**Extended Data Fig. 1: Intact native mono-nucleosomes obtained by hydroxyapatite (HAP) and size-selective purification.**

a, After the HAP purification of MNase-treated chromatin, flow-through and elution samples were run in 2% agarose gel. **b**, Mono-nucleosomes were selected through further size-selective purification of HAP elution. Each purification step and the quality of final product was validated by running the samples in 2% agarose gel (**c**), SDS-PAGE gel showing only four histones without other proteins (**d**), and western blot for histone PTMs (**e**).
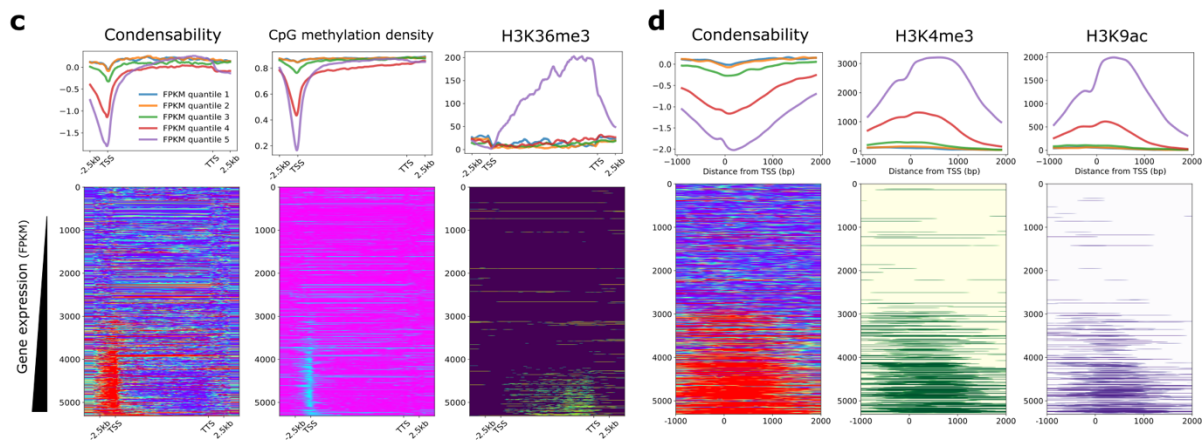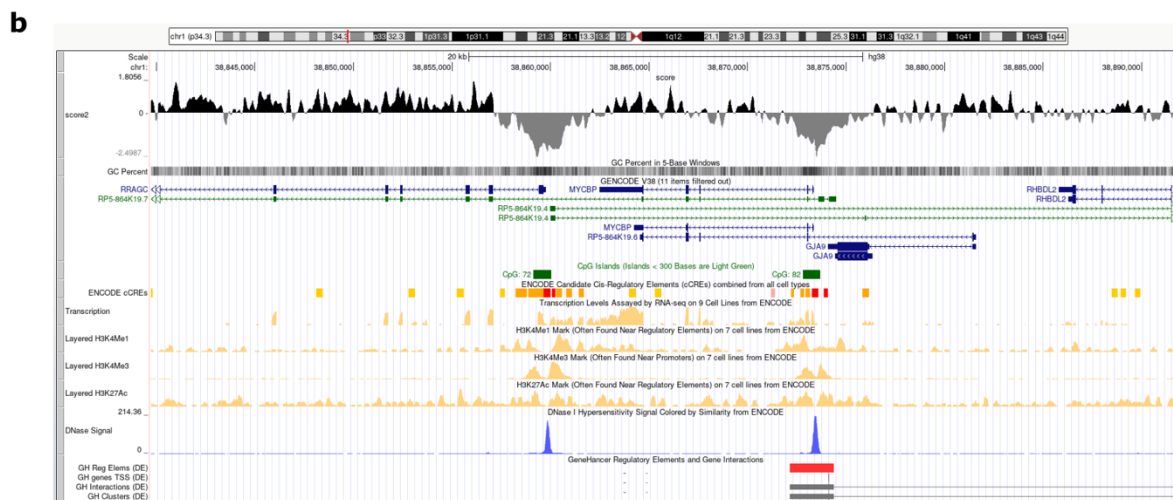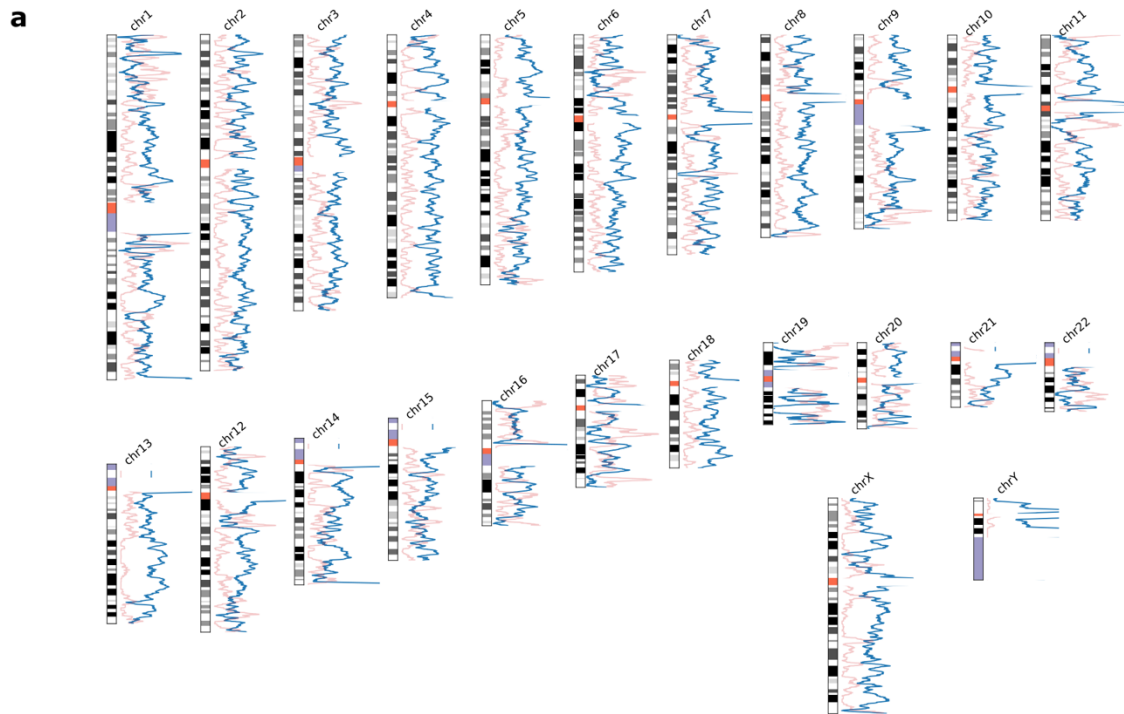
10
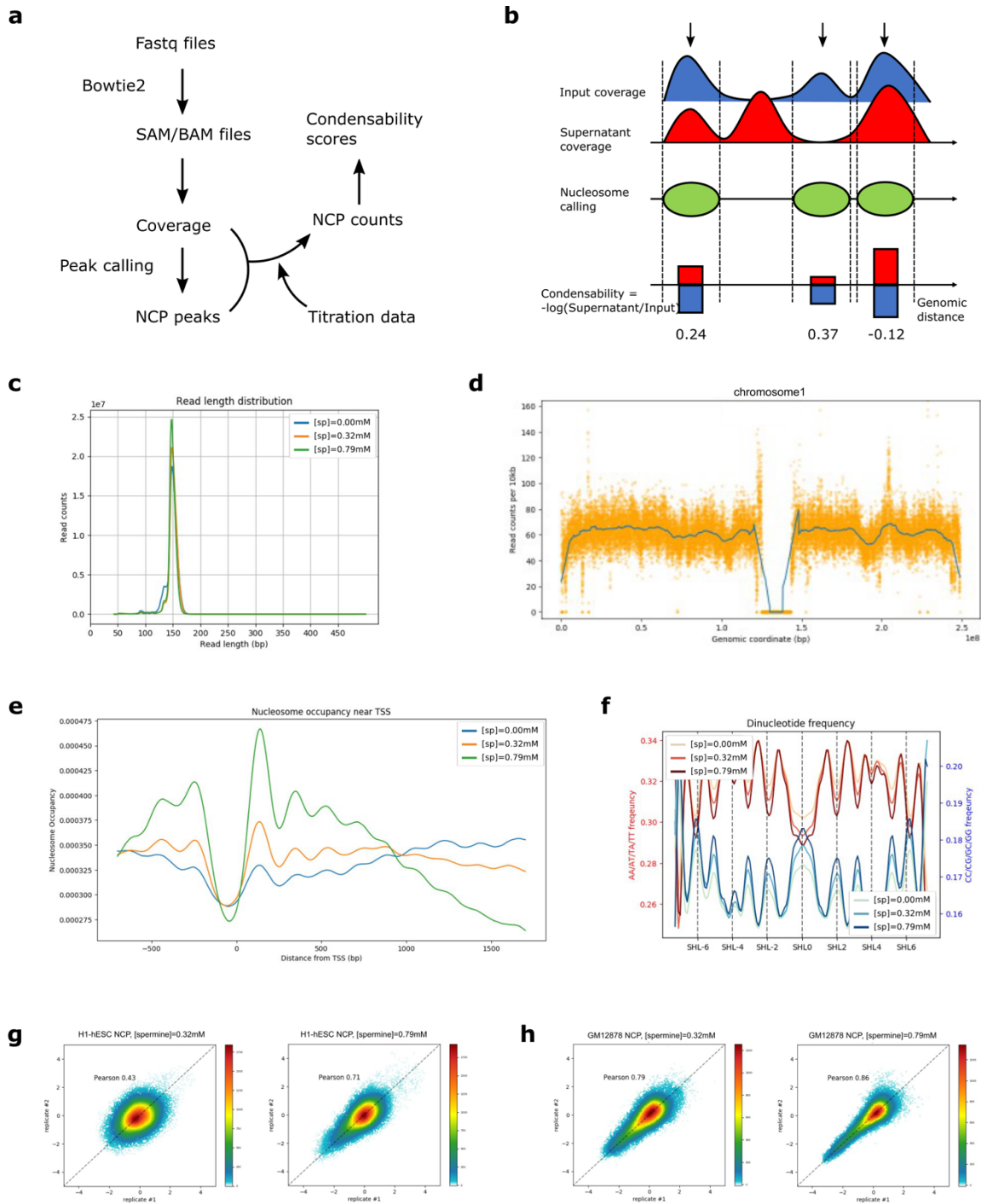
Extended Data Fig. 2: Computational pipeline and data quality controls for condense-seq.

