Title: Multifactorial chromatin regulatory landscapes at single cell resolution

Authors: Michael P. Meers[1], Derek H. Janssens[1], Steven Henikoff[1,2]

Affiliations: [1]Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [2]Howard Hughes Medical Institute, Chevy Chase, MD, USA

Abstract: Chromatin profiling at locus resolution uncovers gene regulatory features that define cell types and developmental trajectories, but it remains challenging to map and compare distinct chromatin-associated proteins within the same sample. Here we describe a scalable antibody barcoding approach for profiling multiple chromatin features simultaneously in the same individual cells, Multiple Target Identification by Tagmentation (MulTI-Tag). MulTI-Tag is optimized to retain high sensitivity and specificity of enrichment for multiple chromatin targets in the same assay. We use MulTI-Tag to resolve distinct cell types using multiple chromatin features on a commercial single-cell platform, and to distinguish unique, coordinated patterns of active and repressive element regulatory usage in the same individual cells. Multifactorial profiling allows us to detect novel associations between histone marks in single cells and holds promise for comprehensively characterizing cell-specific gene regulatory landscapes in development and disease.

Main Text

Single-cell sequencing methods for ascertaining cell type-associated molecular characteristics by profiling the transcriptome[1-3], proteome[4-6], methylome[7,8], and accessible chromatin landscape[9,10], in isolation or in "Multimodal" combinations[11-15], have advanced rapidly in recent years. More recently, methods for profiling the genomic localizations of proteins associated with the epigenome, including Tn5 transposase-based Cleavage Under Targets & Tagmentation (CUT&Tag)[16,17], have been adapted for single cell profiling, overcoming the sparse incidence of such proteins in comparison with other molecular markers. The combinatorial nature of epigenome protein binding and localization[18-20] presents the intriguing possibility that a method for profiling multiple epigenome characteristics at once could both overcome the sparsity issue and derive important information about cell type-specific epigenome patterns at specific loci. However, we still lack methods for profiling multiple epigenome targets simultaneously in the same assay, in what might be considered "Multifactorial" profiling.

Motivated by this gap, and with the knowledge that CUT&Tag profiles chromatin proteins in single cells at high signal-to-noise ratio[16], we explored methods for physical association of a chromatin protein-targeting antibody with an identifying adapter barcode added during tagmentation that could be used to deconvolute epigenome targets directly in sequencing (Fig. 1a). Using antibodies against mutually exclusive H3K27me3 and PolIIS5P in human K562 Chronic Myelogenous Leukemia cells as controls, we systematically tested a variety of protocol conditions for antibody-barcode association with the goal of optimizing both assay efficiency and fidelity of target identification. We

first tested the pre-incubation of barcoded pA-Tn5 complexes with antibodies versus covalent conjugation of barcoded adapters to either primary or secondary antibodies to be loaded into pA-Tn5, and the combined incubation and tagmentation of all antibodies at the same time versus sequential tagmentation of targets one at a time. We found that both pre-incubation and combined tagmentation resulted in high levels of spurious cross-enrichment between targets (Supplementary Fig. 1a), leading us to use adapter-conjugated antibodies loaded into pA-Tn5 to tagment multiple targets in sequence. We next tested conjugating adapters directly to a primary antibody versus a secondary antibody, finding that the former resulted in superior target distinction (Fig. 1b-c, Supplementary Fig. 1a-b), but also variable data quality, likely owing to fewer pA-Tn5 complexes accumulating per target locus in the absence of a secondary antibody. To overcome this obstacle, we (1) Loaded pA-Tn5 onto 1˚ antibody-conjugated i5 forward adapters, (2) Tagmented target chromatin in sequence, and (3) Added a secondary antibody followed by pA-Tn5 loaded exclusively with i7 reverse adapters and carried out a final tagmentation step (Fig. 1a). This resulted in libraries that were as robust as matched CUT&Tag experiments, particularly for H3K27me3 (Supplementary Fig. 1c). We dubbed this combined approach Multiple Targets Identified by Tagmentation (MulTI-Tag) (Fig. 1a).

In K562 cells and H1 human embryonic stem cells (hESCs), we simultaneously profiled three targets that represent distinct waypoints in the temporal trajectory of developmental gene expression: H3K27me3, enriched in developmentally regulated heterochromatin[21,22], H3K4me2, enriched at active enhancers and promoters[23], and

H3K36me3, co-transcriptionally catalyzed during transcription elongation[24,25] (Fig. 1d-e). In comparison with control experiments in which each of the three targets was profiled individually, K562 and H1 MulTI-Tag both retain comparable efficiency of target-specific enrichment in peaks (Supplementary Fig. 2a-c). Moreover, both control and MulTI-Tag experiments exhibit characteristic patterns of enrichment for each mark, including H3K4me2 at promoters, H3K36me3 in gene bodies, and H3K27me3 across both (Fig. 1e, Supplementary Fig. 2d). Of note, in H1 hESCs only we observed overlap between H3K27me3 and H3K4me2 for both control and MulTI-Tag samples consistent with known "bivalent" chromatin[26]. These results show that MulTI-Tag retains high fidelity when scaled to multiple user-defined targets.

Given the successful adaptation of CUT&Tag for single cell profiling[16,27-29], we sought to use multifactorial profiling for single-cell molecular characterization (Fig. 2a). In an experiment profiling a mixture of human K562 cells and mouse NIH3T3 cells in single cell experiments, MulTI-Tag using either individual antibodies or multiple antibodies in combination showed similarly low cross-species collision rates (10%, 12%, 13%, and 17%; for H3K27me3, H3K4me2, and H3K36me3 alone; or all targets, respectively), indicating successful single cell isolation and profiling (Supplementary Fig. 3a). Moreover, pilot experiments in K562 cells profiled by MulTI-Tag yielded a comparable number of unique reads for each of the three targets as single cells profiled using only one antibody, indicating that MulTI-Tag is nearly additive under conditions in which amplification is saturating (Supplementary Fig. 3b). We therefore used MulTI-Tag to simultaneously profile H3K27me3, H3K4me2, and H3K36me3 in 348 K562 cells and 368 H1 cells (Fig.

2b). We found that in the majority of peaks (59.1% for H3K27me3; 95.7% for H3K4me2; 94.5% for H3K36me3), greater than 80% of fragments mapped within the peak were from the same target (Fig. 2c, Supplementary Fig. 4a-b), indicating that MulTI-Tag retains high specificity in single cells. We used Uniform Manifold Approximation and Projection (UMAP)[30,31] to project single cell data into low-dimensional space based on enriched features defined for each of the individual targets profiled and found that H3K27me3 and H3K4me2 were able to distinguish cell types with 99.9% and 95.2% efficiency, respectively, confirming that MulTI-Tag generates data from multiple chromatin targets that are informative for cell type classification (Fig. 2d). Notably, H3K36me3 was insufficient to separate cell types (32.3% efficiency), consistent with previous observations in single cell data and with the relative uniformity of H3K36me3 enrichment across different cell types[32]. An analysis of the 248 most informative enriched features for cell type distinction based on their values in the Singular Value Decomposition (SVD) input to UMAP showed that 100% of them were H3K27me3 features, consistent with its near-perfect distinction of highly dissimilar cell-type specific clusters (Supplementary Fig. 4c). Nevertheless, we were able to identify highly informative target-specific clusters for all three targets that showed cell type-specific patterns of enrichment (Supplementary Fig. 4d). These data show that MulTI-Tag is an effective method for profiling multiple informative chromatin targets in single cells.

Since MulTI-Tag uses barcoding to define fragments originating from specific targets, we can directly ascertain and quantify relative target abundances and instances of their co-occurrence at the same loci in single cells. By quantifying the number of unique reads

occurring for each target, we found that H3K27me3 was highly abundant relative to H3K4me2 and H3K36me3 in single cells, composing a mean of 89.4% of unique reads in K562 cells and 80.0% of unique reads in H1 cells, as compared with H3K4me2 contributing 5.1% and 8.8% and H3K36me3 contributing 5.5% and 11.2% in K562 and H1 cells, respectively (Fig. 3a, Supplementary Fig. 5a). Although several variables could contribute to this difference, including antibody binding efficiency and signal-to-noise ratio, it is consistent with previously reported mass spectrometry quantification of H3K27me3- vs. H3K4me2-containing peptides[33] and of single molecule imaging of H3K27me3- and H3K4me3-modified nucleosomes[34] in hESCs, suggesting the relative balances of target-specific fragments reflect biologically meaningful quantifications.

To quantify co-occurrence of specific target combinations in single cells, we mapped fragments from any target onto genes in a window from 1 kb upstream of the TSS to the gene terminus and clustered cells based on a binarized, low dimensional representation of that signal. We then categorized genes in each cell based on the co-occurrence of target-specific fragments, and finally mapped those categories onto clustered cells (Fig. 3b, Supplementary Fig. 5b). As expected, due to the relative abundances of target-specific fragments, the 363 most informative genes by all-target clustering were highly represented for H3K27me3 enrichment in the absence of other targets (Fig. 3c). However, many notable genes exhibited unique patterns of target co-occurrence that correlated with cell type, including H3K27me3-H3K4me2 co-enriched at the CREB5 and ZNF423 genes in H1 cells in comparison with exclusive H3K27me3 enrichment in K562 cells (Fig. 3d). There are also notable instances of co-enrichment across genes in the same single

cells, including cells with H3K4me2 and/or H3K36me3 enrichment in NR5A2 linked with H3K27me3 enrichment in HOXB3 in H1 cells, and vice-versa in K562 cells (Fig. 3d). These results show that multi-target intra-gene co-enrichment and cross-gene patterns can be directly observed in MulTI-Tag data.

We quantified the proportion of each class of single- or co-enrichment of targets in the same gene in single cells and found that although H3K27me3 in the absence of other targets is the most represented class, K562 cells have a significantly higher share of this class of genes than H1 cells, and a lower frequency of all other target combinations (Fig. 3e). This is consistent with the repeated observation that hESCs have a lower density of heterochromatin than many other cell types[35,36]. To quantify the degree of coordination between each pair of targets, we calculated Cramer's V of association, a variant of Pearson's Chi-square test that is robust to variable effect sizes and is therefore ideal for MulTI-Tag data with highly disparate target-specific counts. We found that H1 cells had a higher degree of association between H3K27me3 and H3K4me2 than K562 cells, consistent with observed bivalency (Fig. 3f). Curiously, the same was true for association between H3K27me3 and H3K36me3, despite previous observations that H3K27me3 and H3K36me3 appear to be antagonistic *in vitro* and *in* vivo[37,38] (Fig. 3F). Nevertheless, in bulk MulTI-Tag and in previously published ENCODE ChIP-seq data from H1 hESCs, we were similarly able to detect co-occurrence of H3K27me3 at the 5' ends and H3K36me3 at the 3' ends of several genes, including many involved in metabolic and developmental signaling and others that appear to have multiple regulated promoter isoforms (Supplementary Fig. 6a-b), suggesting our observation of single-cell co-occurrence is

unlikely to be a technical artifact. In contrast, we find that the association between H3K4me2 and H3K36me3 is low in both cell types, indicating that despite their independent associations with active transcription, the simultaneous enrichment of H3K4me2 at the promoter and H3K36me3 in the body of a gene is not overrepresented relative to chance. Together, these results shed light on patterns of chromatin structure at single cell, single locus resolution.

