Pooled CRISPR screens with joint single-nucleus chromatin accessibility and transcriptome profiling

Rachel E. Yan[1,2,3,4], Alba Corman[1,2], Lyla Katgara[1,2], Xiao Wang[1,2], Xinhe Xue[1,2], Zoran Z. Gajic[1,2], Richard Sam[1,2], Michael Farid[1,2], Samuel M. Friedman[1,2], Jungwook Choo[1,2], Ivan Raimondi[1,6], Shridar Ganesan[3], Eugene Katsevich[5], Jeffrey P. Greenfield[4], Nadia Dahmane[4], and Neville E. Sanjana[1,2,*]

[1] New York Genome Center, New York, NY, USA
[2] Department of Biology, New York University, New York, NY, USA
[3] Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, USA
[4] Department of Neurological Surgery, Weill Cornell Medicine, New York, NY, USA
[5] Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, PA, USA
[6] Present address: Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA

* Correspondence: Neville E. Sanjana (neville@sanjanalab.org)

Abstract

Pooled single-cell CRISPR screens have profiled either gene expression or chromatin accessibility, but not both modalities. Here, we develop MultiPerturb-seq, a high-throughput CRISPR screening platform with joint single-nucleus chromatin accessibility, transcriptome, and guide RNA capture using combinatorial indexing combined with droplet microfluidics to scale throughput and integrate all three modalities. We identify key differentiation genes in a rare pediatric cancer and establish ZNHIT1 as a potential target for reprogramming therapy.

Main text

Recent advances in single-cell perturbation screens have enabled scalable profiling of rich cellular states and phenotypes, particularly with transcriptional phenotypes[1, 2]. Several groups have developed methods that expand single-cell perturbation screens to capture modalities such as protein[3, 4], chromatin accesibility[5-7], and 3D genome conformation[8]. These single-cell screens have included a diverse array of genetic perturbations, including knockout using Cas9 nuclease, transcriptional modulation using CRISPRi and CRISPRa, targeting of RNA using Cas13, precise variant insertion via HDR or base-editing, and overexpression with open-reading frame (ORF) libraries[9].

Here, we introduce MultiPerturb-seq, a method that links pooled CRISPR perturbations with single-cell open chromatin (ATAC-sequencing) and gene expression (RNA-sequencing) profiles at scale (**Fig. 1a, Supplementary Fig. 1**). We then apply this method to drive mechanism-based discovery of differentiation regulators for a rare pediatric brain cancer, atypical teratoid/rhabdoid tumor (AT/RT). While cancer reprogramming therapy (i.e. differentiation therapy) has been curative for patients with malignancies such as acute promyelocytic leukemia[10], success has been limited in other cancers due to a lack of high-throughput methods to identify reprogramming targets. In MultiPerturb-seq, open chromatin provides a broad overview of epigenetic state, capturing many levels of gene regulation, while gene expression provides a robust view of cell state and developmental stage. Together, they link CRISPR perturbations with cell states and putative mechanisms of action for transcriptional reprogramming. We also sought to reduce reagent cost and labor: Recent genome-wide single-cell perturbation screens have required ~100 lanes of commercial single-cell library preparation kits[2]. In MultiPerturb-seq, we combine combinatorial indexing and droplet microfluidics to scale throughput[11-13] — loading 100,000 cells on a single 10X Chromium ATAC lane — which results in significant cost advantages over existing uni- and multimodal single-cell perturbation approaches (**Fig. 1b**).

After cloning CRISPR guide RNA (gRNA) libraries into lentiviral vectors and producing virus, we transduced mammalian cells that already express a second-generation CRISPR repressor[14] at a low multiplicity-of-infection (~0.05) to achieve one guide per cell and selected cells receiving a CRISPR perturbation (**Supplementary Fig. 2**). We waited 7 days to ensure sufficient time for protein depletion and then collected cells for MultiPerturb-seq library preparation (**Supplementary Fig. 3, Supplementary Protocol**). After nuclei isolation and distribution into wells, we tagmented open chromatin using barcoded transposomes (**Supplementary Fig. 4a,**

b)[6]. Next, we performed reverse transcription with a mix of poly-dT and CRISPR gRNA-specific primers and barcoded template switch oligonucleotides (TSO) with matching barcodes (**Supplementary Fig. 4c-f, Supplementary Table 1**). We then pooled cells for second-round barcoding via droplet microfluidics using 10X Chromium ATAC gel beads. Lastly, ATAC, RNA, and CRISPR gRNA libraries were amplified and prepared for sequencing (**Fig. 1c**, **Supplementary Fig. 4g-j, Supplementary Fig. 5a-c**).

To quantify single-cell isolation in MultiPerturb-seq, we performed a species-mixing experiment with 80% human (BT16) and 20% mouse (3T3) cells, and robustly captured ATAC, RNA, and gRNA molecules (**Fig. 1d-g, Supplementary Fig. 6a-d**). We quantified the percent of barcode combinations which contained a mixture of mouse and human fragments (collisions in cell assignment) for each of the three modalities captured. We achieved low barcode collision rates for RNA (6.2%), ATAC (11.6%) and gRNA (6.6%) libraries, despite loading ~10-fold more cells than the standard for a 10x Chromium ATAC lane. We achieved robust detection of expressed genes, open chromatin peaks, and gRNAs (**Fig. 1h, Supplementary Fig. 6e-h**). For the ATAC, we observed characteristic open chromatin enrichment around transcriptional start sites (**Fig. 1i, Supplementary Fig. 6e**) and, for the RNA, we found low mitochondrial reads (**Supplementary Fig. 6f**). The majority of cells only had one gRNA detected and decreased expression of the targeted gene when compared to cells receiving a non-targeting gRNA: 78% of high-quality cells were assigned gRNA identities (**Fig. 1j, Supplementary Fig. 6g,h**). Notably, this does not require the use of any modified CRISPR plasmids or specialized bead oligonucleotides. We also found similar or better RNA and ATAC capture compared to other single-cell RNA-seq and single-cell ATAC-seq technologies, including increased unique molecular identifiers (UMIs) and genes per cell (**Supplementary Fig. 6i-l**), as well as increased ATAC fragments and peaks per cell (**Supplementary Fig. 6m-p**)[6, 11, 15-18].

Though it is not compatible with barcoded superloading, we also utilized the 10X Chromium Multiome kit and the specialized guide RNA plasmid, CROP-seq[19, 20] as an alternate method of multi-modal capture and performed a lower-throughput version of a multiomic CRISPR screen (~10,000 vs. ~100,000 cells per lane) (**Supplementary Fig. 7a**), which we termed CROP-Multiome. Reassuringly, gene expression changes after perturbation were highly correlated between MultiPerturb-seq and CROP-Multiome, supporting the validity of the results on both platforms (**Supplementary Fig. 7b-e**). However, MultiPerturb-seq outperformed CROP-Multiome along several important dimensions, including better gRNA capture (**Supplementary Fig. 7f**) and

higher RNA UMIs per cell (**Supplementary Fig. 7g**), RNA genes per cell (**Supplementary Fig. 7h**), ATAC fragments (**Supplementary Fig. 7i**), and ATAC peaks per cell (**Supplementary Fig. 7j**). Given these differences and the additional advantages of 10-fold increased cell loading, direct guide RNA capture without a specialized plasmid, and 5' capture, we used the MultiPerturb-seq data for all subsequent analyses.

The combination of ATAC and RNA modalities allowed us to detect perturbation-linked changes in open chromatin and gene expression. Despite the sparsity of the single-cell data, we were able to see clear patterns when examining individual genes and groups of genes with shared function. For example, after knockdown of histone methyltransferases (*DOT1L, EHMT2, KDM1A, KDM6A, KMT2B, KMT2D, MECOM, MLLT1, PRDM16, PRMT5, SETD2, SETD5, SETDB1, SUV39H2*), we found increases in open chromatin at the *RFX3* locus and increased *RFX3* gene expression (**Fig. 1k**). We also were able to identify perturbation-specific changes: After knockdown of histone variant *H3F3A*, we found the opposite at the *PPM1B* locus, where we observed decreased chromatin accessibility and expression of *PPM1B* (**Fig. 1l**).

We next sought to apply MultiPerturb-seq to a rare pediatric central nervous system cancer, AT/RT, which is driven by a change in chromatin remodeling. In AT/RT, biallelic loss of *SMARCB1* — an essential subunit of the SWI/SNF chromatin remodeling complex, which is one of the most commonly mutated protein complexes in cancer[21] — prevents complete differentiation of progenitors and drives tumor proliferation[22]. AT/RT is extremely aggressive, and no AT/RT-specific therapies are available: The current standard-of-care is high dose radiation and chemotherapy with autologous stem cell transplant[23]. Despite these intensive (and toxic) therapies, the disease is still nearly uniformly fatal with a median overall survival of four years[23]. Due to the loss of SMARCB1, AT/RT are dependent on alternate epigenetic regulators, such as polycomb[24-26], and SMARCB1-null embryonic stem cell models fail to differentiate into neurons due to altered gene regulation[27]. Therefore, using MultiPerturb-seq, we targeted ~100 epigenetic remodelers in human AT/RT cells (BT16) and sought to discover whether knockdown of specific remodelers can ameliorate the dysfunctional epigenome in AT/RT and restore differentiation (**Fig. 2a**).

Because AT/RT may arise from a variety of lineages, including non-neural lineages[28], we first compared the MultiPerturb-seq transcriptomes to reference developmental and adult atlases of multiple human tissues[29] (cortex, cerebellum, kidney, ovary, testis, and liver) and found the

highest overall similarity with brain cortical tissue (**Supplementary Fig. 8**). To assess the impact of perturbations on differentiation, we measured the correlation in transcriptomic profiles between gene-perturbed cells and primary tissues from different brain developmental stages (**Fig. 2b**). Compared to negative control (non-targeting) perturbations, we found a subset of perturbations with transcriptomes that had greater similarity to late brain stages rather than early ones, such as *ZNHIT1, CTCF, GATAD2B*, and others. These tended to express higher levels of genes correlated with neural differentiation such as *CCND3*[30], *GPM6B*[31], and *SYNJ2*[32, 33] (**Supplementary Fig. 9**).

The chromatin landscape in AT/RT is unusual with broad changes due to loss of SMARCB1, where residual SWI/SNF complexes cannot maintain accessibility to enhancers needed for differentiation[34]. To further focus our analysis, we leveraged the multimodal nature of our assay to find epigenetic remodeler perturbations that may help normalize the AT/RT chromatin landscape (**Fig. 2c**). Using recent ATAC-seq atlases from primary fetal[35] and adult[36] brain tissues, we sought to identify perturbations resulting in open chromatin profiles with greater correlation to mature brain tissue, and found that perturbations of *ZNHIT1, MECOM, CERC2, TRRAP*, and others led to genome-wide chromatin profiles that were more similar to tissue from postnatal brain than fetal brain (**Fig. 2c, Supplementary Fig. 10a**). We also examined ENCODE *cis*-regulatory elements (CREs)[37] and found a greater number of our perturbations triggered changes in chromatin accessibility at promoters with fewer perturbations acting at enhancers (**Supplementary Fig. 10b-f**). Furthermore, when grouping target genes by complex, we found that knockdown of repressor complex (LSD-CoREST/BHC) subunits (*HDAC1, HDAC2, RCOR1*) tended to increase accessibility at ENCODE CREs, while knockdown of CERF complex subunits (*CERC2, SMARCA1*) tended to decrease accessibility (**Supplementary Fig. 10g**).

Next, we computed differentiation scores for gene expression (RNA) and open chromatin (ATAC) that captured relative similarity to postnatal versus prenatal brain tissues (see *Methods*) (**Fig. 2d-e**). Interestingly, we found that RNA and ATAC differentiation score was not always correlated (**Fig. 2f**). For example, we found that most perturbations of BAF complex members led to high ATAC differentiation and low RNA differentiation scores, suggesting that loss of residual BAF complexes can reshape/restore the chromatin landscape but that these perturbations are not sufficient to differentiate cells (**Supplementary Fig. 10h**).

