



Joint single-cell measurements of nuclear proteins and RNA in vivo

Hattie Chung¹✉, Christopher N. Parkhurst^{2,8}, Emma M. Magee^{1,8}, Devan Phillips^{1,7}, Ehsan Habibi¹, Fei Chen^{1,3}, Bertrand Z. Yeung⁴, Julia Waldman¹, David Artis^{2,5} and Aviv Regev^{1,6,7}✉

Identifying gene-regulatory targets of nuclear proteins in tissues is a challenge. Here we describe intranuclear cellular indexing of transcriptomes and epitopes (inCITE-seq), a scalable method that measures multiplexed intranuclear protein levels and the transcriptome in parallel across thousands of nuclei, enabling joint analysis of transcription factor (TF) levels and gene expression in vivo. We apply inCITE-seq to characterize cell state-related changes upon pharmacological induction of neuronal activity in the mouse brain. Modeling gene expression as a linear combination of quantitative protein levels revealed genome-wide associations of each TF and recovered known gene targets. TF-associated genes were coexpressed as distinct modules that each reflected positive or negative TF levels, showing that our approach can disentangle relative putative contributions of TFs to gene expression and add interpretability to inferred gene networks. inCITE-seq can illuminate how combinations of nuclear proteins shape gene expression in native tissue contexts, with direct applications to solid or frozen tissues and clinical specimens.

Single-nucleus (sn)RNA-seq is an essential tool for profiling the heterogeneity of solid tissues, particularly those that are frozen or are challenging to dissociate, or for samples that require preservation of cellular activity states by avoiding nonspecific activation induced by enzymatic cellular dissociation^{1–7}. The nucleus is also a key site of gene regulation by a wide array of proteins, the presence and levels of which shape gene expression. For example, nuclear concentrations of TFs are mechanistic determinants of gene expression that influence the dynamics of TF target binding^{8,9}. Simultaneously measuring quantitative protein levels and the transcriptome inside individual nuclei would enable integrating rich phenotypic and genomic information in tissues and leverage nuclear localization information of proteins. In addition, cellular proteins also provide stable, time-integrated information over mRNA, which is often rapidly degraded^{10–12}.

Methods that use DNA-conjugated antibodies to jointly measure surface protein levels and RNA at single-cell resolution, such as cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)¹³ and the RNA expression and protein sequencing assay (REAP-seq)¹⁴, with recent adaptations to cytoplasmic protein targets^{15–19}, have been applied to circulating immune cells^{20,21}. However, these methods are less suited for non-immune cells and solid tissues where dissociation disrupts the integrity of cellular membranes. While single-nucleus methods are preferred for profiling solid tissues, it remains a challenge to quantitatively measure protein levels with the transcriptome in individual nuclei, as DNA-conjugated antibodies are ‘sticky’ inside the nucleus due to ubiquitous nonspecific binding²².

Aberrations in nuclear levels of specific TFs can be hallmarks of disease and can even be used to predict patient outcomes^{23–25}. Simultaneously measuring nuclear proteins and the transcriptome

in single cells would enable relating levels of nuclear proteins and newly transcribed RNA to reveal genes and pathways involved in cell state changes^{2,10–12,26–28} and how gene networks regulated by TFs vary across contexts and in disease²⁹. Furthermore, nuclear localization of TFs can indicate a change in cellular activity states, such as in the case of TFs that shuttle in and out of the nucleus in response to external signals^{30,31}. The activity-regulated TF complexes nuclear factor (NF)- κ B and activator protein (AP)-1 and their components p65 and c-Fos transiently localize to the nucleus downstream of signal transduction, where they regulate diverse pathways related to inflammation, oncogenesis, apoptosis, cell proliferation and synaptic remodeling^{30,32–36}. Single-nucleus profiling is best suited for studying these activity-regulated TFs because it minimizes spurious expression of these pathways due to technical artifacts such as cellular dissociation⁴.

Current studies that monitor nuclear TF levels and gene expression typically rely on live cell imaging, in situ measurement of protein and RNA levels in tissue by staining and hybridization, or cell sorting and profiling based on fluorescent reporters. Such studies have shown that nuclear localization can vary between individual cells stimulated together due to asynchrony in responses and dynamic shuttling^{37,38}. However, methods relying on reporters or a handful of probes are limited in their ability to relate changes in protein localization to their genome-wide impacts on transcription.

Here, we report inCITE-seq, a method that enables multiplexed and quantitative intranuclear protein measurements using DNA-conjugated antibodies coupled with RNA-seq on a droplet-based profiling platform (Fig. 1a). To allow antibody diffusion across the nuclear membrane, nuclei are lightly fixed with formaldehyde, permeabilized then blocked under optimized

¹Klarman Cell Observatory, Broad Institute of Harvard and MIT, Cambridge, MA, USA. ²Jill Roberts Institute for Research in Inflammatory Bowel Disease, Weill Cornell Medicine, Cornell University, New York, NY, USA. ³Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. ⁴BioLegend Inc, San Diego, CA, USA. ⁵Friedman Center for Nutrition and Inflammation, Joan and Sanford I. Weill Department of Medicine, Department of Microbiology and Immunology, Weill Cornell Medicine, Cornell University, New York, NY, USA. ⁶Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁷Present address: Genentech, South San Francisco, CA, USA. ⁸These authors contributed equally: Christopher N. Parkhurst, Emma M. Magee. ✉e-mail: hchung@broadinstitute.org; aviv.regev.sc@gmail.com

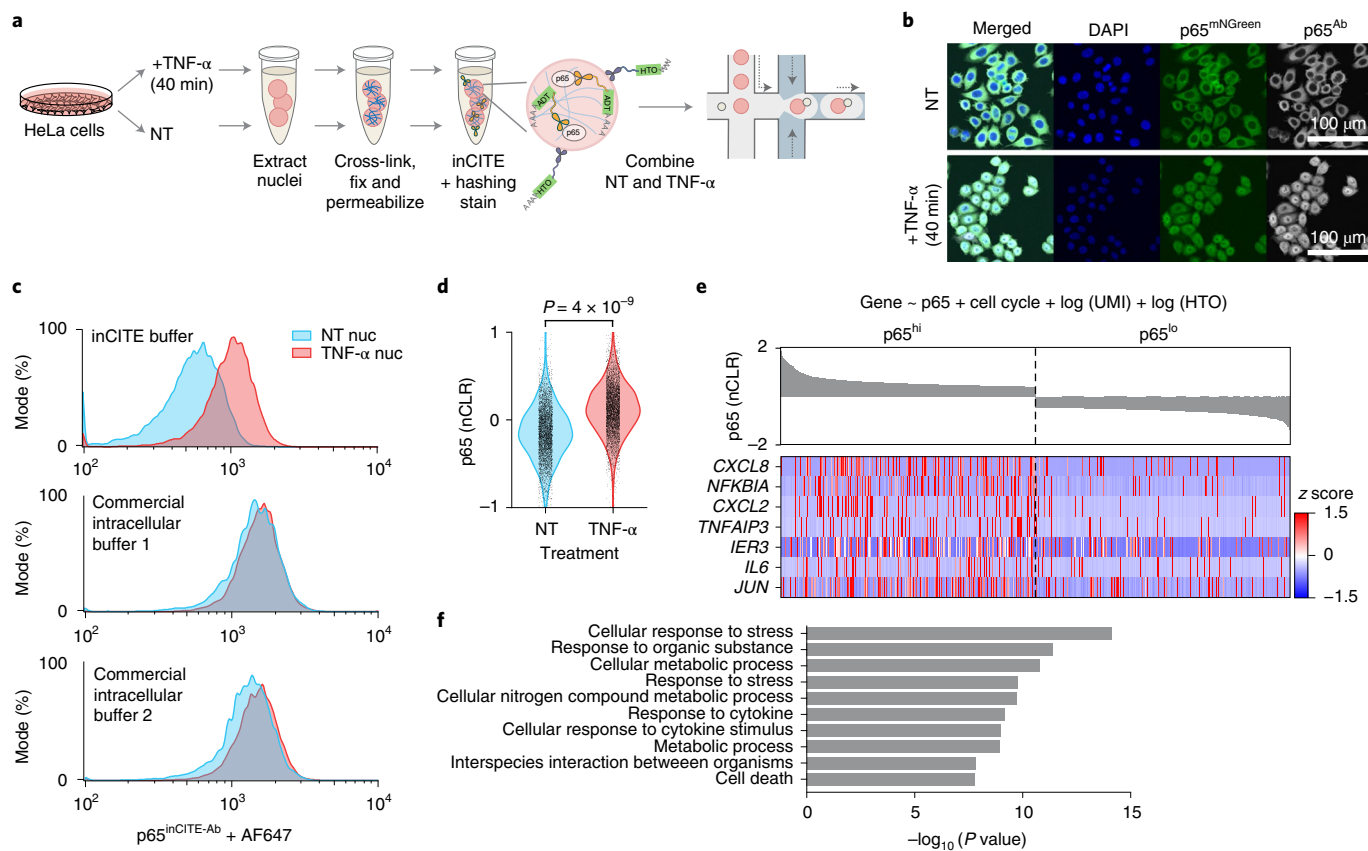


Fig. 1 | inCITE-seq simultaneously measures intranuclear protein and RNA levels at single-nucleus resolution. a, Overview of inCITE-seq for droplet-based profiling of nuclear proteins with nucleus hashing in HeLa cells. **b**, In situ fluorescent images of HeLa cells expressing a p65-mNeonGreen reporter ($p65^{mNGreen}$) stained with p65^{Ab} (followed by Alexa Fluor 657-conjugated secondary antibody), sampled without treatment (NT, top) or 40 min after TNF- α treatment (bottom), representative of four independently conducted experiments. Scale bar, 100 μ m. **c**, Flow cytometry of HeLa nuclei stained with p65^{inCITE-Ab} followed by Alexa Fluor 647 secondary antibody (x axis) sampled from NT (blue) or 40 min after TNF- α treatment (red). Buffers, from top to bottom: optimized inCITE buffer with dextran sulfate, commercial buffer 1, commercial buffer 2 (Methods). **d**, Distribution of p65 levels (nCLR, y axis) in NT (blue) and TNF- α -treated (red) nuclei profiled by inCITE-seq ($P = 4 \times 10^{-9}$, two-sided KS test). **e**, Expression (z score, colored heatmap) of the top seven genes (rows) positively associated with p65 levels identified by a linear model (top, Methods) across nuclei (columns), visualized for the top decile ($p65^{hi}$) and bottom decile ($p65^{lo}$) of p65 nuclear protein levels by inCITE-seq (bar plot, top, nCLR). **f**, Top ten gene ontology terms (y axis) significantly enriched ($-\log_{10}(P \text{ value})$, x axis, hypergeometric test) in 142 genes positively associated with p65 levels.

conditions (blocking buffer, Methods) to minimize nonspecific binding of DNA-conjugated antibodies inside the nucleus. Nucleus-hashing antibodies are simultaneously added to each sample for multiplexing³⁹ (Methods). We then sequence antibody-stained nuclei with droplet-based snRNA-seq for simultaneous capture of antibody DNA tags and the transcriptome. We demonstrate the utility of inCITE-seq for profiling the response to environmental stimuli in cells and tissues, first in a HeLa cell line responding to cytokine stimulation and then in the mouse brain after pharmacological treatment.

Results

inCITE-seq detects nuclear translocation of a TF induced by an extracellular signal. We first developed inCITE-seq to detect cell state changes in HeLa cells, defined as elevated nuclear levels of a TF that translocates into the nucleus in response to an external stimulus. We used a HeLa cell line expressing a p65-mNeonGreen reporter construct^{40,41}, in which p65 localizes to the cytoplasm in untreated cells and translocates to the nucleus upon stimulation with tumor necrosis factor (TNF)- α (Fig. 1b and Methods). At peak nuclear translocation (~40 min after stimulation⁴¹), total p65 levels in the whole cell are constant, with no discernible difference

between untreated and TNF- α -stimulated cells in mNeonGreen signal measured by flow cytometry, but nuclear p65 levels are highly elevated (Extended Data Fig. 1a,b).

To resolve p65 TF levels in the nuclei of untreated versus TNF- α -stimulated cells, we first optimized intranuclear antibody staining conditions using unconjugated anti-p65 antibody (p65^{Ab}) such that flow cytometry detected a clear signal separation that was in agreement with the mNeonGreen reporter signal (Fig. 1c, Extended Data Fig. 1d and Methods). We chose initial fixation and permeabilization parameters based on prior methods⁴² and validated that antibodies successfully diffused across the nuclear membrane by imaging smears of nuclei stained in suspension (Extended Data Fig. 1c). However, the same staining conditions for the DNA-conjugated p65 antibody (p65^{inCITE-Ab}) could not resolve p65 signal between nuclei with no treatment (NT) and TNF- α treatment (Extended Data Fig. 1e), underscoring challenges of nonspecific binding of DNA-conjugated antibodies in the nucleus that occur even in situ²². Adding dextran sulfate to the blocking and staining buffers improved signal separation^{14,22,43}, clearly resolving NT and TNF- α -treated populations, even when compared to two commercial intracellular staining buffers used for cytoplasmic targets (Fig. 1c and Methods).

We used inCITE-seq to profile 10,014 single nuclei from NT and TNF- α -treated HeLa cells that were stained with p65^{inCITE-Ab} (without sorting) and barcoded for multiplexing with nucleus hashing³⁹ (Methods). Antibody levels estimated as counts of antibody-derived tags (ADTs)¹³ were normalized by counts of the nucleus hashtag (hashtag oligonucleotides (HTOs)) to yield nuclear ADT (nADT) units in order to account for differences in poly-dT capture on beads (Extended Data Fig. 1f), which were then log scaled to nuclear centered log ratios (nCLRs)^{13,44} (Methods). Sequencing-derived levels of nuclear p65 differed significantly across NT and TNF- α -treated populations (Fig. 1d; $P=4\times 10^{-9}$, two-sided Kolmogorov–Smirnov (KS) test), confirming that quantitative protein detection by inCITE-seq can distinguish altered cell state due to treatment. The quality of the associated snRNA-seq profiles was comparable to that of snRNA-seq data from the human-derived HEK cell line⁴⁵, as reflected by the number of unique transcripts (UMIs) and genes (Extended Data Fig. 1g; median number of UMIs and genes, 2,655 and 1,158 for inCITE-seq; 1,159 and 812 for snRNA-seq of HEK cells).

Relating genome-wide expression to p65 protein levels. To identify genes for which RNA expression was associated with p65 levels, we used a linear model to fit each gene's expression as a function of continuous p65 levels, while controlling for cell cycle and technical variates (UMIs and HTOs), which identified 142 genes positively associated with p65 levels (false discovery rate (FDR) of 1%; Methods). These genes included well-known NF- κ B targets *CXCL8*, *NFKBIA* and *TNFAIP3* (Fig. 1e), and were enriched for pathways such as cytokine response (Fig. 1f; $P<3\times 10^{-9}$, two-sided hypergeometric test, Methods).

Notably, levels of p65 protein and its encoding transcript *RELA* did not correlate well (Pearson $r^2=0.0008$, $P=0.004$; Extended Data Fig. 1h). This was expected at our stimulation time scale of 40 min, as *RELA* expression peaks ~4 h after induction of the NF- κ B pathway by lipopolysaccharide but is largely unchanged at 40 min, in contrast to other p65 target genes such as *NFKBIA*¹¹ (Extended Data Fig. 1i). *RELA* and p65 levels were also uncorrelated in untreated cells at steady state (NT, $r^2=0.001$, $P=0.032$), which confirmed known differences in mRNA and protein kinetics for this gene. At baseline, *RELA* transcripts are reported to be produced at 12 copies per minute and degraded at 102.4 copies per minute¹⁰, while p65 is translated at 0.9 proteins per mRNA per minute and degraded at 4×10^{-4} proteins per minute¹¹. Our own observation supports prior studies showing fast mRNA degradation but slow protein-degradation kinetics, which further underscores the importance of obtaining protein measurements. In sum, we show that inCITE-seq accurately quantifies nuclear protein and RNA levels that can be integrated to identify putative targets of a TF.

inCITE-seq profiling of the mouse brain after in vivo induction of seizure. Next, we turned to an in vivo setting and applied inCITE-seq to characterize the mouse hippocampus during the early response to seizure, which involves neuroinflammation and oxidative stress, affecting multiple cell types and pathways⁴⁶. Seizures were induced by treatment with kainic acid (KA), a glutamatergic agonist used in models of temporal lobe epilepsy⁴⁷. As the complexes NF- κ B and AP-1 are involved in neuroinflammation, synaptic remodeling and signal transduction of glutamate receptors^{48–50}, we characterized how their components p65 and c-Fos are linked to gene expression in early response to seizure. In addition, we added two nuclear cell type markers, the pan-neuronal marker and regulator of alternative splicing NeuN⁵¹ and the TF and microglial lineage marker PU.1 (ref. ⁵²). Altogether, we used inCITE-seq to profile single nuclei from the hippocampus 2 h after KA treatment with multiplexed measurements of the proteins p65, c-Fos, NeuN and PU.1 (Fig. 2a and Methods).

Antibody validation and optimization for inCITE-seq in the mouse brain. inCITE conjugated antibodies were first validated by flow cytometry on stained nuclei that were extracted from frozen mouse hippocampus. Cell type markers NeuN and PU.1 labeled subpopulations at the expected proportions (58.3% and 2.9% of all nuclei, respectively), with PU.1 specific to microglia (PU.1^{hi} in CD11b^{hi}CX3CR1^{hi} populations but not in CD4^{hi} cells; Extended Data Fig. 2a–c). Levels of p65 were bimodal, and c-Fos^{hi} nuclear subsets were elevated from 0.21% to 48.7% in KA-treated samples as anticipated from neuronal activity (Extended Data Fig. 2d–f). Antibody signals varied across a wide range of concentrations, underscoring the importance of choosing an appropriate concentration regime⁵³ (Extended Data Fig. 3).

