Supplementary Note 1
Quality Assessment of sci-MET human cell line discrimination libraries

Here we include further detail about the quality assessment of the 691 single cells passing quality control filters generated in the human cell line discrimination experiment (Supplementary Figs. 1-3). The genomic coverage of our libraries varied due to raw sequencing depths and individual cell performance, potentially in part due to transposase complex efficiency (Supplementary Fig. 14). On average we achieved a mean unique aligned read count of 403,265 per cell (for cells with ≥ 30,000 unique aligned reads), with 48 cells producing over one million uniquely aligned reads (Supplementary Fig. 2). These data translated to coverages of mappable[28] CG dinucleotides ranging from 0.05% to 7.0% (mean 1.1 ± 0.9%; Supplementary Data 2). We also observed the expected increase in unique CG coverage by unique read alignment (Supplementary Fig. 16), with cells in pools of higher sequencing depth covering a mean of 3.4% of CGs. We note that despite the high alignment rate (reported in Main Text), the percentage of unique, non-PCR duplicate, reads is lower in these sci-MET generated libraries than other existing protocols[7]. Upon projection of libraries generated with four rounds of linear amplification to complete saturation, we cover 2.51 x 10^6 unique CG sites per cell on average (9.30% coverage of CG dinucleotides, Supplementary Figure 5). sci-MET's useable read rates (uniquely aligned reads/all reads assigned to a barcode) were comparable to that of existing scWGBS at 18.5%, with averages ranging from 14.2 to 26.2%[3-7].

On average, our libraries exhibited CG and CH methylation rates of 51 ± 4.0% and 0.77 ± 0.23% respectively (Fig. 1d), consistent with expectations and a high efficiency of bisulfite conversion (99.14% from lambda phage unmethylated control spike-in, Supplementary Table 5). We also observed comparable coverage over annotated regions of the genome, with only a slight bias toward regions of open chromatin over regions with repressive marks (Supplementary Fig. 17), and increased unique CG nucleotide coverage when aggregating multiple cells – with 66 low-coverage cells required to cover approximately 50% of the CG dinucleotides in the haploid genome at our sequencing depth (Supplementary Fig. 18). The methylation rate over functionally annotated regions conformed to expectations in activated and repressed regions (Fig. 1f), and in the promoter region of genes (Fig. 1g), and greater methylation rate change for cells matching the ChIP-seq peak set (Fig. 1h, Supplementary Fig. 20). We did not observe a correlation between alignment rates and methylation level or coverage (Supplementary Fig. 15).

To separate out cells based on their cell line of origin, summarized methylation status[3] was first calculated for each cell across autosomal loci of the Ensembl Regulatory Build[16], which contains known transcription factor binding and other regulatory sites. We then performed Non-negative Matrix Factorization followed by t-distributed Stochastic Neighbor Embedding (NMF-tSNE) to project cells in two-dimensional space (Fig. 2a). Distinct domains containing each respective cell type along with cells from the masked population were observed. Cluster purity was further confirmed by the proportion of reads aligning to the Y-chromosome (specific to the male fibroblast line), as well as minimal bias pertaining to unique read count or global CG methylation percentage (Supplementary Fig. 19). When aggregating cells by cluster identity, we observed ample genome-wide coverage, with the HEK293 cluster (10 cells) covering 3.17×10^6 unique CG sites (11.7%; genome-wide average of 0.15×), the primary fibroblast cluster (32 cells) covering 5.01×10^6 unique CG sites (18.6%; average 0.24×), and the GM12878 cluster covering 26.1×10^6 unique CG sites (97%; average 7.39×; Supplementary Data 2). We next correlated the methylation rates with publically available WGBS datasets[17,18] for the top 1,000 most variable regulatory regions. For each merged cluster, the two most highly correlated

samples were of the same cell type, or the most similar cell line in the case of HEK293 (Fig. 2b, Supplementary Fig. 6,21). Hierarchical clustering on the Pearson correlation coefficients placed the HEK293 cluster in a clade with other aberrant cell lines (HepG2 and K562), the GM12878 cluster in a clade with two GM12878 bulk WGBS samples, and the Primary Fibro. cluster in a clade with two primary forearm fibroblast bulk WGBS samples (Fig. 2c, Supplementary Fig. 21).

In addition to stratifying cells by aggregating over the Ensembl regulatory build, we also aggregated methylation over transcription factor binding sites in the genome as has been previously demonstrated[4]. This analysis also produced cell type separation, though not as cleanly as when using the entire regulatory build (Supplementary Figs. 22,23). Lastly, we sought to characterize intra-cluster variability by determining regions of the genome with non-binary methylation status (*i.e.,* multiple cells covering the CG site with discordant methylation status). We observed a depletion for non-binary CG sites in regions with active histone marks, indicating low cell-cell variability, and an enrichment for loci harboring the repressive mark, H3K27me3, indicating greater cell-cell variability (Supplementary Fig. 24). When extending this analysis to transcription factor binding sites, there was depletion for NRF1 and CTCF and enrichment for AP-1 associated factors, as expected (Supplementary Fig. 25, Supplementary Data 4).

Supplementary Note 2: Sci-MET Protocol

Reagents

- Phosphate Buffer Saline (PBS, Thermo Fisher, Cat. 10010023)
- 0.25% Trypsin (Thermo Fisher, Cat. 15050057)
- Tris (Fisher, Cat. T1503)
- HCl(Fisher, Cat. A144)
- NaCl (Fisher, Cat. M-11624)
- MgCl2 (Sigma, Cat. M8226)
- Igepal CA-630 (Sigma, I8896)
- Protease Inhibitors (Roche, Cat. 11873580001)
- HEPES (Fisher, Cat BP310)
- UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, 10977023
- Lithium 3,5-diiodosalicylic acid (Sigma, Cat. D3635) – LAND Only
- Formaldehyde (Sigma, Cat. F8775) – xSDS Only
- Glycine (Sigma, Cat. G8898) – xSDS Only
- NEBuffer 2.1 (NEB, Cat. B7202) – xSDS Only
- SDS (Sigma, Cat. L3771) – xSDS Only
- Triton-X100 (Sigma, Cat. 9002-93-1) – xSDS Only
- DAPI (Thermo Fisher, Cat. D1306)
- TD buffer from Nextera kit (Illumina, Cat. FC-121-1031)
- 96 Indexed Cytosine-depleted Transposomes (assembled using published methods, sequences shown in Supplementary Data 1)
- 9-Nucleotide Random Primer (Supplementary Table 1)
- 10 mM dNTP Mix (NEB, Cat. N0447)
- Klenow (3'->5' Exo-) Polymerase (Enzymatics, Cat. P7010-LC-L)
- 200 Proof Ethanol (Decon, Cat. 2705)
- Indexed i5 and i7 PCR primers (Supplementary Table 2)
- Kapa HiFi HotStart ReadyMix (Kapa Biosystems, Cat. KK2602)
- SYBR Green (FMC BioProducts, Cat. 50513)
- Qiaquick PCR purification kit (Qiagen, Cat. 28104)
- dsDNA High Sensitivity Qubit (Thermo Fisher, Cat. Q32851)
- High Sensitivity Bioanalyzer kit (Agilent, Cat. 5067-4626)
- Unmethylated Lambda DNA (Promega, Cat. D1521)
- Illumina sequencing kit corresponding to platform
- EZ-96 DNA Methylation MagPrep Kit (Zymo Research, Cat D5040)
- Custom LNA Sequencing primers (Supplementary Table 3)
- Poly(ethylene glycol) BioUltra 8,000(PEG; Sigma, Cat 89510)
- Sera-Mag SpeedBeads (GE, Cat 65152105050250)

Equipment

- 35μM Cell Strainer (BD Biosciences, Cat. 352235)
- Dounce Homogenizer
- 96-well plate compatible magnetic rack
- Sony SH800 cell sorter (Sony Biotechnology, Cat. SH800) or other FACS instrument capable of DAPI based single nuclei sorting

- CFX Connect RT Thermal Cycler (Bio-Rad, Cat. 1855200) or other real time thermocycler
- Thermomixer (Thermo Fisher, Cat. 5382000015) or other heat block.
- Qubit 2.0 Flourometer (Thermo Fisher, Cat. Q32866)
- 2100 Bioanalyzer (Agilent, Cat. G2939A)
- Illumina NextSeq 500, or HiSeq 2500, or HiSeq X

Protocol

1. Preparation of Unmethylated Control Lambda DNA

a. Combine 100 ng of Unmethylated Lambda DNA, 5 uL of 2X TD Buffer, 5 uL NIB (10mM TrisHCl pH7.4, 10MM NaCl, 3mM MgCl2, 0.1% igepal, 1x protease inhibitors), 4 µL 500 nM of uniquely indexed cytosine-depleted transposome.
b. Incubate for 20 min at 55°C.
c. Purify using Qiaquick PCR Purification column and elute in 30 µL of EB.
d. Quantify 2 uL with dsDNA High Sensitivity Qubit 2.0 Fluorometer
e. Dilute to concentration of 17.95 pg/uL (to simulate genomic mass of roughly 5 human cells.)

2. Preparation of 18% PEG SPRI Bead Mixture

a. Aliquot 1mL Sera-Mag beads to a low-bind 1.5 mL tube.
b. Place on magnetic stand until supernatant is cleared.
c. Wash beads with 500 uL 10mM Tris-HCl, pH 8.0 and remove once supernatant is cleared.
d. Repeat the wash in step 2.c. three more times for a total of four washes.
e. Resuspend beads in the following mixture: 18% PEG 8000 (by mass), 1M NaCl, 10mM Tris-HCl, pH 8.0, 1mM EDTA, 0.05% Tween-20.
f. Incubate at room temperature with mild agitation for at least an hour.
g. Store 18% PEG SPRI beads in 4°C. Allow beads to reach room temperature before use.

3. Prepare Nuclei Following either the LAND or xSDS protocol

LAND Nuclei Preparation & Nucleosome Depletion

a. Prepare nuclei in accordance with sample type below:

Suspension Cell Culture:
i. Triturate gently to break up cell clumps.
ii. Pellet cells by spinning at 500xg for 5 minutes at 4°C.
iii. Wash with 500 µL ice cold PBS.

Adherent Cell Culture:
i. Aspirate media and wash cells with 10 mL of PBS at 37°C.
ii. Add enough 0.25% Trypsin at 37° to cover monolayer.
iii. Incubate at 37°C for 5 minutes or until 90% of cells are no longer adhering to surface.
iv. Add 37°C media at 1:1 ratio to quench Trypsin.
v. Pellet cells by spinning at 500xg for 5 minutes at 4°C.
vi. Wash with 500 µL ice cold PBS.

b. Pellet cells by spinning at 500xg for 5 minutes.
c. Resuspend in 200 µL 12.5 mM LIS in NIB buffer (2.5 µL 1M LIS + 197.5 µL NIB buffer).
d. Incubate on ice for 5 minutes.
e. Add 800 µL NIB buffer.
f. Gently pass through 35µM cell strainer.
g. Add 5 µL DAPI (5 mg/mL).

xSDS Nuclei Preparation & Nucleosome Depletion

a. Prepare nuclei in accordance with sample type below:

Suspension Cell Culture:

i. Triturate gently to break up cell clumps.
ii. To 10 mL of cells in media add 406 µL of 37% formaldehyde and incubate at room temp for 10 minutes with gentle shaking.
iii. Add 800 µL of 2.5 M Glycine and incubate on ice for 5 minutes.
iv. Centrifuge at 550xg for 8 minutes at 4°C.
v. Wash with 10 mL of ice cold PBS.
vi. Resuspend cells in 5 mL of ice cold NIB (10mM TrisHCl pH7.4, 10MM NaCl, 3mM MgCl2, 0.1% igepal, 1x protease inhibitors).
vii. Incubate on ice for 20 minutes with gentle mixing.

Adherent Cell Culture:

i. Aspirate media and wash cells with 10 mL of PBS at 37°C.
ii. Add enough 0.25% Trypsin at 37° to cover monolayer.
iii. Incubate at 37°C for 5 minutes or until 90% of cells are no longer adhering to surface.
iv. Add 37°C media at 1:1 ratio to quench Trypsin.
v. Bring volume to 10ml with media.
vi. Resuspend in 10 mL media, add 406 µL of 37% formaldehyde, and incubate at room temp for 10 minutes with gentle shaking.
vii. Add 800 µL of 2.5 M Glycine and incubate on ice for 5 minutes.
viii. Centrifuge at 550xg for 8 minutes at 4°.
ix. Wash with 10 mL of ice cold PBS.
x. Resuspend cells in 5 mL of ice cold NIB.
xi. Incubate on ice for 20 minutes with gentle mixing.

Cortical Samples:

i. Grossly dissect cortical samples and flash freeze in liquid N_2.
ii. Cut samples with a sterile razor blade into smaller pieces on a chilled dish.
iii. Resuspend material in 5mL of ice cold modified NIB solution (NIB-HEPES; 20mM HEPES, 10mM NaCl, 3mM $MgCl_2$, 0.1% Igepal, 1X protease inhibitors).
iv. Give cells 5 minutes to equilibrate to the salt solution before dissociating in the following process:
    a. 5 loose strokes in a dounce homogenize
    b. 5 minutes to equilibrate again
    c. 5 loose strokes and 10 tight strokes.
v. Filter cells by running the solution through a 35um cell strainer.
vi. Spin nuclei in a pre-chilled 4C centrifuge for 5 minutes at 600xg.
vii. Resuspend in 5 mL NIB-HEPES

viii. Add 135 µL of 37% formaldehyde, and incubate at room temp for 10 minutes with gentle shaking.
ix. Add 400 µL of 2.5 M Glycine and incubate on ice for 5 minutes.
x. Centrifuge at 550xg for 8 minutes at 4°C.
xi. Resuspend cells in 5 mL of ice cold NIB (10mM TrisHCl pH7.4, 10MM NaCl, 3mM MgCl2, 0.1% igepal, 1x protease inhibitors).

b. Pellet cells or nuclei by spinning at 500xg for 5 minutes and wash with 900 µL of 1x NEBuffer 2.1
c. Spin 500 x g for 5 minutes.
d. Resuspend in 800 µL 1x NEBuffer 2.1 with 12 µL of 20% SDS and incubate at 42°C with vigorous shaking for 30 minutes.
e. Add 200 µL of 10% Triton-X and incubate at 42°C with vigorous shaking for 30 minutes.
f. Pass through a 35 µm cell strainer.
g. Add 8 µL of 5mg/mL DAPI and proceed to flow sorting.

4. Sorting and Tagmentation

a. Prep tagmentation plate with 10 µL 1x TD buffer (for 1 plate: 500 µL NIB buffer + 500 µL TD buffer)
b. Sort 2500 single nuclei into each well of the tagmentation plate. NOTE: At this step the number of nuclei per well can be varied slightly as long as the number of nuclei per well is consistent for the whole plate. It is also possible to multiplex different samples into different wells of the plate as the transposase index will be preserved.
c. Gate to isolate single, high quality nuclei.
d. Spin down plate at 500 x g for 5 min.
e. Add 4 µL 500 nM of uniquely indexed cytosine-depleted transposome to each well.
f. Seal plate and incubate at 55°C for 15 minutes with gentle shaking.
g. Place plate on ice.
h. Pool all wells, pass through 35µM cell strainer.
i. Add 8 µL 5mg/mL DAPI.

5. Second Sort

a. Prepare a master mix for each well with 5uL Zymo Digestion Reagent (2.5 uL M-Digestion Buffer, 2.25 uL H2O, and 0.25 uL Proteinase K).
b. Sort 10 or 22 single nuclei into each well using the most stringent sort settings.
    a. 10 single nuclei wells to be used for unmethylated control spike-ins.
c. Spin down plate at 600 x g for 5 min at 4˚C.