After examining both differentiation scores, we identified multiple genes with high RNA and ATAC differentiation scores and subsequently focused on *ZNHIT1*, which was the top-ranked gene perturbation for joint ATAC and RNA differentiation score (**Fig. 2f**). ZNHIT1 is a subunit of the SRCAP (SNF-2 related CBP activator protein) complex, which is an INO80 family complex that mediates ATP-dependent exchange of histone H2A.Z, leading to chromatin remodeling and transcriptional modulation (**Supplementary Fig. 11a**). ZNHIT1 has previously been shown to maintain stemness in intestinal stem cells by promoting H2A.Z incorporation[38]. *ZNHIT1* knock-down induced large changes at multiple regulatory elements, including promoters and enhancers, with increased transcriptomic similarity to postnatal — and specifically adult — brain tissues. (**Fig. 2g, Supplementary Fig. 10b-e**)**.** To identify potential mechanisms of action, we examined differentially accessible chromatin in *ZNHIT1*-perturbed cells compared to non-targeting controls. We found that *ZNHIT1* perturbation led to changes in accessibility near genes involved in neuronal differentiation and axonogenesis (**Supplementary Fig. 11b**), as well as increased expression of genes for neuron projection development, cell polarity, and cell growth (**Supplementary Fig. 11c**).

Given the broad changes in chromatin organization and more differentiated transcriptional state upon *ZNHIT1* loss, we wondered whether ZNHIT1 inhibition may be a good candidate to push AT/RT cells toward terminal differentiation. We cloned individual CRISPR guide RNAs to target *ZNHIT1* and measured stemness, proliferation, and expression of differentiated neuronal markers (**Fig. 3a**). Using intracellular antibody labeling and flow cytometry, we found diminished expression of the pluripotency-associated transcription factor SOX2 after knockdown of *ZNHIT1*, compared to non-targeting guide RNA controls (**Fig. 3b**, **Supplementary Fig. 11d,e**). Additionally, the central goal of an AT/RT reprogramming therapy is cessation of cellular proliferation. Because cell cycle arrest occurs during G1, preventing progression to S-phase, we evaluated the relative proportion of cells in S-phase (**Fig. 3c**). We examined genes classified as cell cycle markers[39] and found that *ZNHIT1* perturbation led to a 19% decrease in expression of S-phase genes compared to non-targeting controls. We confirmed this by assaying changes in proliferation via incorporation of the thymidine analogue 5-ethynyl-2′-deoxyuridine (EdU) after a 30 minute pulse and found that *ZNHIT1* knockdown decreased progression through S-phase by 43% relative to non-targeting controls (**Fig. 3d**). Perturbation of related proteins (SRCAP complex co-factor YEATS4 and H2A.Z acetylase KAT5) resulted in similar decreases in EdU incorporation, suggesting that other SRCAP members and enzymes involved in H2A.Z biogenesis are required for normal cell cycle progression (**Supplementary Figure 11f,g**).

In the MultiPerturb-seq data, we also found that target genes of the transcription factor ATOH8 had increased expression in *ZNHIT1*-perturbed cells (~9-fold increase), compared to cells receiving a non-targeting guide RNA (**Fig. 3e**). ATOH8 expression promotes neuronal differentiation and supports neuronal functions[40]. To confirm these findings, we performed immunocytochemistry for ATOH8 in ZNHIT1-perturbed cells and found that *ATOH8* expression was increased (**Fig. 3f**). We also observed increases in early (TUJ1) and more mature (MAP2) neuronal markers in *ZNHIT1*-perturbed cells, further supporting a role for ZNHIT1 in AT/RT differentiation (**Fig. 3g,h, Supplementary Fig. 11h,i**).

Given that ZNHIT1 deposits histone variant H2A.Z and acetylation of H2A.Z is a key epigenetic hallmark of many cancers[41], we also sought to characterize changes in H2A.Z in AT/RT upon *ZNHIT1* loss using CUT&RUN (**Fig. 3i**, **Supplementary Fig. 12a**). *ZNHIT1*-perturbed cells had a large decrease in both the number and magnitude of H2A.Z-bound peaks, including peaks near genes involved in cell cycle and in neuron-related functions such as cytoskeleton-dependent intracellular transport (**Fig. 3j-l**). We observed decreased H2A.Z signal at peaks near neuronal genes such as *SYT4* and *HAP1*, as well as *ATOH8*, *TUJ1*, and *MAP2* (**Supplementary Fig. 12b**), suggesting that decreased H2A.Z deposition secondary to ZNHIT1 loss may facilitate transcription and neuronal differentiation in BT16. As a control, we also measured the promoter-associated histone modification H3K4me3 using CUT&RUN and found virtually no change in peak number or magnitude (**Supplementary Fig. 12c-f**).

To better characterize the role of H2A.Z in cell cycle changes and differentiation, we directly perturbed H2A.Z. Since H2A.Z is encoded by two genes that differ only by three amino acids, we separately perturbed *H2A.Z.1* (encoded by *H2AZ1*) and *H2A.Z.2* (encoded by *H2AZ2*) and measured changes in cell cycle and differentiation. We found a large reduction in cells in S-phase after knock-down of *H2A.Z.2* (74% decrease) and this result was consistent across different AT/RT cell lines (**Fig. 3m,n, Supplementary Fig. 13**), suggesting that the cell cycle arrest mediated by *ZNHIT1* perturbation may work via its role in H2A.Z deposition. Furthermore, we found that loss of *H2A.Z.1* and/or *H2A.Z.2* increases expression of the mature neuronal marker MAP2 across 3 different AT/RT cell lines (**Fig. 3o,p**).

In sum, we have presented MultiPerturb-seq, a multiomic pooled CRISPR screening platform, which captures ATAC, mRNA, and CRISPR perturbations. This method increases throughput more than 10-fold over prior unimodal single-cell perturbation screens and does so with lower cost than other single-cell perturbation methods. Compared to performing separate pooled screens for each modality, MultiPerturb-seq can directly link changes in open chromatin and gene expression, yield multi-modal data without the need for computational integration methods, and provides a better controlled assay with fewer technical and biological confounders. Applied to a rare pediatric brain tumor model, MultiPerturb-seq identified *ZNHIT1* as a potential target for AT/RT reprogramming therapy, which we further confirmed by demonstrating that *ZNHIT1* knockdown pushes cells toward terminal differentiation. We demonstrate the ability of MultiPerturb-seq to perform high-throughput screens with rich phenotypic and mechanistic readout, and the promise of ZNHIT1 and H2A.Z modulation for AT/RT differentiation, though further studies will be needed to understand the therapeutic potential. From a technical viewpoint, there are several ways to further extend this platform. First, MultiPerturb-seq is already compatible with protein capture on the 10X ATAC kit using DNA-barcoded antibodies[42], as well as other types of guide RNAs with a spacer near the 5' end (e.g. CRISPR/Cas9, CRISPRa, prime-editing, base-editing). Second, with two rounds of barcoding, there is an opportunity for a first round of arrayed barcoding in situations where DNA barcoding is challenging, such as different pharmacologic perturbations or processing multiple timepoints in a single experiment. Taken together, MultiPerturb-seq brings together epigenome and transcriptome phenotyping to study the impact of many genetic perturbations.

Acknowledgements

We thank the entire Sanjana laboratory for their support and advice. We also thank P. Smibert, C. Zhu, and S. Hao for sharing single-cell expertise, X. Chen and S. Teichman for single-cell schematics, and the NYU Biology Genomics Core for sequencing resources.

A.C. is supported by the Swedish Research Council (VR). Z.Z.G. is supported by the National Institutes of Health (NIH) T32 Training Grant (grant no. GM136573). N.D. is supported by NIH/Office of the Director (R03OD034499) and N.D. and J.P.G. are supported by the Ty Louis Campbell Foundation. N.E.S. is supported by NIH/National Human Genome Research Institute (DP2HG010099, R01HG012790), NIH/National Cancer Institute (R01CA218668, R01CA279135), the NIH/National Institute of Allergy and Infectious Diseases (R01AI176601), the NIH/National Heart, Lung, and Blood Institute (NHLBI) (R01HL168247),and the Simons Foundation for Autism Research (Genomics of ASD 896724), the MacMillan Center for the Study of the Noncoding Cancer Genome, and New York University and New York Genome Center funds.

Author contributions

R.E.Y. and N.E.S. designed the study. R.E.Y., N.D., and N.E.S. designed the CRISPR library. R.E.Y. performed MultiPerturb-seq experiments and led the analysis. X.X. and A.C. assisted with the pooled screen. N.E.S., E.K., I.R., Z.Z.G., X.W., M.F., and S.F. performed additional data analysis. R.E.Y., L.K., X.W., R.S., X.X., J.C. and A.C. performed arrayed validation. S.G., N.D., J.P.G., and N.E.S. supervised the study. R.E.Y. and N.E.S. wrote the manuscript with input from all authors.

Author information

Correspondence and requests for materials should be addressed to Neville Sanjana (neville@sanjanalab.org).

Competing financial interests

The New York Genome Center and New York University have applied for patents relating to the work in this article. N.E.S. is an adviser to Qiagen and a co-founder and adviser of TruEdit Bio and OverT Bio. The remaining authors declare no competing interests.

Figure legends

Figure 1. MultiPerturb-seq combines single-cell RNA-sequencing and single-cell ATAC-sequencing with pooled CRISPR perturbations for high-throughput functional genomics. **a**, MultiPerturb-seq combines combinatorial indexing with droplet microfluidics for trimodal capture. **b**, Cost comparison for various single-cell CRISPR pooled screens methods. **c**, Capillary electrophoresis of ATAC, RNA, and CRISPR inhibition (CRISPRi) guide RNA (gRNA) libraries from MultiPerturb-seq. All three libraries show expected patterns (ATAC: Nucleosome bands; Tagmented RNA: Range of fragments centered around 400 bp; CRISPR gRNA: Distinct amplicon band at ~200 bp). **d-f,** Single-cell collision rate quantification for ATAC fragments (panel *d*, 11.6%), RNA transcripts (panel *e*, 6.2%), and CRISPR gRNAs (panel *f*, 6,6%) aligning to the human and mouse genomes. ATAC and RNA plots are downsampled for visualization. **g**, Uniform Manifold Approximation and Projection (UMAP) on RNA (transcript) data colored by species. Mouse 3T3 fibroblasts (transduced with the mouse non-targeting gRNA library) constituted 20% of all cells prior to nuclei isolation. **h**, Open chromatin peaks (ATAC), transcripts (RNA) and gRNAs (CRISPR) detected for BT16 (human) cells and 3T3 (mouse) cells. **i**, Distance of ATAC peaks from transcription start sites (TSS). Shaded region represents the 99% confidence interval (n = 10,000 bootstrap samples). **j**, Proportion of single cells with 1, 2, or more than 2 gRNAs detected. **k**, Comparison between cells with histone methyltransferase perturbations (Histone MTs) and cells with non-targeting (NT) control perturbations for gene expression and open chromatin at the *RFX3* locus. **l**, Comparison between cells with perturbations targeting *H3F3A* and cells with non-targeting (NT) control perturbations for gene expression and open chromatin at the *PPM1B* locus. For panels *k* and *l*, reads are normalized to cell number, tracks are binned in 500 bp bins for visualization and scale bars denote 25 kb.

