Title
Identification of astrocyte regulators by nucleic acid cytometry

Journal
Nature, 614(7947)

ISSN
0028-0836

Authors
Clark, Iain C
Wheeler, Michael A
Lee, Hong-Gyun
et al.

Publication Date
2023-02-09

DOI
10.1038/s41586-022-05613-0

Peer reviewed

HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2023 July 01.

Published in final edited form as:

Nature. 2023 February ; 614(7947): 326–333. doi:10.1038/s41586-022-05613-0.

Identification of astrocyte regulators by nucleic acid cytometry

Iain C. Clark[1,2,3,7], **Michael A. Wheeler**[1,4,7], **Hong-Gyun Lee**[1], **Zhaorong Li**[1,4], **Liliana M. Sanmarco**[1], **Shravan Thaploo**[1], **Carolina M. Polonio**[1], **Seung Won Shin**[3], **Giulia Scalisi**[1], **Amy R. Henry**[5], **Joseph M. Rone**[1], **Federico Giovannoni**[1], **Marc Charabati**[1], **Camilo Faust Akl**[1], **Dulce M. Aleman**[1], **Stephanie E. J. Zandee**[6], **Alexandre Prat**[6], **Daniel C. Douek**[5], **Eli A. Boritz**[5], **Francisco J. Quintana**[1,4,8,✉], **Adam R. Abate**[2,8,✉]

[1]Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[2]Department of Bioengineering and Therapeutic Sciences, School of Pharmacy, University of California San Francisco, San Francisco, CA, USA

[3]Department of Bioengineering, College of Engineering, California Institute for Quantitative Biosciences, QB3, University of California Berkeley, Berkeley, CA, USA

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

[5]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

[6]Neuroimmunology Research Lab, Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Quebec, Canada

[7]These authors contributed equally: Iain C. Clark and Michael A. Wheeler

[8]These authors jointly supervised this work: Francisco J. Quintana and Adam R. Abate

Abstract

Multiple sclerosis is a chronic inflammatory disease of the central nervous system[1]. Astrocytes are heterogeneous glial cells that are resident in the central nervous system and participate in the pathogenesis of multiple sclerosis and its model experimental autoimmune encephalomyelitis[2,3].

✉**Correspondence and requests for materials** should be addressed to Francisco J. Quintana, fquintana@rics.bwh.harvard.edu or Adam R. Abate, adam@abatelab.org.

Author contributions I.C.C., M.A.W., E.A.B., F.J.Q. and A.R.A. designed research. I.C.C., M.A.W., H.-G.L., L.M.S., C.M.P., S.T., S.W.S., G.S., A.R.H., M.C., C.F.A., D.M.A., J.M.R. and F.G. performed experiments. I.C.C., M.A.W., Z.L., F.J.Q. and A.R.A. analysed data. D.C.D., S.E.J.Z. and A.P. provided unique materials and/or discussed findings. I.C.C., M.A.W., F.J.Q. and A.R.A. wrote the paper with input from the co-authors. F.J.Q. and A.R.A. directed and supervised the study.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05613-0.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Competing interests The authors declare no competing interests.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05613-0.

However, few unique surface markers are available for the isolation of astrocyte subsets, preventing their analysis and the identification of candidate therapeutic targets; these limitations are further amplified by the rarity of pathogenic astrocytes. Here, to address these challenges, we developed focused interrogation of cells by nucleic acid detection and sequencing (FIND-seq), a high-throughput microfluidic cytometry method that combines encapsulation of cells in droplets, PCR-based detection of target nucleic acids and droplet sorting to enable in-depth transcriptomic analyses of cells of interest at single-cell resolution. We applied FIND-seq to study the regulation of astrocytes characterized by the splicing-driven activation of the transcription factor XBP1, which promotes disease pathology in multiple sclerosis and experimental autoimmune encephalomyelitis[4]. Using FIND-seq in combination with conditional-knockout mice, in vivo CRISPR–Cas9-driven genetic perturbation studies and bulk and single-cell RNA sequencing analyses of samples from mouse experimental autoimmune encephalomyelitis and humans with multiple sclerosis, we identified a new role for the nuclear receptor NR3C2 and its corepressor NCOR2 in limiting XBP1-driven pathogenic astrocyte responses. In summary, we used FIND-seq to identify a therapeutically targetable mechanism that limits XBP1-driven pathogenic astrocyte responses. FIND-seq enables the investigation of previously inaccessible cells, including rare cell subsets defined by unique gene expression signatures or other nucleic acid markers.

Rare cells exert substantial influence on tissue physiology and pathology across many domains of biology[5–7]. In the context of inflammation in the central nervous system (CNS), small cell subsets have important roles in tissue pathology[8–12]. For example, astrocytes are CNS-resident glial cells with multiple roles in development, homeostasis and neurological disorders[2,13,14]; small astrocyte subsets identified by single-cell sequencing have been shown to promote or limit CNS pathology in multiple sclerosis and experimental autoimmune encephalomyelitis[4,15–22] (EAE), a pre-clinical model of multiple sclerosis. However, most astrocyte subsets (and many other cells of interest identified using single-cell genomics) lack unique surface markers, limiting their isolation and further investigation. Moreover, single-cell methods must sequence prohibitively large numbers of cells to sufficiently profile rare target cells. In addition to being costly and wasteful, this results in sparse transcriptome coverage is sparse.

Despite an absence of tools to interrogate the properties of many astrocyte subsets, increasingly sophisticated databases of transcriptional markers are being compiled using single-cell approaches[22–24]. The sorting of astrocytes defined using these datasets is often not possible because their nucleic acid signatures do not map to unique surface markers that enable their isolation. This is especially true when isolation of nuclei is necessary to achieve adequate RNA sequencing (RNA-seq) data quality, as is the case for many cells in the CNS. Fluorescence in situ hybridization linked to fluorescence-activated cell sorting (FACS) enables the isolation of cells based on the direct detection of RNA markers but is insensitive to low copy targets, and the harsh fixatives used in this method degrade mRNA and interfere with the generation of high-quality sequencing data from the sorted cells[25]. Moreover, obtaining a detectable signal often requires the use of multiple tiled probes, limiting the detection of sensitive post-transcriptional events such as RNA splicing[25]. By contrast, droplet methods that use amplification-based nucleic acid detection achieve single-molecule sensitivity and can reliably distinguish transcript splicing isoforms, but mRNA

degradation during lysis and thermocycling limits sequencing data quality even when many cells are pooled[26,27]. An ideal method for studying rare cells defined by nucleic acid markers would combine the specificity and sensitivity of PCR detection with the ability to obtain high-quality post-sorting sequencing data.

In this study, we describe the development and application of FIND-seq, which enables the high-throughput and in-depth sequencing of rare cell subsets defined by RNA or DNA markers. FIND-seq uses poly-dT oligonucleotide-conjugated agarose to capture cell nucleic acids with high throughput. RNA is reverse transcribed on the agarose bead and subjected to sequence-specific PCR with fluorescent TaqMan probes, followed by droplet sorting of cells containing marker RNA and downstream bulk or single-cell analysis. We combined FIND-seq with cell-specific genetic perturbation studies in vivo and bulk and single-cell transcriptional and genomic analyses of EAE and multiple sclerosis samples and identified NR3C2–NCOR2 signalling as a negative regulator of an astrocyte subset driven by the post-transcriptional splicing-dependent activation of the transcription factor XBP1, which promotes CNS pathology in multiple sclerosis and EAE[4].

FIND-seq development

A central challenge in the study of astrocyte subsets is the need to isolate populations of interest for in-depth molecular characterization. For example, we recently identified a subset of astrocytes that promotes CNS pathology in EAE and potentially multiple sclerosis, characterized by the activation of the transcription factor XBP1[22]. Despite the prominent role of this relatively rare subset in promoting CNS pathology[22], $XBP1^+$ astrocytes lack known surface markers that can be used to isolate and investigate them. Compounding these challenges, XBP1 activation is catalysed by an mRNA splicing mechanism, which removes a premature stop codon from the *Xbp1* mRNA transcript[28]. Thus, XBP1 activation cannot be directly quantified by single-cell RNA-seq (scRNA-seq) because 3′ mRNA capture does not consistently detect *Xbp1* mRNA splicing. To analyse XBP1 activity, we assayed *Edem1*, a marker gene associated with XBP1 activation[4] in a recent analysis of 18,047 CNS cells during EAE by scRNA-seq; only 14 out of 1,240 astrocytes expressed *Edem1* (Fig. 1a and Extended Data Fig. 1a,b). These challenges highlight the need for a method to isolate cells of interest, such as $XBP1^+$ astrocytes, on the basis of the expression of mRNA markers.

Efficient nucleic acid detection, sorting and sequencing of cells requires the simultaneous optimization of target marker detection and mRNA stability in many input cells. FIND-seq addresses these challenges by using a multi-step microfluidic workflow that can take millions of cells as input (Fig. 1b,c and Extended Data Fig. 1c–e) and analyse rare cell populations with <1% abundance in a cost-effective manner compared with scRNA-seq (Extended Data Table 1). The first device encapsulates cells in molten agarose with lysis buffer; the air-triggered droplet generation runs at 10 kHz, encapsulating 10^6 cells at limiting dilution in about 20 min (Fig. 1b,c and Extended Data Fig. 2a,b). This speed maximizes cell health and reduces transcriptional changes in cells as they await microfluidic encapsulation. To simultaneously capture cellular genomes and transcriptomes, we developed a method for functionalizing primers directly onto agarose (Extended Data Fig. 2c). After lysis, the agarose is cooled, genomes are sterically trapped in the matrix, and cellular mRNA is

hybridized with the capture probes (Extended Data Fig. 2d,e). The solid gels are removed from oil and transferred to wash buffer to clear cellular debris and residual lysis buffer, thereby retaining only purified nucleic acids from lysed cells. Captured mRNA is reverse transcribed and genomic DNA remains sterically confined in the matrix (Fig. 1d and Extended Data Fig. 3a). Exploiting bulk processing of hydrogels, we performed technical controls to retain high-quality single-cell messenger RNA and optimize cDNA synthesis yield (Extended Data Fig. 3b–g). Thus, we developed a high-throughput workflow to capture, purify and amplify the transcriptomes of millions of cells in parallel.

We next developed a method to detect nucleic acids from single cells using multiplexed droplet TaqMan PCR assays that also preserved the integrity of the transcriptome (Fig. 1d,e). TaqMan enables the use of a diverse inventory of pre-validated assays while enabling user-driven customization when necessary. Moreover, porous hydrogels simplify bulk processing, since the requisite reagents can be introduced by buffer exchange. Specifically, we re-encapsulated 55-μm hydrogels with PCR reagents into oil using a microfluidic device so that, upon thermal cycling, positive signals are generated within 75 μm drops (Fig. 1d,e and Extended Data Fig. 4a,b). We optimized the detection assay to preserve cDNA during the PCR, as normal thermal cycling conditions reduced downstream whole-transcriptome amplification (WTA) yield and data quality (Extended Data Fig. 4c). The resulting droplet PCR also enabled robust detection of single-copy genomic DNA targets (Extended Data Fig. 4d and the accompanying Article[29]). To isolate cells that express the target RNA transcript, we sorted fluorescent drops using a concentric microfluidic device previously optimized for high-fidelity sorting[30] (Fig. 1f,g and Extended Data Fig. 4e,f). Sorted beads were collected and cDNA was amplified and prepared for RNA-seq analysis (Fig. 1f,g). In summary, we established a high-throughput method that sorts cells on the basis of RNA (or DNA) markers.

Analysis of XBP1^{+} astrocytes with FIND-seq

We recently identified an astrocyte subset that promotes CNS pathology in multiple sclerosis and EAE and is driven by the transcription factor XBP1 induced by the unfolded protein response[4]. A similar astrocyte population contributes to prion-induced CNS pathology[31]. These findings suggest that XBP1-driven astrocytes contribute to the pathology of multiple neurological disorders, but little is known about the mechanisms that regulate them. The *Xbp1* gene is first transcribed into an unspliced *Xbp1* mRNA (*Xbp1u*), which harbours a stop codon that results in the synthesis of a truncated, non-functional XBP1[32] (Fig. 2a). In response to endoplasmic reticulum stress, activation of the inositol requiring enzyme 1α (IRE1α) promotes *Xbp1* mRNA splicing, generating spliced *Xbp1* mRNA (*Xbp1s*) lacking the early stop codon, which thus drives the synthesis of the full-length functional transcription factor XBP1[32]. However, neither *Xbp1* mRNA splicing nor the induction of *Edem1*[33], one of its primary target genes, is readily detected by scRNA-seq in astrocytes (Fig. 1a and Extended Data Fig. 1b), presumably because of data sparsity and the 3′ bias of single-cell barcoding methods[34,35]. We therefore applied FIND-seq to study XBP1-driven astrocytes.

We developed a multiplexed digital droplet TaqMan assay that detected the expression of *Edem1* and the astrocyte marker *Aqp4*, which we validated could accurately detect *Aqp4*$^+$ astrocytes with similar sensitivity and specificity (6.6% of cells) to those of flow cytometry (4.7% of cells) and scRNA-seq (6.8% of cells) (Extended Data Fig. 5a). Moreover, we sorted AQP4$^+$ and AQP4$^-$ cells using flow cytometry and analysed each population through the FIND-seq workflow (but before transcriptome sequencing), which showed 80% sensitivity to detect *Aqp4* in the AQP4$^+$ population with a 4.1% detection rate of *Aqp4* in the AQP4$^-$ population (Extended Data Fig. 5b). We then used FIND-seq to study *Aqp4*$^+$*Edem1*$^+$ astrocytes in EAE (Fig. 2b,c and Supplementary Table 1). In agreement with immunohistochemistry analyses[4], we detected an increase of approximately tenfold in the number of XBP1-driven astrocytes in the CNS 18 days after EAE induction by immunization with the myelin oligodendrocyte glycoprotein peptide $MOG_{35–55}$ (Fig. 2d,e). Bulk FIND-seq (cells sorted into 100-cell bins) produced RNA-seq data (300 cells, 26,171 genes) similar in complexity to that of to Smart-seq2 (ref. [36]) (300 cells, 25,076 genes) and showed improved depth and gene diversity compared with techniques such as Probe-seq[25] (22,000 cells, 15,118 genes) and scRNA-seq[37] (14 cells, 8,190 genes) that enable high-throughput identification of cells based on RNA markers (Extended Data Fig. 5c–e). In addition, RNA-seq analyses of *Aqp4*$^+$ astrocytes from EAE mice detected increased expression of pro-inflammatory pathways linked to the promotion of CNS pathology (Fig. 2f). *Aqp4*$^+$*Edem1*$^-$ astrocytes also displayed increased expression of pro-inflammatory transcriptional programmes during EAE (Fig. 2g), consistent with the existence of *Aqp4*$^+$*Edem1*$^-$ astrocyte subsets that also contribute to CNS pathology[16,18,22,38].

FIND-seq enabled the transcriptional analysis of the *Aqp4*$^+$*Edem1*$^+$ astrocyte subset, detecting increased expression of pro-inflammatory signalling pathways during EAE (Fig. 2h). To identify regulatory mechanisms associated with *Aqp4*$^+$*Edem1*$^+$ astrocytes, we tested for signatures of molecules that putatively control *Aqp4*$^+$*Edem1*$^+$ versus *Aqp4*$^+$*Edem1*$^-$ astrocytes using Ingenuity Pathway Analysis (IPA). Notably, we identified *Xbp1* as a regulator of the transcriptional signature we detected in *Aqp4*$^+$*Edem1*$^+$ astrocytes and *Edem1* as a downstream gene in this regulon, suggesting that FIND-seq enabled the analysis of XBP1-driven astrocytes (Fig. 2i). Collectively, these findings underscore the ability of FIND-seq to analyse in depth an astrocyte subset of interest with a role in CNS pathology based on the expression of RNA markers.

NR3C2–NCOR2 signalling controls XBP1$^+$ astrocytes

To investigate the transcriptional heterogeneity within *Aqp4*$^+$*Edem1*$^+$ astrocytes, we developed single-cell FIND-seq (scFIND-seq). We optimized single-bead sorting into microwells of a 96-well plate (Extended Data Fig. 6a) and performed a species-mixing experiment to validate single-drop sorting fidelity (Extended Data Fig. 6b). To analyse RNA expression in naive or EAE *Aqp4*$^+$*Edem1*$^{+/-}$ astrocytes by scFIND-seq, we extended a unique molecular identifier (UMI)-corrected Smart-seq3 approach to individual cells captured in agarose beads (Fig. 3a, Extended Data Fig. 6c–e and Supplementary Table 2). Consistent with our previous analyses validating the quality of bulk FIND-seq data relative to similar technologies, scFIND-seq detected 16,874 unique genes on average compared with Smart-seq2, which detected 25,076 (Extended Data Fig. 5d). scFIND-seq data captured

284 cells representing *Aqp4*$^{+}$*Edem1*$^{-}$ and *Aqp4*$^{+}$*Edem1*$^{+}$ astrocytes isolated from naive and EAE mice; *Aqp4*$^{+}$*Edem1*$^{+}$ astrocytes were predominantly located in a distinct cluster of cells with pro-inflammatory markers that emerged during EAE (Fig. 3b–d). Our bulk FIND-seq datasets identified nine transcriptional regulators predicted to control *Aqp4*$^{+}$*Edem1*$^{+}$ astrocytes in EAE when compared with *Aqp4*$^{+}$*Edem1*$^{-}$ astrocytes (Extended Data Fig. 7a,b), seven of which were also identified using scFIND-seq (Extended Data Figs. 7c–e). Indeed, we identified two subclusters of *Aqp4*$^{+}$*Edem1*$^{+}$ astrocytes (Extended Data Fig. 7c–e) on the basis of single-cell transcriptional analysis, highlighting both the potential heterogeneity within this astrocyte subset and the power of scFIND-seq for the analysis of rare cell populations from complex tissues.

We then performed an in vivo genetic perturbation screen using lentiviral vectors co-expressing CRISP–Cas9 and a single guide RNA (sgRNA) targeting each of the candidate regulators[4,21,22,38] under the control of the *Gfap* promoter (Extended Data Fig. 7f). Control mice were treated with vectors encoding CRISPR–Cas9 and a non-targeting sgRNA (*sgScrmbl*). We delivered the lentiviruses via intracerebroventricular injection to mice and monitored EAE development following immunization with $MOG_{35–55}$ as described[4,18–22,38]. Consistent with our bioinformatic predictions showing downregulation of NR3C2 signalling detected by FIND-seq and scFIND-seq in *Aqp4*$^{+}$*Edem1*$^{+}$ astrocytes (Extended Data Fig. 7a–e), astrocyte-specific knockdown of the ligand-activated nuclear receptor *Nr3c2* resulted in the worsening of EAE (Fig. 3e), suggesting that NR3C2 limits disease-promoting astrocyte responses.

Nr3c2 encodes the mineralocorticoid receptor, and belongs to a family of nuclear receptors with known immunomodulatory effects[39]. *Nr3c2* has been associated with early onset hypertension, severe exacerbation of hypertension in pregnancy as well as other peripheral diseases[40], and neurological disorders such as autism[41]. However, the role of NR3C2 in the control of astrocyte subsets is unknown. Consistent with our in vivo findings, treatment with 1 μM of the NR3C2 antagonist finerenone increased gene expression of *Il6*, *Il1b* and *Ccl2* in primary mouse astrocytes activated in vitro with 1 ng ml^{-1} IL-1β and TNF (Fig. 3f), cytokines that trigger transcriptional responses similar to those detected in astrocytes during the course of EAE and multiple sclerosis, including XBP1 activation[22]. Finerenone treatment also boosted *Xbp1* splicing (Fig. 3f). Similarly, RNA-seq analyses of astrocytes from EAE mice transduced with *Nr3c2*-targeting CRISPR–Cas9 lentiviruses detected inhibition of the *Nr3c2* regulon compared with mice transduced with non-targeting controls (Extended Data Fig. 8a), concomitant with increased IκB phosphorylation (Fig. 3g and Supplementary Table 3) associated with NF-κB activation, which controls disease-promoting transcriptional modules in XBP1-driven astrocytes[4]. *Nr3c2* inactivation in astrocytes was also linked to increased XBP1 activation, as determined by the increased abundance of astrocytes exhibiting spliced XBP1 (Fig. 3h). *Nr3c2* inactivation in astrocytes did not affect the T cell response during EAE (Extended Data Fig. 8b,c).