MulTI-Tag establishes a rigorous baseline for unambiguously profiling multiple epigenome proteins with direct sequence tags, maintaining both exemplary assay efficiency and target-assignment fidelity relative to other similar approaches[39,40]. Single-readout Multifactorial profiling holds a distinct advantage over Multimodal profiling, which often requires highly complicated integration of semi-compatible protocols and analysis methods. MulTI-Tag is theoretically scalable to any combination of user-defined targets in the same assay. Three targets profiled here, H3K27me3, H3K4me2, and H3K36me3, are typically enriched at distinct stages of the gene regulatory cycle that proceeds from developmental repression (H3K27me3) to enhancer and promoter activation (H3K4me2) to productive transcription elongation (H3K36me3). In the future we envision that integrating this temporal information across models of development and differentiation will aid in understanding the causative regulatory events that dictate how developmental transitions proceed at the single cell level prior to the downstream transcriptomic and proteomic outcomes. It will also be instructive to add targets to the same experiments that represent other waypoints in the cycle including RNA PolII pausing (PolIIS5P) and elongation (PolIIS2P), constitutive silencing (H3K9me3), and even direct transcription

factor binding. Furthermore, our analysis of co-occurrence of different targets in the same genes elucidates chromatin structure at single-locus, single-cell resolution in a way that heretofore has been impossible. We anticipate further work to understand intra-locus interactions between different chromatin characteristics to bear on long-standing hypotheses regarding bivalency[26] and hyperdynamic chromatin[35].

Although MulTI-Tag represents a significant advance in chromatin profiling, opportunities for refinement exist. For instance, it is possible that PCR "jackpotting" bias may suppress the equitable amplification of some target combinations. Methods to mitigate target-specific amplification bias could resolve this, though MulTI-Tag unique fragment depth appeared to be additive for the target combinations tested in this study. MulTI-Tag is also more complicated than CUT&Tag, requiring additional steps and the use of adapter-conjugated antibodies. Our emphasis on ensuring both that the efficiency of MulTI-Tag profiling was comparable to CUT&Tag in terms of signal-to-noise, and that each individual target was faithfully assigned with minimal cross-contamination between antibody-assigned adapters, led us to generate antibody-adapter conjugates[41], and to incubate and tagment with antibody-adapter-transposase complexes sequentially rather than simultaneously. By physically excluding the possibility of adapter or Tn5 monomer exchange in the protocol, MulTI-Tag safeguards against potential artifacts originating from adapter crossover, identifying any set of user-defined targets with high fidelity. We anticipate future reagent development and protocol improvements will enable methods in the style of MulTI-Tag to produce reliable multifactorial profiles while also minimizing the barriers to use for any lab seeking to obtain such data.

Methods

Cell culture and nuclei preparation

Human female K562 Chronic Myleogenous Leukemia cells (ATCC) were authenticated for STR, sterility, human pathogenic virus testing, mycoplasma contamination, and viability at thaw. H1 (WA01) male human embryonic stem cells (hESCs) (WiCell) were authenticated for karyotype, STR, sterility, mycoplasma contamination, and viability at thaw. K562 cells were cultured in liquid suspension in IMDM (ATCC) with 10% FBS added (Seradigm). H1 cells were cultured in Matrigel (Corning)-coated plates at 37°C and 5% CO_2 using mTeSR-1 Basal Medium (STEMCELL Technologies) exchanged every 24 hours. K562 cells were harvested by centrifugation for 3 mins at 1000xg, then resuspended in 1x Phosphate Buffered Saline (PBS). H1 cells were harvested with ReleasR (StemCell Technologies) using manufacturer's protocols. Lightly crosslinked nuclei were prepared from cells as described in steps 2-14 of the Bench Top CUT&Tag protocol on protocols.io (https://dx.doi.org/10.17504/protocols.io.bcuhiwt6). Briefly, cells were pelleted 3 minutes at 600xg, resuspended in hypotonic NE1 buffer (20 mM HEPES-KOH pH 7.9, 10 mM KCl, 0.5 mM spermidine, 10% Triton X-100, 20% glycerol), and incubated on ice for 10 minutes. The mixture was pelleted 4 minutes at 1300xg, resuspended in 1xPBS, and fixed with 0.1% Formaldehyde for 2 minutes before quenching with 60 mM glycine. Nuclei were counted using the ViCell Automated Cell Counter (Beckman Coulter) and frozen at -80˚C in 10% DMSO for future use.

Antibodies

Antibodies used for CUT&Tag or MulTI-Tag in this study were as follows: Rabbit Anti-H3K27me3 (Cell Signaling Technologies CST9733S, Lot 16), Mouse anti-RNA PolIIS5P (Abcam ab5408, Lot GR3264297-2), Mouse anti-H3K4me2 (Active Motif 39679, Lot 31718013), Mouse anti-H3K36me3 (Active Motif 61021, Lot 23819012), Guinea Pig anti-Rabbit (Antibodies Online ABIN101961), and Rabbit anti-Mouse (Abcam ab46450). For antibody-adapter conjugation, antibodies were ordered from manufacturers with the following specifications if not already available as such commercially: 1x PBS, no BSA, no Sodium Azide, no Glycerol. For secondary conjugate MulTI-Tag, secondary antibody conjugates from the TAM-ChIP Rabbit and Mouse kits (Active Motif) were used.

CUT&Tag

CUT&Tag was carried out as previously described[17] (https://dx.doi.org/10.17504/protocols.io.bcuhiwt6). Briefly, nuclei were thawed and bound to washed paramagnetic Concanavalin A (ConA) beads (Bangs Laboratories), then incubated with primary antibody at 4˚C overnight in Wash Buffer (10 mM HEPES pH7.5, 150 mM NaCl, 0.5 mM spermidine, Roche Complete Protease Inhibitor Cocktail) with 2mM EDTA. Bound nuclei were washed and incubated with secondary antibody for 1 hour at room temperature (RT), then washed and incubated in Wash-300 Buffer (Wash Buffer with 300 mM NaCl) with 1:200 loaded pA-Tn5 for 1 hour at RT. Nuclei were washed and tagmented in Wash-300 Buffer with 10 mM $MgCl_2$ for 1 hour at 37˚C, then resuspended sequentially in 50 µL 10 mM TAPS and 5 µL 10 mM TAPS with 0.1% SDS, and incubated 1 hour at 58˚C. The resulting suspension was mixed well with 16 µL of 0.9375% Triton X-100, then primers and 2x NEBNext Master Mix (NEB) was added for

direct amplification with the following conditions: 1) 58 ˚C for 5 minutes, 2) 72 ˚C for 5 minutes, 3) 98 ˚C for 30 seconds, 4) 98 ˚C 10 seconds, 5) 60 ˚C for 10 seconds, 6) Repeat steps 4-5 14 times, 7) 72 ˚C for 2 minutes, 8) Hold at 8 ˚C. DNA from amplified product was purified using 1.1x ratio of HighPrep PCR Cleanup System (MagBio) and resuspended in 25 µL 10 mM Tris-HCl with 1 mM EDTA, and concentration quantified using the TapeStation system (Agilent). For sequential and combined CUT&Tag, rather than incubating the secondary antibody and pA-Tn5 separately, pA-Tn5 was pre-incubated with an equimolar amount of secondary antibody in 50 µLWash-300 buffer at 4˚C overnight. For sequential, primary antibody incubation, secondary antibody-pA-Tn5 incubation, and tagmentation were carried out sequentially for each primary-secondary-barcoded pA-Tn5 combination, whereas for combined, all reagents were incubated simultaneously for their respective protocol steps (i.e. primary antibodies together, secondary antibody-pA-Tn5 complexes together), and tagmentation was carried out once for all targets.

Conjugates for MulTI-Tag

Antibody-adapter conjugates were generated by random amino-conjugation between 100 µg antibody purified in PBS in the absence of glycerol, BSA, and sodium azide, and 5' aminated, barcode-containing oligonucleotides (IDT) using Oligonucleotide Conjugation Kit (Abcam) according to manufacturer's protocols. Before conjugation, 200 µM adapter oligos resuspended in 1xPBS were annealed to an equimolar amount of 200 µM Tn5MErev (5'-[phos]CTGTCTCTTATACACATCT-3') in 1xPBS to yield 100 µM annealed adapters. In all cases, primary antibodies were conjugated with an estimated 10:1 molar

excess of adapter to conjugate. The sequences of adapters used are listed in Supplementary Table 1.

Bulk MulTI-Tag protocol

For each target to be profiled in MulTI-Tag, an antibody-i5 adapter conjugate was generated as described above, and 0.5 µg conjugate was incubated with 1 µL of ~5 µM pA-Tn5 and 16 pmol unconjugated, Tn5MErev-annealed i5 adapter of the same sequence in minimal volume for 30 minutes-1 hour at RT to generate conjugate-containing i5 transposomes. In parallel, a separate aliquot of 1 µL pA-Tn5 was incubated with 32 pmol i7 adapter for 30 minutes-1 hour at RT to generate an i7 transposome. Conjugate i5 and i7 transposomes were used in MulTI-Tag experiments within 24 hours of assembly. After transposome assembly, 50000 nuclei were thawed and bound to washed ConA beads, then incubated with the first conjugate transposome resuspended in 50 µL Wash-300 Buffer plus 2 mM EDTA for 1 hour at RT or overnight at 4˚C. After incubation, the nuclei mix was washed 3 times with 200 µL Wash-300 Buffer, then tagmented in 50 µL Wash-300 Buffer with 10 mM $MgCl_2$ for 1 hour at 37˚C. After tagmentation, buffer was removed and replaced with 200 µL Wash-300 with 5 mM EDTA and incubated 5 minutes with rotation. The conjugate incubation and tagmentation protocol was then repeated for the remainder of conjugates to be used, up to the point of incubation with the final conjugate. The optimal order of conjugate tagmentation was ascertained empirically by observing the optimal balance of reads between targets, and in this study were tagmented in the following order: PolIIS5P-H3K27me3; or H3K4me2-H3K36me3-H3K27me3. After incubation, the supernatant was cleared and secondary

antibodies corresponding to the species in which the primary antibody conjugates were raised were added in 100 µL Wash Buffer and incubated for 1 hour at RT. The nuclei were then washed twice with 200 µL Wash Buffer and the i7 transposome was added in 100 µL Wash-300 Buffer, and incubated 1 hour at RT. After three washes with 200 µL Wash-300 Buffer, the final tagmentation is carried out by adding 50 µL Wash-300 Buffer with 10 mM $MgCl_2$ and incubating 1 hour at 37˚C. After tagmentation, the nuclei are resuspended in 10 mM TAPS, denatured in TAPS-SDS, neutralized in Triton X-100, amplified and libraries purified as described above. All nuclei transfers were carried out in low-bind 0.6 mL tubes (Axygen). For combined MulTI-Tag, all antibody conjugate incubation and tagmentation steps were carried out simultaneously.

