulated in microglia-like cells
175 compared with CAM-like cells (**Figures 3A-3C**). TLR7 can specifically recognize SIV single-
176 stranded RNA (ssRNA), making it a key sensor for SIV nucleic acids, and both TLR3 and
177 TLR7 can induce the production of type I IFNs and other antiviral factors. The GO and
178 KEGG enrichment analyses (**Figure S2B and S2C**) for RNA markers detected for CAM-like
179 and microglia-like cells further supported that the microglia-like cells had enhanced Toll-like
180 receptor signaling. Some inflammatory molecules were also highly expressed in microglia-
181 like cells but not in CAM-like cells. For example, we found that the SPP1, IL-6, IL-18, and

182 many CCL chemokines had higher RNA expression and chromatin accessibility in microglia-
183 like cells (**Figure 3C**). The chromatin regions for the SPP1 gene were more open, with more
184 marker peaks ($FDR \leq 0.01$ & $\text{Log}_2\text{FC} \geq 1$) found in microglia-like cells compared to CAM-like
185 cells. Interestingly, we found three distal peaks of the SPP1 gene were also open in CAM-
186 like cells, which were highly correlated with SPP1 gene expression (**Figure 3F**).

187 The correlations between peak accessibility and gene expression allow for the
188 prediction of specific enhancer-promoter links. In our dataset, we identified 63,855 such
189 potential enhancer-promoter interactions (**Figure S2D**). Therefore, the correlations between
190 those upstream regions of the SPP1 gene and SPP1 expression suggested they might be
191 the potential enhancers for the SPP1 gene in brain myeloid cells, especially microglia-like
192 cells in SIVE. The CCL chemokines, including CCL5 (RANTES), CCL2 (MCP-1), and CCL3
193 (MIP-1 α), also had high expression in microglia-like cells (**Figure 3F**) and function as
194 chemoattractants for various immune cells. Like the SPP1 gene, more marker peaks within
195 the CCL5 gene region were found in microglia-like cells. We also found that several peak
196 regions at the CCL5 gene upstream showed strong correlations with CCL5 expression,
197 suggesting the potential enhancer regions. More chromatin accessibility in those enhancer
198 regions coordinated with CCL5 promoter regions, leading to an enhanced expression in
199 microglia-like cells (**Figure 3F**).

200 **Microglia-like cells likely play an essential role in attracting other myeloid cells to the
201 site of inflammation and mediating cell-cell interaction during SIVE.**

202 The initial CNS infection is mediated by the infiltration of peripheral blood leukocytes
203 into the brain, and the subsequent activation of microglia, macrophages, and possibly other
204 cells in the brain could further amplify the neuroinflammation and lead to neurotoxicity. Such
205 inflammation could induce chemoattractant activities and cell-cell interactions of different
206 SIVE-specific myeloid cell clusters. Therefore, we inferred the cell communication probability
207 and cellular communication network for all the myeloid cell clusters identified in this study.
208 Corresponding to the high expression of some CCL chemokines in microglia-like cells, the



209 primary sender in CCL-CCR ligand-receptor interactions, including CCL3-CCR1 and CCL5-
210 CCR1, was microglia-like cells (**Figure 4A**).

211 In addition to CCL chemokines, the microglia-like cells interacted with CAM-like cells
212 through SPP1, APOE, APP, and C3-mediated signaling mechanisms. The interaction of
213 SPP1 and CD44, which can promote cell adhesion, migration, and immune responses,
214 seemed most robust between microglia-like cells and different CAM-like cells in the brains
215 with SIVE. While the interactions of microglia-like cells were dominated by the high
216 expression of SPP1, the interaction initiated by the CAM-like cells appeared driven by
217 fibronectin (FN1), well known for involvement in creating extracellular matrices and wound
218 healing (**Figure 4B**). The myeloid cell interactions in the uninfected normal brain significantly
219 differed from those in the infected brain. Microglial cells in uninfected conditions
220 communicate with CAMs mainly through APP, C3, and PTPRC (CD45), which can interact
221 with CD74, a chaperone for MHC class II molecules on CAMs. The CAMs in uninfected
222 conditions also preferred using APP and PTPRC to communicate with microglia-like cells by
223 binding with MRC1 (CD206) on microglia cells (**Figure 4C**).

224 To further understand the involvement of different cell phenotypes in cell-cell
225 interaction, we ranked them based on their ability to send and receive signals. In general,
226 microglia cells in uninfected animals and microglia-like cells in SIVE animals had a stronger
227 ability to send signals than their ability to receive signals, and the CAM cells and CAM-like
228 cells the opposite (**Figure 4D**). The outgoing and incoming signaling patterns for cell-cell
229 interactions were also very different between the myeloid cells in uninfected conditions and
230 SIVE (**Figure 4E**). In summary, the cell communication analyses further showed the different
231 functions of microglia-like and CAM-like cells in the brains with SIVE, and the microglia-like
232 cells might be the central phenotype to recruit the inflammatory CAM-like cells to amplify the
233 neuroinflammation.

234 **Inflammation-related TFs and TF complexes were enriched and highly expressed in**
235 **SIVE-specific clusters.**

236 To better understand the open chromatin regions of myeloid cell clusters, we
237 performed peak calling on a cell cluster basis. We identified 187,594 peaks from those ten
238 myeloid cell clusters, and more peaks were found in the microglia-like cluster, CAM-like
239 clusters, and Micro_1 cluster compared to other clusters (**Figure S3A**). The marker peaks
240 (FDR ≤ 0.05 and Log₂FC ≥ 1) identified in each cell cluster (**Table S3**) further confirmed that
241 SIVE-specific myeloid cell populations had significantly different open chromatin regions
242 compared to microglia and CAMs in uninfected conditions. Next we evaluated the
243 transcription factor (TF) motif enrichment in cluster-specific peaks (**Figure S3B**) and TF
244 activity on a per-cell basis using two TF databases.³⁰ We identified some highly variable TF
245 complexes between CAM-like and microglia-like cells, which included FOSL/JUND,
246 BACH/NFE, NFKB/RELA, SPI/BCL11A, CEBPB/CEBPE, ELF/SPIB, CTCF/CTCFL, and
247 IRF/STAT (**Figure 5A**). The archetype consensus of the binding sites for those high-variable
248 TF complexes is shown in **Figure 5B**. Including those highly variable TFs, most of marker
249 TFs (FDR ≤ 0.001) enriched in microglia and CAM in uninfected brains were very different
250 from those enriched in SIVE-specific clusters (**Table S4, Figure 5C and S3C**), indicating
251 SIVE dramatically shifted the regulatory events in brain myeloid cells. The enrichment of TFs
252 in brain myeloid cell clusters was also predicted by the Single-Cell rEgulatory Network
253 Inference and Clustering (SCENIC) R package^{31,32}, where only the RNA expression profile
254 was used to predict the regulon activity in given cell clusters, and yielded fewer enriched
255 TFs, but many in common(labeled in red in **Figures 5C and 5D**).

256 For highly variable TFs, we found that the FOSL/JUND, NFKB/RELA, BACH
257 (BACH2) /NFE (NFE2), and CEBPB/CEBPE were more enriched in CAM-like cells,
258 IRF/STAT was enriched in microglia-like cells and microglia and CTCF/CTCFL,
259 SPI(SPI1)/BCL11A, and ELF (ELF3)/SPIB were enriched in microglia and CAMs of
260 uninfected animals (**Figure 5E, 5F and S3D**). To confirm the TFs whose motifs were highly
261 enriched in SIVE-specific clusters also had higher expression, we further compare the RNA
262 expression levels and predicted gene activity of those TFs (**Figure S3E-S3H**). The RNA

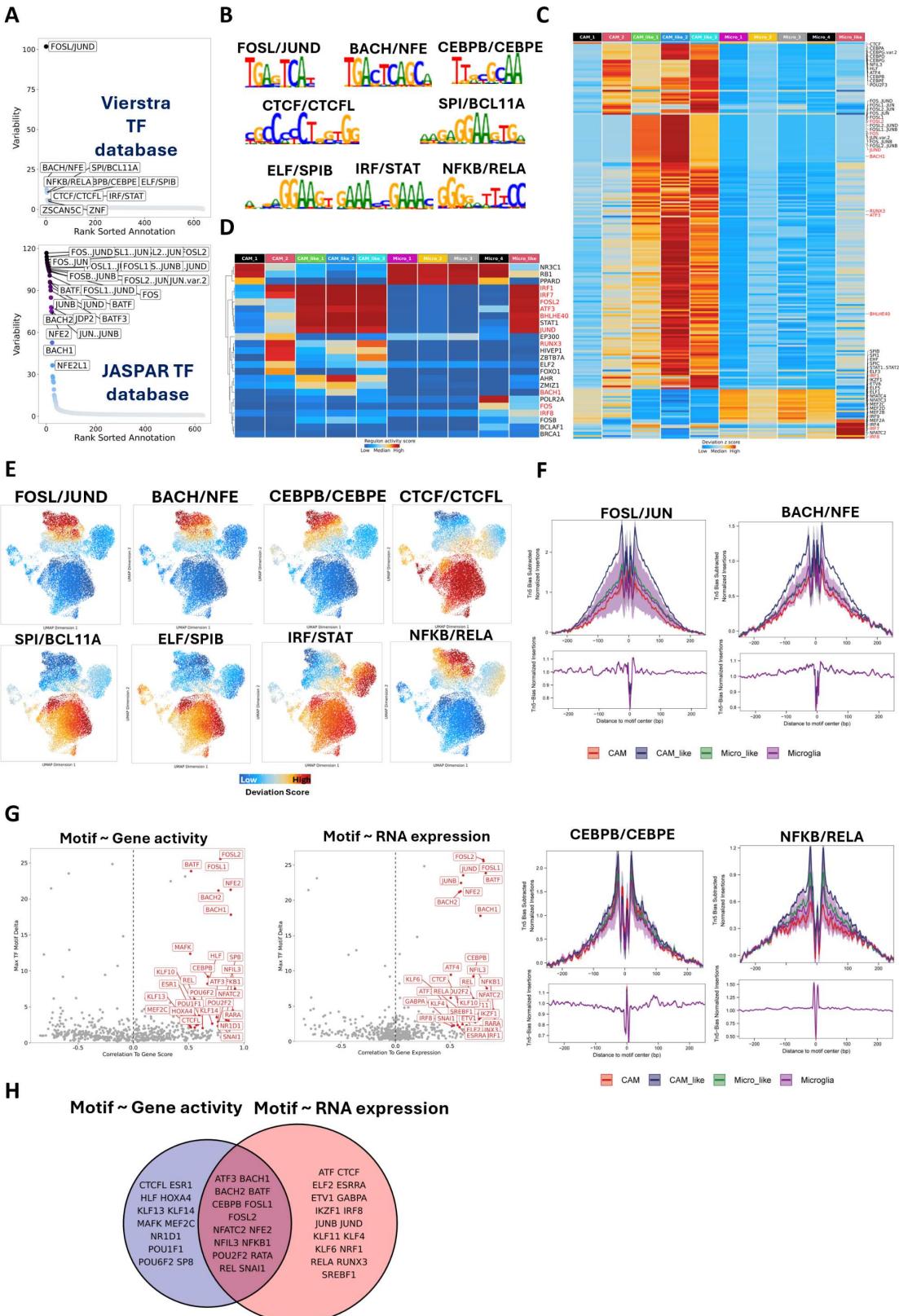


Figure 5. Transcription factor enrichment for different myeloid cell clusters. (A) Variability for enriched TF motifs in Vierstra (upper) and JASPAR (lower) databases. The ranking is based on TF motif variability over all accessible chromatin regions. The top 10 TF motifs from the Vierstra database and the top 25 TF motifs from the JASPAR database were labeled. (B) The Vierstra

archetype consensus of the binding sites for selected high-variable TF complexes. **(C)** Heatmap for motif deviations of TF motifs with an FDR ≤ 0.001 and mean difference ≥ 1 over different cell clusters. The TF motifs were annotated by using the JASPAR database. Top 10 TFs for each cluster were labeled and the TFs that were also predicted by SCENIC (showed in panel D) were labeled in red. **(D)** Heatmap for regulon activity of regulons predicted by SCENIC over different cell clusters. The regulons or TFs that were also found as markers in snATAC motif analyses (panel C) were labeled in red. **(E)** UMAP projection of motif enrichment for selected high-variable TF complexes (annotated by the Vierstra) in different cell clusters. **(F)** The footprints for selected high-variable TF complexes. **(G and H)** Positive TF-regulators whose RNA expression or predicted gene activity positively correlates with the accessibility of their corresponding motif.