6. Digestion and Bisulfite Conversion

a. Spike in ~35 pg (2uL) of Unmethylated Control Lambda DNA Pre-treated with a C-depleted transposome (1) into wells with 10 single nuclei.
b. Incubate plate for 20 minutes at 50°C to digest nuclei.
c. Add 32.5 uL Freshly Prepared Zymo CT Conversion Reagent following manufacturer's protocol.
d. Mix wells by triturating.
e. Spin down plate at 600 x g for 2 min at 4˚C.

f. Place plate on thermocycler for the following steps: 98°C for 8 minutes, 64°C for 3.5 hours, then hold at 4°C for less than 20 hours before continuing.
g. Add 5uL of Zymo MagBinding Beads to each well, and 150 uL of M-Binding Buffer.
h. Mix wells by triturating.
i. Incubate plate at room temperature for 5 minutes.
j. Place plate on 96-well compatible magnetic rack until supernatant is clear.
k. Remove supernatant.
l. Wash wells with fresh 80% Ethanol (by volume).
    i. Remove plate from magnetic rack.
    ii. Add 100 uL of 80% Ethanol to each well, running over bead pellet.
    iii. Place plate back on magnetic rack and remove supernatant once clear.

m. Desulphonation
    i. Add 50 uL M-Desulphonation Buffer to each well.
    ii. Resuspend beads fully by trituration.
    iii. Incubate at room temperature for 15 minutes.
    iv. Place plate on magnetic rack and remove supernatant once clear.

n. Wash wells again repeating step (6.l.).
o. Allow bead pellets to dry for ~10 minutes until pellets begin to visibly crack.
p. Elution
    i. Add 25 uL of Zymo M-Elution Buffer to each well.
    ii. Triturate to fully dissociate pellet.
    iii. Heat the plate at 55°C for 4 minutes.

7. Linear Amplifications

a. Move full elution to a plate prepared with the following reaction mix per well: 16 uL PCR-clean H2O, 5 uL 10X NEBuffer 2.1, 2 uL 10 mM dNTP Mix, and 2 uL 10 uM 9-Nucleotide Random Primer.
b. Perform the following linear amplification:
    i. Render DNA single-stranded by incubating at 95°C for 45 seconds, then flash cool on ice and hold on ice.
    ii. Add 10U Klenow (3'->5' exo-) polymerase to each well once fully cooled.
    iii. Incubate plate at 4°C for 5 minutes, then ramp temperature up at a rate of +1°C/15 sec to 37°C, then hold at 37°C for 90 minutes.

c. Repeat step (7b) three more times for a total of 4 rounds of linear amplification.
    i. Each time, add the following mixture to the reaction in each well: 1 uL 10 uM 9-Nucleotide Random Primer, 1 uL 10 mM dNTP Mix, and 1.25 uL 4X NEBuffer 2.1.
    ii. Note: We have found that four rounds of linear amplification significantly increases the read alignment rate and library complexity compared to fewer rounds.

d. Clean up wells using the prepared 18% PEG SPRI Bead Mixture at 1.1X (concentration by volume compared to well reaction volume).
e. Incubate plate for 5 minutes at room temperature.
    i. Place plate on magnetic rack and remove supernatant once clear.
    ii. Wash bead pellets with 50 uL 80% Ethanol.
    iii. Remove any liquid and allow bead pellet to dry until beginning to crack.

iv. Elute DNA in 21 uL 10 mM Tris-Cl (pH 8.5).

8. Indexing PCR Reaction

a. Move full elution to a plate prepared with the following reaction mix per well: 2 uL of 10 uM i7 index PCR primer, 2 uL of 10 uM i5 index PCR primer, 25 uL of 2X KAPA HiFi HotStart ReadyMix, and 0.5 uL 100X Sybr Green I.
b. Perform the following PCR cycles on a real-time thermocycler: 95°C for 2 minutes, (94°C for 80 seconds, 65°C for 30 seconds, 72°C for 30 seconds [Image].)xN Repeats
c. Stop the reaction once a majority of wells show an inflection of measured SYBR green fluorescence.
d. Note: We observed inflection plateaus between 18-21 PCR cycles for our library preparations.

9. Library Clean Up and Quantification

a. Pool libraries using 5 uL of each well.
b. Clean libraries per-well using the 18% PEG SPRI Bead Mixture at 0.8X (concentration by volume compared to pooled library volume) following the same protocol as step (7d-e).
c. Elute DNA in 25 uL 10 mM Tris-Cl (pH 8.5).
d. Quantify 2 uL with dsDNA High Sensitivity Qubit 2.0 Fluorometer, following manufacturer's protocol.
e. Use Qubit readout to dilute library to ~4 ng/uL and run 1 uL on a High Sensitivity Bioanalyser 2100, following manufacturer's protocol.
f. Quantify library for the 200bp – 1 kbp range to dilute the pool to 1 nM for Illumina Sequencing.

10. Sequencing

a. Set up NextSeq run as per manufacturer's instructions for a 1 nM sample except for the following changes.
b. Library pool should be loaded at a concentration of 0.9 pM and a total volume of 1.5 mL and deposited into catridge position 10.
c. Setup custom primers by diluting 9 µL of 100 µM stock sequencing primer 1 into a total of 1.5 mL of HT1 buffer into cartridge position 7; and 18 µL of each custom index sequencing primer at 100 µM stock concentrations to a total of 3 mL of HT1 buffer into cartridge position 9.
d. Nextseq should be operating in standalone mode. Choose the SCIseq custom chemistry recipe (Amini et. al. 2014). Select dual index. Enter appropriate number of read cycles (150 recommended). And 10 cycles for index 1 and 20 cycles for index 2. Select the custom checkbox for all reads and indices.

Supplementary Table 1: sciMET 9-nulceotide Random Primer (5'-3')

Name	Sequence
sciMET_N9_IPE2	GGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNN

Supplementary Table 2: sciMET PCR primers (5'-3')

Name	Sequence
sciMET_i7_1	CAAGCAGAAGACGGCATACGAGATcaagatgccgGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_2	CAAGCAGAAGACGGCATACGAGATaacgtctagtGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_3	CAAGCAGAAGACGGCATACGAGATaggtatactcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_4	CAAGCAGAAGACGGCATACGAGATttcataggacGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_5	CAAGCAGAAGACGGCATACGAGATggaggcctccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_6	CAAGCAGAAGACGGCATACGAGATttcaatataaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_7	CAAGCAGAAGACGGCATACGAGATacgtcatataGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_8	CAAGCAGAAGACGGCATACGAGATttgaccaggaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_9	CAAGCAGAAGACGGCATACGAGATcggttgcgcgGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_10	CAAGCAGAAGACGGCATACGAGATcaaggaggtcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_11	CAAGCAGAAGACGGCATACGAGATttacgatgaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_12	CAAGCAGAAGACGGCATACGAGATttgctggcatGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_13	CAAGCAGAAGACGGCATACGAGATaatactcttcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_14	CAAGCAGAAGACGGCATACGAGATccaactaaccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_15	CAAGCAGAAGACGGCATACGAGATtatcctcaatGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_16	CAAGCAGAAGACGGCATACGAGATgccgtcgcgtGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_17	CAAGCAGAAGACGGCATACGAGATccgctgcttcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_18	CAAGCAGAAGACGGCATACGAGATtgaccgaatcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_19	CAAGCAGAAGACGGCATACGAGATgtctccagagGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_20	CAAGCAGAAGACGGCATACGAGATaatgctagtcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_21	CAAGCAGAAGACGGCATACGAGATgacgacctgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_22	CAAGCAGAAGACGGCATACGAGATagagccagccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_23	CAAGCAGAAGACGGCATACGAGATccaggccgcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i7_24	CAAGCAGAAGACGGCATACGAGATcaggtatggaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
sciMET_i5_1	AATGATACGGCGACCACCGAGATCTACACgtatcatcgaGGTGTAGTGGGTTTGG
sciMET_i5_2	AATGATACGGCGACCACCGAGATCTACACccgcgattatGGTGTAGTGGGTTTGG
sciMET_i5_3	AATGATACGGCGACCACCGAGATCTACACattcaggtacGGTGTAGTGGGTTTGG
sciMET_i5_4	AATGATACGGCGACCACCGAGATCTACACatggaattggGGTGTAGTGGGTTTGG
sciMET_i5_5	AATGATACGGCGACCACCGAGATCTACACgacgaagcgtGGTGTAGTGGGTTTGG
sciMET_i5_6	AATGATACGGCGACCACCGAGATCTACACcttgcagtagGGTGTAGTGGGTTTGG
sciMET_i5_7	AATGATACGGCGACCACCGAGATCTACACcttggtaatgGGTGTAGTGGGTTTGG
sciMET_i5_8	AATGATACGGCGACCACCGAGATCTACACcaagtcgaccGGTGTAGTGGGTTTGG

Supplementary Table 3: sciMET Sequencing Primers (LNA, 5'-3')

Name	Sequence
sciMET_Read1	TGGTAGAGAGGGTG AGATGTGTATAAGAGATAG
sciMET_Index1	CTATCTCTTATACACATCT CACCCTCTCTACCA

Supplementary Table 4: sci-MET Library Preparation Cost Breakdown

Cost Per Plate:

Preparation Stage	Reagent	Vol. (x1;uL)	Vol. (x96;uL)	Catalog	Cost Of Total Reagent	Frac. Used	Specific Cost
Brain Dissociation	Sterile Razor Blade	NA	1	VWR, Cat 55411-050	$42.75	0.01	$0.43
	Sterile Petri Dish	NA	1	VWR, Cat 25384-092	$259.65	0.002	$0.52
	35um cell strainer	NA	1	BD Biosciences, Cat. 352235	$668.55	0.002	$1.34
xSDS Treatment	Formaldehyde	NA	135	Sigma, Cat. F8775	$43.30	0.0014	$0.06
	2.5M Glycine	NA	400	Sigma, Cat. G8898	$52.70	0.0002	$0.01
	NEB 2.1 (10X)	NA	170	NEB, Cat. B7202	$21.00	0.0034	$0.07
	20% SDS	NA	12	Sigma, Cat. L3771	$65.40	2E-05	$0.00
	10% Triton-X	NA	200	Sigma, Cat. 9002-93-1	$36.70	0.0002	$0.01
	5mg/mL DAPI	NA	8	Thermo Fisher, Cat. D1306	$98.25	0.004	$0.39
	35um cell strainer	NA	1	BD Biosciences, Cat. 352235	$668.55	0.002	$1.34
First Sort	1X TD Buffer	5	480	Illumina, Cat. FC-121-1031	$23.70	0.003	$0.07
	NIB Buffer	5	480	See Reagent Preparation Table	$76.08	0.0096	$0.73
	Low Profile Strip Tubes	NA	12	BioRad, TLS0801	$110.00	0.1	$11.00
	500nM C-Depleted tn5	4	384	(Cost based on Picelli et. al. 2014)	$450.00	0.01	$4.50
	35 um cell strainer	NA	1	BD Biosciences, Cat. 352235	$668.55	0.002	$1.34
	5 mg/mL DAPI	NA	8	Thermo Fisher, Cat. D1306	$98.25	0.004	$0.39
	p10 Pipettor Tips (Full Box)	NA	1	USA Scientific, Cat 1120-3710	$76.70	0.1	$7.67
	Plate Seal	NA	1	BioRad, MSB1001	$170.00	0.01	$1.70
Second Sort	Zymo M-Digestion Buffer	2.5	240	Zymo D5020-9	$10.00	0.03	$0.30
	UltraPure DNase/RNase-Free Distilled Water	2.25	216	Invitrogen, 10977023	$164.00	4E-05	$0.01
	Proteinase K	0.25	24	Zymo D3001-2-5	$21.00	0.0231	$0.48
	Low Profile Strip Tubes	NA	12	BioRad, TLS0801	$110.00	0.1	$11.00
Digestion and Bisulfite Conversion	Zymo CT Conversion Reagent	32.5	3120	Zymo D5003-1, D5021-7,D5021-8,D5006-2	$122.00	0.2496	$30.45
	Zymo MagBinding Beads	5	480	Zymo D4100-2-8	$92.00	0.06	$5.52
	Zymo M-Binding Buffer	150	14400	Zymo D5006-3	$58.00	0.1152	$6.68
	Ethanol	160	15360	Decon, Cat. 2705	$167.50	0.0041	$0.68
	UltraPure DNase/RNase-Free Distilled Water	40	3840	Invitrogen, 10977023	$164.00	0.0008	$0.13
	Zymo M-Desulphonation Buffer	50	4800	Zymo D5002-5	$42.00	0.12	$5.04
	Zymo M-Elution Buffer	25	2400	Zymo D5041-6	$10.00	0.3	$3.00
	96-Well Plate	NA	1	Eppendorf, Cat. 951020427	$118.00	0.04	$4.72
	Plate Seal	NA	1	BioRad, MSB1001	$170.00	0.01	$1.70
Linear Amplifications	UltraPure DNase/RNase-Free Distilled Water	16	1536	Invitrogen, 10977023	$164.00	0.0003	$0.05
	10X NEBuffer 2.1	6.5	624	NEB, Cat. B7202	$21.00	0.0125	$0.26
	10mM dNTP Mix	5	480	NEB, Cat. N0447	$236.00	0.12	$28.32
	10uM 9NT-Random Primer	5	480	IDT Custom	$10.85	0.0302	$0.33
	Klenow(3'->5' exo-) polymerase	10U	960U	Enzymatics, Cat. P7010-LC-L	$393.00	0.096	$37.73
	18% PEG SPRI Bead Mixture	74.5	7152	See Reagent Preparation Table	$37.39	0.143	$5.35
	Ethanol	40	3840	Decon, Cat. 2705	$167.50	0.001	$0.17
	UltraPure DNase/RNase-Free Distilled Water	10	960	Invitrogen, 10977023	$164.00	0.0002	$0.03
	10mM Tris-Cl	21	2016	See Reagent Preparation Table	$13.29	0.0002	$0.00
	96-Well Plate	NA	1	Eppendorf, Cat. 951020427	$118.00	0.04	$4.72
	Plate Seal	NA	5	BioRad, MSB1001	$170.00	0.05	$8.50
PCR Reaction	10 uM i7 Index PCR Primer	2	192	IDT Custom	$139.92	0.0024	$0.34
	10 uM i5 Index PCR Primer	2	192	IDT Custom	$93.28	0.0036	$0.34
	KAPA HiFi HotStart ReadyMix	25	2400	Kapa Biosystems, Cat. KK2602	$664.13	0.384	$255.03
	100X Sybr Green I	0.5	48	FMC BioProducts, Cat. 50513	$250.00	0.0096	$2.40
Library Cleanup	18% PEG SPRI Bead Mixture	40	3840	See Reagent Preparation Table	$37.39	0.0768	$2.87
	10mM Tris-Cl	25	2400	See Reagent Preparation Table	$13.29	0.0002	$0.00
	Ethanol	40	3840	Decon, Cat. 2705	$167.50	0.001	$0.17
	UltraPure DNase/RNase-Free Distilled Water	10	960	Invitrogen, 10977023	$164.00	0.0002	$0.03
	dsDNA High Sensitivity Qubit 2.0 Fluorometer	NA	4	Thermo Fisher, Cat. Q32851	$89.00	0.04	$3.56
	DNA HS Bioanalyzer Chip	NA	1	Agilent, Cat. 5067-4626	$342	0.1	$34.20
	Plate Seal	NA	1	BioRad, MSB1001	$170.00	0.05	$8.50
	96-Well Plate	NA	1	Eppendorf, Cat. 951020427	$118.00	0.04	$4.72

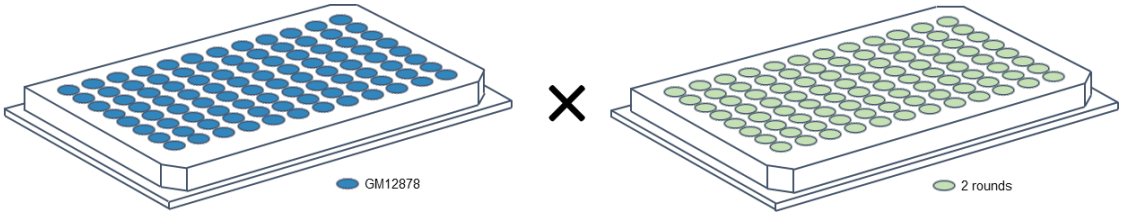
Total Cost Per Plate: $498.90

Supplementary Table 5: Unmethylated lambda control for bisulfite conversion calculation