Single cell MulTI-Tag

Single cell MulTI-Tag was carried out as described in Bulk MulTI-Tag protocol up to the completion of the final tagmentation step, with the following modifications: 250 µL paramagnetic Streptavidin T1 Dynabeads (Sigma-Aldrich) were washed 3 times with 1 mL 1x PBS and resuspended in 1 mL 1x PBS with 0.01% Tween-20, 240 µL of Biotin-Wheat Germ Agglutinin (WGA) (Vector Labs) combined with 260 µL 1x PBS was incubated with dynabeads for 30 minutes and resuspended in 1 mL 1x PBS with 0.01% Tween-20 to generate WGA beads, and 100 µL of washed beads were pre-bound with 6 million nuclei. For each experiment, 15 µg H3K4me2 and H3K36me3 conjugate and 7.5 µg H3K27me3 conjugate were used, loaded into transposomes at the ratios described above. All incubations were carried out in 200 µL, and washes in 400 µL. After final conjugate and secondary antibody incubation, nuclei were distributed equally across i7

transposomes containing 96 uniquely barcoded adapters (Supplementary Table 1). After the final tagmentation step, nuclei were reaggregated into a single tube, washed twice in 100 µL 10 mM TAPS, and transferred to a cold block chilled to 0˚C on ice. Supernatant was removed and nuclei were incubated in ice cold DNase reaction mix (10 µL RQ1 DNase, (Promega), 10 µL 10x DNase buffer, 80 µL ddH_2O) for 10 minutes in cold block. The reaction was stopped by adding 100 µL ice cold RQ1 DNase Stop Buffer. Nuclei were immediately washed once in 100 µL 10mM TAPS and then resuspended in 650 µL TAPS. Two 20-micron cell strainers (Fisher Scientific) were affixed to fresh 1.5 mL low bind tubes, and 325 µL nuclei mix was added to the top of each. Tubes were spun 10 minutes at 300 xg to force nuclei through strainer, flowthrough was combined, and resuspended in 640 µL 10 mM TAPS. To the final nuclei mix, 16 µL 100x DAPI and 8 µL ICELL8 Second Diluent (Takara) were added and incubated 10 minutes at RT. The entire nuclei mix was dispensed into an ICELL8 microfluidic chip according to manufacturer's protocols, and SDS denaturation, Triton X-100 neutralization, and amplification were carried out in microwells as described previously[32]. After amplification, microwell contents were reaggregated and libraries were purified with two rounds of cleanup with 1.3x HighPrep beads and resuspended in 20 µL 10 mM Tris-HCl with 1 mM EDTA.

Sequencing and data preprocessing

Libraries were sequenced on an Illumina HiSeq instrument with paired end 25x25 reads. Sequencing data were aligned to the UCSC hg19 genome build using bowtie2[42], version 2.2.5, with parameters --end-to-end--very-sensitive--no-mixed--no-discordant -q–phred33 -I 10 -X 700. Mapped reads were converted to paired-end BED files containing

coordinates for the termini of each read pair, and then converted to bedgraph files using bedtools genomecov with parameter –bg[43]. For single cell experiments, mapped reads were converted to paired-end CellRanger-style bed files, in which the fourth column denotes cell barcode combination, and the fifth column denotes the number of fragment duplicates. Raw read counts and alignment rates for all sequencing datasets presented in this study are listed in Supplementary Table 2.

Data Availability Statement

All primary sequence data and interpreted track files for sequence data generated in this study have been deposited at Gene Expression Omnibus (GEO): GSE179756. Publicly available CUT&Tag and ChIP-seq data analyzed in this study are found at GSE124557, and at the ENCODE Portal at the UCSC Genome Browser (http://hgdownload.soe.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeBroadHistone/), respectively.

Data Analysis

Code necessary for the analyses performed in this study are available on Github (https://github.com/mpmeers/MeersEtAl_MulTI-Tag). Single cell MulTI-Tag pre-processing, feature selection, dimensionality reduction and UMAP projection were carried out as follows: for each target, a unique fragments per cell cutoff (500 for H3K27me3, 200 for H3K4me2, 200 for H3K36me3) was selected based on knee plot analysis, and cells were retained only if they met unique read count criteria for all three targets. For bulk MuLTI-Tag, peaks were called using SEACR v1.4[44] with the following settings: -n norm,

-m stringent, -e 0.1 (https://github.com/FredHutch/SEACR). For single cell MulTI-Tag, peaks were called from aggregate profiles from unique read count-filtered cells using SEACR v1.4 with the following settings: -n norm, -m stringent, -e 5. Peak calls presented in this study are listed in Supplementary Table 3. Cell-specific unique reads were intersected with peaks using Bedtools[43] to generate bed files in which each line contained a unique peak-cell-read count instance. In R (https://www.r-project.org), these bed files were cast into peak (rows) by cell (columns) matrices, which were filtered for the top 40% of peaks by aggregate read counts, scaled by term frequency-inverse document frequency (TF-IDF), and log-transformed. For the gene-centric analysis presented in Figure 3, fragments were mapped to genes in a window extending from 1 kb upstream of the fathest distal annotated TSS to the annotated TES, and matrices were binarized before filtering and TF-IDF processing. Transformed matrices were subjected to Singular Value Decomposition (SVD), and SVD dimensions for which the values in the diagonal matrix ($d as output from the "svd" command in R) were greater than 0.2% of the sum of all diagonal values were used as input to the "umap" command from the umap library in R. Variable features for heatmap plotting were defined as those in the 99th percentile of absolute value in the first or second component of the SVD transformation, or those in the 99th percentile of target-specific features. For genic co-occurrence analysis, the statistical significance of cell-specific, target-specific fragment accumulation in genes was verified by calculating the probability of X fragment-gene overlaps in cell i based on a poisson distribution with a mean μ_i defined by the cell-specific likelihood of a fragment overlap with any base pair in the hg19 reference genome:

$$p = Poisson(X \geq x, \mu_i); \text{ where } x = \frac{r * L_i}{L_{gene}} \text{ and } \mu_i = \frac{L_i * f_i}{L_{genome}}$$

Where L_i = median fragment size in cell i, f_i = number of fragments mapping in cell i, L_{gene} = length of the gene being tested, and L_{genome} = length of the reference genome. All gene-fragment overlaps considered in this study were determined to be statistically significant at a $p < 0.01$ cutoff after Benjamini-Hochberg multiple testing correction. P-values comparing target combination proportions in single cells were calculated using two-sided t-tests. All underlying statistics associated with statistical comparisons presented in this study are listed in Supplementary Table 4. Genome browser screenshots were obtained from Integrative Genomics Viewer (IGV)[45]. CUT&Tag/MulTI-Tag enrichment heatmaps and average plots were generated in DeepTools[46]. UMAPs, violin plots, scatter plots and knee plots were generated using ggplot2 (https://ggplot2.tidyverse.org). Single cell enrichment heatmap displays and hierarchical clustering were generated using the "heatmap" utility and base R graphics. Ternary plots were generated using the ggtern library (http://www.ggtern.com/).

Acknowledgments

We thank Trizia Llagas, Terri Bryson, and Christine Codomo for technical support and Jorja Henikoff and Matthew Fitzgibbon for bioinformatics support for the experiments described in this manuscript. We also thank Kami Ahmad and members of the Henikoff Lab for manuscript critiques, Manu Setty for crucial advice on statistical validation, and Hatice Kaya-Okur for early inspiration and continuing advice throughout the development of this study. This work was supported by the Howard Hughes Medical Institute, an NIH Postdoctoral Fellowship to MPM (F32 GM129954) and an NIH R01 to SH (R01 HG010492).

Author Contributions

MPM conceived the study, carried out the experiments, analyzed the data, and wrote the manuscript. DHJ developed and advised on methods for single cell isolation on the Takara ICELL8 microfluidic platform. SH provided funding, guidance on experiments, and critical and editing support for the manuscript.

Competing Interests Statement

The authors declare no competing interests.

References

1 Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187-1201, doi:10.1016/j.cell.2015.04.044 (2015).

2 Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202-1214, doi:10.1016/j.cell.2015.05.002 (2015).

3 Cao, J. *et al.* Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661-667, doi:10.1126/science.aam8940 (2017).

4 Bandura, D. R. *et al.* Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* **81**, 6813-6822, doi:10.1021/ac901049w (2009).

5 Bendall, S. C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687-696, doi:10.1126/science.1198704 (2011).

6 Shahi, P., Kim, S. C., Haliburton, J. R., Gartner, Z. J. & Abate, A. R. Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci Rep* **7**, 44447, doi:10.1038/srep44447 (2017).

7 Guo, H. *et al.* Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res* **23**, 2126-2135, doi:10.1101/gr.161679.113 (2013).

8 Smallwood, S. A. *et al.* Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods* **11**, 817-820, doi:10.1038/nmeth.3035 (2014).

9 Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486-490, doi:10.1038/nature14590 (2015).

10 Cusanovich, D. A. *et al.* Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910-914, doi:10.1126/science.aab1601 (2015).

11 Cao, J. *et al.* Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380-1385, doi:10.1126/science.aau0730 (2018).

12 Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **14**, 865-868, doi:10.1038/nmeth.4380 (2017).

13 Hu, Y. *et al.* Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol* **17**, 88, doi:10.1186/s13059-016-0950-z (2016).

14 Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* **9**, 781, doi:10.1038/s41467-018-03149-4 (2018).

15 Swanson, E. *et al.* Simultaneous trimodal single-cell measurement of transcripts, epitopes, and chromatin accessibility using TEA-seq. *Elife* **10**, doi:10.7554/eLife.63632 (2021).

16 Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, 1930, doi:10.1038/s41467-019-09982-5 (2019).

17 Kaya-Okur, H. S., Janssens, D. H., Henikoff, J. G., Ahmad, K. & Henikoff, S. Efficient low-cost chromatin profiling with CUT&Tag. *Nat Protoc* **15**, 3264-3283, doi:10.1038/s41596-020-0373-x (2020).

18 Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* **28**, 817-825, doi:10.1038/nbt.1662 (2010).

19 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).

20 Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**, 613-626, doi:10.1038/nrg3207 (2012).

21 Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-1043, doi:10.1126/science.1076997 (2002).

22 Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-353, doi:10.1038/nature04733 (2006).

23 Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837, doi:10.1016/j.cell.2007.05.009 (2007).

24 Bannister, A. J. *et al.* Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem* **280**, 17732-17736, doi:10.1074/jbc.M500796200 (2005).

25 Wagner, E. J. & Carpenter, P. B. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol* **13**, 115-126, doi:10.1038/nrm3274 (2012).

26 Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326, doi:10.1016/j.cell.2006.02.041 (2006).

27 Bartosovic, M., Kabbe, M. & Castelo-Branco, G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat Biotechnol*, doi:10.1038/s41587-021-00869-9 (2021).

28 Wu, S. J. *et al.* Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression. *Nat Biotechnol*, doi:10.1038/s41587-021-00865-z (2021).

29 Zhu, C. *et al.* Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat Methods* **18**, 283-292, doi:10.1038/s41592-021-01060-3 (2021).

30 McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 (2018). <https://ui.adsabs.harvard.edu/abs/2018arXiv180203426M>.

31 Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*, doi:10.1038/nbt.4314 (2018).

32 Janssens, D. H. *et al.* Automated CUT&Tag profiling of chromatin heterogeneity in mixed-lineage leukemia. *bioRxiv*, 2020.2010.2006.328948, doi:10.1101/2020.10.06.328948 (2021).

33 Leroy, G. *et al.* A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics Chromatin* **6**, 20, doi:10.1186/1756-8935-6-20 (2013).

34 Shema, E. *et al.* Single-molecule decoding of combinatorially modified nucleosomes. *Science* **352**, 717-721, doi:10.1126/science.aad7701 (2016).

35 Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**, 105-116, doi:10.1016/j.devcel.2005.10.017 (2006).

36 Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479-491, doi:10.1016/j.stem.2010.03.018 (2010).

37 Yuan, W. *et al.* H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem* **286**, 7983-7989, doi:10.1074/jbc.M110.194027 (2011).

38 Lu, C. *et al.* Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* **352**, 844-849, doi:10.1126/science.aac7272 (2016).

39 Handa, T. *et al.* Chromatin integration labeling for mapping DNA-binding proteins and modifications with low input. *Nat Protoc* **15**, 3334-3360, doi:10.1038/s41596-020-0375-8 (2020).