Antibodies suitable for inCITE-seq should detect epitopes in frozen tissue after minimal fixation. We therefore expect antibodies that are optimized for flow cytometry, immunocytochemistry or even immunoprecipitation to be compatible with inCITE-seq, while those optimized for heavily fixed tissues (for example, formalin-fixed paraffin-embedded tissue) may not be suitable. As an example, we compared two different antibodies targeting each of NeuN, PU.1 and c-Fos in tissues that were frozen immediately versus tissues after overnight fixation in 4% paraformaldehyde (PFA), which revealed a stark contrast in epitope detection (Extended Data Fig. 4); for example, one version of a c-Fos-specific antibody was unable to detect epitopes in frozen tissue, while another version used for inCITE-seq exhibited clean epitope detection in frozen tissue but not in tissue fixed overnight. To validate antibodies for inCITE-seq, we recommend conducting flow cytometry of antibody stains on nuclei isolated from frozen tissues of interest and in situ immunofluorescence of frozen sections (post-fixed) to help determine cell types and features that express protein targets (Extended Data Figs. 3 and 4).

RNA profiles from inCITE-seq reveal key cell subsets of the mouse hippocampus. Profiling 24,444 nuclei from control (PBS) and KA-treated mice with inCITE-seq yielded RNA profiles that captured all major cell types of the hippocampus. Compared to standard snRNA-seq of the mouse hippocampus from a matched experiment and from another study using the same nuclear extraction protocol⁵⁴, RNA profiles from inCITE-seq were reduced in quality, with a 6.2-fold and 4.8-fold reduction in the median UMI and gene counts, respectively (Extended Data Fig. 5a). Despite the loss in complexity, unsupervised clustering using snRNA-seq profiles from inCITE-seq alone still discerned major cell types of the hippocampus by post hoc annotation with known cell type markers^{54,55} (after addressing ambient RNA⁵⁶, batch correction, and regressing out treatment; Extended Data Fig. 5b,c and Methods).

Cell cluster separation was substantially improved by jointly embedding single-nucleus RNA profiles from inCITE-seq and snRNA-seq, including data from a published study⁵⁴. Unsupervised clustering with variable genes identified jointly across 37,767 high-quality nuclei (22,260 nuclei from inCITE-seq and 15,507 nuclei from snRNA-seq) showed 16 well-delineated clusters, each with contributions from both assays, with robust mixing across batches and treatment, and well annotated post hoc using known cell type markers (Fig. 2b,c, Extended Data Fig. 5d–j and Methods).

Protein levels match cell type-specific and condition-specific expression in RNA-based clusters. Nuclear protein levels measured by inCITE-seq differed across RNA-defined cell types as expected. As before, protein counts were normalized by nuclear HTO counts and then scaled to centered log ratios within each batch (that is, using batch-specific geometric means) to account for systematic batch differences (Extended Data Fig. 6). NeuN levels were elevated in neuronal clusters as expected (Fig. 2d,e; $P=0.005$, two-sided KS test); ambient NeuN levels in subpopulations of other clusters may be due to high expression of NeuN in this tissue or

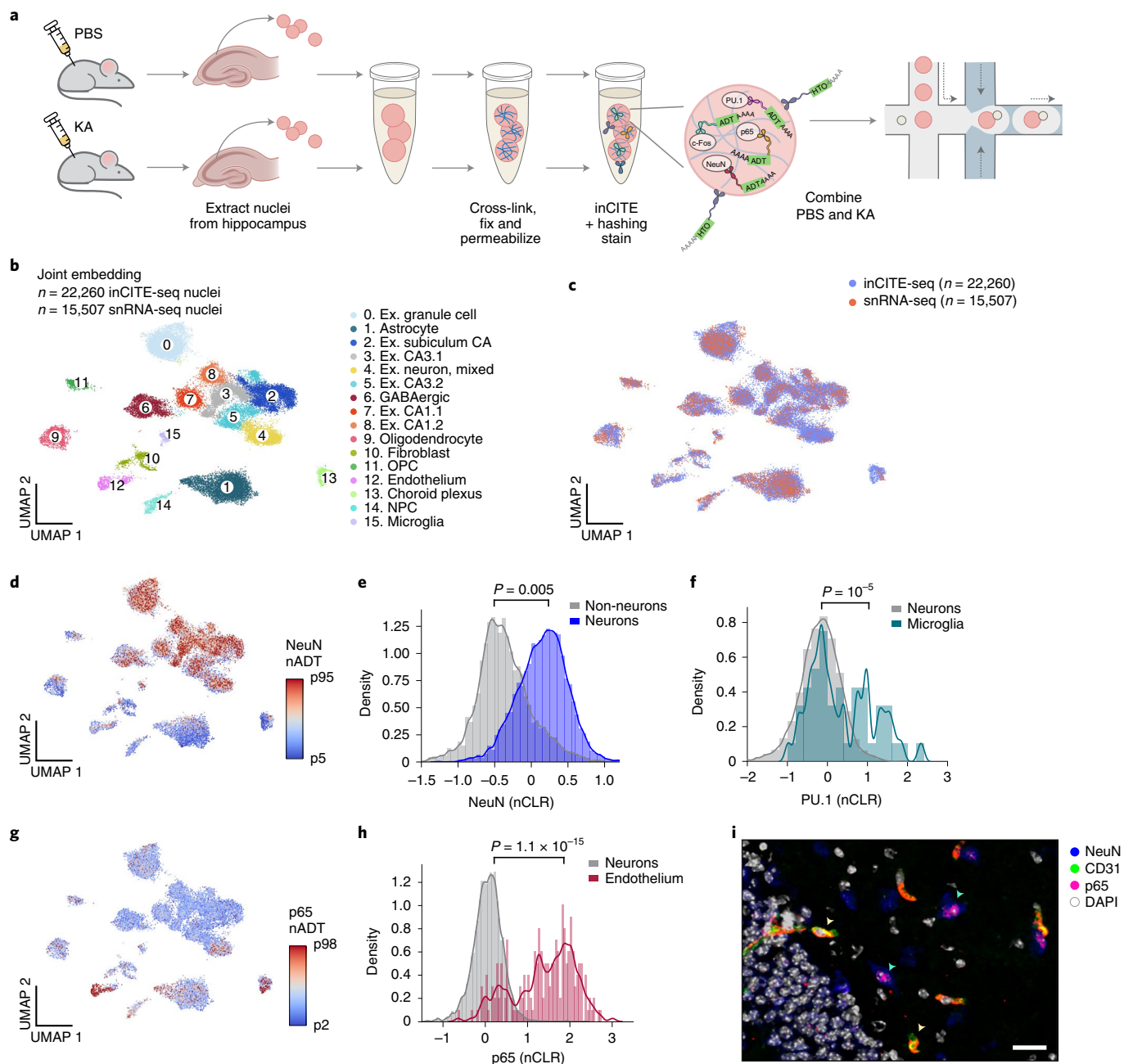


Fig. 2 | In vivo application of inCITE-seq shows cell type-specific protein expression in the mouse hippocampus. **a**, inCITE-seq of the mouse hippocampus after KA or PBS (control) treatment, with nucleus hashing. **b**, Cell types from the adult mouse hippocampus identified by joint embedding of inCITE-seq and snRNA-seq. Uniform manifold approximation and projection (UMAP) embedding of single-nucleus RNA profiles from two batches of inCITE-seq ($n = 22,260$) and two snRNA-seq experiments (this study and Habib et al.⁵⁴, $n = 15,507$) of the mouse hippocampus after regressing out treatment and batch (Methods), colored by cluster and annotated post hoc (color legend). Ex., excitatory EX neuron clusters; OPC, oligodendrocyte precursor cells; NPC, neuronal precursor cells. **c**, Integration of inCITE-seq and snRNA-seq profiles. UMAP embedding as in **b**, colored by assay type (inCITE-seq, blue; snRNA-seq, pink). **d,g**, UMAP embeddings as in **b** but showing only inCITE-seq nuclear profiles colored by protein levels (nADT) for NeuN (fifth to 95th percentile (**d**)) and p65 (color scale from the second to 98th percentile (**g**)). **e,f,h**, Distribution of protein levels (nCLR, x axis) for NeuN in neuronal (blue) and non-neuronal (gray) nuclei ($P = 0.005$, two-sided KS test (**e**)), PU.1 in microglial (turquoise) and neuronal (gray) nuclei ($P = 10^{-5}$, two-sided KS test (**f**)) and p65 in endothelial (fuchsia) and neuronal (gray) nuclei ($P = 1.1 \times 10^{-15}$, two-sided KS test (**h**)) from one batch. Curve, kernel density estimate. **i**, Immunofluorescence stain of the hippocampus with endothelial marker CD31 (green), NeuN (blue), p65 (pink) and 4,6-diamidino-2-phenylindole (DAPI) (white), representative of three independently conducted experiments. Yellow arrowheads, colocalization of CD31 and p65. Blue arrowheads, lowly expressed p65 in neurons. Scale bar, 50 μm .

reflect background signal (such as ambient RNA or doublets that were missed in filtering). Levels of PU.1, the microglial marker and a lineage-specifying TF, were significantly higher in microglia than

those in neurons (Fig. 2f; $P = 10^{-5}$, two-sided KS test). Expression of p65 was enriched in endothelial nuclei (Fig. 2g,h; $P = 1.1 \times 10^{-15}$, two-sided KS test), which was confirmed by immunohistochemistry

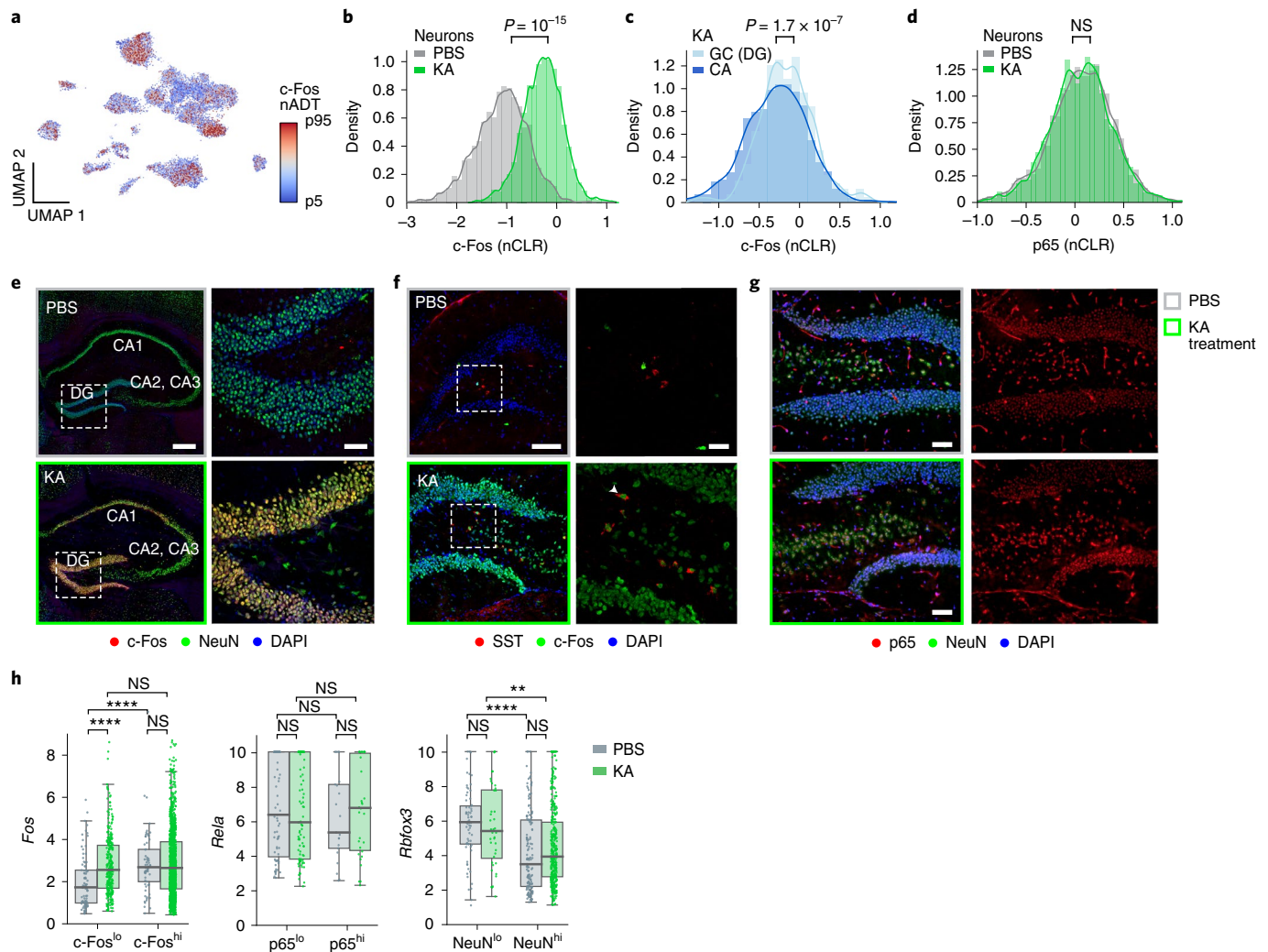


Fig. 3 | inCITE-seq measures changes in nuclear TF levels after stimulation of the mouse hippocampus. a, UMAP embedding of nuclei from inCITE-seq (as in Fig. 2b), colored by c-Fos protein levels (nADT, color scale from the fifth to 95th percentile). Distribution of c-Fos (**b,c**) or p65 (**d**) protein levels (nCLR, x axis) shown as a kernel density estimate in neurons of KA-treated versus PBS-treated mice ($P=10^{-15}$, two-sided KS test (**b**)), in GCs versus CA neurons in KA-treated mice ($P=1.7 \times 10^{-7}$, two-sided KS test (**c**)), or in neurons of KA-treated versus PBS-treated mice (not significant; $P=0.15$, two-sided KS test (**d**)). **e–g**, Immunofluorescence stain of the hippocampus after PBS (gray border) or KA (green border) treatment; representative of three independent experiments. Major hippocampal features denoted are the DG and the CA. **e**, Stain of c-Fos (red), NeuN (green) and DAPI (blue). Left, scale bar, 600 μm . Right, close-up of the DG (dashed box) shows heterogeneity in c-Fos intensity; scale bar, 100 μm . **f**, Stain of SST (red), c-Fos (green) and DAPI (blue). Left, scale bar, 100 μm . Right, close-up of the DG (dashed box); scale bar, 30 μm . **g**, Immunofluorescence stains of p65 (red), NeuN (green) and DAPI (blue); PBS or KA treatment. Left, all stains. Right, p65 only. Scale bar, 100 μm . **h**, Distribution of mRNA levels (z score of log-normalized counts, y axis) in nuclei with high or low levels (defined in Extended Data Fig. 6) of the encoded protein (x axis) after PBS (gray) or KA (green) treatment. For the box plot, the center line indicates the median, box bounds represent first and third quartiles, and whiskers span from each quartile to the minimum or the maximum (1.5 \times interquartile range below 25% or above 75% quartiles). Dots, nuclei with non-zero mRNA levels measured across $n=2$ biologically independent samples, with 1,696 nuclei, 214 nuclei and 653 nuclei shown for *Fos*, *Rela* and *Rbfox3*, respectively. Significance from bottom-left to top-right: $P=2 \times 10^{-6}$, $P=9 \times 10^{-5}$, $P=4 \times 10^{-6}$, $P=9.7 \times 10^{-3}$, two-sided Mann-Whitney test. NS, not significant.

with the endothelial marker CD31 (Fig. 2i, CD31⁺p65⁺, yellow arrowheads), and p65 was also expressed at lower levels in neurons (Fig. 2i, NeuN⁺p65⁺, green arrowheads).

KA treatment induced c-Fos expression, with variations across neuronal subtypes. We observed widespread expression of inCITE-seq-derived nuclear c-Fos (Fig. 3a), with significant upregulation in neurons after KA treatment (Fig. 3b; $P=10^{-15}$, two-sided KS test). Subsets of neurons differed in c-Fos levels, such that nuclei from cornu ammonis (CA) neurons had lower levels than those from granule cells of the dentate gyrus (DG) (Fig. 3c; $P=1.7 \times 10^{-7}$, two-sided KS test). By contrast, p65 levels did not change after KA

treatment at this time scale (Fig. 3d) as expected⁴⁹. These patterns were confirmed by immunofluorescence, showing that c-Fos is expressed in multiple neuronal types, including DG granule cells, CA neurons and somatostatin (SST⁺) interneurons, with higher c-Fos intensity in granule cells than that in CA neurons (Fig. 3e,f) and no change in p65 levels due to treatment (Fig. 3g). Overall, inCITE-seq quantitatively measured nuclear protein levels that reflected diverse levels of activity-regulated TFs across cell types and treatment.

Relating protein and mRNA levels of inCITE target genes. Relating mRNA and protein levels of inCITE target genes revealed

a wide range of regulatory dynamics. We compared transcript and protein levels across nuclei with ‘high’ or ‘low’ protein levels that were categorized using batch-specific thresholds (Extended Data Fig. 6a–d) and assessed how treatment impacted their relationship (Fig. 3h). Transcript levels of the activity-regulated gene *Fos* were elevated in populations with high protein levels (c-Fos^{hi}) compared to those in c-Fos^{lo} populations ($P = 9 \times 10^{-5}$, two-sided Mann–Whitney test) but only after PBS treatment, whereas *Fos* mRNA was highly expressed after KA treatment regardless of c-Fos protein levels. This is consistent with a model where c-Fos positively regulates its own transcript, expression of which is already saturated after treatment at this time scale⁵⁷ and when the downstream effect of mRNA induction on protein levels is not yet observable. By contrast, NeuN and its encoding transcript *Rbfox3* were inversely associated ($P < 0.01$, two-sided Mann–Whitney test); although unspliced, intron-retaining pre-mRNA levels of *Rbfox3* (estimated by Velocyto²⁷, Methods) were upregulated by KA treatment (Extended Data Fig. 7a). This is consistent with other RNA-binding proteins that negatively regulate their own expression via intron retention⁵⁸. *Rela* mRNA levels did not differ across p65 protein levels (Fig. 3h, $P = 1$, two-sided Mann–Whitney test), and *Sp1* transcripts (encoding PU.1) were not detected, underscoring the importance of protein measurements as a complement to RNA measurements, particularly for TFs for which the corresponding RNA is often lowly expressed⁵⁹.