	Human Mouse (xSDS)	Human cell types	Mouse Cortex
Number of De-duplicated, >Q30, Alignments to Lambda Genome	69,074	43,600,172	219,854
Total number of Cytosines analyzed	1,969,768	844,025,788	5,563,564
Total unmethylated Cytosines	1,955,068	836,742,874	5,521,569
Total methylated Cytosines	14,700	7,282,904	41,995
Bisulfite Conversion Efficiency	99.25%	99.14%	99.25%

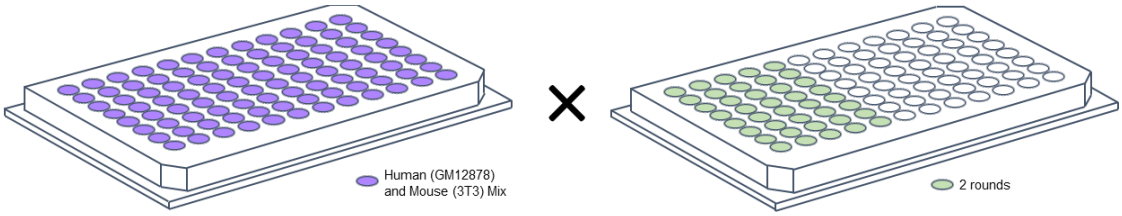
a

Expected Cell Count	$\frac{22\ Cells}{Well} x\ 96\ Wells$	2,112 Cells Expected
Observed Cell Count		708 Cells Produced



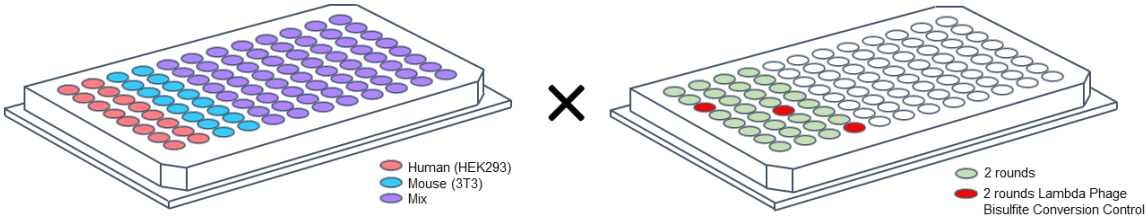
b

Expected Cell Count	$\frac{22\ Cells}{Well} x\ 40\ Wells$	880 Cells Expected
Observed Cell Count		711 Cells Produced 706 ≥ 10,000 Unique Reads



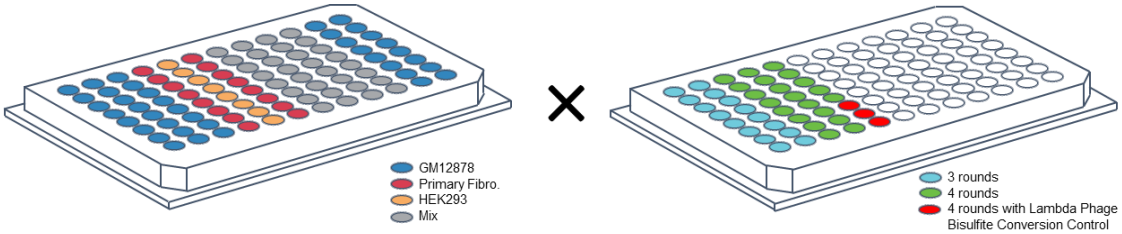
c

Expected Cell Count	$(\frac{22\ Cells}{Well} x\ 29\ Wells) +$	668 Cells Expected
Observed Cell Count	$\left(\frac{10\ Cells}{Control\ Well} \text{x 3 Control Wells}\right) =$	566 Cells Produced



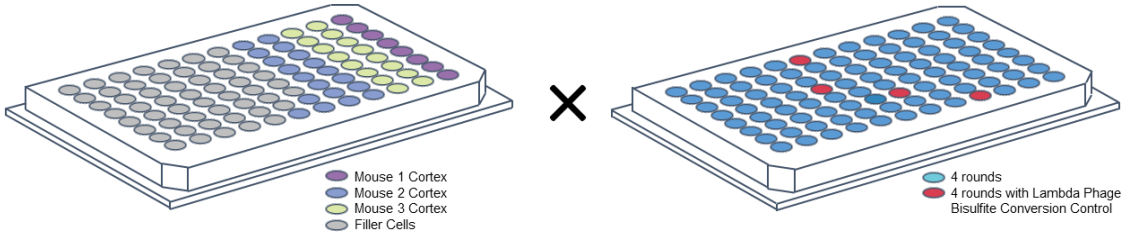
d

Expected Cell Count	$\left(\frac{22\ Cells}{Well} x\ 37\ Wells\right) + \left(\frac{10\ Cells}{Control\ Well} \text{x } 3 \text{ Control Wells}\right) =$	844 Cells Expected
Observed Cell Count		611 Cells Produced 683 ≥ 30,000 Unique Reads

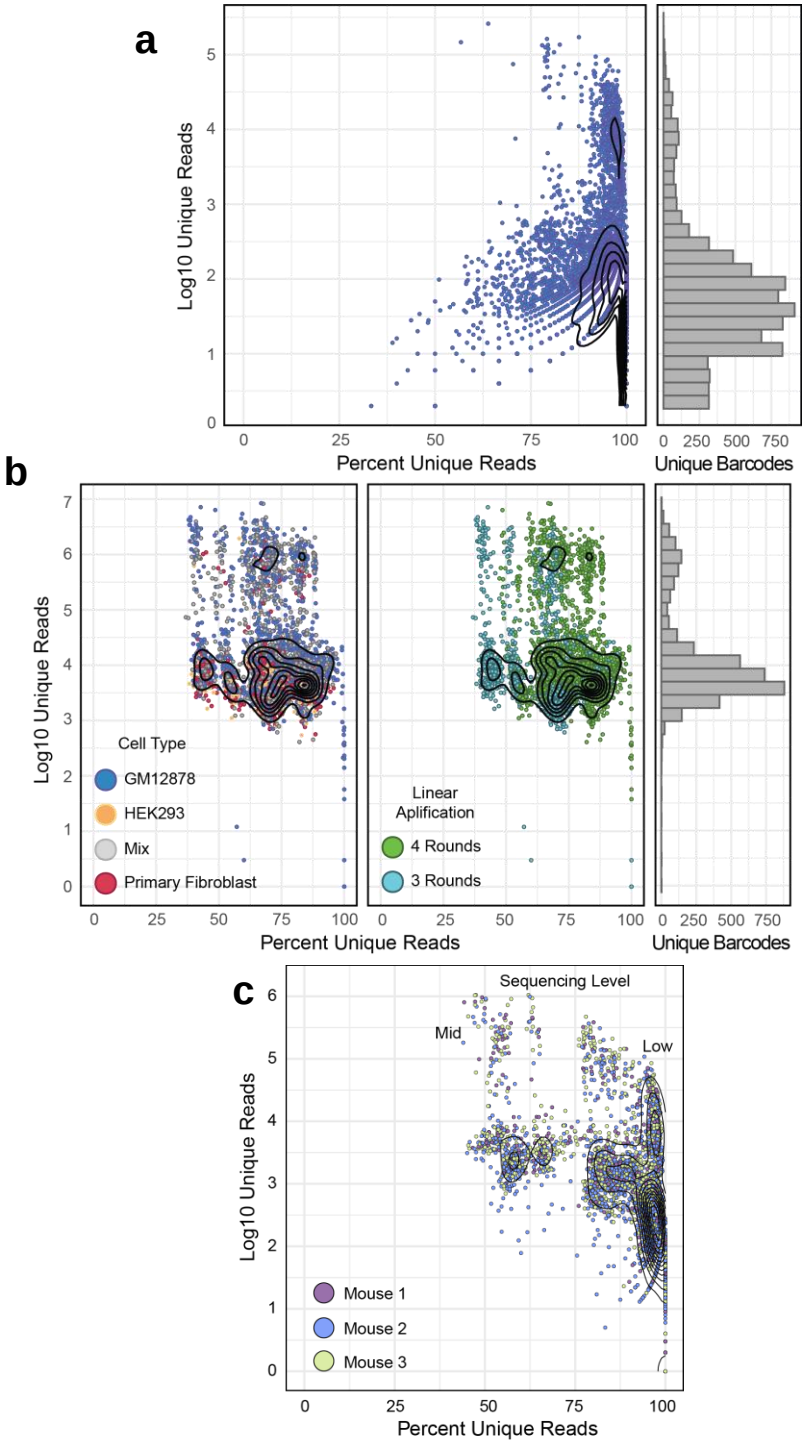


e

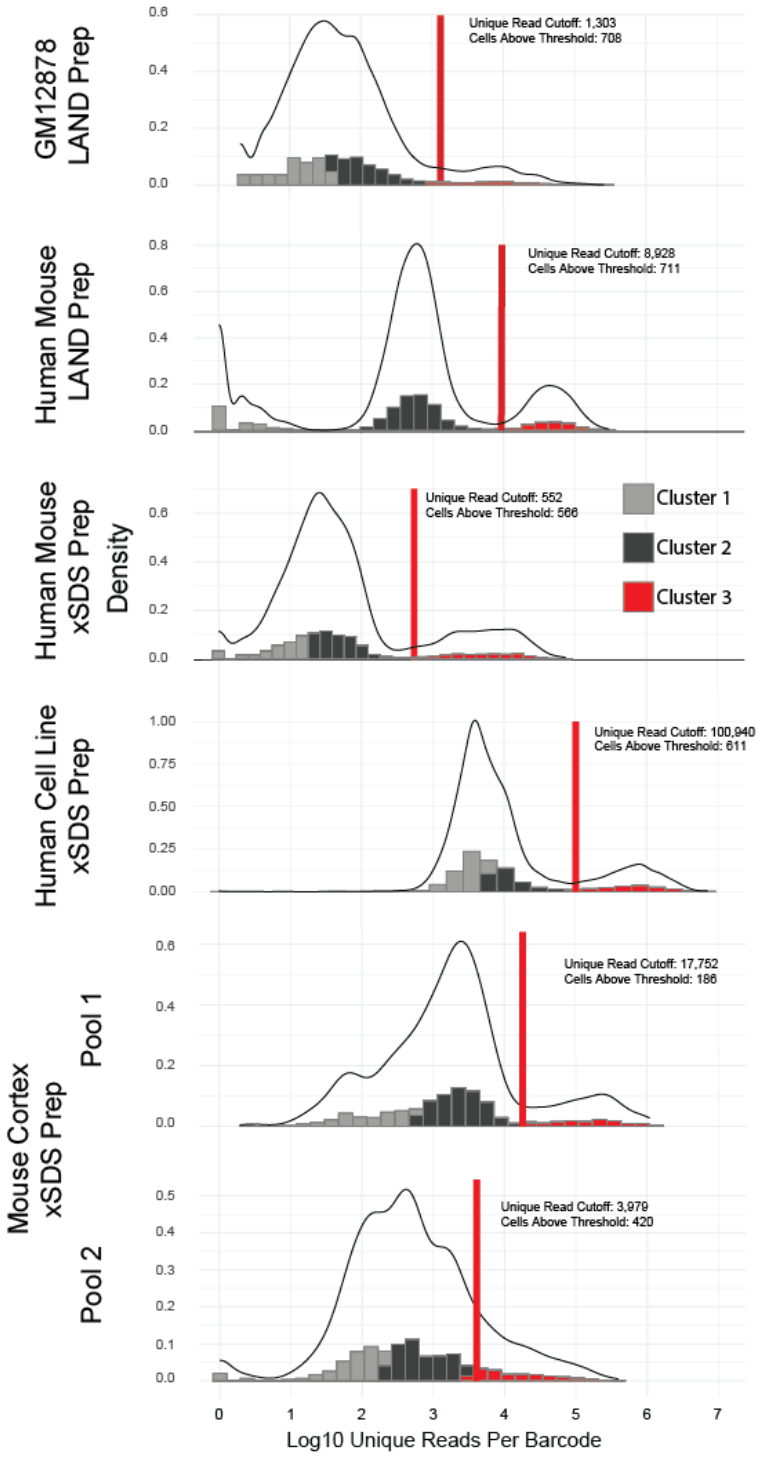
Expected Cell Count	$\left(\frac{22\ (43/96)\ Cells}{Well} x\ 92\ Wells\right) + \left(\frac{10\ (43/96)\ Cells}{Control\ Well} \text{x } 4 \text{ Control Wells}\right) =$	925 Cells Expected (Cortex Only)
Observed Cell Count		606 Cells Produced 285 ≥ 50,000 Unique Reads



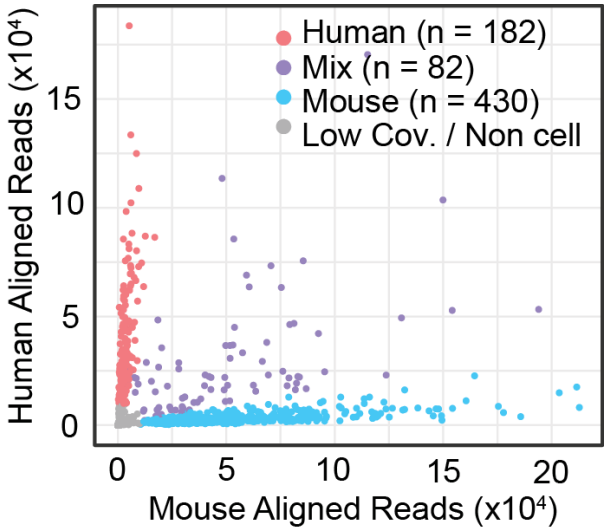
Supplementary Figure 1 | Summary of experiments and indexing design. **a.** GM12878 initial experiment using lithium-based nucleosome depletion. **b.** Human and mouse mix experiment using lithium-based nucleosome depletion. **c.** Human and mouse mix experiment using crosslinking and SDS nucleosome depletion. **d.** Human cell type mixing experiment using crosslinking and SDS nucleosome depletion. **e.** Mouse cortex crosslinking and SDS nucleosome depletion experiment. Filler cells indicate work not related to this study. Rounds indicates the number of rounds of linear amplification post-bisulfite treatment.