Figure 2. MultiPerturb-seq identifies genetic perturbations that trigger differentiation in atypical teratoid/rhaboid tumor (AT/RT). a, Overview of differentiation challenge in AT/RT brain tumors and design of pooled CRISPR library to identify chromatin remodelers for cancer reprogramming therapy. **b,** Correlation between gene-perturbed human AT/RT cells and gene expression over developmental stages from 4 weeks post-conception (wpc) to senior adulthood.[29] The Pearson correlation is computed on the top 1000 highly variable genes and values are normalized such that cells receiving a non-targeting perturbation display as zero on the colorscale. **c,** Correlation between gene-perturbed human AT/RT cells and open chromatin peaks

in developmental[35] and adult[36] brain atlases (*left*) and sum of fold-changes (log_2) at peaks overlapping ENCODE regulatory elements[37] (*right*). The Pearson correlation is computed on the top 1000 highly variable promoter-adjacent peaks and values are normalized such that cells receiving a non-targeting perturbation display as zero on the colorscale. PLS: promoter-like sequence, pELS: proximal enhancer-like sequence, dELS: distal enhancer-like sequence, DNase-H3K4me3: poised elements.[37] **d-e,** Ranked CRISPRi gene perturbations by RNA differentiation score (panel *d*) and ATAC differentiation score (panel e). Higher values indicate greater similarity to postnatal primary brain tissues (see *Methods*). **f**, RNA and ATAC differentiation scores for all CRISPRi gene perturbations. **g,** Normalized difference in correlations of gene expression between *ZNHIT1*-perturbed cells and cells receiving NT (negative control) perturbations. For each cell population (ZNHIT1, NT), we computed the Pearson correlation of gene expression with human brain developmental expression (n = 53 primary cerebrum samples at the indicated developmental timepoints). Line denotes LOESS fit and shaded region indicates the 95% confidence interval.

Figure 3. ZNHIT1 loss drives AT/RT cell cycle arrest and differentiation via decreased H2A.Z deposition. a, CRISPRi validation in AT/RT cells to assess stemness, proliferation and differentiation after ZNHIT1 loss. **b,** SOX2 expression in cells receiving *ZNHIT1*, *SOX2* or non-targeting (negative control, NT) guide RNAs (gRNAs). **c,** Proportion of S-phase genes[39] as a fraction of expression of all cell-cycle genes (n = 262 *ZNHIT1*-perturbed cells and 4,808 NT cells with at least 100 RNA UMI counts). Error bars indicate the 95% confidence interval (bootstrap resampling). **d**, EdU incorporation in cells with *ZNHIT1*-targeting gRNAs compared to NT gRNAs (n = 3 biological replicates). Treatment with the topoisomerase II inhibitor doxorubicin (Doxo) serves as a positive control for cell cycle arrest. Significance was determined via a one-way ANOVA with Tukey's post-hoc test. **e,** ATOH8 transcription factor signature in MultiPerturb-seq. Transcription factor signatures were calculated by aggregating counts of ATOH8 target genes (n = 262 *ZNHIT1*-perturbed cells and 4,808 NT cells with at least 100 RNA UMI counts) Significance was determined using conditional resampling (SCEPTRE). **f-h,** Expression and quantification of **f,** ATOH8, **g,** TUJ1, and **h,** MAP2 in BT16 cells with *ZNHIT1*-targeting or NT gRNAs (n = 3 biological replicate gRNAs per condition and 3 technical replicates per gRNA). Open circles represent the median for each gRNA. Scale bar: 50μm. **i,** *Above*: H2A.Z CUT&RUN in BT16 cells. *Below*: CUT&RUN signal at H2A.Z peaks in cells with ZNHIT1-targeting or NT gRNAs (n = 5 biological replicates). A representative replicate is shown for visualization. **j,** Change in reproducible H2A.Z CUT&RUN peaks after ZNHIT1 loss (n = 5 biological replicates per condition

with peaks present in at least 4 of out 5 replicates for either cells with *ZNHIT1*-targeting or NT gRNAs). **k,** Change in peak height for reproducible H2A.Z CUT&RUN peaks. For visualization, outliers beyond the 99th percentile are omitted. Significance was determined with a two-sided paired *t*-test. **l,** Enriched Gene Ontology Biological Processes for nearest genes to decreased H2A.Z-bound peaks in ZNHIT1-perturbed cells. *P*-values were computed using a one-sided Fisher's exact test. **m,** Cell cycle analysis in CHLA06 AT/RT cells transduced with *ZNHIT1*-, *H2AZ1*-, or *H2AZ2*-targeting (or NT) gRNAs (n = 2 - 3 guide RNAs per perturbed gene). **n,** Quantification of S-phase cells from panel *n* and significance determined via two-sided χ^2-test (n = 2 - 3 guide RNAs per perturbed gene). **o,** Representative immunofluorescence images of MAP2 expression in BT16, BT12, and CHLA06 AT/RT cells with *H2AZ1*- or *H2AZ2*-targeting (or NT) gRNAs. Scale bar: 50μm. **p,** Quantification of MAP2 expression in BT16, BT12, and CHLA06 AT/RT cells with *H2AZ1*- or *H2AZ2*-targeting (or NT) gRNAs (n = 3 biological replicate gRNAs per condition and 3 technical replicates per gRNA). Open circles represent the median for each gRNA. For all panels, significance levels: *n.s.*, not significant, **, $p < 10^{-2}$ and ****, $p < 10^{-4}$. Unless specified otherwise, significance was determined via a two-sided Mann-Whitney *U* test with a Bonferroni correction for multiple comparisons. Barplots in panels *d,e,* and *n* represent mean values +/- s.e.m. Boxplots show median and interquartile range with whiskers indicating 1.5× interquartile range.

References

1. Dixit, A. et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853-1866.e1817 (2016).
2. Replogle, J.M. et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell* (2022).
3. Frangieh, C.J. et al. Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nature genetics* **53**, 332-341 (2021).
4. Mimitou, E.P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nature Methods* **16**, 409-412 (2019).
5. Rubin, A.J. et al. Coupled Single-Cell CRISPR Screening and Epigenomic Profiling Reveals Causal Gene Regulatory Networks. *Cell* **176**, 361-376 e317 (2019).
6. Liscovitch-Brauer, N. et al. Profiling the genetic determinants of chromatin accessibility with scalable single-cell CRISPR screens. *Nature Biotechnology* (2021).
7. Pierce, S.E., Granja, J.M. & Greenleaf, W.J. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nature Communications* **12**, 2969 (2021).
8. Yubao, C. et al. Perturb-tracing enables high-content screening of multiscale 3D genome regulators. *bioRxiv*, 2023.2001.2031.525983 (2023).
9. Morris, J.A., Sun, J.S. & Sanjana, N.E. Next-generation forward genetic screens: uniting high-throughput perturbations with single-cell analysis. *Trends in Genetics* (2023).
10. Huang, M.E. et al. Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood* **72**, 567-572 (1988).
11. Datlinger, P. et al. Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. *Nature Methods* **18**, 635-642 (2021).
12. Lareau, C.A. et al. Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nature Biotechnology* **37**, 916-924 (2019).
13. Hao, Z. et al. txci-ATAC-seq, a massive-scale single-cell technique to profile chromatin accessibility. *bioRxiv*, 2023.2005.2011.540245 (2023).
14. Morris, J.A. et al. Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science* **380**, eadh7699 (2023).
15. Ma, S. et al. Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103-1116.e1120 (2020).
16. Cao, J. et al. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380-1385 (2018).
17. Chen, S., Lake, B.B. & Zhang, K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nature Biotechnology* **37**, 1452-1457 (2019).
18. Zhu, C. et al. An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nature Structural & Molecular Biology* **26**, 1063-1070 (2019).
19. Datlinger, P. et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nature Methods* **14**, 297-301 (2017).
20. Hill, A.J. et al. On the design of CRISPR-based single-cell molecular screens. *Nature Methods* **15**, 271-274 (2018).
21. Kadoch, C. & Crabtree, G.R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Science Advances* **1**, e1500447 (2015).
22. Jackson, E.M. et al. Genomic analysis using high-density single nucleotide polymorphism-based oligonucleotide arrays and multiplex ligation-dependent probe amplification provides a

comprehensive analysis of INI1/SMARCB1 in malignant rhabdoid tumors. *Clinical Cancer Research* **15**, 1923-1930 (2009).

23. Reddy, A.T. et al. Efficacy of High-Dose Chemotherapy and Three-Dimensional Conformal Radiation for Atypical Teratoid/Rhabdoid Tumor: A Report From the Children's Oncology Group Trial ACNS0333. *Journal of Clinical Oncology* **38**, 1175-1185 (2020).

24. Wilson, B.G. et al. Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer Cell* **18**, 316-328 (2010).

25. Wang, X. et al. BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nature Communications* **10**, 1881 (2019).

26. Nakayama, R.T. et al. SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nature genetics* **49**, 1613-1623 (2017).

27. Langer, L.F., Ward, J.M. & Archer, T.K. Tumor suppressor SMARCB1 suppresses super-enhancers to govern hESC lineage determination. *Elife* **8**, e45672 (2019).

28. Jessa, S. et al. Stalled developmental programs at the root of pediatric brain tumors. *Nature genetics* **51**, 1702-1713 (2019).

29. Cardoso-Moreira, M. et al. Gene expression across mammalian organ development. *Nature* **571**, 505-509 (2019).

30. Lacomme, M., Liaubet, L., Pituello, F. & Bel-Vialar, S. NEUROG2 drives cell cycle exit of neuronal precursors by specifically repressing a subset of cyclins acting at the G1 and S phases of the cell cycle. *Molecular Cell Biology* **32**, 2596-2607 (2012).

31. Bayat, H. et al. CRISPR/Cas9-mediated deletion of a GA-repeat in human GPM6B leads to disruption of neural cell differentiation from NT2 cells. *Scientific reports* **14**, 2136 (2024).

32. Jovanovic, V.M. et al. A defined roadmap of radial glia and astrocyte differentiation from human pluripotent stem cells. *Stem Cell Reports* **18**, 1701-1720 (2023).

33. Chuang, Y.Y. et al. Role of synaptojanin 2 in glioma cell migration and invasion. *Cancer research* **64**, 8271-8275 (2004).

34. Wang, X. et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nature genetics* **49**, 289-295 (2017).

35. Domcke, S. et al. A human cell atlas of fetal chromatin accessibility. *Science* **370**, eaba7612 (2020).

36. Zhang, K. et al. A single-cell atlas of chromatin accessibility in the human genome. *Cell* **184**, 5985-6001.e5919 (2021).

37. Moore, J.E. et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699-710 (2020).

38. Zhao, B. et al. Znhit1 controls intestinal stem cell maintenance by regulating H2A.Z incorporation. *Nature Communications* **10**, 1071 (2019).

39. Viner-Breuer, R., Yilmaz, A., Benvenisty, N. & Goldberg, M. The essentiality landscape of cell cycle related genes in human pluripotent and cancer cells. *Cell Division* **14**, 15 (2019).

40. Divvela, S.S.K., Saberi, D. & Brand-Saberi, B. Atoh8 in Development and Disease. *Biology (Basel)* **11** (2022).

41. Valdés-Mora, F. et al. Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer. *Genome Research* **22**, 307-321 (2012).

42. Mimitou, E.P. et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nature Biotechnology* **39**, 1246-1258 (2021).

43. Medvedeva, Y.A. et al. EpiFactors: a comprehensive database of human epigenetic factors and complexes. *Database (Oxford)* **2015**, bav067 (2015).

44. Corces, M.R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nature Methods* **14**, 959-962 (2017).

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

46. Nørholm, M.H.H. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnology* **10**, 21 (2010).

47. Timothy, B., Kaishu, M., Kathryn, R. & Eugene, K. Robust differential expression testing for single-cell CRISPR screens at low multiplicity of infection. *bioRxiv*, 2023.2005.2015.540875 (2023).

48. D'Amato, L. et al. ARHGEF3 controls HDACi-induced differentiation via RhoA-dependent pathways in acute myeloid leukemias. *Epigenetics* **10**, 6-18 (2015).

49. Nenasheva, V.V. & Tarantul, V.Z. Many Faces of TRIM Proteins on the Road from Pluripotency to Neurogenesis. *Stem Cells and Development* **29**, 1-14 (2019).

50. Hong, F. et al. Dissecting Early Differentially Expressed Genes in a Mixture of Differentiating Embryonic Stem Cells. *PLOS Computational Biology* **5**, e1000607 (2009).

51. Liu, T. et al. EphB4 Regulates Self-Renewal, Proliferation and Neuronal Differentiation of Human Embryonic Neural Stem Cells in Vitro. *Cellular Physiology and Biochemistry* **41**, 819-834 (2017).