40 Gopalan, S., Wang, Y., Harper, N. W., Garber, M. & Fazzio, T. G. Simultaneous profiling of multiple chromatin proteins in the same cells. *bioRxiv*, 2021.2004.2027.441642, doi:10.1101/2021.04.27.441642 (2021).

41 Harada, A. *et al.* A chromatin integration labelling method enables epigenomic profiling with lower input. *Nat Cell Biol* **21**, 287-296, doi:10.1038/s41556-018-0248-3 (2019).

42 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

43 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

44 Meers, M. P., Tenenbaum, D. & Henikoff, S. Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin* **12**, 42, doi:10.1186/s13072-019-0287-4 (2019).

45 Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26, doi:10.1038/nbt.1754 (2011).

46 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).

Figure Legends

Figure 1: MulTI-Tag directly identifies user-defined chromatin targets in the same cells. a) Schematic describing the MulTI-Tag methodology: 1) Antibody-oligonucleotide conjugates are used to physically associate forward-adapter barcodes with targets, and are loaded directly into pA-Tn5 transposomes for sequential binding and tagmentation; 2) pA-Tn5 loaded exclusively with reverse adapters are used for a secondary CUT&Tag step to efficiently introduce the reverse adapter to conjugate-bound loci; 3) Target-specific profiles are distinguished by barcode identity in sequencing. b) Genome browser screenshot showing individual CUT&Tag profiles for H3K27me3 (first row) and RNA PolIIS5P (second) in comparison with MulTI-Tag profiles for the same targets probed individually in different cells (third and fourth rows) or sequentially in the same cells (fifth and sixth). c) Heatmaps describing the enrichment of H3K27me3 (red) or RNA PolIIS5P (blue) signal from sequential MulTI-Tag profiles at CUT&Tag-defined H3K27me3 peaks (left) or RNA PolIIS5P peaks (right). d) Genome browser screenshot showing H3K27me3 (red), H3K4me2 (purple), and H3K36me3 (teal) MulTI-Tag signal from experiments in H1 hESCs using an individual antibody (rows 1, 3, and 5) or all three antibodies in sequence (rows 2, 4, and 6). e) Top: Schematic showing characteristic enrichment of H3K27me3 (red), H3K4me2 (purple), and H3K36me3 (teal) across genes. Bottom: Normalized CUT&Tag (light colors) and MulTI-Tag (dark colors) enrichment of H3K27me3, H3K4me2, and H3K36me3 across genes in H1 hESCs.

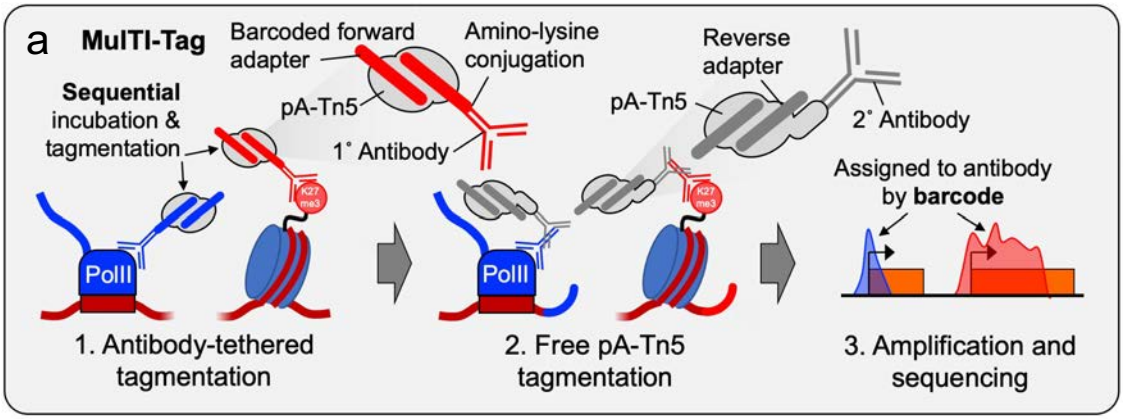
Figure 2: MulTI-Tag in single cells. a) Schematic describing single cell MulTI-Tag experiments. H1 hESCs (red) and K562 cells (green) were profiled separately in bulk, then individual cells were dispensed into nanowells on a Takara ICELL8 microfluidic

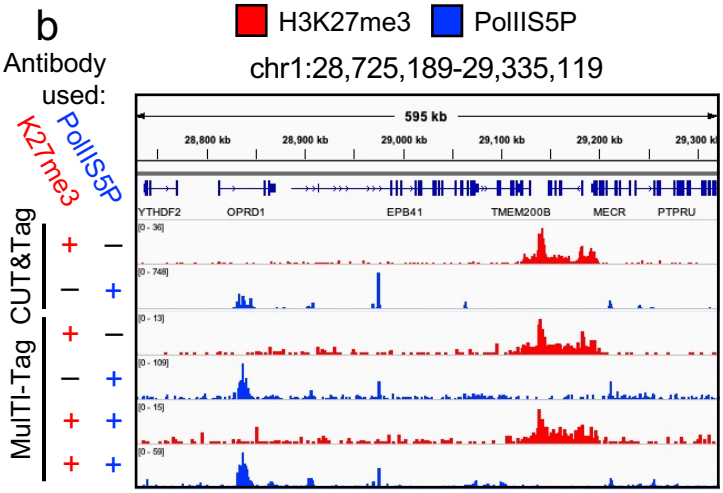
device for single cell barcoding via amplification. b) Genome browser screenshot showing single cell MulTI-Tag data from K562 (top) and H1 (bottom) cells. Target-specific enrichment is shown in aggregate in first three columns, and cell-specific fragments for the top 50 cells by unique fragments per cell are shown below. c) Ternary plot describing the percentage of normalized fragments originating from H3K27me3 (left axis), H3K4me2 (right axis), or H3K36me3 (bottom axis) in each peak (open circles) called from aggregate K562 cell MulTI-Tag. Peak identity is denoted by circle color, and total normalized fragments is denoted by circle size. d) UMAP plots for single cell MulTI-Tag data in H1 and K562 cells. Projections based on H3K27me3 (left), H3K4me2 (center left), H3K36me3 (center right), or all features combined (right) are shown.

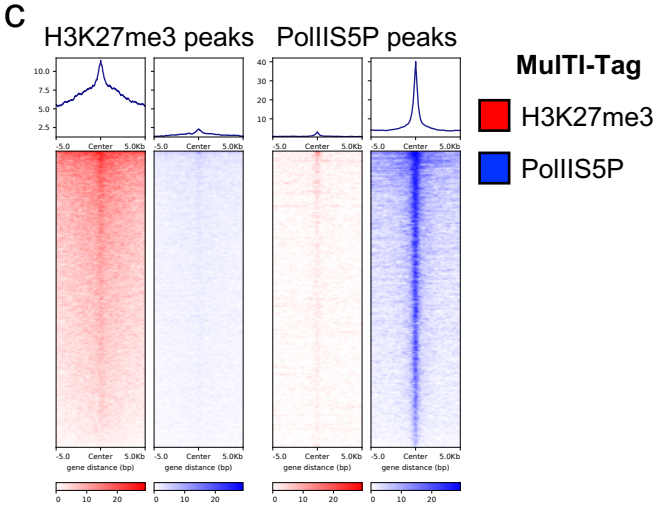
Figure 3: Coordinate multifactorial analysis in the same cells using MulTI-Tag. a) Violin plots describing the distribution of the proportions of MulTI-Tag H3K27me3 (red), H3K4me2 (purple), or H3K36me3 (teal) unique reads out of total unique reads in individual cells, with points representing each single cell value. Lines connect points that represent the same individual cell. b) Schematic describing coordinated multifactorial analysis strategy for MulTI-Tag. Genes in individual cells are analyzed for the enrichment of all MulTI-tag targets, and gene-cell target combinations are mapped onto a matrix for clustering and futher analysis. c) Heatmap describing co-occurrence of MulTI-tag targets in 363 highly variable genes in each of 716 single cells. The heatmap is hierarchically clustered on the column side by normalized all-factor enrichment, and on the row side by displayed co-occurrence values. Heatmap color code describes co-occurrence representations; color coding above columns indicates cell identity. d) Heatmap column-

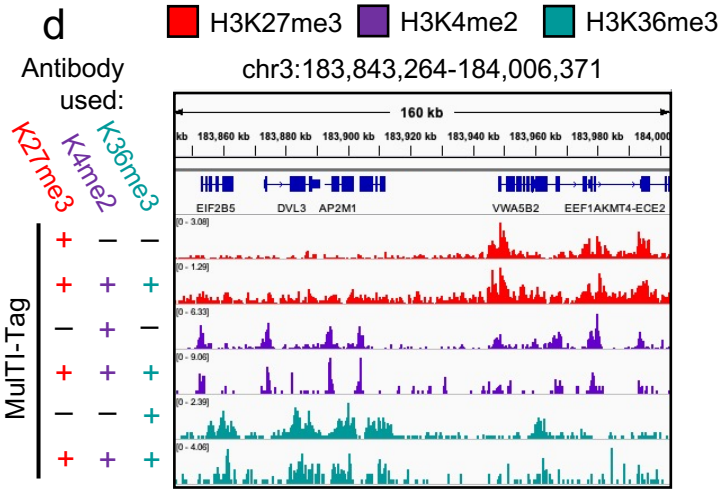
clustered as in c) describing co-occurrence in 6 genes of interest. e) Violin plots describing the distributions of proportions of each co-occurrence state as described below the plot in individual H1 (red) or K562 (green) cells, with points denoting individual cell values. The last six co-occurrence states are rescaled and inset at top right; p-values derived from two-sided student's t-test comparing distributions between cell types are listed above violins (not corrected for multiple hypothesis testing). f) Violin plots describing calculated Cramer's V of Association between target combinations listed at bottom in individual H1 (red) or K562 (green) cells, with points denoting individual cell values.

Figure 1


a
MulTI-Tag
Sequential incubation & tagmentation
Barcoded forward adapter
Amino-lysine conjugation
pA-Tn5
1˚ Antibody
Reverse adapter
pA-Tn5
2˚ Antibody
Assigned to antibody by barcode
PolII
PolII
1. Antibody-tethered tagmentation
2. Free pA-Tn5 tagmentation
3. Amplification and sequencing


b
H3K27me3
PolIIS5P
chr1:28,725,189-29,335,119
Antibody used:
K27me3
PolIIS5P
595 kb
YTHDF2
OPRD1
EPB41
TMEM200B
MECR
PTPRU
CUT&Tag
MulTI-Tag


c
H3K27me3 peaks
PolIIS5P peaks
MulTI-Tag
H3K27me3
PolIIS5P


d
H3K27me3
H3K4me2
H3K36me3
chr3:183,843,264-184,006,371
Antibody used:
K27me3
K4me2
K36me3
160 kb
EIF2B5
DVL3
AP2M1
VWA5B2
EEF1AKMT4-ECE2
MulTI-Tag