Modeling genome-wide association with each protein recovers known TF targets. We devised an approach to infer the putative genome-wide impact of each TF on gene expression based on nuclear protein levels. As concentrations of TFs in the nucleus shape gene expression⁶⁰, we modeled gene expression as a function of protein levels. To first identify global impacts of TFs, we modeled each gene’s RNA as a linear combination of the four proteins (c-Fos, p65, PU.1, NeuN) after regressing out contributions of cell type (cluster), treatment and their interaction to account for collinearity (especially between KA treatment and c-Fos levels; Methods). Genes significantly associated with each of the three TFs were interpreted as putative TF-regulated genes, with the effect size estimated by their coefficients. TF-associated genes comprised known targets and pathways (Fig. 4a–c and Extended Data Fig. 7b), with PU.1-associated genes including direct regulatory targets and known microglial markers, for example, *Trem2*, *Tyrbp* and *C1qa*⁶⁰, and c-Fos-associated genes including known targets *Npas4*, *Nr4a1*, *Homer1* and its own transcript *Fos*, that reflect activity-induced upregulation³⁵. Although our model also identified NeuN-associated genes (Extended Data Fig. 7c), we reasoned that these genes could reflect direct changes in transcript levels via splicing or indirect effects of a generally transcriptionally active state of the nucleus, as NeuN is associated with decondensed chromatin and enlarged nuclei⁶¹.

TF-associated genes are coexpressed as distinct modules in excitatory neurons. Next, we probed how expression profiles of TF-associated genes relate to each other by their coexpression patterns. Focusing on the broad type of excitatory neurons (EX neurons), we again modeled each gene as a linear combination of the four proteins after regressing out treatment within only EX neurons (Methods and Fig. 4d). We then clustered the differentially expressed genes (DEGs) associated with c-Fos and p65 levels based on the correlation of their coexpression across all nuclei of EX neurons. Genes associated with c-Fos or p65 levels were coexpressed as distinct modules, such that each module reflected a unique set of inferred TF effects (Fig. 4e). Specifically, modules 3 and 4 corresponded to a positive c-Fos effect, modules 1 and 2 corresponded to a negative and positive p65 effect, respectively, and module 5 corresponded to a mixture of both c-Fos and p65 effects.

We hypothesized that widespread weak correlations between c-Fos- and p65-associated genes (Fig. 4e, dashed gray box) may reflect interactions between these two TFs, as their complexes AP-1 and NF- κ B are known to interact^{62,63}. To test for protein–protein interactions, we re-implemented our model with interaction terms (c-Fos*p65, c-Fos*NeuN and p65*NeuN), which uncovered 56 genes associated with c-Fos*p65 that were also coexpressed as a module (Extended Data Fig. 8a,b and Methods). Of these 56 genes, only 11 were associated with c-Fos or p65 alone in the previous model (genes marked by a red asterisk in Fig. 4e), suggesting that these genes may uniquely reflect direct or indirect regulation that requires both c-Fos and p65 to be highly expressed in a cell. Although we caution against overinterpretation of the specific genes due to the limited size of our data and subsequently underpowered analyses, we demonstrate that, as a proof of concept, we can use protein-combination measurements to infer gene modules reflecting TF combinations versus individual TFs.

The association between gene modules and inferred TF effect types prompted us to ask whether gene programs that are normally identified from expression alone⁶⁴ align with protein effects. Using non-negative matrix factorization (NMF) on RNA profiles alone, we identified five gene programs in EX neurons (Fig. 4f and Methods), some of which were upregulated in response to treatment (Extended Data Fig. 8c). Coexpression patterns of the top ten genes of each program revealed that each program also coincided with the different types of TF effects (Fig. 4g). Altogether, our approach allows quantifying the associations between TFs and gene expression modules or pathways, and even disentangling contributions of TF combinations, which can increase the interpretability of gene expression programs.

To assess whether each TF could play a direct or indirect role in regulating their associated DEGs, and identify other TFs that may be involved in co-regulation, we analyzed the *cis* regulatory TF motifs enriched in the enhancer regions of c-Fos-associated and p65-associated genes. Using enhancers defined by differentially accessible regions (DARs) that were profiled in the hippocampus of saline- or KA-treated mice⁶⁵, we identified TF motifs significantly enriched each gene set (c-Fos and p65 DEGs from the additive model; c-Fos*p65 DEGs from the interaction model; Methods). Observed TF motifs included components of the AP-1 complex (Fos, JunB and JunD), activity-regulated TFs Egr1 and Egr3, and Atf2 which is phosphorylated by c-Jun in KA-induced seizure models⁶⁶ (Fig. 4h and Extended Data Fig. 9a). Additionally, the enrichment of Smarcc1, part of the chromatin remodeling BAF (SWI/SNF) complex, in c-Fos DEGs was consistent with the key role of AP-1 in recruiting the BAF complex for chromatin remodeling⁶⁷. RNA transcripts of these TFs were lowly expressed in our data (Fig. 4i). As the widespread presence of AP-1 motifs likely reflected the role of AP-1 itself in chromatin remodeling following KA treatment^{65,67}, we additionally examined TF motifs enriched in DEGs compared to other enhancers upon KA treatment (Extended Data Fig. 9b,c). Notably, motifs of NF- κ B components (encoded by *Rel*, *Rela*, *Relb*, *Nfkb1*, *Nfkb2*) were not significant. This suggests that TF-associated genes in our analysis may also reflect indirect effects of each TF, or that direct effects occur at a different time scale in the case of p65, or depend on the direction of the TF effect (that is, downregulated versus upregulated).

Inferred TF impact on genes depends on treatment context and cell type. Finally, we analyzed whether TF impact on genome-wide expression depended on the treatment context. To assess the global treatment impact, we modeled RNA levels as a linear combination of the four proteins separately within each treatment (PBS or KA) and compared their effects on each gene across treatments (Methods and Extended Data Fig. 9d). The impact of p65 on gene expression was largely consistent across PBS and KA treatments as reflected

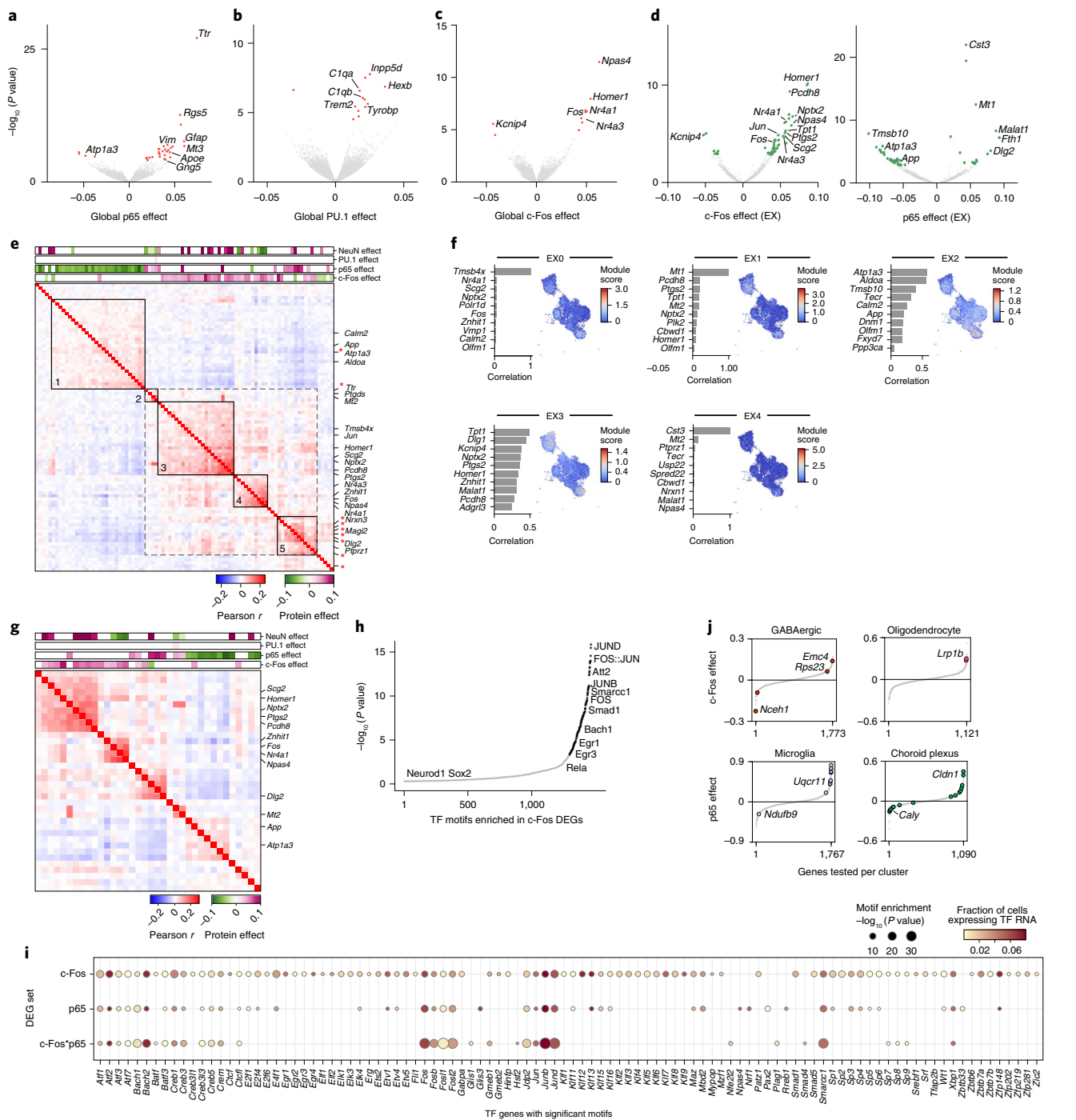


Fig. 4 | Inferring TF effects on gene and module expression using joint protein and transcriptome measurements. a–c, Global association of TFs with genes. Significance (y axis, $-\log_{10}(P \text{ value})$) and effect size (x axis) for genes (dots) associated with p65 (**a**), PU.1 (**b**) and c-Fos (**c**) protein levels across all nuclei, by a model of gene expression as a linear combination of proteins. Colored dots, Benjamini–Hochberg FDR < 5%; select genes are labeled. **d**, Genes associated with each TF in EX neurons. Volcano plot axes for c-Fos (left) or p65 (right) are the same as in **a–c**. Colored dots, Benjamini–Hochberg FDR < 5%; select genes are labeled. **e**, Pearson correlation coefficient (red- and blue-colored bar) of pairwise gene expression across EX neurons (rows and columns) for genes that are positively (purple) or negatively (green) associated with c-Fos or p65 levels, ordered by hierarchical clustering. Top bars, effect size of each protein. Black boxes, coexpression modules. Red asterisk, DEGs associated with c-Fos*p65 in the interaction model (Extended Data Fig. 8b). **f**, NMF programs of EX neurons. Right, UMAP embedding of the EX neuron subset (as in Fig. 2b), colored by the NMF program score. Left, top ten program genes (y axis) and their Pearson correlation with program scores (x axis). **g**, Pearson correlation coefficient (red- and blue-colored bar) of pairwise gene expression across EX neurons (rows and columns) using the top ten genes of each program, ordered by hierarchical clustering. Top bars (purple or green), significant effect sizes of each protein from the linear model. **h**, Significance ($-\log_{10}(P \text{ value})$, y axis) and rank order (x axis) of TF motifs (dots) in enhancers of c-Fos DEGs in EX neurons. Black, significant motifs ($P < 10^{-3}$, hypergeometric test); gray, not significant. **i**, Enriched TF motifs (columns; dot size, $-\log_{10}(P \text{ value})$) and their corresponding RNA expression in EX neurons (dot color), identified in enhancers of DEGs associated with each of the following (rows): c-Fos (additive model), p65 (additive model) or c-Fos*p65 (interaction model). **j**, Cell type-specific DEGs of c-Fos or p65 after KA treatment. Effect size (y axis) of c-Fos (top) or p65 (bottom), sorted by rank order (x axis) in select cell types (top). Color, significant genes (Benjamini–Hochberg FDR < 5%).

by correlated effect sizes ($R^2 = 0.03$, $P = 6 \times 10^{-82}$). By contrast, c-Fos effects varied by treatment ($R^2 = 0.0001$, $P = 0.21$). Genes associated with c-Fos only upon KA treatment included *Ptgs2* (encoding cyclooxygenase 2, COX-2), a responder to oxidative stress after traumatic brain injury⁶⁸. Additionally, we show that cell type-specific TF effects upon KA treatment can be identified, as a proof of concept^{33,36} (Methods and Fig. 4j). Altogether, our results demonstrate that the direct or indirect regulatory impact of TFs on individual genes can depend on cell type and environmental context.

Discussion

inCITE-seq reliably measures quantitative protein and RNA levels in individual nuclei, offering a means to better understand the relationship between regulatory proteins and their target genes during dynamic responses in the native context of tissues. Nuclear protein levels can be used as an interpretable and mechanistic link bridging regulatory proteins and their genome-wide effects, and enable cell type-specific studies of signaling pathways in complex tissues in vivo. While numerous multimodal methods measure intracellular protein targets^{15,16,19}, proteins in the cytoplasm carry different information than those in the nucleus, as in the case of regulatory proteins and translocating TFs; buffers used for intracellular and cytoplasmic targets may result in high background levels for nuclear targets (Fig. 1c). Furthermore, nucleus-based multimodal profiling surmounts key technical challenges to enable characterization of cells from solid tissues that are either difficult to dissociate or archived in frozen form, especially clinical specimens from human disease studies such as cancer and neurodegeneration. inCITE-seq is particularly well suited for studying proteins in pathways that are affected by cellular dissociation protocols, for example, the activity-regulated TF c-Fos.

Antibodies compatible with inCITE-seq should recognize epitopes in frozen tissues after light fixation. We recommend validating antibodies by flow cytometry and in situ immunofluorescence on frozen tissues (Extended Data Figs. 3 and 4). Jointly embedding RNA profiles from inCITE-seq and standard snRNA-seq should help in discerning clusters despite the currently lower RNA complexity of inCITE-seq data. In the future, computational methods that tackle ambient RNA and batch effects through principled modeling, such as CellBender⁵⁶ and totalVI⁶⁹, could be adapted to denoise ambient nuclear protein expression. Transfer learning from inCITE-seq to snRNA-seq may also be used to predict protein levels in snRNA-seq datasets collected without protein measurements.

For refined understanding of gene regulation in tissues during dynamic response, inCITE-seq can be applied with antibodies targeting phosphorylated forms of TFs in samples collected across time. Future studies can combine inCITE-seq with other modalities, such as metabolic labeling^{70–72}, and joint RNA and chromatin accessibility profiles⁷³, coupled with spatial inference⁷⁴. Our intranuclear staining conditions could also be readily adapted to enrich for snRNA-seq of subpopulations based on marker proteins⁷⁵. By measuring TFs simultaneously, inCITE-seq opens the way to decipher complex phenotypes and regulatory mechanisms in development when TF combinations are key to defining cell type diversity or to disentangle interacting pathways. Multiplexed profiling of signaling proteins will enable deciphering changes in activity states, which could be used to recover the impact of ligands acting on multiple receptors across cell types in tissues, and where genome-wide association studies highlight the role of variants in regulatory regions by monitoring protein targets alongside gene expression changes.

Online content

Any methods, additional references, Nature Research 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/s41592-021-01278-1>.

Received: 11 January 2021; Accepted: 19 August 2021;
Published online: 4 October 2021

References

- Habib, N. et al. Div-seq: single-nucleus RNA-seq reveals dynamics of rare adult newborn neurons. *Science* **353**, 925–928 (2016).
- Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
- Slyper, M. et al. A single-cell and single-nucleus RNA-seq toolbox for fresh and frozen human tumors. *Nat. Med.* **26**, 792–802 (2020).
- Lacar, B. et al. Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* **7**, 11022 (2016).
- van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
- Drokhlyansky, E. et al. The human and mouse enteric nervous system at single-cell resolution. *Cell* **182**, 1606–1622 (2020).
- Hwang, W. L. et al. Single-nucleus and spatial transcriptomics of archival pancreatic cancer reveals multi-compartment reprogramming after neoadjuvant treatment. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.08.25.267336> (2020).
- Chen, J. et al. Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274–1285 (2014).
- Liu, Z. & Tjian, R. Visualizing transcription factor dynamics in living cells. *J. Cell Biol.* **217**, 1181–1191 (2018).
- Rabani, M. et al. High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. *Cell* **159**, 1698–1710 (2014).
- Jovanovic, M. et al. Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347**, 1259038 (2015).
- Rabani, M., Pieper, L., Chew, G.-L. & Schier, A. F. A massively parallel reporter assay of 3' UTR sequences identifies in vivo rules for mRNA degradation. *Mol. Cell* **68**, 1083–1094 (2017).
- Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
- Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).
- Gerlach, J. P. et al. Combined quantification of intracellular (phospho-) proteins and transcriptomics from fixed single cells. *Sci. Rep.* **9**, 1469 (2019).
- Reimegård, J. et al. A combined approach for single-cell mRNA and intracellular protein expression analysis. *Commun. Biol.* **4**, 624 (2021).
- Katzenelenbogen, Y. et al. Coupled scRNA-seq and intracellular protein activity reveal an immunosuppressive role of TREM2 in cancer. *Cell* **182**, 872–885 (2020).
- Rivello, F. et al. Single-cell intracellular epitope and transcript detection revealing signal transduction dynamics. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.12.02.408120> (2020).
- Mimitou, E. P. et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-021-00927-2> (2021).
- Mimitou, E. P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).
- Frangieh, C. J. et al. Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nat. Genet.* **53**, 332–341 (2021).
- Wang, Y. et al. Multiplexed in situ protein imaging using DNA-barcoded antibodies with extended hybridization chain reactions. Preprint at *bioRxiv* <https://doi.org/10.1101/274456> (2018).
- Weichert, W. et al. High expression of RelA/p65 is associated with activation of nuclear factor- κ B-dependent signaling in pancreatic cancer and marks a patient population with poor prognosis. *Br. J. Cancer* **97**, 523–530 (2007).
- Lehmann, A. et al. High class I HDAC activity and expression are associated with RelA/p65 activation in pancreatic cancer in vitro and in vivo. *BMC Cancer* **9**, 395 (2009).
- Yang, S.-H. et al. Nuclear expression of glioma-associated oncogene homolog 1 and nuclear factor- κ B is associated with a poor prognosis of pancreatic cancer. *Oncology* **85**, 86–94 (2013).
- Kim, H. D., Shay, T., O'Shea, E. K. & Regev, A. Transcriptional regulatory circuits: predicting numbers from alphabets. *Science* **325**, 429–432 (2009).
- La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
- Buccitelli, C. & Selbach, M. mRNAs, proteins and the emerging principles of gene expression control. *Nat. Rev. Genet.* **21**, 630–644 (2020).
- Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
- Nelson, D. E. et al. Oscillations in NF- κ B signaling control the dynamics of gene expression. *Science* **306**, 704–708 (2004).