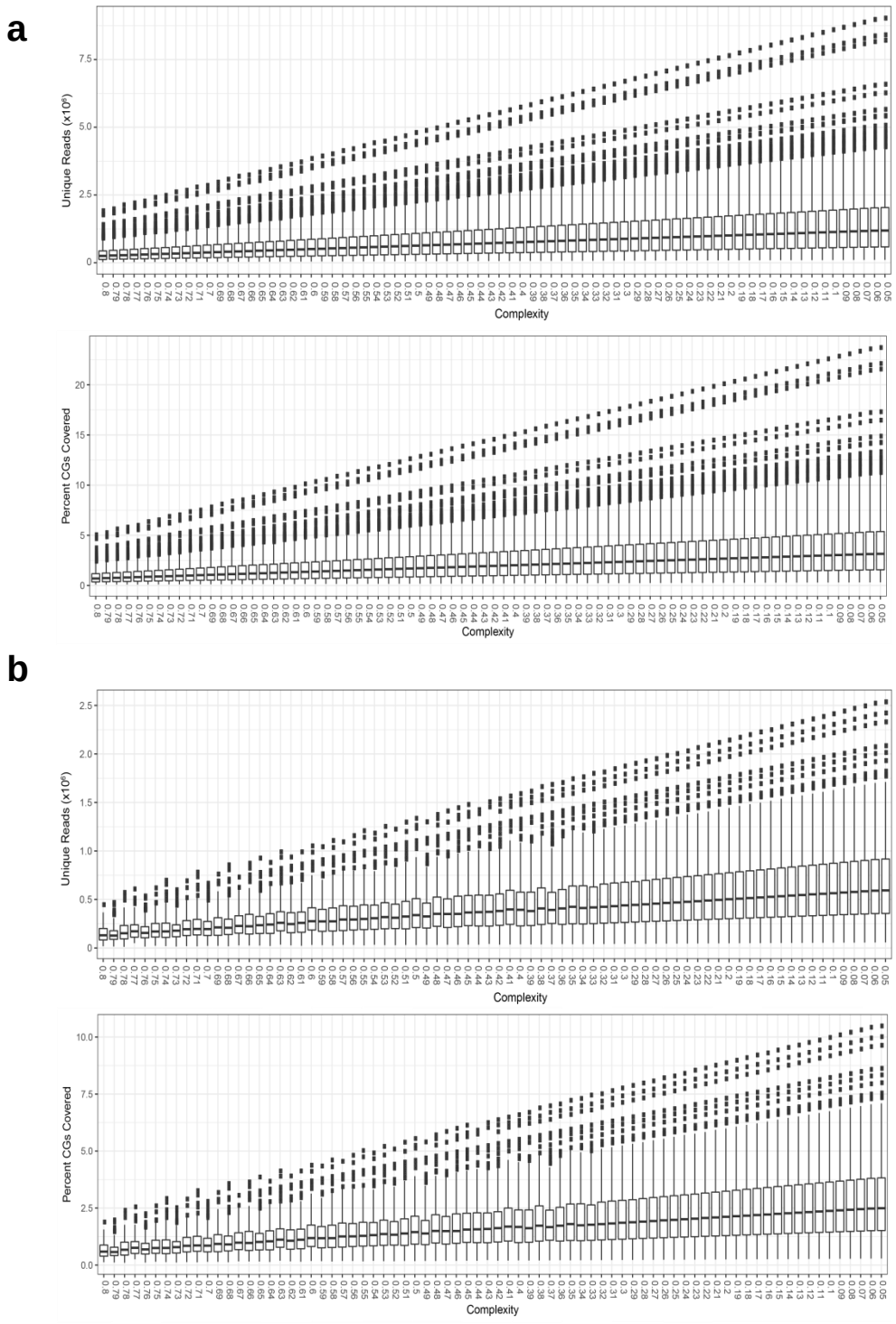
Supplementary Figure 2 | Library read count and unique read percentages for (**a**) the initial GM12878 only experiment which was sequenced to low depth, and (**b**) the human cell line mix experiment and (**c**) the mouse cortex experiment. X-axis indicates the percent of unique reads out of total aligned reads, Y-axis is the log10 unique read count for each cell (each represented as a point). Two dimension density overlay is illustrated by black lines within plot space. The high read count distribution indicates true single cell libraries and the lower distribution is the background noise. In the middle panel of **b**, rounds indicates the number of rounds of linear amplification. No histogram is shown for (**c**) due to the mix of distributions from various depths of sequencing.



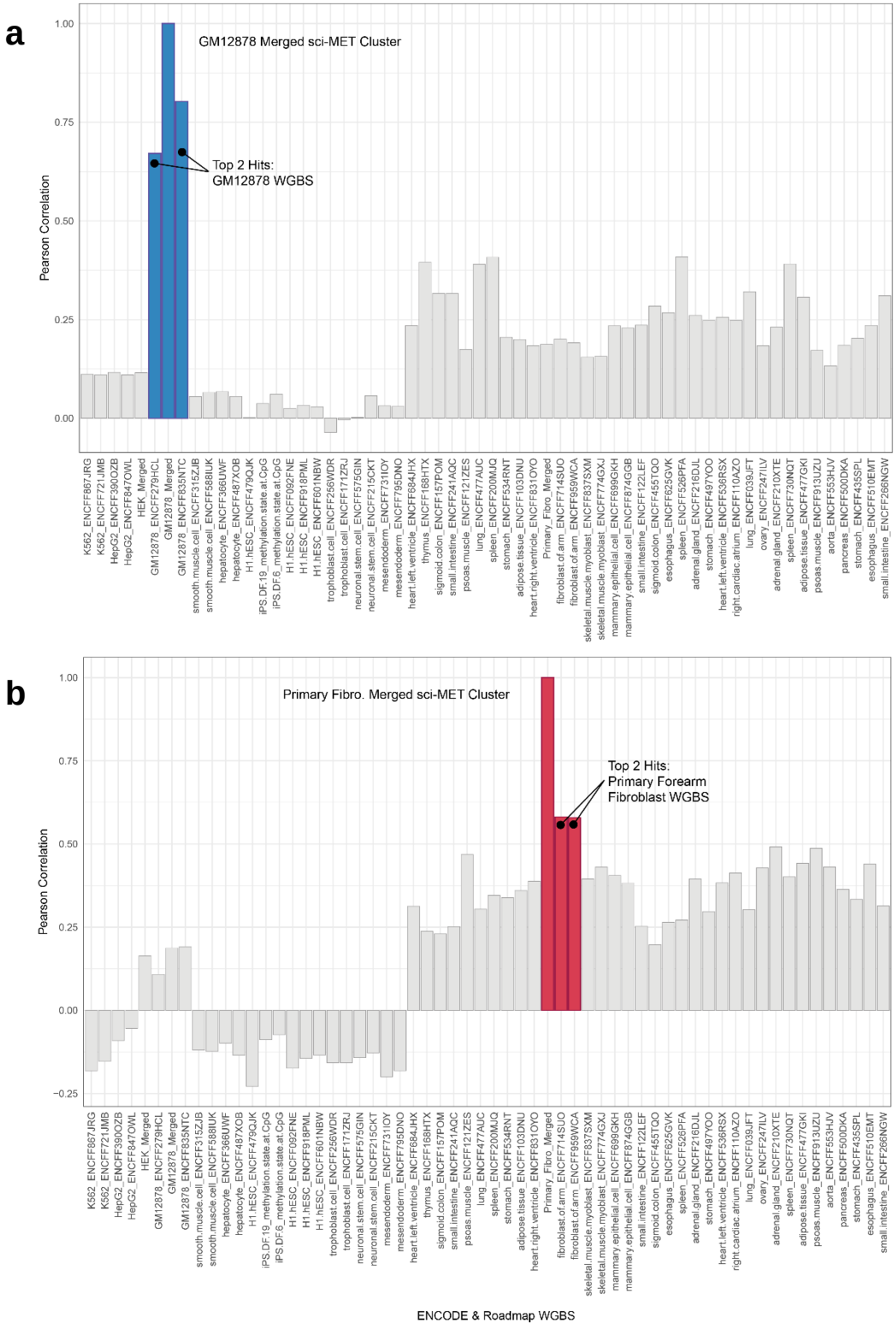
Supplementary Figure 3 | Discrimination of single-cell libraries through library read counts. Histograms and density plots of the unique aligned single-cell library preps. The clusters (k=3) are shown in gray, black, and red. The red vertical lines mark the read cutoff based on the 95% confidence interval of the cluster with the highest unique aligned reads. LAND = lithium assisted nucleosome depletion. xSDS = crosslinking and SDS nucleosome depletion. n=708, 711, 566, 611, 186, 420 cells defined by threshold cut-off (in order of top panel to bottom panel).

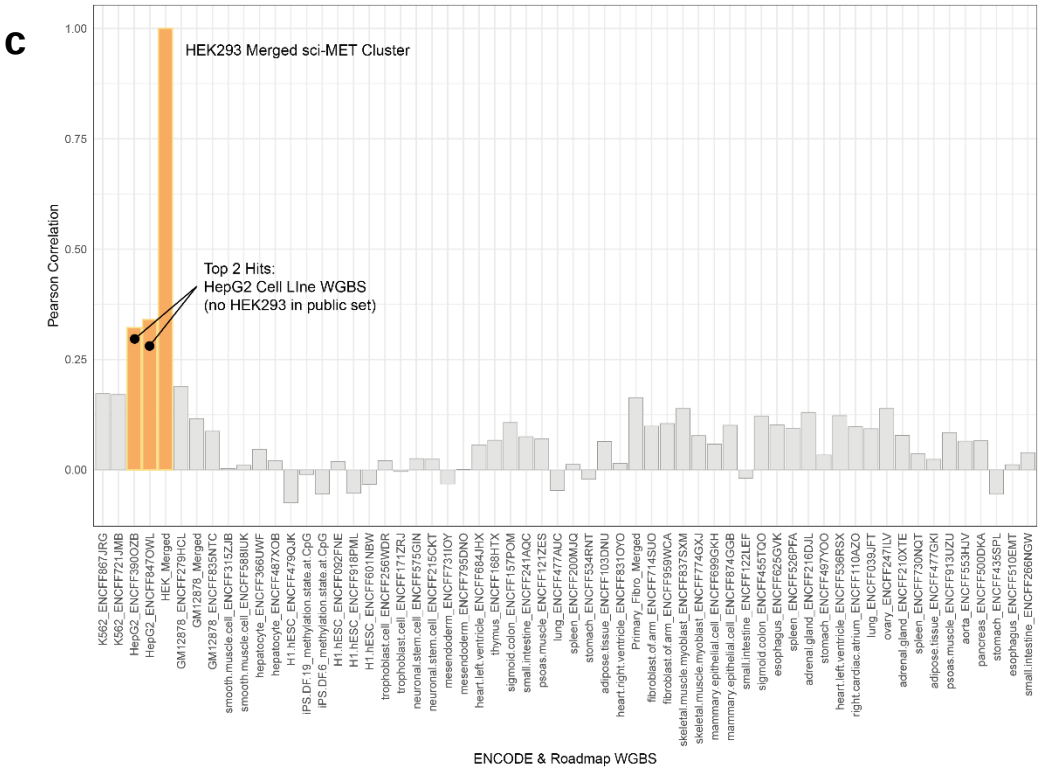


Supplementary Figure 4 | sci-MET assay on a mixture of human (GM12878) and mouse (NIH/3T3) cells using lithium-based nucleosome depletion. Red reflects barcode libraries with >90% of unique reads aligning to the human genome (hg19). Blue reflects barcode libraries with >90% of unique reads aligning to the mouse genome (mm10). Purple reflects barcode libraries with unique alignments exceeding >10% to both reference genomes. Gray-colored barcode libraries did not read count threshold of $1x10^4$ and were not counted in the collision estimate. A high collision rate (22% total estimated collision rate) was observed, possibly due to library molecule leakage after transposition.

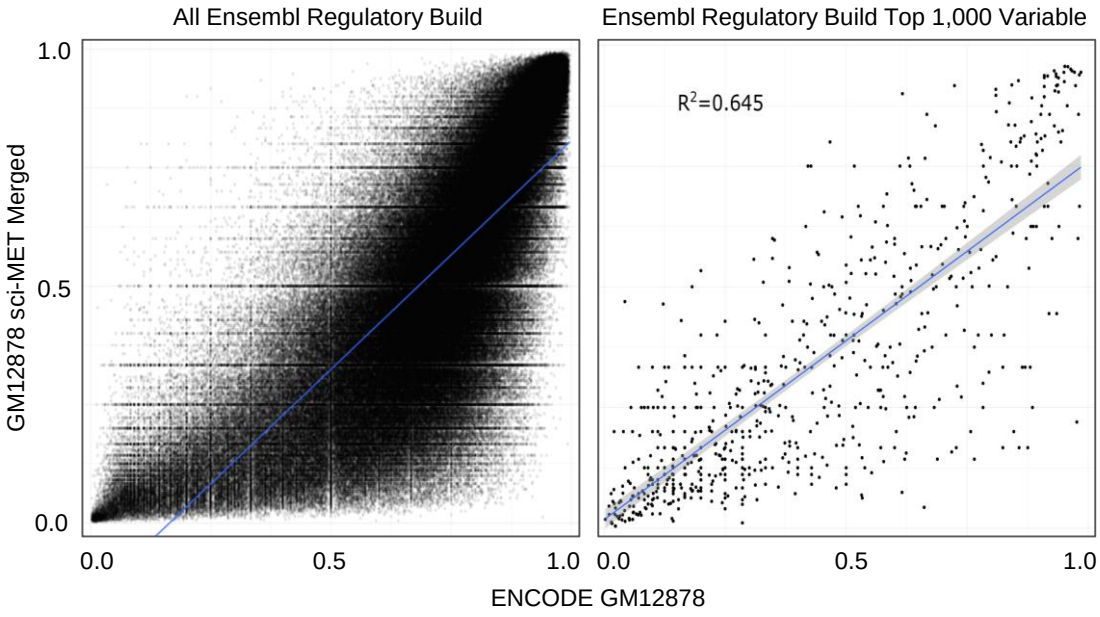


Supplementary Figure 5 | Library saturation projections for the human cell type mix (**a**) and the mouse cortex experiment (**b**). As libraries are projected to additional raw sequencing reads, the overall complexity is reduced. Projections are out to 5% unique molecules and considered saturated. Box plots describe median, with hinges for lower 25% quantile and upper 75% quantile, whiskers describe smallest observation greater than or equal to lower hinge – 1.5 x the interquartile region, and the largest observation less than or equal to upper hinge + 1.5 x the interquartile region.

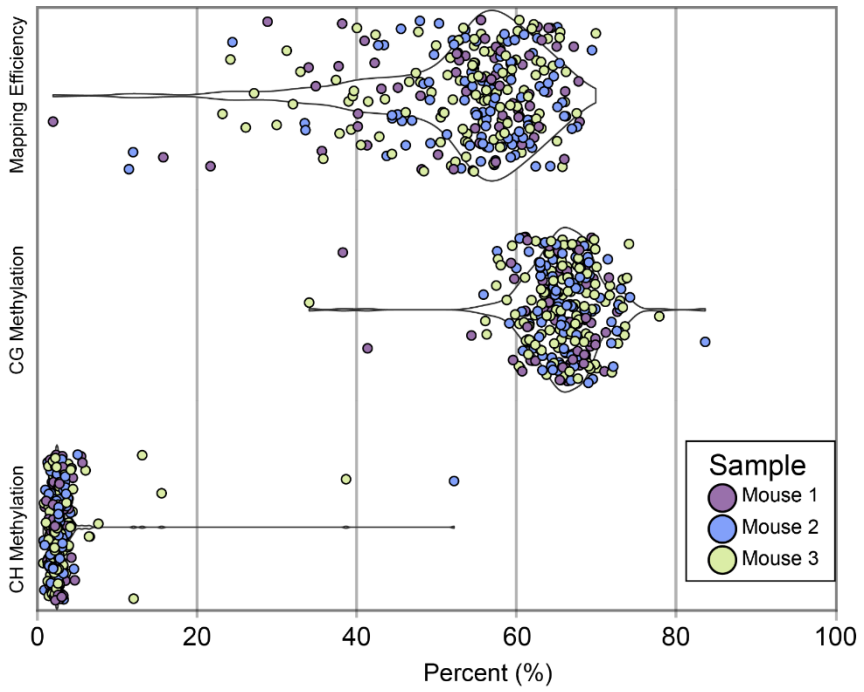

a
GM12878 Merged sci-MET Cluster
Top 2 Hits:
GM12878 WGBS
Pearson Correlation
1.00
0.75
0.50
0.25
0.00
b
Primary Fibro. Merged sci-MET Cluster
Top 2 Hits:
Primary Forearm
Fibroblast WGBS
Pearson Correlation
1.00
0.75
0.50
0.25
0.00
−0.25
ENCODE & Roadmap WGBS