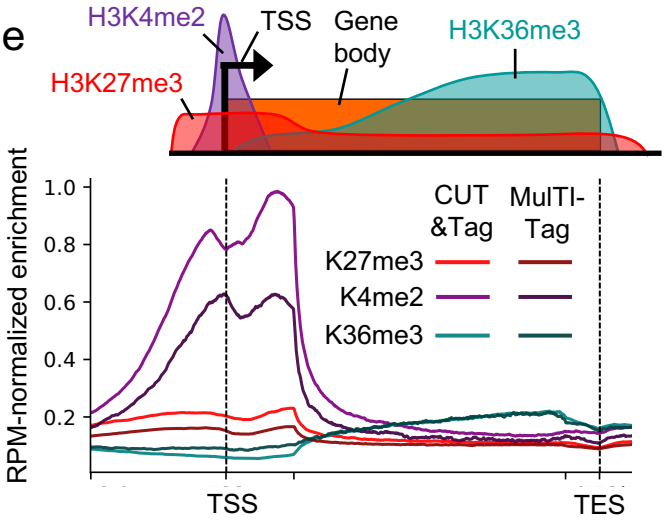
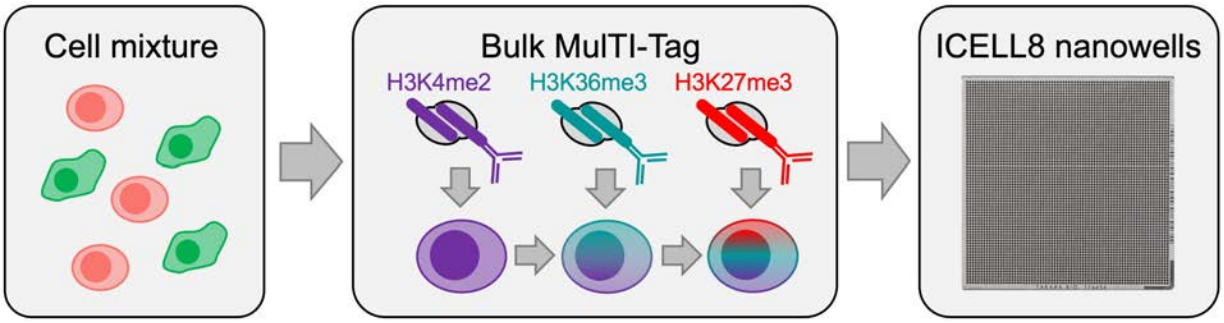
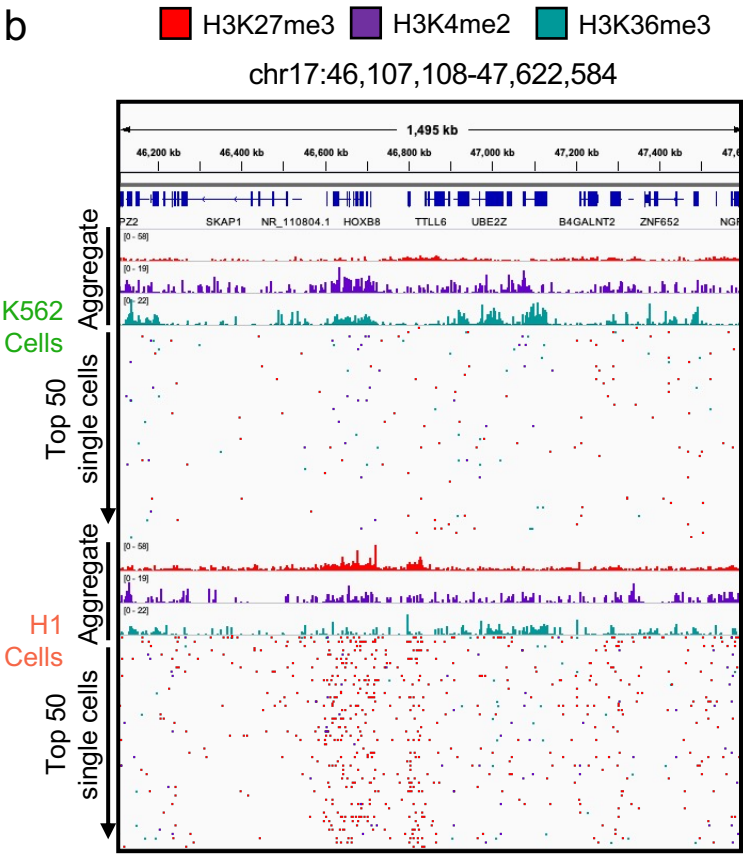
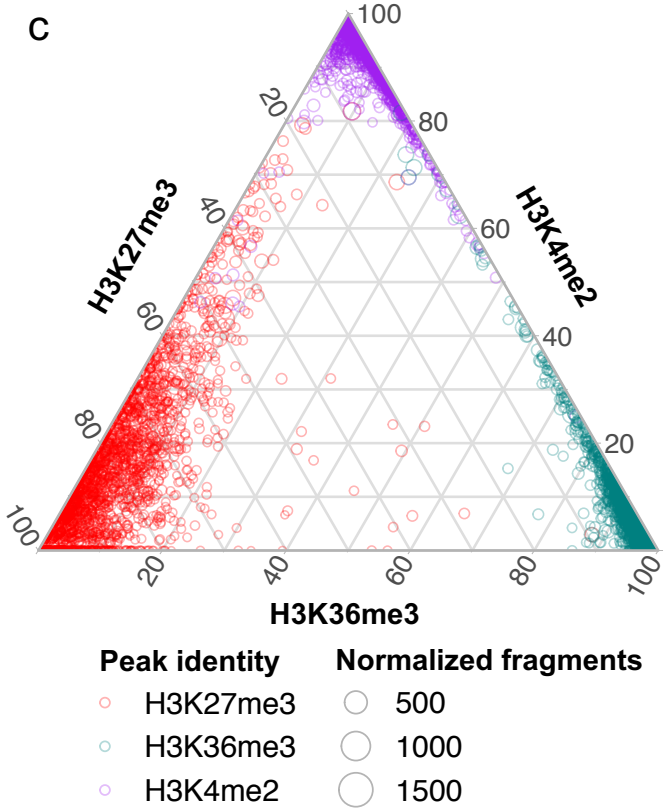

e
H3K4me2
TSS
Gene body
H3K36me3
H3K27me3
RPM-normalized enrichment
1.0
0.8
0.6
0.4
0.2
CUT &Tag
MulTI-Tag
K27me3
K4me2
K36me3
TSS
TES

Figure 2


a
Cell mixture
Bulk MulTI-Tag
H3K4me2
H3K36me3
H3K27me3
ICELL8 nanowells


b
H3K27me3
H3K4me2
H3K36me3
chr17:46,107,108-47,622,584
1,495 kb
K562 Cells
Aggregate
Top 50 single cells
H1 Cells
Aggregate
Top 50 single cells


c
H3K27me3
H3K4me2
H3K36me3
Peak identity
H3K27me3
H3K36me3
H3K4me2
Normalized fragments
500
1000
1500

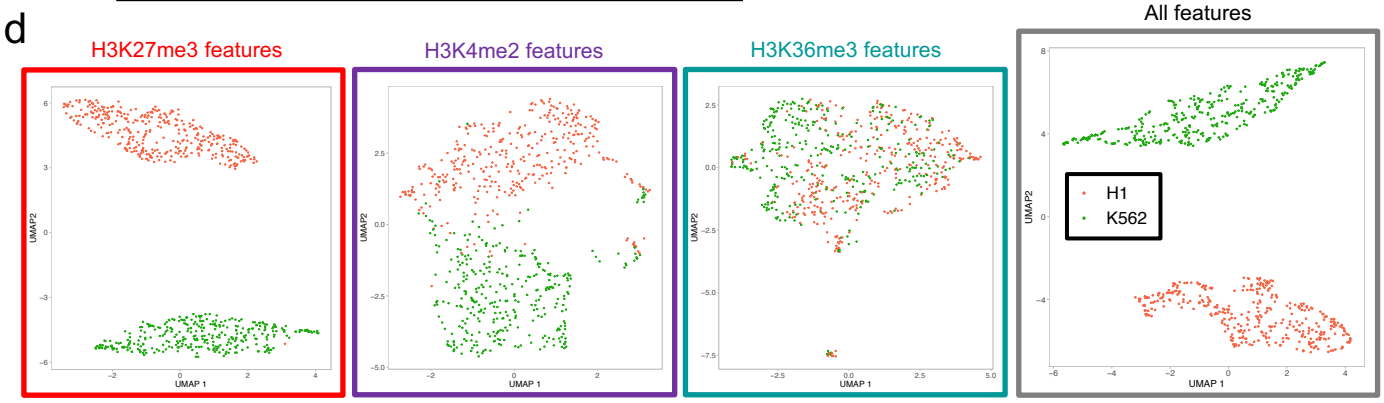
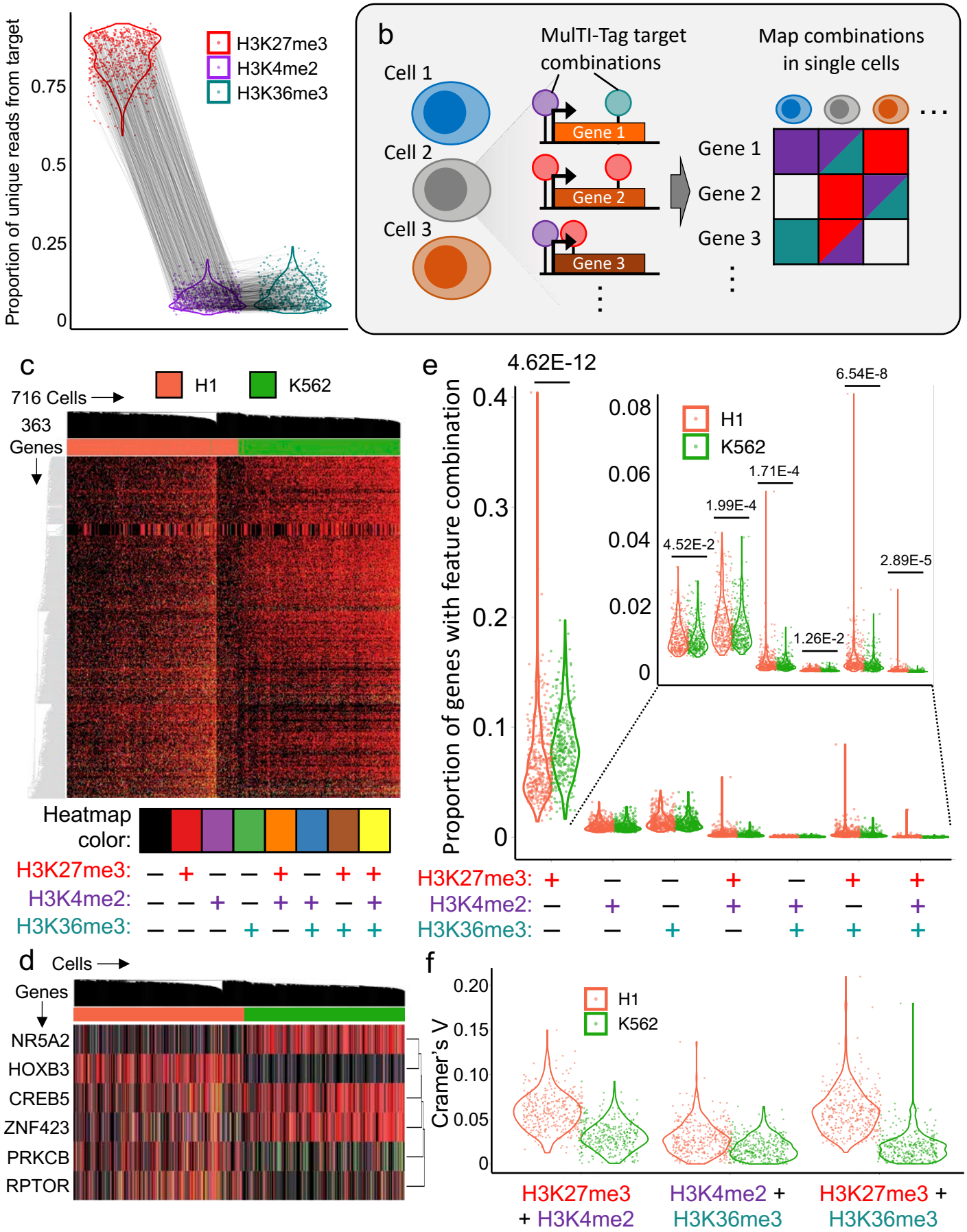
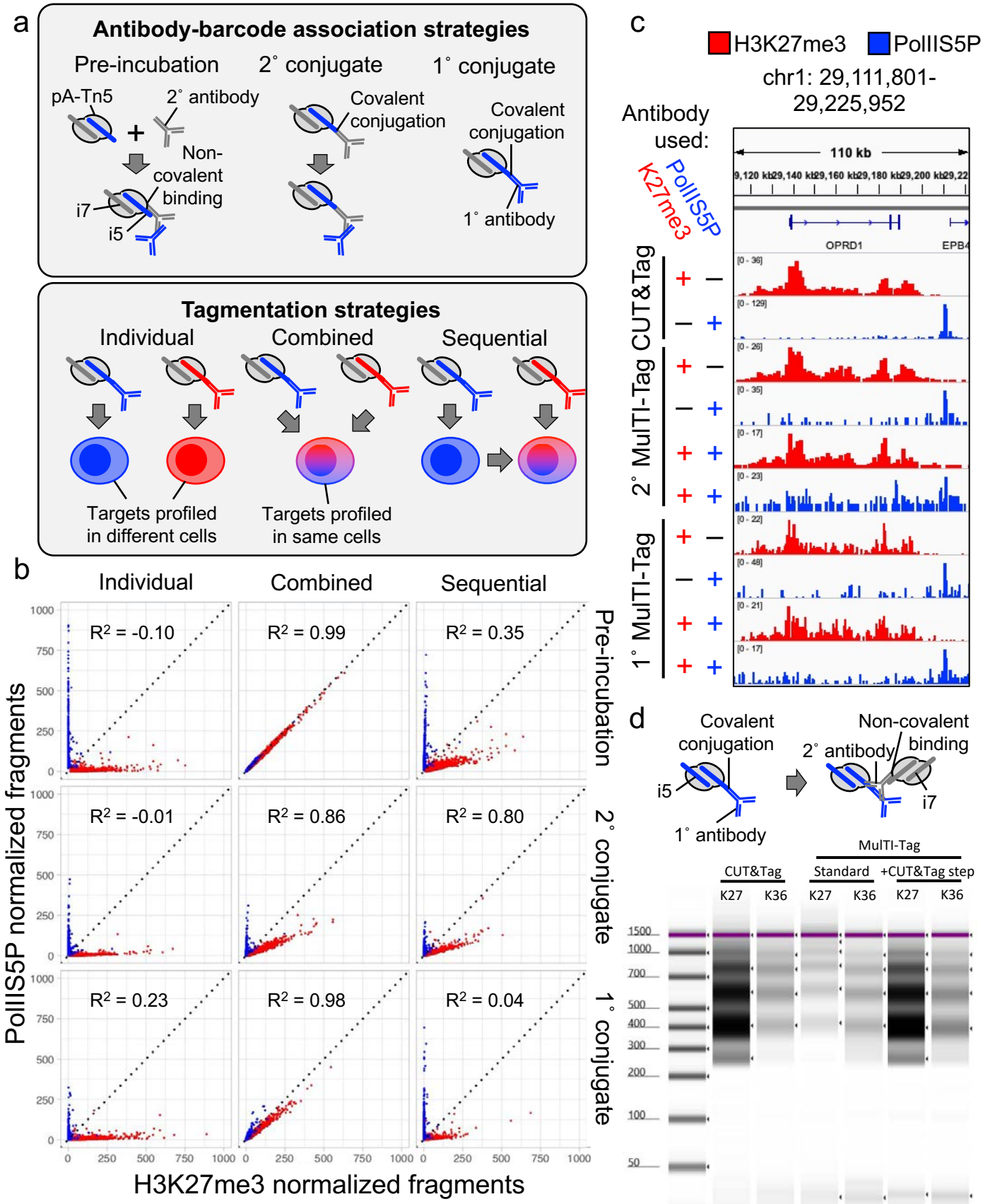