Following its ligand-dependent activation, NR3C2 translocates to the nucleus, where it binds responsive elements in target genes to modulate their expression. However, the transcriptional regulation of target genes by NR3C2 requires additional coactivators and corepressors[39]. The analysis of the RNA-seq dataset of astrocytes following *Nr3c2*

inactivation identified the nuclear receptor corepressor 2 (NCOR2) as a potential contributor to the suppression of XBP1-driven astrocyte pro-inflammatory responses by NR3C2 (Fig. 3i), which was also predicted as a regulator of *Aqp4*$^{+}$*Edem1*$^{+}$ cells in cluster 1 using scFIND-seq (Fig. 3i). Indeed, NCOR2 inactivation in astrocytes using lentiviral vectors co-expressing CRISPR–Cas9 and a targeting sgRNA under the control of the *Gfap* promoter resulted in exacerbated EAE (Fig. 3j), concomitant with the increased expression of pro-inflammatory transcriptional modules driven by XBP1, but no changes in the T cell response during EAE (Fig. 3k,l, Extended Data Fig. 8d–f and Supplementary Table 4). Together, these data suggest that NR3C2–NCOR2 signalling limits XBP1-driven pro-inflammatory transcriptional responses in astrocytes.

XBP1 limits NR3C2–NCOR2 signalling in astrocytes

XBP1 is activated in grey and white matter astrocytes in the spinal cord and brain during EAE[4] (Fig. 2d,e). To investigate the cross-regulation between XBP1 and NR3C2–NCOR2 signalling, we bred mice harbouring a floxed *Xbp1* allele[42] with mice expressing tamoxifen-inducible Cre recombinase in astrocytes under the control of the *Aldh1l1* promoter[43], generating *Xbp1*^{Astro} mice in which *Xbp1* is knocked out in astrocytes following tamoxifen administration (Fig. 4a). *Xbp1* inactivation in astrocytes resulted in a significant amelioration of EAE (Fig. 4b), but did not affect the T cell response or the number of CNS-resident cells during EAE (Extended Data Fig. 9a–d). However, analysis of the transcriptional programme of *Xbp1*-deficient astrocytes by RNA-seq showed an increased activation of NR3C2-driven transcriptional modules (Fig. 4c and Supplementary Table 5) concomitant with the decreased activation of disease-promoting transcriptional programmes such as cytokine signalling and the unfolded protein response (Fig. 4c and Extended Data Fig. 9d,e). Consistent with these data, *Xbp1* inactivation resulted in increased chromatin accessibility of genes driven by NR3C2 mineralocorticoid signalling, as shown by assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) in bulk sorted astrocytes (Extended Data Fig. 10a), as well as increased chromatin accessibility of *Nr3c2* and *Ncor2* (Extended Data Fig. 10b). These data suggest that XBP1 activation limits anti-inflammatory NR3C2 signalling in astrocytes, and that XBP1- and NR3C2-driven signalling pathways engage in an antagonistic cross-talk that regulates astrocyte pathogenic activities.

To evaluate the clinical relevance of XBP1–NR3C2 antagonism in human astrocytes, we first analysed CNS scRNA-seq datasets from patients with multiple sclerosis and controls matched for sex, age and brain region[22], detecting the expression of *NR3C2* and *NCOR2* in astrocytes (Fig. 4d). We next analysed the activation of XBP1-driven transcriptional modules in *NR3C2*$^{+}$ or *NCOR2*$^{+}$ astrocytes, and found that the expression of *NR3C2* or *NCOR2* in astrocytes was associated with a decreased activation of the XBP1-driven unfolded protein response and the production of reactive oxygen species (Fig. 4e).

To further evaluate these findings, we used immunostaining to analyse CNS samples from patients with multiple sclerosis and controls. We detected decreased numbers of NR3C2^{+} astrocytes in white matter from patients with multiple sclerosis relative to healthy controls (Fig. 4f). Moreover, the analysis of NR3C2- and NCOR2-driven transcriptional responses

in scRNA-seq astrocyte datasets showed a reduction in NR3C2 and NCOR2 signalling in astrocytes from lesions of patients with multiple sclerosis compared with control donors (Fig. 4g). A similar reduction in NR3C2 and NCOR2 signalling was also detected in normal-appearing white matter from patients with multiple sclerosis (Fig. 4g). Together, these findings suggest that the balance between XBP1 and NR3C2–NCOR2 signalling controls disease-promoting astrocyte activities that contribute to CNS pathology in EAE and potentially in multiple sclerosis.

Discussion

Here we describe FIND-seq–a high-throughput method to isolate and sequence rare cells defined by nucleic acid markers–and its use to study an XBP1-driven astrocyte subset associated with CNS pathology in multiple sclerosis and EAE[4]. FIND-seq is a unique tool for the investigation of cell subsets identified in single-cell genomic studies because it can directly target the nucleic acid markers that define them. FIND-seq captures rare cell populations from millions of input cells, produces high-quality transcriptome data, and is compatible with scRNA-seq. FIND-seq eliminates the need for, and the associated high cost of, sequencing large numbers of background cells when focusing on important rare cell populations. Because FIND-seq can sort cells on the basis of RNA markers, including spliced transcripts, non-coding RNA or the presence of viral RNA, it has the potential to accelerate the study of cells for which transgenic animals or antibodies are unavailable. In addition, FIND-seq's flexible processing of millions of input cells in porous, biocompatible and thermally reversable hydrogels makes it compatible with additional multi-omic readouts in future iterations, including genomic, epigenomic or proteomic measurements. However, further optimization is required to combine FIND-seq detection of transcripts in subcellular compartments such as astrocyte endfeet.

New high-dimensional, genome-wide and single-cell technologies are rapidly uncovering astrocyte subsets associated with a broad array of physiologic functions and neurological disorders[22,23,44,45]. However, without techniques to isolate these subsets, it remains challenging to define the molecular mechanisms that control them and discover candidate targets for their therapeutic modulation[2,17]. For example, FIND-seq enabled the investigation of rare UPR-driven astrocytes associated with CNS pathology in multiple sclerosis[4] and prion disease[31], resulting in the identification of an NR3C2–NCOR2 regulatory pathway that antagonizes XBP1-driven stimulus-induced astrocytes[2] in the context of CNS inflammatory conditions. To our knowledge, NR3C2 expression and function in astrocytes has not been investigated previously. Our data complement previous reports on the regulatory roles of ligand-activated transcription factors in CNS-resident cells in the context of multiple sclerosis and other neurological diseases. For example, in response to its activation by tryptophan-derived agonists controlled by the commensal flora, the ligand-activated transcription factor aryl hydrocarbon receptor suppresses pro-inflammatory responses in astrocytes and microglia during EAE and potentially, multiple sclerosis[19,20]. Similarly, NR4A1 and NR4A2 limit the pro-inflammatory responses of astrocytes, microglia and CNS-recruited macrophages[46,47]. On the basis of the potential for selective modulation of NR3C2 activity by small molecules[48], our findings identify NR3C2 as a candidate target for the therapeutic control of pathogenic, disease-associated astrocytes[2,17,22,23].

NCOR2 facilitates the recruitment of histone deacetylases to target genes recognized by interacting nuclear receptors, introducing epigenetic chromatin modifications that silence gene expression. NR4A2 has been shown to limit NF-κB-driven pro-inflammatory responses in astrocytes and microglia in Parkinson's disease through epigenetic mechanisms mediated by the histone deacetylase complex CoREST[46]. Moreover, the silencing of antioxidant loci in $MAFG^{+}$ astrocytes contributes to CNS pathology in EAE and multiple sclerosis[22]. Collectively, these data highlight the potential role of epigenetic modifications in long-term changes in astrocytes that may amplify or perpetuate responses relevant for the development of CNS pathology[3], resembling recent reports of trained microglial responses in the context of Alzheimer's disease[49,50]. Thus, the pharmacological targeting of epigenetic programmes associated with NCOR2 may offer novel approaches for the therapeutic modulation of CNS inflammation and neurodegeneration.

In summary, we report the development of FIND-seq, a method that can isolate rare cell subsets defined by nucleic acid signatures. Using FIND-seq, we revealed a NR3C2–NCOR2 regulatory mechanism controlling an astrocyte subset that promotes CNS pathogenesis in EAE and, potentially, multiple sclerosis.

Methods

Mice

All experiments were reviewed and approved by the Brigham & Women's Hospital Institutional Animal Care and Use Committee. Adult mice aged 8–10 weeks and pups aged from postnatal day (P)0–P3 were used on a C57Bl/6J background (000664, The Jackson Laboratory). *Aldh1l1-cre/ERT2* mice[43] (The Jackson Laboratory, 029655) were bred to $Xbp1^{fl/fl}$ mice[42] (a gift from L. Glimcher) to generate *Aldh1l1-creERT2;*$Xbp1^{fl/fl}$ mice. Conditional deletion of *Xbp1* was induced at 6–10 weeks of age by gavage with 225 mg kg^{-1} tamoxifen (Sigma-Aldrich, T5648), diluted in corn oil (Sigma-Aldrich, C8267); EAE was induced 4 weeks later.

Analysis of previous data

Previously published data were accessed using the following accession codes: EMBL E-MTAB-9268[52], GSE135572[25], and GSE130119[22].

Fabrication of microfluidic devices

Masters were fabricated using soft lithography. Silicon wafers were spin coated with SU-8 2025 photoresist (MicroChem), UV-patterned using a mask aligner (SUSS MJB3), developed for 10 min, and baked at 65 °C overnight. Polydimethylsiloxane (PDMS) prepolymer and curing agent (Momentive; RTV 615) were combined at 10:1, mixed by hand, degassed, poured onto fabricated masters, degassed again, and baked at 65 °C overnight. Cured PDMS was removed from the master mould, punched with a 0.75-mm biopsy punch (Harris Uni-Core 0.75), and plasma bonded (Technics 500-II Plasma Etcher) to a glass slide (75 × 50 × 1.0 mm, Fisher Scientific, 12-550C). Aquapel was flushed through the microfluidic device and allowed to react for 5 min. Channels were purged with

air and flushed with Fluorinert FC-40 oil (Sigma, F9755), before being baked at 65 °C for 30 min.

FIND-seq

Generation of oligonucleotide-conjugated agarose.—A 5′-acrydited polyT capture primer was covalently attached to allyl agarose as follows. The primer (/5Acryd/ TTTTTTTAAGCAGTGGTATCAACGCAGAGTACT 30VN, Acrydite-T5-SMART-T30VN, IDT) was resuspended in 0.375 M Tris-HCl Buffer pH 8.3 (Teknova) to a concentration of 1,000 μM. SFR allyl agarose (Lucidant Polymers) was resuspended in 0.375 M Tris-HCl buffer (pH 8.3) to a final concentration of 0.5% and heated to 95 °C. Agarose was mixed by vortexing every 15 min for 2 h. Agarose was cooled to 45 °C and allowed to sit under vacuum for 30 min. Acrydite-T5-SMART-T30VN (50 μM), ammonium persulfate (Promega, V3131, 0.1% w/v), and TEMED (Invitrogen, 15524-010, 0.1% v/v) were added to the molten agarose, vortexed, and placed back under vacuum. After four hours, APS and TEMED were added again at their original concentrations, vortexed, and placed under vacuum overnight at 45 °C. Following completion of the reaction, the agarose was heated to 95°C, filtered through a 0.45 μm syringe filter, and ultra-low gelling temperature agarose type IX-A (Sigma-Aldrich, A2576) was added to a final concentration of 2% type IX-A agarose (0.5% SFR and 1.5% agarose). The final 2% conjugated agarose was hardened and allowed to wash overnight in 500 ml of ultra-pure water. Agarose-bound oligonucleotides were quantified by QuBit ssDNA assay (Invitrogen, Q10212) and normalized to 4 μM polyT with 2% ultra-low gelling type IX-A agarose.

Encapsulation and lysis of cells within agarose hydrogels.—Oligo-dT conjugated agarose, cells, lysis buffer, and fluorinated oil were co-flowed onto a bubble triggered droplet microfluidic device to generate 55-μm diameter droplets[53]. Agarose was heated to 95 °C and vortexed until completely homogenous. Molten agarose was loaded into a 3 ml syringe (BD Biosciences, 309657) and placed into a custom-built syringe heater set to 85 °C. Cells were prepared as previously described[22]. In brief, mice were perfused with 1× PBS (Gibco, 10010023), the CNS was digested with papain suspension (Worthington, LS003126), mechanically dissociated, filtered through a 70 μm cell strainer (Fisher Scientific, 22363548), centrifuged at 500*g* for 5 min, resuspended in 10 ml of 30% Percoll solution (GE Healthcare Biosciences, 17-5445-01), and centrifuged at 500*g* for 25 min with no braking. The cell pellet was washed using HBSS with calcium or magnesium (Gibco, 14170112) and cells were counted manually using a haemocytometer. Cells were resuspended to 6.11×10^6 cells per ml in HBSS with 18% (v/v) OptiPrep (Sigma-Aldrich, D1556-250ML) and 1% Pluronic F-68 non-ionic surfactant (v/v) (Thermo Fisher, 24040032). Lysis buffer containing 20 mM Tris-HCl pH 7.5, 1,000 mM LiCl, 1% LiDS (Sigma, L9781-5G), 10 mM EDTA (VWR, E177), 10 mM DTT (Sigma, 43816), and 0.4 μg μl^{-1} proteinase K (NEB, P8107S) was prepared and loaded into a 3 ml syringe. A 10 ml syringe was loaded with droplet generation oil for probes (Bio-Rad, 186-3005). All syringes were connected to the microfluidic device using PE/2 tubing (Scientific Commodities, BB31695-PE/2). To make drops, pumps were run at 600 μl h^{-1} for cells, 1,200 μl h^{-1} for agarose, at 600 μl h^{-1} for lysis buffer, and 5,000 μl h^{-1} for oil. Air was controlled manually

at 20 psi and tuned to generate 55-μm droplets. Droplets were collected in a 50 ml Falcon tube, incubated at 55 °C for 2 h, and placed at 4 °C for at least 2 h.

Drop breaking and washes.—After the encapsulation workflow, solid agarose microspheres were removed from oil using a drop-breaking procedure. All steps were carried out at 4 °C to prevent dissociation of mRNA from the polyT oligonucleotides. Oil was removed from the bottom of tubes and 20% ^{1}H,^{1}H,^{2}H,^{2}H-perfluoro-1-octanol (PFO) (Sigma-Aldrich, 370533) in HFE-7500 oil was added at a ratio of 3:1 (v/v) PFO:emulsion. Tubes were gently mixed by hand and centrifuged for 1 min at 2,000*g*. Oil was removed and discarded and the remaining beads were washed in 50 ml of 0.1% SPAN-80 in hexane, shaken gently by hand, and centrifuged at 2,000*g* for 1 min. Next, beads were subjected to five washes. For each wash, beads were incubated in wash buffer for 5 min on ice, spun down at 4,700 rpm for 10 min, and aspirated before the next wash. Beads were washed 1× with wash buffer 1: 20 mM Tris-HCl pH 7.5, 500 mM LiCl, 0.1% LiDS, and 0.1 mM EDTA. Beads were then washed 2× with wash buffer 2: 20 mM Tris-HCl pH 7.5, and 500 mM NaCl. Beads were then washed 2× with wash buffer 3: 5× reverse transcription buffer (Thermo Fisher); this buffer contains 250 mM Tris-HCl pH 8.3, 375 mM KCl, 15 mM $MgCl_2$, and 50 mM DTT. Beads in 5× reverse transcription buffer were then filtered using a 100 μm cell strainer.

Synthesis of cDNA using bead-conjugated primers.—Cellular RNA captured on polyT primers conjugated to the agarose hydrogel was reverse transcribed to make cDNA using 1 mM dNTPs (Thermo Scientific, R1121), 2 μM template switch oligo (AAGCAGTGGTATCAACGCAGAGTACATrGrGrG), 6 mM $MgCl_2$ (Sigma-Aldrich, M1028), 1 M betaine (Sigma, 61962), 7.5% PEG-8000 (Promega, V3011), 2 U μl^{-1} Maxima H-minus reverse transcriptase, and 0.5 U μl^{-1} RNAse inhibitor (Lucigen, 30281-2). Reverse transcription was performed on an inverter at room temperature for 30 min, followed by 90 min at 42 °C. Following reverse transcription, agarose beads were washed 5 times in 0.1% Tween-20 (Sigma, P9416) in nuclease-free ultra-pure water (Invitrogen, 10977015). After the final wash, cDNA was amplified from a small aliquot of beads, to ensure reverse transcription was successful. In brief, pelleted agarose beads were diluted fourfold and stained with 10× SYBR green (Sigma, S9430) for 30 min in the dark. Beads were loaded into a haemocytometer using a Leica DMi8 fluorescence microscope and the concentration of genomes was determined based on fluorescence. WTA was performed in a 25 μl reaction containing 750 genomes using 1× KAPA HiFi master mix (Roche, KK2601) and 0.4 μM Smart-seq2 PCR primer (AAGCAGTGGTATCAACGCAGAGT). Reactions were thermocycled under the following conditions: 95 °C for 3 min, 14, 16 or 18 cycles of (98 °C for 15 s, 67 °C for 20 s, 68 °C for 4 min), followed by 72 °C for 5 min, then held at 4 °C. PCR was purified using a 2:1 ratio of AMPure XP beads (Beckman Coulter, A63881), and eluted in 20 μl. The size of the amplified cDNA was validated by a high-sensitivity DNA chip (Agilent, 5067-4626) with an average size of 1.5–2.0 kb and a negligible fraction of fragments below 300 bp.

Validation of TaqMan assays in drops.—To develop a multiplex assay targeting *Edem1*-expressing astrocytes, we tested combinations of assays targeting astrocyte marker

genes (Bio-Rad, dMmuCPE5095061; Thermo Fisher, Mm03048957_m1), and *Edem1* (Bio-Rad, dMmuCNS573941944; Thermo Fisher, Mm00551797_m1) with quantitative PCR (qPCR). We performed qPCR on agarose hydrogels containing purified genomes and transcriptomes of primary mouse cells captured from EAE mice, prepared as described above. The reactions consisted of 2 μl of hydrogels, 1× Taqman fast universal PCR master mix (Life Technologies, 4367846), and 1× of each TaqMan assay containing primers and probes. We chose *Aqp4* (Bio-Rad, dMmuCPE5095061) and *Edem1* (Thermo Fisher, Mm00551797_m1) because single-plex qPCR curves were equivalent to the multiplex curves for both assays.

Single-cell multiplexed digital droplet PCR.—Agarose hydrogels containing purified genomes and transcriptomes from single cells were mixed with PCR reagents to achieve a final concentration of 1× TaqPath master mix, 2.5% PEG, 2.5% Tween-20, and 1× of each TaqMan assay. The bead solution was mixed in the dark for an hour on a shaker, transferred to a 3 ml syringe, and spun at 4,700 rpm for 10 min in a centrifuge (Thermo Sorvall XTR) in custom 3D-printed syringe holders. The supernatant was removed, transferred to a second 3 ml syringe, and the remaining beads in the first syringe were pushed into PE/5 tubing (Scientific Commodities, BB31695-PE/5). Both syringes, and fluorinated oil (Bio-Rad, 1863006) were re-injected into a microfluidic device at 400 μl h^{-1} for beads, 400 μl h^{-1} for 1× PCR mix, and 2,000 μl h^{-1} for oil, and collected in 30 μl aliquots in PCR tubes. Drops in PCR tubes were then thermocycled under the following conditions: 88 °C for 10 min, then 50 cycles of (88 °C for 15 s and 60 °C for 1 min). After PCR amplification of the TaqMan targeted genes, the emulsions were pooled into a 3 ml syringe and sorted by droplet cytometry.

Microfluidic droplet cytometry.—Droplet microfluidic sorting was performed using a custom droplet cytometer controlled by LabView software. Three lasers (OptoEngine, MLL-FN-473, MLL-III-532, MRL-III-640) were aligned through dichroic mirrors (Semrock, FF552, R488), diverted into an inverted microscope (Motic, AE31) using a quad-pass dichroic mirror (Semrock, Di01-R405/488/532/635) and focused onto the microfluidic chip using a 20× objective. Photons from fluorescent droplets were collected by the same objective, filtered using a series of dichroic mirrors and bandpass filters (GFP channel FF552 and FF01-517/20; RFP channel FF593 and FF01-572/28; Cy5 channel FF01-673/11), and detected with photomultiplier tubes (PMTs) (Thorlabs, PMM01, PMM02). An embedded controller (National Instruments, cRIO-9056) programmed in LabView FPGA detected droplet peaks, integrated time series fluorescence data, and triggered on-chip sorting with a high-voltage amplifier (Trek, 609E-6). Gates, sorting parameters, and data acquisition occurred on a host computer graphical user interface built using LabView. Thermocycled droplets were re-injected into the concentric sorter[30] with the following flow rates: 100–150 μl h^{-1} for drops, 400 μl h^{-1} for drop spacing oil (Bio-Rad, 1863006), 2,000 μl h^{-1} for HFE used as additional spacer between drops, 3,000 μl h^{-1} for bias oil. The sorting parameters were 20 cycles of 1,000 V at 4,000 Hz. Drop sorting was verified by inspecting slow-motion videos taken of each sort using a high-speed camera (Vision Research, Miro110). For bulk FIND-seq, drops were collected in aliquots of 100 in 1.5 ml microcentrifuge tubes.

