nature biotechnology

Brief Communication

https://doi.org/10.1038/s41587-024-02475-x

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

Received: 24 January 2024

Accepted: 17 October 2024

Published online: 21 November 2024

Check for updates

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] & Neville E. Sanjana[1,2]

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 cancer reprogramming therapy.

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 three-dimensional genome conformation[8]. These single-cell screens have included a diverse array of genetic perturbations, including knockout using clustered regularly interspaced short palindromic repeats (CRISPR) nucleases like Cas9, transcriptional modulation using CRISPR inhibition (CRISPRi) and activation (CRISPRa), targeting of RNA using Cas13, precise variant insertion through homology-directed repair or base editing, and overexpression with open reading frame libraries[9].

Here, we introduce MultiPerturb-seq, a method that links pooled CRISPR perturbations with single-cell open chromatin (assay for transposase-accessible chromatin with sequencing, ATAC-seq) and gene expression (RNA sequencing, RNA-seq) profiles at scale (Fig. 1a and 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 (that is, differentiation therapy) has been curative for patients with malignancies such as acute promyelocytic leukemia[10], success has been limited in other cancers because of a lack of high-throughput methods to identify reprogramming targets. In MultiPerturb-seq, open chromatin provides a broad overview of the epigenetic state, capturing many levels of gene regulation, while gene expression provides a robust view of the cell state and developmental stage. Together, they link CRISPR perturbations with cell states and putative mechanisms of action for transcriptional and epigenetic 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 notable cost advantages over existing unimodal 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 and Supplementary Protocol). After nuclear isolation and

[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]Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. e-mail: neville@sanjanalab.org

https://doi.org/10.1038/s41587-024-02475-x

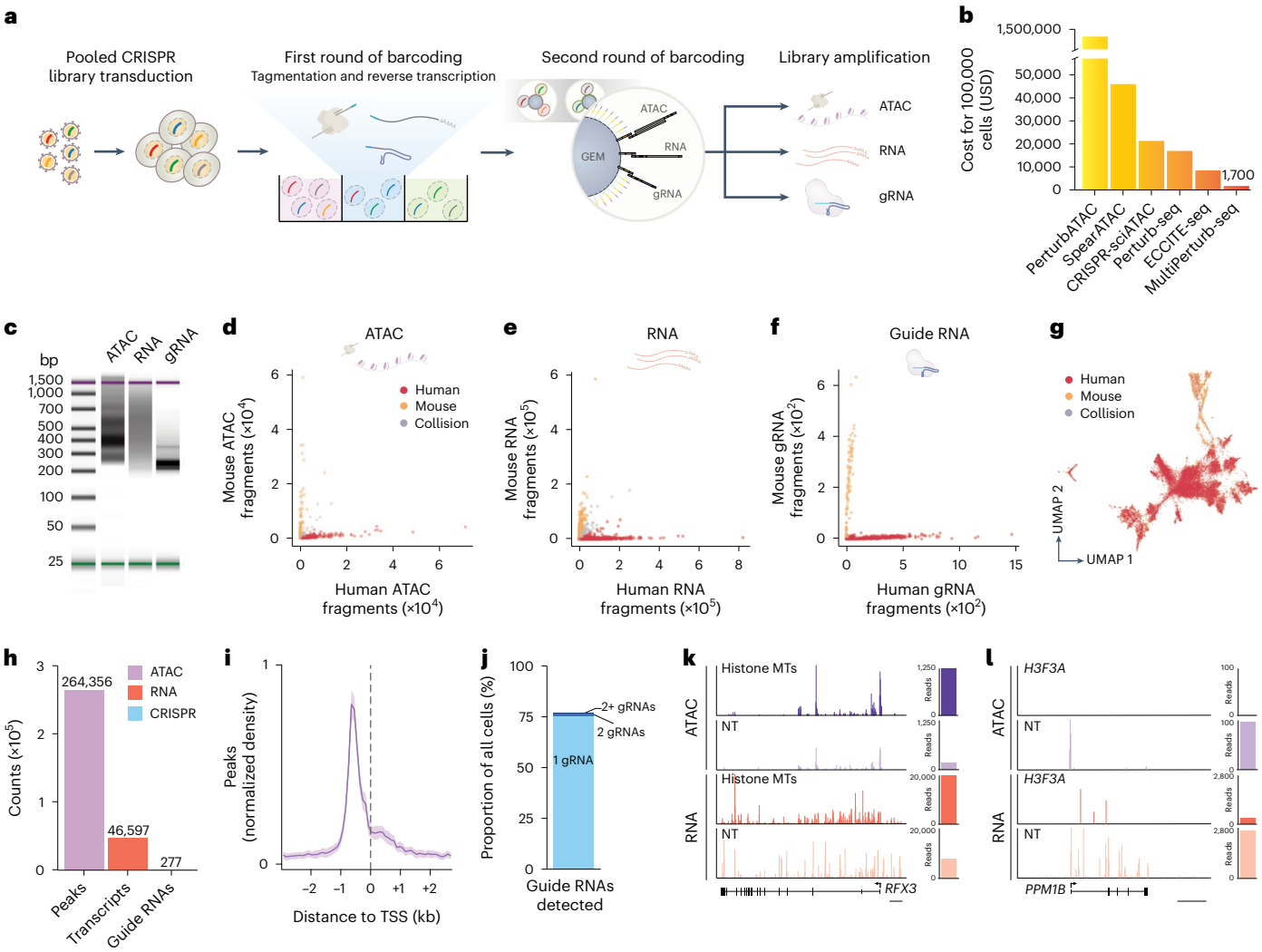


Fig. 1 | MultiPerturb-seq combines single-cell RNA-seq and single-cell ATAC-seq with pooled CRISPR perturbations for high-throughput functional genomics. a, MultiPerturb-seq combines combinatorial indexing with droplet microfluidics for trimodal capture. **b**, Cost comparison for various single-cell CRISPR pooled screen methods. **c**, Capillary electrophoresis of ATAC, RNA and CRISPRi 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 (**d**; 11.6%), RNA transcripts (**e**; 6.2%) and CRISPR gRNAs (**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 NT gRNA library) constituted 20% of all cells before nuclear 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 TSSs. The shaded region represents the 99% confidence interval (n = 10,000 bootstrap samples). **j**, Proportion of single cells with one, two or more than two gRNAs detected. **k**, Comparison between cells with histone methyltransferase perturbations (histone MTs) and cells with NT control perturbations for gene expression and open chromatin at the *RFX3* locus. **l**, Comparison between cells with perturbations targeting *H3F3A* and cells with NT control perturbations for gene expression and open chromatin at the *PPM1B* locus. In **k** and **l**, reads are normalized to cell number, tracks are binned in 500-bp bins for visualization and scale bars denote 25 kb. PerturbATAC, perturbation-indexed single-cell ATAC-seq; SpearATAC, single-cell perturbations with an accessibility readout using scATAC-seq; CRISPR-sciATAC, pooled CRISPR screens with single-cell combinatorial indexing ATAC; Perturb-seq, pooled, combinatorial CRISPR screens with scRNA-seq readout; ECCITE-seq, expanded CRISPR-compatible cellular indexing of transcriptomes and epitopes by sequencing.

distribution into wells, we tagmented open chromatin using barcoded transposomes (Supplementary Fig. 4a,b)[6]. Next, we performed reverse transcription (RT) with a mix of poly(dT) and CRISPR gRNA-specific primers and barcoded template switch oligonucleotides (TSOs) with matching barcodes (Supplementary Fig. 4c–f and Supplementary Table 1). We then pooled cells for second-round barcoding using droplet microfluidics with 10x Chromium ATAC gel beads. Lastly, ATAC, RNA and CRISPR gRNA libraries were amplified and prepared for sequencing (Fig. 1c and Supplementary Figs. 4g–j and 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 and Supplementary Fig. 6a–d). We quantified the percentage of barcode combinations that 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 and Supplementary Fig. 6e–h). For the ATAC, we observed characteristic open chromatin enrichment around transcription start sites (TSSs) (Fig. 1i and 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 nontargeting (NT) gRNA: 78% of high-quality cells were assigned gRNA identities (Fig. 1j and 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) and increased ATAC fragments and peaks per cell (Supplementary Fig. 6m–p)[6,11,15–18].