263

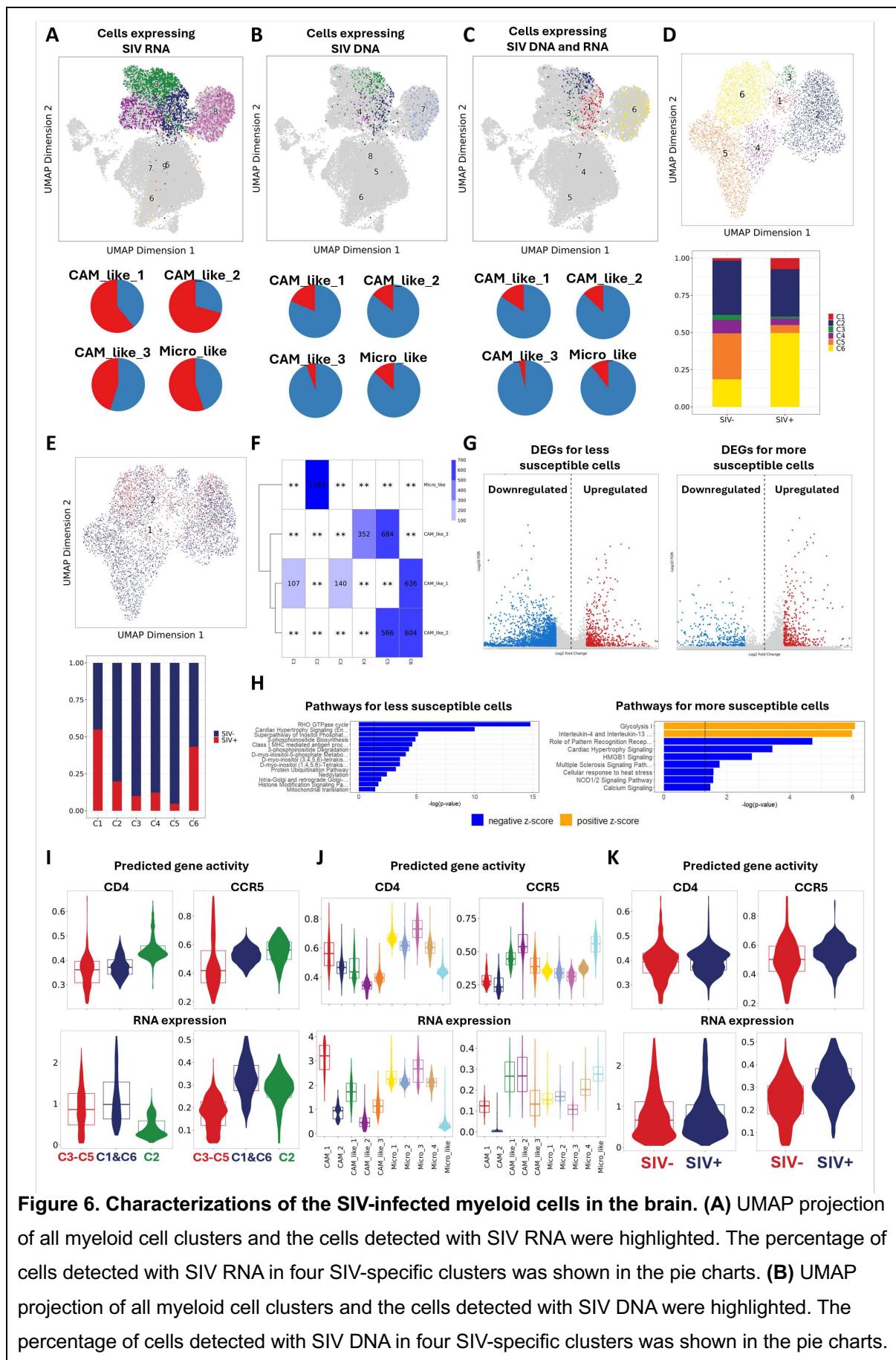
264 expression and predicted gene activity of most of those TF genes (e.g. NFKB1, NFKB2,
265 RELA, BACH1, BACH2, NFE2, and CEBPB) were positively correlated with their binding
266 motif enrichments, suggesting they might be positive TF regulators. However, the RNA
267 expression and predicted gene activity of STAT1, STAT2, IRF1, and CEBPE were negatively
268 correlated with their binding motif enrichment,³³ which reflected more complicated regulatory
269 mechanisms. To identify more positive TF regulators whose gene expression or activity is
270 positively correlated with the accessibility of their binding motifs, we performed correlation
271 analyses of TF gene expression or predicted gene activity and their corresponding motif's
272 enrichment (**Figure 5G**). The positive TF regulators that were identified by both correlations
273 (**Figure 5H**) included many TFs mentioned above predominantly upregulated in CAM-like
274 cells (e.g. BACH1, BACH2, NFE2, CEBPB, FOSL1, FOSL2, NFKB1, and REL), highlighting
275 their critical roles in inducing inflammation in this cell type.

276 Enriched TFs in SIVE-specific clusters also correlated with some inflammatory and
277 anti-viral genes' RNA expression or predicated gene activity. For example, the RNA
278 expression of IL1 β , IL10, and CXCL1 that were regulated by NFKB/RELA (p50/p65) showed
279 a strong positive correlation ($r \geq 0.95$) with the motif deviation Z-score of NFKB/RELA
280 (**Figure S4**). The CAM-like and Microglia-like cells with high enrichment of NFKB/RELA had
281 high expression of IL1 β , IL10, and CXCL1. A strong positive correlation was also found for
282 type I IFNs (i.e., IFN α 1, IFN α 8, and IFN ϵ) and their TFs STAT1/STAT2 (**Figure S4**). The

283 Microglia-like and Microglia cells with high enrichment of STAT1/STAT2 also had high
284 predicted gene activity of IFN α 1, IFN α 8, and IFN ϵ . Those results further provide the
285 evidence that the inflammation observed in SIVE-specific clusters was caused by TF shifting
286 in the brain myeloid cells.

287 **Brain myeloid cell susceptibility to SIV infection is linked to certain enriched cellular
288 pathways and CCR5 expression.**

289 We added the SIV proviral genome as an additional “chromosome” in the sequence
290 searches to enable the identification of both DNA with accessible open chromatin and RNA
291 gene expression. In the two infected animals with SIVE (i.e. 21T and 34T), SIV DNA
292 sequences (**Figure 6A**) were identified in 8.9% of the brain myeloid cells in 21T and 15.9%
293 in 34T. More SIV RNA transcripts (**Figure 6B**) were found in myeloid cells in those two SIV-
294 infected animals: 44.9% in 21T and 67.7% in 34T. Neither DNA nor RNA SIV sequences
295 were found in the two uninfected animals (104T and 106T). However, although ATAC
296 analysis contains more sparse data than RNA analysis, the high expression proportion of
297 SIV-infected myeloid cells raised the possibility of false positives. This is in keeping with the
298 analyses of Plaza-Jennings et al., who examined HIV expression in nuclear preps from
299 brains with HIVE and found that the high expression of HIV in myeloid cells in the
300 encephalitis condition led to contamination of uninfected cells with viral messages, indeed
301 leading to false positives.³⁴ Since there were no definitive means to determine which cells
302 truly expressed SIV RNA, we chose a conservative method, requiring both SIV DNA
303 fragments from snATAC-seq and RNA transcripts from snRNA-seq in the same cell to call a
304 cell SIV-positive. (**Figure 6C**) This restriction for defining SIV-positive cells resulted in 6.7%
305 SIV+ cells in 21T and 13.5% SIV+ cells in 34T. The percentage of SIV+ cells found in 21T
306 and 34T was close to the numbers we previously reported using scRNA-seq (7.6% and
307 12.8%, respectively) from FACS-sorted intact brain myeloid cells,²³ which should contain
308 little ambient contaminating RNA, thus giving us confidence in this method.



(C) The UMAP projection of all myeloid cell clusters and the cells detected with SIV RNA and DNA were highlighted. The percentage of cells detected with both SIV RNA and DNA in four SIV-specific clusters was shown in the pie charts. **(D and E)** The UMAP projection for reclustered cells only included SIV⁺ and SIV⁻ cells in two infected animals. The distribution of SIV⁺ and SIV⁻ cells in the six new clusters was shown in aggregated bar charts. **(F)** Distribution of the cells in new clusters (i.e., C1-C6) and previously defined clusters (i.e., CAM_like_1, CAM_like_2, CAM_like_3, Micro_like). The number of cells was labeled. **(G)** The upregulated marker genes (red) and downregulated marker genes (blue) for the less susceptible clusters (C3-C5) and the more susceptible clusters (C1 and C6). The statistical test used for identifying markers was the Wilcoxon rank test, and the cut-off was set as FDR ≤ 0.01 and absolute log₂ fold change ≥ 1. **(H)** The top upregulated or downregulated pathways detected for the cells in the less susceptible cluster (left) and the more susceptible clusters (right) by IPA. The cut-off is set as -log(p-value) > 1.3 and |z| > 4.0 for less susceptible clusters and |z| > 2.5 for more susceptible clusters. **(I)** Violin plots of RNA expression and predicted gene activity of CD4 and CCR5 receptors between the more susceptible and less susceptible clusters. **(J)** Violin plots of RNA expression and predicted gene activity of CD4 and CCR5 receptors among all myeloid cell clusters. **(K)** Violin plots of RNA expression and predicted gene activity of CD4 and CCR5 receptors between SIV⁺ cells and SIV⁻ cells.