a,**b**, The pipeline of Condense-seq analysis is composed of (i) reads alignment by Bowtie2, (ii) coverage calculations, (iii) mono-nucleosome peak calling for each local maximum of input coverage, (iv) absolute nucleosome count estimation using coverage area and soluble fraction changes, and (v) compute condensability score as minus log of soluble fraction after condensation for each nucleosome. For the quality control of mono-nucleosome samples, we checked that the length distribution of nucleosomal DNA is mostly at 150 bp (**c**), reads are covered evenly along the genome (**d**), nucleosome positioning shows the signature properties such as the nucleosome free region (NFR) near the transcription start site (**e**), and the

19   periodicity of AT-rich versus GC-rich dinucleotides in nucleosomal DNAs (**f**). Correlation between replicates
20   of Condense-seq measurements on H1-hESC (**g**) and GM12878 (**h**) are shown for condensabilities
21   determined for 0.32 mM spermine and 0.79 mM spermine.

22

a Genome-wide condensability profiles across chromosomes chr1–chrY.

b Genome browser view of chr1 region (chr1:38,840,000–38,890,000, hg38) showing score2, GC Percent, GENCODE V38 gene annotations, CpG Islands, ENCODE cCREs, Transcription, Layered H3K4Me1, H3K4Me3, H3K27Ac, DNase Signal, and GeneHancer Regulatory Elements.

c Condensability, CpG methylation density, and H3K36me3 profiles across gene bodies (TSS to TTS, ±2.5kb), grouped by FPKM quantile.

d Condensability, H3K4me3, and H3K9ac profiles around TSS (distance from TSS in bp), ordered by gene expression (FPKM).