31. Hafner, A. et al. p53 pulses lead to distinct patterns of gene expression albeit similar DNA-binding dynamics. *Nat. Struct. Mol. Biol.* **24**, 840–847 (2017).
32. O'Neill, L. A. J. & Kaltschmidt, C. NF- κ B: a crucial transcription factor for glial and neuronal cell function. *Trends Neurosci.* **20**, 252–258 (1997).
33. Spiegel, I. et al. Npas4 regulates excitatory–inhibitory balance within neural circuits through cell-type-specific gene programs. *Cell* **157**, 1216–1229 (2014).
34. Hu, P. et al. Dissecting cell-type composition and activity-dependent transcriptional state in mammalian brains by massively parallel single-nucleus RNA-seq. *Mol. Cell* **68**, 1006–1015 (2017).
35. Hrvatin, S. et al. Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nat. Neurosci.* **21**, 120–129 (2018).
36. Yap, E.-L. et al. Bidirectional perisomatic inhibitory plasticity of a Fos neuronal network. *Nature* **590**, 115–121 (2020).
37. Sigal, A. et al. Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nat. Methods* **3**, 525–531 (2006).
38. Purvis, J. E. et al. p53 dynamics control cell fate. *Science* **336**, 1440–1444 (2012).
39. Gaublomme, J. T. et al. Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat. Commun.* **10**, 2907 (2019).
40. McKinley, K. L. & Cheeseman, I. M. Large-scale analysis of CRISPR/Cas9 cell-cycle knockouts reveals the diversity of p53-dependent responses to cell-cycle defects. *Dev. Cell* **40**, 405–420 (2017).
41. Feldman, D. et al. Optical pooled screens in human cells. *Cell* **179**, 787–799 (2019).
42. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018).
43. Saka, S. K. et al. Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol.* **37**, 1080–1090 (2019).
44. Aitchison, J. Measures of location of compositional data sets. *Math. Geol.* **21**, 787–790 (1989).
45. McGinnis, C. S. et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).
46. Vezzani, A., French, J., Bartfai, T. & Baram, T. Z. The role of inflammation in epilepsy. *Nat. Rev. Neurol.* **7**, 31–40 (2011).
47. Lévesque, M. & Avoli, M. The kainic acid model of temporal lobe epilepsy. *Neurosci. Biobehav. Rev.* **37**, 2887–2899 (2013).
48. Lerea, L. S. & McNamara, J. O. Ionotropic glutamate receptor subtypes activate c-Fos transcription by distinct calcium-requiring intracellular signaling pathways. *Neuron* **10**, 31–41 (1993).
49. Shen, W., Zhang, C. & Zhang, G. Nuclear factor κ B activation is mediated by NMDA and non-NMDA receptor and L-type voltage-gated Ca^{2+} channel following severe global ischemia in rat hippocampus. *Brain Res.* **933**, 23–30 (2002).
50. Kaltschmidt, B., Widera, D. & Kaltschmidt, C. Signaling via NF- κ B in the nervous system. *Biochim. Biophys. Acta* **1745**, 287–299 (2005).
51. Dammer, E. B. et al. Neuron enriched nuclear proteome isolated from human brain. *J. Proteome Res.* **12**, 3193–3206 (2013).
52. Kierdorf, K. et al. Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat. Neurosci.* **16**, 273–280 (2013).
53. Buus, T. B. et al. Improving oligo-conjugated antibody signal in multimodal single-cell analysis. *eLife* **10**, e61973 (2021).
54. Habib, N. et al. Disease-associated astrocytes in Alzheimer's disease and aging. *Nat. Neurosci.* **23**, 701–706 (2020).
55. Saunders, A. et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030 (2018).
56. Fleming, S. J., Marioni, J. C. & Babadi, M. CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. Preprint at *bioRxiv* <https://doi.org/10.1101/791699> (2019).
57. Morgan, J. I., Cohen, D. R., Hempstead, J. L. & Curran, T. Mapping patterns of c-Fos expression in the central nervous system after seizure. *Science* **237**, 192–197 (1987).
58. Bergeron, D., Pal, G., Beaulieu, Y. B., Chabot, B. & Bachand, F. Regulated intron retention and nuclear pre-mRNA decay contribute to PABPN1 autoregulation. *Mol. Cell Biol.* **35**, 2503–2517 (2015).
59. Hughes, T. K. et al. Second-strand synthesis-based massively parallel scRNA-seq reveals cellular states and molecular features of human inflammatory skin pathologies. *Immunity* **53**, 878–894 (2020).
60. Crotti, A. et al. Mutant Huntingtin promotes autonomous microglia activation via myeloid lineage-determining factors. *Nat. Neurosci.* **17**, 513–521 (2014).
61. Lind, D., Franken, S., Kappler, J., Jankowski, J. & Schilling, K. Characterization of the neuronal marker NeuN as a multiply phosphorylated antigen with discrete subcellular localization. *J. Neurosci. Res.* **79**, 295–302 (2005).
62. Stein, B. et al. Cross-coupling of the NF- κ B p65 and Fos/Jun transcription factors produces potentiated biological function. *EMBO J.* **12**, 3879–3891 (1993).
63. Fujioka, S. et al. NF- κ B and AP-1 connection: mechanism of NF- κ B-dependent regulation of AP-1 activity. *Mol. Cell Biol.* **24**, 7806–7819 (2004).
64. Cleary, B., Cong, L., Cheung, A., Lander, E. S. & Regev, A. Efficient generation of transcriptomic profiles by random composite measurements. *Cell* **171**, 1424–1436 (2017).
65. Fernandez-Albert, J. et al. Immediate and deferred epigenomic signatures of in vivo neuronal activation in mouse hippocampus. *Nat. Neurosci.* **22**, 1718–1730 (2019).
66. Gupta, S., Campbell, D., Dérjard, B. & Davis, R. J. Transcription factor ATF2 regulation by the JNK signal transduction pathway. *Science* **267**, 389–393 (1995).
67. Vierbuchen, T. et al. AP-1 transcription factors and the BAF complex mediate signal-dependent enhancer selection. *Mol. Cell* **68**, 1067–1082 (2017).
68. Strauss, K. I. et al. Prolonged cyclooxygenase-2 induction in neurons and glia following traumatic brain injury in the rat. *J. Neurotrauma* **17**, 695–711 (2000).
69. Gayoso, A. et al. Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18**, 272–282 (2021).
70. Hartmann, F. J. et al. Single-cell metabolic profiling of human cytotoxic T cells. *Nat. Biotechnol.* **39**, 186–197 (2020).
71. Qiu, Q. et al. Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. *Nat. Methods* **17**, 991–1001 (2020).
72. Boon, R., Silveira, G. G. & Mostoslavsky, R. Nuclear metabolism and the regulation of the epigenome. *Nat. Metab.* **2**, 1190–1203 (2020).
73. Ma, S. et al. Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* **183**, 1103–1116 (2020).
74. Biancalani, T. et al. Deep learning and alignment of spatially-resolved whole transcriptomes of single cells in the mouse brain with Tangram. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.08.29.272831> (2020).
75. Amamoto, R. et al. FIN-seq: transcriptional profiling of specific cell types from frozen archived tissue of the human central nervous system. *Nucleic Acids Res.* **48**, e4 (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

Mice. C57BL/6J (JAX 000664) mice were purchased from Jackson Laboratory and bred in house. Male mice were used at ~8 weeks of age. All mice were maintained under SPF conditions on a 12-h light–dark cycle at an ambient temperature of $21.5 \pm 1^\circ\text{C}$ with relative humidity between 30% and 70% and were provided food and water ad libitum. All mouse experiments were approved by and performed in accordance with the Institutional Animal Care and Use Committee guidelines at Weill Cornell Medicine.

Cell culture. HeLa-TetR-Cas9 cells expressing a p65–mNeonGreen reporter construct (gift from J. Schmid-Burgk (Broad Institute, currently at Universität Bonn); original cell line source, I. Cheeseman, MIT^{40,41}) were cultured at 37°C with 5% CO_2 in DMEM with high glucose, pyruvate, GlutaMAX (Thermo Fisher Scientific, 10569010), heat-inactivated FBS (Thermo Fisher Scientific, 16000044) and 100 U ml^{-1} penicillin–streptomycin (Thermo Fisher Scientific, 15140163). For immunohistochemistry, cells were seeded on poly-L-lysine-treated #1.5 glass coverslips (Thomas Scientific, 1217N81) in six-well plates at a density of $\sim 5 \times 10^4$ cells per ml 24 h before TNF- α stimulation. For inCITE-seq, HeLa cells were seeded in 10-cm Petri dishes at least 24 h before TNF- α stimulation and were assayed at 70–80% confluence.

Stimulation of HeLa cells with TNF- α . HeLa cells were stimulated by adding medium containing TNF- α at final concentration of 30 ng ml^{-1} and were incubated at 37°C for 40 min to induce translocation of p65 into the nucleus. TNF- α -containing medium was aspirated, and cells were washed with $1 \times$ PBS. We then added 2 ml EZ lysis buffer (Sigma-Aldrich, N3408), scraped cells, and transferred lysates into a 15-ml Falcon tube for nuclear extraction.

Injection of mice with kainic acid. Eight-week-old male mice were acclimated in the procedure room for 1 h beforehand and then injected i.p. with either PBS or 20 mg per kg KA (Sigma, K0250) dissolved in PBS. All animals were observed continuously for 2 h and scored using a modified Racine scale (stages 0–6)⁷⁶; any mice that did not reach at least stage 1 (immobility and rigidity) by 30 min after KA treatment were given an additional injection of 10 mg per kg KA to facilitate seizure activity. After 2 h, mice were euthanized with CO_2 , 20 ml PBS was perfused through the left ventricle, and the entire brain was removed. The hippocampus was then dissected on ice before freezing on dry ice.

Immunohistochemistry of HeLa cells. Coverslips (#1.5, 18 mm, Thomas Scientific, 1217N81) treated with poly-L-lysine were seeded with HeLa cells in six-well plates and stimulated with TNF- α as described previously. Wells were washed with PBS, fixed with 4% PFA at room temperature (RT) for 15 min then washed three times with PBS. For subsequent steps, coverslips were placed with cells facing down on a sheet of Parafilm, such that solutions (100 μl) were sandwiched between the coverslip and Parafilm. Cells were blocked and permeabilized at RT for 30 min ($1 \times$ PBS, 5% normal goat serum, 0.3% Triton X-100), incubated with 1:200 p65^{Ab} (BioLegend, 622601) in antibody solution (5% BSA, 0.02% Tween-20 in $1 \times$ PBS) for 1 h at RT and washed three times with PBST ($1 \times$ PBS, 0.02% Tween-20). Anti-rabbit Alexa Fluor 647 secondary antibody (Invitrogen, A27040) was added at 1:1,000 in PBST for 1 h in the dark at RT, and then samples were washed four times with PBST, and the final wash contained 1:1,000 DAPI. Coverslips were mounted onto Superfrost slides (Fisher Scientific, 22-037-246) with antifade (Thermo Fisher Scientific, S36937) and sealed with nail polish. Slides were stored at 4°C until imaging.

Immunohistochemistry of the mouse hippocampus. Mice were euthanized 2 h after injection of PBS or KA and perfused through the left ventricle with 20 ml PBS. The entire brain tissue (CNS) was removed and then immersed in OCT and quickly frozen on dry ice. Tissue was sectioned at $10 \mu\text{m}$ using a cryotome and collected on slides, which were frozen on dry ice and stored at -30°C until further use. Slides were removed from storage and fixed with 4% PFA for 10 min at RT, washed with PBS and then blocked (PBS containing 0.1% Triton X-100, 5% normal donkey serum and 5% normal goat serum) for 30 min at RT. Sections were incubated with the following primary antibodies at the indicated dilution in blocking buffer overnight at 4°C : anti-NeuN (BioLegend, 1B7, 834502, 1:1,000), anti-p65 (BioLegend, Poly6226, 622601, 1:200), anti-c-Fos (BioLegend, Poly6414, 641401, 1:200), anti-CD31 (eBioscience, 390, 14-0311-82, 1:200). Sections were then washed three times with PBS before incubation with secondary antibodies (Jackson ImmunoResearch, donkey anti-rabbit AF647, 711-605-152; donkey anti-rabbit AF594, 711-585-152; donkey anti-mouse AF488, 715-545-150; donkey anti-rat AF647, 712-605-153) at 1:500 in blocking buffer for 1 h at RT. Sections were washed once with PBS, once with PBS containing DAPI and a final time with PBS before mounting (ProLong Diamond Antifade, Thermo Fisher). Slides covered with coverslips were dried overnight, sealed with clear nail polish and imaged.

For immunohistochemistry of mouse hippocampal sections from fixed brains, mice were perfused with 20 ml PBS followed by 20 ml 4% PFA. The entire brain was removed, and brain tissue was incubated in 4% PFA overnight before washing three times with 10 ml PBS. Brain tissue was dehydrated in 30% sucrose in PBS overnight and then embedded and frozen in OCT. Sections were processed in the

same manner as that described above but without the 10-min post-cut fixation in 4% PFA. Antibodies used for the fixed versus frozen brain comparison in Extended Data Fig. 4 are anti-c-Fos (Abcam, ab190289), anti-NeuN (Abcam, ab190565) and anti-PU.1 (Cell Signaling Technology, 2258).

Microscopy. HeLa cells and mouse hippocampal sections were imaged on an Olympus Fluoview FV1200 biological confocal scanning microscope at $20 \times$ or $40 \times$ (Olympus, LUCPLFN) with sequential laser emission and Kalman filtering. Images were processed with ImageJ.

Flow cytometry analysis of PU.1 staining in microglial suspensions. Mice were euthanized with CO_2 and perfused through the left ventricle with 20 ml PBS. The whole brain was removed and placed in 2.5 ml digestion buffer (PBS, 5% FCS, 1 mM HEPES) before being finely chopped. Collagenase D (Roche, 400 U) was added to the mixture, which was then incubated at 37°C for 30 min before adding $50 \mu\text{l}$ 0.5 M EDTA, followed by a 5-min incubation. Digested tissue was mashed through a $40\text{-}\mu\text{m}$ cell strainer, pelleted at 700g in a swinging-bucket centrifuge and then resuspended in 10 ml 38% isotonic Percoll and centrifuged at 2,000 r.p.m. for 30 min with no brake. The myelin debris layer was removed by aspiration, and the pellet was washed with PBS. Cells were then blocked with 1:100 FcX (BioLegend, 156604) before a 15-min incubation with the following antibodies at a 1:200 dilution in PBS: anti-CD45.2-FITC (eBioscience, 104, 11-0454-82), anti-CD4-BUV395 (BD, GK1.5, 563790), anti-CD11b-BV421 (BioLegend, M1/70, 101235), anti-CX3CR1-APC (BioLegend, SA011F11, 149008). Cells were washed with PBS and then fixed and permeabilized using the FcX3/Transcription Factor Staining Buffer Set (eBioscience) before staining with anti-PU.1-PE antibody (BioLegend, 7C2C34, 681307) or rat anti-IgG2a-PE isotype (BioLegend, RTK2758, 400507) for 30 min at RT. Cells were washed once, resuspended, run on an LSRFortessa cytometer (BD Biosciences) and analyzed with FlowJo software (Tree Star).

Nuclear extraction. Nuclei from tissue or cell lines were extracted using EZ Prep (Sigma-Aldrich, N3408) and the Glass Dounce kit (Sigma-Aldrich, D8938) as previously described⁷. Briefly, cells or frozen tissue were placed in 2 ml EZ lysis buffer containing Recombinant RNase Inhibitor (Takara Bio, 2313A) and dounced 24 times with pestle A and then 24 times with pestle B. Nuclear suspensions were transferred to 15-ml Falcon tubes, an additional 3 ml EZ lysis buffer was added, and suspensions were incubated on ice for 5 min, pelleted (500g for 5 min at 4°C) with a swinging-bucket centrifuge, resuspended in 5 ml EZ lysis buffer with a P1000 pipette, incubated on ice for 5 min and pelleted as in the previous step. Nuclei were then resuspended in 1 ml prechilled buffer ($1 \times$ PBS, 3 mM MgCl_2 , Recombinant RNase Inhibitor (Takara Bio, 2313A)) and filtered through a 35- μm flow cytometry tube (Falcon, 352235).