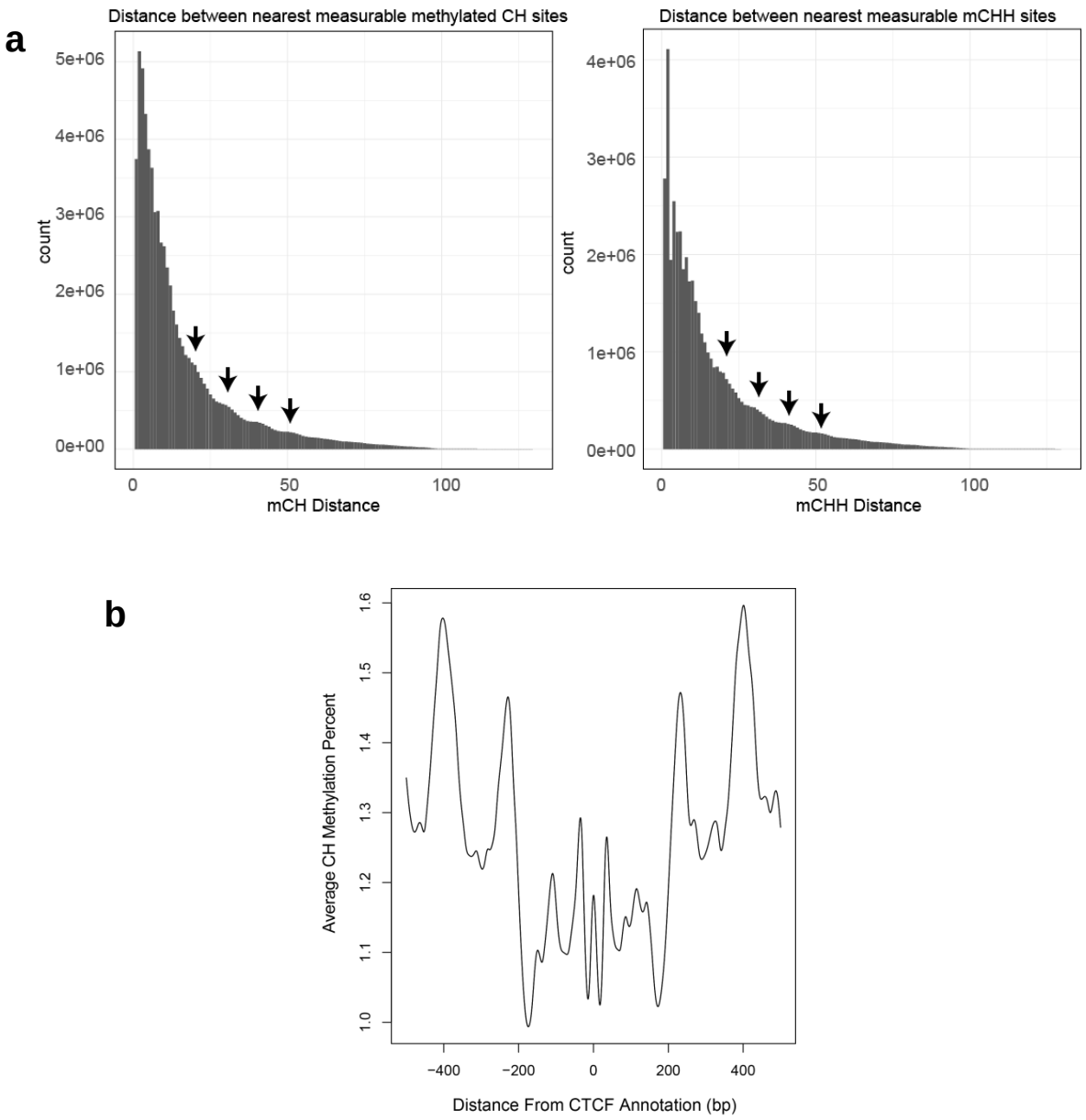
Supplementary Figure 6 | Correlation of aggregate clusters to ENCODE WGBS data sets for the top 1,00 most variable Ensembl regulatory regions for the (**a**) GM12878 (599 single-cell libraries), (**b**) Primary Fibro. (n=32 single-cell libraries), and (**c**) HEK293 merged cell clusters (n=10 single-cell libraries). Datasets are in the order of the hierarchical clustering.



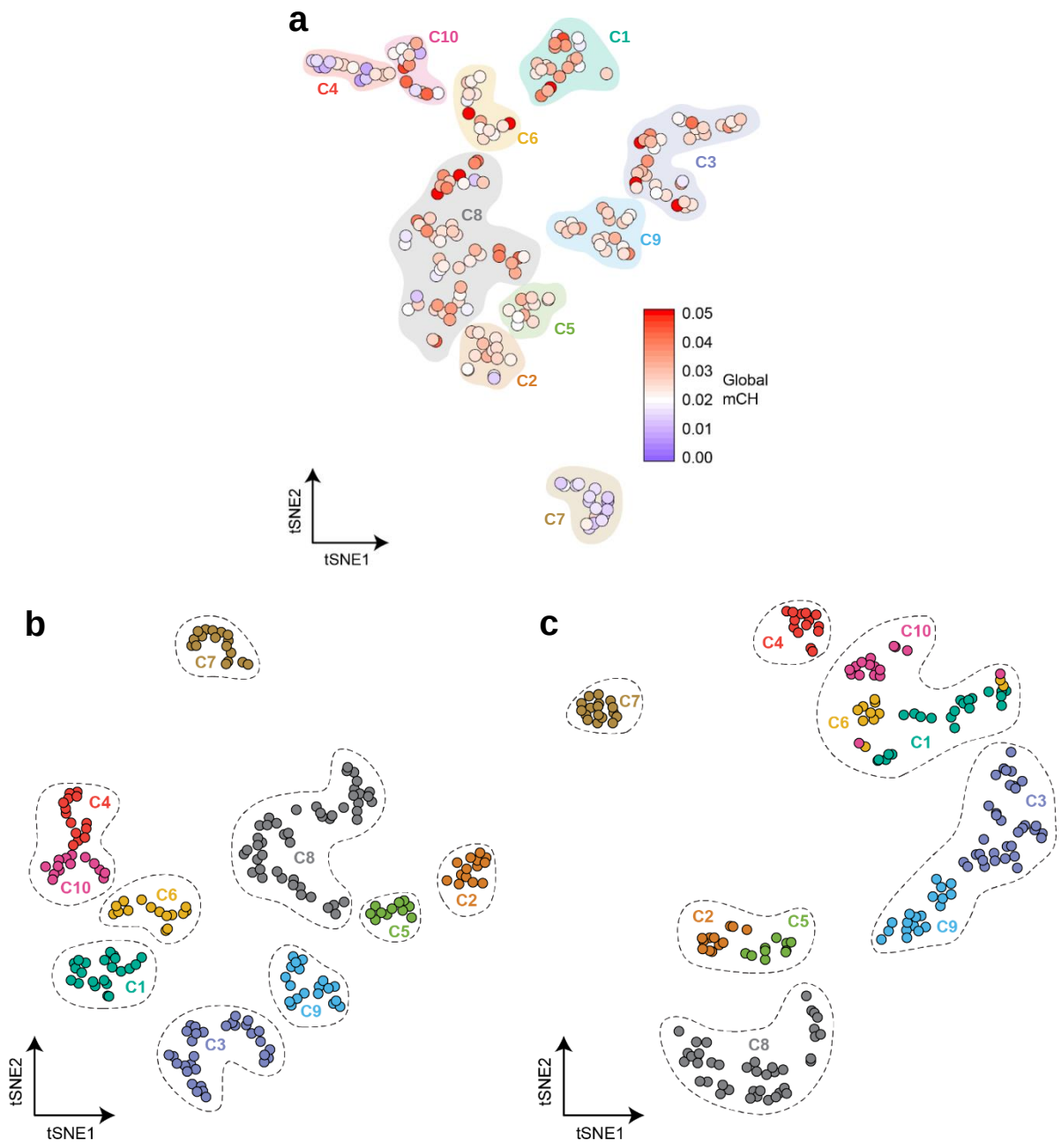
Supplementary Figure 7 | Pairwise methylation percent across windows with ≥30 CG measurements within the human cell line experiment. All windows with sufficient measurement (left) and top 1,000 most variable across all single cell libraries windows (right). Comparison was to the most highly correlated ENCODE Bulk data set to the merged GM12878 cluster (ENCFF835NTC).



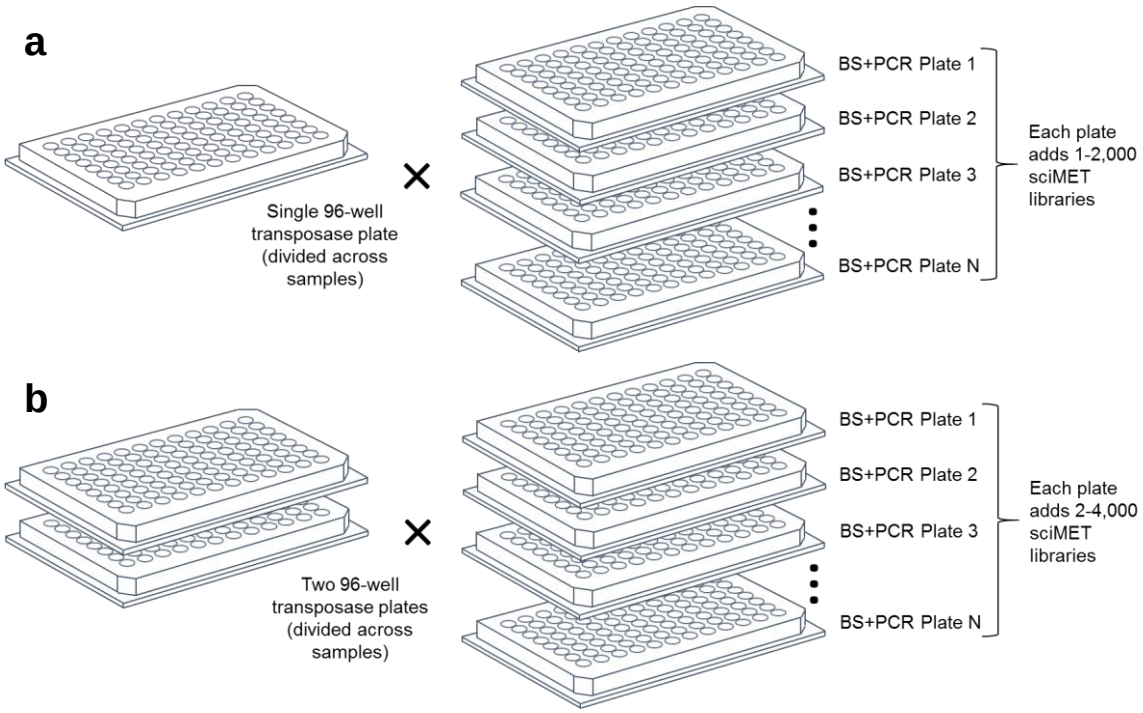
Supplementary Figure 8 | Quality control metrics for mouse cortex libraries (n=285 single-cell libraries). Colors reflect mouse cortex source (n= 64, 99, 122 single-cell libraries, respectively). Median alignment rate across libraries is 56.2%. CG methylation percentage is 66% ± 4.8%. CH methylation percentage is 2.5% ± 0.8%. Libraries with mapping efficiency below 20% were excluded from downstream analysis. Violin plot s are scaled to be equal width, and reflect point density along the x-axis.



Supplementary Figure 9 | **a)** Spacing analysis of adjacent mCH and mCHH marks across mouse cortical sample reads. Neuronal populations demonstrate a previously reported CH methylation periodicity hypothesized to reflect mCH regulation. **b)** Normalized CH methylation relative to distance from Homer-annotated CTCF binding sites. Methylation percentages follow a complex waveform reflecting nucleosome positioning and mCH regulation.

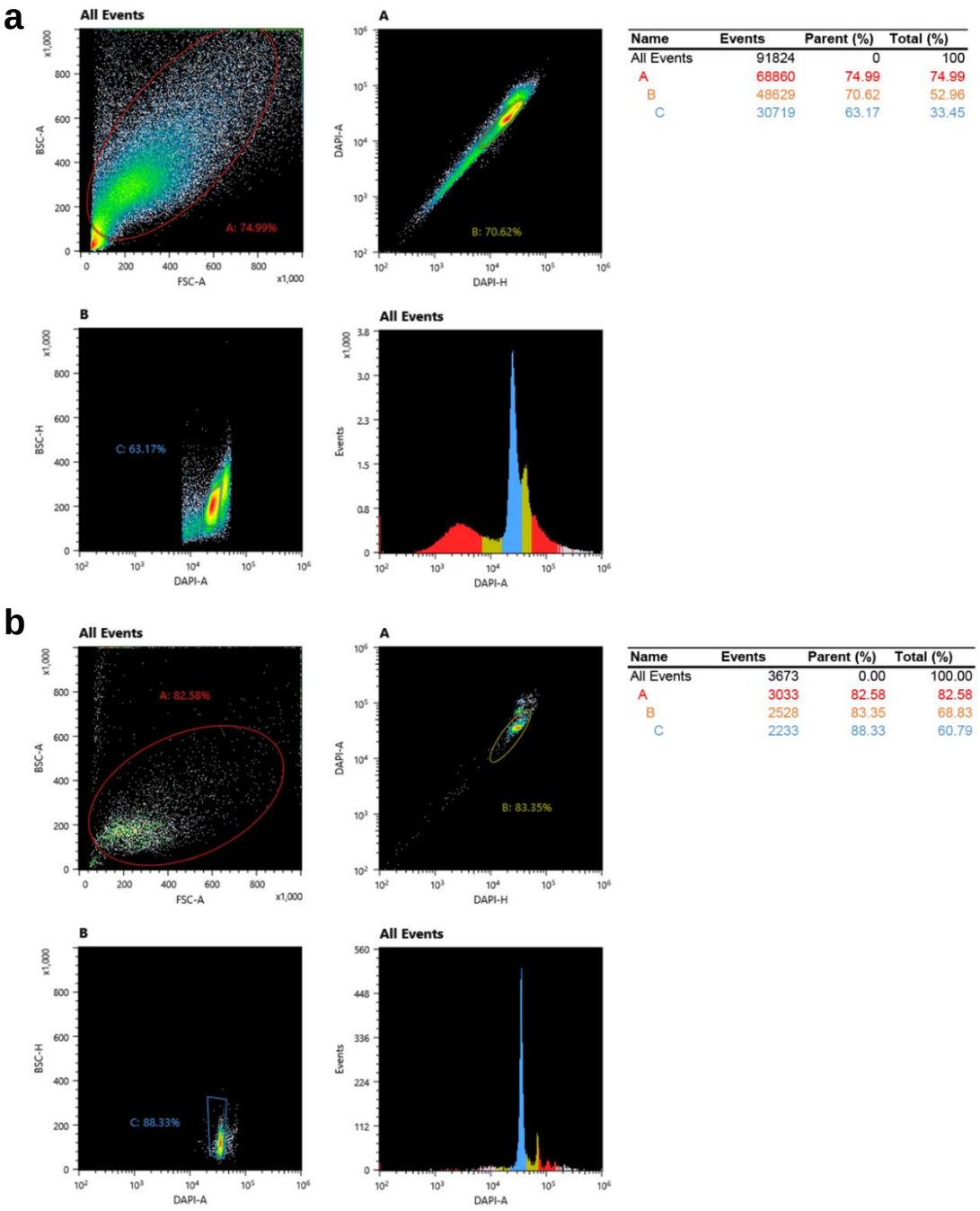


Supplementary Figure 10 | **a.** Global CH methylation levels of individual mouse cortex cells projected into the combined tSNE coordinates as in Figure 2d. **b.** CG methylation of Ensembl regulatory build sites. Cells are colored according to their final cluster identity (as in Figure 2d). Dashed lines represent clusters that were identified using DBSCAN on the CH tSNE only. **c.** CH tSNE and Clustering of cells on 100 kbp windows. Coloring and dashed lines are as in panel **b**.