309

310 We then examined the distribution of cells found with SIV RNA and DNA fragments in
311 different myeloid cell clusters. To examine whether there might be a potential myeloid cell
312 phenotype in the brain that was more susceptible to SIV infection, we identified SIV⁺ cells,
313 defined as above (both DNA+ and RNA+), as well as SIV⁻ cells, which we defined as those
314 without SIV DNA or RNA detected. To eliminate the effect of the SIV viral expression itself on
315 clustering, we excluded the whole SIV genome when we reduced dimensionality and
316 subsequently clustering for this subset dataset. The clustering resolution was set as 0.2,
317 which gave us six cell clusters (**Figure 6D and 6E**). Interestingly, most SIV⁺ cells were found
318 in C2 and C6 (**Figure 6D**), but the C2 cluster had even more SIV⁻ cells, which resulted in
319 a low infection rate in this cluster (**Figure 6E**). On the other hand, we found that over 50% of
320 cells in the C1 cluster were SIV⁺, although fewer cells were found in this cluster. Given the
321 high infection rate found in C1 and C6 clusters, they were deemed more susceptible to SIV
322 infection. Then we compared these newly identified clusters (i.e. C1-C6) with the previously
323 identified SIVE-specific clusters (i.e. Micro_like, CAM_like_1, CAM_like_2, CAM_like_3) to
324 assess the phenotypes of these clusters (**Figure 6F**). We found that the cells in the

325 microglia-like cluster were exclusively distributed in the C2 cluster, indicating that the C2
326 cluster had a microglia-like phenotype, and the other 5 clusters had a CAM-like phenotype.
327 The cells in the CAM_like_1 cluster were clustered in C1, C3, and C6 clusters, and the cells
328 in the CAM_like_2 cluster were clustered in C5 and C6 clusters. In summary, the
329 CAM_like_1, and possibly CAM_like_2, cluster, might be more susceptible to infection.

330 To better understand what genes might cause the cells in C1 and C6 to be more
331 susceptible to SIV infection, we found the genes that were upregulated and downregulated
332 in the C3-C5 clusters and the C1 and C6 clusters by comparing RNA expression (**Figure**
333 **6G**). Enriching those upregulated or downregulated genes into pathway analyses, we found
334 that the RHO GTPase cycle was downregulated in less susceptible clusters, as were the
335 pathways related to cell activity, including biosynthesis, metabolism, and phosphorylation,
336 (**Figure 6H**). The more susceptible cells had enhanced glycolysis and IL4 and IL13
337 signaling, but lower PPR signaling via TLRs and NODs.

338 We also found that the principal receptors (i.e. CD4 and CCR5) assisting virus entry
339 into the cells were upregulated in the C1 and C6 clusters (**Figure 6I**), which helps explain
340 the reasons why there were more SIV+ cells found in those two clusters. Additionally,
341 compared to microglia and CAMs found in the uninfected condition, the CAM-like and
342 microglia-like cells found in SIVE downregulated the expression and chromatin accessibility
343 of CD4 (**Figure 6J**). HIV-induced downregulation of CD4 mRNA (as well as the CD4 protein)
344 has been found in T cells.³⁵ When we compared the expression and chromatin accessibility
345 of CD4 and CCR5 in SIV+ cells and SIV- cells in the infected animals, we only found that the
346 RNA expression of CCR5 was slightly higher in SIV+ cells (**Figure 6K**). Increased
347 expression of the CCR5 protein has been found in microglia and CAM in the brains of those
348 with HIVE.³⁶ In total the changes from the uninfected state were likely due to the global effect
349 of inflammation.

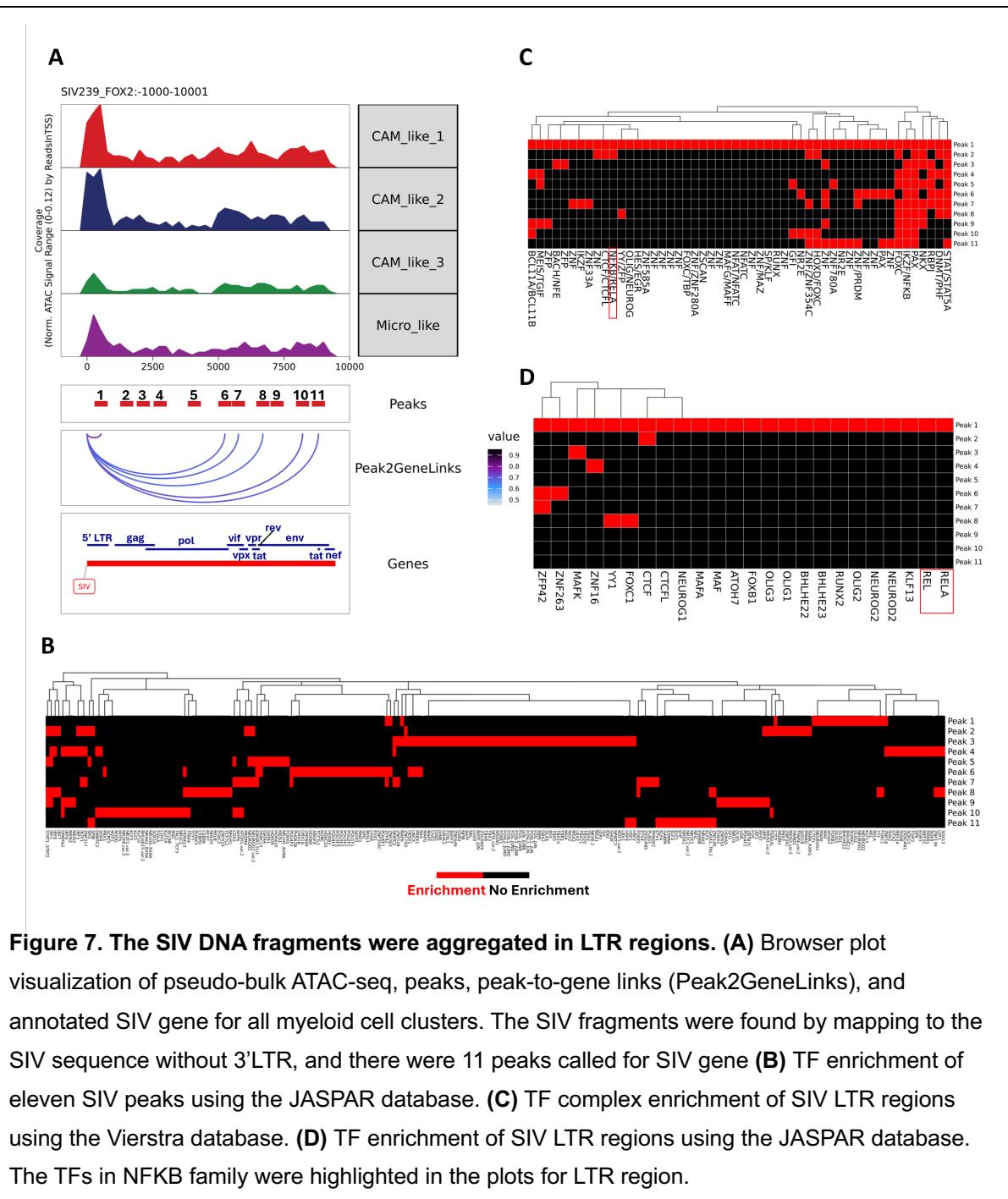
350 **The LTR regions of the SIV DNA found in brain myeloid cells were more accessible**
351 **than other SIV gene regions.**

352 We included the genome of SIV in our annotation as an extra chromosome to identify
353 SIV-infected cells. As discussed in the Methods section, we tested different methods to map
354 the SIV genome. Using the method that kept most SIV fragments, we found that the LTR
355 regions of the SIV DNA were more accessible compared to other regions. (**Figure 7A and**
356 **S6A**). The more infection-susceptible CAM_like_1 and CAM_like_2 clusters had higher peak
357 signals at LTR regions than CAM_like_3 and microglia-like clusters. The regions encoding
358 Pol, Vif, Vpr, Vpx, and envelope protein (peaks 6, 7, 8, 10, and 11) also showed more
359 accessibility and a strong association with the SIV RNA expression (**Figure 7A**).

360 To understand what host TFs myeloid cells might use for SIV transcription, we
361 performed TF motif enrichment specifically for the SIV peaks. While Peak 3 (coordinate:
362 1891-2391) had a large number of enriched motifs, with most related to FOS/JUN (**Figure**
363 **7B**), in the LTR there was enrichment for transcription factors related to NF- κ B (e.g. REL,
364 RELA, NFKB/RELA), which were very specific for the LTR region (**Figure 7C and 7D**). This
365 further validates that the SIV in the infected brain myeloid cells could use host NF- κ B TFs to
366 express its genes, as previously reported.³⁷⁻³⁹

367 **Neuroinflammation in animals with SIVE might further differentiate myeloid cells
368 toward the CAM-like phenotype.**

369 Given the plasticity nature of myeloid cells in response to the inflammation, we
370 performed the trajectory analyses to find the differentiation hierarchy and the potential genes
371 and TFs driving the differentiation. We first used unsupervised trajectory analyses for
372 pseudo-time prediction. The homeostatic microglia cluster found in uninfected conditions and
373 with most microglia (Micro_1) was set as the root node (**Figure S5A**). The cells in the
374 CAM_like_2 and microglia-like clusters had the highest pseudo-time values, indicating they
375 might be at the late stage of differentiation. The end of the trajectory path with Micro_1 as
376 the root node pointed to microglia-like cells but not CAM-like cells. We then set CAM_1, the
377 classical macrophage population in the uninfected brain, as the root node for prediction
378 (**Figure S5B**). Again, the end of the trajectory pathway pointed to microglia-like cells but not



379

380 CAM-like cells, which suggests CAM-like cells might not arise from microglia or CAMs found
381 in uninfected brains. In addition, we found that the gene expression and predicted gene
382 activity for the CEBPB gene, which becomes active during the differentiation process, were
383 higher in CAM-like cells compared to microglia-like cells (**Figure S5A and S5B**), indicating
384 CAM-like cells might be more polarized than microglia-like cells. Therefore, it is possible that

385 during SIV infection, the homeostatic microglia first become microglia-like cells by
386 suppressing the transcription of homeostatic microglial core genes. However, the chromatin
387 accessibility of homeostatic microglial core genes did not decrease at this stage of
388 differentiation (**Figures 2D and 2F**). With the progression of neuroinflammation, it is possible
389 that some activated microglia-like cells might further differentiate into CAM-like cells, which
390 completely shut down the chromatin accessibility for homeostatic microglial core genes and
391 show more pathogenic.