52. Powell, G.T. et al. Cachd1 interacts with Wnt receptors and regulates neuronal asymmetry in the zebrafish brain. *Science* **384**, 573-579 (2024).

53. Sadek, C.M. et al. TACC3 expression is tightly regulated during early differentiation. *Gene Expr Patterns* **3**, 203-211 (2003).

54. Takebe, A. et al. Microarray analysis of PDGFRα+ populations in ES cell differentiation culture identifies genes involved in differentiation of mesoderm and mesenchyme including ARID3b that is essential for development of embryonic mesenchymal cells. *Developmental Biology* **293**, 25-37 (2006).

55. Zhang, L.-h. et al. TRIM24 promotes stemness and invasiveness of glioblastoma cells via activating Sox2 expression. *Neuro-Oncology* **22**, 1797-1808 (2020).

56. Kobayashi, K., Era, T., Takebe, A., Jakt, L.M. & Nishikawa, S.-I. ARID3B Induces Malignant Transformation of Mouse Embryonic Fibroblasts and Is Strongly Associated with Malignant Neuroblastoma. *Cancer research* **66**, 8331-8336 (2006).

Materials and methods

Cell lines

BT12 and BT16 cells were a gift from Peter Houghton, Rintaro Hashizume and Charles Roberts. NIH-3T3 (CRL-1658) and CHLA06 (CRL-3038) were acquired from ATCC. HEK293FT cells were acquired from ThermoFisher (R70007). BT12 and BT16 cells were validated by STR profiling; other lines were authenticated by the vendor. All cell lines were maintained at 37 °C and 5% CO2 in D10 medium: DMEM with high glucose and stabilized L-glutamine (Caisson DML23) supplemented with 10% Serum Plus II (Sigma 14009C). Monoclonal CRISPRi-expressing BT16 cell lines were generated by transducing cells with lentiCRISPRi(v2)-Blast (Addgene 170068)[14], selecting with 10µg/ml Blasticidin S (ThermoFisher A1113903), and plating at a low density for colony picking. Several clones were selected and monitored for growth. A clone maintaining normal BT16 growth patterns and CRISPRi(v2) expression by Cas9 immunocytochemistry was selected for the MultiPerturb-seq screen. NIH-3T3, BT12, and CHLA06 cells were also transduced with lentiCRISPRi(v2)-Blast and selected with 10µg/ml blasticidin for 1 week.

Guide RNA design for pooled library and array validation

To identify factors involved in reprogramming AT/RT cells, we constructed a library of 109 epigenomic remodelers with 3 guide RNAs (gRNAs) per gene. The AT/RT library targeted genes that encode proteins with roles in DNA modification, histone modification, histone chaperones, transcription factors, chromatin remodelers, and structural factors. We also included 17 non-targeting controls that do not target anywhere in the human genome. The library was designed using gRNAs from the Dolcetto CRISPRi library and CRISPick[57]. Three gRNAs were selected per gene and homopolymers were excluded. Oligonucleotides were ordered and synthesized by Twist Biosciences in pooled format. For the mouse spike-in, mouse non-targeting gRNAs were ordered individually through Integrated DNA Technologies (IDT) and pooled for library cloning. For both pooled and arrayed guide RNA sequences, please see **Supplementary Table 2**.

Pooled CRISPR library cloning and quality control

Oligonucleotides were diluted, and a PCR cycle test was performed to ascertain the minimum cycles needed for library amplification to preserve integrity. Following this, oligonucleotides were amplified using a two-step nested PCR, then cloned in lentiGuideFE-Puro (Addgene 170069) with Gibson cloning using Gibson mix (NEB E2611L) and precipitated with ethanol. The library was then transformed into Endura cells (Biosearch 60242-2). Bacteria were then grown on plates, maxi-prepped (IBI Scientific IB47125), and then sequenced. For quality control, libraries were sequenced on Illumina MiSeq. Reads were demultiplexed using bcl2fastq (version 2.20), guide spacers were extracted using cutadapt[58] (version 4.0), and aligned with bowtie[59] (version 1.1.2). For the epigenomic remodeler library, we recovered 98% of the designed guide RNAs and, using the read distribution, we computed that the 90^{th}:10^{th} quantile ratio of guide RNAs was 1.8. For the non-targeting library (mouse), we recovered 100% of the designed gRNA and the 90^{th}:10^{th} quantile ratio was 6.5.

Lentivirus production

Lentiviral libraries were prepared in T225 flasks. Each flask was seeded with $27x10^6$ cells the day before in 30 ml of antibiotic-free D10 media to achieve 80-90% confluence before transfection. The transfection mix was 24.9µg of the transfer plasmid (including the epigenetic remodelers or mouse non-targeting library), 13.7µg pMD2.G (Addgene 12260), 19.9µg psPAX2 (Addgene 12259), 2490µl OptiMEM (Invitrogen 51985-091) and 138µl 1 mg/ml polyethylenimine linear MW 25000 (Polysciences 23966). The mixture was mixed and allowed to incubate for 10 minutes at room temperature. After removing 15 ml media from the cells, the mixture was added dropwise. Six hours following transfection, an additional 15 ml fresh media with 1% BSA (VWR AAJ65097-18) was added. Viral supernatants were collected 72 hours following transfection, spun down, filtered with a 0.45mm filter (Millipore SE1M003M00). Lentivirus for the pooled library was concentrated using 2 ml of a 20% sucrose cushion by ultracentrifugation for two hours at 4°C (Beckman JS24.38 swinging bucket rotor, Avanti JXN30), then resuspended in PBS, aliquoted, and stored at -80°C.

Pooled library transduction

Pooled libraries were transduced into BT16 and NIH-3T3 cells with the corresponding libraries with variable viral volumes to determine the appropriate multiplicity of infection for a high single-infection rate, as determined by puromycin survival ($p_{survival}$). We aimed for a $p_{survival}$ of 1 - 5% to ensure single-guide integration. Based on this titration, cells were infected with the appropriate volume of virus. Forty-eight hours following transduction, BT16 and NIH-3T3 cells were lifted and selected with 1 µg/ml and 2µg/ml puromycin respectively (Invivogen ant-pr-1). At the same time, we performed in-line controls in 6-well plates and confirmed that $p_{survival}$ was within the 1 - 5% target. Seven days following infection, cells were lifted, counted, and pooled with 80% BT16 (human) cells and 20% mouse cells (3T3) as a spike-in control for the MultiPerturb-seq library preparation workflow.

MultiPerturb-seq library preparation

For a detailed protocol, please see **Supplementary Protocol 1**. For primer sequences, see **Supplementary Table 1.**

Part 1: Nuclei isolation, tagmentation, and reverse transcription

Overall, our ATAC protocol is similar to a previous, well-validated ATAC method[45] and our transposomes are assembled as in Picelli et al.[60] with MEDS-A (MPSprimer_01), pMENT (MPSprimer_02), and 48 barcoded MEDS-B (MPSprimer_03 – MPSprimer_50) for a 48-well barcoded transposome plate. Of note, although we used standard unsalted oligonucleotides (Integrated DNA Technologies), we found that HPLC-purified oligonucleotides can lead to increased fragments captured per cell. MultiPerturb-seq may also be performed without combinatorial indexing, in which case we advise use of HPLC-purified oligonucleotides since only one MEDS-B is required.

2.4 million human cells and 600k mouse cells were combined and lysed in 1ml Omni lysis buffer (10mM Tris-HCl, pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% NP-40 (ThermoFisher 85124), 0.1% Tween-20 (Sigma P1379), 0.01% digitonin (Promega G9441))[44]. Cells were lysed for 10 minutes on ice. After lysis, nuclei were spun down, pooled, resuspended in 450µl PBS and combined with

tagmentation mix: 240µl 5X TD-TAPS (50mM TAPS-NaOH buffer, pH 8.5 [Boston BioProducts BB-2375], 25mM $MgCl_2$, 50% DMF [Sigma 494488]), 120µl 10% Tween-20, 300µl dilution buffer (10mM Tris-HCl, pH 7.4, 100mM NaCl, 50% glycerol, 1mM DTT), 30µl RiboLock RNase inhibitor (ThermoFisher EO0381). The nuclei were then split among wells of barcoded transposomes for tagmentation.

Cells were then incubated for 30 minutes at 37°C in tagmentation mix while shaking at 350rpm on a ThermoMixer. Following tagmentation, 1µl 126mM EDTA was added to each well and mixed to stop tagmentation. After this, 50µl PBS was added, and nuclei were spun at 400rcf for 4 minutes at 4°C. 53µl of supernatant was then removed, leaving 17µl and the nuclei pellet undisrupted. For the reverse transcription (RT), we added a master mix of 8µl 5X RT buffer (Thermo EP0742: 250 mM Tris-HCl, 375 mM KCl, 15 mM $MgCl_2$, 50 mM DTT), 2µl dNTPs, 2µl MPSprimer_51 (10µM), 4µl MPSprimer_52 (10µM), 2µl Maxima RT H-minus (ThermoFisher EP0753), and 1µl Ribolock (ThermoFisher EO0381) per well. We then added 4µl of barcoded TSO (sequences for the 48 barcoded TSOs are MPSprimer_53 to MPSprimer_100) to match the ATAC barcodes to individual wells. This plate was then incubated for 90 minutes at 53°C, while shaking at 450rpm on a ThermoMixer. An alternative reverse transcription protocol using thermal cycling (50 °C for 10 min; then three cycles of 8 °C for 12 s, 15 °C for 45 s, 20 °C for 45 s, 30 °C for 30 s, 42 °C for 2 min and 50 °C for 3 min followed by a final step at 50 °C for 5 min) as previously used in ISSAAC-seq[61] improves both ATAC and RNA capture, and we recommend this cycling instead of the fixed temperature RT. Nuclei were then resuspended well by triturating with a narrowed pipette tip and all wells were pooled into 2 x 1.5mL tubes, spun down, and re-pooled in a 1.5mL tube. The narrowed pipette tip was produced using a standard plastic 20µl pipette tip (Rainin) melted to narrow gauge using an infrared sterilizer (Joanlab DS-900S). After observing nuclei to avoid clumps and counting, nuclei were resuspended in diluted nuclei buffer to achieve the desired loading amount (100,000 nuclei in 8µl) and combined with 7µl ATAC buffer B (10x Genomics PN2000193).

Part 2: 10X ATAC GEM generation, barcoding, and cleanup

The nuclei suspension was prepared for second-round barcoding using droplet microfluidics (10X Genomics ATAC kit PN1000176) following the manufacturer's instructions. Briefly, nuclei were mixed with the master mix (56.5µl Barcoding reagent B (PN2000194), 1.5µl Reducing agent A (PN2000087), 2µl Barcoding enzyme (PN2000125/139), and loaded onto the Chromium Next GEM Chip H (PN1000162) with glycerol, gel beads, and partitioning oil. Following the run on the Chromium Controller, 100µl GEMs were collected and transferred to a PCR tube for GEM incubation. 15 cycles were substituted for 12 cycles during the linear amplification. GEMs were then cleaned with Dynabeads per the manufacturer's instructions, and libraries were split into 20µl ATAC and 20µl RNA libraries for final library prep. We recovered ~3.6 cells per droplet on average.