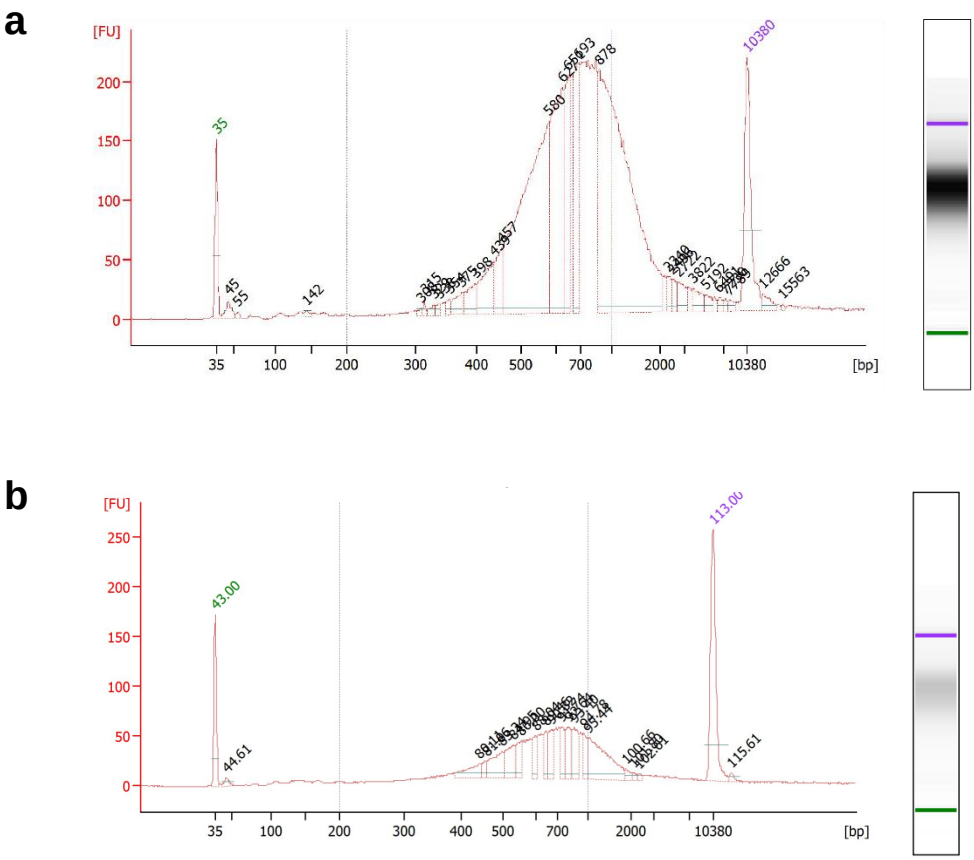


Supplementary Figure 11 | Scalability of sciMET platform. **a.** Scalability of sciMET using a single initial 96-well plate of indexed transposomes. **b.** Scalability of sciMET when increasing the number of initial indexes which increases libraries produced for each subsequent plate. The expected cell count can be represented as:

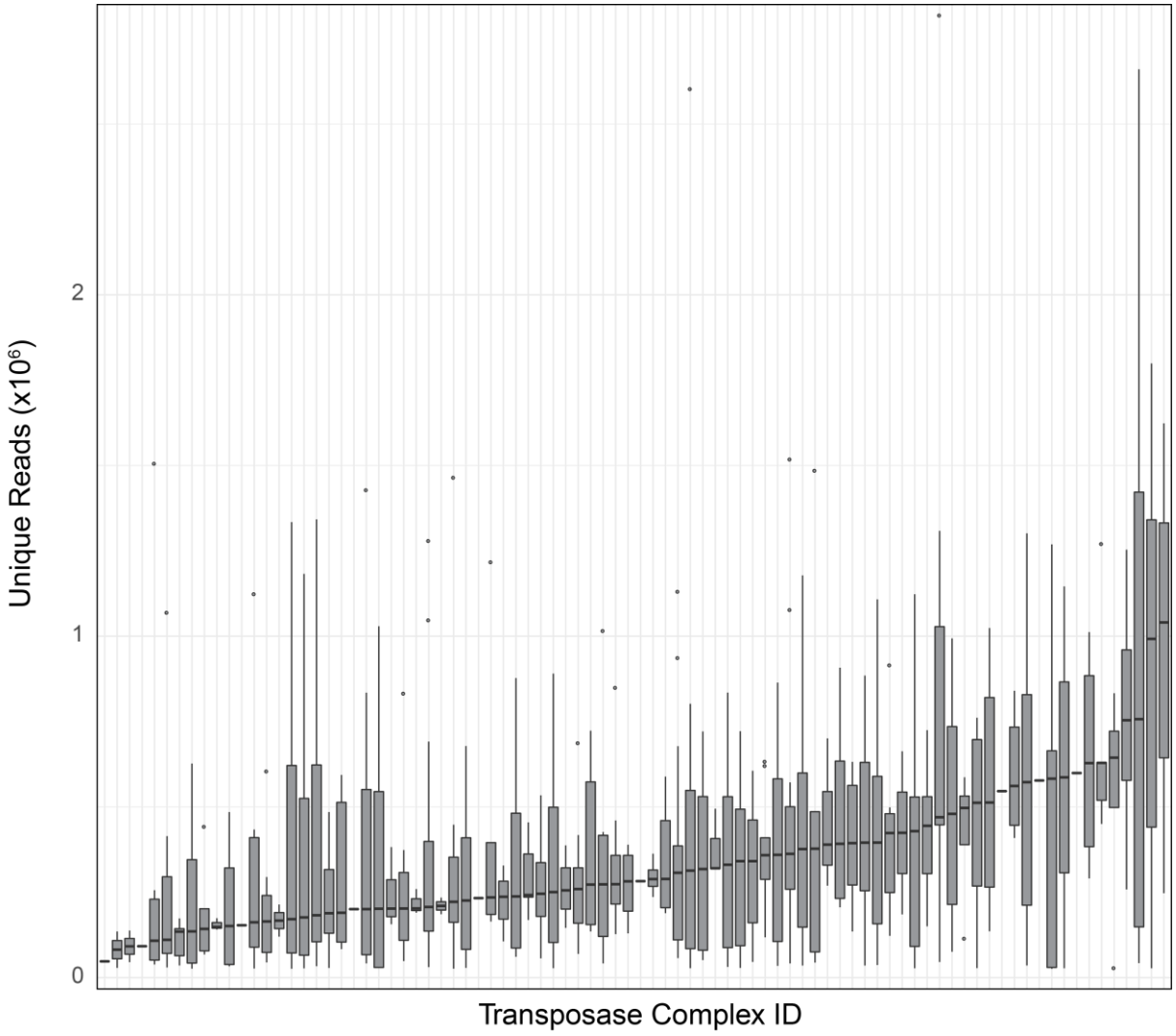
$$Cells_{expected} = \left(0.2 \times Wells_{transposase\ (min=96)}\right) \times Wells_{BS+PCR}$$



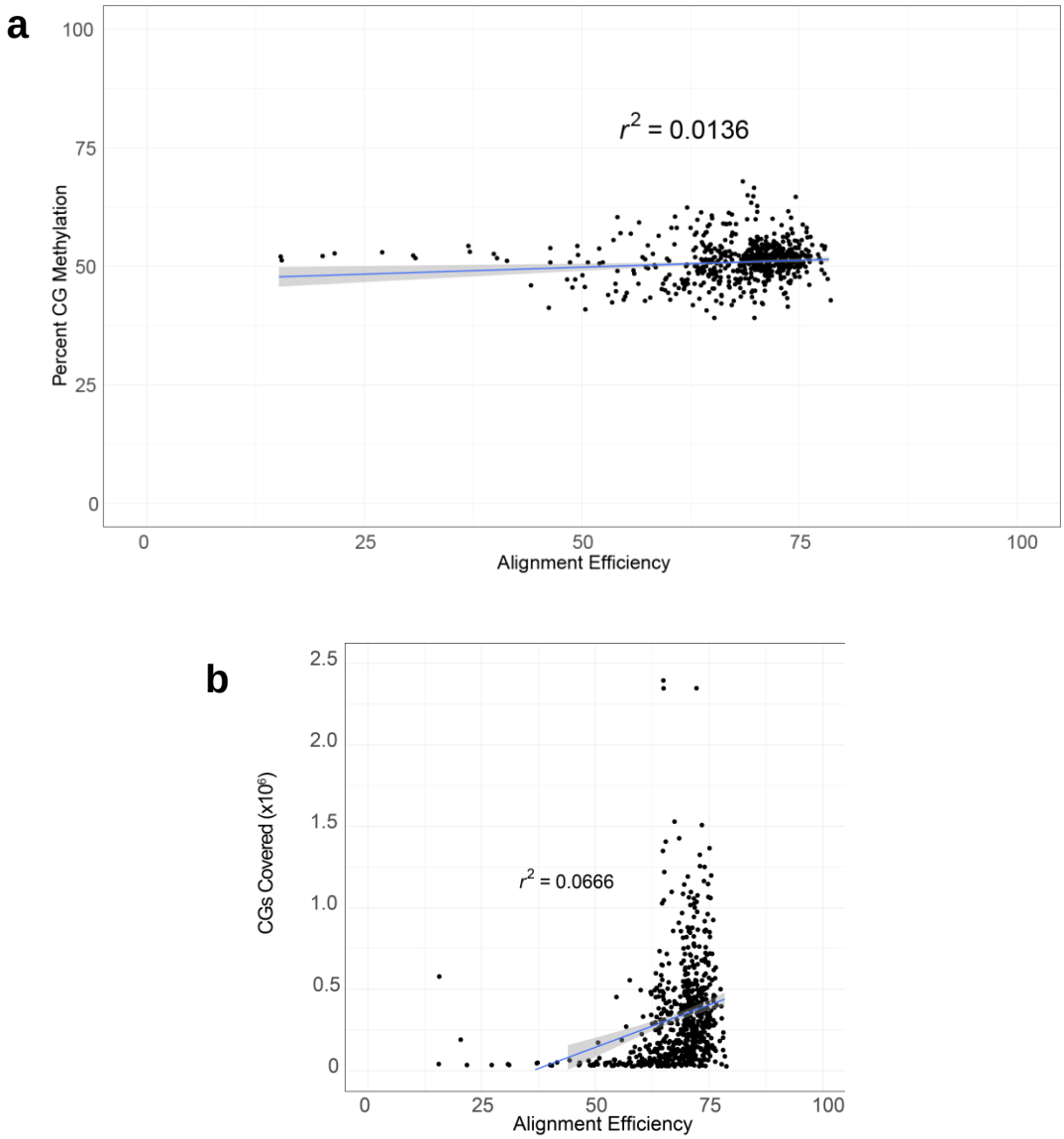
Supplementary Figure 12 | Gating strategy used for single-cell sorting. Data provided are from the human and mouse split library preparation with xSDS treatment. This is an exemplar case of one of the five total sci-MET library preparations. a) Each preparation of a sci-MET library consisted of two FANS procedures which followed the same strategy to isolate pure single nuclei. First, events were gated on forward scatter-area by back scatter-area to remove small debris (Gate A). Using that defined gate, events were then assessed by gating on DAPI staining of height by area. Singlet nuclei are expected to populate the diagonal (Gate B). Gate B was then purified to singlet nuclei by gating on DAPI-area by backscatter-height. Multiple sub-populations of events formed reflecting nuclei doublets and triplets. Gate C defined singlet nuclei and was used for the final gating for sorting. For the first FANS procedure, 2,000 nuclei were placed into each well using the "normal" sorting setting. b) For the second FANS procedure, we placed 22 or 10 nuclei into each well using the "single-cell" sorting setting.



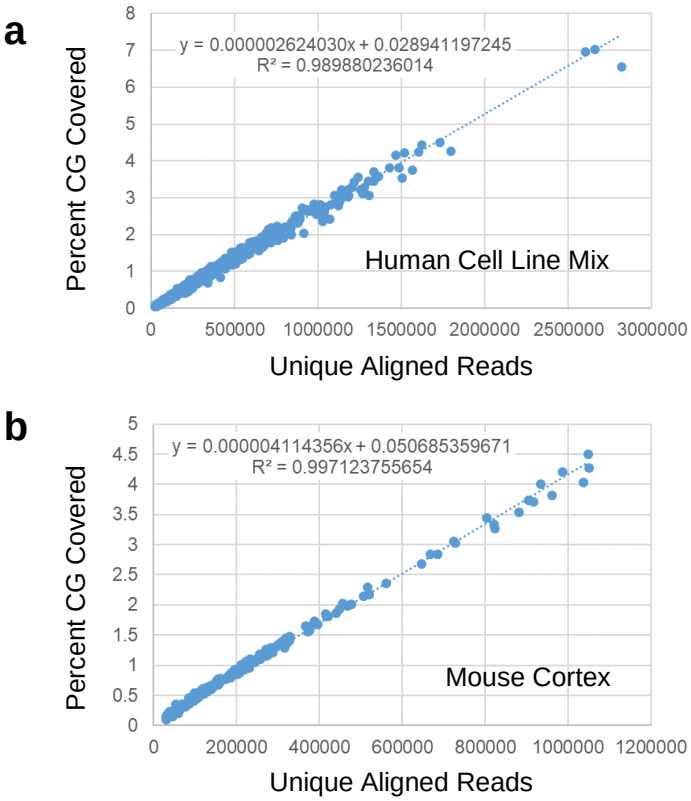
Supplementary Figure 13 | Two example bioanalyzer traces of **a) a** human cell line mix library pool and **b)** a mouse cortex library pool. These two cases are exemplar for the five separate sci-MET experiments performed in total.



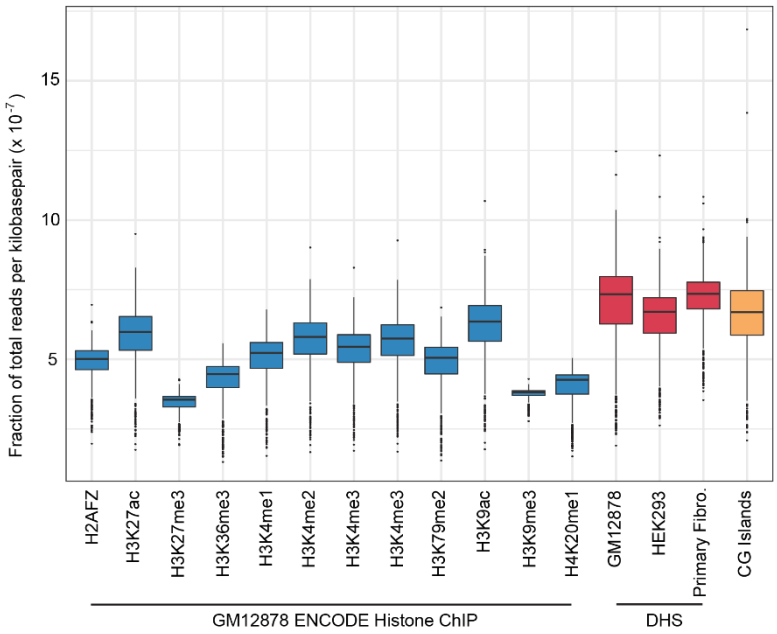
Supplementary Figure 14 | Breakdown of uniquely aligned reads by transposase complex for the human cell line mix experiment. Box plots reflect the distribution of unique aligned reads assigned to a transposase complex in the human cell type mix experiment for cell libraries containing a minimum of 30,000 unique aligned reads (n=86 transposase complexes). Median and variance illustrate that some complexes outperform others by number of reads per cell assigned and consistency across cell libraries. Box plots describes median, with hinges for lower 25% quantile and upper 75% quantile, whiskers describe smallest observation greater than or equal to lower hinge – 1.5 x the interquartile region, and the largest observation less than or equal to upper hinge + 1.5 x the interquartile region.