Although it is not compatible with barcoded superloading, we also used the 10x Chromium Multiome kit and a specialized gRNA plasmid for CRISPR droplet sequencing (CROP-seq)[19,20] as an alternate method of multimodal capture and performed a lower-throughput version of a multiomic CRISPR screen (~10,000 versus ~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 per cell (Supplementary Fig. 7i) and ATAC peaks per cell (Supplementary Fig. 7j). Given these differences and the additional advantages of tenfold increased cell loading, direct gRNA 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*, and *SUV39H2*), we found increases in open chromatin at the *RFX3* locus and increased *RFX3* gene expression (Fig. 1k). We were also 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 mutant 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 4 years[23]. Because of the loss of *SMARCB1*, AT/RT is dependent on alternate epigenetic regulators, such as Polycomb[24–26], and *SMARCB1*-null embryonic stem cell models fail to differentiate into neurons because of 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] (cerebrum, cerebellum, kidney, ovary, testis and liver) and found the highest overall similarity with brain cerebral 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 the negative controls (NT 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* and *GATAD2B*. These tended to express higher levels of genes correlated with neural differentiation such as *CCND3* (ref. 30), *GPM6B* (ref. 31) and *SYNJ2* (refs. 32,33) (Supplementary Fig. 9).

The chromatin landscape in AT/RT is unusual with broad changes due to the 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 and 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 (Methods and Fig. 2d,e). Interestingly, we found that RNA and ATAC differentiation scores were 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 or 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 was previously shown to maintain stemness in intestinal stem cells by promoting H2A.Z incorporation[38]. *ZNHIT1* knockdown induced large changes at multiple regulatory elements, including promoters and enhancers, with increased transcriptomic similarity to postnatal and specifically adult brain tissues. (Fig. 2g and Supplementary Fig. 10b–e). To identify potential mechanisms of action, we examined differentially accessible chromatin in *ZNHIT1*-perturbed cells compared to NT controls. We found that *ZNHIT1* perturbation led to changes in accessibility near genes involved in neuronal differentiation and axonogenesis (Supplementary Fig. 11b) and 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 might be a good candidate to push AT/RT cells toward terminal differentiation. We cloned individual CRISPR gRNAs to target *ZNHIT1* and measured stemness, proliferation and expression of differentiated neuronal markers (Fig. 3a).