Validation of transcriptome amplification by FIND-seq.—Before amplifying sorted samples, the correct number was determined to ensure that drops were not over- or under-cycled during WTA. We sorted 100 genomes based on the percent of agarose hydrogels that contained cells, determined using fluorescence microscopy as described above. A 50 μl aqueous overlay of distilled nuclease-free water was added before spinning the tubes at 20,000*g* for 5 min. Tubes were then frozen at −80 °C for at least 2 h and 100 μl PCR reactions were set up with a final concentration of 1× KAPA HiFi master mix (Roche, KK2601) and 0.4 μM Smart-seq2 primer (AAGCAGTGGTATCAACGCAGAGT). For 100-aliquot sorts, we amplified triplicate aliquots of 100-cell sorts using the following PCR conditions: 95 °C for 3 min, *N* cycles of (98 °C for 15 s, 67 °C for 20 s, and 68 °C for 4 min), followed by 72 °C for 5-minutes, then held at 4 °C, where *N* varied as 18, 20, or 22 for primary mouse CNS cells to identify the optimal cycle number. Libraries were purified using AMPure XP magnetic beads (Beckman Coulter, A63881) at 1.2× ratio according to the manufacturer's protocol, then eluted in 20 μl nuclease-free water, and run on a Bioanalyzer 2100 high sensitivity DNA chip (Agilent Technologies, 5067-4626) to assess size and yield. The minimum cycle number that resulted in a concentration of greater than 2 ng μl^{-1} cDNA was then performed on the sorted samples.

Whole-transcriptome processing by FIND-seq.—After these validation steps, sorted samples were processed as above: 100-drop sorts were overlayed with 50 μl of distilled nuclease-free water, centrifuged at 20,000*g* for 5 min, frozen at −80 °C for 2 h, heated to 60 °C for 10 min, mixed gently by pipette and PCR amplified using the appropriate cycle number that was identified on cells from the same batch. After amplification, cDNA yield was quantified using a Qubit 4 fluorometer (Thermo Fisher) and DNA size distribution was assayed using a Bioanalyzer 2100 high sensitivity DNA chip (Agilent Technologies, 5067-4626). Libraries were prepared using the Nextera XT library preparation kit with v2 indices (Illumina, FC-131-1096) according to the manufacturer's instructions. Individual sample libraries were pooled in equimolar amounts to produce a single library. The library was quantified using KAPA fast universal qPCR kit (Kapa Biosystems, KK4824). Library concentration was validated by Bioanalyzer. The library was diluted and denatured in accordance with standard Illumina protocols. The library was sequenced by paired-end 75-bp reads on one lane of NovaSeq S2.

Bulk FIND-seq analysis.—The reads from the raw fastq files were trimmed using Trim Galore[54] in order to remove adapter contents and low-quality bases. The trimmed reads were aligned to GRCm38 mouse genome using STAR (v2.7.3)[55] with the default parameters for paired-end RNA-seq samples; the alignment process generated transcriptome-level alignment bam files. The alignment results were used by the quantification software RSEM[56] to yield transcript-level quantification, and tximport[57] was used to aggregate the gene-level quantification from the transcript-level quantification into a count matrix. FIND-seq detected an average of 17,566 genes per 100 cells. DESeq2[58] was used to complete the differential expression analysis, and ApeGLM[59] was used to shrink the log2 fold change results to remove noise from the differential expression analysis. The results of differential expression analysis were analysed using GSEA[60] and IPA[61]. To search for DNA binding motifs in the upstream elements of genes, the upstream 2000 base pair sequences were

extracted from the transcriptional start site of each gene using UCSC Genome Browser. Then the sequences were analysed using the FIMO algorithm[62] from MEME suite[63]. The motifs in the JASPAR database[64] were used in the analysis.

scFIND-seq.—A Smart-seq3 approach[65] was used for scFIND-seq. The protocol remained largely identical, with the following modifications. A Smart-seq3 bead-side mRNA capture and reverse transcription primer (/5Acryd/ TTTTTTACGAGCATCAGCAGCATACGATTTTTTTTTTTTTTTTTTTTTTTTTTV) was conjugated onto allyl agarose as described above. Cell encapsulation and washing were identical, but a Smart-seq3 template switch oligonucleotide (/5Biosg/ AGAGACAGATTGCGCAATGNNNNNNNNrGrGrG) containing an 8-bp UMI was used during reverse transcription. Bead re-injection and ddPCR followed the 100-cell sort procedures described above. Thereafter, single-drop sorts were performed using an automatic *xyz* stage to accurately deposit single sorted drops into a 96-well plate. Drops were overlayed with 25 μl of nuclease-free water, and frozen at −80 °C overnight. The following day, cells were defrosted for 10 min at 60 °C (Eppendorf, Thermomixer C), and centrifuged for 5 min at 1,000*g*. Whole-transcriptome cDNA amplification was performed in a 50 μl reaction for single cells using KAPA HiFi master mix (Roche, KK2601) and 1 μM Smart-seq3 forward primer (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTGCGCAA*T*G) and 0.5 μM Smart-seq3 reverse primer (ACGAGCATCAGCAGCATAC*G*A). Reactions were thermocycled at 95 °C for 3 min, then 26 cycles of (98 °C for 15 s, 65 °C for 20 s, and 68 °C for 4 min), followed by 72 °C for 5 min, and held at 4 °C. PCR reactions were purified using a 2× ratio of AMPure XP beads, followed by a second cleanup using 0.7× AMPure XP beads followed by elution in 10 μl of nuclease-free water. The size of the amplified cDNA was randomly sampled from individual wells of the 96-well plate by Bioanalyzer (Agilent Technologies, 5067-4626). Libraries were prepared using the Nextera XT DNA library preparation kit for 96 samples (Illumina, FC-131-1096) with 10-bp dual indices. Reaction volume was scaled down tenfold and automated with robotics (SPTLabtech Dragonfly). Individual sample libraries were pooled, purified by AMPure XP beads at a 0.9× ration, and sequenced on two separate NextSeq runs with 150-bp paired-end reads.

scFIND-seq analysis.—Kallisto[66] was used to align the reads from scFIND-seq by specifying the technology as Smartseq3 using '–technology SmartSeq3'. The 8-bp UMI in read 1 was used to assign unique transcript molecules to individual cells. The alignment results were quantified using bustools[67] by default parameters, and the quantification results were analysed in Seurat[68] as follows. The top 2,000 variable genes were used to conduct the principal component analysis reduction of the dataset, and the top 25 PCs were used to perform the nearest neighbourhood construction and UMAP reduction. The Leiden community clustering algorithm was used to cluster the dataset using resolution of 1.4. scFIND-seq had an average of 520 genes per cell, and 5,749 transcripts per cell.

Cell line validation studies

The JLat full-length human T cell line (6.3, ARP-9846) or Jurkat human T cells (Jurkat Clone E6-1, ATCC, TIB-152) were cultured in Gibco RPMI Media 1640 (Thermo Fisher,

11875093) with penicillin and streptomycin (Thermo Fisher, 15140122) and 10% fetal bovine serum (FBS). Mouse fibroblast cells (NIH/3T3, ATCC, CRL-1658) were cultured in DMEM with penicillin and streptomycin (Thermo Fisher, 15140122) and 10% FBS. 3T3 cells were dissociated using 0.25% trypsin-EDTA (Thermo Fisher, 25200-072) and neutralized in DMEM with 10% FBS. Cells were washed twice with HBSS, no calcium, no magnesium (Thermo Fisher, 14170112) and counted.

Mouse–human mixing studies.—Cells were mixed at a 1:100 JLat:3T3 ratio and resuspended in HBSS containing 18% OptiPrep density gradient medium (Sigma-Aldrich) for microfluidic encapsulation. Mouse–human mixing studies used agarose conjugated with acrydite-T5-SMART-T30VN and the Smart-seq2 TSO (AAGCAGTGGTATCAACGCAGAGTACATrGrGrG) for reverse transcription. The digital droplet genomic PCR was performed using 900 nM HIV *gag* forward primer (CACTGTGTTTAGCATGGTGTTT), 900 nM HIV *gag* reverse primer (TCAGCCCAGAAGTAATACCCATGT) and 250 nM HIV *gag* TaqMan probe (CY5-ATTATCAGAAGGAGCCACCCCACAAGA-3′ Iowa Black RQ). Drops were thermocycled under the following conditions: 88 °C for 10 min, then 55 cycles of (88 °C for 15 s and 60 °C for 1 min) and sorted into single wells. PCR strip tubes were frozen at −80 °C for at least 2 h and 50 μl WTA reactions were set up with a final concentration of 1× KAPA HiFi master mix (Roche, KK2601) and 0.4 μM Smart-seq2 primer (AAGCAGTGGTATCAACGCAGAGT). Library preparation was performed as described in the scFIND-seq section.

Validation of detection of a single genomic target in single cells.—JLat (contains a near full-length HIV provirus sequence) and Jurkat (HIV-negative) human T cells were independently cultured, washed, and encapsulated. FIND-seq was completed as described, with the following changes: No reverse transcription was performed to ensure that only a single copy of the HIV genomic target was present in the JLat cell line; detection PCR was performed as described above, using HIV *gag*-specific droplet TaqMan PCR. Quantification of positive droplets used a custom microfluidic droplet cytometer.

Generation of lentivirus

Constructs were generated using the pLenti-U6-sgScrmbl-Gfap-Cas9-2A-EGFP-WPRE lentiviral backbone, described previously[4]. Cas9 was driven by expression of the ABC_1D *gfa2 GFAP* promoter[69]. Two PCR products were generated using Phusion flash HF 2× master mix (Thermo Fisher, F548L) and the pLenti-U6-sgScrmbl-Gfap-Cas9-2A-EGFP-WPRE plasmid template. The U6-PCR-F and U6-PCR-R amplicon was purified (Qiagen, 28104) and digested with BsaI-HF (NEB, R3535/R3733), AscI (NEB, R0558) and DpnI (NEB, R0176S). The cr-RNA-F and cr-RNA-R amplicon, which included the sgRNA sequence, was purified and digested with BsaI-HF (NEB, R3535/R3733), SbfI-HF (NEB, R3642) and DpnI (NEB, R0176S). The backbone was digested with SbfI-HF (NEB, R3642) and AscI (NEB, R0558). All digestions were purified (Qiagen, 28104) and a three-way ligation performed overnight at 16 °C using T4 DNA Ligase (NEB, M0202L). Ligations were transformed into NEB Stable Cells (NEB, C3040), recovered at 30 °C for 1 h, and plated at 30 °C overnight. Six single colonies per

sgRNA construct were picked, grown in SOB media, and purified (QIAprep Spin Miniprep, 27104). The genes targeted for knockdown using the lentivirus delivery system were: *E2f1*(GCAGCACGTCAGAATCGCGA), *Rgs19*(TCATGGCTGGACATCGAGGG), *Etv5* (ACAGGACGACAACTCGGAGG), *Bcl3*(TGAGTAGGCAGGTTCAGCAG), *Thrb*(GCAGCGATAGTGGTACCCTG), *Ss18*(GCTCACCTTCTGGATGGCGG), *Nr3c2*(CAAGGAACTTTCAGCCACGG), *Hira*(CATGTGTGAACTGTGTGCGG), and *Angpt1* (CGCCGAAATCCAGAAAACGG). Lentiviral plasmids were co-transfected into HEK293FT cells with helper plasmids (pLP1, pLP2, and pLP/VSVG) according to the manufacturer's protocol (Thermo Fisher Scientific, K497500), concentrated overnight at 4 °C (Clontech, 631231) and resuspended in 150 μl of nuclease-free water.

In vivo CRISPR transduction

Intracerebroventricular injection was used to deliver concentrated lentivirus to mice as previously described[4]. Mice were anaesthetized with 1–3% isoflurane in oxygen. Heads were shaved and cleaned with ethanol and betadine (Thermo Fisher, 19-027132) and a small incision was made in the skin. Two injections of 10 μl were made bilaterally with a 25 μl Hamilton syringe (Sigma-Aldrich, 20787) at ±1.0 (lateral), −0.44 (posterior), −2.2 (ventral) relative to Bregma using a stereotaxic alignment system (Kopf, 1900). Mice were sutured and allowed to recover in a clean cage on a heating pad for 1 h. Mice were further recovered for 6 days before the induction of EAE.

Experimental autoimmune encephalomyelitis

Induction of EAE was performed as previously described[4]. In brief, 150 μg $MOG_{35–55}$ (Genemed Synthesis, 110582) per mouse was prepared from a 10 mg ml^{-1} stock diluted in 1× PBS. Complete Freund's adjuvant (CFA) was prepared by mixing 100 mg *Mycobacterium tuberculosis* H-37Ra (BD Biosciences, 231141) with 20 ml Incomplete Freund's adjuvant (BD Biosciences, BD263910). An emulsion was prepared by mixing MOG and CFA at 1:1 v/v using 20 ml Hamilton syringes until the emulsion was a thick, milky white solution. The MOG/CFA was delivered to mice with two subcutaneous injections of 100 μl each. *Pertussis* toxin (List Biological Laboratories, 180) was delivered with an intraperitoneal injection at a concentration of 2 ng μl^{-1} in 200 μl of PBS. EAE clinical scores were defined as follows: 0, no signs; 1, fully limp tail; 2, hindlimb weakness; 3, hindlimb paralysis; 4, forelimb paralysis; 5, moribund–as described previously[4,19–22,38].

Bulk RNA-seq of sorted astrocytes

Bulk RNA isolated from flow cytometry-sorted cells was processed using the PicoPure kit (Thermo Fisher, KIT0204) and used as input with the kit (NEB, E6420) according to the manufacturer's protocol. Reverse transcription was performed according to the Smart protocol using a template switching oligo. Then, cDNA was amplified and cleaned using Ampure XP beads (Beckman Coulter, A63881) and quantified using a Bioanalyzer DNA HS assay (Agilent, 50674626). Libraries were then fragmented, end-repaired, and ligated to Illumina compatible adapters followed by sample barcoding using NEBNext Multiplex Oligos for Illumina (E7335S, E7500S). Samples were selected again using Ampure XP beads and quantified using a Bioanalyzer. Libraries were quantified using a Kapa library quantification kit (Kapa Biosystems, KK4824) and run on an Illumina NextSeq550 as 1 ×

75 bp reads with 6-bp index read and de-multiplexed into FASTQ files. Alternatively, 5 ng of RNA was resuspended 5 μl of nuclease-free H2O and processed using the Smart-seq2 workflow at Broad Technology Labs and the Broad Genomics Platform. Processed RNA-seq data was filtered, removing genes with low read counts. The fastq files of each RNA-seq data sample were aligned to the *Mus musculus* GRCm38 transcriptome using Kallisto (v0.46.1), and the same software was used to quantify the alignment results. Differential expression analysis was conducted using DESeq2, and the log2 fold change was adjusted using apeGLM for downstream analysis.

Pathway analysis

Genes ranked by fold change were analysed using GSEA pre-ranked analyses[60]. GSEA analysis used the KEGG/Reactome/Biocarta (c2.cp.all), Gene ontology (c5.cp.all), and Hallmark (h.all) gene sets from Molecular Signatures Database v7.1. ENRICHR was used on lists of differentially expressed genes[70,71]. Upstream regulator analysis was completed using IPA (Qiagen)[61]. Transcriptional signatures were developed based on several publicly available gene sets: ‘GOBP response to mineralocorticoid’, ‘GOBP cellular response to mineralocorticoid stimulus’ and ‘Jepsen smrt targets’.

Primary astrocyte cultures and gene expression analysis—Procedures were performed largely as described previously[4,21,22,38]. Brains of mice aged P0–P3 were dissected into PBS on ice. Brains were pooled, centrifuged at 500*g* for 10 min at 4 °C and resuspended in 0.25% Trypsin-EDTA (Thermo Fisher Scientific, 25200-072) at 37 °C for 10 min. DNase I (Thermo Fisher Scientific, 90083) was then added to the solution, and the brains were digested for 10 more minutes at 37 °C. Trypsin was neutralized by adding DMEM/F12 + GlutaMAX (Thermo Fisher Scientific, 10565018) supplemented with 10% FBS (Thermo Fisher Scientific, 10438026) and 1% penicillin/streptomycin (Thermo Fisher Scientific, 15140148), and cells were passed through a 70-μm cell strainer. Cells were centrifuged at 500*g* for 10 min at 4 °C, resuspended in DMEM/F12+GlutaMAX with 10% FBS/1% penicillin/streptomycin and cultured in T-75 flasks (Falcon, 353136) at 37 °C in a humidified incubator with 5% CO_2, for 7–10 days until confluency was reached. Astrocytes were shaken for 30 min at 180 rpm, the supernatant was aspirated and the media was changed, then astrocytes were shaken for at least 2 h at 220 rpm and the supernatant was aspirated and the media was changed again. Medium was replaced every 2–3 days. Primary astrocytes were lysed in Buffer RLT (Qiagen) and RNA was isolated from cultured astrocytes using the RNeasy Mini kit (Qiagen, 74106). cDNA was transcribed using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies, 4368813). Gene expression was then measured by qPCR using TaqMan fast universal PCR master mix (Life Technologies, 4367846). TaqMan probes used in this study are: *Actb* (Mm02619580_g1), *Il1b* (Mm00434228_m1), *Ccl2* (Mm00441242_m1), *Il6* (Mm00446190_ m1), and *Xbp1* (Mm03464496_m1, Mm00457357_m1). qPCR data were analysed by the ddCt method by normalizing the expression of each gene for each replicate to the *Actb* housekeeping gene and then to the control group. Compound treatment was performed for 24 or 4 h with compounds diluted in DMEM/F12+GlutaMAX (Life Technologies, 10565042) that was supplemented with 10% FBS (Life Technologies, 10438026) and 1% penicillin/streptomycin (Life Technologies, 15140122). Compounds used in these studies are: 1 ng ml^{-1} IL-1β

(R&D Systems, 401-ML-005, 100 μg ml^{-1} stock in PBS), 1 ng ml^{-1} TNF (R&D Systems, 410-MT-010, 100 μg ml^{-1} stock in PBS), and 1 μM finerenone (MedChem Express, HY-111372, 1mM stock in DMSO).

Isolation of cells from adult mouse CNS

Astrocytes were isolated by flow cytometry as described[4,21,22,38]. In brief, mice were perfused with 1× PBS and the CNS was isolated into 10 ml of enzyme digestion solution consisting of 75 μl papain suspension (Worthington, LS003126) diluted in enzyme stock solution (ESS) and equilibrated to 37 °C. ESS consisted of 10 ml 10× EBSS (Sigma-Aldrich, E7510), 2.4 ml 30% D(+)-Glucose (Sigma-Aldrich, G8769), 5.2 ml 1 M $NaHCO_3$ (VWR, AAJ62495-AP), 200 μl 500 mM EDTA (Thermo Fisher Scientific, 15575020), and 168.2 ml ddH2O, filter-sterilized through a 0.22 μm filter. Samples were shaken at 80 rpm for 30–40 min at 37 °C. Enzymatic digestion was stopped with 1 ml of 10× hi ovomucoid inhibitor solution and 20 μl 0.4% DNase (Worthington, LS002007) diluted in 10 ml inhibitor stock solution (ISS). 10× hi ovomucoid inhibitor stock solution contained 300 mg BSA (Sigma-Aldrich, A8806), 300 mg ovomucoid trypsin inhibitor (Worthington, LS003086) diluted in 10 ml 1× PBS and filter-sterilized using at 0.22 μm filter. ISS contained 50 ml 10× EBSS (Sigma-Aldrich, E7510), 6 ml 30% D(+)-Glucose (Sigma-Aldrich, G8769), 13 ml 1 M $NaHCO_3$ (VWR, AAJ62495-AP) diluted in 170.4 ml ddH2O and filter-sterilized through a 0.22 μm filter. Tissue was mechanically dissociated using a 5 ml serological pipette and filtered through a 70 μm cell strainer (Fisher Scientific, 22363548) into a fresh 50 ml conical. Tissue was centrifuged at 500*g* for 5 min and resuspended in 10 ml of 30% Percoll solution (9 ml Percoll (GE Healthcare Biosciences, 17-5445-01), 3 ml 10× PBS, 18 ml ddH2O). Percoll suspension was centrifuged at 500*g* for 25 min with no brakes. Supernatant was discarded and the cell pellet was washed 1× with 1× PBS, centrifuged at 500*g* for 5 min and prepared for downstream applications.