392 Based on this hypothesis, we performed trajectory analyses only for the SIVE-
393 specific clusters (**Figure 8A**). The trajectory pathway started from microglia-like cells and
394 ended in the CAM_like_2 cluster, which was found to be more pathogenic in the above
395 analyses. The predicted pseudo-time for the SIVE-specific clusters positively correlated with
396 the RNA expression and chromatin accessibility (predicted gene activity) of the CEBPB gene
397 (**Figure 8A**), confirming that this differentiation path is reasonable. Then, we further
398 visualized the changes in RNA expression (**Figure 8B**), chromatin accessibility (**Figure**
399 **S5C**), peaks (**Figure S5D**), and motif enrichment (**Figure 8C**) across this differentiation.
400 From the RNA expression changes (**Figure 8B**), we found that the differentiation process
401 decreased the pathogen recognition and alarm responsiveness of the brain myeloid cells
402 since the RNA expression for complements and scavenger receptors was absent in CAM-
403 like cells. The RNA expression of heat shock proteins was also downregulated during the
404 differentiation from microglia-like cells to CAM-like cells. On the other hand, this
405 differentiation leads to enhanced expression of MHC class II molecules, S100 proteins and
406 interferon-inducible proteins. From the changes in enriched TFs (**Figure 8C**), the
407 differentiation shifted the motif enrichment from STAT and NFATC families to CEBP, BATF,
408 FOS/JUN, and KLF families. The rise in binding sites for inflammatory-linked TF found in
409 end-differentiated cells suggested that the differentiation of myeloid cells from microglia-like
410 to CAM-like is accompanied by increasing regulatory events for genes contributing to
411 inflammation. To better understand the TF regulation during this differentiation/activation, we

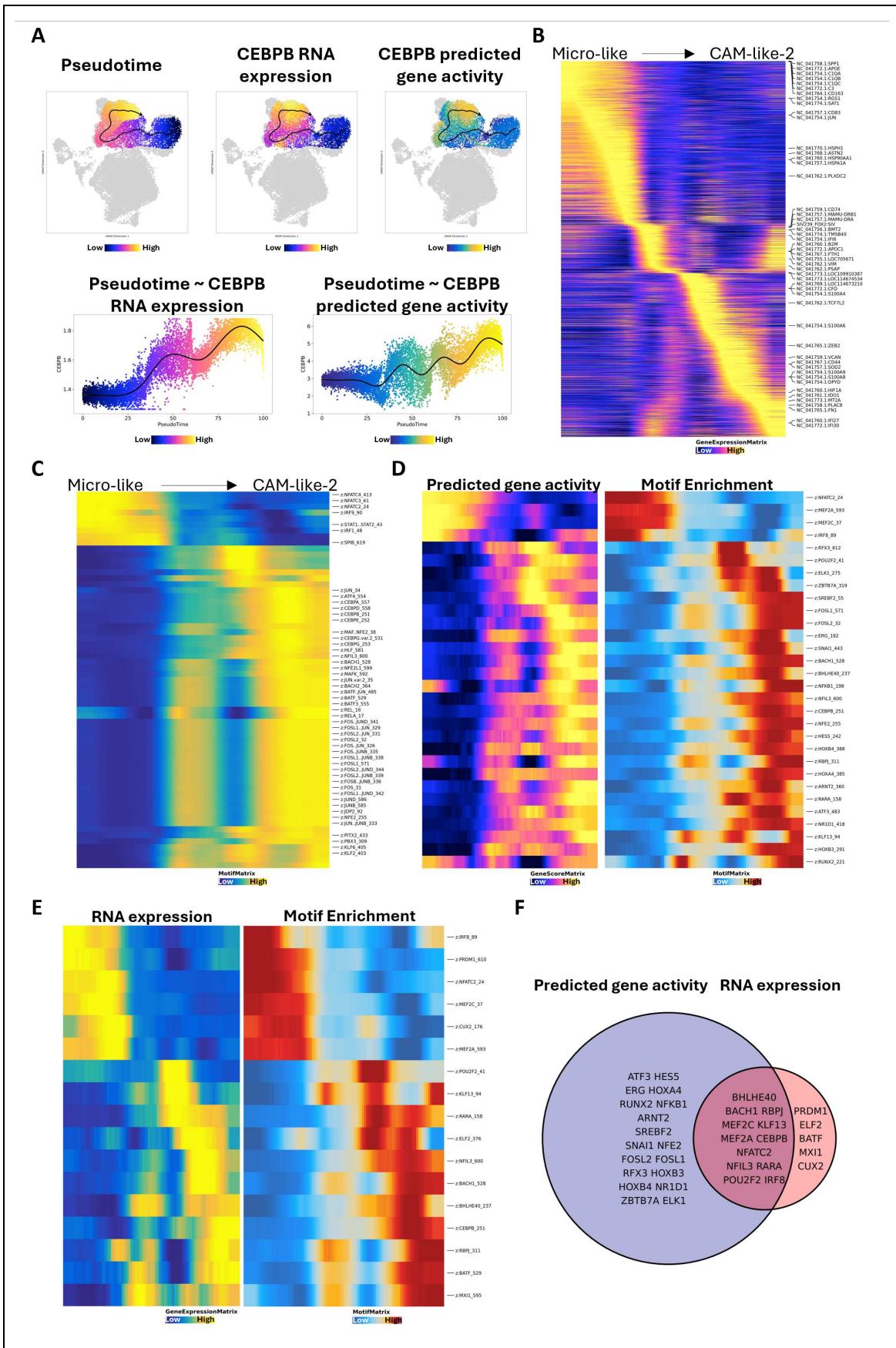


Figure 8. Identifying the genes and transcription factors driving the differentiation of brain myeloid cells during SIVE. **(A)** The pseudo-time for the brain myeloid cells was mainly found in SIVE animals (left), and the correlation between cell pseudo-time and predicted gene activity (middle) or RNA expression (right) of the CEBPB gene. **(B)** The changes of the genes regarding their RNA expression across pseudo-time. **(C)** The changes of the transcription factors (TFs) regarding their motif enrichment across pseudo-time. **(D)** The positive TF regulators with high predicted gene activity and motif enrichment across pseudo-time. **(E)** The positive TF regulators with high RNA expression and motif enrichment across pseudo-time. **(F)** The positive TF regulators were identified by both correlations (shown in D and E), which might drive the differentiation of brain myeloid cells in SIVE.

412

413 correlated the RNA expression or predicted gene activity (chromatin accessibility) with the
414 TF motif accessibility across pseudo-time. We identified the positive TF regulators that might
415 drive this differentiation process (**Figure 8D and 8E**). Many common positive TF regulators
416 were potentially found to drive this differentiation (**Figure 8F**). For example, IRF8, NFATC2,
417 MEF2A, and MEF2C had higher RNA expression, predicted gene activity, and motif
418 accessibility in the cells at an early stage of differentiation, suggesting the activity of those
419 TFs might relate to the polarization of microglia-like cells to CAM-like cells. However,
420 microglia-like cells and CAM-like cells may be from very different progenitor cells that are not
421 interchangeable. Such cells may come from blood myeloid cells that enter the brain in
422 inflammatory conditions such as SIVE. It is also possible that CAM-like cells and microglia-
423 like cells can be polarized to each other, responding to different inflammatory milieu.

424

425 **DISCUSSION**

426 Myeloid cells in the brain are essential in inducing neurocognitive disorders once
427 activated. In HIV infection, the invading virus can reshape the immune milieu in the brain,
428 which further restructures the myeloid cell constitution. This study found that the myeloid cell
429 phenotypes in the brain with SIVE significantly differed from those in the uninfected brains.
430 Given the advantages of multiomic sequencing, we found four different phenotypes in the

431 infected encephalitic brain, broadly classified into two primary phenotypes: microglia-like and
432 CAM-like. The CAM-like myeloid cells were more pathogenic and wholly lost the expression
433 and chromatin accessibility of homeostatic microglial core accessibility genes. On the other
434 hand, microglia-like myeloid cells suppressed the RNA expression of homeostatic microglial
435 core genes but surprisingly kept the chromatin accessibility of those genes.

436 By assessing the microglia-like and CAM-like phenotypes using snRNA-seq and
437 snATAC-seq information, we found that the microglia-like cells (relative to microglia cells)
438 had increased RNA expression and chromatin of SPP1 and chemokines, and CAM-like cells
439 (relative to CAM cells) upregulated S100 proteins. Cell communication analyses further
440 revealed the potential biological consequences for upregulating SPP1 and chemokines in
441 microglia-like cells. The SPP1 highly expressed in microglia-like cells was likely to bind with
442 the CD44 on CAM-like cells. The interaction between SPP1 and CD44 has been widely
443 reported in cancer for immune cell infiltration^{40,41}, and those studies also highlighted the
444 ability of SPP1-CD44 interaction to modulate cell adhesion and movement. Combining the
445 facts that chemokines also actively induce infected/activated myeloid and lymphoid cell
446 migration in HIV infection,^{42,43} the microglia-like phenotype might be the main initiator of
447 inflammatory cell chemotaxis in the brain with SIVE. CCL3 and CCL5 are natural ligands for
448 the primary HIV-1 coreceptor CCR5, which may prevent HIV from entering cells.^{44,45} Thus,
449 microglia-like cells might also contribute to preventing SIV infection by secreting those two
450 chemokines. The CAM-like cells might possess a completely different biological function in
451 SIVE.²⁹ Although S100A8/A9 is reported to suppress HIV replication in macrophages,⁴⁶ their
452 ability to promote inflammatory conditions such as lethal endotoxin-induced shock²⁹ can lead
453 to brain damage. The other S100 proteins were also found to have higher expression in
454 CAM-like cells but not microglia-like cells, and their function is to serve as alarmins to induce
455 and amplify the immune response through degranulation.