d
H3K27me3 features
H3K4me2 features
H3K36me3 features
All features
H1
K562
UMAP 1
UMAP2

Figure 3


a
Proportion of unique reads from target
0.75
0.5
0.25
0
H3K27me3
H3K4me2
H3K36me3
b
Cell 1
Cell 2
Cell 3
MulTI-Tag target combinations
Gene 1
Gene 2
Gene 3
Map combinations in single cells
Gene 1
Gene 2
Gene 3
c
716 Cells
363 Genes
H1
K562
Heatmap color:
H3K27me3: – + – – + – + +
H3K4me2: – – + – + + – +
H3K36me3: – – – + – + + +
e
4.62E-12
6.54E-8
1.71E-4
1.99E-4
4.52E-2
2.89E-5
1.26E-2
Proportion of genes with feature combination
0.4
0.3
0.2
0.1
0
0.08
0.06
0.04
0.02
H1
K562
H3K27me3: + – – + – + +
H3K4me2: – + – + + – +
H3K36me3: – – + – + + +
d
Cells
Genes
NR5A2
HOXB3
CREB5
ZNF423
PRKCB
RPTOR
f
Cramer's V
0.20
0.15
0.10
0.05
0
H1
K562
H3K27me3 + H3K4me2
H3K4me2 + H3K36me3
H3K27me3 + H3K36me3

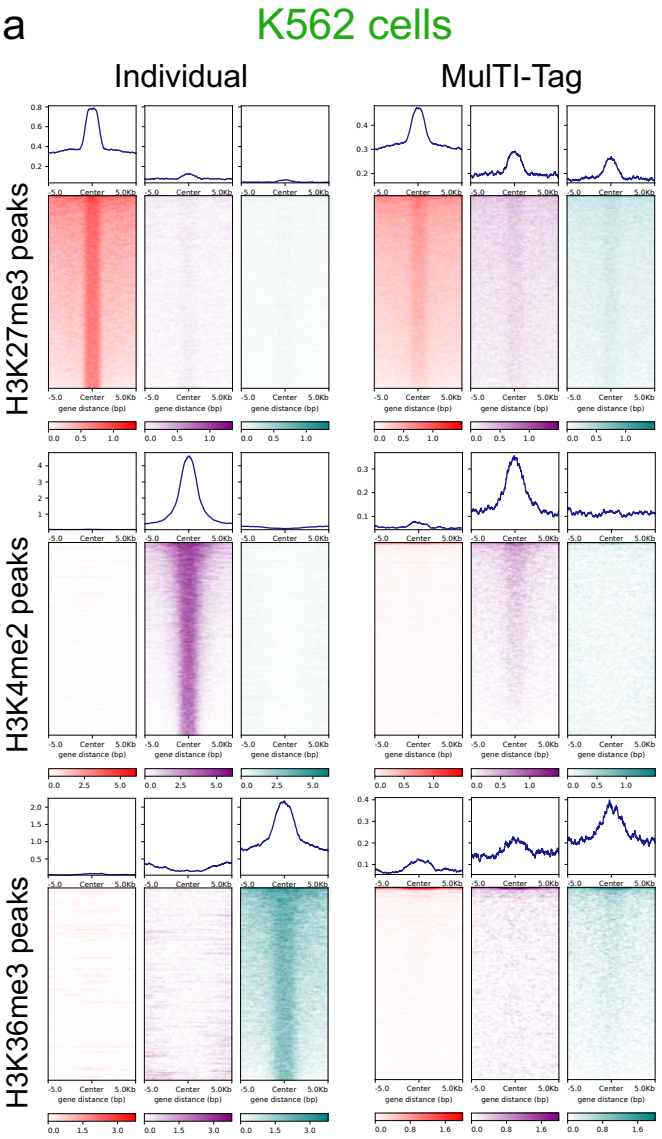
Supplementary Figure 1



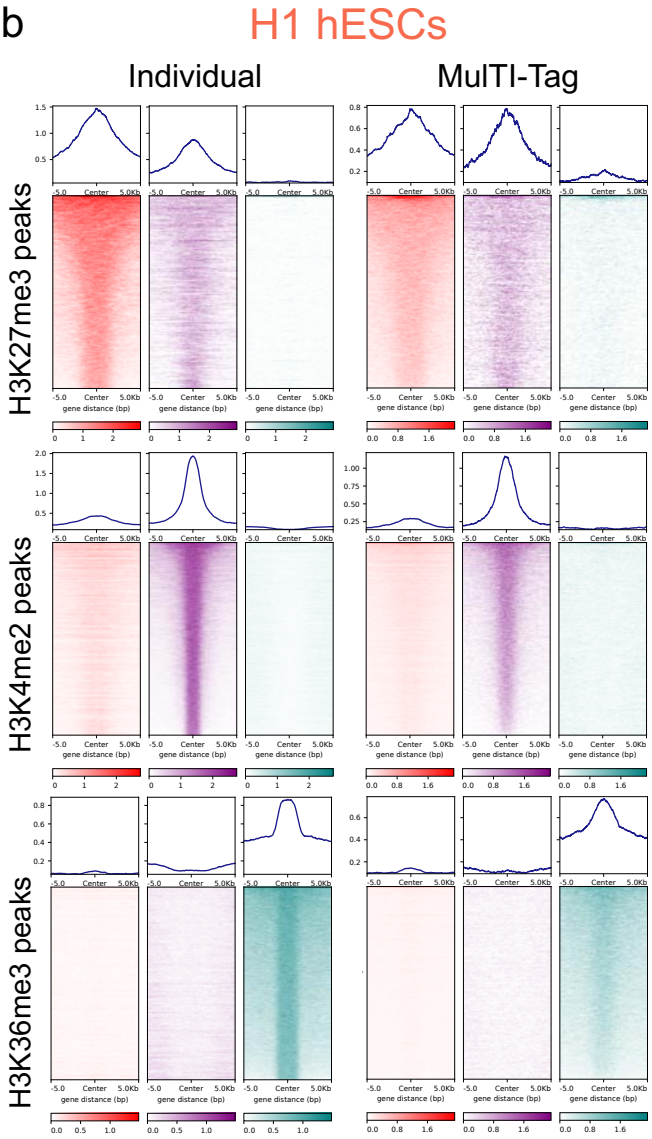
Supplementary Figure 1: a) Schematic of protocol variations tested for distinguishing CUT&Tag targets by sequencing barcode. Top: Approaches for pairing barcodes with antibodies, either by pre-incubation of barcoded pA-Tn5 with a secondary antibody ("Pre-incubation", left), or covalent conjugation of barcode-containing adapters to secondary ("2˚ conjugate", center) or primary ("1˚ conjugate", right) antibodies. Bottom: Approaches for tagmenting multiple targets, either in separate cells ("Individual", left), in the same cells simultaneously ("Combined", center), or in the same cells sequentially ("Sequential", right). b) Scatterplots describing the enrichment of H3K27me3 (X-axis) and PolIIS5P (Y-axis) in H3K27me3 (red points) or PolIIS5P (blue points) peaks for combinations of experimental conditions described in 2a. Pearson's R^2 of all data points is denoted for each of the nine protocol conditions. c) Genome browser screenshot showing individual CUT&Tag profiles for H3K27me3 (first row) and RNA PolIIS5P (second) in comparison with MulTI-Tag profiles for the same targets probed individually in different cells (third and fourth rows secondary conjugate MulTI-Tag; seventh and eighth rows primary conjugate MulTI-Tag) or sequentially in the same cells (fifth and sixth rows secondary conjugate MulTI-Tag; ninth and tenth rows primary conjugate MulTI-Tag). d) Top: Schematic of MulTI-Tag with additional CUT&Tag step, in which 1˚ antibody conjugates are loaded into pA-Tn5 along with free i5 adapter (left), and secondary antibody and pA-Tn5 loaded only with i7 adapter are added before tagmentation (right). Bottom: TapeStation HSD1000 trace describing DNA size and enrichment from libraries produced from CUT&Tag (lanes 1 and 2), "standard" MulTI-Tag with conjugate-only tagmentation (3 and 4), or MulTI-Tag with a secondary CUT&Tag step as described in methods (5 and 6), targeting H3K27me3 (1, 3, and 5) or H3K36me3 (2, 4, and 6) in K562 cells.