Flow cytometry

Cells were stained in the dark on ice for 15 min with flow cytometry antibodies. Cells were then washed once with 1× PBS and resuspended in 1× PBS for sorting as described previously[4,21,22,38]. Antibodies used in this study were: PE anti-mouse CD45R/B220 (BD Biosciences, 553089, 1:100), PE anti-mouse TER-119 (Biolegend, 116207, 1:100), PE anti-O4 (R&D Systems, FAB1326P, 1:100), PE anti-CD105 (eBioscience, 12-1051-82, 1:100), PE anti-CD140a (eBioscience, 12-1401-81, 1:100), PE anti-Ly-6G (Biolegend, 127608, 1:100), PerCP anti-Ly-6C (Biolegend, 128028, 1:100), APC anti-CD45 (eBioscience, 17-0451-83, 1:100), APC-Cy7 anti-CD11c (BD Biosciences, 561241, 1:100), and FITC anti-CD11b (eBioscience, 11-0112-85, 1:100). All cells were gated on the following parameters: $CD105^{neg}CD140a^{neg}O4^{neg}Ter119^{neg}Ly\text{-}6G^{neg}CD45R^{neg}$. Astrocytes were subsequently gated on: $CD11b^{neg}CD45^{neg}Ly\text{-}6C^{neg}CD11c^{neg}$. Microglia were subsequently gated on: $CD11b^{high}CD45^{low}Ly\text{-}6C^{low}$. Pro-inflammatory monocytes were subsequently gated on: $CD11b^{high}CD45^{high}Ly\text{-}6C^{high}$. Compensation was performed on single-stained samples of cells and an unstained control. Cells were sorted on a FACS Aria IIu (BD Biosciences). Sorted astrocytes were subsequently stained for intracellular XBP1S using the antibody PE Mouse anti-XBP-1S (BD, 562642). For AQP4 sorting experiments, isolated CNS cells were stained with AmCyan dye (Thermo Fisher, L34957, 1:2,000) for 30 min on ice, washed

1× with 1× PBS, then stained with anti-AQP4-PE (Bioss, BS-0634R-PE, 1:100) for 20 min on ice. Thereafter, cells were washed with 1× PBS and subjected to flow cytometry. For FIND-seq analysis of *Aqp4*$^{+}$ cells, bead loading was estimated at 70% to determine number of positive cells within identified droplets.

T cell FACS analysis—To analyse T cell populations, CNS and splenic cell suspensions were stimulated with 50 ng ml^{-1} phorbol 12-myristate 13-acetate (PMA, Sigma-Aldrich, P8139), 1 μM ionomycin (Sigma-Aldrich, I3909-1ML), GolgiStop (BD Biosciences, 554724, 1:1500) and GolgiPlug (BD Biosciences, 555029, 1:1500) diluted in T cell culture medium (RPMI (Life Technologies, 11875119) containing 10% FBS, 1% penicillin/streptomycin, 50 μM 2-metcaptoethanol (Sigma-Aldrich, M6250), and 1% non-essential amino acids (Life Technologies, 11140050)). After 4 h, cell suspensions were washed with 0.5% BSA, 2 mM EDTA in 1× PBS and incubated with surface antibodies and a live/dead cell marker on ice. After 30 min, cells were washed with 0.5% BSA, 2 mM EDTA in 1× PBS and fixed according to the manufacturer's protocol of an intracellular labeling kit (eBiosciences, 00-5523-00). Surface antibodies used in this study were: BUV661 anti-mouse CD45 (BD Biosciences, 565079, 1:100), BV750 anti-mouse CD3 (Biolegend, 100249, 1:50), PE-Cy7 anti-mouse CD4 (eBioscience, 25-0041-82, 1:100), BUV737 anti-mouse CD11b (BD Biosciences, 564443, 1:100), BV570 anti-mouse Ly6C (Biolegend, 128030, 1:100), BUV805 anti-mouse CD8a (BD Bioscience, 564920, 1:100), BUV563 anti-mouse Ly6G (BD Biosciences, 565707, 1:100), BUV737 anti-mouse CD11c (BD Biosciences, 612797, 1:100), Pe/Cy5 anti-mouse CD44 (BioLegend, 103010, 1:100). Intracellular antibodies were: APC anti-mouse IFN-γ (BD Biosciences, 554413, 1:100), PE anti-mouse IL-17A (eBiosciences, 12-7177-81, 1:100), BV421 anti-mouse GM-CSF (BD Biosciences, 564747, 1:100), PE/Dazzle 594 anti-mouse IL-10 (BioLegend 505033), and FITC anti-mouse FoxP3 (eBiosciences, 11-5773-82, 1:100). Gating of CNS and splenic cells was performed on >50,000 live CD4^{+} cells. FACs was performed on a Symphony A5 (BD Biosciences).

Analysis of human brain tissue by immunofluorescence

Use of tissue from patients with multiple sclerosis was approved by the Neuroimmunology Research Laboratory, Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM) under ethical approval number BH07.001. Human brain tissue was obtained from patients with clinical and neuropathological multiple sclerosis diagnosis according to the revised 2010 McDonald's criteria[72]. Tissue samples were collected from healthy donors and patients with multiple sclerosis with full ethical approval (BH07.001) and informed consent as approved by the local ethics committee. Flash frozen 7 μm tissue sections were fixed in 4% PFA for 10 min at room temperature. Next, sections were washed 3× with 1× PBS, followed by three washes with 0.3% PBS-T. Sections were blocked with 5% donkey serum (D9663, Sigma-Aldrich) in 0.3% PBS-T at room temperature for 30 min. Sections were then incubated with primary antibody (mouse anti-NR3C2, Abcam, ab2774, Clone: H10E4C9F, 1:100) diluted in blocking buffer overnight at 4 °C. Following primary antibody incubation, sections were washed 3× with 0.3% PBS-T and incubated with secondary antibody (Alexa Fluor 488 AffiniPure Fab Fragment Goat Anti-Mouse IgG (H+L), Jackson ImmunoResearch, 115-547-003, 1:500) diluted in blocking buffer for

2 h at room temperature. Sections were then washed 6× with 0.3% PBS-T and stained with pre-conjugated GFAP-CY3 (Sigma-Aldrich, C9205, 1:100). Sections were then dried and mounted using Fluoromount-G with DAPI (SouthernBiotech, 0100-20) and imaged using a Zeiss LSM710 confocal microscope. Staining was quantified by first selecting $DAPI^+GFAP^+$ cells then quantifying the number of those cells expressing NR3C2.

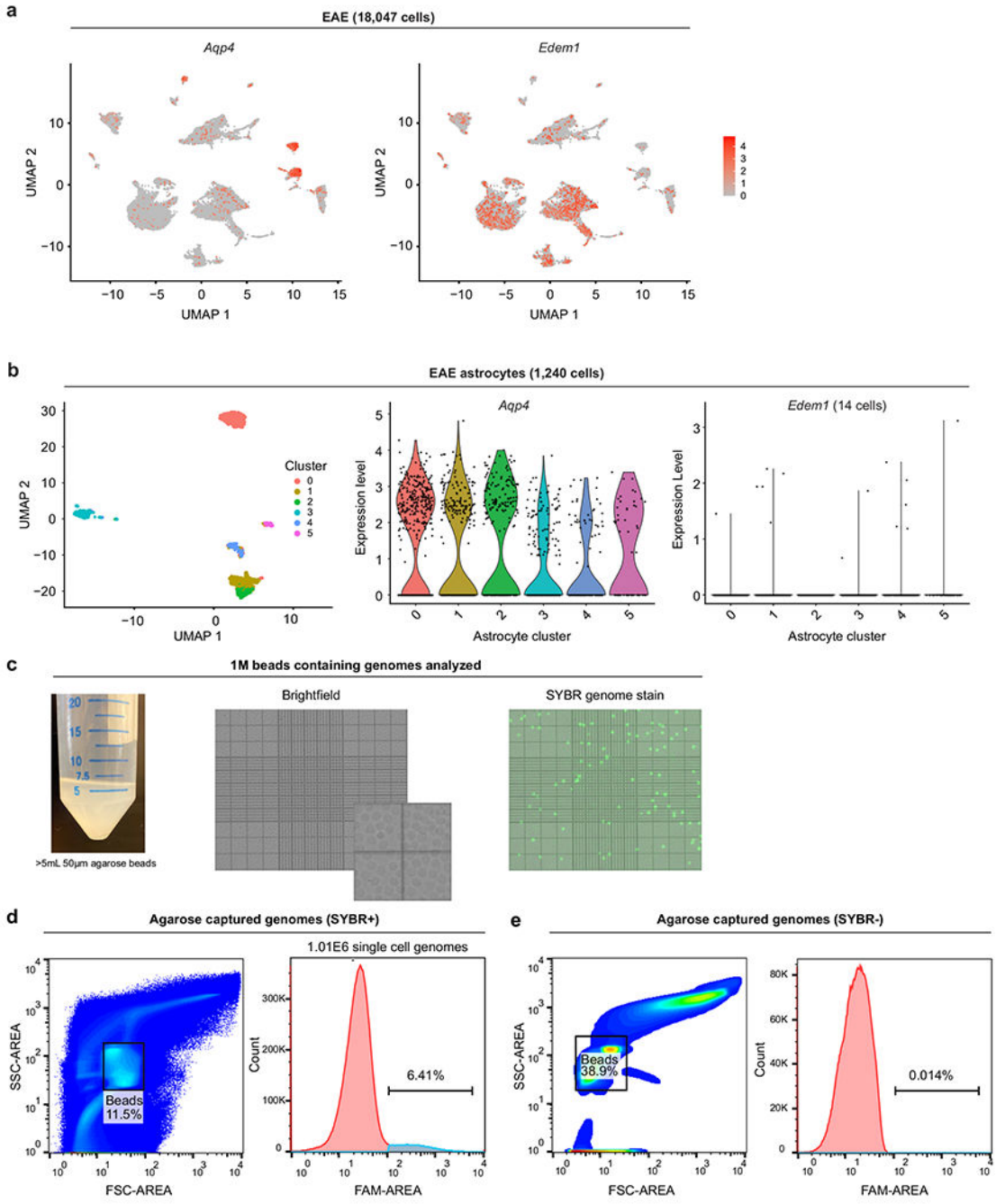
Analysis of mouse CNS tissue by immunostaining—Mice were intracardially perfused with ice cold 1× PBS and CNS tissue was excised from the mouse then post-fixed in 4% PFA overnight at 4 °C, then the tissue was dehydrated with 30% sucrose for one week at 4 °C. If whole spinal columns were collected containing bone, they were de-calcified in 20% EDTA (pH = 7.4) for one week at 4 °C with inversion prior to sucrose dehydration. Tissue was frozen in OCT (Sakura, 4583) and 20 μm sections were prepared by cryostat on SuperFrost Plus slides (Fisher Scientific, 15-188-48). Sections were permeabilized with 1× permeabilization buffer (BD Biosciences, 554723) for 10 min at room temperature, then blocked using serum-free protein block (Agilent, X0909) for 10 min at room temperature. Sections were then incubated with primary antibodies diluted in 1× permeabilization buffer overnight at 4 °C. Following primary antibody incubation, sections were washed 3× with 1× permeabilization buffer and incubated with secondary antibodies diluted in 1× permeabilization buffer for 2 h at room temperature. After secondary and conjugated antibody incubation, sections were stained with 1 μg ml^{-1} DAPI (Sigma-Aldrich, D9542) diluted in 1× permeabilization buffer for 5 min at room temperature, then washed 3× with 1× permeabilization buffer and mounted with ProLong Gold Antifade Mountant (Fisher Scientific, P36930) or alternatively, Fluoromount-G with DAPI (SouthernBiotech, 0100-20). Primary antibodies used in this study were: rabbit anti-XBP1S (Biolegend, 619502, 1:50), mouse anti-NR3C2, (Abcam, ab2774, clone: H10E4C9F, 1:100), rabbit anti-NCOR2 (Abcam, ab5802, 1:100), chicken anti-GFAP (Abcam, ab4674, 1:100), or mouse anti-GFAP (Millipore, MAB360, 1:500). Secondary antibodies used in this study were: Rhodamine Red-X-AffiniPure Fab Fragment donkey anti-mouse IgG (H+L) (715-297-003, Jackson ImmunoResearch, 2 mg ml^{-1} stock), Goat anti-Chicken IgY (H+L) Secondary Antibody Alexa Fluor 647 (Jackson Immunoresearch, 703-475-155), Goat anti-Chicken IgY (H+L) Alexa Fluor 488 (Life Technologies, A11039), and Goat Anti-Rabbit IgG H&L Alexa Fluor 488 (Abcam, ab150077) all at 1:500 working dilution. Sections were imaged on an LSM-880-AiryScan confocal microscope (Zeiss) or LSM710 confocal microscope (Zeiss) and image analysis was performed with FIJI[73] first by selecting $GFAP^+DAPI^+$ cells and then quantifying the proportion positive for a given marker protein.

Statistical analysis

Statistical analyses were performed with Prism software (GraphPad, v9.4.0). Data plotted in Extended Data Table 1 were generated in Excel (Microsoft, v16.64). All data represented shown are represented as mean ± s.e.m., unless indicated otherwise. The specific statistical test used for each comparison is indicated in the respective panel in the figure legend and the exact *P* values are indicated in each panel. In brief, for the comparison of two datasets, an unpaired two-sided *t*-test was used. To assess changes in a dependent variable across a range of conditions a one-way ANOVA was used. For analyses involving two variables, such as EAE time courses based on genotype, a two-way repeated measures ANOVA was used. Data

were tested for normality by the D'Agostino and Pearson test, but if the number of replicates per experiment was insufficient to test for normality, a Gaussian distribution was assumed. Testing for activation of pathways was determined by *z*-score measured using IPA, which is a metric that represents predicted relative activity of the indicated pathway in an input gene list. The *z*-score for each pathway is specifically calculated using a list of curated genes assigned to each pathway; a simplified *z*-score is calculated as follows: $z = \frac{N^+ - N^-}{\sqrt{N}}$, where N is the total number of genes and N^+ and N^- are the number of genes where the direction of regulation is concordant or discordant, respectively, with predictions from the literature within the curated pathway. A positive *z*-score suggests pathway activation while a negative *z*-score suggests pathway inhibition. Statistical significance of the enrichment of a pathway or control by an upstream regulator was determined using a Fisher's exact test based on overlap between the curated gene list and the input gene list as previously described[61].

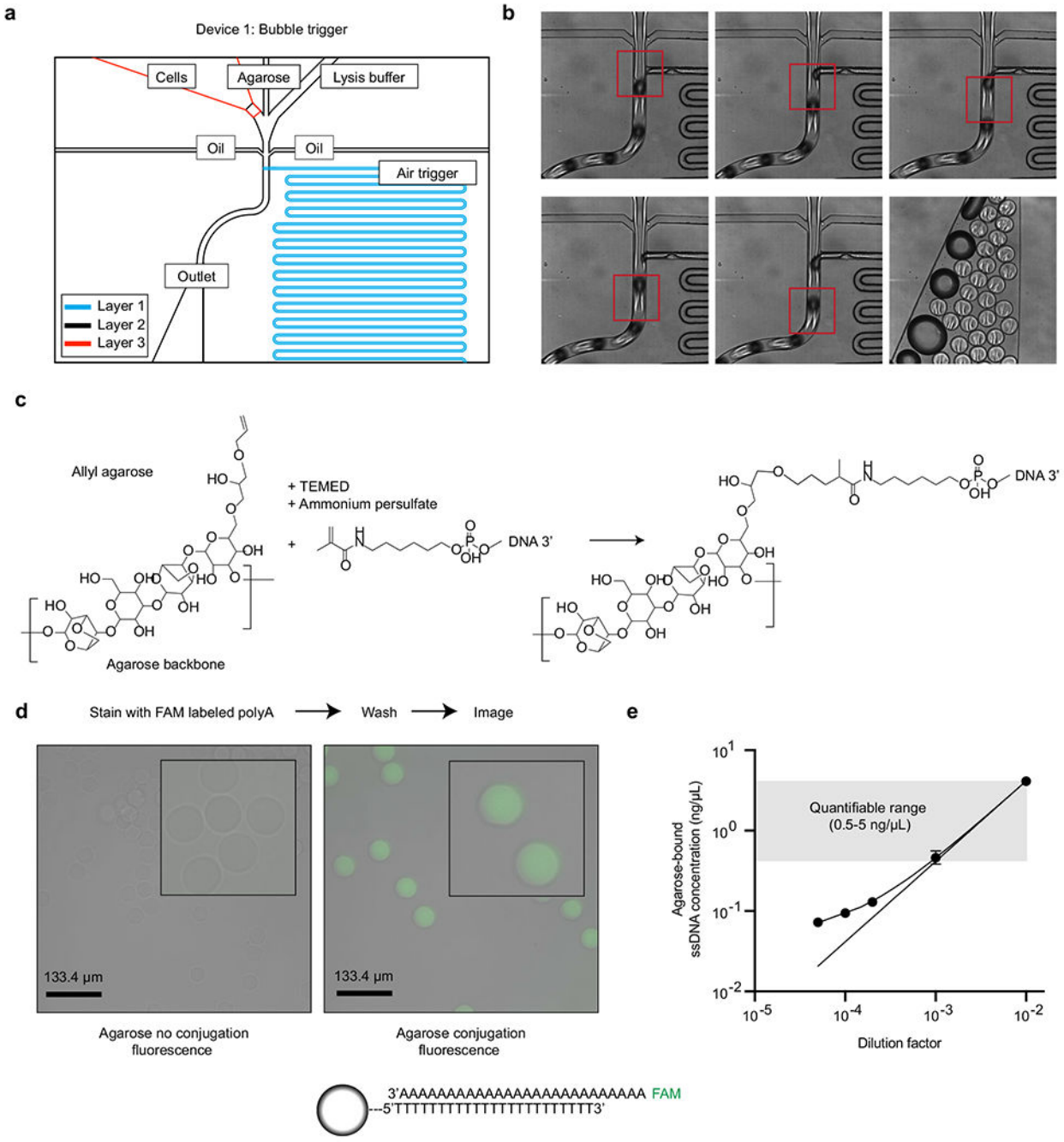
Extended Data



Extended Data Fig. 1 |. Analysis of rare cells by FIND-seq.