456 The TF motif enrichment in CAM-like and microglia-like cells further indicated their
457 differences. Although most TFs enriched in SIVE-specific clusters were significantly different

458 from TFs enriched in CAM or microglia, the CAM-like cells were enriched with much more
459 inflammatory TF motifs than microglia-like cells. The TFs in the IRF/STAT family enriched in
460 microglia-like cells were also enriched in microglial clusters (found in uninfected brains) but
461 were barely enriched in CAM-like cells. Correspondingly, some type I IFNs were also highly
462 expressed in microglia-like and microglia cells but not CAM-like cells. The TFs in the
463 IRF/STAT family regulated the expression of type I interferons, essential for anti-viral
464 activity.⁴⁷⁻⁴⁹ Additionally, the anti-inflammatory properties of type I IFNs⁵⁰⁻⁵² further suggested
465 the potential immunoregulatory roles of microglia-like cells in SIVE brains. On the other
466 hand, the CAM-like cells were highly enriched with FOS/JUN (AP-1) and NFKB/RELA, the
467 TF complexes mediating immune activation in brain myeloid cells.^{53,54}

468 While the microglia-like and CAM-like cells appeared activated to defend against the
469 invading virus in SIVE, many infected myeloid cells could still be found in the brain. Our
470 estimation based on SIV RNA transcripts and DNA fragments indicates that the infection rate
471 in brain myeloid cells was ~ 10% in SIVE. Furthermore, we found that some cells in CAM-
472 like clusters, mainly in the CAM_like_3 cluster, were less likely to be infected. However,
473 other CAM-like cells were more likely to be infected, mainly in the CAM_like_1 and
474 CAM_like_2 clusters. By assessing the upregulated and downregulated genes and pathways
475 between more susceptible and less susceptible brain myeloid cells in SIVE, we found that
476 the less susceptible cells significantly downregulated many cellular activities. However, we
477 still do not know if the suppressed cellular activities caused the lower susceptibility to SIV
478 infection or if the lower levels of SIV infection led to those cells being comparatively less
479 active. Regarding infection, the more susceptible cells upregulated the expression of CCR5
480 but not CD4. Both CD4 and CCR5 can assist the HIV/SIV entering the cells, but their
481 changes in response to SIV infection were very different (**Figure 6J**). The downregulation of
482 CD4 has been widely reported during HIV infection, and this downregulation is thought to be
483 caused by Nef proteins for stimulating HIV-1 production and infectivity.⁵⁵⁻⁵⁷ Upregulation of

484 CCR5 may be important in susceptibility to infection because it serves as a critical receptor
485 for HIV and SIV infection, and its modulation correlated with infectability.^{58,59}

486 The advantages of ATAC-seq inclusion enabled us to understand SIV DNA in the
487 host nucleus. Most of the accessible chromatin found in the SIV DNA were at the SIV long
488 terminal repeats (LTRs), which are essential regions for viral gene expression and latency.
489 The proviral genomes of HIV are flanked by two LTRs. Each LTR consists of three regions:
490 U3, R, and U5.⁶⁰ Among those LTR regions, the U3 region contains promoter, enhancer, and
491 modulatory elements that regulate the expression of viral genes. The promoter region
492 contains three Sp1 binding sites, and the enhancer region contains two NF- κ B and one
493 NFAT binding site⁶¹⁻⁶⁴, which were also enriched at the SIV LTR region in our dataset. We
494 also found more de novo host TF binding sites that have not been reported, and their
495 potential involvement in modulating HIV/SIV gene expression needs further investigation. In
496 addition to the LTR, five additional peaks within the SIV genome identified by snATAC-seq
497 were strongly correlated with expression. As open chromatin is more accessible for
498 modification by methods such as CRISPR-Cas9, this may provide new regions to target to
499 shut down or eliminate HIV/SIV from myeloid cells.

500 In conclusion, the high-throughput multiomic sequencing technique highly promotes
501 the identification of heterogeneous myeloid cell populations in uninfected brains and brains
502 with SIVE. Including simultaneous snATAC-seq with snRNA-seq enabled a deeper
503 understanding of the gene regulatory events in brain myeloid cells, especially during SIVE.
504 However, due to the many considerations for the investigations related to NHPs, the results
505 and conclusions in this study, which were generated from a limited number of SIVE animals,
506 will need further study. The deposition of sequence data and metadata in publicly accessible
507 databases from our studies and others enables the building of larger analyses with more
508 subjects. These data can be useful in meta-analyses across models and disease states.

509

510

511 **MATERIALS AND METHODS**

512 **Animals**

513 The four male adult rhesus macaques used in this study tested negative for the indicated
514 viral pathogens: SIV, SRV, STLV1, Herpes B-virus, and measles; and bacterial pathogens:
515 salmonella, shigella, campylobacter, yersinia, and vibrio. Macaques were housed in
516 compliance with the Animal Welfare Act and the Guide for the Care and Use of Laboratory
517 Animals in the NHP facilities of the Department of Comparative Medicine, University of
518 Nebraska Medical Center (UNMC). The American Association for Accreditation of Laboratory
519 Animal Care International has accredited the primate facility at UNMC. The UNMC
520 Institutional Animal Care and Use Committee (IACUC) reviewed and approved this study
521 under protocols 19-145-12-FC and 16-073-07-FC. Animals were maintained in a
522 temperature-controlled ($23 \pm 2^\circ \text{ C}$) indoor climate with a 12-h light/dark cycle. They were fed
523 Teklad Global 25% protein primate diet (Envigo, Madison, WI) supplemented with fresh fruit,
524 vegetables, and water ad libitum. The animal care and veterinary personnel observed the
525 monkeys twice daily to check their health status. Two of the four animals (21T and 34T) were
526 intravenously inoculated with a stock of SIVmac251 and developed SIV-induced encephalitis
527 (SIVE). The other two macaques (104T and 106T) were uninfected and used as controls.
528 Results from scRNA-seq of microglia for 21T (who had been treated with methamphetamine)
529 and 34T have been previously reported.²².

530 **Viral loads**

531 To determine the terminal viral load in plasma, the blood of infected animals (21T and 34T)
532 was collected. The EDTA-anticoagulated plasma was separated from blood by
533 centrifugation. Brain and lymphoid organ specimens were taken to determine viral load in
534 tissues. Plasma and tissue SIV RNA levels were determined SIV RNA was measured using
535 the branched DNA assay by Siemens (Emeryville, CA). The viral load in plasma, lymphatic
536 tissues, and brain of those two animals could be found in our previously published paper.²²

537 **Immunohistochemical staining**

538 Brains were fixed in 10% neutral buffered formalin, embedded in paraffin, cut into 5 µm
539 sections, and mounted on glass slides. For immunohistochemistry, sections were
540 rehydrated, and endogenous peroxidase activity was blocked by a 3% hydrogen peroxide
541 treatment in absolute methanol. Following that, a heat treatment with 0.1 M citrate pH 6.39
542 was performed for antigen exposure. Sections were blocked with 5% Normal Horse Serum
543 (Vector Labs, Burlingame, CA, USA) in PBS and incubated with the primary antibody diluted
544 in the same buffer. Antibodies were targeted against S100A9 (Cat# PA5-79949, dThermo
545 Fisher Scientific, Rockford, IL, USA) at a 1-10,000 dilution. Biotinylated secondary antibodies
546 (horse anti-rabbit IgG Cat# MP-7401, Vector Labs, Burlingame, CA, USA) were used.
547 Visualization was achieved using Nova Red (Vector Labs, Burlingame, CA, USA).
548 Counterstaining was done using Gill 2 Hematoxylin (StatLab, McKinney, TX, USA).

549 **Isolation of myeloid cells in the brain**

550 92 days (for 21T) and 49 days (for 34T) after viral inoculation, a necropsy was performed
551 due to symptoms in the animals consistent with simian AIDS. Uninfected animals 104T and
552 106T were necropsied per approved animal protocol. Animals were deeply anesthetized with
553 ketamine plus xylazine and blood cleared from the brain and other organs by intracardial
554 perfusion with sterile PBS containing 1 U/ml heparin. Brains were harvested, and
555 approximately half of the brain was taken for microglia/macrophage isolation.
556 Microglia/macrophage-enriched cellular isolation was performed using our previously
557 described procedure.⁶⁵ Briefly, the brain was minced and homogenized in cold Hank's
558 Balanced Salt Solution (HBSS, Invitrogen, Carlsbad, CA). After being centrifuged, the brain
559 tissue was digested at 37° C in HBSS containing 28 U/ml DNase I and 8 U/ml papain for 30
560 minutes. After digestion, the enzymes were inactivated by adding 2.5% FBS, and the cells
561 were centrifuged and resuspended in cold HBSS. The cell suspension was mixed with 90%
562 Percoll (GE HealthCare, Pittsburg, PA) to a final concentration of 20% Percoll and
563 centrifuged at 4° C for 15 minutes at 550 x g. The microglia/macrophage pellet at the bottom
564 was resuspended in HBSS and passed through a 40 µm strainer to remove cell clumps and

565 aggregates. Cells were again pelleted by centrifugation and resuspended in RBC lysis buffer
566 for 3 minutes to eliminate contaminating red blood cells. A final wash was performed before
567 the resulting cells were quantified on a hemocytometer and Coulter Counter Z1. The cells
568 were resuspended in 10% DMSO, 90% fetal bovine serum and subjected to slow controlled
569 freezing followed by storage in liquid N₂. The methods for cryopreservation followed our
570 previous study, which was found to maintain the vast majority of the transcriptomic features
571 of fresh isolated microglia/macrophages.⁶⁵

572 **Single nuclei preparation and multiomic sequencing (ATAC and gene expression)**

573 Cryopreserved cell isolates were rapidly thawed in a 37° C water bath. The cell
574 recovery procedures were well described in our previous publications.⁶⁵ After the recovery,
575 cells were washed and counted by Coulter Counter Z1. Once cell concentration was known,
576 cells were transferred to ice-cold PBS and stained with CD11b (Biolegend 101257) and UV-
577 blue live/dead (Invitrogen L23105). Cells were washed, resuspended in flow cytometry
578 staining buffer (e-bioscience), and sorted on an Aria2 flow cytometer (BD Biosciences, San
579 Jose, CA, USA).

580 Nuclei were isolated from microglia cells using the 10XGenomics protocol. Nuclei
581 were then quantified on a hemocytometer and concentrated to approximately 2200-2400
582 nuclei per μL. Based on 10× Genomics parameters targeting 8000 nuclei, the ideal volume
583 of cells was loaded onto the 10× Genomics (Pleasanton, CA, USA) Chromium Next GEM
584 Chip J and placed into the Chromium Controller for nuclei capturing and library preparation.
585 The prepared libraries were sequenced using Illumina (San Diego, CA, USA) Novaseq6000
586 sequencers. The sequences have been deposited in NCBI GEO (accession number
587 GSE272669).

588 **Pre-processing of multiomic data**

589 Sequenced samples were processed using the 10× Genomics Cell Ranger ARC pipelines
590 (2.0.2). Specifically, the multiomic data was demultiplexed and aligned to a customized
591 genome combining the rhesus monkey reference genome (NCBI RefSeq assembly

592 Mmul_10) and an SIV genome that we constructed by sequencing our virus stock (NCBI
593 GenBank, accession number PP236443). Since SIV had two LTR regions with repeating
594 sequences and the Cell Ranger arc count pipeline discarded the reads mapped to two
595 identical regions (MAPQ mapping score < 30), we built three testing reference genomes
596 containing the same rhesus macaque genome but different SIV genome (i.e., remove 3' LTR
597 region, remove 5' LTR region, and keep the whole length of SIV genome) for Cell Ranger
598 ARC count to test on two infected samples. After comparing the percentage of reads
599 mapped to the SIV genome but with low MAPQ mapping scores (<30) and the number of
600 SIV DNA fragments in fragment files generated by Cell Ranger among the three tests, we
601 found that removing 3'LTR and keeping 5'LTR regions allowed minimized the low-quality SIV
602 reads (**Figure S6B**) and maximized the SIV DNA fragments (**Figure S6C**) to pass the
603 filtering of Cell Ranger ARC count. Therefore, we decided to use the genome containing the
604 Mmul10 rhesus macaque genome and the SIV genome without 3'LTR for aligning and
605 counting all the samples. After the Cell Ranger ARC count pipeline, the filtered feature
606 barcode matrices containing the RNA count and fragment files containing the ATAC fragment
607 count were obtained. The counting summary statistics generated by 10x Genomics for each
608 sample are shown in **Table S5**.