23

24   **Extended Data Fig. 3: Condensability scores of H1-hESC across all chromosomes and**
25   **correlations with gene expression and other epigenetic marks near the TSS.**

26   **a**, Comparison between condensability (blue) and transcription level (red) along all chromosomes. **b**,
27   Snapshot of UCSC genome browser for the condensability profile along with many other cis-regulatory
28   elements. **c**,**d**, All genes in chromosome 1 were grouped into five quantiles according to the transcription
29   level (quantile 1 through 5 for increasing transcription). (top) Condensability, meCpG fraction, H3K36me3,
30   H3K4me4 and H3K9ac along the transcription unit coordinate averaged for each quantile. (bottom) Heat
31   maps show the same quantities for each gene, rank ordered with increasing condensability.   TSS
32   (transcription start site), TTS (transcription termination site). Views zoomed around TSS are shown in (**d**) .

33

34
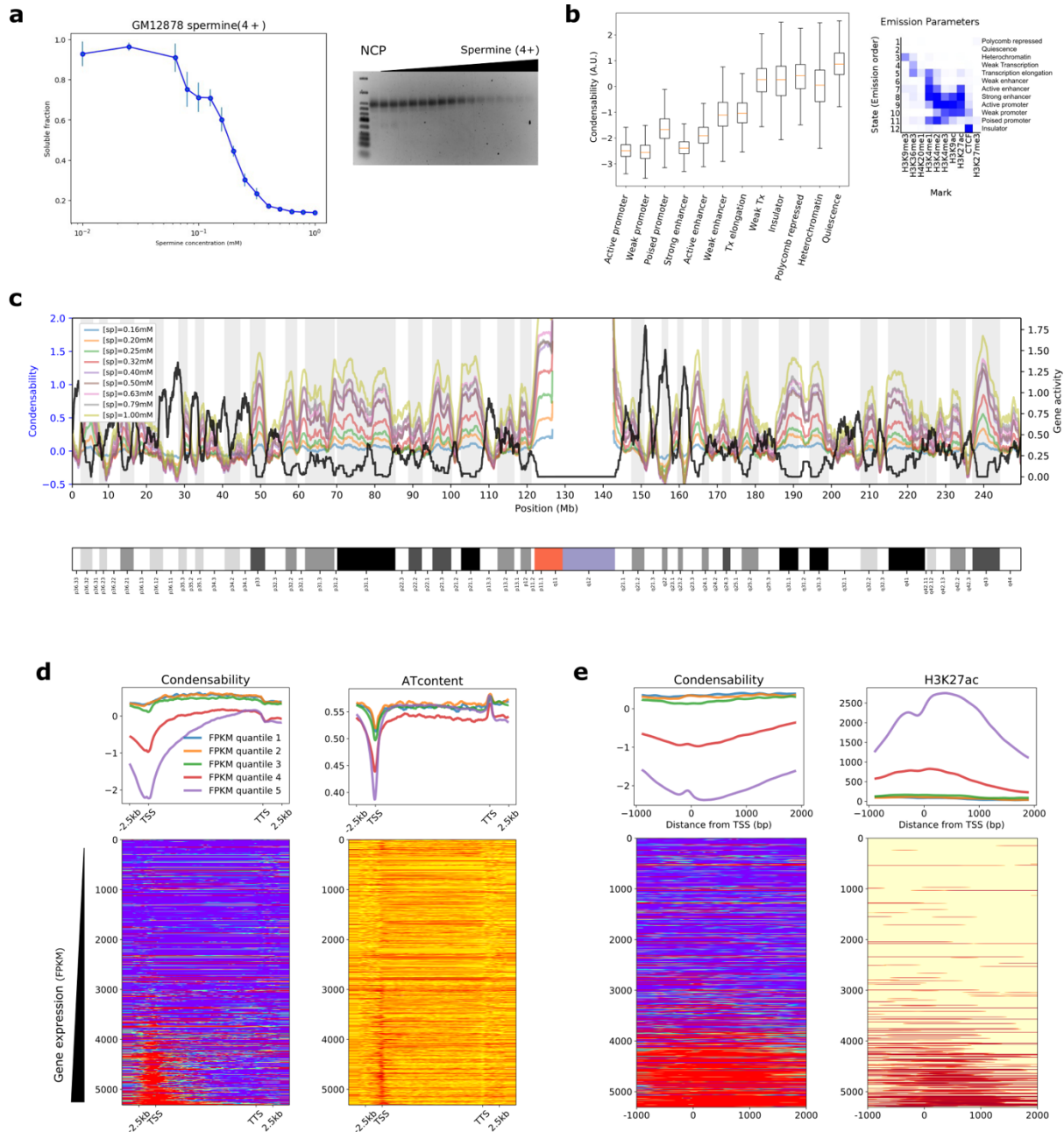
Extended Data Fig. 4: Condense-seq measurement of native mono-nucleosomes from GM12878 cells and the comparison between nucleosome condensability and their chromatin states.