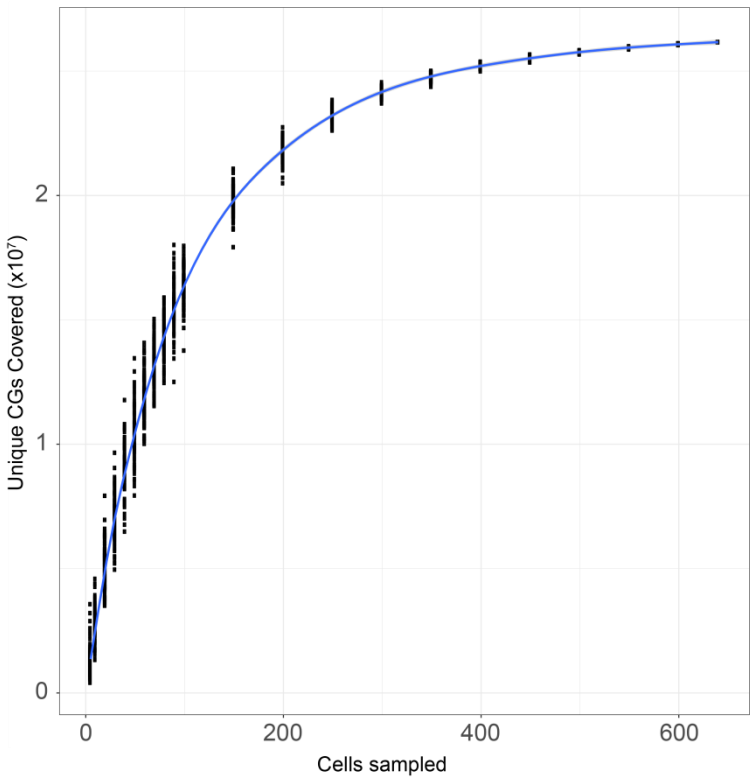
Supplementary Figure 15 | Low correlation is observed between alignment efficiency and percent CG methylation (**a**) or number of CG sites covered (**b**). Values represent human cell type mix experiment for cell libraries containing a minimum of 30,000 unique aligned reads (n=617 single-cell libraries).



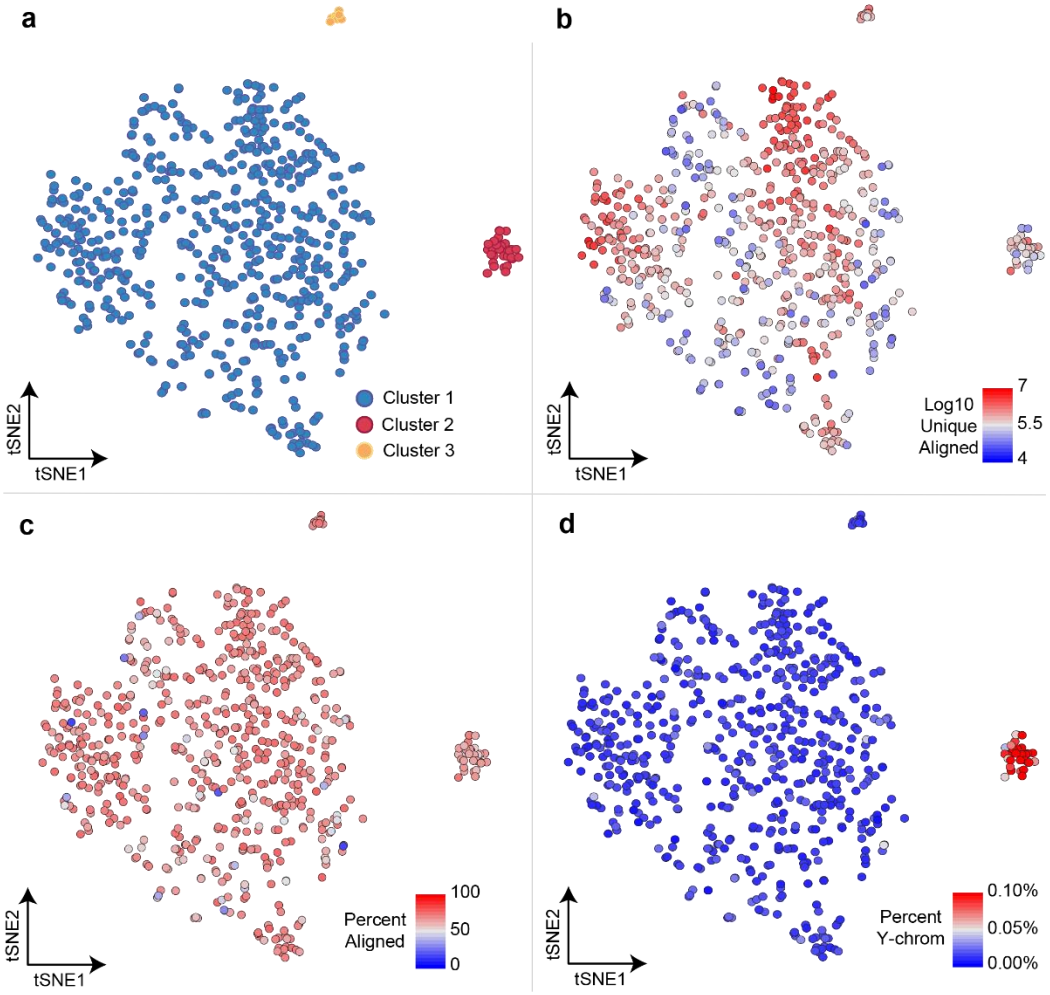
Supplementary Figure 16 | A high correlation is observed between the number of uniquely aligned reads per cell and the percent of CG sites covered for both the human cell line mixing experiment (n=641 single-cell libraries) (**a**) and the mouse cortex experiment (n=285 single-cell libraries) (**b**), indicating a broadly uniform distribution of library construction throughout the genome.



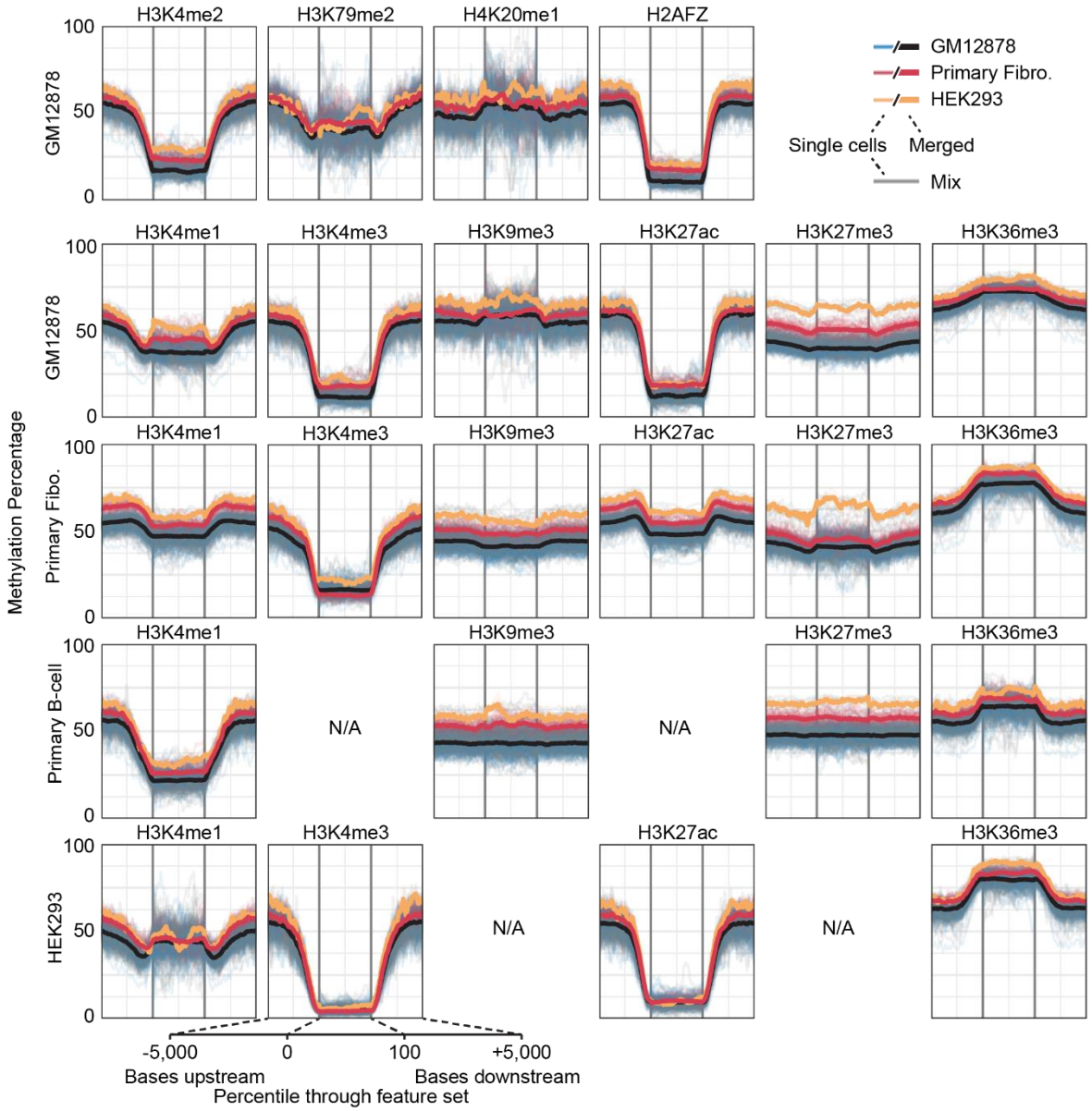
Supplementary Figure 17 | Bias of coverage for sci-MET libraries constructed for the cell type mix experiment. Each cell had its read count within feature windows normalized to the total count for the cell and the total combined size of the features. DHS = DNase-seq hypersensitivity. All ChIP-seq and DHS feature sets were obtained from the ENCODE data portal. Box plots describe median, with hinges for lower 25% quantile and upper 75% quantile, whiskers describe smallest observation greater than or equal to lower hinge – 1.5 x the interquartile region, and the largest observation less than or equal to upper hinge + 1.5 x the interquartile region.



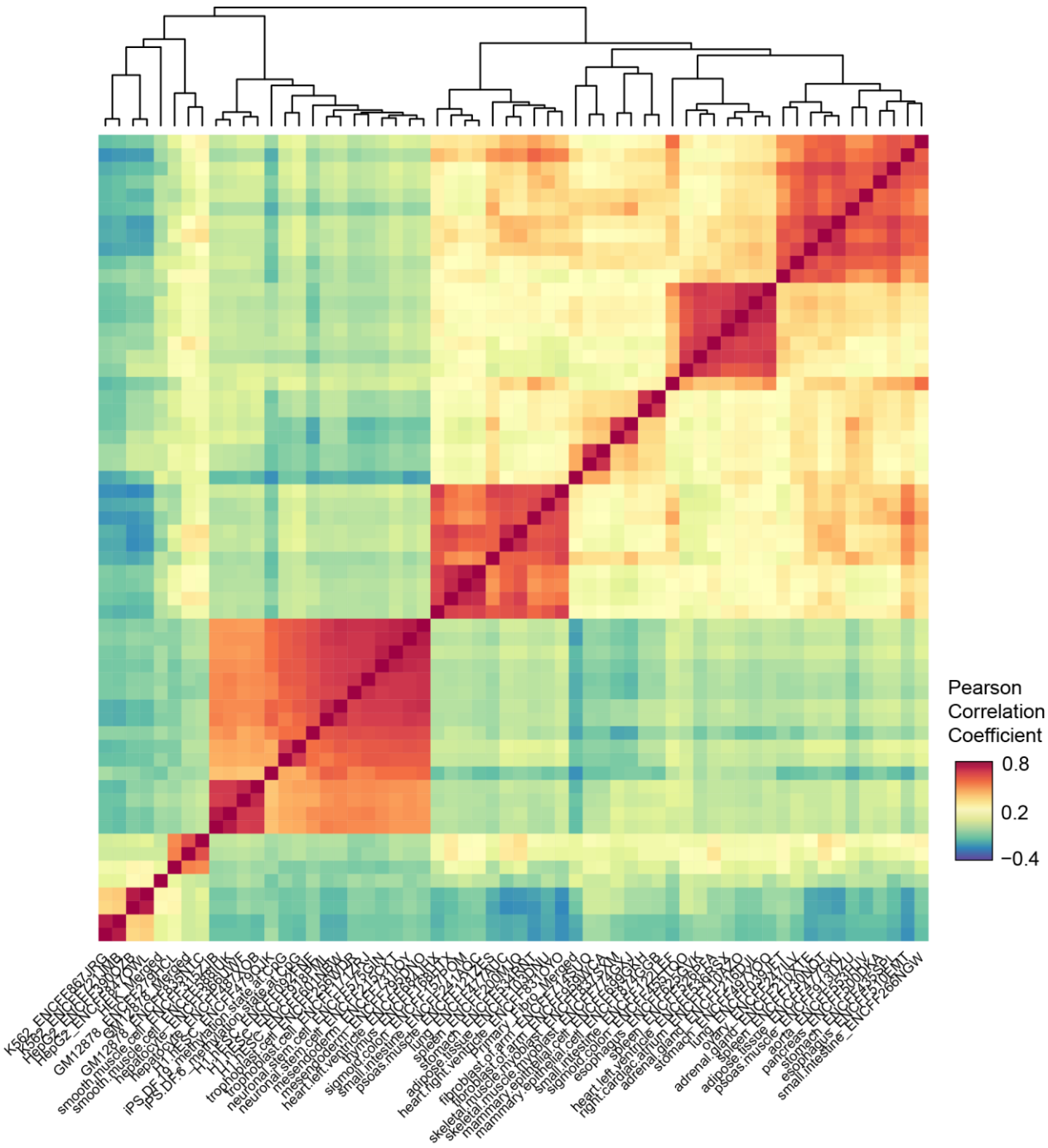
Supplementary Figure 18 | Total unique haploid CG dinucleotides covered as a function of the number of sciMET single cell libraries merged. 100 random subsets of cells were sampled at each increment from the human cell type mix experiment for cell libraries containing a minimum of 30,000 unique aligned reads.



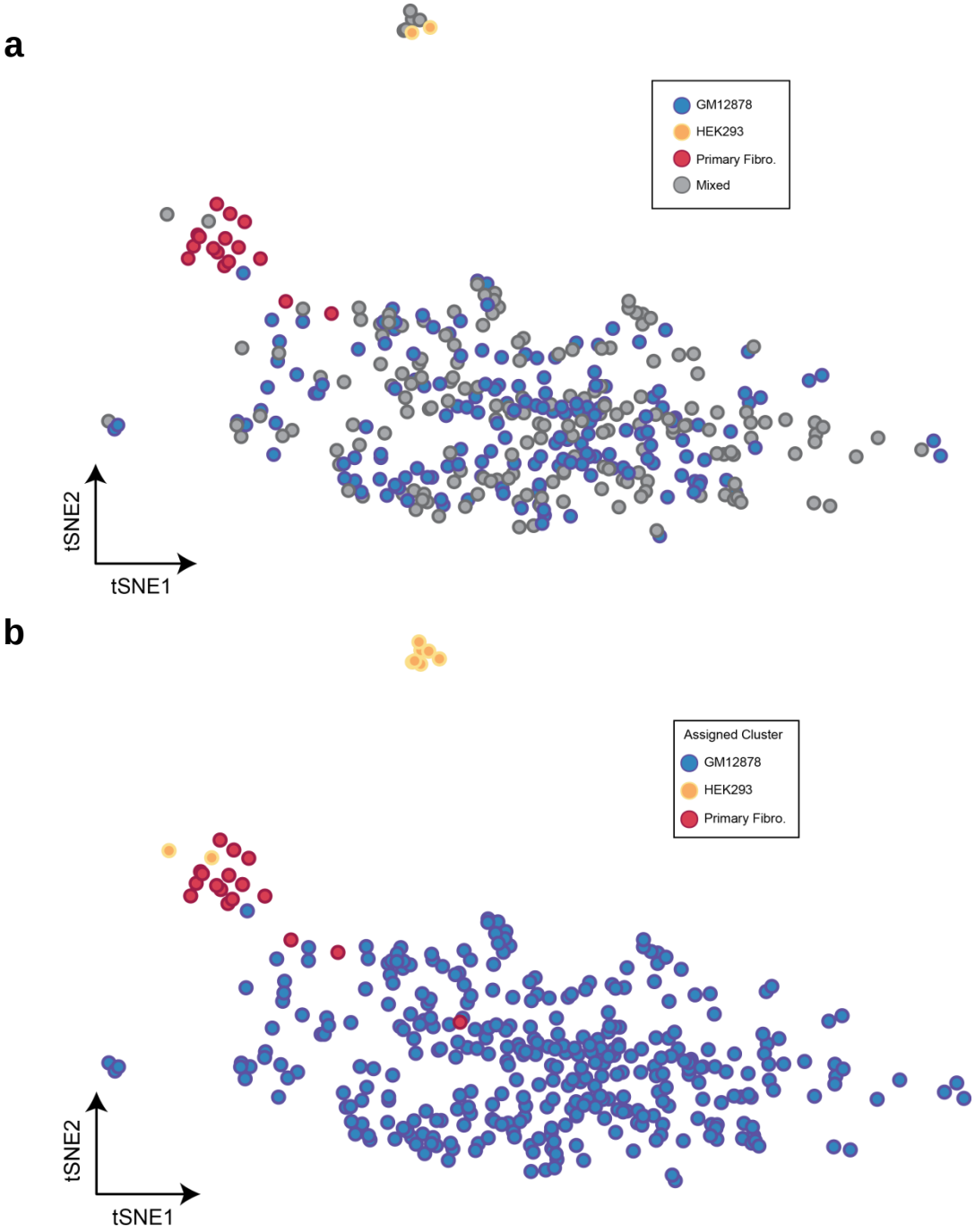
Supplementary Figure 19 | Clustering and library metrics. **a.** After NMF-tSNE, cell coordinates were used to identify clusters using the density-based clustering algorithm dbscan with an epsilon value of 4 and a minimum of 4 cells per seed. Clusters are colored based on their corresponding cell types. **b.** Log10 unique aligned reads with an alignment quality score ≥ 10. **c.** Percent of reads aligned for each library. **d.** Percent of reads that align to the Y-chromosome. BJ cells (foreskin fibroblast cell line) are the only cells derived from a male and form the far right cluster.