Supplementary Figure 2

a



b

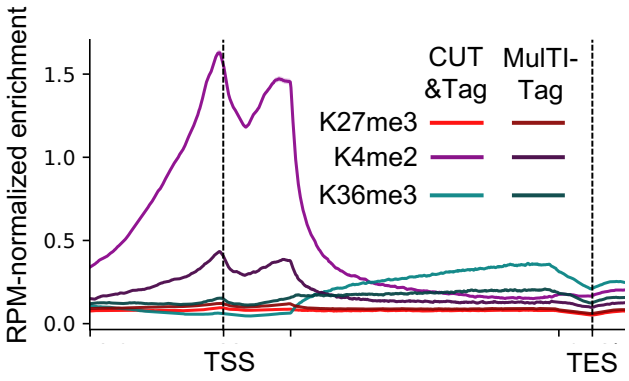


c

$$\text{Enrichment efficiency} = \frac{\text{MulTI-Tag reads in target peaks/reads in all peaks}}{\text{CUT\&Tag reads in target peaks/reads in all peaks}}$$

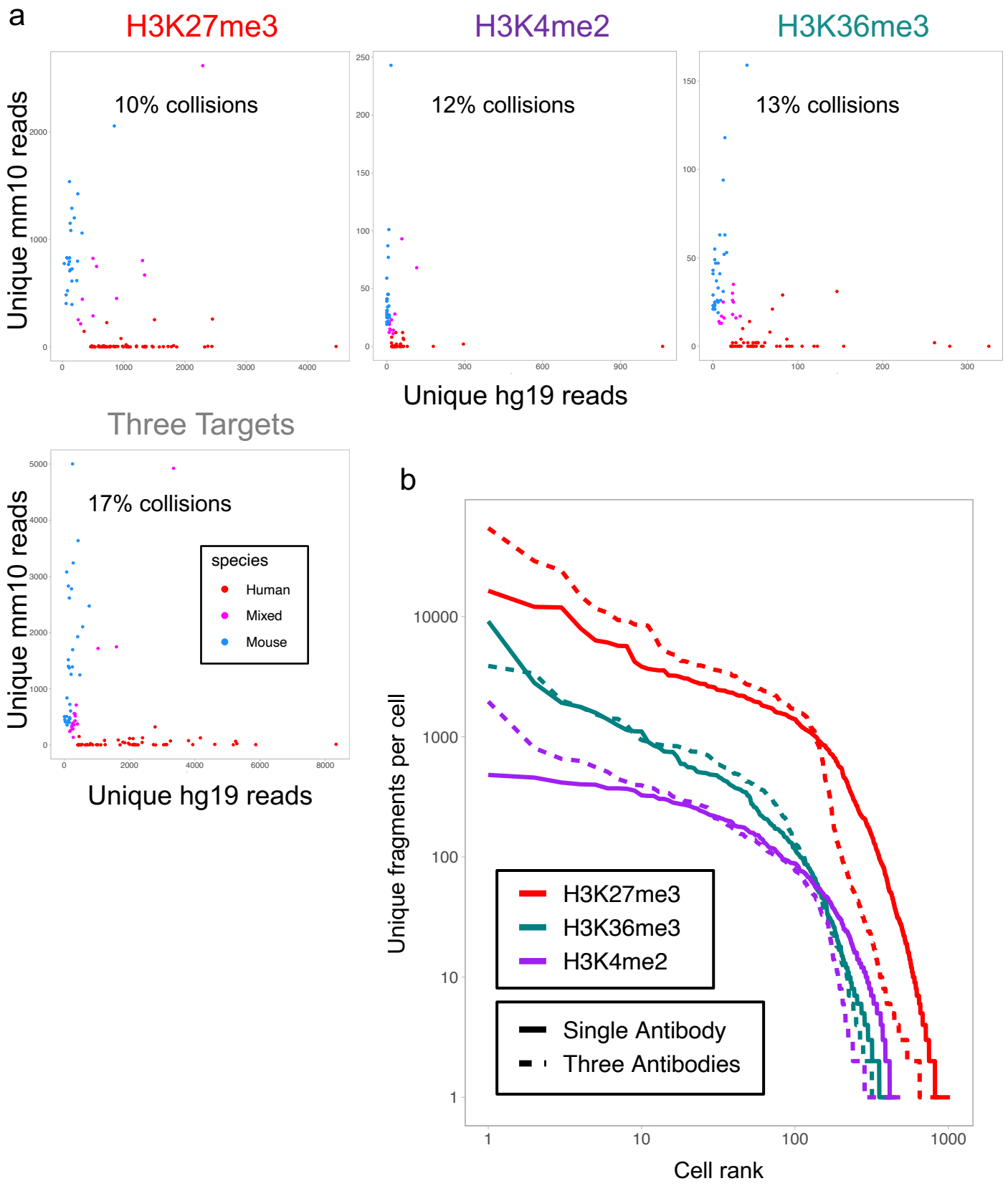
	H3K27me3	H3K4me2	H3K36me3
K562	0.5325151	0.5870659	0.8923193
H1	0.7886033	0.6160182	1.107574

d



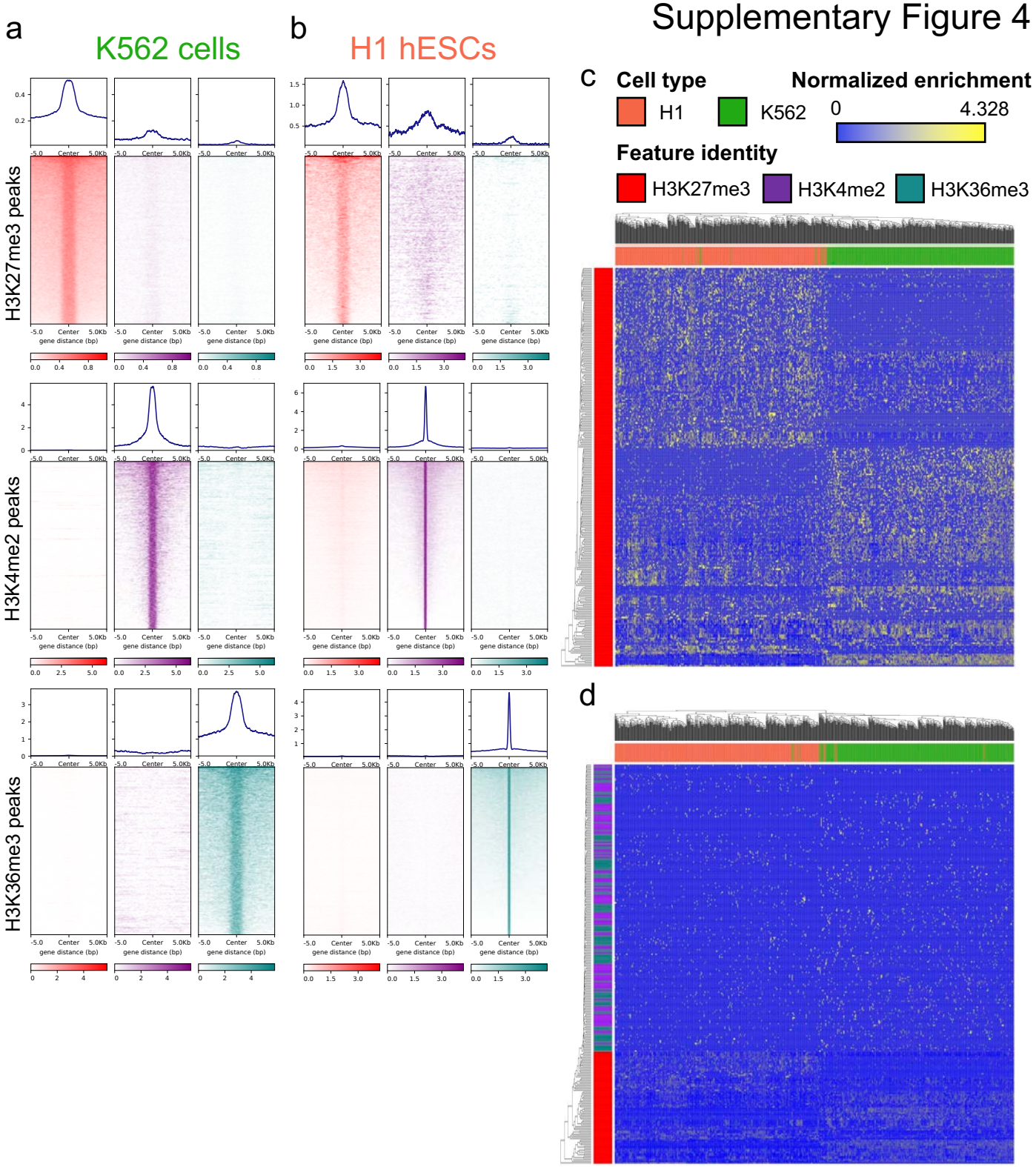
Supplementary Figure 2: a) Heatmaps describing the enrichment of H3K27me3 (red), H3K4me2 (purple), or H3K36me3 (teal) signal from K562 cell MulTI-Tag profiles using single antibodies (left) or three antibodies sequentially (right) in H3K27me3 (top), H3K4me2 (middle), or H3K36me3 (bottom) peaks. b) Heatmaps as in a) from H1 hESC MulTI-Tag experiments. c) Enrichment efficiency of MulTI-Tag vs. CUT&Tag in peaks from cell types as denoted. d) Normalized CUT&Tag (light colors) and MulTI-Tag (dark colors) enrichment of H3K27me3, H3K4me2, and H3K36me3 across genes in K562 cells.