https://doi.org/10.1038/s41587-024-02475-x

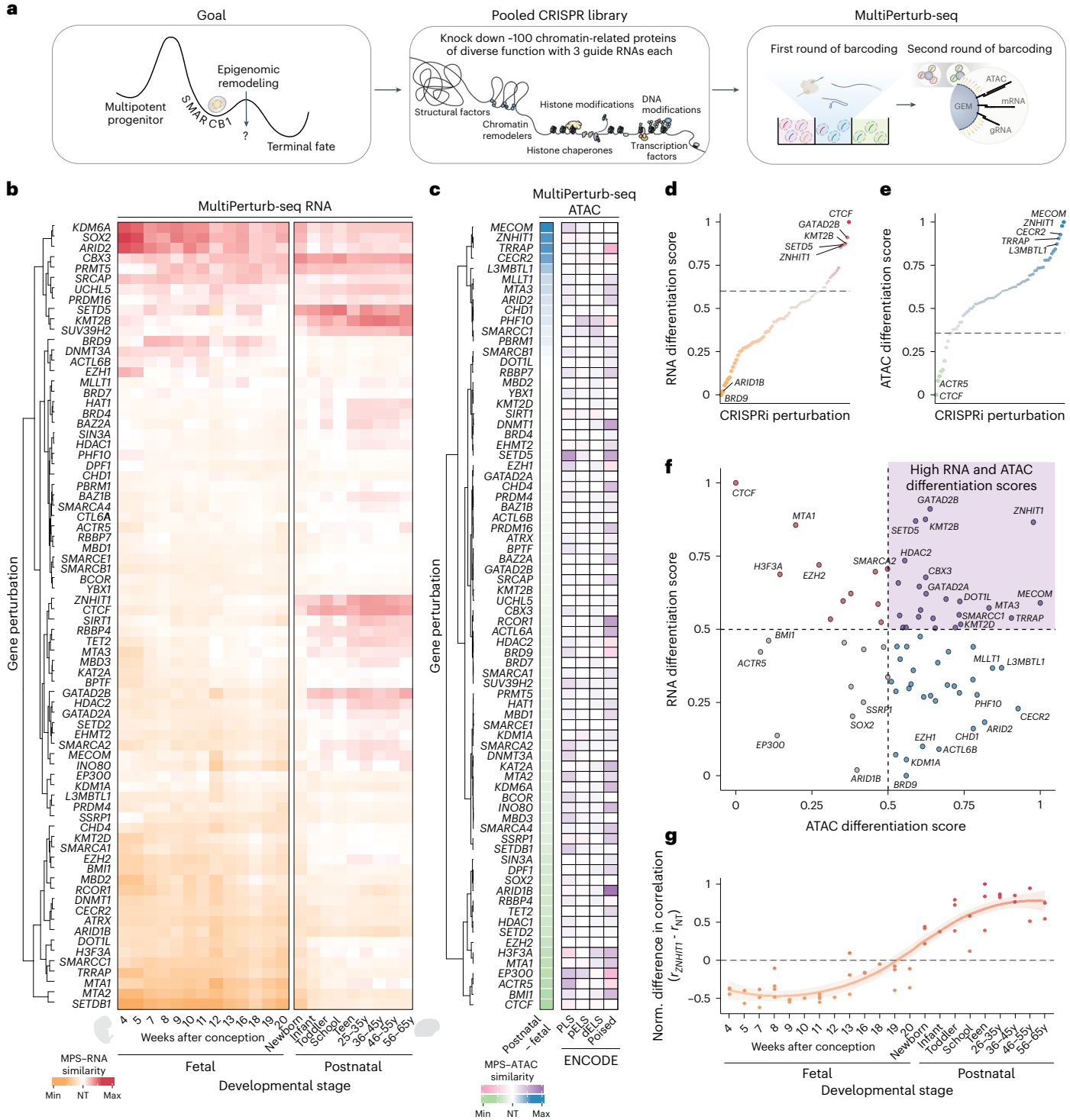
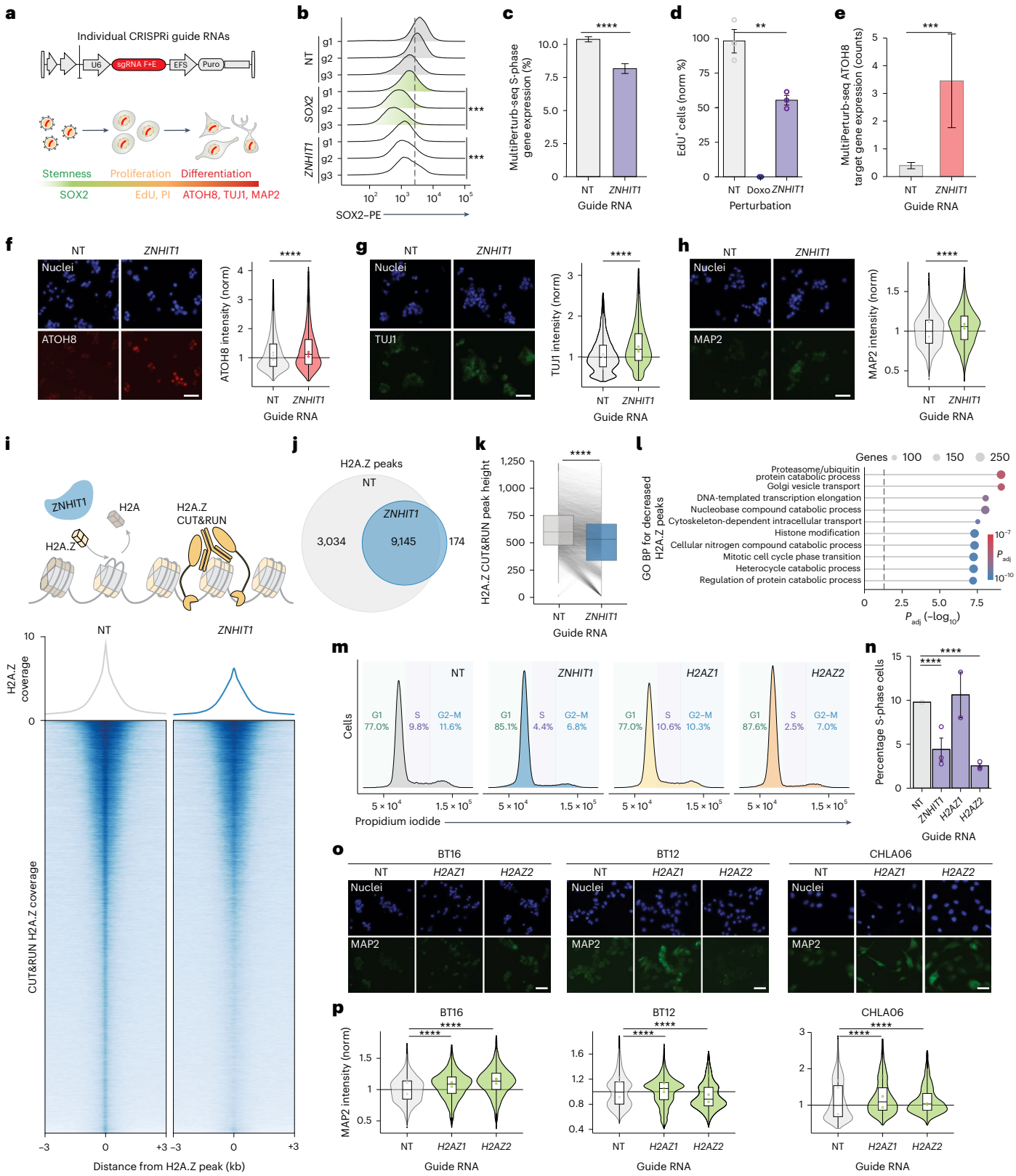


Fig. 2 | MultiPerturb-seq identifies genetic perturbations that trigger differentiation in AT/RT. a, Overview of differentiation challenge in AT/RT brain tumors and design of pooled CRISPR library to identify chromatin remodelers for cancer reprogramming therapy. **b**, Correlation between gene-perturbed human AT/RT cells and gene expression over developmental stages from 4 weeks after conception to senior adulthood[29]. The Pearson correlation is computed on the top 1,000 HVGs and values are normalized such that cells receiving an NT perturbation display as zero on the color scale. **c**, Left, correlation between gene-perturbed human AT/RT cells and open chromatin peaks in developmental[35] and adult[36] brain atlases. Right, sum of fold changes ($\log_2$) at peaks overlapping ENCODE regulatory elements[37]. The Pearson correlation is computed on the top 1,000 HVPPs and values are normalized such that cells receiving an NT perturbation display as zero on the color scale. 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 (**d**) and ATAC differentiation score (**e**). Higher values indicate greater similarity to postnatal primary brain tissues (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* and NT), we computed the Pearson correlation of gene expression with human brain developmental expression ($n = 53$ primary cerebrum samples at the indicated developmental time points). The line denotes the locally estimated scatterplot smoothing (LOESS) fit and the shaded region indicates the 95% confidence interval.

https://doi.org/10.1038/s41587-024-02475-x



Using intracellular antibody labeling and flow cytometry, we found diminished expression of the pluripotency-associated transcription factor SOX2 after knockdown of *ZNHIT1* compared to NT gRNA controls (Fig. 3b and 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 the S phase, we evaluated the relative proportion of cells in the 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 NT controls. We confirmed this by assaying changes in proliferation through incorporation of the thymidine analog 5-ethynyl-2′-deoxyuridine (EdU) after a 30-min pulse and found

Fig. 3 | *ZNHIT1* loss drives AT/RT cell-cycle arrest and differentiation through decreased H2A.Z deposition. a, CRISPRi validation in AT/RT cells to assess stemness, proliferation and differentiation after *ZNHIT1* loss. EFS, EF1α-short. **b**, SOX2 expression in cells receiving *ZNHIT1*, *SOX2* or NT 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) served as a positive control for cell-cycle arrest. Significance was determined using a one-way analysis of variance (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 ATOH8 (**f**), TUJ1 (**g**) and MAP2 (**h**) in BT16 cells with *ZNHIT1*-targeting or NT gRNAs ($n = 3$ biological replicate gRNAs per condition, with three technical replicates per gRNA). Open circles represent the median for each gRNA. Scale bar, 50 µm. **i**, Top, H2A.Z CUT&RUN in BT16 cells. Bottom, 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 four of out five 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 using a two-sided paired *t*-test. **l**, Enriched gene ontology biological processes (GO BP) 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 NT gRNAs or gRNAs targeting *ZNHIT1*, *H2AZ1* or *H2AZ2* ($n = 2$–3 gRNAs per perturbed gene). **n**, Quantification of S-phase cells; significance was determined using a two-sided χ^2-test ($n = 2$–3 gRNAs per perturbed gene). **o**, Representative immunofluorescence images of MAP2 expression in BT16, BT12 and CHLA06 AT/RT cells with NT gRNAs or gRNAs targeting *H2AZ1* or *H2AZ2*. Scale bar, 50 µm. **p**, Quantification of MAP2 expression in BT16, BT12 and CHLA06 AT/RT cells with NT gRNAs or gRNAs targeting *H2AZ1* or *H2AZ2* ($n = 3$ biological replicate gRNAs per condition, with three technical replicates per gRNA). Open circles represent the median for each gRNA. $^{**}P < 10^{-2}$, $^{***}P < 10^{-3}$ and $^{****}P < 10^{-4}$. Unless specified otherwise, significance was determined using a two-sided Mann–Whitney *U*-test with a Bonferroni correction for multiple comparisons. Bar plots in **d**, **e** and **n** represent mean values ± s.e.m. Box plots show the median and interquartile range with whiskers indicating 1.5× the interquartile range.