609 **Characterization of cell phenotypes in muti-omics dataset**

610 Pre-processed ATAC-seq data from multiomic sequencing processed with Cell
611 Ranger arc were read with the ArchR R package (version: 1.0.2). We built our own gene and
612 genome annotation used in ArchR by the BsgeomeForge R package and the
613 GenomicFeature R package, and the input files for those two packages were the gtf and fa
614 files that were used to build the reference genome in Cell Ranger. After successfully reading
615 the fragment files, we performed several steps of quality controls for the cell barcodes in the
616 fragment files. Firstly, the cell barcodes were filtered based on the ATAC data, and the cell
617 barcodes with less than 1000 fragments per cell and a TSS enrichment score of less than
618 four were removed. Then, we integrated ATAC-seq data with filtered RNA count matrices by

619 using the import10xFeatureMatix function in the archR package, and the cell barcodes with
620 both RNA data and ATAC-seq data were retained. After integration, we filtered the cell
621 barcodes based on the RNA data. The cell barcodes with gene count and UMI count of less
622 than 400 and mitochondria percentage of more than 15% were removed. Finally, we
623 identified and removed the inferred doublets using addDoubletScore and filterDoublets
624 functions in archR. After filtering, 24,633 cells were left for downstream analyses. The QC
625 plots showed that all four samples had a similar distribution of their TSS enrichment score,
626 number of unique fragments, UMI counts, and gene counts (**Figure S7A and S7B**). After
627 quality control, dimensionality reductions were performed separately using the iterative
628 Latent Semantic Indexing (LSI) approach for RNA and ATAC data. The parameters set for
629 implementing LSI in archR were iterations = 2, res = 0.2, maxClusters =6, varFeatures =
630 25000 (for ATAC), varFeatures = 2500 (for RNA) dimsToUse = 1:30, LSI method = log(tf-idf),
631 and the other unmentioned parameters followed the default. The addCombinedDims function
632 in archR combined the dimensionality reductions from ATAC and RNA. The batch effect
633 between different samples was addressed using the Harmony implementation of archR.
634 Then, clustering was performed using dimensionalities combining RNA and ATAC data, and
635 the resolution was set at 0.2. Finally, UMAP dimensionality reduction was performed to
636 visualize data, and again, the combined dimensionalities were used. The following
637 characterizations of each cell cluster were performed in two ways. Firstly, we screened the
638 normalized RNA expression and predicted gene activity of the gene markers for microglia
639 (e.g. P2RY12, GPR34, CX3CR1), CNS-associated macrophages (e.g. MAMU-DRA, MAMU-
640 DRB1, MAMU-DRB5), and lymphocytes (e.g. CD3D, CD3E, GZMB). The predicted gene
641 activity was the gene scores estimated by archR based on chromatin accessibility
642 information. The predicted gene activity scores and RNA expression were imputed for
643 visualization using MAGIC to smooth signals across nearby cells. Another way to visualize
644 chromatin accessibility is through cluster-based genome browser tracks using the
645 plotBrowserTrack function. In the visualization of chromatin accessibility for the SIV gene,

646 the parameters for plotting were set as upstream = 500, downstream = 10000, and for other
647 genes the parameters were set as default. After the initial screening, we found some
648 lymphocyte contaminations, which were removed from our analyses. The marker genes
649 were subsequently identified in only myeloid cell populations using normalized RNA
650 expression and predicted gene activity (**Table S1**). The Wilcoxon statistical test was
651 implemented by the getMarkerFeatures function in archR, and the marker genes were then
652 identified by setting the cut-off as FDR ≤ 0.01 and \log_2 fold change ≥ 1.25 . The marker genes
653 were plotted in a heatmap using the markerHeatmap function in arhcR for visualization.

654 From the UMAP projection and initial characterization, we found two central myeloid
655 populations in the infected brains. Given their RNA expression and predicted gene activity
656 profiles, we annotated the two populations as CAM-like cells and microglia-like (Micro-like)
657 cells. To find their differences, we subset those cells and merged each population's
658 subclusters (i.e., CAM_like_1, CAM_like_2, and CAM_like_3) into one cluster to compare
659 their RNA expression and predicted gene activity. As mentioned, the statistical test used for
660 identifying markers was the Wilcoxon rank test, and the cut-off was set as FDR ≤ 0.01 and
661 \log_2 fold change ≥ 1.25 . To better understand those changes, we further enriched the CAM-
662 like and microglia-like cell markers into Gene Ontology (GO) pathway analyses (Biological
663 Process, BP was selected), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway
664 analyses, and Ingenuity Pathway Analyses (IPA). The GO and KEGG analyses were
665 implemented by the clusterProfiler R package (version 4.0.2).⁶⁶

666 **Cell-to-cell communication analyses**

667 The cell-cell communication network's inference, analyses, and visualization were
668 implemented in the CellChat v2 R package (version 2.1.0).^{67,68} To obtain the input file for
669 CellChat analyses, we preprocessed the filtered feature barcode matrix from each sample in
670 the Seurat R package (version 4.4.0).⁶⁹ We first created and merged the Seurat objects of
671 different samples, and then we removed the cells not included in the ArchR object by
672 matching the cell barcode. After this filtering, all the cells in the ArchR object were also

673 included in the Seurat Object. In addition, the annotation for each cell was also transferred
674 from the ArchR object to the Seurat object. Then, we normalized the RNA counts using the
675 log normalization method in Seurat, and the log-normalized data was further used to create
676 a CellChat object. The ligand-receptor interaction database provided by CellChat only
677 includes the mouse and human, so we selected the human database, which is more relative
678 to rhesus monkeys for prediction. After subsetting signaling genes and identifying over-
679 expressed ligand-receptor interactions among cell clusters, the computeCommunProb
680 function in the CellChat computed the communication probability. We selected the “trimean”
681 method to compute the communication probability, which produced fewer but stronger
682 interactions. We obtained the communication probability on the signaling pathway level by
683 summarizing the communication probabilities of all ligand-receptor interactions associated
684 with each signaling pathway, and this was achieved by the computeCommunProbPathway
685 function. Finally, we used the aggregateNet function to calculate the aggregated cell-cell
686 communication network. To visualize the involvement of different cell clusters in the CCL
687 signaling pathway, the chord diagram was used, and the cells were grouped into four
688 categories, including microglia, CAM, microglia-like, and CAM-like. We also used the chord
689 diagram to visualize the interactions of myeloid cells in uninfected and SIV-infected brains.
690 We grouped the cells into CAM and microglia for uninfected brains and CAM-like and
691 microglia-like for SIV-infected brains. Then, we plotted all the interactions sent from either of
692 the two groups to another group for uninfected brains and SIV-infected brains separately.
693 Then, we identified the dominant senders and receivers in the intercellular communication
694 network using the netAnalysis_comuteCentrality function, and the total outgoing or incoming
695 communication probability associated with each cell cluster was shown in a 2D scatter plot.

696 **Peak Calling and transcription factor (TF) motif enrichment**

697 Single-cell chromatin accessibility data were used to generate the pseudobulk
698 replicates by the ArchR function, addGroupCovages, for peak calling with MACS2.⁷⁰ The
699 implementation of MACS2 in the ArchR was through the addReproduciblePeakSet function,

700 in which we set genome size as 2.2e9, which, as suggested by MACS2 documentation, is
701 78% of the total genome size of rhesus macaques.⁷¹ The pseudobulk replicates and peak
702 calling were based on the ten myeloid cell clusters identified in this study. Marker peaks for
703 each myeloid cell cluster were further found using the getMarkerFeatures function on the
704 peak matrix embedded in the ArchR object. The statistical test used for finding marker peaks
705 was the Wilcoxon rank test, and the cut-off was set as FDR ≤ 0.01 and log₂ fold change ≥ 1.
706 Then, we determined the enrichment of transcription factor binding sites for the marker
707 peaks by the peakAnnoEnrichment function. Before the enrichment, we used two human TF
708 motif databases, including the JASPAR 2020 database and the Vierstra non-redundant TF
709 (version: 2.1) database,³⁰ to annotate the peaks. We extracted the binarized motif
710 enrichment information for 11 SIV peaks using the getMatches function in ArchR to
711 understand the host TF enrichment for SIV DNA fragments. We used the chromVAR R
712 package⁷² embedded in ArchR to calculate the TF enrichment per-cell basis. A background
713 peak set controlling for total accessibility, and GC-content was generated by the
714 addBgdPeaks function before ChromVAR was run with the addDeviationsMatix function,
715 using JASPAR or Vierstra motif set to calculate enrichment of chromatin accessibility at
716 different TF motif sequences in a single cell. The motif deviation z-scores were used to find
717 marker motifs in each myeloid cell cluster and for visualization after being imputed with
718 MAGIC. The statistical test used for finding marker motifs was the Wilcoxon rank test, and
719 the cut-off was set as FDR ≤ 0.001 and MeanDiff ≥ 1. Transcription factor footprinting was
720 performed and visualized using the ArchR functions, which included getFootprints and
721 plotFootprints (normMethod = “subtract,” smoothWindow =6).

722 **SCENIC analysis**

723 SCENIC (version 1.1.2-01) (Single-Cell Regulatory Network Inference)^{31,32} was to examine
724 the transcription factors controlling the gene regulatory networks in the identified cell
725 clusters. This method examines the coexpression of genes with transcription factors followed
726 by *cis*-regulatory motif analysis to identify significant motif enrichment of the correct

727 upstream regulator, termed “regulons.” Gene expression RNA counts for each cluster were
728 used as input, and analysis was conducted on enriched regulons. SCENIC uses a random
729 forest⁷³ to calculate the importance of each regulon found in the expression matrix. After
730 being ranked, the top 50 regulons annotated by RcisTarget, were selected and activity
731 scores for regulons were calculated using AUCCell. Regulons with less than 20 genes and low
732 confidence (termed “extended”) regulons were filtered out.

733 **Integrated chromatin accessibility and RNA expression**

734 As mentioned, the DNA fragments and RNA transcripts from multiomic sequencing were
735 initially integrated. We performed two integrative analyses involving the features' correlation
736 to predict the gene regulatory interactions. Peak-to-gene links (PGLs) were calculated by
737 correlations between peak accessibility and snRNA expression data using the
738 addPeak2GeneLinks function. The dimensionality that combined ATAC and RNA information
739 after batch effect correction was used, and other parameters were set as default for PGLs.
740 All PGLs were plotted in a peak-to-gene heatmap using the plotPeak2GeneHeatmap
741 function for visualization. We also identified positive TF regulators by correlating the gene
742 expression or predicted gene activity of TFs and the accessibility of their corresponding
743 motifs. This analysis started by identifying deviant TF motifs using the getGroupSE function,
744 then correlating TF motifs and TF gene expression or predicted gene activity using the
745 correlateMatrices function. The positive TF regulators were considered as TFs whose
746 correlation between motif and gene expression (or predicted gene activity) is more than 0.5
747 with an adjusted p-value less than 0.01 and a maximum inter-cluster difference in deviation
748 Z-score that was in the top quartile. The TF motif database used for these correlations was
749 JASPAR. The ggplot2 R package achieved the data visualization for positive TF regulators,
750 and the ggvenn R package generated the Venn diagram.