Part 3: Library preparation

The ATAC fraction (20µl) was cleaned up with 1.2X SPRI (Illumina) and amplified with an 100µl reaction using NEBNext: 50µl 2X High-Fidelity 2X Master Mix (NEB M0541S), 5µl MPSprimer_101 (10µM), MPSprimer_102 (10µM), 20µl ATAC fraction and 20µl water (30

seconds 98°C, (10 seconds 98°C, 30 seconds 63°C, 1 minute 72°C) x 10-15 cycles, 2 minutes 72°C, hold 4°C), then cleaned with double-sided SPRI (0.45X, 1.8X) in order to isolate fragments of lengths 50-1000 bp. The RNA (cDNA and gRNA) fraction (20µl) was cleaned by incubation with 8µl ExoSAP for 15 minutes at 37°C and then 15 minutes at 80°C. To make 100µl of ExoSAP, we combine 1µl of Exonuclease I (NEB M0293), 20µl of Shrimp Alkaline Phosphatase (NEB M0371), and 79µl water. The cleaned RNA product was amplified with an 100µl KAPA HiFi reaction (Roche 07958935001): 50µl 2X Master Mix, 2.5µl MPSprimer_101 (10µM), 2.5µl MPSprimer_103 (10µM), 2.5µl MPSprimer_104 (10µM), 28µl cleaned RNA product, and 14.5µl water (3 minutes 95°C, (20 seconds 95°C, 30 seconds 66°C, 1 minute 72°C) x 10 cycles, 2 minutes 72°C, hold 4°C). Following amplification, the mRNA and gRNA fractions were split using a two-sided SPRI[4]. The mRNA was collected with a 0.6X SPRI and the gRNA was isolated from the supernatant using an additional 1.4X SPRI. Each fraction was then resuspended in 10µl water. The mRNA may then be amplified with 3-9 additional cycles of a 50µl reaction if there is less than 1ng of product: 25µl 2X KAPA HiFi Master Mix, 1.25µl MPSprimer_101 (10µM), 1.25µl MPSprimer_103 (10µM), 10µl cleaned RNA product, and 12.5µl water (3 minutes 95°C, (20 seconds 95°C, 30 seconds 66°C, 1 minute 72°C) x 3-9 cycles, 2 minutes 72°C, hold 4°C).

After this, the 10µl mRNA fraction was tagmented with Tn loaded with MPSprimer_107 in 20µl of tagmentation buffer for 5 minutes at 55°C. This was then cleaned with DNA Clean & Concentrator-5 (Zymo D4014), resuspended in 33.5µl water and PCR amplified with 50µl PfuX7 reaction[46]: 10µl 5X GC buffer, 1µl dNTPs, 2.5µl MPSprimer_101 (10µM), 2.5µl MPSprimer_108 (10µM), 0.5µl PfuX7 polymerase, and 33.5µl mRNA fraction using the following program: 5 minutes 72°C, 30 seconds 98°C, (10 seconds 98°C, 30 seconds 61°C, 1 minute 72°C) x 10 cycles, 2 minutes 72°C, hold 4°C. The 10µl gRNA fraction was cleaned with 4µl 0.2U/µl ExoSAP and amplified with a 50µl intermediate PCR: 25µl 2X KAPA HiFi Master Mix with 1.25µl biotinylated guide scaffold primer (MPSprimer_105, 10µM), 1.25µl MPSprimer_101 (10µM), 10µl gRNA fraction, and 8.5µl water (3 minutes 95°C, (20 seconds 95°C, 30 seconds 64°C, 1 minute 72°C) x 10 cycles, 2 minutes 72°C, hold 4°C), then cleaned again with 1.8X SPRI, resuspended in 10µl water, and incubated with 4µl ExoSAP. Following cleanup, the gRNA was pulled down with Dynal MyOne Dynabeads Streptavidin C1 (ThermoFisher 65001), resuspended in 45µl water, then amplified with a final inner (guide library) PCR using KAPA HiFi Master Mix: 50µl Master Mix, 2.5µl MPSprimer_101 (10µM), 2.5µl MPSprimer_106 (10µM), and 45µl gRNA pulldown product (3 minutes 95°C, (20 seconds 95°C, 30 seconds 57°C, 1 minute 72°C) x 10 cycles, 2 minutes 72°C, hold 4°C). Samples were evaluated with Tapestation High Sensitivity D1000 ScreenTape and Reagents (Agilent 5067), quantified with Qubit (ThermoFisher Q33231), and sequenced on both Illumina MiSeq and Illumina NovaSeq 6000 v1.5 platforms with 16bp index 1, 8bp index 2, and 50 (MiSeq) or 100bp (NovaSeq) read 1 and 2.

MultiPerturb-seq optimization

MultiPerturb-seq was developed incrementally, first incorporating ATAC, then mRNA and gRNA capture, ensuring preservation of each modality throughout the process (several key examples shown in **Supplementary Fig. 4**). In brief, we built off of our previous work[6], adapting it to the 10X ATAC kit using a mock gel bead oligonucleotide (MPSprimer_109, **Supplementary Table 1**). We further optimized ATAC conditions based on protocols including[44, 45, 63, 64] (**Supplementary**

Fig. 4a-b). Both Tn5[65] and TnY[6] were used in these experiments. We then adapted the direct guide capture technique from [4], also described in [66]. We designed a template switch oligonucleotide (TSO)[67] with barcode and unique molecular identifier (UMI) (**Supplementary Fig. 4c**), and tested PCR[62, 68] and cleanup conditions to achieve mRNA and gRNA capture (**Supplementary Fig. 4c-h**). We also tested several variants of TSO (MPSprimer 110 to MPSprimer_112) (**Supplementary Fig. 4e**). We tested different methods to amplify or enrich the mRNA and gRNA, such as biotin pulldown. Finally, we ensured trimodality integrity, confirming that tagmentation was stopped before reverse transcription, to avoid tagmenting the RNA-DNA heteroduplex[69] (**Supplementary Fig. 4i-j**). Agarose gels in **Supplementary Fig. 4** are 1-2% with 1kb Plus DNA ladder (NEB N3200L) unless otherwise noted. For cost estimates, we used the method's calculated cost when provided or estimated it based on major cost drivers (e.g. 10X Genomics Kits). Sequencing cost was not included in these estimates.

Read alignment and pre-processing

For alignment and pre-processing (**Supplementary Fig. 5a**), we demultiplexed reads using bcl2fastq (version 2.20) with FASTQs for index reads. Reads were then trimmed with cutadapt[58] (version 4.0) to extract barcode 1 (well barcode), barcode 2 (droplet barcode), ATAC reads, mRNA reads, gRNA reads, and UMIs based on position (**Supplementary Fig. 5b**), then aligned separately (**Supplementary Fig. 5c**). Barcodes and gRNA spacers were aligned with bowtie[59] (version 1.1.2) with the settings -v 2 -m 1 –norc –best –strata. The barcode 1 reference was derived from oligonucleotide sequences and the barcode 2 reference was constructed from the whitelist provided by cellranger-atac (10X Genomics). ATAC reads were aligned with bowtie2[70] (version 2.5.1) with default parameters to the joint human (hg38, GENCODE v32/Ensembl98) and mouse (mm10, GENCODE vM23/Ensembl98) genome reference provided by 10X Genomics (2020-A) at https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCh38-and-mm10-2020-A.tar.gz. Open chromatin peaks were called using macs2[71] callpeak (version 2.2.7.1) with the parameters -f BED -g hs -p 0.01 –nomodel –shift 37 –extsize 73 -B –SPMR –keep-dup all –call-summits then reads were assigned to peaks based on loci with bedtools window (version 2.30.0) with a 100 bp window around the start position. mRNA reads were aligned with STAR[72] (version 2.7.3a) using the settings –quantMode GeneCounts –soloFeatures GeneFull_Ex50pAS, then annotated with subread[73] featureCounts (version 2.0.4) using a joint human and mouse gtf with the settings -t gene -R SAM. Aligned reads were then joined to create a list of cell barcodes (barcode 1 and barcode 2), unique molecular identifiers (UMIs) if applicable, and aligned/annotated reads. These were then deduplicated using awk based on barcode, UMI, and position, then imported into R (version 4.2.3), reformatted as a counts matrix using DropletUtils[74] (version 1.18.1), and stored as a SingleCellExperiment[75] object (version 1.20.1). Counts and features were summed with scuttle (version 1.8.4) and peaks were annotated with ChIPseeker (version 1.34.1). We proceeded with the intersection of all three modalities — that is, cell barcodes with all 3 modalities captured (429,139 cell barcodes (**Supplementary Fig. 5c**). Next, we performed additional filtering for cell barcodes with at least 100 RNA UMIs or 100 ATAC unique fragments. This yielded 121,651 cell barcodes, which was the dataset used in all downstream analyses. For barcode collision rate calculations, we defined a collision in any modality as having <66% of the primary species. Each modality was evaluated independently using the same threshold. Cells with at least 500 RNA or ATAC fragments were considered for barcode collision analysis.

Guide RNA assignment

We implemented an algorithm that collapsed highly similar UMIs within the same cell. We did this because, within individual cells, we sometimes identified guide RNA UMIs that differed by only one or two bases. This phenomenon likely arose from sequencing or PCR error, rather than representing genuine biological diversity among UMIs. Consequently, these errors could lead to inflated UMI counts for certain guide-UMI combinations, ultimately skewing the guide assignment and biasing our analysis towards overamplified reads. The algorithm first ranked UMIs based on their read count, assuming that the UMI with the most reads represented the original molecule, that was then mutated during sequencing or PCR. Subsequently, the algorithm recursively removed UMIs that were within a Levenshtein distance[76] of 2 from any remaining UMI with a higher read count or any UMI previously removed. This approach allowed us to account for UMIs that underwent multiple perturbations, such as mutations in both PCR and sequencing stages. Furthermore, we occasionally encountered instances where a single UMI with a high read count was associated with multiple guide RNA, with one association typically dominating in read support. In these cases, we only retained the UMI-guide pairing with the highest read count.

Correlations with primary tissues atlases and differentiation scores

Perturbed cells were separated (pseudo-bulk) by perturbation and compared to published transcriptomic[29] and accessible chromatin[35, 36] atlases by computing the Pearson correlation across the top 1000 highly variable genes or peaks. Correlations were computed between each perturbation-specific pseudo-bulk and previously published primary tissue gene expression or open chromatin. For all correlations and differentiation scores, we only used cells with at least 200 fragments per cell and perturbations with at least 100 cells captured.

For analysis of MultiPerturb-seq gene expression, we first identified highly variable genes (HVGs). We defined HVGs as those genes with the largest standard deviation across cerebrum samples (n = 53 samples from 4 weeks post-conception [wpc] to adulthood with 1-4 donors per developmental stage for that tissue). To compute correlations between MPS and the transcriptomic developmental atlas at specific timepoints, we take the Pearson correlations using only the top 1000 HVGs.

For analysis of MultiPerturb-seq open chromatin, we first identified highly variable promoter-adjacent peaks (HVPPs). We defined HVPPs as those peaks within 2 kb of a protein-coding gene transcription start site with the greatest standard deviation over a unified sample of the MPS ATAC-seq dataset (n = 77 perturbation pseudo-bulk samples) and accessible chromatin pre- or postnatal primary tissues (n = 8 prenatal samples of different brain cell types and 1 postnatal sample from frontal cortex). To compute correlations between MPS and the accessible chromatin developmental atlases, we take the Pearson correlations using only the top 1000 HVPPs.

We computed normalized differentiation scores for either gene expression or open chromatin by taking the difference between correlations (Pearson) with late (postnatal) timepoints and early (prenatal) timepoints to identify those perturbations that increased similarity to mature tissues. This difference was computed using the mean of the correlations over each post- or pre-natal

timepoint. That is, we computed one mean correlation across timepoints prenatal and one mean correlation across timepoints postnatal, normalized each mean correlation, and then took the difference between these normalized means. For the normalization (over perturbations), for each stage (pre-natal or post-natal), we computed maximum and minimum values over perturbations and then assigned each perturbation a normalized $r_i^{norm} = (r_i - \min(\boldsymbol{r})) / (\max(\boldsymbol{r}) - \min(\boldsymbol{r}))$.

Differentially expressed genes, peaks, and signatures

In order to identify differentially expressed genes and peaks, we used SCEPTRE[47], a nonparametric tool that resamples perturbations to infer associations with gene expression[47] with features per cell and counts per cell as covariates. We included barcodes with at least 100 fragments as cells and genes with at least 10 cells captured (n = 106,424 cells). We also applied SCEPTRE to other analyses beyond gene expression, such as the ATAC nearest gene (any distance), ATAC TSS (+/-2kb), and RNA transcription factor transcription factor signatures from msigdb. Gene Ontology (GO) enrichment analyses were performed using clusterProfiler[77] enrichGO (version 4.6.2).