that *ZNHIT1* knockdown decreased progression through the S phase by 43% relative to NT controls (Fig. 3d). Perturbation of related proteins (SRCAP complex cofactor 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 Fig. 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 (approximately a ninefold increase), compared to cells receiving an NT gRNA (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 and 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 cleavage under targets and release using nuclease (CUT&RUN) (Fig. 3i and Supplementary Fig. 12a). *ZNHIT1*-perturbed cells showed decreases in both the number and the magnitude of H2A.Z-bound peaks, including peaks near genes involved in the 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. Because 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 the S phase after knockdown of *H2AZ2* (74% decrease) and this result was consistent across different AT/RT cell lines (Fig. 3m,n and Supplementary Fig. 13), suggesting that the cell-cycle arrest mediated by *ZNHIT1* perturbation may work through its role in H2A.Z deposition. Furthermore, we found that loss of *H2AZ1* and/or *H2AZ2* increased expression of the mature neuronal marker MAP2 across three different AT/RT cell lines (Fig. 3o,p).

In summary, we present MultiPerturb-seq, a multiomic pooled CRISPR screening platform that captures ATAC, RNA and CRISPR perturbations. This method increases throughput more than tenfold 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 multimodal data without the need for computational integration methods and provide 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 we show the promise of ZNHIT1 and H2A.Z modulation for AT/RT differentiation, although 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 gRNAs with a spacer near the 5′ end (for example, CRISPR–Cas9, CRISPRa, prime editing and 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 time points in a single experiment. Taken together, MultiPerturb-seq brings together epigenome and transcriptome phenotyping to study the impact of many genetic perturbations.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-024-02475-x.

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.e17 (2016).

2. Replogle, J. M. et al. Mapping information-rich genotype–phenotype landscapes with genome-scale Perturb-seq. *Cell* **185**, 2559–2575.e28 (2022).
3. 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).
4. Mimitou, E. P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. 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.e17 (2019).
6. Liscovitch-Brauer, N. et al. Profiling the genetic determinants of chromatin accessibility with scalable single-cell CRISPR screens. *Nat. Biotechnol.* **39**, 1270–1277 (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. *Nat. Commun.* **12**, 2969 (2021).
8. Cheng, Y. et al. Perturb-tracing enables high-content screening of multiscale, global 3D genome regulators. *Biophys. J.* **123**, 83a (2024).
9. Morris, J. A., Sun, J. S. & Sanjana, N. E. Next-generation forward genetic screens: uniting high-throughput perturbations with single-cell analysis. *Trends Genet.* **40**, 118–133 (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. *Nat. Methods* **18**, 635–642 (2021).
12. Lareau, C. A. et al. Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat. Biotechnol.* **37**, 916–924 (2019).
13. Zhang, H. et al. txci-ATAC-seq: a massive-scale single-cell technique to profile chromatin accessibility. *Genome Biol.* **25**, 78 (2024).
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.e20 (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. *Nat. Biotechnol.* **37**, 1452–1457 (2019).
18. Zhu, C. et al. An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat. Struct. Mol. Biol.* **26**, 1063–1070 (2019).
19. Datlinger, P. et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301 (2017).
20. Hill, A. J. et al. On the design of CRISPR-based single-cell molecular screens. *Nat. 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. *Sci. Adv.* **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. *Clin. Cancer Res.* **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. *J. Clin. Oncol.* **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. *Nat. Commun.* **10**, 1881 (2019).
26. Nakayama, R. T. et al. SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nat. Genet.* **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. *Nat. Genet.* **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. *Mol. Cell. Biol.* **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. *Sci. Rep.* **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 Rep.* **18**, 1701–1720 (2023).
33. Chuang, Y. Y. et al. Role of synaptojanin 2 in glioma cell migration and invasion. *Cancer Res.* **64**, 8271–8275 (2004).
34. Wang, X. et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat. Genet.* **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.e19 (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. *Nat. Commun.* **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 Div.* **14**, 15 (2019).
40. Divvela, S. S. K., Saberi, D. & Brand-Saberi, B. Atoh8 in development and disease. *Biology* **11**, 136 (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 Res.* **22**, 307–321 (2012).
42. Mimitou, E. P. et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* **39**, 1246–1258 (2021).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2024

Methods

Cell lines

BT12 and BT16 cells were gifts from P. Houghton (University of Texas Health, San Antonio), R. Hashizume (University of Alabama, Birmingham) and C. Roberts (St. Jude Children's Research Hospital). NIH-3T3 (CRL-1658) and CHLA06 (CRL-3038) cells were acquired from the American Type Culture Collection. HEK293FT cells were acquired from Thermo Fisher Scientific (R70007). BT12 and BT16 cells were validated by short tandem repeat profiling; other lines were authenticated by the vendor. All cell lines were maintained at 37 °C and 5% CO_2 in D10 medium (DMEM with high glucose and stabilized L-glutamine (Caisson, DML23) supplemented with 10% Serum Plus II (Sigma-Aldrich, 14009C)). Monoclonal CRISPRi-expressing BT16 cell lines were generated by transducing cells with lentiCRISPRi(v2)-Blast (Addgene, 170068)[14], selecting with 10 μg ml^{-1} blasticidin S (Thermo Fisher Scientific, 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 10 μg ml^{-1} blasticidin for 1 week.

gRNA 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 three 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 NT controls that do not target anywhere in the human genome. The library was designed using gRNAs from the Dolcetto CRISPRi library and CRISPick[43]. Three gRNAs were selected per gene and homopolymers were excluded. Oligonucleotides were ordered and synthesized by Twist Biosciences in a pooled format. For the mouse spike-in, mouse NT gRNAs were ordered individually through Integrated DNA Technologies (IDT) and pooled for library cloning. Pooled and arrayed gRNA sequences are provided in 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 (New England Biolabs (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) and guide spacers were extracted using cutadapt[44] (version 4.0) and aligned with Bowtie (version 1.1.2)[45]. For the epigenomic remodeler library, we recovered 98% of the designed gRNAs and, using the read distribution, we computed that the 90th:10th quantile ratio of gRNAs was 1.8. For the NT library (mouse), we recovered 100% of the designed gRNA and the 90th:10th quantile ratio was 6.5.