751 **Pseudotime trajectory analysis**

752 To understand the dynamic changes of brain myeloid cells' differentiation in the uninfected
753 and infected brains, we performed pseudo-time analyses using the Monocle⁷⁴

754 implementation of ArchR. The harmony-corrected ArchR object of the core analysis was
755 subjected to the getMonocleTrajectories and addMonocleTrajectory functions for the
756 unsupervised trajectory analysis. In this analysis, all the myeloid cell clusters identified were
757 used, and the root node was set as the Micro_1 cluster or CAM_1 cluster for prediction.
758 Then, we performed the analyses for the four clusters mainly found in infected brains using
759 the addTrajectory function, which followed a user-defined trajectory to guide the supervised
760 trajectory analysis from the “Micro_like” to the “CAM_like_2” cluster. To visualize the
761 changes in RNA expression, predicted gene activity, peak accessibility, and ChromVAR
762 deviation scores across pseudo-time, we used the getTrajectory to retrieve the
763 corresponding matrix and used the plotTrajectoryHeatmap functions to plot. Then, we
764 performed integrative pseudo-time analyses to identify positive TF regulators driving the
765 differentiation. This was achieved by integrating predicted gene activity or gene expression
766 with motif accessibility across pseudo-time using the correlateTrajectories function. The TF
767 motif database used for those correlations was JASPAR.

768 **Analyses for SIV infected cells**

769 To better understand the infected myeloid cell population, we subsetted the SIV+ and SIV-
770 cells. The SIV+ cells were defined as the cells with at least one SIV RNA transcript and one
771 DNA fragment found, and the SIV- cells were defined as the cells with neither SIV RNA
772 transcript nor DNA fragment found from the two SIVE animals, this resulted in 1113 SIV+
773 cells and 3920 SIV- cells. Then, we redid the dimensionality reduction without the SIV
774 genome (set the excludeChr in the addIterativeLSI function to the SIV genome) for RNA
775 expression and DNA fragments separately. We also redid the batch effect correction, UMAP,
776 and clustering with the combined dimensionalities for the subset dataset. After getting the
777 new cluster (6 clusters) information for each cell, we compared it with the previous
778 annotations using the pheatmap R package. To find the upregulated and downregulated
779 genes in more susceptible clusters, we set the useGroups in the getMarkerFeatures function
780 as C1 and C6 clusters and set the bgdGroups as C2, C3, C4, and C5 clusters. On the other

781 hand, we set the useGroups as C3, C4, and C5 clusters and bgdGroups as C1, C2, and C6
782 clusters to find the differentially expressed genes in less susceptible clusters. The statistical
783 test method and cut-off were the same as mentioned. The detected upregulated and
784 downregulated genes for more susceptible and less susceptible clusters were further used
785 as input for IPA.

786

787 **ACKNOWLEDGEMENTS**

788 This work was supported by National Institutes of Health (NIH) grants U01DA053624,
789 R21MH128057, and P30MH062261. We want to acknowledge UMNC Genomics Core
790 Facility and Flow Cytometry Research Facility for the excellent technical support of the
791 Multiome procedure and FACS used in this study, and the University of Nebraska Lincoln
792 (UNL) Holland Computing Center for high-performance supercomputer access. This
793 publication's contents are the authors' sole responsibility and do not necessarily represent
794 the official views or policies of the NIH or Department of Health and Human Services, nor
795 does mention of trade names, commercial products, or organizations imply endorsement by
796 the U.S. Government.

797

798 **COMPETING INTERESTS**

799 The authors declare no competing interests.

800

801 **REFERENCES**

- 802 1 *HIV data and statistics*, <<https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics>> (July 17, 2023).
- 803 2 Data tables : Quality of life and HIV stigma—indicators for the National HIV/AIDS
804 Strategy, 2022–2025, CDC Medical Monitoring Project, 2017–2020 cycles. (2022).
- 805 3 Vasudev Malik Daliparty, R. B. *HIV Encephalitis*,
806 <<https://www.ncbi.nlm.nih.gov/books/NBK555894/>> (September 20, 2022).
- 807 4 Antinori, A. et al. Updated research nosology for HIV-associated neurocognitive
808 disorders. *Neurology* **69**, 1789–1799, doi:10.1212/01.WNL.0000287431.88658.8b
809 (2007).
- 810

- 811 5 Heaton, R. K. *et al.* HIV-associated neurocognitive disorders before and during the
812 era of combination antiretroviral therapy: differences in rates, nature, and predictors.
813 *Journal of NeuroVirology* **17**, 3-16, doi:10.1007/s13365-010-0006-1 (2011).
- 814 6 Heaton, R. K. *et al.* Neurocognitive Change in the Era of HIV Combination
815 Antiretroviral Therapy: The Longitudinal CHARTER Study. *Clinical Infectious
816 Diseases* **60**, 473-480, doi:10.1093/cid/ciu862 (2015).
- 817 7 Hatzlioannou, T. & Evans, D. T. Animal models for HIV/AIDS research. *Nature
818 Reviews Microbiology* **10**, 852-867, doi:10.1038/nrmicro2911 (2012).
- 819 8 Fox, H. S., Gold, L. H., Henriksen, S. J. & Bloom, F. E. Simian Immunodeficiency
820 Virus: A Model for NeuroAIDS. *Neurobiology of Disease* **4**, 265-274,
821 doi:<https://doi.org/10.1006/nbdi.1997.0159> (1997).
- 822 9 Grant, I. *et al.* Asymptomatic HIV-associated neurocognitive impairment increases
823 risk for symptomatic decline. *Neurology* **82**, 2055-2062,
824 doi:10.1212/WNL.0000000000000492 (2014).
- 825 10 Gartner, S. HIV Infection and Dementia. *Science* **287**, 602-604,
826 doi:10.1126/science.287.5453.602 (2000).
- 827 11 Zink, W. E., Zheng, J., Persidsky, Y., Poluektova, L. & Gendelman, H. E. The
828 neuropathogenesis of HIV-1 infection. *FEMS Immunology & Medical Microbiology* **26**,
829 233-241, doi:10.1111/j.1574-695X.1999.tb01394.x (1999).
- 830 12 Hou, J. & Major, E. O. Direct and indirect mechanisms of HIV-1 neuropathogenesis in
831 the human central nervous system. *Adv Exp Med Biol* **493**, 29-34, doi:10.1007/0-306-
832 47611-8_3 (2001).
- 833 13 Kaul, M., Garden, G. A. & Lipton, S. A. Pathways to neuronal injury and apoptosis in
834 HIV-associated dementia. *Nature* **410**, 988-994, doi:10.1038/35073667 (2001).
- 835 14 Cinque, P. *et al.* Elevated cerebrospinal fluid levels of monocyte chemotactic protein-
836 1 correlate with HIV-1 encephalitis and local viral replication. *AIDS* **12** (1998).
- 837 15 Conant, K. *et al.* Induction of monocyte chemoattractant protein-1 in HIV-1 Tat-
838 stimulated astrocytes and elevation in AIDS dementia. *Proceedings of the National
839 Academy of Sciences* **95**, 3117-3121, doi:10.1073/pnas.95.6.3117 (1998).
- 840 16 Weiss, J. M., Nath, A., Major, E. O. & Berman, J. W. HIV-1 Tat induces monocyte
841 chemoattractant protein-1-mediated monocyte transmigration across a model of the
842 human blood-brain barrier and up-regulates CCR5 expression on human monocytes.
843 *J Immunol* **163**, 2953-2959 (1999).
- 844 17 Gonzalez, E. *et al.* HIV-1 infection and AIDS dementia are influenced by a mutant
845 MCP-1 allele linked to increased monocyte infiltration of tissues and MCP-1 levels.
846 *Proceedings of the National Academy of Sciences* **99**, 13795-13800,
847 doi:10.1073/pnas.202357499 (2002).
- 848 18 Lipton, S. A. Requirement for macrophages in neuronal injury induced by HIV
849 envelope protein gp120. *NeuroReport* **3** (1992).
- 850 19 Giulian, D., Wendt, E., Vaca, K. & Noonan, C. A. The envelope glycoprotein of human
851 immunodeficiency virus type 1 stimulates release of neurotoxins from monocytes.
852 *Proceedings of the National Academy of Sciences* **90**, 2769-2773,
853 doi:10.1073/pnas.90.7.2769 (1993).
- 854 20 Kaul, M. & Lipton, S. A. Chemokines and activated macrophages in HIV gp120-
855 induced neuronal apoptosis. *Proceedings of the National Academy of Sciences* **96**,
856 8212-8216, doi:10.1073/pnas.96.14.8212 (1999).
- 857 21 Kramer-Hämmerle, S., Rothenaigner, I., Wolff, H., Bell, J. E. & Brack-Werner, R.
858 Cells of the central nervous system as targets and reservoirs of the human
859 immunodeficiency virus. *Virus Research* **111**, 194-213,
860 doi:<https://doi.org/10.1016/j.virusres.2005.04.009> (2005).