(**a**) Featureplots of *Aqp4* and *Edem1* expression in cells isolated from the mouse CNS during EAE reanalyzed from ref.[22]. (**b**) Violin plots of *Aqp4* and *Edem1* by cluster from astrocytes isolated from the EAE CNS in a dataset re-analyzed from ref. [22]. (**c**) Micrograph images of agarose beads containing captured nucleic acids from encapsulated cells. (**d**)

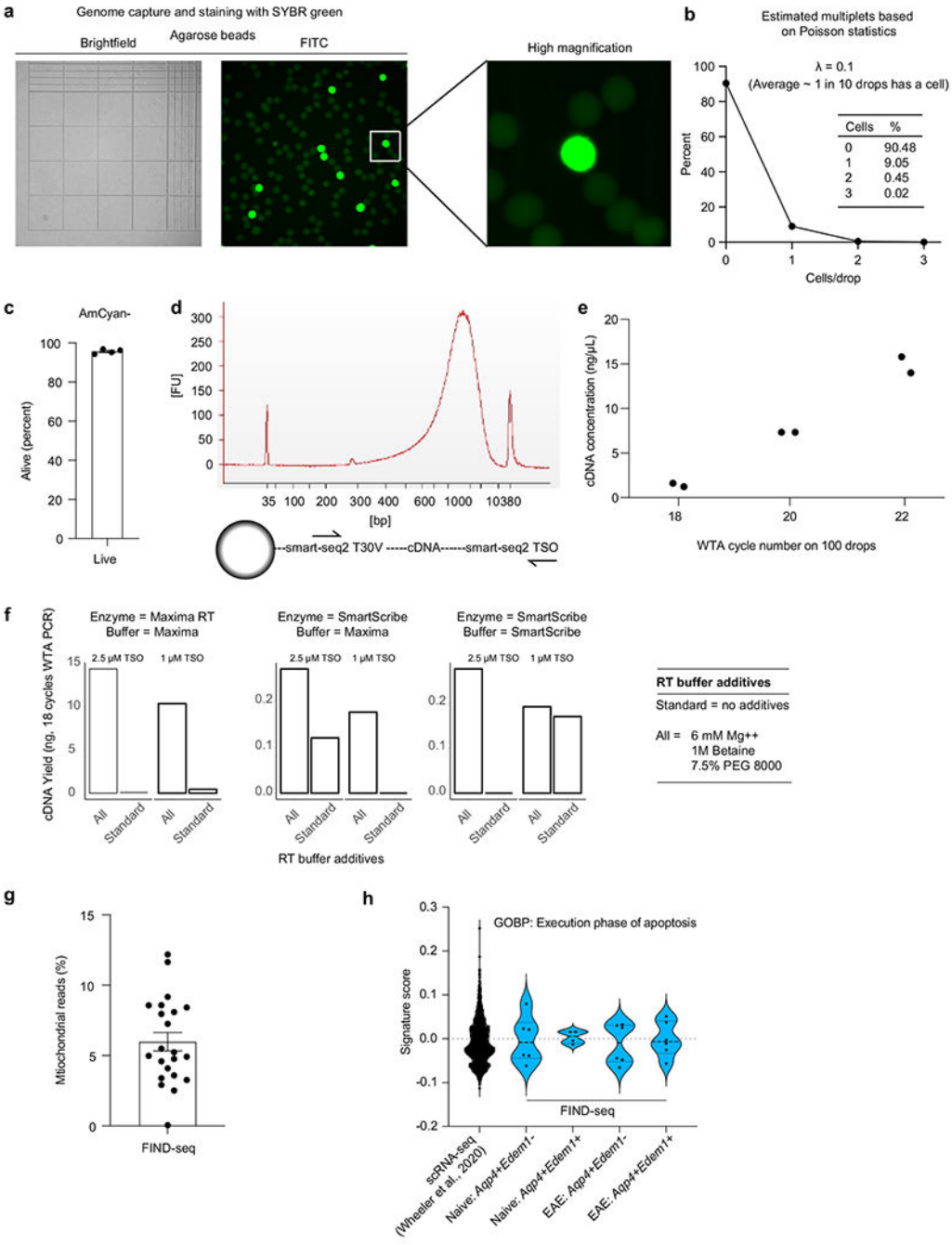
Droplet cytometry plots showing an estimate of the number of cells captured (>1 million) in the bead volume shown in (**c**). (**e**) Negative control of droplet cytometry SYBR fluorescence.



Extended Data Fig. 2 |. Droplet formation and capture of mRNA by FIND-seq.

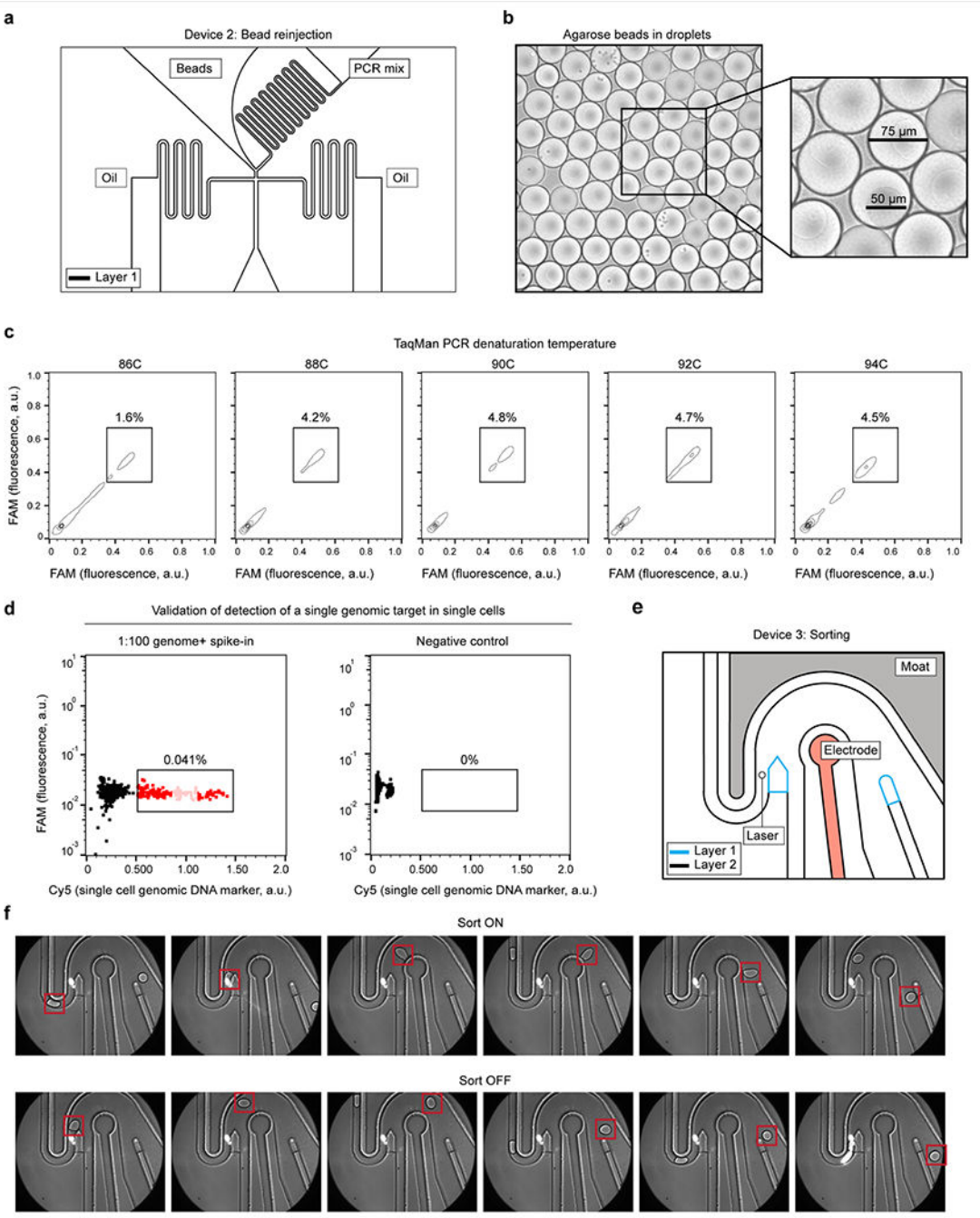
(**a**) Schematic of bubble trigger device that drives air-triggered droplet formation during cell encapsulation. (**b**) Air-triggered droplet formation enables kHz generation of beads from viscous molten agarose. Images from time-lapse videos of agarose jet breakup. (**c**) Functionalization of agarose with allyl groups is used to directly link acrydited primers. (**d**)

Microscope images demonstrating successful conjugation of polyT primers to agarose with polyA-FAM probes. (**e**) Quantification of agarose-bound polyT.



Extended Data Fig. 3 |. Optimization of first-strand synthesis of agarose captured mRNA.
(**a**) Steric capture of cellular genomic DNA inside the agarose matrix. Left: brightfield image of agarose gels. Right: SYBR green fluorescence of stained genomes inside agarose gels. Cells are loaded at limiting dilution to ensure single-cell encapsulation; approximately one in every ten agarose hydrogels contains a cell. (**b**) Estimated number of cells per drop

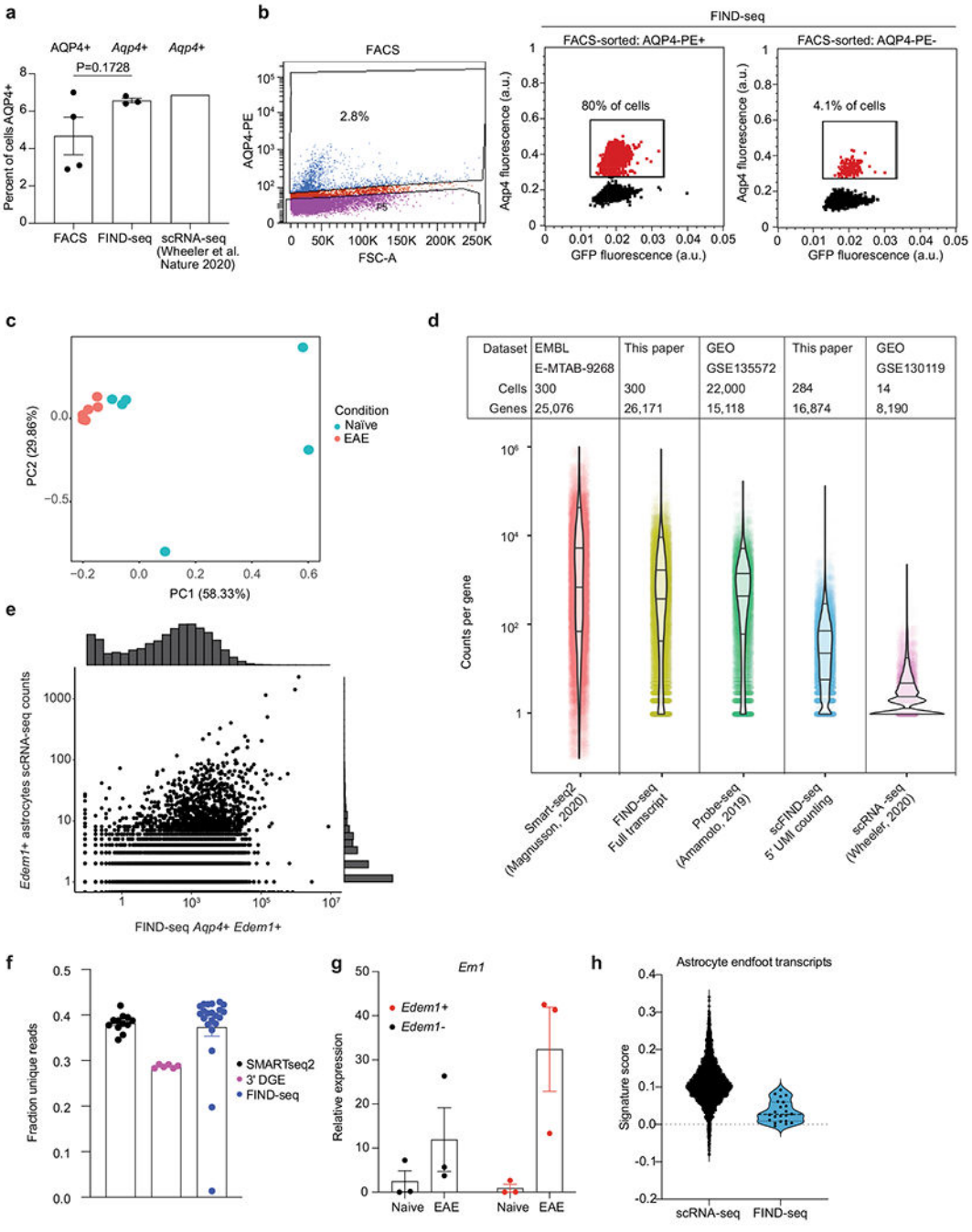
based on Poisson statistics for microfluidic loading during FIND-seq. (**c**) Quantification of live cells by flow cytometry using AmCyan live/dead cell dye. n = 4 mice. (**d**) Whole transcriptome amplification (WTA) of cDNA covalently attached to agarose beads shows full length material is captured and reverse transcribed. (**e**) WTA yield as a function of PCR cycle number. (**f**) Optimization of cDNA capture with buffer composition, enzyme, template switch oligonucleotide concentration and additives (6 mM Mg^{2+}, 1M betaine, and 7.5% PEG-8000). (**g**) Quantification of the percent mitochondrial reads in bulk FIND-seq data for each replicate. n = 22 samples. (**h**) Calculation of score per cell from astrocytes derived from scRNA-seq from ref. [22] or each bulk FIND-seq replicate for the pathway GOBP: Execution phase of apoptosis (GO: 0097194).



Extended Data Fig. 4 |. Agarose bead re-injection and sorting for single-cell detection of nucleic acid markers.

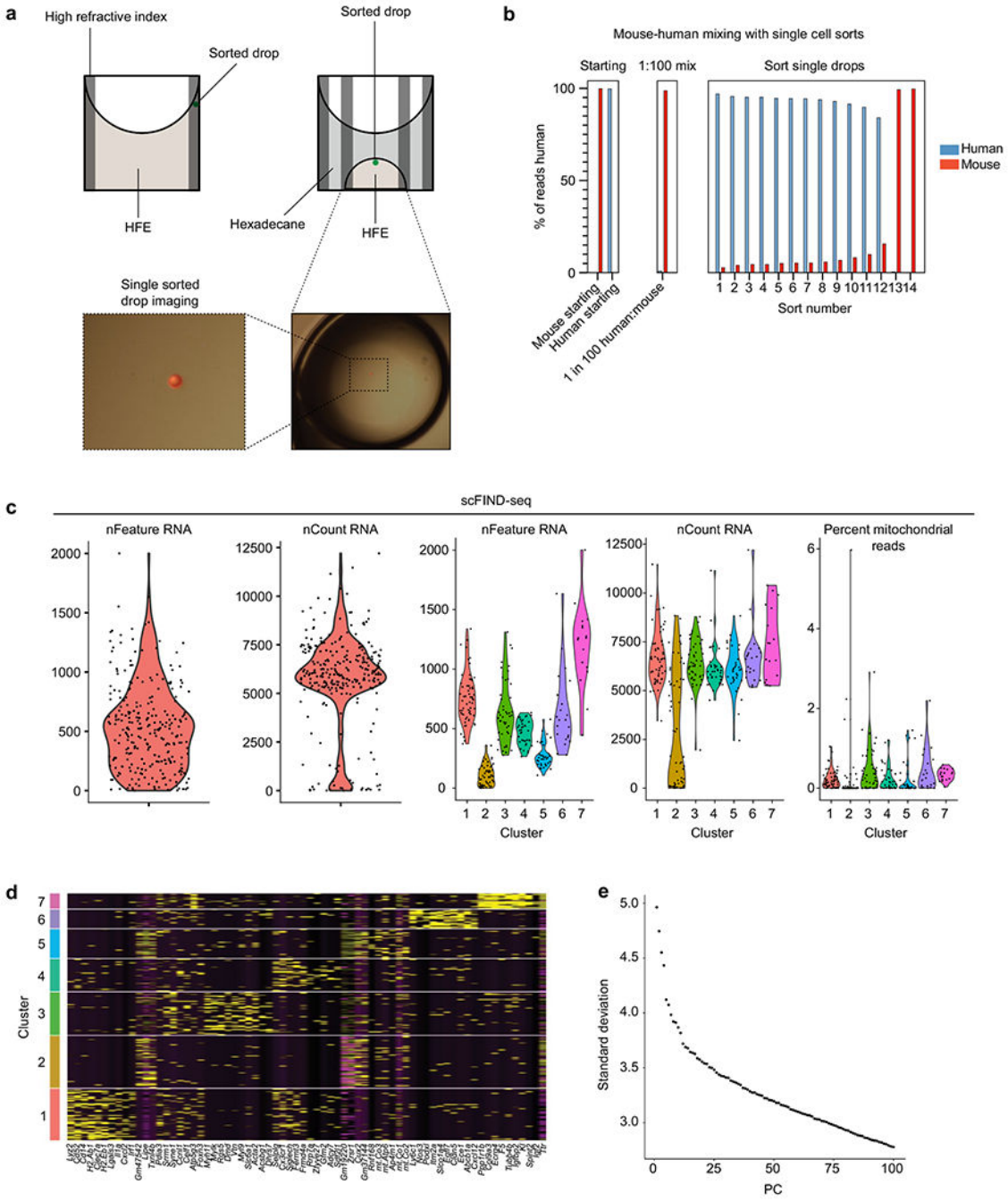
(**a**) Schematic of microfluidic for re-injection of agarose-captured genomes. (**b**) Microscope image of single agarose hydrogel beads inside droplets with measured size for each droplet and agarose sphere. (**c**) Optimization of droplet detection PCR to preserve cDNA quality during thermocycling. (**d**) Detection of a single-copy HIV genomic DNA target by FIND-seq in infected human JLat cells, but not in uninfected Jurkat control cells, as a proof-of-concept experiment testing FIND-seq sensitivity and specificity. The DNA target

was amplified using TaqMan PCR in the FIND-seq workflow (Step 3 in Fig. 1d) followed by detection by droplet cytometry (Step 4 in Fig. 1f). (**e**) Schematic of microfluidic for droplet sorting using a concentric dielectrophoretic design. (**f**) Micrographs of droplet sorting. Top: time-lapse images from a droplet sorting video showing droplet deflection into the collection channel. Bottom: In the absence of FPGA sort-triggering, droplets are maintained, via bias oil flow, in the outer waste channel.



Extended Data Fig. 5 |. Benchmarking FIND-seq as a technology.

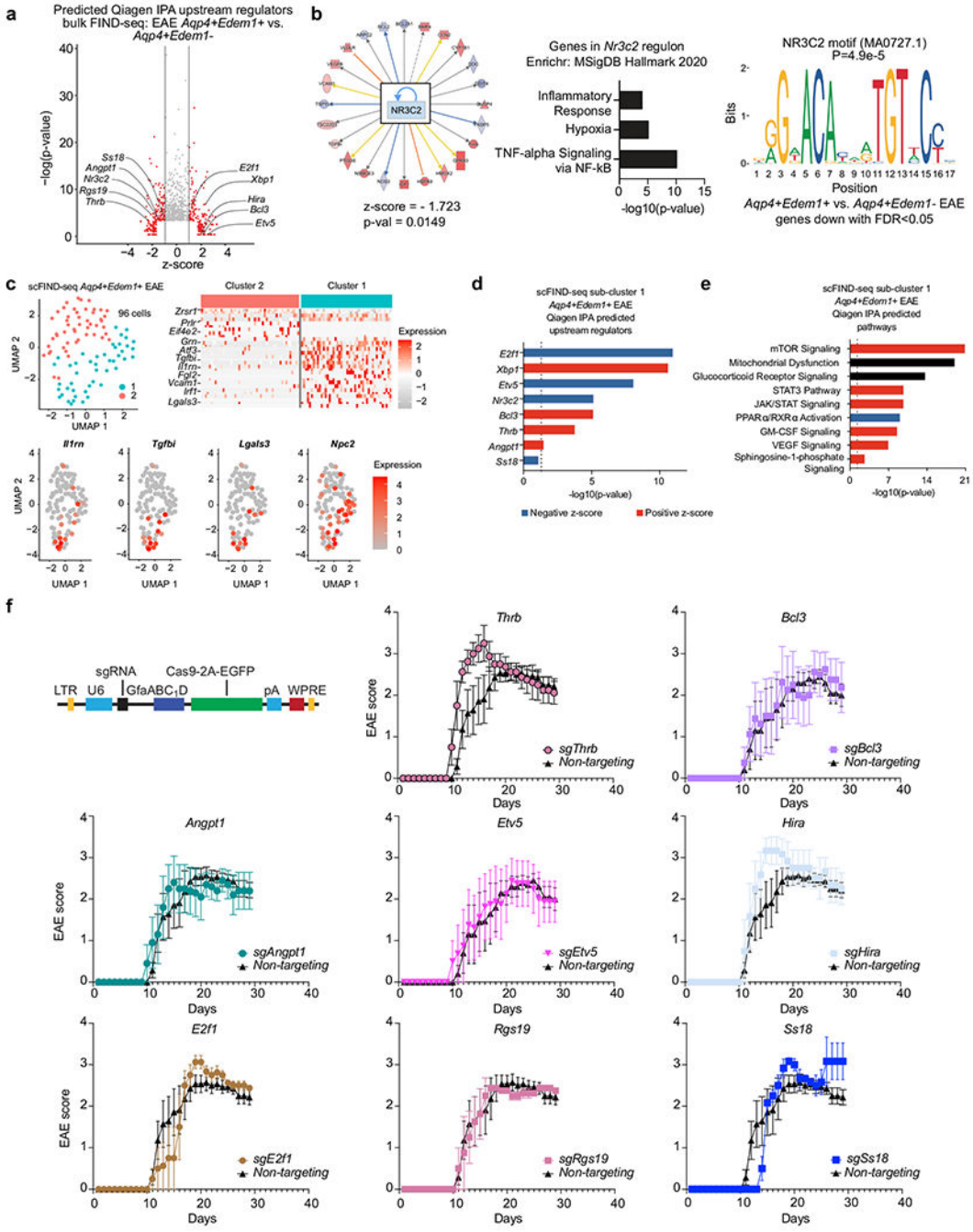
(**a**) Quantification of percent aquaporin-4-expressing cells in naïve mice by antibody-based flow cytometry (n = 4); FIND-seq (n = 3), or scRNA-seq reanalyzed from ref. [22]. (**b**) Validation of FIND-seq specificity and sensitivity through the analysis of AQP4+ and AQP4− cells isolated by flow cytometry and subjected to FIND-seq. Percentages show number of cells expressing *Aqp4* in each population based on 70% bead loading. (**c**) PCA plot of bulk FIND-seq analysis of *Aqp4*+*Edem1*− cells. (**d**) Comparison of FIND-seq detection sensitivity with comparable technologies. (**e**) Correlation of raw expression counts per gene between bulk FIND-seq-sorted *Aqp4*+*Edem1*+ cells and *Edem1*+ astrocytes extracted from droplet-based scRNA-seq data that we previously reported in[22]. (**f**) Quantification of fraction of duplicate reads across bulk RNA-seq platforms. (**g**) Quantification of *Ern1* in bulk FIND-seq data as a function of EAE and *Edem1* expression. (**h**) Calculation of a signature score for transcripts enriched in astrocyte endfeet as reported by ref. [51], analyzed in astrocytes from ref. [22] and bulk FIND-seq.