CROP-Multiome

We recloned our epigenomic remodeler library into CROP-seq-opti[20] (Addgene 106280), a vector that places the guide RNA within a polyadenylated mRNA transcript, thus allowing capture by the 3' polyA tail[19]. We then transduced the same BT16 clone expressing CRISPRi-v2 with the CROP-seq library, and prepared snATAC-seq and snRNA-seq libraries using the 10X Chromium Single Cell Multiome ATAC + Gene Expression kit (10X Genomics 1000285). Library cloning, virus production, titration, transduction, and selection was performed as described above for MultiPerturb-seq. We loaded 10,000 cells on one 10X Multiome lane, per manufacturer's instructions. In brief, four days after infection, 200,000 cells (80% BT16 cells and 20% NIH-3T3) were trypsinzied, washed, and lysed in 500µl chilled lysis buffer (10X Genomics) with 12.5µl Ribolock RNase inhibitor (ThermoFisher EO0381). Cells were washed 3 times with 1mL wash buffer (10X Genomics) with 12.5µl Ribolock, and 16,100 cells were resuspended in 10µl transposition mix (10X Genomics) and incubated for 60 minutes at 37°C. Following tagmentation, the mix was loaded on the GEM chip as instructed and run on the Chromium Controller X (10X Genomics). Following incubation, 5 µl quenching agent was added to stop the reaction before proceeding to post-GEM cleanup and library preparation per the manufacturer's instructions (10X Genomics). Samples were sequenced on the Illumina NovaSeq 6000 v1.5 platform with 34bp index 1, 24bp index 2, and 125bp read 1 and 2 and counts matrices were generated with cellranger-arc (version 2.0.2, 10X Genomics). Polyadenylated guide RNA identities aligned with bowtie and joined with barcodes as described above for MultiPerturb-seq with the barcode whitelist provided with cellranger-arc.

CUT&RUN for H2A.Z and H3K4me3

For CUT&RUN[78], we used the CUTANA ChIC/CUT&RUN Kit (EpiCypher 14-1048) with antibodies against H2A.Z (Abcam ab4174), H3K4me3 (EpiCypher 14-1048), and IgG (EpiCypher 14-1048). BT16 cells were transduced with a ZNHIT1-targeting or a non-targeting (negative control) guide RNA (n = 5 biological replicate transductions per guide RNA). Two days later, cells were lifted and selected with 1 µg/ml puromycin. An in-line control was used to ensure complete

selection. Five days following transduction, cells were collected for CUT&RUN. 500,000 cells were used per condition. Cells were lifted, washed, and bound to 10 µl activated Concanavalin A-conjugated paramagnetic beads (EpiCypher), then resuspended with 0.5 µg of the antibody of interest and incubated overnight at 4°C on a rotator. The next day, the beads were washed twice with permeabilization buffer and incubated with 2.5 µl pAG-MNase (Epicypher) for 10 minutes at room temperature. Following binding, the beads were washed, and 2mM $CaCl_2$ was added to begin digestion. Digestion was allowed to proceed for 2 hours at 4°C, then the reaction was terminated by adding 33 µl Stop Buffer (Epicypher) and incubating the reactions at 37°C for 10 minutes. We included a 0.5 ng *E. coli* DNA (Epicypher 18-1401) spike-in. DNA was purified with bead cleanup provided (EpiCypher). Libraries were then prepared using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs E7645S), pooled, and sequenced on an Illumina NovaSeq S1 6000 v1.5 with 2 x 90 bp paired-end reads.

Files were trimmed and with Trim Galore (version 0.6.10) with options –fastqc --paired, then aligned to hg38 (GRCh38.p14, downloaded from the UCSC Genome Browser at https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz) using bowtie2[70] (version 2.5.1) with options --local --very-sensitive-local --no-unal --no-mixed --no-discordant --dovetail -I 10 -X 700. Paired reads were sorted and indexed with samtools (version 1.14). Reads were deduplicated with sambamba[79] (version 0.7.0) view with the options -f bam -F "[XS] == null and not unmapped and not duplicate". Peaks were called with macs2[71] callpeak (version 2.2.7.1) with options –f BAMPE -g hs –bdg with IgG as the control file (-c).

Coordinates (chromosome, start, end, and peak pileups (height at peak summit) from macs2 outputs were used for further analysis. Peak pileups were adjusted by read depth. When combining biological replicates, we sought to only consider peaks that were reproducibly present between replicates. To do this, we called a master peak set on all 10 samples from both conditions. Using valr[80] (version 0.7.0), we only retained peaks called by at least 4 biological replicates of the same condition (ZNHIT1-targeting or non-targeting [NT]) and termed these reproducible peaks. Gene ontology (GO) enrichment was computed using clusterProfiler[77] enrichGO (version 4.6.2) on decreased reproducible peaks.

For visualization, bigwig files were created using deeptools[81] bamCoverage (version 3.4.2) with the options --extendReads --binSize 10 --effectiveGenomeSize 2913022398 --normalizeUsing RPGC. For the pileup visualizations for H2A.Z and H3K4me3, one representative biological replicate is shown: We selected the pair of samples (ZNHIT1-targeting, NT) with the median change in mean coverage at the peak maximum (i.e. median over all 25 possible pairings of 5 ZNHIT1 replicates x 5 NT replicates). For H2A.Z, we used all peaks from the NT samples. For H3K4me3, we used all peaks from the NT samples within 3 kb of the transcription start site of all protein-coding genes expressed at 10 transcripts per million (TPM) or more in BT16 cells[82]. Binding scores were calculated by deeptools computeMatrix reference-point with the input file (H2A.Z or H3K4me3) and IgG control and the parameters -a 3000 -b 3000 –skipZeros --missingDataAsZero --sortRegions descend --sortUsing mean with the blacklist file ENCODE Blacklist v2 (https://www.encodeproject.org/annotations/ENCSR636HFF/) for hg38[83] as --blackListFileName to filter out reads aligning to problematic genome regions and then plotted using plotHeatmap.

Arrayed CRISPRi validation

For arrayed validation, BT16, BT12, and/or CHLA06 cells with lentiCRISPRi(v2)-Blast were transduced with guide RNAs (gRNAs) in lentiGuideFE-Puro (Addgene 170069). The gRNAs were designed using the Dolcetto CRISPRi library and CRISPick[57] then synthesized by Integrated DNA Technologies (IDT) (**Supplementary Table 2**). The backbone was digested with *BsmB*I (ThermoFisher FD0454) and oligos were annealed, phosphorylated and ligated into the lentiGuideFE-Puro backbone. Lentivirus was produced as described in *Lentivirus production* above (scaled to 6-well format) and stored at -80°C. For arrayed validations, sufficient lentivirus was added to the cells to achieve 20 – 50% cell transduction. After 48 hours, cells were replated in media with puromycin (1 µg/ml) and selected for at least 2 days with confirmation of complete selection via an in-line selection control.

SOX2 staining and flow cytometry

To label and quantify SOX2-positive cells, cells were lifted, washed, and stained with LIVE/DEAD Violet (ThermoFisher L34963) (diluted 1:400, 15µl for $1x10^6$ cells) for 5 minutes at room temperature, then washed with PBS and fixed with 1% formaldehyde and incubated at room temperature for 10 minutes on rotator (ThermoFisher Digital Tube Revolver 88881101)[84]. Following fixation, they were quenched with 0.125M glycine (by addition of 2.5 M glycine), washed with PBS, and lysed with 100µl of a previously optimized lysis buffer[84] (10mM Tris-HCl pH 7.5, 10mM NaCl, 3mM $MgCl_2$, 0.1% NP-40, 1% BSA) on ice for 5 minutes. Then they were washed with 1 ml wash buffer (same as lysis without NP-40) and blocked in 1 ml PBS with 3% BSA for 30 minutes at room temperature. Following blocking, they were washed and resuspended in 100 µl PBS-3% BSA and antibody (1:100, 1µg for $5x10^6$ cells, anti-SOX2 Biolegend 656104) for 60 minutes at room temperature. They were then washed twice more (PBS with 3% BSA and 1% Tween) and resuspended in PBS with 1% BSA and 2mM EDTA for flow cytometry on the flow cytometer (Sony SH800). Sequential gating was performed as follows: exclusion of debris on the basis of forward (FSC-A) and side scatter (SSC-H) cell parameters followed by exclusion of dead cells based on LIVE/DEAD and analyzed with FlowJo (version 10.10.0).

Immunofluorescence

Cells were plated in 96-well plates with 5,000 cells per well in triplicate. The next day, media was aspirated, and cells were washed and fixed with 4% paraformaldehyde (diluted 1:4 from 16%, Electron Microscopy Sciences 15710-S) for 15 minutes, and washed with PBS. Cells were then permeabilized with 0.2% Tween-20 for 5 minutes and blocked with PBS with 0.2% Tween-20 and 3% BSA for 1 hour. Cells were then incubated with primary antibodies: TUJ1 at a 1:1000 dilution (BioLegend 801201), MAP2 at a 1:500 dilution (SYSY 188004), or ATOH8 (ThermoFisher PA5-65024) at a 1:400 dilution overnight at 4°C. The following day, cells were washed three times for 5 minutes with cold PBS. The corresponding secondary antibody was added at a 1:800 dilution (ThermoFisher A-21202 for TUJ1 (mouse), ThermoFisher A-11073 for MAP2 (guinea pig), ThermoFisher 31572 for ATOH8 (rabbit)) with 2mM Hoechst (Sigma B2261) and incubated for 1 hour at room temperature. Cells were then washed with PBS for an additional 3 washes. All steps were performed at room temperature on a rocker unless otherwise noted. Images were acquired

with a 20X objective using an epifluorescence microscope (Keyence BZ-X800). Five images were acquired per well.

Quantitative image analysis was run in CellProfiler[85] (version 4.2.6). Primary objects were identified based on the nucleus (Hoechst) with a threshold calculated via Otsu's method. Secondary objects (cytoplasm) were defined by extension from the nucleus (Distance-B method with a threshold calculated via Otsu's method). After segmentation, images were manually examined and images with segmentation artifacts were discarded. ATOH8 signal (nuclear) was quantified using integrated intensity (sum) per cell/object. TUJ1 and MAP2 signal (cytoplasmic) were quantified using mean intensity per cell/object. For MAP2 images, we also applied a flatfield illumination correction. Normalization was performed to the median intensity of cells/objects receiving non-targeting (NT) gRNAs. Cells/objects with an assigned intensity (integrated or mean depending on the protein) greater than 3 standard deviations from the NT mean were excluded as fluorescent debris.

EdU incorporation and cell cycle analysis

Cells were labeled with 5-ethynyl-2'-deoxyuridine (EdU) using the Click-iT EdU Cell Proliferation Kit (ThermoFisher C10337). 2,000 cells/well were plated on 96 well plates in triplicate. Cells were incubated with 10 mM EdU for 30 minutes, fixed with 4% PFA for 15 minutes, and permeabilized with 0.5% Triton X-100 for 10 minutes at room temperature. Cells were then washed and incubated with the Click-iT reaction cocktail for 30 minutes. As a positive control, untransduced BT16 cells were exposed to 1μM doxorubicin (MedChemExpress HY-15142) to inhibit proliferation and EdU incorporation. After EdU staining, nuclei were stained with 2mM Hoechst 3342 (Sigma 4533) for 15 minutes, washed with PBS, and images were acquired with a 20X objective using an epifluorescence microscope (Keyence BZ-X800). The images were processed for display using FIJI (version 2.1.0) and quantitative image analysis was run in CellProfiler[85] (version 4.2.6). Cells were quantified based on Hoechst staining and binned into EdU positive and EdU negative cells based on the integrated intensity (sum) per cell/object, using the ClassifyObjects module.