Lentivirus production

Lentiviral libraries were prepared in T225 flasks. Each flask was seeded with 27 × 10^6 cells the day before in 30 ml of antibiotic-free D10 medium to achieve 80–90% confluence before transfection. The transfection mix was 24.9 μg of the transfer plasmid (including the epigenetic remodelers or mouse NT library), 13.7 μg of pMD2.G (Addgene, 12260), 19.9 μg of psPAX2 (Addgene, 12259), 2,490 μl of OptiMEM (Invitrogen, 51985-091) and 138 μl of 1 mg ml^{-1} polyethylenimine linear (molecular weight: 25,000; Polysciences, 23966). The mixture was mixed and allowed to incubate for 10 min at room temperature. After removing 15 ml of the medium from the cells, the mixture was added dropwise. Then, 6 h after transfection, an additional 15 ml of fresh medium with 1% BSA (VWR, AAJ65097-18) was added. Viral supernatants were collected 72 h after transfection, spun down and filtered with a 0.45-mm filter (Millipore, SE1M003M00). Lentivirus for the pooled library was concentrated using 2 ml of a 20% sucrose cushion by ultracentrifugation for 2 h 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. On the basis of this titration, cells were infected with the appropriate volume of virus. Then, 48 h after transduction, BT16 and NIH-3T3 cells were lifted and selected with 1 μg ml^{-1} and 2 μg ml^{-1} puromycin, respectively (Invivogen, ant-pr-1). At the same time, we performed inline controls in six-well plates and confirmed that $p_{survival}$ was within the 1–5% target. Then, 7 days after 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

A detailed protocol is provided in the Supplementary Protocol. Primer sequences are provided in Supplementary Table 1.

Part 1: nuclear isolation, tagmentation and RT. Overall, our ATAC protocol is similar to a previous, well-validated ATAC method[46] and our transposomes were assembled as in Picelli et al.[47] 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 (IDT), we found that high-performance liquid chromatography (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 the use of HPLC-purified oligonucleotides because only one MEDS-B is required.

Human cells (2.4 million) and mouse cells (600,000) were combined and lysed in 1 ml of Omni lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP-40 (Thermo Fisher Scientific, 85124), 0.1% Tween-20 (Sigma-Aldrich, P1379) and 0.01% digitonin (Promega, G9441))[48]. Cells were lysed for 10 min on ice. After lysis, nuclei were spun down, pooled, resuspended in 450 μl of PBS and combined with tagmentation mix (240 μl of 5× TD-TAPS (50 mM TAPS–NaOH buffer pH 8.5 (Boston BioProducts, BB-2375), 25 mM $MgCl_2$ and 50% DMF (Sigma-Aldrich, 494488)), 120 μl of 10% Tween-20, 300 μl of dilution buffer (10 mM Tris-HCl pH 7.4, 100 mM NaCl, 50% glycerol and 1 mM DTT) and 30 μl of RiboLock RNase inhibitor (Thermo Fisher Scientific, EO0381). The nuclei were then split among wells of barcoded transposomes for tagmentation.

Cells were then incubated for 30 min at 37 °C in tagmentation mix while shaking at 350 rpm on a ThermoMixer. Following tagmentation, 1 μl of 126 mM EDTA was added to each well and mixed to stop tagmentation. After this, 50 μl of PBS was added and nuclei were spun at 400 rcf for 4 min at 4 °C. Then, 53 μl of supernatant was removed, leaving 17 μl and the nuclear pellet undisrupted. For RT, we added a master mix of 8 μl of 5× RT buffer (Thermo EP0742: 250 mM Tris-HCl, 375 mM KCl, 15 mM $MgCl_2$ and 50 mM DTT), 2 μl of dNTPs, 2 μl of MPSprimer_51 (10 μM), 4 μl of MPSprimer_52 (10 μM), 2 μl of Maxima RT H-minus (Thermo Fisher Scientific, EP0753) and 1 μl of Ribolock (Thermo Fisher Scientific, EO0381) per well. We then added 4 μl of barcoded TSOs (sequences for the 48 barcoded TSOs were MPSprimer_53 to MPSprimer_100) to match the ATAC barcodes to individual wells.

The plate was then incubated for 90 min at 53 °C, while shaking at 450 rpm on a ThermoMixer. An alternative RT 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 (in situ sequencing hetero RNA–DNA-hybrid after ATAC-seq)[49] 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 two 1.5-ml tubes, spun down and repooled in a 1.5-ml tube. The narrowed pipette tip was produced using a standard plastic 20-µl pipette tip (Rainin) melted to a 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 8 µl) and combined with 7 µl of ATAC buffer B (10x Genomics, PN2000193).

Part 2: 10X ATAC GEM generation, barcoding and cleanup. The nuclear 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 (56.5 µl of barcoding reagent B (PN2000194), 1.5 µl of reducing agent A (PN2000087) and 2 µl of 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, 100 µl of GEMs were collected and transferred to a PCR tube for GEM incubation. For the linear amplification step, we performed 15 cycles instead of 12 cycles. GEMs were then cleaned with Dynabeads per the manufacturer's instructions and libraries were split into 20-µl ATAC and 20-µl RNA libraries for final library prep. We recovered ~3.6 cells per droplet on average.

Part 3: library preparation. The ATAC fraction (20 µl) was cleaned up with 1.2× solid-phase reversible immobilization (SPRI; Illumina), amplified with a 100-µl reaction using NEBNext (50 µl of 2× High-Fidelity 2X master mix (NEB, M0541S), 5 µl of MPSprimer_101 (10 µM), 5 µl of MPSprimer_102 (10 µM), 20 µl of ATAC fraction and 20 µl of water; protocol: 30 s at 98 °C, then 10–15 cycles of 10 s at 98 °C, 30 s at 63 °C and 1 min at 72 °C, followed by 2 min at 72 °C and a hold at 4 °C) and then cleaned with double-sided SPRI (0.45×, 1.8×) to isolate fragments of lengths 50–1,000 bp. The RNA (complementary DNA and gRNA) fraction (20 µl) was cleaned by incubation with 8 µl of ExoSAP for 15 min at 37 °C and then 15 min at 80 °C. To make 100 µl of ExoSAP, we combined 1 µl of exonuclease I (NEB, M0293), 20 µl of shrimp alkaline phosphatase (NEB, M0371) and 79 µl of water. The cleaned RNA product was amplified with a 100-µl KAPA HiFi reaction (Roche, 07958935001) using 50 µl of 2X master mix, 2.5 µl of MPSprimer_101 (10 µM), 2.5 µl of MPSprimer_103 (10 µM), 2.5 µl of MPSprimer_104 (10 µM), 28 µl of cleaned RNA product and 14.5 µl of water (protocol: 3 min at 95 °C, then ten cycles of 20 s at 95 °C, 30 s at 66 °C and 1 min at 72 °C), followed by 2 min at 72 °C and a hold at 4 °C). Following amplification, the RNA and gRNA fractions were split using a two-sided SPRI[4]. The RNA was collected with a 0.6× SPRI and the gRNA was isolated from the supernatant using an additional 1.4× SPRI. Each fraction was then resuspended in 10 µl of water. The RNA could then be amplified with 3–9 additional cycles of a 50-µl reaction if there was less than 1 ng of product (25 µl of 2× KAPA HiFi master mix, 1.25 µl of MPSprimer_101 (10 µM), 1.25 µl of MPSprimer_103 (10 µM), 10 µl of cleaned RNA product and 12.5 µl of water; protocol: 3 min at 95 °C, then 3–9 cycles of 20 s at 95 °C, 30 s at 66 °C and 1 min at 72 °C, followed by 2 min at 72 °C and a hold at 4 °C).