Extended Data Fig. 6 |. Single-cell analysis by FIND-seq.

(**a**) Visualization of a single droplet in flat-bottom microwell plates. Refraction of light at well edges obscures imaging. This is solved by sorting directly into hexadecane. HFE oil sinks, forming a convex shape that forces droplets to the center so that they can be imaged. (**b**) The percentage of reads mapping to mouse or human cells after single-cell sorting cell mixtures. Mouse (3T3) and human (JLat) cells were mixed 1:100, sorted based on a TaqMan PCR targeting JLat cells, and the transcriptome was sequenced. (**c**) Quality control analyses for scFIND-seq. (**d**) Marker genes of *Aqp4*+*Edem1*+/− cells from naïve or

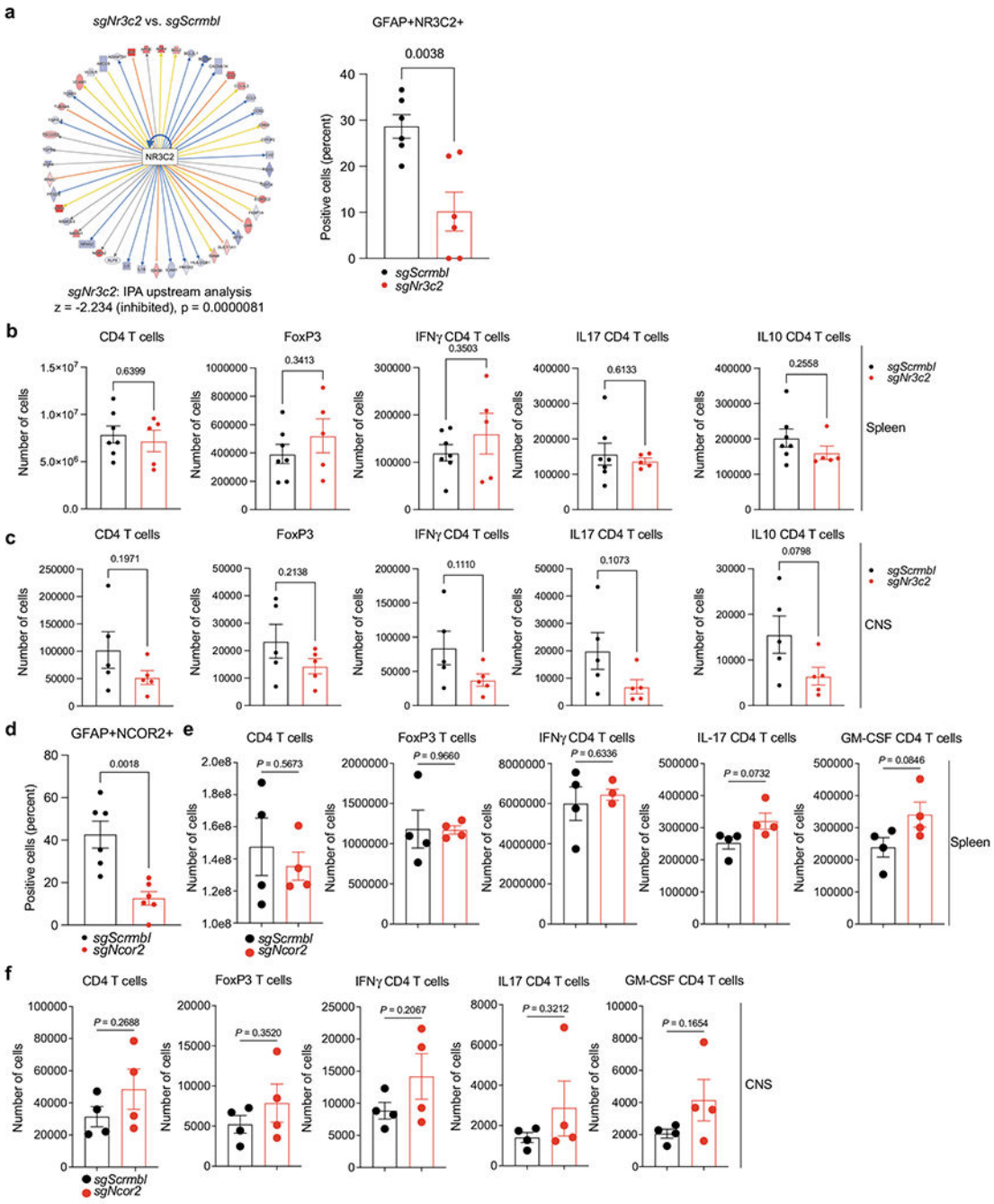
EAE mice analyzed by scFIND-seq. (**e**) Elbow plot of principal components detected by scFIND-seq.



Extended Data Fig. 7 |. *In vivo* screening of FIND-seq-identified candidate regulators of XBP1+ astrocytes.

(**a**) Predicted upstream regulator analysis showing *Nr3c2* and *Xbp1* from bulk FIND-seq data using Qiagen IPA. Differentially expressed genes were used as input and the overlap with the regulon controlled by each molecule was computed. Fisher's exact test. (**b**) Left: Prediction of *Nr3c2* as an upstream regulator in *Aqp4+Edem1+* cells analyzed by FIND-seq

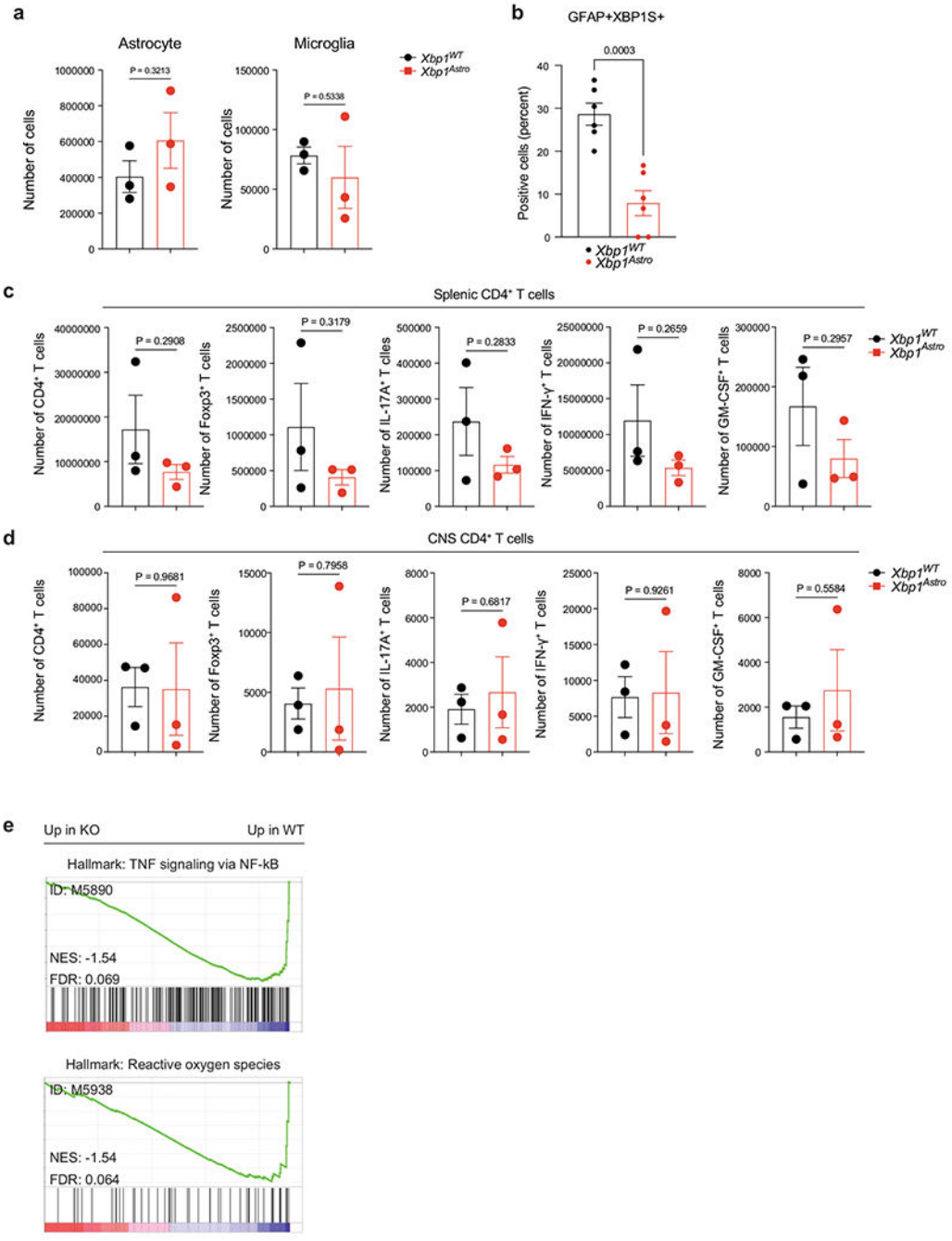
during EAE using Qiagen IPA as in (**a**). Fisher's exact test. Right: Identification of an NR3C2 motif by SeqPos in genes downregulated in *Aqp4+Edem1+* versus *Aqp4+Edem1−* cells in EAE. (**c**) UMAP plot of *Aqp4+Edem1+* cells from EAE mice analyzed by scFIND-seq. (**d**–**e**) Prediction of upstream regulators (d) and pathway analysis (e) based on Qiagen IPA in Cluster 1 astrocytes analyzed from *Aqp4+Edem1+* cells in EAE mice shown in (**c**). Fisher's exact test. (**f**) Schematic of lentiviral vector containing a sgRNA targeting candidate genes and *spCas9* under the control of a *Gfap* promoter. EAE disease progression in mice transduced with *sgScrmbl* and candidate sgRNA lentiviruses. n = 4–5 mice per condition.



Extended Data Fig. 8 |. Control analyses of *Nr3c2* and *Ncor2* knockdown.

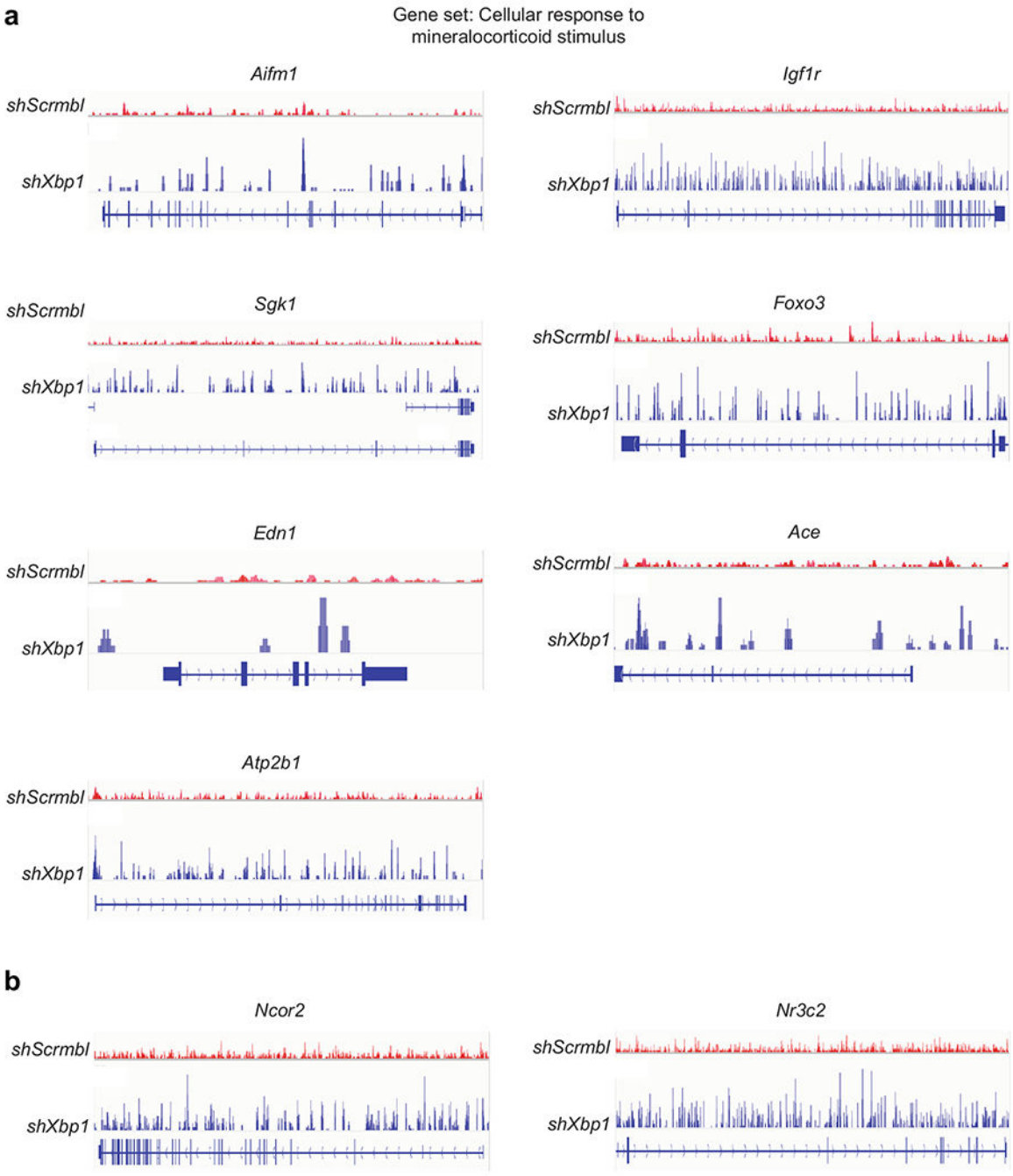
(**a**) Left: Upstream regulator analysis of RNA-seq data by Qiagen IPA from *sgNr3c2*-targeted versus *sgScrmble*-targeted mice shows NR3C2 downregulation. Fisher's exact test. Right: Validation of NR3C2 knockdown by immunostaining quantification. n = 6 images per group from n = 3 mice. Unpaired two-sided t-test. (**b**–**c**) FACS analysis of total CD4+ cells or FoxP3+, IFNγ+, IL17+, and IL10+ CD4 cell subsets in the (**b**) the spleen and (**c**) CNS. (**d**) Validation of NCOR2 knockdown by immunostaining. n = 6 images per group from n =

3 mice. Unpaired two-sided t-test. (**e**–**f**) FACS analysis of total CD4+ cells or FoxP3+, IFNγ+, IL17+, and GM-CSF+ CD4 cell subsets in the (**e**) the spleen and (**f**) CNS.



Extended Data Fig. 9 |. Control analysis of cell subsets from *Xbp1*^{WT} and *Xbp1*^{Astro} mice.
(**a**) FACS analysis of astrocytes and microglia in the CNS. n = 3 per group. (**b**) Validation of XBP1 KO by immunostaining. n = 6 images from n = 3 mice per group. Unpaired two-sided t-test. (**c**–**d**) FACS analysis of total CD4+ cells or FoxP3+, IFNγ+, and GM-CSF+ CD4+ T cell subsets from the (**c**) the spleen and (**d**) CNS of *Xbp1*^{WT} and *Xbp1*^{Astro} mice.

(**e**) Pathways analyzed by pre-ranked gene set enrichment analysis (GSEA) of genes from RNA-seq data comparing *Xbp1*^{WT} and *Xbp1*^{Astro} mice.



Extended Data Fig. 10 |. Analysis of chromatin accessibility in genes responsive to mineralocorticoid signaling as a function of XBP1 expression in astrocytes.

(**a**–**b**) Re-analysis of ATAC-seq data on bulk flow cytometry-sorted astrocytes that we reported in ref. [4], showing increased chromatin accessibility in NR3C2 responsive genes (**a**) and in *Nr3c2* and *Ncor2* (**b**) as a function of *Gfap*-specific shRNA-driven lentiviral knockdown of *Xbp1*.

Extended Data Table 1 |
FIND-seq practical considerations

(**a**) Table of the molecular reagents required to perform a FIND-seq experiment on 1M cells itemized by cost, part number, and stage of the workflow. (**a**) Cost-benefit analysis of performing FIND-seq relative to 10X Genomics 3′ scRNA-seq. The comparison shown represents the cost associated with sequencing all target cells in a sample at a sequencing

depth of 20k reads per cell. 10X Genomics requires sequencing every cell to find rare target cells whereas FIND-seq sorts and sequences only target cells.

a

Per cell estimates	
55	µm drop size
87.11	pL drop size
65%	agarose packing
0.1	cell loading
1340.21	pL agarose needed per cell
1,000,000	cells pre run
1.3402	mL agarose

Microfluidics	Part Number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Silicone Elastomer Kit	Dow Silicones Cat. 04019862	$334.16	500	g	$ 0.09	1	1	-	90	$ 8.10	
Glass slides (75 × 50 mm)	Corning Cat. 2947-75X50	$53.20	72	slides	$ 0.74	1	1	-	3	$ 2.22	
Master device device 1	Custom fabrication	$400.00	100	mold	$ 4.00	1	1	-	4	$ 16.00	
Master device device 2	Custom fabrication	$400.00	100	mold	$ 4.00	1	1	-	4	$ 16.00	
Master device device 3	Custom fabrication	$400.00	100	mold	$ 4.00	1	1	-	4	$ 16.00	
Bubble trigger Droplet Oil	BioRad Cat. 1863005	$350.00	70	ml	$ 5.00	1	1	-	6.701	$ 33.51	
Re-injection Droplet Generation Oil	BioRad Cat. 1864006	$350.00	70	ml	$ 5.00	1	1	-	5.361	$ 26.80	
											$ 118.63

Agarose functionalization	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Ultra-low melt Agarose IX-A	Sigma-Aldrich Cat. A2576	$94.40	5	g	$ 18.88	10	1.5	%	0.201	$ 0.038	
Allyl agarose	Lucidant Polymers	$349.00	5	g	$ 69.80	10	0.5	%	0.067	$ 0.047	
Ammonium Persulfate	Promega Cat. V3131	$25.20	25	g	$ 1.01	10	0.1	%	0.013	$ 0.000	
TEMED	Invitrogen Cat. 15524-010	$61.50	30	ml	$ 2.05	100	0.1	%	0.001	$ 0.003	
Acrydited oligo dT primer	IDT	$250.00	0.2	umoles	$ 1,250	1000	20	uM	0.027	$ 33.51	
											$ 33.59

Lysis	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Tris-HCl pH 7.5	Teknova Cat. T5075	$56.40	500	ml	$ 0.11	1000	20	mM	0.01	$ 0.002	
Lithium Chloride	Sigma Life Science Cat. L7026	$46.40	100	ml	$ 0.46	8000	1000	mM	0.08	$ 0.039	
Lithium dodecyl sulfate	Sigma Cat. L4632-5G	$53.70	5	g	$ 10.74	10%	1.00%	wt./v	0.07	$ 0.720	
EDTA	VWR Cat. E177	$118.14	500	ml	$ 0.24	500	10	mM	0.01	$ 0.003	
DTT (1M)	Sigma Life Science Cat. 43816	$36.80	10	ml	$ 3.68	1000	10	mM	0.01	$ 0.025	
Proteinase K	NEB P8107S	$87.00	2	ml	$ 43.50	20	2	ug/ul	0.07	$ 2.915	
											$ 3.70