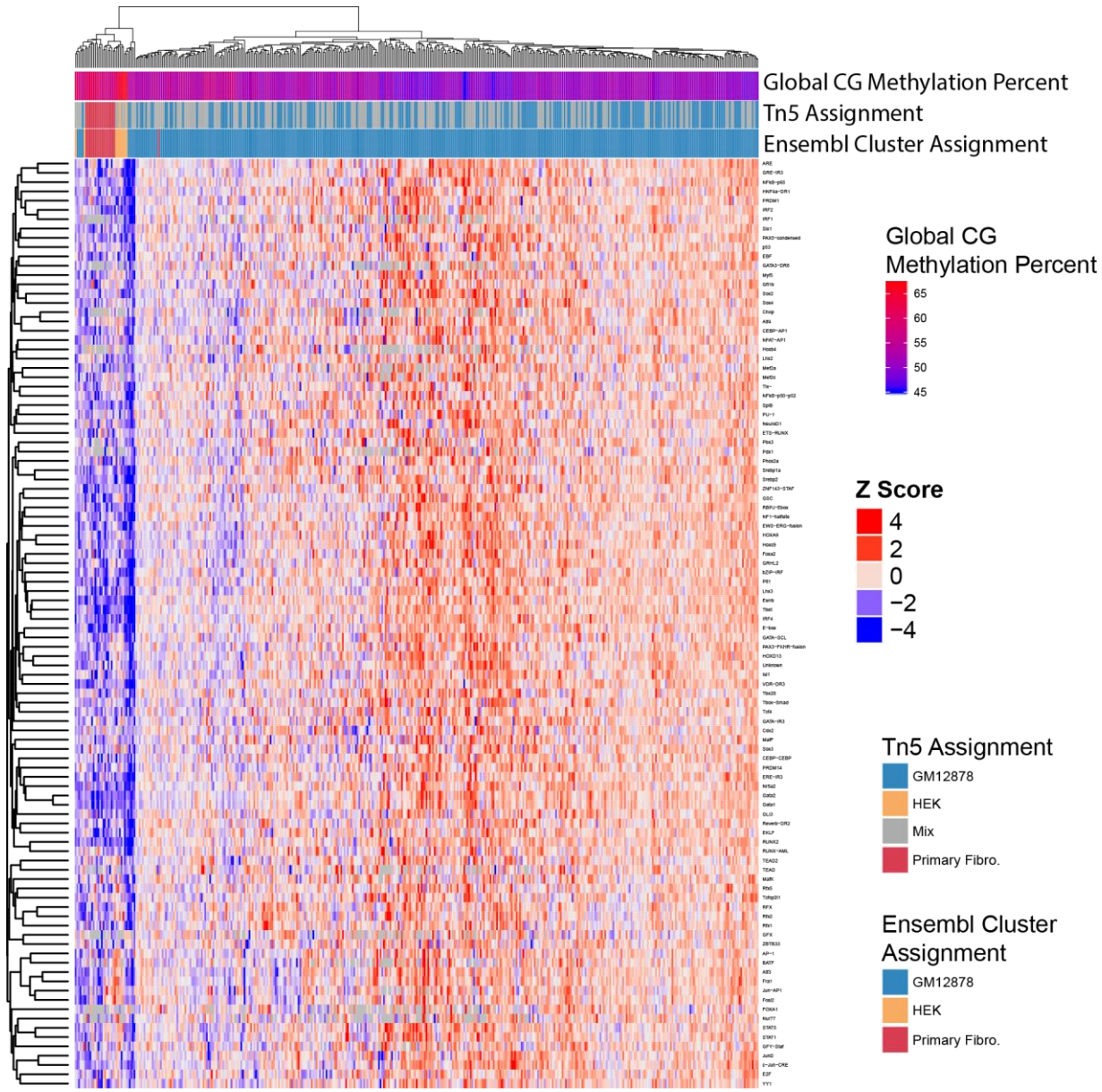
Supplementary Figure 20 | Methylation rate aggregated over ENCODE features identified by ChIP-seq. Methylation rate was calculated for the upstream 5,000 bp of each feature, throughout the feature (defined as percentile of progression through the feature), and 5,000 bp downstream and then aggregated across all features for each single cell (transparent lines) and then averaged across the pure cell types (solid lines).



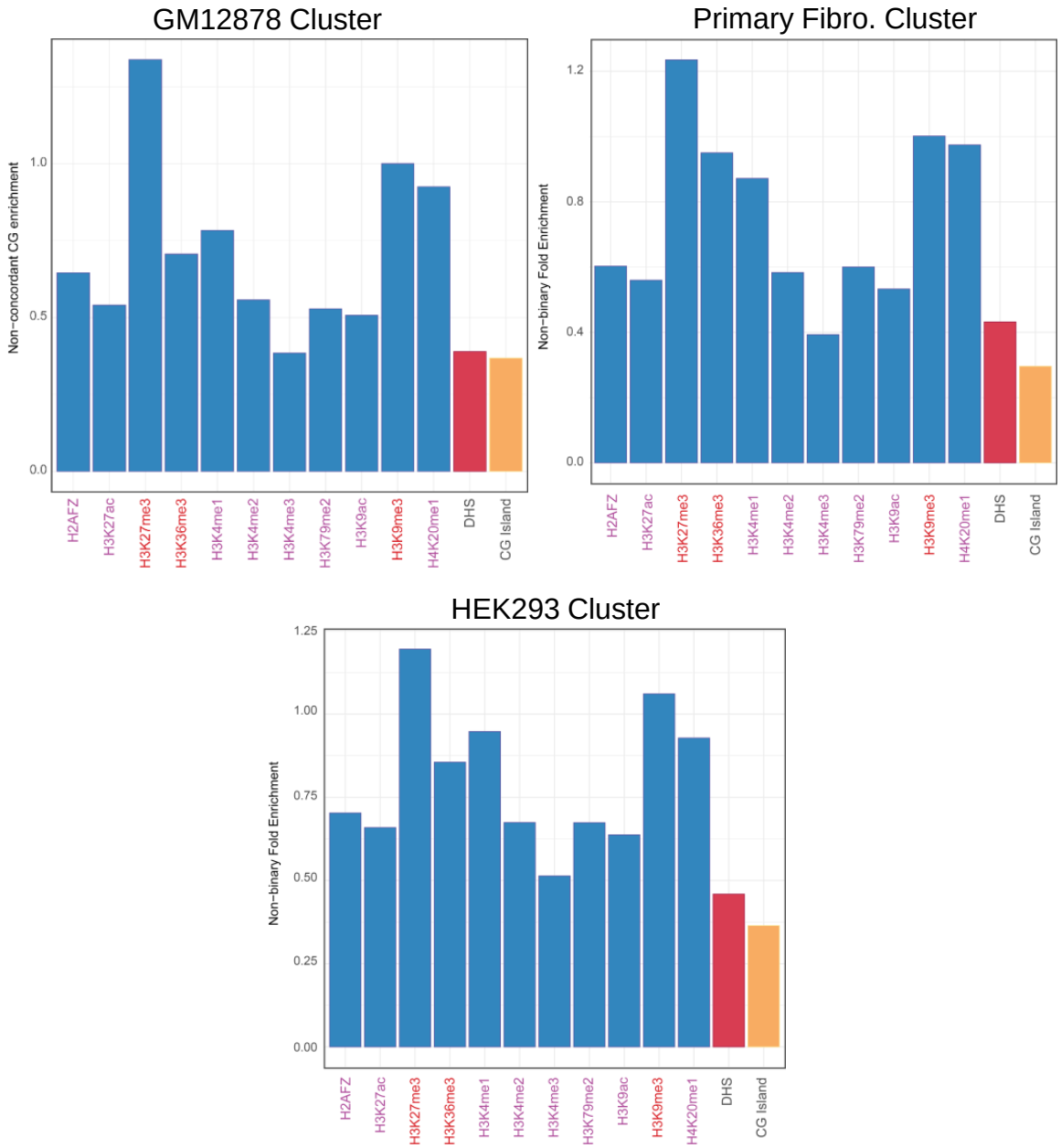
Supplementary Figure 21 | Biclustering of correlation coefficients for publically available WGBS data and the merged sci-MET clusters of the top 1,000 most variable regions in the Ensembl regulatory build. Of note, the GM12878 sci-MET cluster ("GM12878_Merged") is in a clade with two GM12878 WGBS samples, the HEK293 sci-MET cluster ("HEK_Merged") is in a clade with HepG2 and K562 cell lines (there is no HEK293 public dataset used), and the Primary Fibroblast sci-MET cluster ("Primary_Fibro_Merged") is in a clade with two primary forearm fibroblast WGBS samples.



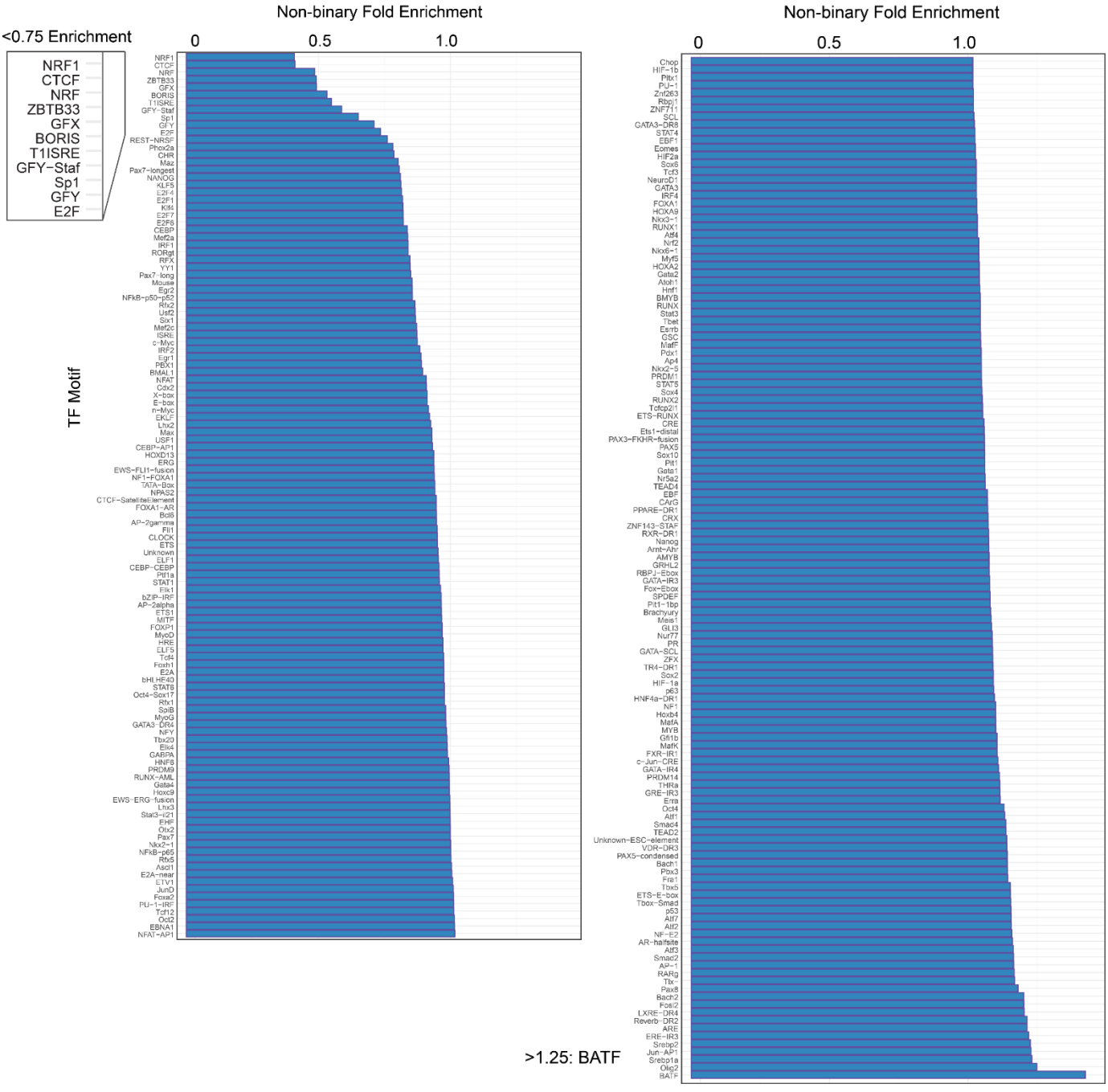
Supplementary Figure 22 | tSNE projection of cells exceeding 30,000 unique reads (n=617 single-cell libraries) from the human cell type mixing experiment based on aggregate methylation over a set of transcription factor motifs. All overlapping CG measurements were collapsed per motif, percent methylation was then Z-score normalized and plotted via tSNE. **a)** Color reflects retained cell line identity. **b)** Color reflects cluster assignment from Ensembl Regulatory Build clustering strategy.



Supplementary Figure 23 | Hierarchical clustering of cells exceeding 30,000 unique reads (n=617 single-cell libraries) from the human cell line mixing experiment (x axis) based on aggregate methylation over a set of transcription factor motifs (y axis; Z scored). All overlapping CG measurements were collapsed per motif, and CG methylation percentage was used for biclustering.



Supplementary Figure 24 | Fold enrichment of non-binary CG methylation status over functional marks (ENCODE) for cells assigned to respective clusters. Enrichment is observed for the repressive mark, H3K27me3, indicating high cell-cell variability across all cell types. In contrast, activating regulatory elements are depleted for non-binary CGs, indicating tighter control and less cell-cell variability.



Supplementary Figure 25 | Fold enrichment of non-binary CG methylation status over transcription factor binding motifs for cells in the GM12878 cluster. NRF1, NRF, and CTCF have the lowest levels of enrichment, indicating tighter control and reduced cell-cell variability of DNA methylation. In contrast BATF and AP1-related binding sites tend to be enriched for nonbinary CG sites, indicating greater cell-cell variability.