Intranuclear antibody staining of nuclear suspensions. Nuclei were simultaneously fixed and permeabilized by adding 3 ml 1.33% FA–NT (1.33% formaldehyde, 0.2% NP-40, 0.1% Tween-20, 3 μl glacial acetic acid) to 1 ml nuclei suspended in PBS with 3 mM MgCl_2 . Samples were incubated for 10 min at 4°C with rocking. Fixation was quenched by adding 3 μl of 1 M glycine and then immediately filtering through a 20- μm strainer (pluriSelect, 431002040). Nuclei were pelleted in a swinging-bucket centrifuge at 850g for 5 min at 4°C (centrifuge condition for all subsequent spins) and then resuspended in 500 μl blocking buffer (see below) and incubated for 15 min at 4°C with rocking and pelleted. Pellets were resuspended in 200 μl blocking buffer containing primary antibodies and incubated at 4°C for 1 h with rocking. Primary antibody concentrations were as follows: p65^{Ab} (raised in rabbit) at 1:400, p65^{inCITE-Ab} (raised in rabbit) at 1:400, c-Fos^{inCITE-Ab} (raised in rabbit) at 1:400, NeuN^{inCITE-Ab} (raised in mouse) at 1:500, PU.1^{inCITE-Ab} (raised in rat) at 1:200. Nucleus-hashing antibodies (BioLegend, 682213, 682215) were simultaneously added to each sample at 1:200. After incubation, nuclei were pelleted, washed twice with 500 μl 0.2% PBST with Recombinant RNase Inhibitor, incubated for 5 min and repelleted. Nuclei were then either resuspended in 300 μl $1 \times$ PBS to prepare for loading on the 10x Genomics Chromium instrument (below) or resuspended in 200 μl blocking buffer with secondary antibodies at 1:1,000 and $10 \times$ DAPI, incubated in the dark at 4°C for 30 min, washed twice as previously described, resuspended in 0.2% PBST with Recombinant RNase Inhibitor and filtered through a 20- μm strainer (pluriSelect, 431002040) for flow cytometry.

Antibodies used were p65^{Ab} or p65^{inCITE-Ab} (BioLegend, Poly6226, 622601), c-Fos^{inCITE-Ab} (BioLegend, Poly6414, 641401), NeuN^{inCITE-Ab} (BioLegend, 1B7, 834502) and PU.1^{inCITE-Ab} (BioLegend, 7C2C34, 681307).

Blocking buffers used are described as follows:

- Optimized inCITE-seq buffer: 1:100 FcX (BioLegend, 156604), 1% UltraPure BSA (Thermo Fisher Scientific, AM2618), 0.05% dextran sulfate, 0.2% Tween-20 and Recombinant RNase Inhibitor in $1 \times$ PBS. Dextran sulfate may be substituted with a 1:200 dilution of HCR probe hybridization buffer (tissue-section format) from Molecular Instruments.
- Intranuclear stain buffer used for p65^{Ab} in Extended Data Fig. 1c: 1:100 FcX (BioLegend, 156604), 1% UltraPure BSA (Thermo Fisher Scientific, AM2618), 0.2% Tween-20 and Recombinant RNase Inhibitor in $1 \times$ PBS.

- Commercial intracellular buffer 1 in Fig. 1c: Intracellular Staining Permeabilization Wash Buffer (BioLegend, 42100), used according to the manufacturer's instructions.
- Commercial intracellular buffer 2 in Fig. 1c: eBioscience Permeabilization Buffer (Thermo Fisher, 00-8333-56), used according to the manufacturer's instructions.

inCITE antibodies. Pure clones of all antibodies were conjugated with the TotalSeq-A format (BioLegend).

inCITE-seq. Antibody-stained nuclei were resuspended in PBS with 3 mM MgCl₂, filtered through a 10-μm filter, counted in a hemocytometer and promptly loaded onto a Chromium single-cell V3 3' chip (10x Genomics) according to the manufacturer's protocol for GEM formation. For the HeLa experiment, a single V3 3' 10x channel was loaded with 10,000 NT nuclei and 10,000 TNF-α-treated nuclei with nucleus hashing. For the mouse hippocampus experiment, two V3 3' 10x channels were loaded, each channel with 30,000 nuclei of a 1:1 mix of nucleus-hashed PBS-treated and KA-treated samples, such that, in total, $n = 2$ PBS-treated samples and $n = 2$ KA-treated samples were loaded across two channels.

After GEM formation, simultaneous reverse cross-linking and reverse transcription were conducted by incubating GEMs at 53 °C for 45 min, followed by incubation at 85 °C for 5 min. Samples were stored at -20 °C until GEM recovery according to the manufacturer's instructions. cDNA was amplified using the standard 10x Genomics single-cell 3' V3 protocol (10x Genomics) with both HTO and ADT PCR additive primers included in the AMP mix at 0.1 μM and 0.2 μM, respectively. After amplification, antibody-oligonucleotide-derived cDNA fragments were separated from mRNA-derived cDNA through solid phase reversible immobilization (SPRI)-based size selection by incubating the cDNA amplification product in 0.6× SPRIselect (Beckman Coulter, B23319) for 5 min at RT. At this stage, antibody-oligonucleotide-derived cDNA is contained in the supernatant, while mRNA-derived cDNA remains on SPRIselect beads. Supernatant containing antibody-oligonucleotide-derived cDNA was removed and separately stored for HTO and ADT library construction. SPRIselect containing mRNA-derived cDNA (WTA) was washed two times with 80% ethanol, eluted into 40 μl elution buffer and stored for gene expression library construction.

Gene expression libraries were constructed from cleaned WTA (10 μl) using manufacturer-specific enzymatic fragmentation, adaptor ligation and sample index attachment and then eluted in 30 μl elution buffer according to the standard 10x Genomics single-cell 3' V3 protocol. Samples were stored at -20 °C until library quantification and sequencing.

For HTO and ADT library construction, antibody-oligonucleotide-derived cDNA was mixed with 1.4× SPRIselect, incubated at RT for 5 min, incubated on a magnet for 5 min, washed twice with standard washes with 80% ethanol and eluted into 24 μl sterile ddH₂O. Afterward, 6 μl eluted solution was added to one of two PCR solutions containing either a unique HTO or ADT index primer mix and NEBNext 2× Master Mix (New England Biolabs, M0541L) to construct separate HTO and ADT libraries. Libraries were constructed by PCR amplification with the following conditions: 98 °C for 5 min, 21 cycles at 98 °C for 2 s and 72 °C for 15 s and a final step at 72 °C for 1 min. PCR products were purified with 2.0× SPRI beads with a 10-min incubation, a 5-min magnetic separation and two washes with 80% ethanol. Purified products were eluted into 20 μl EB. Samples were stored at 4 °C until library quantification and sequencing.

Gene expression and HTO and ADT libraries were quantified using a standard Qubit instrument (Thermo Fisher, Q32853) and the Agilent TapeStation (Agilent, G2991AA) to check for library size. Optimal library sizes were ~420 bp for gene expression and ~180 bp for ADT and HTO. Libraries were pooled and sequenced on the NextSeq 500 platform (Illumina) using a 75-cycle kit (read 1, 28 cycles; index 1, eight cycles; read 2, 55 cycles).

snRNA-seq of the mouse hippocampus. Nuclei were extracted from the hippocampus of PBS-treated ($n = 1$) and KA-treated ($n = 1$) mice as described above, resuspended in 1 ml chilled 1× PBS with 3 mM MgCl₂ and filtered through a 10-μm filter, and nuclei were hashed per treatment as previously described and promptly loaded onto a Chromium single-cell V3.1 3' chip (10x Genomics) according to the manufacturer's protocol. A single V3.1 3' 10x channel was loaded with 10,000 nuclei in a 1:1 mix of nucleus-hashed PBS and KA samples. cDNA and HTO libraries were generated and sequenced as previously described³⁹.

inCITE-seq and snRNA-seq data preprocessing. Sequencing data were processed with Cell Ranger version 4.0.0 on Cumulus version 1.0 (ref. ⁷⁷). Reads from demultiplexed FASTQ files were aligned to pre-mRNA annotated genomes of the mouse mm10 or human GRCh38 reference genome as previously described⁷⁸. Hashed nuclei were demultiplexed using DemuxEM³⁹ with parameters 'min_num_genes=10, min_num_umis=1, min_signal_hashtag=1'; nuclei with ambiguous treatment assignment (that is, nuclei not assigned to NT or TNF-α groups among HeLa nuclei or to PBS or KA groups among mouse hippocampal nuclei) were discarded (5.9% and 2.7% for HeLa and mouse, respectively); barcode collision rates are discussed elsewhere^{33,39,79}. For mouse hippocampus data, raw counts across both

channels were combined. Unspliced pre-mRNA and spliced mRNA counts were generated from BAM files using Velocyto version 0.17.17 (ref. ⁷⁷). All gene expression matrices were analyzed by SCANPY (version 1.6.0)⁸⁰. For inCITE-seq data, protein ADT counts corresponding to each nucleus or cell barcode were also added.

Analysis of gene expression. For HeLa data, genes in at least ten nuclei and nuclei with at least 500 genes and at most 5,000 UMI counts were retained, resulting in a matrix of 10,014 nuclei with 13,942 genes across both NT and TNF-α groups. For mouse hippocampus data, genes in at least three nuclei, and nuclei with at least 50 genes and at most 900 or 1,200 UMI counts per replicate (batch 1 or 2, respectively) were retained; we then removed nuclei with mitochondrial gene content >5%, HTO counts greater than 5,000, and anti-c-Fos antibody counts exceeding 300 ADT counts, resulting in a matrix of 41,332 nuclei with 20,679 genes. Gene counts were normalized within each nucleus and then log normalized as $\ln(x + 1)$.

Normalizing protein expression. Protein abundances measured by ADT were normalized by nuclear HTO counts after adding a pseudocount as in the equation $nADT = (ADT + 1)/HTO$ and then scaled to centered log ratios (nuclear; 'nCLR') as in the equation $nCLR = nADT / (\prod_i nADT_i)^{1/n}$, where the denominator is the geometric mean calculated as the n^{th} root of the product \prod_i across individual nuclei i . For mouse hippocampus data, nCLR scaling was conducted within each batch, such that batch-specific geometric means were used in the last term.

Clustering mouse hippocampus data using single-nucleus RNA profiles from inCITE-seq. 5,194 variable genes on log-normalized counts were selected using SCANPY's 'highly_variable_genes' function ('min_mean=0.004, max_min=0.08, min_disp=0.3'), log counts were scaled, and UMI counts and mitochondrial content were regressed out using SCANPY's 'regress_out' function. Dimensionality reduction was performed with PCA on variable genes in SCANPY, followed by a PyTorch implementation of Harmony⁸¹ to correct for batch based on 10x channels and treatment (PBS and KA). A nearest-neighbor graph was constructed with $k = 10$ neighbors and the top 40 principal components, clustered with the Leiden algorithm and embedded using UMAP⁸², all in SCANPY.

Joint embedding of inCITE-seq and snRNA-seq mouse hippocampus profiles. snRNA-seq profiles of fixed nuclei with inCITE-seq data ($n = 24,444$ nuclei) were jointly embedded with snRNA-seq data from this study ($n = 1,887$) and from a published study⁵⁴ ($n = 15,001$, WT only). In total, 7,541 variable genes were identified in a single joint analysis from log-normalized counts across all datasets using SCANPY's 'highly_variable_genes' function ('min_mean=0.004, max_min=0.08, min_disp=0.3'). Log counts were then scaled; UMI counts, mitochondrial content and sequencing assay type were regressed out with SCANPY's 'regress_out' function. Dimensionality reduction was performed on variable genes via PCA in SCANPY, followed by a Python implementation of Harmony⁸¹ to correct for batch based on 10x channels, to regress out differences due to treatment (PBS and KA) and to further correct for assay type. A k -nearest-neighbor graph was constructed with $k = 10$ neighbors and the top 40 principal components, clustered with the Leiden algorithm⁸³ in SCANPY and embedded using UMAP⁸² in SCANPY. Unsupervised clusters were identified after accounting for treatment, batch and sequencing assay type. A cluster with high mitochondrial content ($n = 327$ nuclei) and a subcluster with low transcript abundance ($n = 179$) were removed; 3,059 individual suspected doublets identified by Scrublet⁸⁴ were removed. The final embedding consisted of $n = 22,260$ nuclear RNA profiles from inCITE-seq and $n = 15,507$ nuclear RNA profiles from standard snRNA-seq.

Genes associated with p65 protein levels in HeLa samples. Gene expression was modeled using a generalized linear model with a negative binomial fit as follows:

$$Y_i \sim p65 + G2M_{score} + S_{score} + \log(UMI) + \log(HTO),$$

where Y_i is log normalized, unscaled $\ln(x + 1)$ counts for gene i , p65 is the p65 protein level in units of nCLR, $G2M_{score}$ and S_{score} are cell cycle scores calculated with 'score_genes_cell_cycle' in SCANPY using previously defined genes⁸⁵, and $\log(UMI)$ and $\log(HTO)$ are natural log counts of unique RNA molecular identifiers and nuclear HTOs, respectively. Significance was established at an FDR of 1% after Benjamini-Hochberg correction using the statsmodels package in Python.

Gene ontology analysis. Gene sets were queried using SCANPY's 'queries.enrich' module, a wrapper around g:Profiler⁸⁶, to identify gene ontology biological processes. P values were calculated using a two-sided hypergeometric test, corrected for term and query size.

Genes globally associated with protein levels in the mouse hippocampus. We implemented a two-step mixed linear model using the statsmodel package in Python to account for collinearity, specifically between treatment and c-Fos protein levels. First, we modeled gene expression as shown in the following equation to regress out the effects of treatment and cell type:

$$Y_i \sim C(cluster) + C(treatment) + C(cluster) * C(treatment) + \log(UMI) + \log(HTO) + (1|B),$$

where Y_i is the scaled z score of $\ln(x+1)$ counts for gene i across all nuclei, $C(\text{cluster})$ is a categorical (C) variable indicating cluster membership (cell type), $C(\text{treatment})$ is a categorical variable indicating PBS or KA treatment, $\log(\text{UMI})$ and $\log(\text{HTO})$ are natural log counts of unique RNA molecular identifiers and nuclear HTOs, and B is a categorical variable denoting the 10x channel batch. We then fit the residuals of each gene, rY_i , as a linear combination of the four proteins, accounting for batch:

$$rY_i \sim \text{NeuN} + c\text{-Fos} + p65 + \text{PU.1} + (1|B).$$

Genes in at least 15 nuclei were used for analysis, and significance was established at an FDR of 5% after Benjamini–Hochberg correction. For identifying treatment-dependent global effects, we implemented the two-step model separately within PBS-treated or KA-treated nuclei.

Genes associated with protein levels in the mouse hippocampus in excitatory neurons. Similar to the global approach, we first regressed out treatment:

$$Y_i \sim C(\text{treatment}) + \log(\text{UMI}) + \log(\text{HTO}) + (1|B),$$

where all variates are the same as those previously described. We then modeled residuals of each gene as a linear combination of the four proteins, accounting for batch. In the additive model, we used

$$rY_i \sim \text{NeuN} + c\text{-Fos} + p65 + \text{PU.1} + (1|B)$$

and in the interaction model, we used

$$rY_i \sim \text{NeuN} + c\text{-Fos} + p65 + \text{PU.1} + c\text{-Fos} * p65 \\ + c\text{-Fos} * \text{NeuN} + p65 * \text{NeuN} + (1|B).$$

Genes found in $\geq 3\%$ of EX neuronal nuclei were used for analysis (1,757 genes tested across 15,404 EX neuronal nuclei); significance was established at an FDR of 5% with Benjamini–Hochberg correction.

Identifying cell type-specific gene programs with non-negative matrix factorization. Gene programs of EX neurons were identified by NMF (Python sklearn package, NMF function, 'random_state=0, L1 regularization with l1_ratio=1, alpha=0') of their RNA profiles on a subset of genes (a combination of highly variable genes identified for clustering and c-Fos-associated and p65-associated genes, with manual removal of highly expressed and variable genes *Ttr*, *Fth1*, *Ptgds*). Five programs were identified for EX neurons.

Cell type-specific genes associated with protein levels in the mouse hippocampus after KA treatment. Similar to methods used before, we implemented the following mixed linear model for each cell cluster c using the statsmodels package in Python:

$$Y_{i,c} \sim \text{NeuN} + c\text{-Fos} + p65 + \text{PU.1} + \log(\text{UMI}) + \log(\text{HTO}) + (1|B),$$

where $Y_{i,c}$ is the scaled z score of $\ln(x+1)$ counts for gene i in cluster c , NeuN, c-Fos, p65 and PU.1 are protein levels in units of nCLR, $\log(\text{UMI})$ and $\log(\text{HTO})$ are natural log counts of unique RNA molecular identifiers and nuclear HTOs, and B is a categorical variable denoting the 10x channel batch. Protein levels (nCLR) were first scaled using the Python package sklearn preprocessing. Genes found in $\geq 3\%$ nuclei of each cluster were used for analysis; significance was established at an FDR of 5% with Benjamini–Hochberg correction.

Transcription factor motif enrichment in DEGs. We used DARs of the mouse hippocampus profiled in saline-treated and KA-treated samples (1 h after treatment) by Fernandez-Albert et al.⁶⁵. Nearest DARs located more than 1 kb from the transcriptional start site (upstream) or transcriptional termination site (downstream) of DEGs associated with c-Fos or p65 (additive model) or c-Fos*p65 (interactional model) were considered as enhancers and used for motif enrichment analysis. To find enriched motifs, we scanned a given set of differentially accessible peaks for all DNA-binding motifs in CIS-BS (<http://cisbp.ccbbr.utoronto.ca>) and JASPAR2018_CORE vertebrates_non-redundant (<http://jaspar2018.genereg.net>) databases. We then tested for the probability of our observed motif frequency among KA-treated DARs using a hypergeometric test, with a null model based on random sampling of all ATAC-seq peaks matching for GC content. Motifs with $P < 10^{-3}$ were considered significant.

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

Data availability

Raw gene expression count matrices of all inCITE-seq data, BAM files of mapped reads and the matrix of mouse hippocampus inCITE-seq data jointly embedded with snRNA-seq data are available on Gene Expression Omnibus under the

accession GSE163480. Data from Habib et al.⁵⁴ are available under GSE143758. Data from MULTI-seq used to compare RNA complexity in HEK cells are available under GSE129578. Databases of TF motifs (CIS-BS and JASPAR2018_CORE vertebrates_non-redundant) are available at <http://cisbp.ccbbr.utoronto.ca> and <http://jaspar2018.genereg.net>, respectively. Source data are provided with this paper.

Code availability

Code used for analyses is available at <https://github.com/klarman-cell-observatory/inCITE-seq>.

References

- Racine, R. J. Modification of seizure activity by electrical stimulation: II. Motor seizure. *Electroencephalogr. Clin. Neurophysiol.* **32**, 281–294 (1972).
- Li, B. et al. Cumulus provides cloud-based data analysis for large-scale single-cell and single-nucleus RNA-seq. *Nat. Methods* **17**, 793–798 (2020).
- Bakken, T. E. et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE* **13**, e0209648 (2018).
- Stoeckius, M. et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
- Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
- Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
- McInnes, L. et al. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018).
- Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
- Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291 (2019).
- Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
- Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).