For propidium iodide (PI) staining, cells were pelleted in 1.5 mL tubes, washed once with 1 mL PBS, and resuspended well in 300 μl PBS. Then, 700 μl of ice cold 100% ethanol was added to fix cells at a final concentration of 70%. Fixed cells were then incubated on ice at 4°C overnight. Next, cells were spun down at 1000g for 4 minutes and ethanol was removed. Cells were washed with 1mL PBS and stained with 0.5mL FxCycle PI/RNAse solution (ThermoFisher F10797) per 1 million cells. Pellets were resuspended and incubated for 15 minutes at room temperature before being resuspended for flow cytometry (Sony SH800 or MACSQuant10). Sequential gating was performed as follows: exclusion of debris on the basis of forward (FSC-A) and side (SSC-H) scatter cell parameters followed by getting on singlets with FSC-A – FSC-H. The cell cycle profile was modeled, and gates were generated based on the PI-A signal of the cell population by FlowJo (version 10.10.0) using a Watson model.

Data availability

MultiPerturb-seq data can be downloaded from BioProject (accession number PRJNA1160410)[86]. The human genome hg38 (GRCh38.p14) is from the UCSC Genome Browser: https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz. The joint human (hg38, GENCODE v32/Ensembl98) and mouse (mm10, GENCODE vM23/Ensembl98) genome (2020-A) is from 10X Genomics: https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCh38-and-mm10-2020-A.tar.gz). Reference developmental and adult atlases were downloaded from https://apps.kaessmannlab.org/evodevoapp/[29], https://descartes.brotmanbaty.org/[35], and http://catlas.org/humanbrain/[36]. Data from previously published studies are from SRA/GEO: CRISPR-sciATAC[6] (PRJNA674902), scifiRNA-seq[11] (PRJNA713314), sci-CAR-seq[16] (PRJNA481032), SNARE-seq[17] (PRJNA520914), Paired-seq[18] (PRJNA539985), and SHARE-seq[15] (PRJNA588784).

Code availability

Code for data processing and visualization are available at https://gitlab.com/sanjanalab/mps[87].

Methods-only References

57. Sanson, K.R. et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications* **9**, 5416 (2018).

58. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* **17**, 3 (2011).

59. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25 (2009).

60. Picelli, S. et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Research* **24**, 2033-2040 (2014).

61. Xu, W. et al. ISSAAC-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. *Nature Methods* (2022).

62. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**, 171-181 (2014).

63. Grandi, F.C., Modi, H., Kampman, L. & Corces, M.R. Chromatin accessibility profiling by ATAC-seq. *Nature Protocols* **17**, 1518-1552 (2022).

64. Preissl, S. et al. Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nature Neuroscience* **21**, 432-439 (2018).

65. Adey, A.C. Tagmentation-based single-cell genomics. *Genome Research* **31**, 1693-1705 (2021).

66. Replogle, J.M. et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nature Biotechnology* **38**, 954-961 (2020).

67. Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R. & Siebert, P.D. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* **30**, 892-897 (2001).

68. Bagnoli, J.W. et al. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nature Communications* **9**, 2937 (2018).

69. Di, L. et al. RNA sequencing by direct tagmentation of RNA/DNA hybrids. *Proceedings of the National Academy of Sciences* **117**, 2886 (2020).

70. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359 (2012).

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

72. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

73. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).

74. Griffiths, J.A., Richard, A.C., Bach, K., Lun, A.T.L. & Marioni, J.C. Detection and removal of barcode swapping in single-cell RNA-seq data. *Nature Communications* **9**, 2667 (2018).

75. Amezquita, R.A. et al. Orchestrating single-cell analysis with Bioconductor. *Nature Methods* **17**, 137-145 (2020).

76. Levenshtein, V.I. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady* **10**, 707-710 (1966).

77. Xu, S. et al. Using clusterProfiler to characterize multiomics data. *Nature Protocols* (2024).

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

79. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-2034 (2015).

80. Riemondy, K. et al. valr: Reproducible genome interval analysis in R. *F1000Research* **6** (2017).

81. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* **44**, W160-W165 (2016).

82. Dharia, N.V. et al. A first-generation pediatric cancer dependency map. *Nature genetics* **53**, 529-538 (2021).

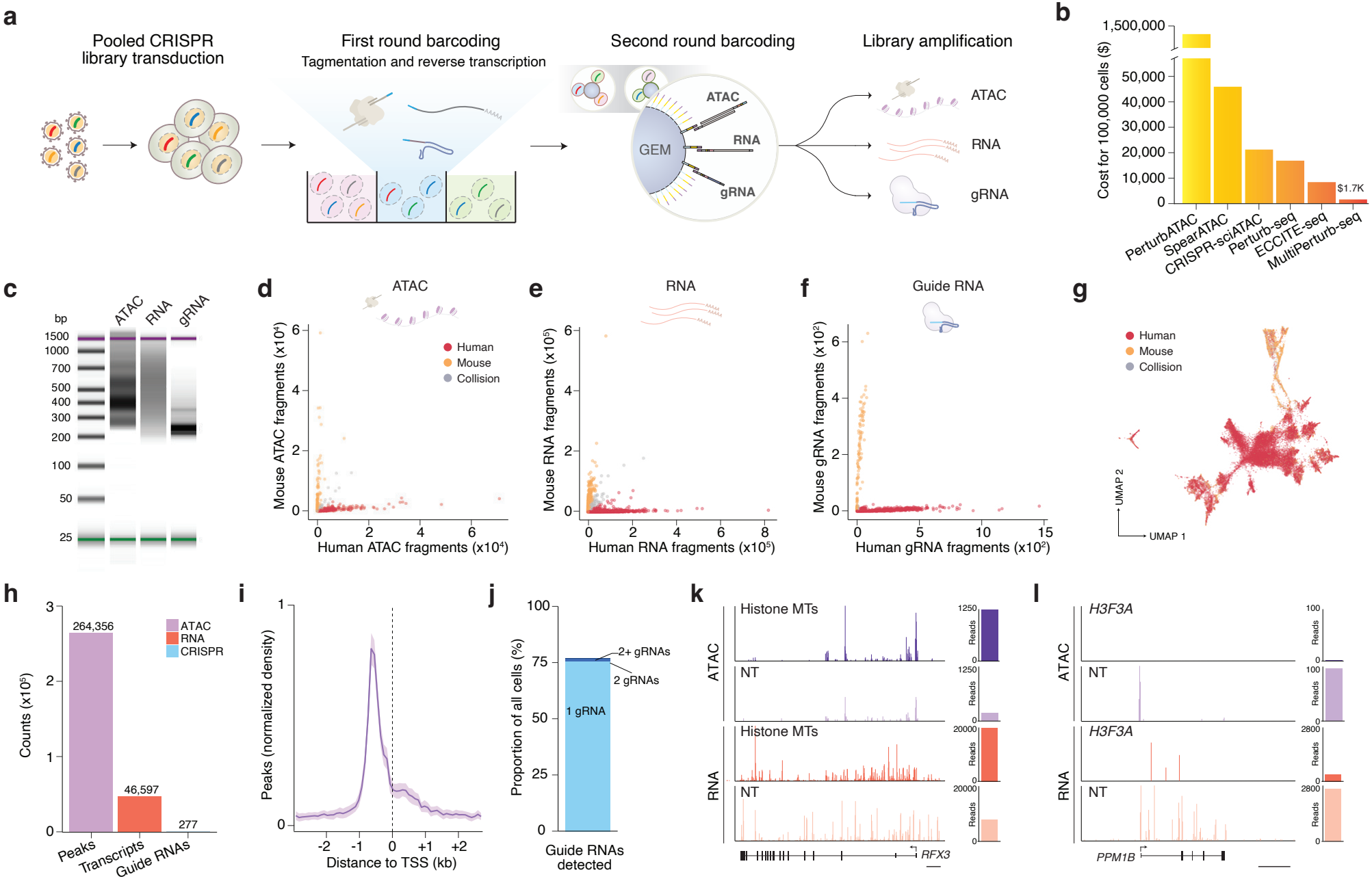
83. Amemiya, H.M., Kundaje, A. & Boyle, A.P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Scientific reports* **9**, 9354 (2019).

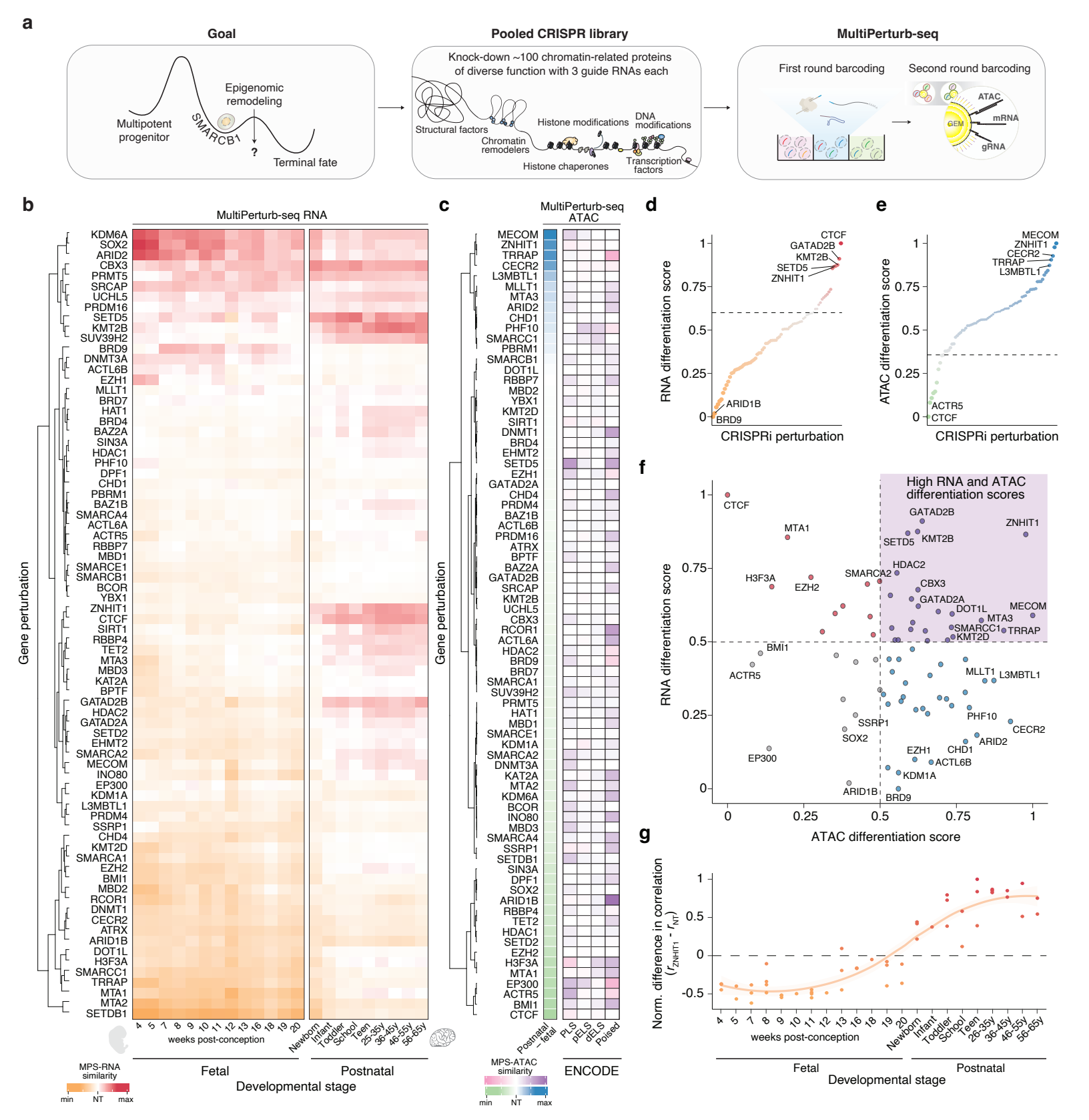
84. Fiskin, E. et al. Single-cell profiling of proteins and chromatin accessibility using PHAGE-ATAC. *Nature Biotechnology* **40**, 374-381 (2022).