Wash Buffer	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Tris-HCl pH 7.5	Teknova Cat. T5075	$56.40	500	ml	$ 0.11	1000	20	mM	0.27	$ 0.03	
Lithium Chloride	Sigma Life Science Cat. L7026	$46.40	100	ml	$ 0.46	8000	500	mM	0.84	$ 0.39	
Lithium dodecyl sulfate	Sigma Cat. L4632-5G	$53.70	5	g	$ 10.74	10%	0.10%	wt./v	0.13	$ 0.00	
EDTA	VWR Cat. E177	$118.14	500	ml	$ 0.24	500	0.1	mM	0.00	$ 0.00	
											$ 0.42

Wash Buffer	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Tris-HCl pH 7.5	Teknova Cat. T5075	$56.40	500	ml	$ 0.11	1000	20	mM	0.54	$ 0.06	
Sodium Chloride	Invitrogen Cat. AM9759	$94.75	500	ml	$ 0.19	8000	500	mM	1.68	$ 0.32	
											$ 0.38

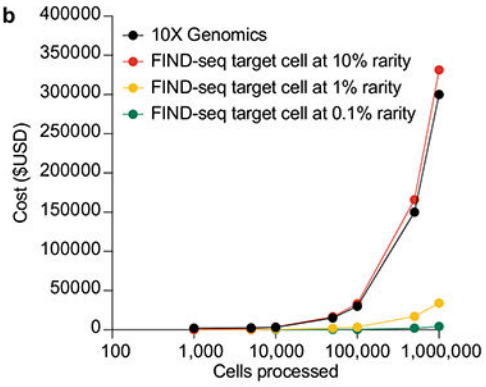
RT Buffer wash	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Tris-HCl pH 8.3	Teknova Cat. T5083	$ 47.00	500	ml	$ 0.09	1000	250	mM	6.70	$ 0.63	
Potassium Chloride	ThermoFisher Cat. AM9640G	$ 49.36	100	ml	$ 0.49	2000	375	mM	5.03	$ 2.48	
MgCl2	ThermoFisher Cat. AM9530G	$ 53.70	50.5	ml	$ 1.06	1000	15	mM	0.40	$ 0.43	
DTT (1M)	Sigma Life Science Cat. 43816	$ 36.80	10	ml	$ 3.68	1000	50	mM	1.34	$ 4.93	
											$ 8.47

RT Buffer	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
Maxima H- Reverse Transcriptase	ThermoFisher Cat. EP0753	$746	40000	U	$ 0.02	200	2	U/ul	0.067	$ 249.95	
Magnesium Chloride	Sigma-Aldrich Cat. M1028-100ML	$67	10	ml	$ 6.71	1000	6	mM	0.040	$ 0.27	
Betaine	Sigma Life Science Cat. 61962	$79	4e5	umoles	$ 0.00	5	1	mM	1.340	$ 0.25	
PEG8k	Sigma-Aldrich Cat. P2139-500G	$79	50	g	$ 1.58	20	7.5%	%	0.025	$ 0.00	
dNTP	Thermo Scientific Cat. R1121	$249	1	mL	$ 249.00	25	1	mM	0.268	$ 66.74	
NxGen RNAse Inhibitor	Lucigen Cat. 30281-2	$224	10000	U	$ 0.02	40	0.5	U/ul	0.084	$ 75.05	
Template switch oligonucleotide	IDT	$200	0.3	umoles	$ 666.67	1000	2	uM	0.013	$ 8.93	
											$ 401.20

ddPCR	Cat number	Cost	Qty	Units	Cost/qty	Stock	Final	Units	Volume	Cost	Total
TaqPATH ProAmp MasterMix	ThermoFisher Cat. A30866	$786	10	ml	$ 78.60	2	1	x	0.670	$ 52.67	
Tween-20	Sigma Life Science Cat. P9416	$25.2	50	ml	$ 0.50	100%	2.5%	%	0.034	$ 0.02	
PEG-6000	Alfa Aesar Cat. A17541	$79.0	50	g	$ 1.58	50%	2.5%	%	0.067	$ 0.11	
Assay 1	Custom	$192.1	4	mL	$ 48.03	20	1	x	0.067	$ 3.22	
Assay 1	Custom	$192.1	4	mL	$ 48.03	20	1	x	0.067	$ 3.22	
											$ 59.23

WTA and library prep	Cat number										
Kapa HIFI	Kapa HiFi Master Mix (2x)	$535.44	6.25	ml	$ 85.67	2	1	x	0.025	$ 2.14	
Nextera XT DNA Kit (24 sample)	FC-131-1024	$769.25	24	rxns	$ 32.05	-	-	4	rxns	$ 128.21	
											$ 130.35
										TOTAL	$ 755.97

b



Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by grants 1U01AI129206, UM1AI126611, NS087867, ES025530, ES032323, AI126880 and AI149699 from the National Institutes of Health, by the Chan Zuckerberg Biohub, by Wellcome Leap, by the NMSS and the Progressive MS Alliance. I.C.C. was supported by the NIH (K22AI152644, DP2AI154435). M.A.W. was supported by the NIH (1K99NS114111, F32NS101790 and 1R01MH130458-01), a training grant from the NIH and Dana-Farber Cancer Institute (T32CA207201), a travelling neuroscience fellowship from the Program in Interdisciplinary Neuroscience at the Brigham and Women's Hospital and the Women's Brain Initiative at the Brigham and Women's Hospital. H.-G.L. was supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1A6A3A14039088). C.M.P. was supported by a fellowship from FAPESP BEPE (2019/13731-0). We thank L. Glimcher for providing *Xbp1*$^{fl/fl}$ mice; all members of the Quintana and Abate laboratories for helpful advice and discussions; and the NeuroTechnology Studio at Brigham and Women's Hospital for providing instrument access. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Data availability

Genomic data have been deposited in the GEO database under accession number GSE198971. All other data are available from the corresponding authors on reasonable request. Correspondence and request for materials should be addressed to F.J.Q. or A.R.A. Source data are provided with this paper.

References

1. Baecher-Allan C, Kaskow BJ & Weiner HL Multiple sclerosis: mechanisms and immunotherapy. Neuron 97, 742–768 (2018). [PubMed: 29470968]
2. Lee H-G, Wheeler MA & Quintana FJ Function and therapeutic value of astrocytes in neurological diseases. Nat. Rev. Drug Discovery 21, 339–358 (2022). [PubMed: 35173313]
3. Linnerbauer M, Wheeler MA & Quintana FJ Astrocyte crosstalk in cns inflammation. Neuron 108, 608–622 (2020). [PubMed: 32898475]
4. Wheeler MA et al. Environmental control of astrocyte pathogenic activities in CNS inflammation. Cell 176, 581–596.e518 (2019). [PubMed: 30661753]
5. Börner K et al. Anatomical structures, cell types and biomarkers of the human reference atlas. Nat. Cell Biol 23, 1117–1128 (2021). [PubMed: 34750582]
6. Ginhoux F, Yalin A, Dutertre CA & Amit I Single-cell immunology: past, present, and future. Immunity 55, 393–404 (2022). [PubMed: 35263567]
7. Rozenblatt-Rosen O et al. The human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. Cell 181, 236–249 (2020). [PubMed: 32302568]
8. Cugurra A et al. Skull and vertebral bone marrow are myeloid cell reservoirs for the meninges and CNS parenchyma. Science 373, eabf7844 (2021). [PubMed: 34083447]
9. Giladi A et al. Cxcl10^{+} monocytes define a pathogenic subset in the central nervous system during autoimmune neuroinflammation. Nat. Immunol 21, 525–534 (2020). [PubMed: 32313246]
10. Grigg JB et al. Antigen-presenting innate lymphoid cells orchestrate neuroinflammation. Nature 600, 707–712 (2021). [PubMed: 34853467]
11. Hiltensperger M et al. Skin and gut imprinted helper T cell subsets exhibit distinct functional phenotypes in central nervous system autoimmunity. Nat. Immunol 22, 880–892 (2021). [PubMed: 34099917]
12. Jordão MJC et al. Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. Science 363, eaat7554 (2019). [PubMed: 30679343]
13. Khakh BS & Deneen B The emerging nature of astrocyte diversity. Annu. Rev. Neurosci 42, 187–207 (2019). [PubMed: 31283899]
14. Sofroniew MV Astrocyte reactivity: subtypes, states, and functions in cns innate immunity. Trends Immunol. 41, 758–770 (2020). [PubMed: 32819810]

15. Absinta M et al. A lymphocyte–microglia–astrocyte axis in chronic active multiple sclerosis. Nature 597, 709–714 (2021). [PubMed: 34497421]
16. Chao CC et al. Metabolic control of astrocyte pathogenic activity via cpla2-mavs. Cell 179, 1483–1498.e1422 (2019). [PubMed: 31813625]
17. Escartin C et al. Reactive astrocyte nomenclature, definitions, and future directions. Nat. Neurosci 24, 312–325 (2021). [PubMed: 33589835]
18. Mayo L et al. Regulation of astrocyte activation by glycolipids drives chronic cns inflammation. Nat. Med 20, 1147–1156 (2014). [PubMed: 25216636]
19. Rothhammer V et al. Microglial control of astrocytes in response to microbial metabolites. Nature 557, 724–728 (2018). [PubMed: 29769726]
20. Rothhammer V et al. Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor. Nat. Med 22, 586–597 (2016). [PubMed: 27158906]
21. Sanmarco LM et al. Gut-licensed IFNγ^+ NK cells drive LAMP1$^+$TRAIL$^+$ anti-inflammatory astrocytes. Nature 590, 473–479 (2021). [PubMed: 33408417]
22. Wheeler MA et al. Mafg-driven astrocytes promote cns inflammation. Nature 578, 593–599 (2020). [PubMed: 32051591]
23. Habib N et al. Disease-associated astrocytes in alzheimer's disease and aging. Nat. Neurosci 23, 701–706 (2020). [PubMed: 32341542]
24. Hasel P, Rose IVL, Sadick JS, Kim RD & Liddelow SA Neuroinflammatory astrocyte subtypes in the mouse brain. Nat. Neurosci 24, 1475–1487 (2021). [PubMed: 34413515]
25. Amamoto R et al. Probe-seq enables transcriptional profiling of specific cell types from heterogeneous tissue by rna-based isolation. eLife 8, e51452 (2019). [PubMed: 31815670]
26. Eastburn DJ, Sciambi A & Abate AR Ultrahigh-throughput mammalian single-cell reverse-transcriptase polymerase chain reaction in microfluidic drops. Anal. Chem 85, 8016–8021 (2013). [PubMed: 23885761]
27. Eastburn DJ, Sciambi A & Abate AR Identification and genetic analysis of cancer cells with pcr-activated cell sorting. Nucleic Acids Res. 42, e128 (2014). [PubMed: 25030902]
28. Calfon M et al. Ire1 couples endoplasmic reticulum load to secretory capacity by processing the *XBP-1* mRNA. Nature 415, 92–96 (2002). [PubMed: 11780124]
29. Clark IC et al. HIV silencing and cell survival signatures in infected T cell reservoirs. Nature 10.1038/s41586-022-05556-6 (2023).
30. Clark IC, Thakur R & Abate AR Concentric electrodes improve microfluidic droplet sorting. Lab Chip 18, 710–713 (2018). [PubMed: 29383336]
31. Smith HL et al. Astrocyte unfolded protein response induces a specific reactivity state that causes non-cell-autonomous neuronal degeneration. Neuron 105, 855–866.e855 (2020). [PubMed: 31924446]
32. Glimcher LH, Lee AH & Iwakoshi NN Xbp-1 and the unfolded protein response (UPR). Nat. Immunol 21, 963–965 (2020). [PubMed: 32616861]
33. Lee AH, Iwakoshi NN & Glimcher LH Xbp-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. Mol. Cell. Biol 23, 7448–7459 (2003). [PubMed: 14559994]
34. Arzalluz-Luque A & Conesa A Single-cell RNAseq for the study of isoforms—how is that possible? Genome Biol. 19, 110 (2018). [PubMed: 30097058]
35. Buen Abad Najar CF, Yosef N & Lareau LF Coverage-dependent bias creates the appearance of binary splicing in single cells. eLife 9, e54603 (2020). [PubMed: 32597758]
36. Picelli S et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods 10, 1096–1098 (2013). [PubMed: 24056875]
37. Macosko EZ et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161, 1202–1214 (2015). [PubMed: 26000488]
38. Clark IC et al. Barcoded viral tracing of single-cell interactions in central nervous system inflammation. Science 372, eabf1230 (2021). [PubMed: 33888612]

39. Glass CK & Saijo K Nuclear receptor transrepression pathways that regulate inflammation in macrophages and t cells. Nat. Rev. Immunol 10, 365–376 (2010). [PubMed: 20414208]
40. Geller DS et al. Activating mineralocorticoid receptor mutation in hypertension exacerbated by pregnancy. Science 289, 119–123 (2000). [PubMed: 10884226]
41. Ruzzo EK et al. Inherited and de novo genetic risk for autism impacts shared networks. Cell 178, 850–866.e826 (2019). [PubMed: 31398340]
42. Hetz C et al. Unfolded protein response transcription factor XBP-1 does not influence prion replication or pathogenesis. Proc. Natl Acad. Sci. USA 105, 757–762 (2008). [PubMed: 18178615]
43. Srinivasan R et al. New transgenic mouse lines for selectively targeting astrocytes and studying calcium signals in astrocyte processes in situ and in vivo. Neuron 92, 1181–1195 (2016). [PubMed: 27939582]
44. Anderson MA et al. Astrocyte scar formation aids central nervous system axon regeneration. Nature 532, 195–200 (2016). [PubMed: 27027288]
45. John Lin CC et al. Identification of diverse astrocyte populations and their malignant analogs. Nat. Neurosci 20, 396–405 (2017). [PubMed: 28166219]
46. Saijo K et al. A Nurr1/CoREST pathway in microglia and astrocytes protects dopaminergic neurons from inflammation-induced death. Cell 137, 47–59 (2009). [PubMed: 19345186]
47. Shaked I et al. Transcription factor Nr4a1 couples sympathetic and inflammatory cues in CNS-recruited macrophages to limit neuroinflammation. Nat. Immunol 16, 1228–1234 (2015). [PubMed: 26523867]
48. Clarisse D, Deng L, de Bosscher K & Lother A Approaches towards tissue-selective pharmacology of the mineralocorticoid receptor. Br. J. Pharmacol 179, 3235–3249 (2021). [PubMed: 34698367]
49. Ayata P et al. Epigenetic regulation of brain region-specific microglia clearance activity. Nat. Neurosci 21, 1049–1060 (2018). [PubMed: 30038282]
50. Wendeln AC et al. Innate immune memory in the brain shapes neurological disease hallmarks. Nature 556, 332–338 (2018). [PubMed: 29643512]
51. Boulay AC et al. Translation in astrocyte distal processes sets molecular heterogeneity at the gliovascular interface. Cell Discov. 3, 17005 (2017). [PubMed: 28377822]
52. Magnusson JP et al. Activation of a neural stem cell transcriptional program in parenchymal astrocytes. eLife 9, e59733 (2020). [PubMed: 32744501]
53. Yan Z, Clark IC & Abate AR Rapid encapsulation of cell and polymer solutions with bubble-triggered droplet generation. Macromol. Chem. Phys 218, 1600297 (2017).
54. Martin M Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 10.14806/ej.17.1.200 (2011).
55. Dobin A et al. Star: Ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
56. Li B & Dewey CN RSEM: accurate transcript quantification from rna-seq data with or without a reference genome. BMC Bioinf. 12, 323 (2011).
57. Soneson C, Love MI & Robinson MD Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res 4, 1521 (2015). [PubMed: 26925227]
58. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014). [PubMed: 25516281]
59. Zhu A, Ibrahim JG & Love MI Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics 35, 2084–2092 (2019). [PubMed: 30395178]
60. Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl Acad. Sci. USA 102, 15545–15550 (2005). [PubMed: 16199517]
61. Kramer A, Green J, Pollard J Jr. & Tugendreich S Causal analysis approaches in ingenuity pathway analysis. Bioinformatics 30, 523–530 (2014). [PubMed: 24336805]
62. Grant CE, Bailey TL & Noble WS FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017–1018 (2011). [PubMed: 21330290]

63. Bailey TL, Johnson J, Grant CE & Noble WS The meme suite. Nucleic Acids Res. 43, W39–W49 (2015). [PubMed: 25953851]

64. Sandelin A, Alkema W, Engstrom P, Wasserman WW & Lenhard B JASPAR: An open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res. 32, D91–D94 (2004). [PubMed: 14681366]

65. Hagemann-Jensen M et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. Nat. Biotechnol 38, 708–714 (2020). [PubMed: 32518404]

66. Bray NL, Pimentel H, Melsted P & Pachter L Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol 34, 525–527 (2016). [PubMed: 27043002]

67. Melsted P, Ntranos V & Pachter L The barcode, UMI, set format and BUStools. Bioinformatics 35, 4472–4473 (2019). [PubMed: 31073610]

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

69. Lee Y, Messing A, Su M & Brenner M *GFAP* promoter elements required for region-specific and astrocyte-specific expression. Glia 56, 481–493 (2008). [PubMed: 18240313]

70. Chen EY et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinf. 14, 128 (2013).

71. Kuleshov MV et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44, W90–W97 (2016). [PubMed: 27141961]

72. Polman CH et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol 69, 292–302 (2011). [PubMed: 21387374]

73. Schindelin J et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012). [PubMed: 22743772]

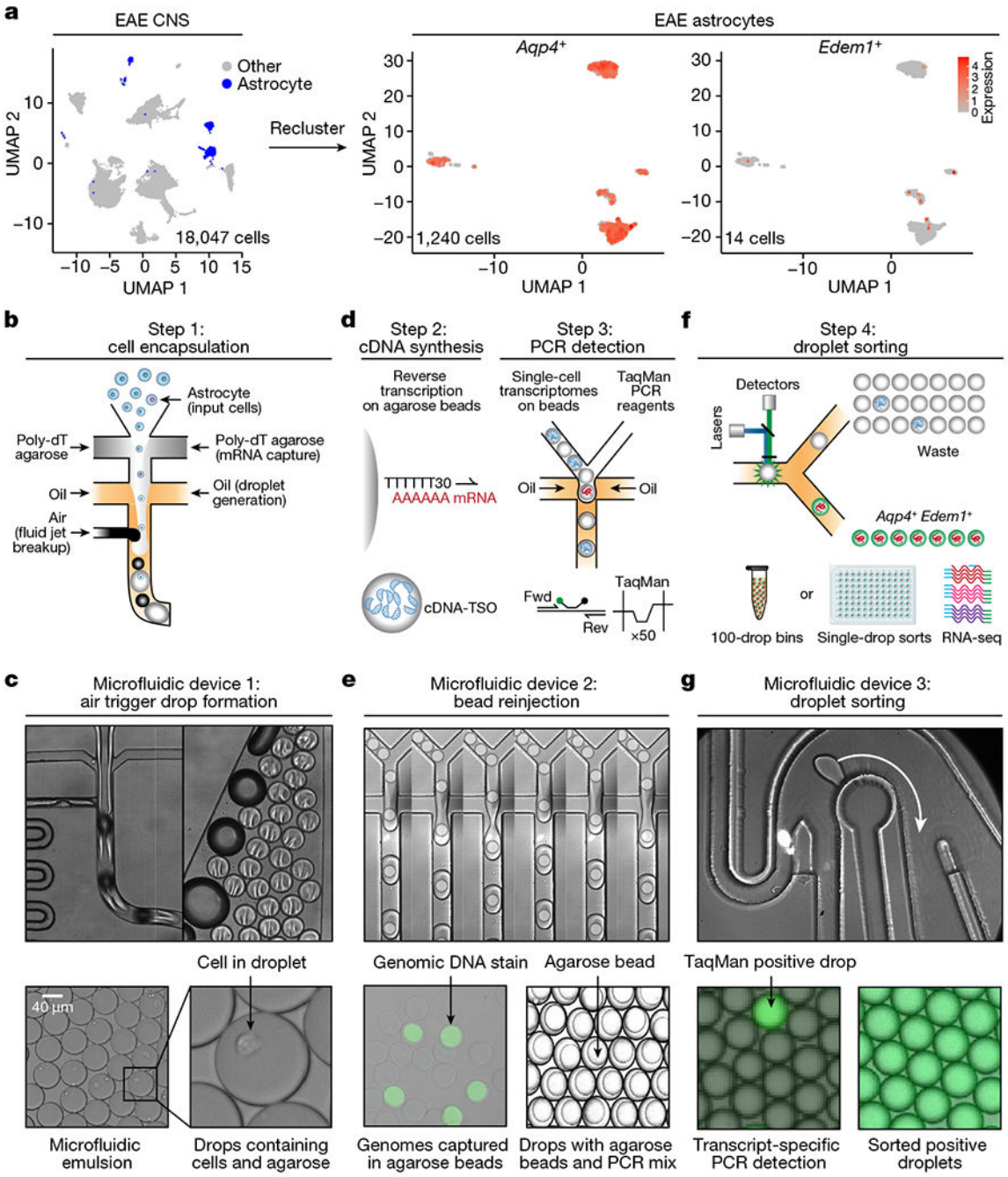


Fig. 1 |. Development of FIND-seq to study rare astrocyte subsets.

a, Left, clustering of scRNA-seq data from the CNS of EAE mice (re-analysed from ref. [22]). Right, extraction and re-clustering of EAE astrocytes showing *Aqp4* and *Edem1* expression superimposed on all clusters. **b**, Schematic of air-triggered droplet generation of molten agarose functionalized with polyT oligonucleotides (step 1). **c**, Micrographs of cells encapsulated in molten agarose droplets. **d**, Schematic of reverse transcription of cDNA on agarose beads (left, step 2) followed by agarose bead re-injection for digital PCR detection (right, step 3). **e**, Bead re-injection. Bottom left, micrographs of agarose hydrogel beads after

solidification and removal from oil. Genomes are visualized with SYBR staining. Bottom right, agarose beads re-encapsulated in oil. Droplets maintain single-cell resolution during PCR detection. **f**, Schematic of dielectrophoretic sorting of droplets into either 100-cell bins or 96-well plates for single-cell analysis (step 4). **g**, Micrographs of single-cell droplet PCR detection before sorting (bottom left) and after sorting (bottom right). Data are mean ± s.e.m.