Acknowledgements

We thank J. Schmid-Burgk and I. Cheeseman for the HeLa p65–mNeonGreen reporter line, L. Gaffney for assistance with figure preparation, P. Thakore for coining the acronym inCITE-seq, A. Rubin for critical feedback on the manuscript, C. McGinnis for helpful sharing of data, the Broad Institute Flow Cytometry Core facility, and all members of the Regev laboratory for helpful discussions. This research was supported by NIH/NHGRI CEGRS grant 5RM1 HG006193. A.R. was a Howard Hughes Medical Institute Investigator (until 31 July 2020). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

H.C. conceived and designed the study with guidance from A.R. C.N.P. designed and performed mouse experiments with guidance from D.A. H.C. and E.M.M. developed and performed inCITE-seq experiments, with buffer optimization input from F.C. and early-stage experimental support from J.W. C.N.P. and E.M.M. conducted immunohistochemistry. D.P. conducted 10x experiments and constructed sequencing libraries. B.Z.Y. conjugated inCITE antibodies. H.C. analyzed and interpreted data with help from E.H. on *cis* regulatory motif enrichment analysis and supervision from A.R. A.R. provided project oversight and funding. H.C. and A.R. wrote the manuscript with input from all authors.

Competing interests

A.R. is a founder and equity holder of Celsius Therapeutics, an equity holder in Immunitas Therapeutics and, until 31 August 2020, was an SAB member of Syros Pharmaceuticals, Neogene Therapeutics, Asimov and Thermo Fisher Scientific. From 1 August 2020, A.R. is an employee of Genentech. From May 2021, D.P. is an employee of Genentech. B.Z.Y. was formerly an employee of BioLegend and is now an employee of Spatial Genomics. The remaining authors declare no competing interests.

Additional information

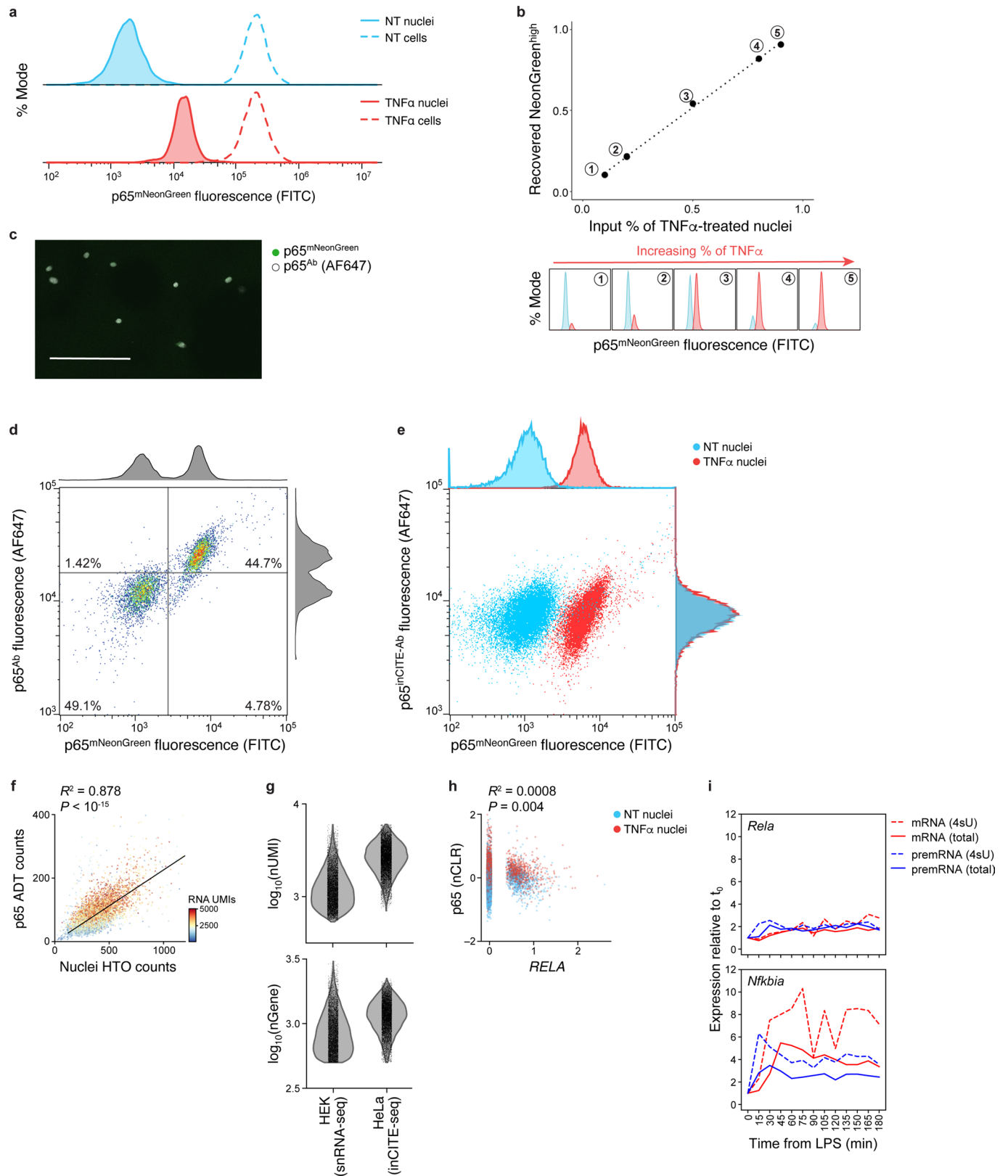
Extended data is available for this paper at <https://doi.org/10.1038/s41592-021-01278-1>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41592-021-01278-1>.

Correspondence and requests for materials should be addressed to Hattie Chung or Aviv Regev.

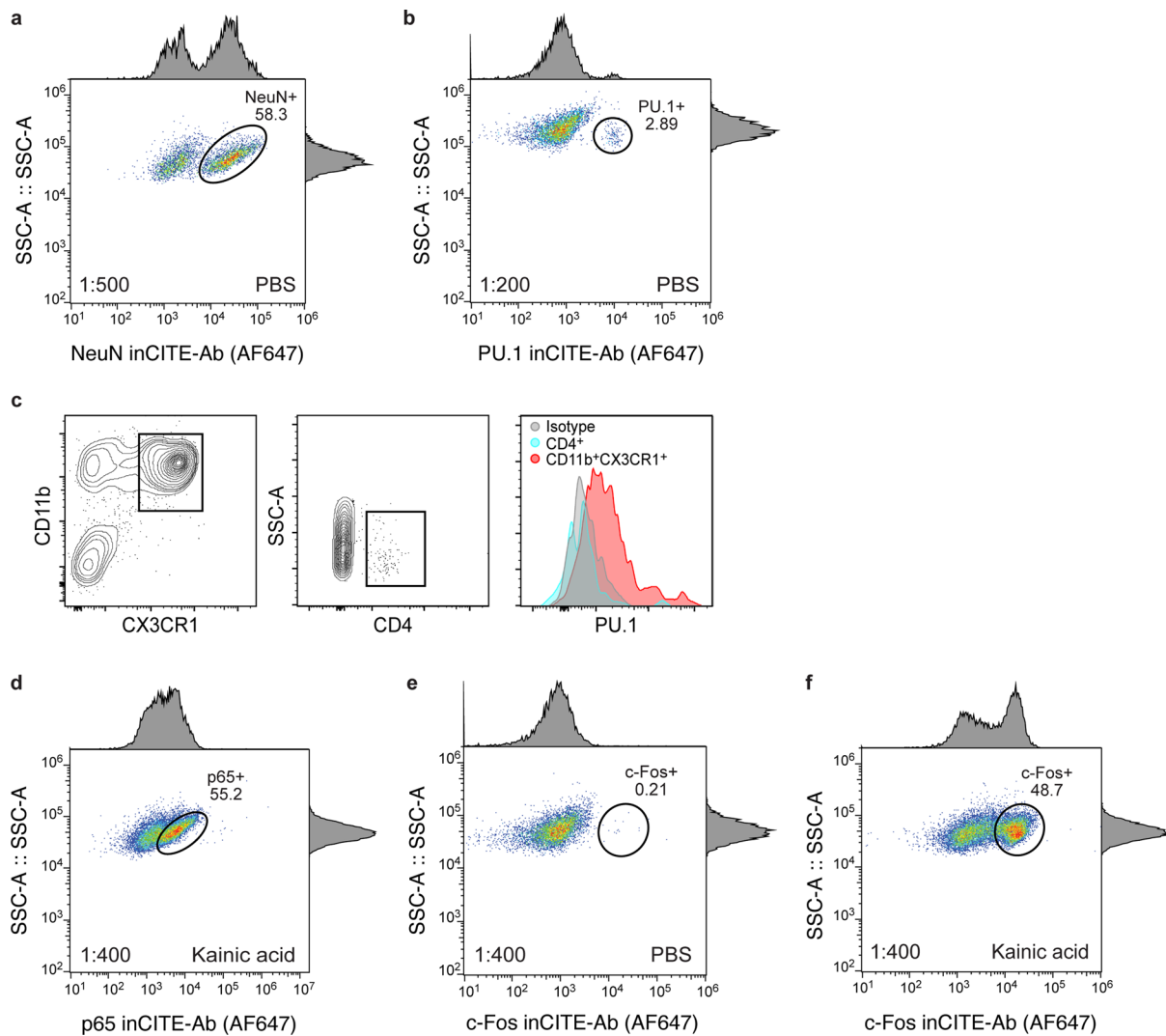
Peer review information *Nature Methods* thanks Dominic Grun, Bing Ren and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Lei Tang was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

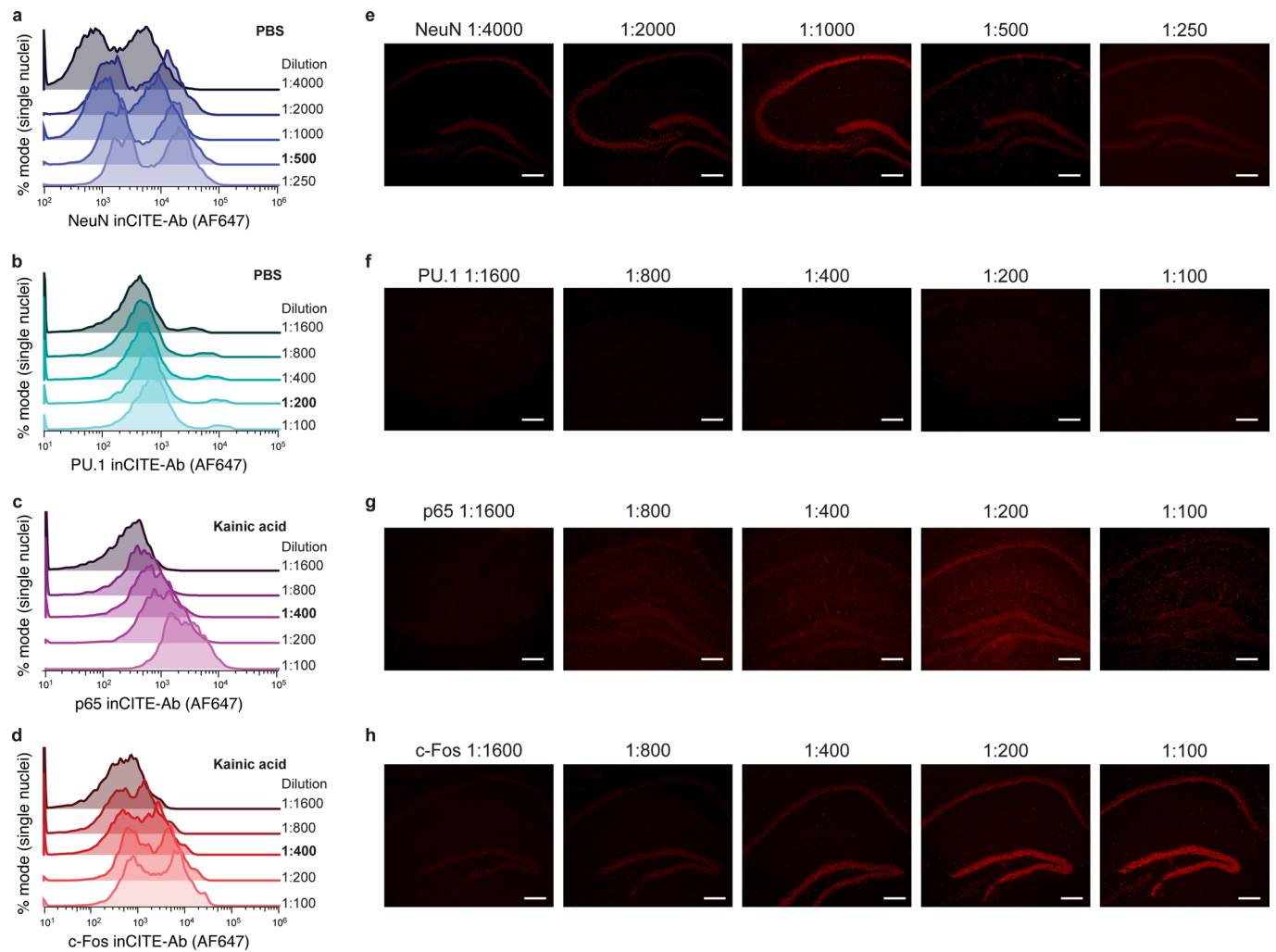


Extended Data Fig. 1 | See next page for caption.

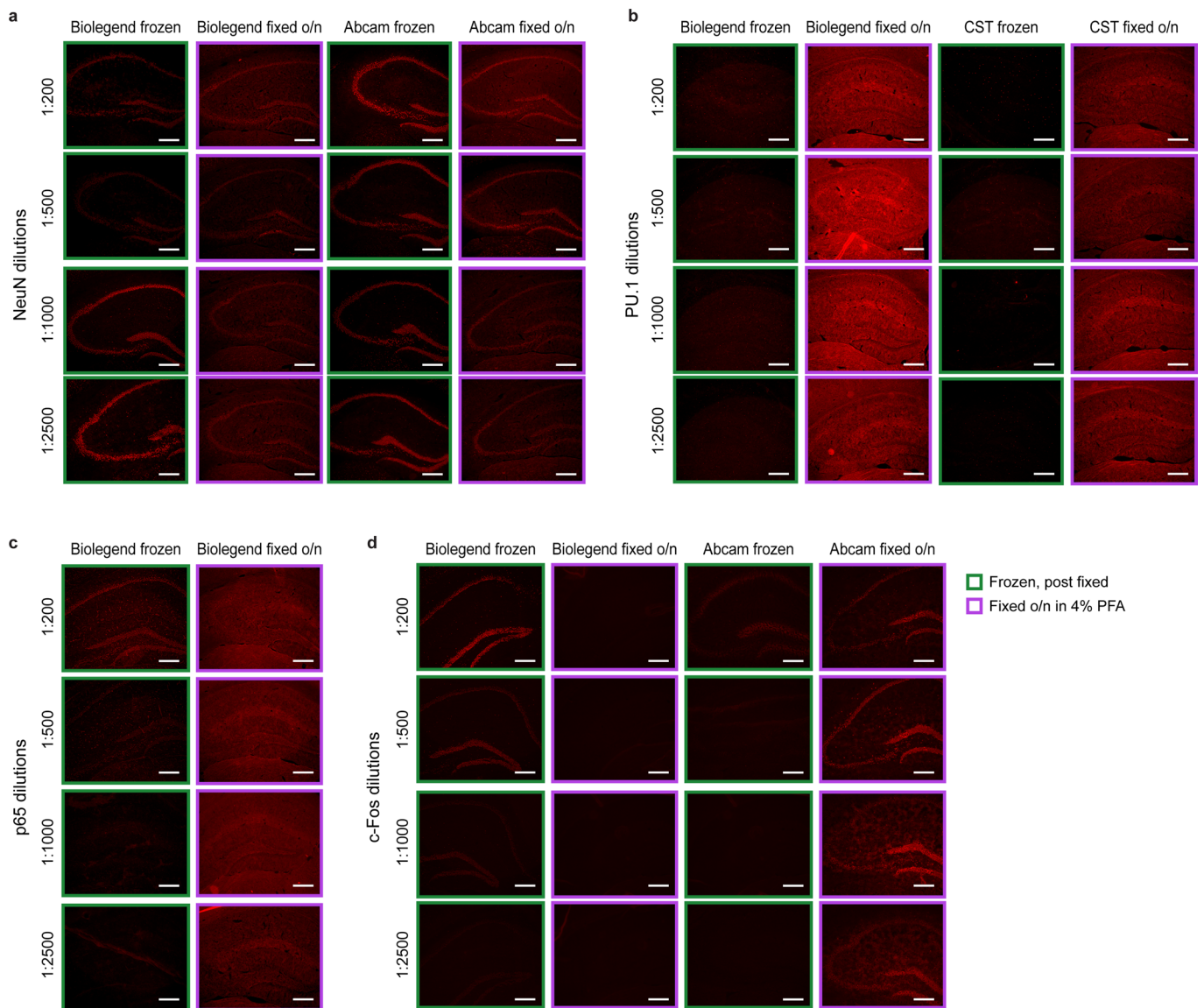
Extended Data Fig. 1 | Optimization of intranuclear antibody staining in HeLa cells. **a.** Nuclear p65 levels change after TNF α treatment, while total p65 in cells remains unchanged. Distribution of p65-mNeonGreen reporter fluorescence (x axis; % mode of singlet nuclei, y axis) measured by flow cytometry of nuclei (solid line) vs. cells (dashed line) from untreated ('NT', blue) or TNF α treated cells (red). **b.** Flow cytometry distinguishes p65-mNeonGreen signals across mixtures of NT and TNF α . Top: Flow cytometry measures of mNeonGreen^{high} fraction (x axis) match the input fraction of TNF α nuclei (x axis). Bottom: Corresponding high (red) and low (blue) mNeonGreen distributions. **c.** Immunofluorescence of nuclei smeared onto a slide after intranuclear p65 stain in suspension, showing complete antibody diffusion into the nucleus; representative of 3 experiments. Scale: 100 μ m. **d,e.** Comparing antibody- and fluorescence reporter-derived p65 levels. Antibody (from Alexa Fluor 647 secondary, y axis) and mNeonGreen (x axis) signal of p65 in an equal mixture of NT and TNF α stimulated nuclei. Histograms: marginal distributions. **d.** Agreement between unconjugated p65 antibody and mNeonGreen signal. **e.** No relationship between DNA-conjugated p65^{inCITE-Ab} and mNeonGreen signal using standard intranuclear staining buffer (pre-optimization). **f.** Relation between nuclei hashtag oligonucleotide (HTO; x axis) counts and p65 antibody-derived tag (ADT; y axis) counts, shown across 10,014 NT and TNF α nuclei, colored by the number of RNA UMIs. Top left: Pearson R² and associated P-value (two-sided t-test). To control for this relation, we normalize protein ADT counts by nuclei HTO counts (Methods). **g.** Comparing RNA complexity from inCITE-seq (fixed HeLa nuclei) and MULTI-seq (unfixed HEK nuclei, from McGinnis et al.⁴⁵) by the distribution of the number of detected transcripts (UMIs; top) and genes (bottom). **h.** Low correlation between p65 protein (y axis, nCLR) and *RELA* RNA levels (x axis, log normalized), with Pearson R² and associated P-value (two-sided t-test). Dots: nuclei colored by treatment (NT, blue; TNF α , red). **i.** Dynamics of gene expression after LPS stimulation in mouse dendritic cells, from Rabani et al.¹⁰, measured across time (x axis). Relative expression to steady state, t0 (y axis): pre-mRNA precursor (blue) and mRNA (red) for total (solid) vs. 4sU labeled (dashed) RNA, shown for *Rela* (top) and *Nfkb* (bottom), a p65 target as in Fig. 1e.