After this, the 10-µl RNA fraction was tagmented with Tn loaded with MPSprimer_107 in 20 µl of tagmentation buffer for 5 min at 55 °C. The RNA was then purified with DNA Clean and Concentrator 5 (Zymo D4014), resuspended in 33.5 µl of water and PCR-amplified with 50 µl of PfuX7 reaction[50] (10 µl of 5× GC buffer, 1 µl of dNTPs, 2.5 µl of MPSprimer_101 (10 µM), 2.5 µl of MPSprimer_108 (10 µM), 0.5 µl of PfuX7 polymerase and 33.5 µl of RNA fraction; protocol: 5 min at 72 °C, 30 s at 98 °C, then ten cycles of 10 s at 98 °C, 30 s at 61 °C and 1 min at 72 °C, followed by 2 min at 72 °C and a hold at 4 °C). The 10-µl gRNA fraction was cleaned with 4 µl of 0.2 U μl^{-1} ExoSAP and amplified with a 50-µl intermediate PCR (25 µl of 2× KAPA HiFi master mix with 1.25 µl of biotinylated guide scaffold primer (MPSprimer_105, 10 µM), 1.25 µl of MPSprimer_101 (10 µM), 10 µl of gRNA fraction and 8.5 µl of water; protocol: 3 min at 95 °C, then ten cycles of 20 s at 95 °C, 30 s at 64 °C and 1 min at 72 °C), followed by 2 min at 72 °C and a hold at 4 °C), before cleaning again with 1.8× SPRI, resuspending in 10 µl of water and incubating with 4 µl of ExoSAP. Following cleanup, the gRNA was pulled down with Dynal MyOne Dynabeads Streptavidin C1 (Thermo Fisher Scientific, 65001), resuspended in 45 µl of water and amplified with a final inner (guide library) PCR using KAPA HiFi master mix (50 µl of master mix, 2.5 µl of MPSprimer_101 (10 µM), 2.5 µl of MPSprimer_106 (10 µM) and 45 µl of gRNA pulldown product; protocol: 3 min at 95 °C, then ten cycles of 20 s at 95 °C, 30 s at 57 °C and 1 min at 72 °C, followed by 2 min at 72 °C and a hold at 4 °C). Samples were evaluated with Tapestation high-sensitivity D1000 ScreenTape and reagents (Agilent, 5067), quantified with Qubit (Thermo Fisher Scientific, Q33231) and sequenced using both Illumina MiSeq and Illumina NovaSeq 6000 v1.5 platforms with a 16-bp index 1, 8-bp index 2 and 50-bp (MiSeq) or 100-bp (NovaSeq) reads 1 and 2.

MultiPerturb-seq optimization

MultiPerturb-seq was developed incrementally, first incorporating ATAC and then RNA and gRNA capture, ensuring preservation of each modality throughout the process (several key examples are 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 on the basis of previously published protocols[46,48,51,52] (Supplementary Fig. 4a,b). Both Tn5 (ref. 53) and TnY[6] were used in these experiments. We then adapted the direct guide capture technique from a previous study[4], as also described in another study[54]. We designed a TSO[55] with barcode and UMI (Supplementary Fig. 4c) and tested PCR[56,57] and cleanup conditions to achieve RNA and gRNA capture (Supplementary Fig. 4c–h). We also tested several variants of TSO (MPSprimer_110 to MPSprimer_112) (Supplementary Fig. 4e). Additionally, we tested different methods to amplify or enrich the RNA and gRNA, such as biotin pulldown. Lastly, we ensured trimodality integrity, confirming that tagmentation was stopped before RT, to avoid tagmenting the RNA–DNA heteroduplex[58] (Supplementary Fig. 4i,j). Agarose gels in Supplementary Fig. 4 are 1–2% with a 1-kb Plus DNA ladder (NEB, N3200L) unless otherwise noted. For cost estimates, we used the method's calculated cost, when provided, or estimated the cost on the basis of major cost drivers (for example, 10x Genomics kits). Sequencing cost was not included in the estimates.

Read alignment and preprocessing

For alignment and preprocessing (Supplementary Fig. 5a), we demultiplexed reads using bcl2fastq (version 2.20) with FASTQs for index reads. Reads were then trimmed with cutadapt[44] (version 4.0) to extract barcode 1 (well barcode), barcode 2 (droplet barcode), ATAC reads, RNA reads, gRNA reads and UMIs according to position (Supplementary Fig. 5b), then aligned separately (Supplementary Fig. 5c). Barcodes and gRNA spacers were aligned with Bowtie (version 1.1.2)[45] 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 Bowtie 2 (ref. 59) (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; https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCh38-and-mm10-2020-A.tar.gz). Open chromatin peaks were called using MACS2 (ref. 60) 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' and then reads were assigned to peaks using bedtools window (version 2.30.0) with a 100-bp window around the start position. RNA reads were aligned with STAR[61] (version 2.7.3a) using the settings '--quantMode GeneCounts --soloFeatures GeneFull_Ex50pAS' and then annotated with subread[62] 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), UMIs if applicable and aligned/annotated reads. These were then deduplicated using awk on the basis of barcode, UMI and position before importing into R (version 4.2.3), reformatting as a count matrix using DropletUtils[63] (version 1.18.1) and storing as a SingleCellExperiment[64] 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 three 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.

gRNA assignment

We implemented an algorithm that collapsed highly similar UMIs within the same cell. We did this because, within individual cells, we sometimes identified gRNA 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 toward overamplified reads. The algorithm first ranked UMIs on the basis of their read count, assuming that the UMI with the most reads represented the original molecule, which was then mutated during sequencing or PCR. Subsequently, the algorithm recursively removed UMIs that were within a Levenshtein distance[65] 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 gRNA, 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 (pseudobulk) by perturbation and compared to published transcriptomic[29] and accessible chromatin[35,36] atlases by computing the Pearson correlation across the top 1,000 highly variable genes (HVGs) or highly variable promoter-adjacent peaks (HVPPs). Correlations were computed between each perturbation-specific pseudobulk 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 HVGs. We defined HVGs as those genes with the largest s.d. across cerebrum samples ($n = 53$ samples from 4 weeks after conception to adulthood with 1–4 donors per developmental stage for that tissue). To compute correlations between MPS and the transcriptomic developmental atlas at specific time points, we took the Pearson correlations using only the top 1,000 HVGs.

For analysis of MultiPerturb-seq open chromatin, we first identified HVPPs. We defined HVPPs as those peaks within 2 kb of a protein-coding gene TSS with the greatest standard deviation over a unified sample of the MPS ATAC-seq dataset ($n = 77$ perturbation pseudobulk samples) and accessible chromatin prenatal or postnatal primary tissues ($n = 8$ prenatal samples of different brain cell types and $n = 1$ postnatal sample from frontal cortex). To compute correlations between MPS and the accessible chromatin developmental atlases, we took the Pearson correlations using only the top 1,000 HVPPs.