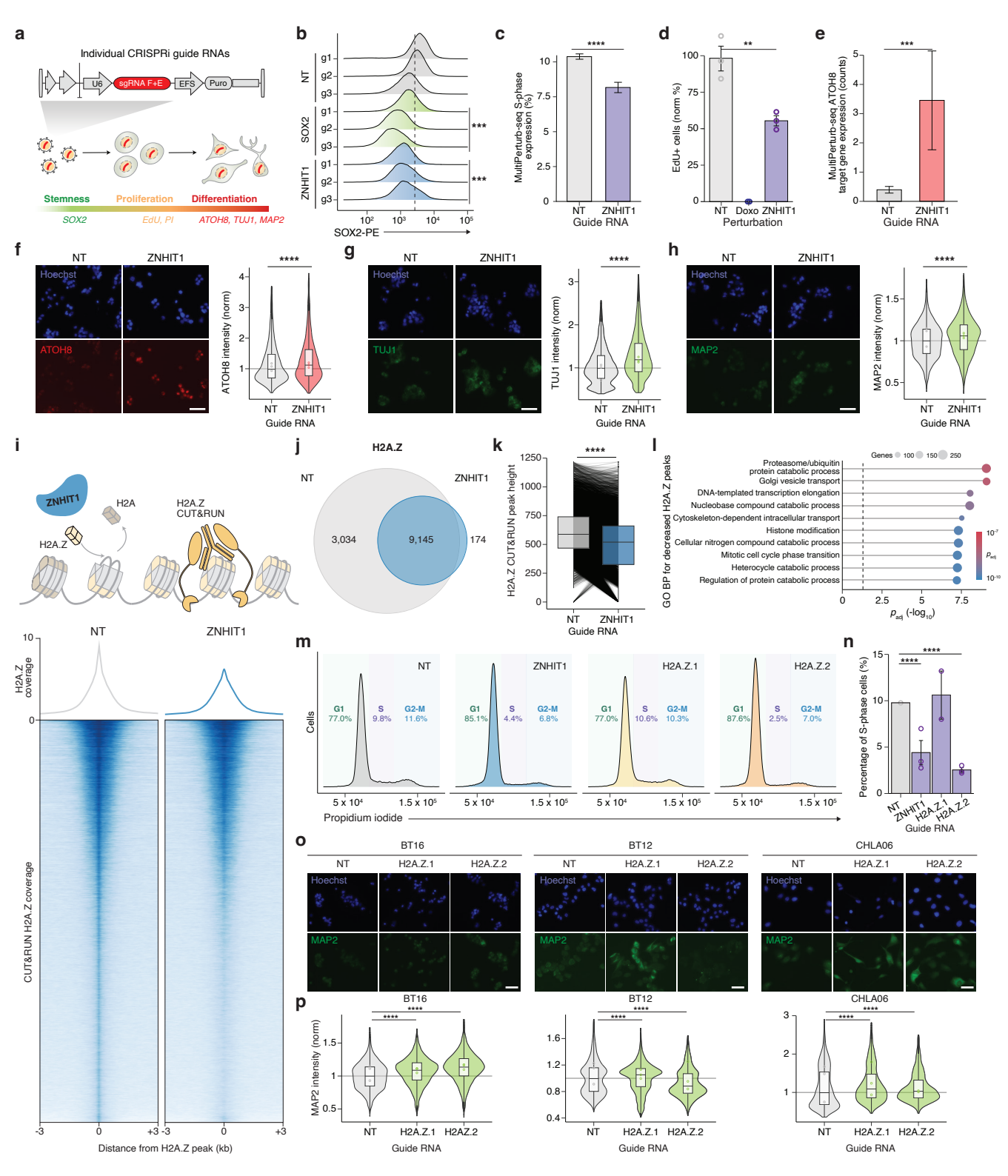
85. Stirling, D.R. et al. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021).

86. Yan, R.E. et al. Pooled CRISPR screens with joint single-nucleus chromatin accessibility and transcriptome profiling. *Datasets. Gene Expression Omnibus* (2024).

87. Yan, R.E. et al. Pooled CRISPR screens with joint single-nucleus chromatin accessibility and transcriptome profiling. *Gitlab* (2024).


a
Pooled CRISPR library transduction
First round barcoding
Tagmentation and reverse transcription
Second round barcoding
Library amplification
GEM
ATAC
RNA
gRNA
b
Cost for 100,000 cells ($)
1,500,000
50,000
40,000
30,000
20,000
10,000
0
$1.7K
PerturbATAC
SpearATAC
CRISPR-sciATAC
Perturb-seq
ECCITE-seq
MultiPerturb-seq
c
bp
1500
1000
700
500
400
300
200
100
50
25
d
Mouse ATAC fragments (x10⁴)
Human ATAC fragments (x10⁴)
Human
Mouse
Collision
e
Mouse RNA fragments (x10⁵)
Human RNA fragments (x10⁵)
f
Guide RNA
Mouse gRNA fragments (x10²)
Human gRNA fragments (x10²)
g
UMAP 1
UMAP 2
h
Counts (x10⁵)
264,356
46,597
277
Peaks
Transcripts
Guide RNAs
i
Peaks (normalized density)
Distance to TSS (kb)
j
Proportion of all cells (%)
2+ gRNAs
2 gRNAs
1 gRNA
Guide RNAs detected
k
Histone MTs
NT
Reads
RFX3
l
H3F3A
PPM1B


a
Goal
Multipotent progenitor
SMARCB1
Epigenomic remodeling
?
Terminal fate
Pooled CRISPR library
Knock-down ~100 chromatin-related proteins of diverse function with 3 guide RNAs each
Structural factors
Chromatin remodelers
Histone modifications
DNA modifications
Histone chaperones
Transcription factors
MultiPerturb-seq
First round barcoding
Second round barcoding
ATAC
mRNA
gRNA
GEM
b
MultiPerturb-seq RNA
Gene perturbation
KDM6A
SOX2
ARID2
CBX3
PRMT5
SRCAP
UCHL5
PRDM16
SETD5
KMT2B
SUV39H2
BRD9
DNMT3A
ACTL6B
EZH1
MLLT1
BRD7
HAT1
BRD4
BAZ2A
SIN3A
HDAC1
PHF10
DPF1
CHD1
PBRM1
BAZ1B
SMARCA4
ACTL6A
ACTR5
RBBP7
MBD1
SMARCE1
SMARCB1
BCOR
YBX1
ZNHIT1
CTCF
SIRT1
RBBP4
TET2
MTA3
MBD3
KAT2A
BPTF
GATAD2B
HDAC2
GATAD2A
SETD2
EHMT2
SMARCA2
MECOM
INO80
EP300
KDM1A
L3MBTL1
PRDM4
SSRP1
CHD4
KMT2D
SMARCA1
EZH2
BMI1
MBD2
RCOR1
DNMT1
CECR2
ATRX
ARID1B
DOT1L
H3F3A
SMARCC1
TRRAP
MTA1
MTA2
SETDB1
4 5 7 8 9 10 11 12 13 16 18 19 20
weeks post-conception
Newborn Infant Toddler School Teen 25-35y 36-45y 46-55y 56-65y
Fetal
Postnatal
Developmental stage
MPS-RNA similarity
min NT max
c
MultiPerturb-seq ATAC
Gene perturbation
MECOM
ZNHIT1
TRRAP
CECR2
L3MBTL1
MLLT1
MTA3
ARID2
CHD1
PHF10
SMARCC1
PBRM1
SMARCB1
DOT1L
RBBP7
MBD2
YBX1
KMT2D
SIRT1
DNMT1
BRD4
EHMT2
SETD5
EZH1
GATAD2A
CHD4
PRDM4
BAZ1B
ACTL6B
PRDM16
ATRX
BPTF
BAZ2A
GATAD2B
SRCAP
KMT2B
UCHL5
CBX3
RCOR1
ACTL6A
HDAC2
BRD9
BRD7
SMARCA1
SUV39H2
PRMT5
HAT1
MBD1
SMARCE1
KDM1A
SMARCA2
DNMT3A
KAT2A
MTA2
KDM6A
BCOR
INO80
MBD3
SMARCA4
SSRP1
SETDB1
SIN3A
DPF1
SOX2
ARID1B
RBBP4
TET2
HDAC1
SETD2
EZH2
H3F3A
MTA1
EP300
ACTR5
BMI1
CTCF
Postnatal – fetal
PLS
pELS
dELS
Poised
ENCODE
MPS-ATAC similarity
min NT max
d
CTCF
GATAD2B
KMT2B
SETD5
ZNHIT1
ARID1B
BRD9
RNA differentiation score
CRISPRi perturbation
e
MECOM
ZNHIT1
CECR2
TRRAP
L3MBTL1
ACTR5
CTCF
ATAC differentiation score
CRISPRi perturbation
f
High RNA and ATAC differentiation scores
CTCF
GATAD2B
ZNHIT1
MTA1
SETD5
KMT2B
HDAC2
SMARCA2
H3F3A
EZH2
CBX3
GATAD2A
DOT1L
MECOM
MTA3
SMARCC1
TRRAP
KMT2D
BMI1
ACTR5
MLLT1
L3MBTL1
PHF10
CECR2
SSRP1
SOX2
ARID2
EZH1
CHD1
ACTL6B
EP300
KDM1A
ARID1B
BRD9
RNA differentiation score
ATAC differentiation score
g
Norm. difference in correlation (r_{ZNHIT1} - r_{NT})
4 5 7 8 9 10 11 12 13 16 18 19 20
weeks post-conception
Newborn Infant Toddler School Teen 26-35y 36-45y 46-55y 56-65y
Fetal
Postnatal
Developmental stage


a
Individual CRISPRi guide RNAs
U6
sgRNA F+E
EFS
Puro
Stemness
Proliferation
Differentiation
SOX2
EdU, PI
ATOH8, TUJ1, MAP2
b
NT
SOX2
ZNHIT1
g1
g2
g3
***
SOX2-PE
c
MultiPerturb-seq S-phase expression (%)
****
NT
ZNHIT1
Guide RNA
d
EdU+ cells (norm %)
**
NT
Doxo
ZNHIT1
Perturbation
e
MultiPerturb-seq ATOH8 target gene expression (counts)
***
NT
ZNHIT1
Guide RNA
f
NT
ZNHIT1
Hoechst
ATOH8
ATOH8 intensity (norm)
****
Guide RNA
g
Hoechst
TUJ1
TUJ1 intensity (norm)
****
Guide RNA
h
Hoechst
MAP2
MAP2 intensity (norm)
****
Guide RNA
i
ZNHIT1
H2A
H2A.Z
H2A.Z CUT&RUN
j
H2A.Z
NT
ZNHIT1
3,034
9,145
174
k
H2A.Z CUT&RUN peak height
****
NT
ZNHIT1
Guide RNA
l
GO BP for decreased H2A.Z peaks
Genes 100 150 250
Proteasome/ubiquitin protein catabolic process
Golgi vesicle transport
DNA-templated transcription elongation
Nucleobase compound catabolic process
Cytoskeleton-dependent intracellular transport
Histone modification
Cellular nitrogen compound catabolic process
Mitotic cell cycle phase transition
Heterocycle catabolic process
Regulation of protein catabolic process
p_{adj}
p_{adj} (-log10)
NT
ZNHIT1
H2A.Z coverage
CUT&RUN H2A.Z coverage
Distance from H2A.Z peak (kb)
m
NT
ZNHIT1
H2A.Z.1
H2A.Z.2
Cells
G1 77.0% S 9.8% G2-M 11.6%
G1 85.1% S 4.4% G2-M 6.8%
G1 77.0% S 10.6% G2-M 10.3%
G1 87.6% S 2.5% G2-M 7.0%
Propidium iodide
n
Percentage of S-phase cells (%)
****
NT
ZNHIT1
H2A.Z.1
H2A.Z.2
Guide RNA
o
BT16
BT12
CHLA06
NT
H2A.Z.1
H2A.Z.2
Hoechst
MAP2
p
BT16
BT12
CHLA06
MAP2 intensity (norm)
****
NT
H2A.Z.1
H2AZ.2
H2A.Z.2
Guide RNA