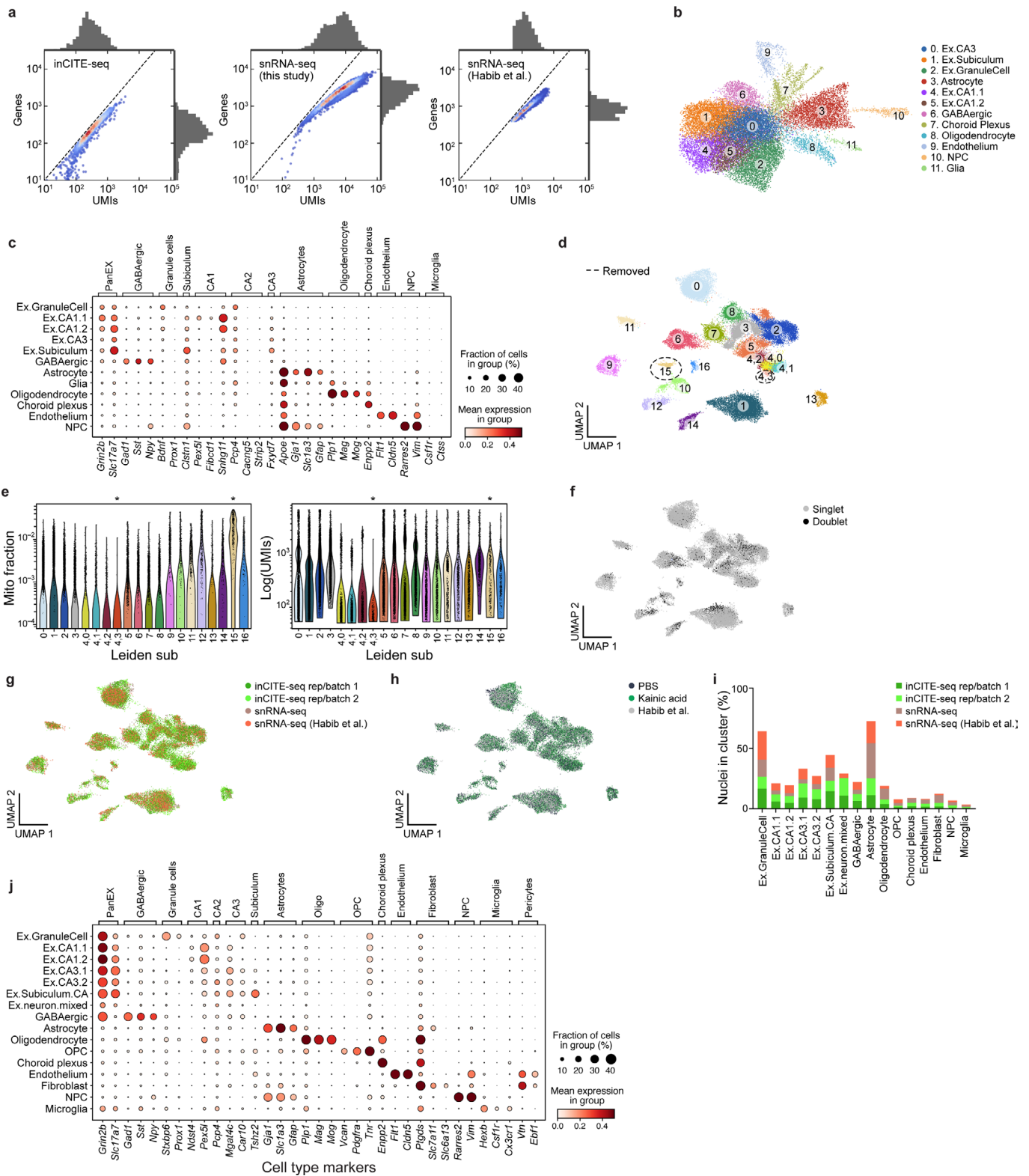
Extended Data Fig. 2 | Flow cytometry of inCITE targets on nuclei or cells extracted from frozen mouse hippocampus. Flow cytometry of nuclei populations from the mouse hippocampus after intranuclear stains with inCITE antibodies, followed by Alexa Fluor 647-conjugated secondary stain: NeuN in PBS (**a**), PU.1 in PBS (**b**), p65 in kainic acid (KA) (**d**), and c-Fos in PBS (**e**) and KA (**f**) treated mice. Axes show fluorescence signal (x axis) and side scatter (y axis) of singlet nuclei (dots); histograms show marginal distributions. Oval gates show NeuN^{high} (**a**, 58.3%), PU.1^{high} (**b**, <3%), p65^{high} (**d**, 55.2%), c-Fos^{high} (0.21% in PBS (**e**), and 48.7% after KA treatment (**f**)). **c**. Right: Distribution of PU.1 in microglia (CD11b⁺ CX3CR1⁺, red), CD4⁺ cells (blue) and isotype (gray) cells measured by flow cytometry (left and middle panels) after simultaneous surface protein and intracellular protein stains (Methods).



Extended Data Fig. 3 | Antibody signal varies across concentration regimes. Antibody stains of the mouse hippocampus (extracted nuclei or *in situ*) with inCITE antibodies across a wide range of dilutions, targeting NeuN in PBS (**a,e**), PU.1 in PBS (**b,f**), p65 in kainic acid (**c,g**), and c-Fos in kainic acid (**d,h**) treated mice. Antibody-derived fluorescence measured by Alexa Fluor 647-conjugated secondary antibody stain. **a-d**. Histograms are normalized as % mode of nuclei singlets. Antibody dilutions are indicated to the right of each axis, with dilutions used for inCITE-seq in bold (NeuN 1:500, PU.1 1:200, p65 1:400, c-Fos 1:400). **e-h**. *In situ* immunofluorescence of frozen mouse hippocampus with inCITE antibodies across different dilutions, matching the concentrations used in flow cytometry; representative of 2 independently conducted experiments. Scale bars, 100 μm .

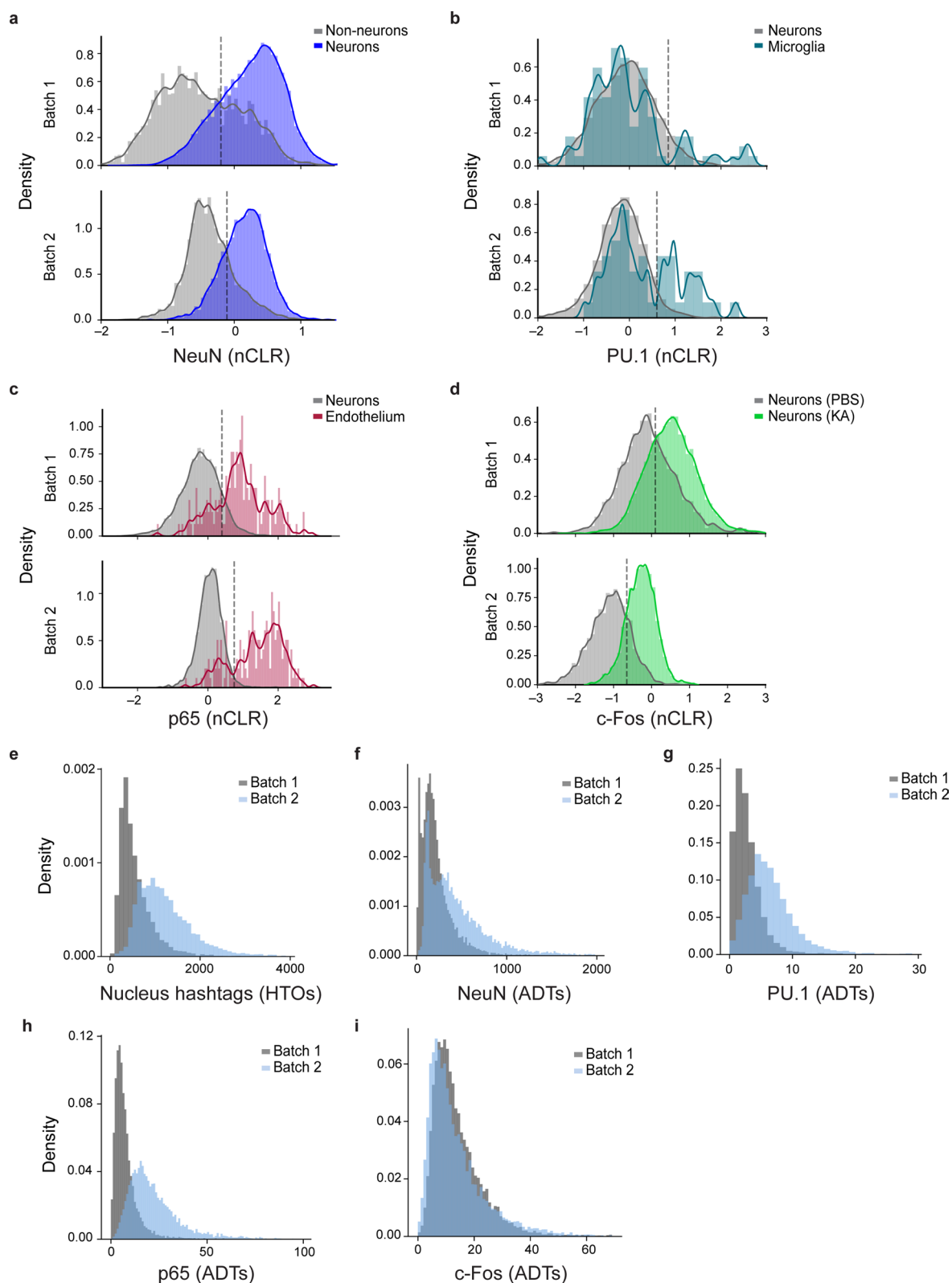


Extended Data Fig. 4 | Impact of tissue preparation on epitope detection by antibodies. Comparing *in situ* immunofluorescence of antibody stains (followed by Alexa Fluor 647-conjugated secondary stain) in mouse hippocampus tissue that were immediately frozen (green box) or frozen after overnight fixation in 4% PFA (purple box, Methods) across a wide range of antibody dilutions. Images are representative of 2 independent experiments. **a.** NeuN in PBS. Biologend NeuN antibody (clone 1B7) used for inCITE and Abcam NeuN antibody (clone EPR12763). **b.** PU.1 in PBS. Biologend PU.1 antibody (clone 7C2C34) used for inCITE and Cell Signaling Technology PU.1 antibody (clone 9G7). **c.** p65 in KA. Biologend p65 antibody (clone Poly6226) used for inCITE. **d.** c-Fos in KA treated mice. Biologend c-Fos antibody (clone Poly6414) used for inCITE and Abcam c-Fos antibody (ab190289). Scale bars, 100 μ m.

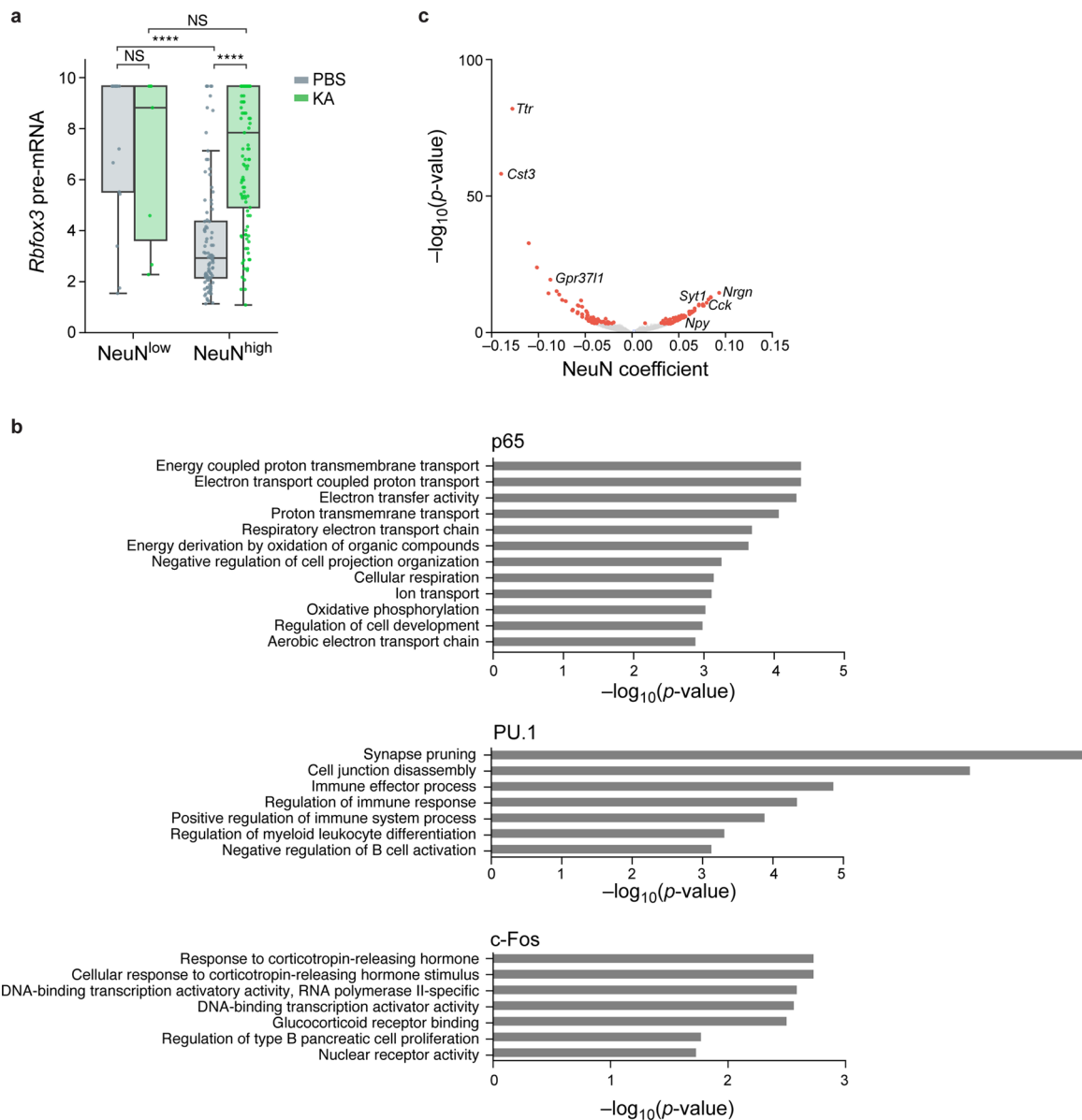


Extended Data Fig. 5 | See next page for caption.