We computed normalized differentiation scores for either gene expression or open chromatin by taking the difference between correlations (Pearson) with late (postnatal) time points and early (prenatal) time points to identify those perturbations that increased in the same way as mature tissues. This difference was computed using the mean of the correlations over each postnatal or prenatal time point. That is, we computed one mean correlation across prenatal time points and one mean correlation across postnatal time points, normalized each mean correlation and then took the difference between these normalized means. For the normalization (over perturbations), for each stage (prenatal or postnatal), we computed maximum and minimum values over perturbations and then assigned each perturbation a normalized $r_i^{\text{norm}} = (r_i - \min(r))/(\max(r) - \min(r))$.

Differentially expressed genes, peaks and signatures

To identify differentially expressed genes and peaks, we used SCEPTRE[66], a nonparametric tool that resamples perturbations to infer associations with gene expression[66] 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 ten 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 (±2 kb) and RNA transcription factor transcription factor signatures from msigdb. Gene ontology enrichment analyses were performed using enrichGO in clusterProfiler[67] (version 4.6.2).

CROP-Multiome

We recloned our epigenomic remodeler library into CROP-seq-opti[20] (Addgene, 106280), a vector that places the gRNA within a polyadenylated RNA transcript, thus allowing capture by the 3′ poly(A) 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 were performed as described above for MultiPerturb-seq. We loaded 10,000 cells on one 10x Multiome lane per the manufacturer's instructions. In brief, 4 days after infection, 200,000 cells (80% BT16 cells and 20% NIH-3T3) were trypsinzied, washed and lysed in 500 µl of chilled lysis buffer (10x Genomics) with 12.5 µl of Ribolock RNase inhibitor (Thermo Fisher Scientific, EO0381). Cells were washed three times with 1 ml of wash buffer (10x Genomics) with 12.5 µl of Ribolock and 16,100 cells were resuspended in 10 µl of transposition mix (10x Genomics) and incubated for 60 min at 37 °C. Following tagmentation, the mix was loaded onto a GEM chip as instructed and run on the Chromium Controller X (10x Genomics). Following incubation, 5 µl of 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 using the Illumina NovaSeq 6000 version 1.5 platform with a 34-bp index 1, 24-bp index 2 and 125-bp reads 1 and 2 and count matrices were generated with cellranger-arc (version 2.0.2, 10x Genomics). Polyadenylated gRNA identities were 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[68], we used the CUTANA ChIC/CUT&RUN Kit (EpiCypher, 14-1048) with antibodies to H2A.Z (Abcam, ab4174), H3K4me3 (EpiCypher, 14-1048) and IgG (EpiCypher, 14-1048). BT16 cells were transduced with a *ZNHIT1*-targeting or an NT (negative control) gRNA ($n = 5$ biological replicate transductions per gRNA). Then, 2 days later, cells were lifted and selected with 1 μg ml^{-1} puromycin. An inline control was used to ensure complete selection. Then, 5 days after transduction, cells were collected for CUT&RUN and 500,000 cells were used per condition. Cells were lifted, washed and bound to 10 μl of activated concanavalin A-conjugated paramagnetic beads (EpiCypher), before resuspending with 0.5 μg of the antibody of interest and incubating overnight at 4 °C on a rotator. The next day, the beads were washed twice with permeabilization buffer and incubated with 2.5 μl of pAG-MNase (Epicypher) for 10 min at room temperature. After binding, the beads were washed and 2 mM $CaCl_2$ was added to begin digestion. Digestion was allowed to proceed for 2 h at 4 °C; then, the reaction was terminated by adding 33 μl Stop Buffer (Epicypher) and incubating the reactions at 37 °C for 10 min. We included a 0.5-ng *Escherichia 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 (NEB, E7645S), pooled and sequenced using Illumina NovaSeq S1 6000 version 1.5 with 2× 90-bp paired-end reads.

Files were trimmed with Trim Galore (version 0.6.10) with options '--fastqc --paired' and then aligned to hg38 (GRCh38.p14; https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz) using Bowtie 2 (ref. 59) (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[69] (version 0.7.0) view with the options '-f bam -F [XS] == null and not unmapped and not duplicate'. Peaks were called with MACS2 (ref. 60) 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 ten samples from both conditions. Using valr[70] (version 0.7.0), we only retained peaks called by at least four biological replicates of the same condition (*ZNHIT1*-targeting or NT) and termed these reproducible peaks. Gene ontology enrichment was computed using clusterProfiler[67] enrichGO (version 4.6.2) on decreased reproducible peaks.

For visualization, bigwig files were created using deepTools[71] 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 and NT) with the median change in mean coverage at the peak maximum (that is, median over all 25 possible pairings of 5 *ZNHIT1*-targeting replicates × 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 TSS of all protein-coding genes expressed at ten transcripts per million or more in BT16 cells[72]. 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 (ref. 73) as '--blackListFileName' to filter out reads aligning to problematic genome regions before plotting using plotHeatmap.

Arrayed CRISPRi validation

For arrayed validation, BT16, BT12 and/or CHLA06 cells with lentiCRISPRi(v2)-Blast were transduced with gRNAs in lentiGuideFE-Puro (Addgene, 170069). The gRNAs were designed using the Dolcetto CRISPRi library and CRISPick[43] and then synthesized by IDT (Supplementary Table 2). The backbone was digested with *Bsm*BI (Thermo Fisher Scientific, FD0454) and oligos were annealed, phosphorylated and ligated into the lentiGuideFE-Puro backbone. Lentivirus was produced as described above (scaled to six-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 h, cells were replated in medium with puromycin (1 μg ml^{-1}) and selected for at least 2 days with confirmation of complete selection using an inline selection control.

SOX2 staining and flow cytometry

To label and quantify SOX2-positive cells, cells were lifted, washed and stained with live–dead violet (Thermo Fisher Scientific, L34963) (diluted 1:400, 15 μl for 1×10^6 cells) for 5 min at room temperature, before washing with PBS, fixing with 1% formaldehyde and incubating at room temperature for 10 min on a rotator (Thermo Fisher Scientific digital tube revolver, 88881101)[74]. After fixation, cells were quenched with 0.125 M glycine (addition of 2.5 M glycine), washed with PBS and lysed with 100 μl of a previously optimized lysis buffer[74] (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP-40 and 1% BSA) on ice for 5 min. Then, they were washed with 1 ml of wash buffer (same as for lysis but without NP-40) and blocked in 1 ml of PBS with 3% BSA for 30 min at room temperature. Following blocking, they were washed and resuspended in 100 μl of PBS with 3% BSA and antibody (1:100, 1 μg for 5×10^6 cells, anti-SOX2; Biolegend, 656104) for 60 min at room temperature. They were then washed twice more (PBS with 3% BSA and 1% Tween) and resuspended in PBS with 1% BSA and 2 mM EDTA for flow cytometry (Sony, SH800). Sequential gating was performed as follows: exclusion of debris on the basis of forward scatter (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, the medium was aspirated and cells were washed, fixed with 4% paraformaldehyde (diluted 1:4 from 16%; Electron Microscopy Sciences, 15710-S) for 15 min and then washed with PBS. Cells were then permeabilized with 0.2% Tween-20 for 5 min and blocked with PBS with 0.2% Tween-20 and 3% BSA for 1 h. Cells were then incubated overnight at 4 °C with the following primary antibodies: anti-TUJ1 at 1:1,000 dilution (BioLegend, 801201), anti-MAP2 at a 1:500 dilution (SYSY, 188004) or anti-ATOH8 (Thermo Fisher Scientific, PA5-65024) at a 1:400 dilution. The next day, cells were washed three times for 5 min with cold PBS. The corresponding secondary antibody was added at 1:800 dilution (Thermo Fisher Scientific, A-21202 for TUJ1 (mouse); Thermo Fisher Scientific, A-11073 for MAP2 (guinea pig); Thermo Fisher Scientific, 31572 for ATOH8 (rabbit)) with 2 mM Hoechst (Sigma-Aldrich, B2261) and incubated for 1 h at room temperature. Cells were then washed with PBS for an additional three washes. All steps were performed at room temperature on a rocker unless otherwise noted. Images were acquired with a ×20 objective using an epifluorescence microscope (Keyence, BZ-X800). Five images were acquired per well.