- 861 22 Niu, M. *et al.* Methamphetamine Increases the Proportion of SIV-Infected
862 Microglia/Macrophages, Alters Metabolic Pathways, and Elevates Cell Death
863 Pathways: A Single-Cell Analysis. *Viruses* **12** (2020).
- 864 23 Trease, A. J. *et al.* Antiretroviral therapy restores the homeostatic state of microglia in
865 SIV-infected rhesus macaques. *J Leukoc Biol* **112**, 969-981, doi:10.1002/jlb.3hi0422-
866 635r (2022).
- 867 24 Xu, X. *et al.* Microglia and macrophages alterations in the CNS during acute SIV
868 infection: a single-cell analysis in rhesus macaques. *bioRxiv*,
869 2024.2004.2004.588047, doi:10.1101/2024.04.04.588047 (2024).
- 870 25 Li, Q. & Barres, B. A. Microglia and macrophages in brain homeostasis and disease.
871 *Nat Rev Immunol* **18**, 225-242, doi:10.1038/nri.2017.125 (2018).
- 872 26 Geissmann, F., Jung, S. & Littman, D. R. Blood monocytes consist of two principal
873 subsets with distinct migratory properties. *Immunity* **19**, 71-82, doi:10.1016/s1074-
874 7613(03)00174-2 (2003).
- 875 27 Auffray, C. *et al.* Monitoring of blood vessels and tissues by a population of
876 monocytes with patrolling behavior. *Science* **317**, 666-670,
877 doi:10.1126/science.1142883 (2007).
- 878 28 Wang, S. *et al.* S100A8/A9 in Inflammation. *Front Immunol* **9**, 1298,
879 doi:10.3389/fimmu.2018.01298 (2018).
- 880 29 Ehrchen, J. M., Sunderkötter, C., Foell, D., Vogl, T. & Roth, J. The endogenous Toll-
881 like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection,
882 autoimmunity, and cancer. *J Leukoc Biol* **86**, 557-566, doi:10.1189/jlb.1008647
883 (2009).
- 884 30 Vierstra, J. *et al.* Global reference mapping of human transcription factor footprints.
885 *Nature* **583**, 729-736, doi:10.1038/s41586-020-2528-x (2020).
- 886 31 Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering.
887 *Nature Methods* **14**, 1083-1086, doi:10.1038/nmeth.4463 (2017).
- 888 32 Van de Sande, B. *et al.* A scalable SCENIC workflow for single-cell gene regulatory
889 network analysis. *Nature Protocols* **15**, 2247-2276, doi:10.1038/s41596-020-0336-2
890 (2020).
- 891 33 Kiani, K., Sanford, E. M., Goyal, Y. & Raj, A. Changes in chromatin accessibility are
892 not concordant with transcriptional changes for single-factor perturbations. *Mol Syst
893 Biol* **18**, e10979, doi:10.15252/msb.202210979 (2022).
- 894 34 Plaza-Jennings, A. L. *et al.* HIV integration in the human brain is linked to microglial
895 activation and 3D genome remodeling. *Mol Cell* **82**, 4647-4663.e4648,
896 doi:10.1016/j.molcel.2022.11.016 (2022).
- 897 35 Salmon, P. *et al.* Loss of CD4 membrane expression and CD4 mRNA during acute
898 human immunodeficiency virus replication. *J Exp Med* **168**, 1953-1969,
899 doi:10.1084/jem.168.6.1953 (1988).
- 900 36 Vallat, A. V. *et al.* Localization of HIV-1 co-receptors CCR5 and CXCR4 in the brain of
901 children with AIDS. *Am J Pathol* **152**, 167-178 (1998).
- 902 37 Chan, J. K. & Greene, W. C. NF-κB/Rel: agonist and antagonist roles in HIV-1
903 latency. *Curr Opin HIV AIDS* **6**, 12-18, doi:10.1097/COH.0b013e32834124fd (2011).
- 904 38 Heusinger, E. & Kirchhoff, F. Primate Lentiviruses Modulate NF-κB Activity by
905 Multiple Mechanisms to Fine-Tune Viral and Cellular Gene Expression. *Front
906 Microbiol* **8**, 198, doi:10.3389/fmicb.2017.00198 (2017).
- 907 39 Nabel, G. & Baltimore, D. An inducible transcription factor activates expression of
908 human immunodeficiency virus in T cells. *Nature* **326**, 711-713,
909 doi:10.1038/326711a0 (1987).

- 910 40 He, C. et al. Single-Cell Transcriptomic Analysis Revealed a Critical Role of
911 SPP1/CD44-Mediated Crosstalk Between Macrophages and Cancer Cells in Glioma.
912 *Front Cell Dev Biol* **9**, 779319, doi:10.3389/fcell.2021.779319 (2021).
- 913 41 Rao, G. et al. Reciprocal interactions between tumor-associated macrophages and
914 CD44-positive cancer cells via osteopontin/CD44 promote tumorigenicity in colorectal
915 cancer. *Clin Cancer Res* **19**, 785-797, doi:10.1158/1078-0432.Ccr-12-2788 (2013).
- 916 42 Suresh, P. & Wanchu, A. Chemokines and chemokine receptors in HIV infection: role
917 in pathogenesis and therapeutics. *J Postgrad Med* **52**, 210-217 (2006).
- 918 43 Wang, Z., Shang, H. & Jiang, Y. Chemokines and Chemokine Receptors:
919 Accomplices for Human Immunodeficiency Virus Infection and Latency. *Front*
920 *Immunol* **8**, 1274, doi:10.3389/fimmu.2017.01274 (2017).
- 921 44 Hudspeth, K. et al. Engagement of NKp30 on V δ 1 T cells induces the production of
922 CCL3, CCL4, and CCL5 and suppresses HIV-1 replication. *Blood* **119**, 4013-4016,
923 doi:10.1182/blood-2011-11-390153 (2012).
- 924 45 Lambert, J. S. et al. Production of the HIV-Suppressive Chemokines CCL3/MIP-1 α
925 and CCL22/MDC Is Associated with More Effective Antiretroviral Therapy in HIV-
926 Infected Children. *The Pediatric Infectious Disease Journal* **26** (2007).
- 927 46 Oguariri, R. M. et al. Short Communication: S100A8 and S100A9, Biomarkers of
928 SARS-CoV-2 Infection and Other Diseases, Suppress HIV Replication in Primary
929 Macrophages. *AIDS Res Hum Retroviruses* **38**, 401-405, doi:10.1089/aid.2021.0193
930 (2022).
- 931 47 Mogensen, T. H. IRF and STAT Transcription Factors - From Basic Biology to Roles
932 in Infection, Protective Immunity, and Primary Immunodeficiencies. *Front Immunol* **9**,
933 3047, doi:10.3389/fimmu.2018.03047 (2018).
- 934 48 Lukhele, S., Boukhaled, G. M. & Brooks, D. G. Type I interferon signaling, regulation
935 and gene stimulation in chronic virus infection. *Semin Immunol* **43**, 101277,
936 doi:10.1016/j.smim.2019.05.001 (2019).
- 937 49 Plataniolas, L. C. Mechanisms of type-I- and type-II-interferon-mediated signalling.
938 *Nature Reviews Immunology* **5**, 375-386, doi:10.1038/nri1604 (2005).
- 939 50 Billiau, A. Anti-inflammatory properties of Type I interferons. *Antiviral Res* **71**, 108-
940 116, doi:10.1016/j.antiviral.2006.03.006 (2006).
- 941 51 Benveniste, E. N. & Qin, H. Type I Interferons as Anti-Inflammatory Mediators.
942 *Science's STKE* **2007**, pe70-pe70, doi:doi:10.1126/stke.4162007pe70 (2007).
- 943 52 Axtell, R. C. & Steinman, L. Type 1 interferons cool the inflamed brain. *Immunity* **28**,
944 600-602, doi:10.1016/j.jimmuni.2008.04.006 (2008).
- 945 53 Kaminska, B., Mota, M. & Pizzi, M. Signal transduction and epigenetic mechanisms
946 in the control of microglia activation during neuroinflammation. *Biochim Biophys Acta*
947 **1862**, 339-351, doi:10.1016/j.bbadiis.2015.10.026 (2016).
- 948 54 Lanzillotta, A. et al. NF- κ B in Innate Neuroprotection and Age-Related
949 Neurodegenerative Diseases. *Front Neurol* **6**, 98, doi:10.3389/fneur.2015.00098
950 (2015).
- 951 55 Lundquist, C. A., Tobiume, M., Zhou, J., Unutmaz, D. & Aiken, C. Nef-mediated
952 downregulation of CD4 enhances human immunodeficiency virus type 1 replication in
953 primary T lymphocytes. *J Virol* **76**, 4625-4633, doi:10.1128/jvi.76.9.4625-4633.2002
954 (2002).
- 955 56 Ross, T. M., Oran, A. E. & Cullen, B. R. Inhibition of HIV-1 progeny virion release by
956 cell-surface CD4 is relieved by expression of the viral Nef protein. *Curr Biol* **9**, 613-
957 621, doi:10.1016/s0960-9822(99)80283-8 (1999).

- 958 57 Cortés, M. J., Wong-Staal, F. & Lama, J. Cell surface CD4 interferes with the
959 infectivity of HIV-1 particles released from T cells. *J Biol Chem* **277**, 1770-1779,
960 doi:10.1074/jbc.M109807200 (2002).
- 961 58 Wang, J., Roderiquez, G., Oravecz, T. & Norcross, M. A. Cytokine regulation of
962 human immunodeficiency virus type 1 entry and replication in human
963 monocytes/macrophages through modulation of CCR5 expression. *J Virol* **72**, 7642-
964 7647, doi:10.1128/jvi.72.9.7642-7647.1998 (1998).
- 965 59 Louboutin, J. P. & Strayer, D. S. Relationship between the chemokine receptor CCR5
966 and microglia in neurological disorders: consequences of targeting CCR5 on
967 neuroinflammation, neuronal death and regeneration in a model of epilepsy. *CNS*
968 *Neurol Disord Drug Targets* **12**, 815-829, doi:10.2174/18715273113126660173
969 (2013).
- 970 60 Varmus, H. Retroviruses. *Science* **240**, 1427-1435, doi:10.1126/science.3287617
971 (1988).
- 972 61 Schiralli Lester, G. M. & Henderson, A. J. Mechanisms of HIV Transcriptional
973 Regulation and Their Contribution to Latency. *Mol Biol Int* **2012**, 614120,
974 doi:10.1155/2012/614120 (2012).
- 975 62 Rohr, O., Marban, C., Aunis, D. & Schaeffer, E. Regulation of HIV-1 gene
976 transcription: from lymphocytes to microglial cells. *J Leukoc Biol* **74**, 736-749,
977 doi:10.1189/jlb.0403180 (2003).
- 978 63 Pereira, L. A., Bentley, K., Peeters, A., Churchill, M. J. & Deacon, N. J. A compilation
979 of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids*
980 *Res* **28**, 663-668, doi:10.1093/nar/28.3.663 (2000).
- 981 64 Wu, Y. & Marsh, J. W. Gene transcription in HIV infection. *Microbes Infect* **5**, 1023-
982 1027, doi:10.1016/s1286-4579(03)00187-4 (2003).
- 983 65 Morsey, B. *et al.* Cryopreservation of microglia enables single-cell RNA sequencing
984 with minimal effects on disease-related gene expression patterns. *iScience* **24**,
985 102357, doi:10.1016/j.isci.2021.102357 (2021).
- 986 66 Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics
987 data. *The Innovation* **2**, 100141, doi:<https://doi.org/10.1016/j.xinn.2021.100141>
988 (2021).
- 989 67 Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell-cell
990 communication from single-cell and spatially resolved transcriptomics. *bioRxiv*,
991 2023.2011.2005.565674, doi:10.1101/2023.11.05.565674 (2023).
- 992 68 Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nature*
993 *Communications* **12**, 1088, doi:10.1038/s41467-021-21246-9 (2021).
- 994 69 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-
995 3587.e3529, doi:10.1016/j.cell.2021.04.048 (2021).
- 996 70 Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**,
997 R137, doi:10.1186/gb-2008-9-9-r137 (2008).
- 998 71 Gibbs, R. A. *et al.* Evolutionary and Biomedical Insights from the Rhesus Macaque
999 Genome. *Science* **316**, 222-234, doi:10.1126/science.1139247 (2007).
- 1000 72 Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring
1001 transcription-factor-associated accessibility from single-cell epigenomic data. *Nature*
1002 *Methods* **14**, 975-978, doi:10.1038/nmeth.4401 (2017).
- 1003 73 Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks
1004 from expression data using tree-based methods. *PLoS One* **5**,
1005 doi:10.1371/journal.pone.0012776 (2010).
- 1006 74 Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis.
1007 *Nature* **566**, 496-502, doi:10.1038/s41586-019-0969-x (2019).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [XuetalSuppNatureComm.pdf](#)