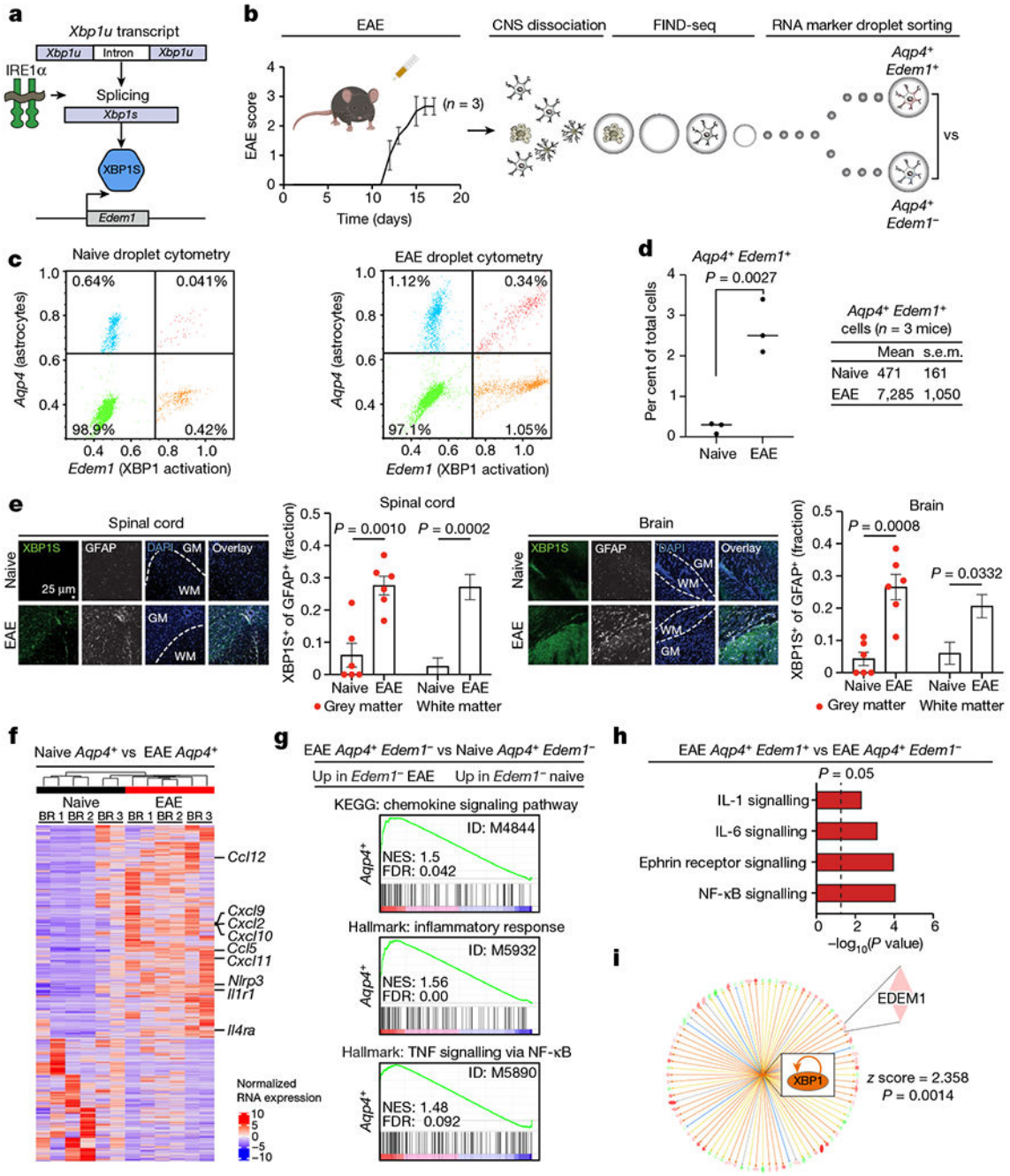


Fig. 2 |. FIND-seq analysis of XBP1-driven astrocytes.

a, Schematic showing the generation of functional XBP1S and *Edem1* expression following *Xbp1u* mRNA splicing. **b**, Naive and EAE mice were analysed by FIND-seq using *Aqp4* and *Edem1* expression as markers of astrocytes and *Xbp1* mRNA splicing, respectively. **c**, Representative droplet cytometry plots. Numbers in each quadrant are the percentage of positive drops. Top quadrants contain $Aqp4^+$ cells, and the top-right quadrant contains $Aqp4^+Edem1^+$ cells. **d**, Percentage and absolute numbers of $Aqp4^+Edem1^+$ cells in naive and EAE mice. $n = 3$ mice per group. Unpaired two-sided *t*-test. **e**, Immunostaining analysis

of XBP1S^{+} GFAP^{+} astrocytes in the grey and white matter of the brain and spinal cord in naive and EAE mice. $n = 6$ images per group from $n = 3$ mice. Unpaired two-sided *t*-test. **f**, Heat map of genes that are differentially expressed between naive and EAE *Aqp4*$^{+}$ astrocytes (defined by $P < 0.05$ and $|[\log_2(\text{fold change})]| > 2$). BR, biological replicate. Two technical replicates were analysed for each mouse in each condition. **g**, Pathways analysed by pre-ranked gene set enrichment analysis (GSEA) of genes from RNA-seq data comparing EAE *Aqp4*$^{+}$*Edem1*$^{-}$ with naive *Aqp4*$^{+}$*Edem1*$^{-}$ cells. **h**, Canonical inflammatory pathways in EAE *Aqp4*$^{+}$*Edem1*$^{-}$ cells relative to EAE *Aqp4*$^{+}$*Edem1*$^{-}$ cells. **i**, IPA analysis of FIND-seq data comparing *Aqp4*$^{+}$*Edem1*$^{+}$ to *Aqp4*$^{+}$*Edem1*$^{-}$ cells during EAE showing prediction of XBP1 as an upstream regulator in EAE *Aqp4*$^{+}$*Edem1*$^{+}$ cells. This analysis is performed by computing the overlap in the differentially expressed genes of the FIND-seq dataset with the genes regulated by XBP1 by Fisher's exact test. Data are mean ± s.e.m.

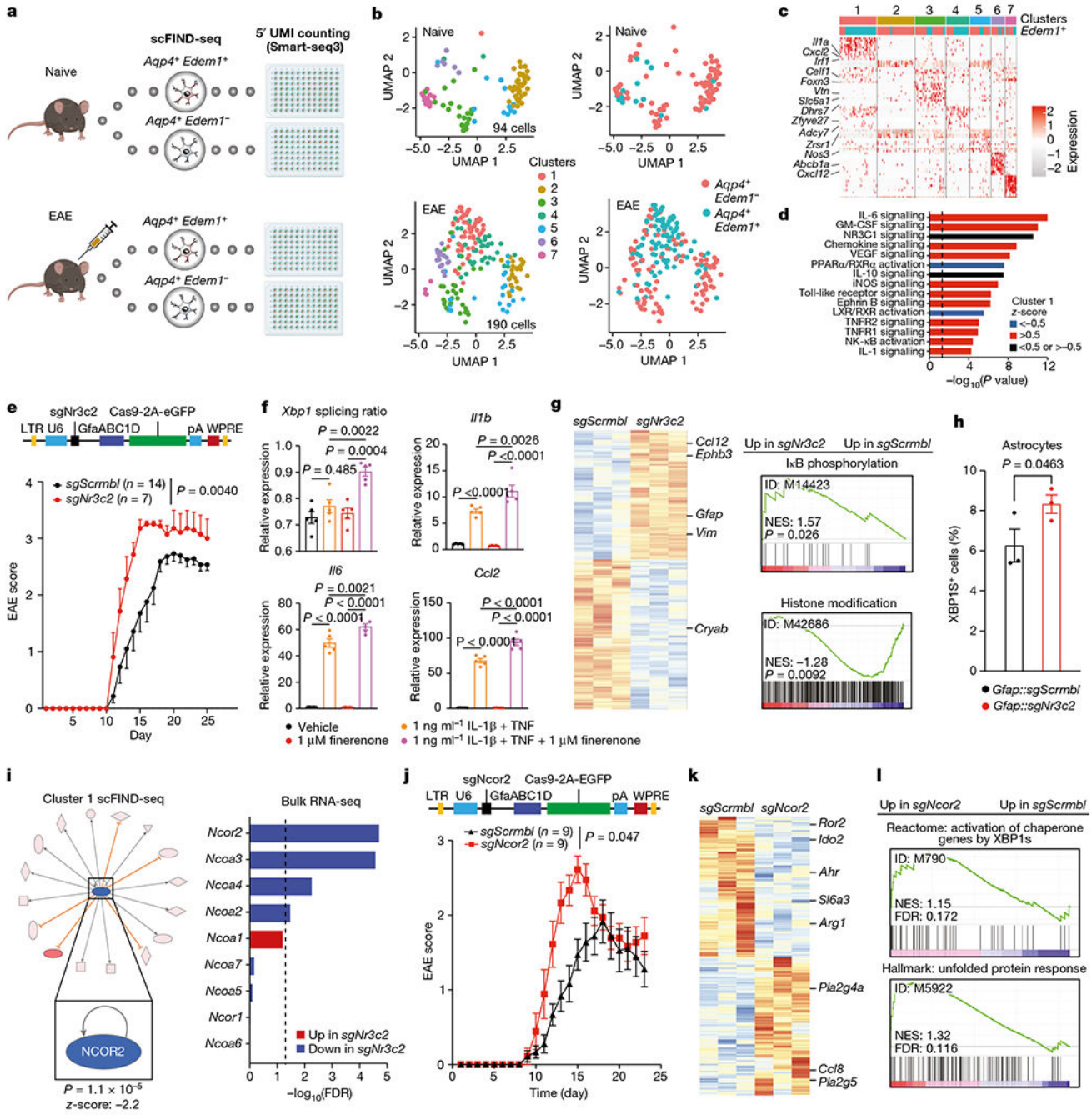


Fig. 3 |. NR3C2–NCOR2 signalling limits disease-promoting astrocyte responses.

a, Schematic of scFIND-seq applied to *Aqp4*$^{+}$*Edem1*$^{+}$ astrocytes in EAE. **b**, Uniform manifold approximation and projection (UMAP) plots of *Aqp4*$^{+}$*Edem1*$^{+}$ and *Aqp4*$^{+}$*Edem1*$^{-}$ astrocytes from naive and EAE mice analysed by scFIND-seq. **c**, Analysis of marker genes in *Aqp4*$^{+}$*Edem1*$^{+}$ or *Aqp4*$^{+}$*Edem1*$^{-}$ astrocytes from naive and EAE mice. **d**, Pathway analysis of cluster 1 astrocytes analysed by scFIND-seq using IPA. Fisher's exact test. **e**, In vivo knockdown of *Nr3c2* in astrocytes. Top, schematic of the lentiviral vector containing sgRNA targeting *Nr3c2* and *Streptococcus pyogenes Cas9* under the control of a *Gfap*

promoter. Bottom, EAE disease progression in mice transduced with *sgScrmbl* ($n = 14$) or *sgNr3c2* ($n = 7$) lentiviruses. Two-way repeated measures ANOVA. **f**, Quantitative PCR with reverse transcription *Il1b*, *Il6* and *Ccl2* after 24 h and *Xbp1* after 4 h in primary mouse astrocytes treated with or without 1 ng ml^{-1} IL-1β and TNF in the presence of 1 μM finerenone or vehicle. $n = 5$ per group. Unpaired two-sided *t*-test. **g**, Left, heat map of RNA-seq data from of astrocytes isolated from *sgScrmbl* and *sgNr3c2* mice. $n = 3$ per group. Right, GSEA pre-ranked analysis of RNA-seq data comparing astrocytes isolated from *sgNr3c2* versus *sgScrmbl* mice. **h**, FACS analysis of spliced XBP1 ($XBP1S^{+}$) in astrocytes from *sgScrmbl* and *sgNr3c2* mice ($n = 3$ per group). Unpaired two-sided *t*-test. **i**, Left, prediction of decreased NCOR2 activation by IPA in cluster 1 scFIND-seq data compiled from all cells. Fisher's exact test. Right, differential gene expression of candidate co-regulators of NR3C2 in *sgNr3c2* versus *sgScrmbl* astrocytes showing *Ncor2* as the top candidate. **j**, Analysis of EAE mice transduced with lentivirus co-expressing *Gfap::Cas9* and *sgNcor2* ($n = 9$) or *sgScrmbl* ($n = 9$). Top, schematic of the lentiviral vector containing sgRNA targeting *Ncor2* and *S. pyogenes Cas9* under the control of a *Gfap* promoter. Bottom, disease progression in *sgScrmbl* versus *sgNcor2* mice. Two-way repeated measures ANOVA. **k**, Heat map of RNA-seq data from astrocytes in *sgScrmbl* versus *sgNcor2* mice. $n = 3$ per group. **l**, GSEA pre-ranked analysis of scRNA-seq data in astrocytes from *sgScrmbl* versus *sgNcor2* mice. Data are mean ± s.e.m.

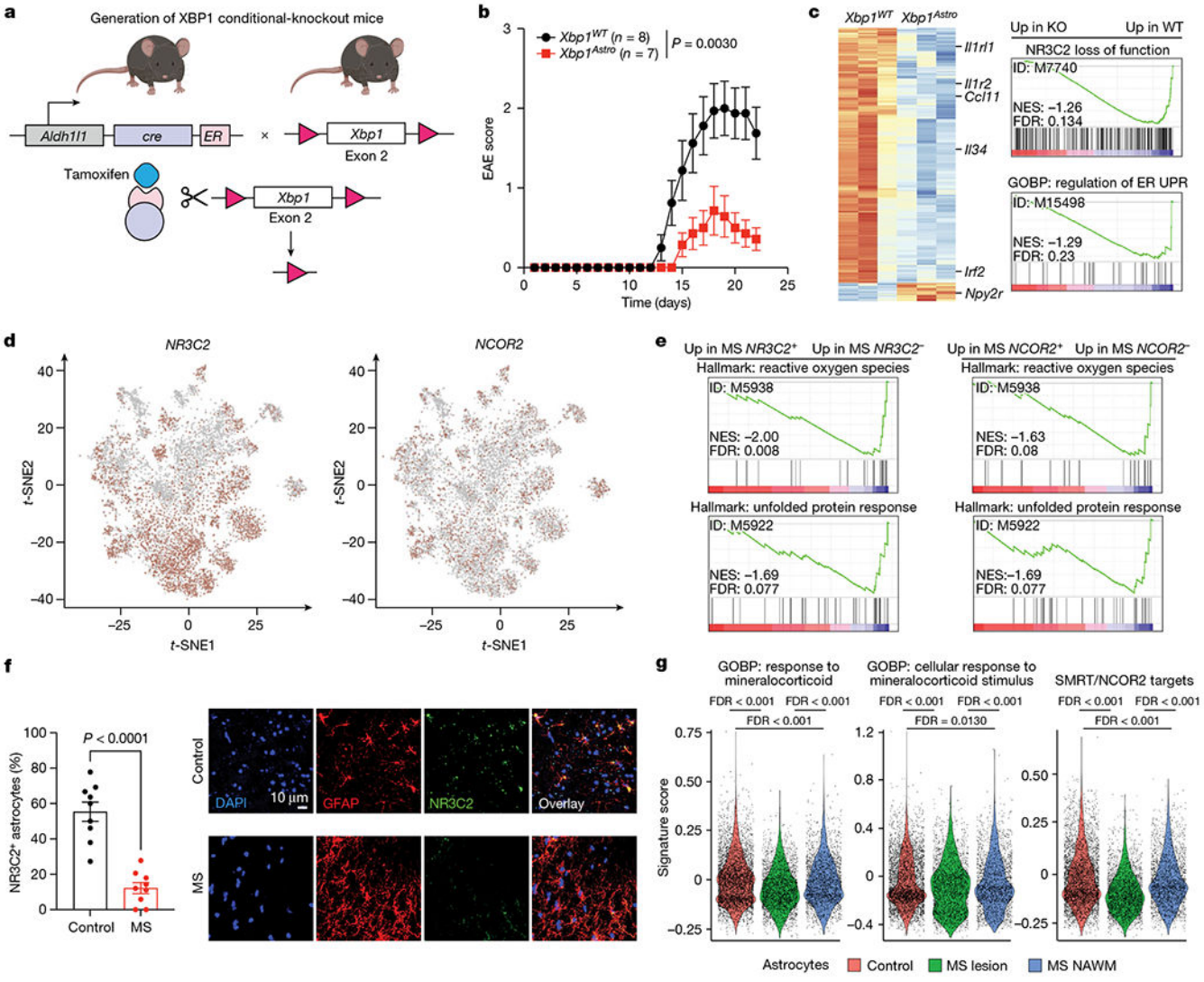


Fig. 4 |. XBP1 limits NR3C2–NCOR2 signalling in EAE and multiple sclerosis.

a, Mice with a floxed *Xbp1* allele were crossed to mice expressing *creERT2* under control of the *Aldh1l1* promoter to generate tamoxifen-inducible *Xbp1* astrocyte (*Xbp1*^{Astro}) conditional-knockout mice. **b**, EAE disease progression in wild-type *Xbp1* (*Xbp1*^{WT}) (n = 8) and *Xbp1*^{Astro} knockout (n = 7) mice. **c**, Analysis of RNA-seq data comparing astrocytes isolated from *Xbp1*^{WT} and *Xbp1*^{Astro} mice. n = 3 per group. ER UPR, endoplasmic reticulum unfolded protein response; GOBP, Gene Ontology biological process; KO, knockout. **d**, Re-analysis of scRNA-seq datasets from patients with multiple sclerosis (previously reported in ref. [22]), showing expression of *NR3C2* and *NCOR2* in human astrocytes. **e**, GSEA pre-ranked analysis comparing *NCOR2*$^{+}$ astrocytes to *NCOR2*$^{-}$ astrocytes and *NR3C2*$^{+}$ astrocytes to *NR3C2*$^{-}$ astrocytes in patients with multiple sclerosis. **f**, Quantification of NR3C2^{+} astrocytes identified by immunostaining of tissue from patients with multiple sclerosis. n = 3 patients and n = 9 images per group. Unpaired two-sided t-test. **g**, Violin plots depicting the signature score calculated for the indicated gene set in

astrocytes isolated from controls or patients with multiple sclerosis, derived from data in **d**. Data are mean ± s.e.m.