Quantitative image analysis was run in CellProfiler[75] (version 4.2.6). Primary objects were identified on the basis of the nucleus (Hoechst) with a threshold calculated using Otsu's method. Secondary objects (cytoplasm) were defined by extension from the nucleus (distance B method with a threshold calculated using 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 or object. TUJ1 and MAP2

signals (cytoplasmic) were quantified using the mean intensity per cell or object. For MAP2 images, we also applied flatfield illumination correction. Normalization was performed to the median intensity of cells or objects receiving NT gRNAs. Cells or objects with an assigned intensity (integrated or mean depending on the protein) greater than 3 s.d. from the NT mean were excluded as fluorescent debris.

EdU incorporation and cell-cycle analysis

Cells were labeled with EdU using the Click-iT EdU cell proliferation kit (Thermo Fisher Scientific, C10337). A total of 2,000 cells per well were plated on 96-well plates in triplicate. Cells were incubated with 10 µM EdU for 30 min, fixed with 4% PFA for 15 min and permeabilized with 0.5% Triton X-100 for 10 min at room temperature. Cells were then washed and incubated with the Click-iT reaction cocktail for 30 min. As a positive control, untransduced BT16 cells were exposed to 1 µM doxorubicin (MedChemExpress, HY-15142) to inhibit proliferation and EdU incorporation. After EdU staining, nuclei were stained with 2 µM Hoechst 3342 (Sigma-Aldrich, 4533) for 15 min and washed with PBS; then, images were acquired with a ×20 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[75] (version 4.2.6). Cells were quantified on the basis of Hoechst staining and binned into EdU-positive and EdU-negative cells according to the integrated intensity (sum) per cell or object using the ClassifyObjects module.

For propidium iodide (PI) staining, cells were pelleted in 1.5-ml tubes, washed once with 1 ml of PBS and resuspended in 300 µl of 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 1,000*g* for 4 min and ethanol was removed. Cells were washed with 1 ml PBS and stained with 0.5 ml of FxCycle PI/RNAse solution (Thermo Fisher Scientific, F10797) per 1 million cells. Pellets were resuspended and incubated for 15 min 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 FSC-A and SSC-H parameters followed by gating on singlets with FSC-A–FSC-H. The cell-cycle profile was modeled and gates were generated on the basis of the PI-A signal of the cell population with FlowJo (version 10.10.0) using a Watson model.

Reporting summary

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

Data availability

MultiPerturb-seq data can be downloaded from BioProject (accession number PRJNA1160410). The human genome hg38 (GRCh38.p14) was from the University of California, Santa Cruz 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) was 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/ (ref. 29), https://descartes.brotmanbaty.org/ (ref. 35) and http://catlas.org/humanbrain/ (ref. 36). Data from previously published studies were from the Sequence Read Archive or Gene Expression Omnibus: 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 is available from GitLab (https://gitlab.com/sanjanalab/mps).

References

43. Sanson, K. R. et al. Optimized libraries for CRISPR–Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).
44. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
45. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
46. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
47. Picelli, S. et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).
48. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
49. Xu, W. et al. ISSAAC-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. *Nat. Methods* **19**, 1243–1249 (2022).
50. Nørholm, M. H. H. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol.* **10**, 21 (2010).
51. Grandi, F. C., Modi, H., Kampman, L. & Corces, M. R. Chromatin accessibility profiling by ATAC-seq. *Nat. Protoc.* **17**, 1518–1552 (2022).
52. Preissl, S. et al. Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432–439 (2018).
53. Adey, A. C. Tagmentation-based single-cell genomics. *Genome Res.* **31**, 1693–1705 (2021).
54. Replogle, J. M. et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* **38**, 954–961 (2020).
55. 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).
56. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
57. Bagnoli, J. W. et al. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun.* **9**, 2937 (2018).
58. Di, L. et al. RNA sequencing by direct tagmentation of RNA/DNA hybrids. *Proc. Natl Acad. Sci. USA* **117**, 2886–2893 (2020).
59. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
60. Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
61. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
62. 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).
63. 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. *Nat. Commun.* **9**, 2667 (2018).
64. Amezquita, R. A. et al. Orchestrating single-cell analysis with Bioconductor. *Nat. Methods* **17**, 137–145 (2020).
65. Levenshtein, V. I. Binary codes capable of correcting deletions, insertions, and reversals. *Sov. Phys. Dokl.* **10**, 707–710 (1966).
66. Barry, T., Mason, K., Roeder, K. & Katsevich, E. Robust differential expression testing for single-cell CRISPR screens at low multiplicity of infection. *Genome Biol.* **25**, 124 (2024).
67. Xu, S. et al. Using clusterProfiler to characterize multiomics data. *Nat. Protoc.* https://doi.org/10.1038/s41596-024-01020-z (2024).

68. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856 (2017).
69. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
70. Riemondy, K. et al. valr: reproducible genome interval analysis in R. *F1000Res.* **6**, 1025 (2017).
71. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
72. Dharia, N. V. et al. A first-generation pediatric cancer dependency map. *Nat. Genet.* **53**, 529–538 (2021).
73. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).
74. Fiskin, E. et al. Single-cell profiling of proteins and chromatin accessibility using PHAGE-ATAC. *Nat. Biotechnol.* **40**, 374–381 (2022).
75. Stirling, D. R. et al. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021).

Acknowledgements

We thank the entire Sanjana laboratory for their support and advice. We also thank P. Smibert, C. Zhu and S. Hao for sharing their 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. Z.Z.G. is supported by the National Institutes of Health (NIH) T32 training grant (grant no. GM136573). E.K. is partially supported by the National Science Foundation (NSF) (DMS-2113072 and DMS-2310654). N.D. is supported by the NIH–Office of the Director (R03OD034499). N.D. and J.P.G. are supported by the Ty Louis Campbell Foundation. N.E.S. is supported by the NIH–National Human Genome Research Institute (DP2HG010099 and R01HG012790), NIH–National Cancer Institute (R01CA218668 and R01CA279135), NIH–National Institute of Allergy and Infectious Diseases (R01AI176601), NIH–National Heart, Lung, and Blood Institute (R01HL168247), the Simons Foundation for Autism Research (Genomics of ASD 896724), the MacMillan Center for the Study of the Non-coding 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 paper with input from all authors.

Competing interests

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

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-024-02475-x.

Correspondence and requests for materials should be addressed to Neville E. Sanjana.

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