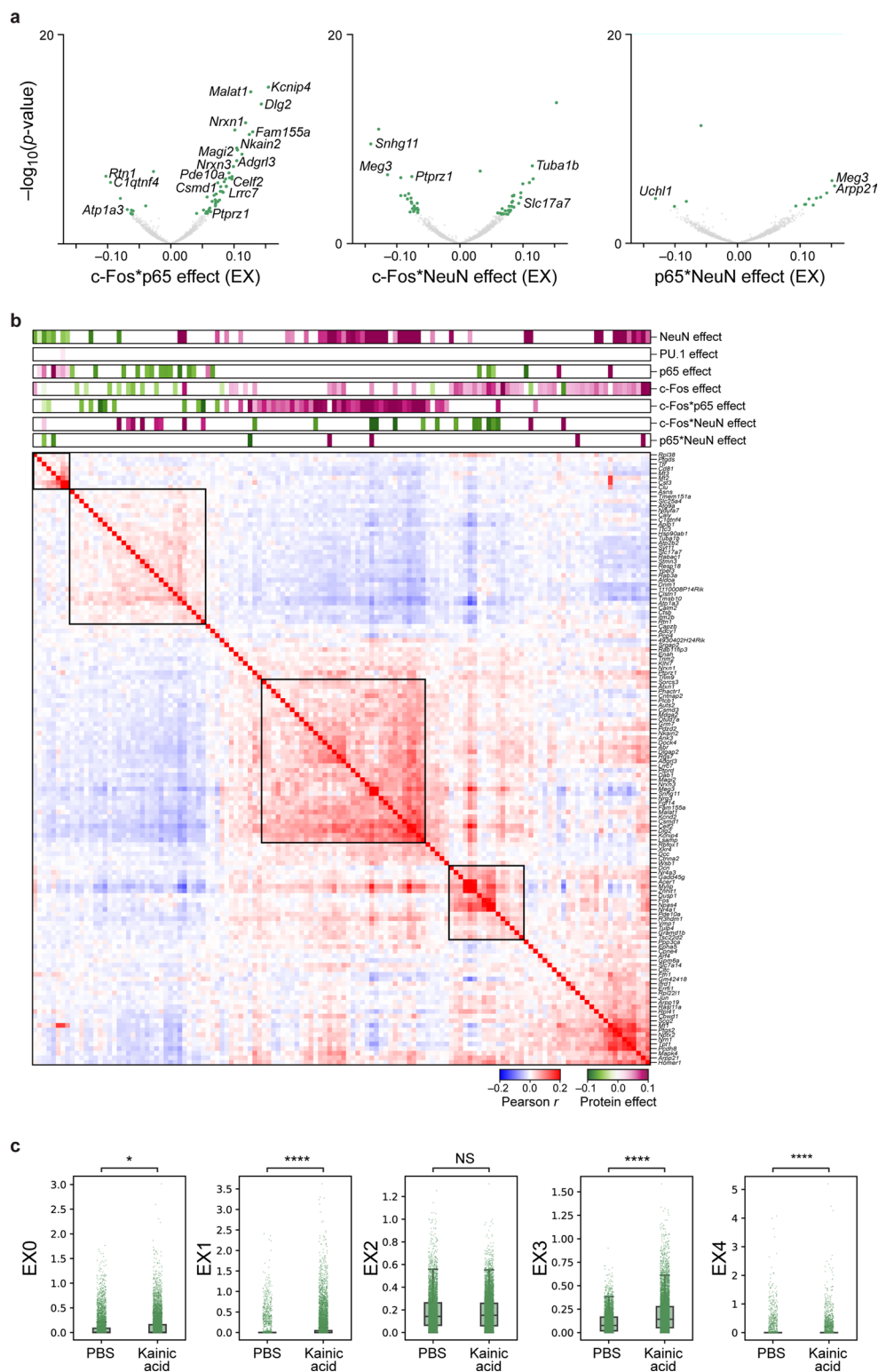
Extended Data Fig. 5 | Comparing and combining single nucleus RNA profiles from inCITE-seq and snRNA-seq of mouse hippocampus. **a.** Comparing the complexity of RNA profiles from inCITE-seq and standard snRNA-seq of the mouse hippocampus. Distributions (marginals) of the number of UMIs (x axis) and genes (y axis) from inCITE-seq (left), matching mouse hippocampus snRNA-seq in this study (middle), and previously published snRNA-seq (right). Scatter plot shows the density of individual nuclei (dots) calculated with a Gaussian kernel estimate. **b,c.** Major cell types from the adult mouse hippocampus identified from inCITE-seq RNA profiles alone. **b.** UMAP embedding of 24,444 single nucleus inCITE-seq RNA profiles (dots) colored by annotated cluster (number). **c.** Expression of marker genes (columns) used for annotating cell type clusters (rows), showing mean expression of log normalized counts (dot color) and proportion of expressing cells (dot size). **d-j.** Enhanced cell type distinctions and annotation by combining RNA profiles from inCITE-seq and snRNA-seq. Joint UMAP embedding of 22,260 inCITE-seq and 15,507 snRNA-seq RNA profiles (dots) colored by unsupervised leiden clusters or subcluster of leiden group 4 (numbers) (Methods). **e.** Distribution of mitochondrial fraction of total gene content (y axis, left) and total transcript counts (y axis, right) in each leiden cluster or subcluster of leiden group 4 (x axis, both). Asterisks indicate cluster 15 ($n=327$ nuclei) and subcluster 4,3 ($n=179$ nuclei) that were removed for high mitochondrial content and for low RNA complexity, respectively. **f-h.** UMAP embedding as in Fig. 2d colored by doublets that were removed from subsequent analyses ($n=3,059$ doublets, **f**), batch and assay (**g**), or condition (**h**). **i.** Percent of nuclei (y axis) from each batch/assay (color) in each cluster (x axis). **j.** Mean expression of log normalized counts (dot color) and proportion of expressing cells (dot size) of marker genes (columns) used for annotating cell type clusters in **d** (rows).



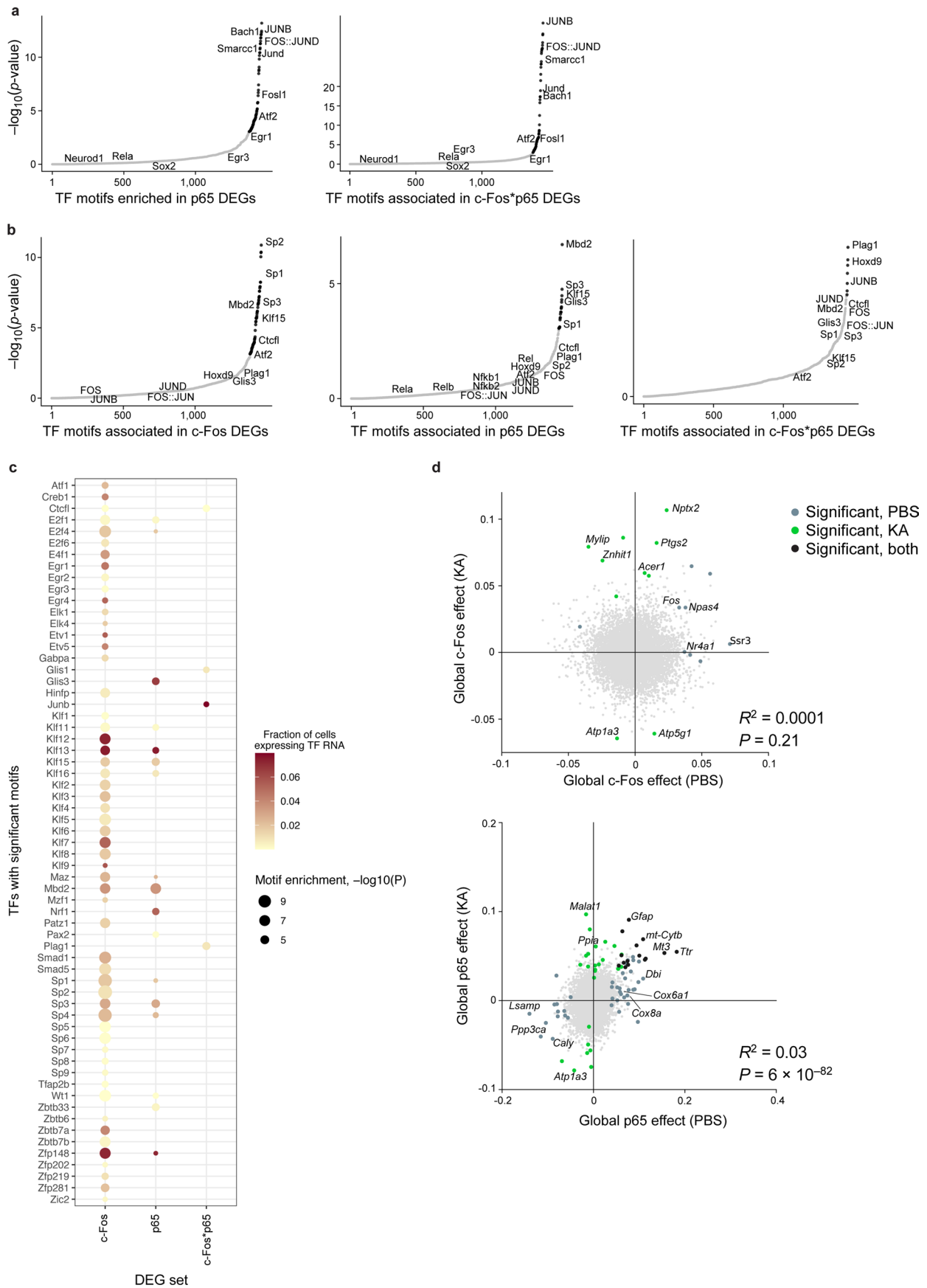
Extended Data Fig. 6 | Protein levels by inCITE-seq batch (replicate). **a-d.** Distribution of protein levels (x axis, nCLR) shown as kernel density estimates of NeuN (**a**), PU.1 (**b**), p65 (**c**), or c-Fos (**d**) in each batch (top: batch 1; bottom: batch 2) in biologically relevant subsets as foreground (color) and appropriate background set of nuclei (grey). Dashed line: Batch-specific threshold used to partition protein level as high vs. low. **e-i.** Density distribution of (**e**) nucleus hashtag counts (x axis, HTOs) or (**f-i**) antibody-derived tags (x axis, ADTs) of inCITE target proteins, colored by batch (batch 1, gray; batch 2, blue).



Extended Data Fig. 7 | Protein effects on global gene expression. **a**, Relation between unspliced pre-mRNA expression of *Rbfox3* and nuclear protein levels of NeuN. Distribution of pre-mRNA levels (Z score of log-normalized counts, y axis) in nuclei with high or low levels of NeuN (x axis) after PBS (gray) or KA (green) treatment (NeuN thresholds in Extended Data Fig. 6). Boxplots show the median (centre line), box bounds represent first and third quartiles, and whiskers span from each quartile to the minimum or the maximum (1.5 interquartile range below 25% or above 75% quartiles). Dots correspond to 227 individual nuclei with non-zero pre-mRNA levels measured across $n = 2$ biologically independent samples. Significance, from left: $P = 5 \times 10^{-15}$, $P = 9 \times 10^{-5}$ two-sided Mann-Whitney test. NS - not significant. **b**, Functional gene sets enriched in TF associated genes. Enrichment ($-\log_{10}(P\text{-value})$, x axis, hypergeometric test) of Gene Ontology (GO) terms (y axis) in genes significantly associated (from top to bottom) with p65 (33 genes), PU.1 (13 genes), and c-Fos (10 genes). **c**, Genes associated with NeuN. Effect size (x axis) and associated significance (y axis, $-\log_{10}(P\text{-value})$) for the association of each gene (dots) with NeuN by a model of gene expression as a linear combination of the four inCITE-seq target proteins after regressing out treatment and cell type (Methods). Select genes are labeled. Colored dots: Benjamini-Hochberg FDR < 5%.



Extended Data Fig. 8 | Genes and modules associated with TFs within excitatory (EX) neurons. a. Genes associated with protein-protein pairs in the interaction model, identified by modeling gene expression across excitatory neurons as a linear combination of individual proteins and their pairwise interactions after regressing out treatment. Effect size (x axis) and significance (y axis, $-\log_{10}(\text{P-value})$) for DEGs (dots) associated with each protein-protein interaction term: p65 and c-Fos (left), c-Fos and NeuN (middle), and p65 and NeuN (right). Colored dots: Benjamini-Hochberg FDR < 5%. **b.** Pearson correlation coefficient (red/blue colorbar) of pairwise gene expression profiles (rows and columns) significantly (FDR < 5%) associated positively (purple) or negatively (green), with c-Fos (additive model), p65 (additive model), or c-Fos*p65 (interaction model), ordered by hierarchical clustering. Top bars: Effect size of each protein or protein-protein pair. **c.** Treatment effect on gene programs. Program scores (y axis) for 5 EX programs (in Fig. 4f) of 15,226 individual nuclei (dots) from PBS or KA treated mice (x axis) measured across 2 biologically independent experiments. Boxplots show the median (centre line), box bounds represent first and third quartiles, and whiskers span from each quartile to the minimum or the maximum (1.5 interquartile range below 25% or above 75% quartiles). Significance, from left: $P = 0.049$, $P = 2.7 \times 10^{-271}$, $P = 2.2 \times 10^{-199}$, $P = 6.1 \times 10^{-7}$, two-sided Mann-Whitney test. NS - not significant.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Treatment-dependent *cis*-regulatory elements and TF-associated genes. a-c. Prediction of co-regulatory patterns by TF motif enrichment in DEGs associated with c-Fos or p65 (additive model), or their interaction c-Fos*p65 (interaction model). **a,b.** Significance ($-\log_{10}(P\text{-value})$, y axis) and rank order (x axis) of TF motifs (dots) enriched in enhancers of DEGs associated with each protein (additive model) or protein-protein (interaction model) term in excitatory neurons, using enhancers of PBS (**a**) or KA (**b**) treated sample as background. Black: significant motifs ($P < 10^{-3}$, hypergeometric test); gray: not significant. **c.** TF motif enrichment (columns; dot size, $-\log_{10}(P\text{-value})$) and proportion of excitatory neuron nuclei expressing the RNA (color) of significant TFs (rows) in the enhancers of c-Fos (additive model), p65 (additive model), or c-Fos*p65 (interaction model) DEGs, compared to other enhancers within the KA treated sample. **d.** Treatment-dependence of gene association with c-Fos and p65. Global effect size of genes (dots) associated with c-Fos (left) and p65 (right), after PBS (x axis) or KA treatment (y axis) (Methods). Colored dots: genes with significant coefficients (Benjamini-Hochberg FDR < 5%) in PBS (gray), KA (green), or both (black). Select genes are labeled. Bottom right: linear correlation R^2 and associated P value (two-sided t-test).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-----|-----------|
| n/a | Confirmed |
|-----|-----------|
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
 - A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
 - The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
 - A description of all covariates tested
 - A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
 - A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
 - For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
 - For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
 - For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
 - Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Sequencing data was collected on a NextSeq500

Data analysis Cumulus v1.0, Cellranger v4.0.0, DemuxEM v0.1, Python 3.7.9, scanpy 1.7.2, Anndata 0.7.6, Leidenalg 0.8.4, harmony-pytorch 0.1.4, numpy 1.20.2, matplotlib 3.4.1, seaborn 0.11.1, pandas 1.2.4, velocyto v0.17.17, statsmodels 0.12.2, sklearn 0.24.2.

Jupyter notebooks with custom Python code and the complete list of used python packages is available on GitHub: <https://github.com/klarman-cell-observatory/inCITE-seq>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw (BAM files) and processed data (gene and protein count matrices) are available at GEO accession GSE163480.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were conducted for either the HeLa or mouse hippocampus experiment. As we sequenced tens of thousands of single nuclei for each experiment, we determined that these measurements provided a robust sample size to discern cell state and cell type-specific differences in protein levels as a proof of principle.
Data exclusions	No data were excluded from the study.
Replication	All technical replicates of inCITE antibody stains were successful, evaluated by flow cytometry to verify clean signal separation after intranuclear stain with DNA-conjugated antibodies. We conducted roughly 12 technical replicates for NT vs. TNF α treated HeLa cells, and 2 technical replicates for the mouse hippocampus (not including the data used in the paper).
Randomization	HeLa populations seeded on two separate dishes were randomly selected to be treated with TNF α or no treatment. 8-week-old male mice were randomly designated to PBS and kainic acid treatment.
Blinding	Blinding was not possible, as the operators conducting the experiments and the analyses were the same individual.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	p65 (Biolegend Poly6226, cat #622601), NeuN (Biolegend 1B7, cat #834502), c-Fos (Biolegend Poly6414, cat #834502), PU.1 (Biolegend 7C2C34, cat #681307), CD31 (eBioscience 390, cat #14-0311-82), CD4-BUV395 (BD GK1.5, cat #563790), CD45.2-FITC (eBioscience 104, cat #11-0454-82), CD11b-BV421 (BioLegend M1/70 cat #101235, CX3CR1-APC (BioLegend SA011F11 cat #149008), rat IgG2a-PE isotype (Biolegend RTK2758, cat #400507), anti-rabbit Alexa Fluor 647 secondary (Invitrogen A27040), donkey anti-rabbit AF647 (Jackson ImmunoResearch #711-605-152), donkey anti-rabbit AF594 (Jackson ImmunoResearch #711-585-152), donkey anti-mouse AF488 (Jackson ImmunoResearch #715-545-150), donkey anti-rat AF647 (Jackson ImmunoResearch #712-605-153).
Validation	All primary antibodies used in this study were established clones validated by the vendor (provider). We provide additional validation of each antibody used for inCITE-seq by the following: p65 signal matching that of the internal mNeonGreen reporter in HeLa cells, NeuN antibody signal enriched in neuronal clusters determined by sequencing, microglia-specific PU.1 antibody signal verified by flow cytometry with CX3CR1, and c-Fos flow cytometry signal that match widespread neuronal excitation measured by immunohistochemistry of KA treated mice.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s) Gift from Lab of Iain Cheeseman, MIT (McKinley and Cheeseman, Developmental Cell 2017).

Authentication	Cell lines were not authenticated by the authors in this study, as only the species and the signal from p65-mNeonGreen reporter were important for the study. We confirmed that these HeLa cells were human-derived, based on the aligned reads from single nucleus RNA sequencing.
Mycoplasma contamination	Cell lines were tested monthly for mycoplasma contamination and always tested negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	C57BL/6J (Jax 000664) mice were purchased from The Jackson Laboratory and bred in-house. Male mice were used at ~8 weeks of age. All mice were maintained under SPF conditions on a 12-h light–dark cycle, at ambient temperature 21.5 + 1°C and relative humidity between 30% and 70%, and provided food and water ad libitum. All mouse experiments were approved by, and performed in accordance with, the Institutional Animal Care and Use Committee guidelines at Weill Cornell Medicine.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All mouse experiments were approved by, and performed in accordance with, the Institutional Animal Care and Use Committee guidelines at Weill Cornell Medicine.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Nuclei extractions from HeLa and the mouse hippocampus were isolated using the EZ prep lysis buffer and glass douncers (elaborated in Methods). Nuclei suspensions were lightly fixed, blocked, then stained with antibodies.
Instrument	Flow cytometry data was collected on a Beckman CytoFLEX Analyzer. No sorting was involved.
Software	Data were collected with CytExpert on the CytoFLEX Analyzer, and subsequently analyzed with FlowJo v10.7.1.
Cell population abundance	No sorting was involved in this study.
Gating strategy	FSC/SSC values were first gated to exclude debris. DAPI signal was used to identify and gate on singlet nuclei peaks. Subsequent flow cytometry analyses were conducted on only single nuclei subpopulations.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.