Supplementary Figure 3



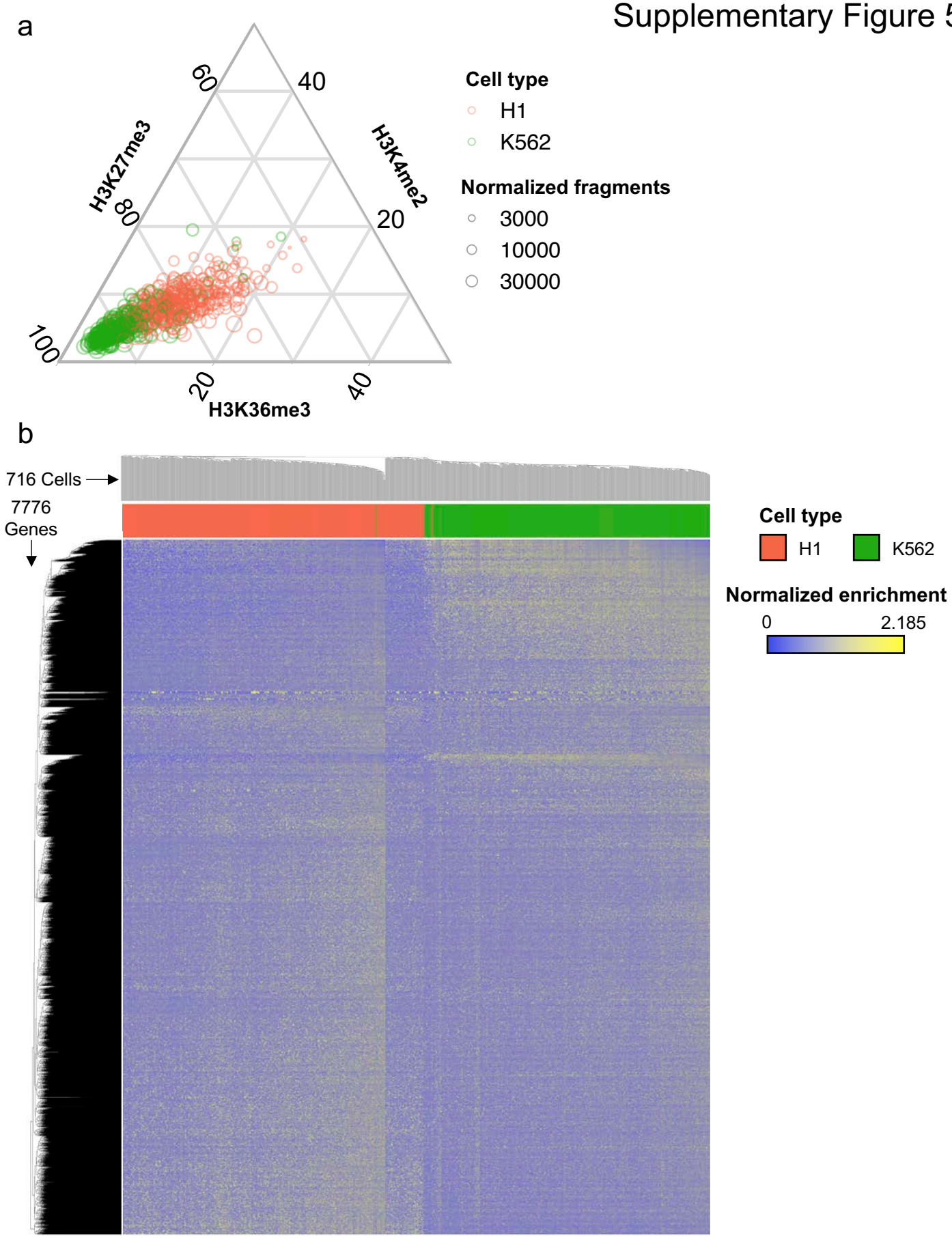
Supplementary Figure 3: a) Barnyard plots describing the number of unique fragments exclusively mapping to the hg19 genome build (X-axis) vs. mm10 (Y-axis) in the top 100 cells for each of the denoted experiments. Points are colored by the cell identity as human (red; > 70% of unique reads mapping to hg19), mouse (blue; >70% mapping to mm10), or mixed (magenta; < 70% mapping to either), and collision rate, defined as the percentage of cells classified as “mixed”, is denoted for each experiment. b) Knee plot describing unique fragments per cell (Y-axis) for cells ranked by the number of unique fragments per cell (X-axis) from MulTI-Tag experiments profiling H3K27me3 (red), H3K4me2 (purple), or H3K36me3 (teal) individually (solid line) or together (dashed line) in MulTI-Tag experiments.

Supplementary Figure 4



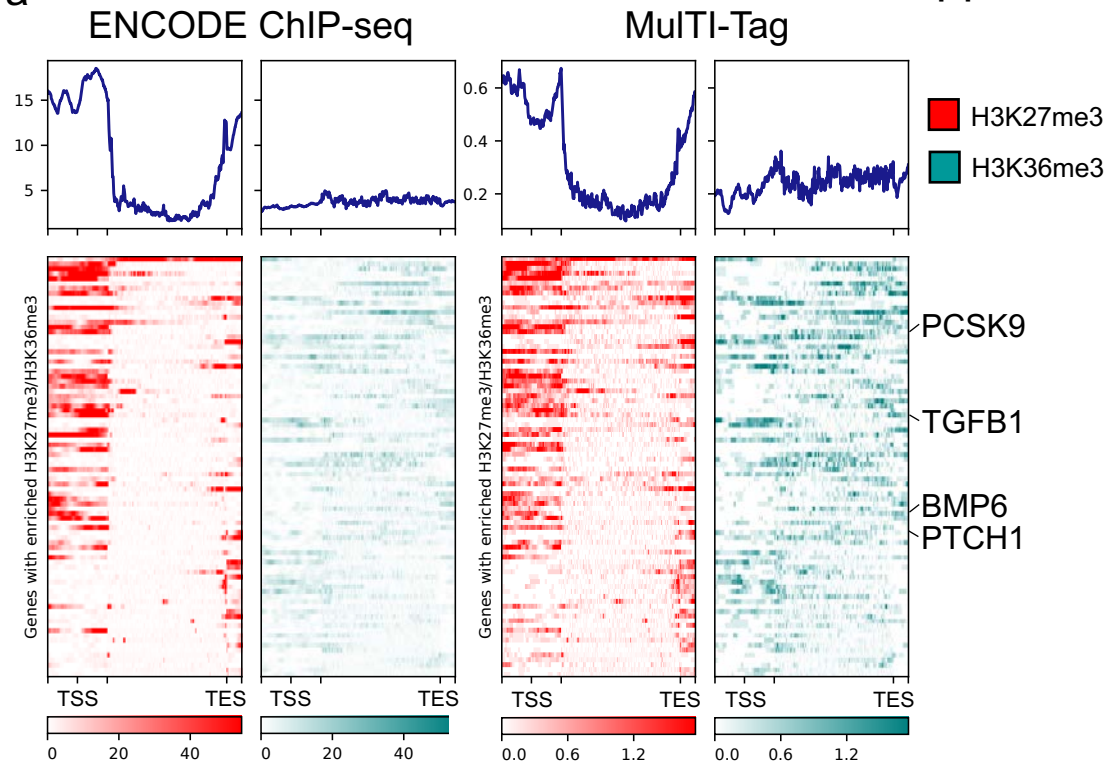
Supplementary Figure 4: a) Heatmaps describing the enrichment of H3K27me3 (red), H3K4me2 (purple), or H3K36me3 (teal) signal from aggregate single cell MulTI-Tag profiles from K562 cellsin H3K27me3 (top), H3K4me2 (middle), or H3K36me3 (bottom) peaks. b) Heatmaps as in a) for H1 hESC MulTI-Tag experiments. c) Hierarchically clustered heatmap describing MulTI-Tag enrichment from single cells (columns) in the 248 most variable defined features (rows) in the experiment. Color coding above columns indicates cell type, and color coding to the left of rows indicates feature identity. d) Hierarchically clustered heatmap describing MulTI-Tag enrichment from single cells (columns) in 164 features (rows) defined as most variable relative to other features of the same target identity in the experiment. Color coding above columns indicates cell type, and color coding to the left of rows indicates feature identity.

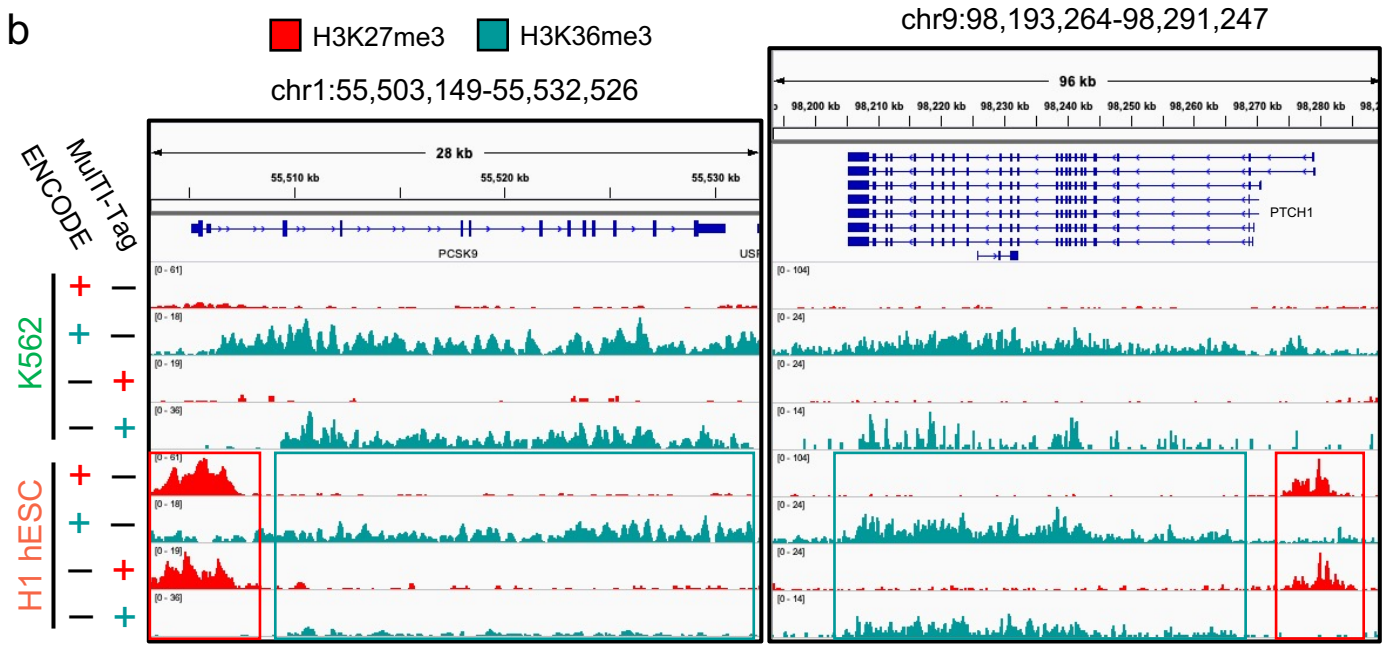
Supplementary Figure 5



Supplementary Figure 5: a) Ternary plot describing the percentage of fragments originating from H3K27me3 (left axis), H3K4me2 (right axis), or H3K36me3 (bottom axis) in each cell (open circles) in single cell MulTI-Tag. Cell identity is denoted by circle color, and total fragment count is denoted by circle size. b) Hierarchically clustered heatmap describing MulTI-Tag enrichment for all three targets from single cells (columns) in 7776 genes (rows) that are in the top 40% of genes by total fragments mapped. Color coding above columns indicates cell type.

Supplementary Figure 6





Supplementary Figure 6: a) Heatmaps describing the enrichment of H3K27me3 (red) and H3K36me3 (teal) signal from ENCODE ChIP-seq (left) or bulk MulTI-Tag (right) in H1 hESCs in 86 genes for which 1) a MulTI-Tag H3K27me3 peak overlapped a 2 kb window surrounding the TSS, and 2) a MulTI-Tag H3K36me3 peak overlapped the gene body. Selected genes of interest, including those involved in metabolic and developmental signaling, are highlighted at right. b) Genome browser screenshots showing H3K27me3 (red) and H3K36me3 (teal) enrichment from ENCODE ChIP-seq (rows 1, 2, 5, and 6) or bulk MulTI-Tag (rows 3, 4, 7, and 8) in K562 cells (rows 1-4) or H1 hESCs (rows 5-8) at the PCSK9 (left) and PTCH1 (right) genes. Colored boxes indicate co-enrichment of H3K27me3 and H3K36me3 in the same gene in H1 hESCs. Gene model for PTCH1 is expanded to visualize alternative promoter structure.