a, Native mono-nucleosomes were purified from GM12878 cell line, and the condensation was induced by adding spermine. Soluble fractions were measured using UV-VIS spectroscopy (left) and ran in the 2% agarose gel (right). **b**, Genomic nucleosomes were grouped according to their ChromHMM chromatin states based on the combination of various PTMs. Their condensabilities are shown in box plot for each chromatin state. **c**, The condensability score was plotted over human chromosome1 for different spermine titration points (colored lines) and compared with the gene expression level (black line). The condensability profiles from TSS to TTS (**d**) near TSS (**e**) for five quantiles based on the gene expression levels in the GM12878 cell line. Also shown at AT content and H3K27ac profiles.

45

46
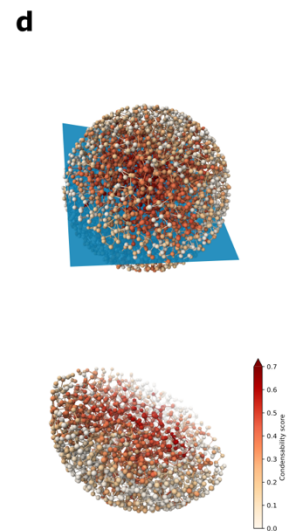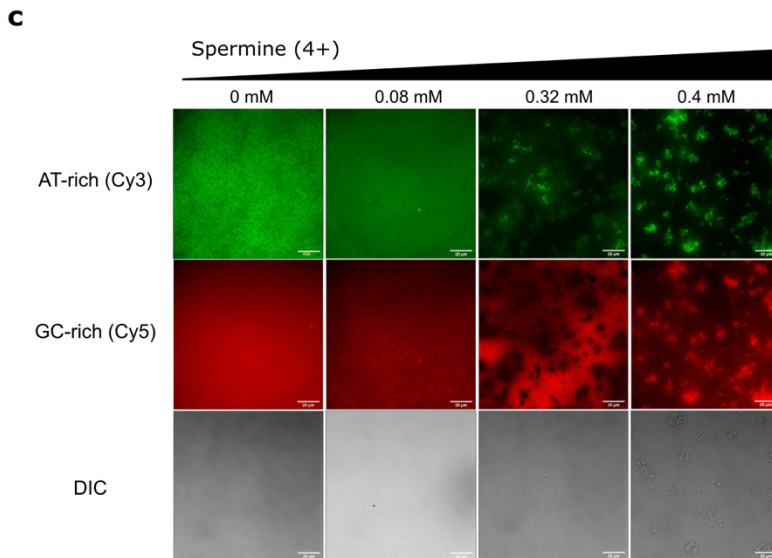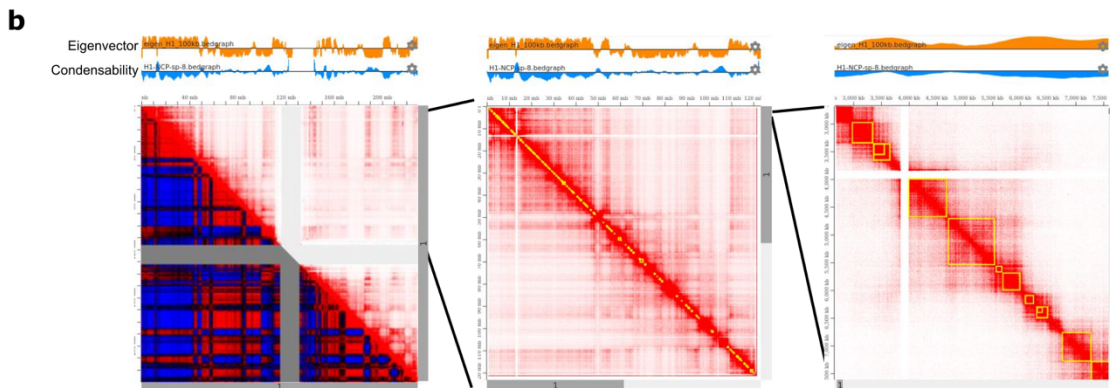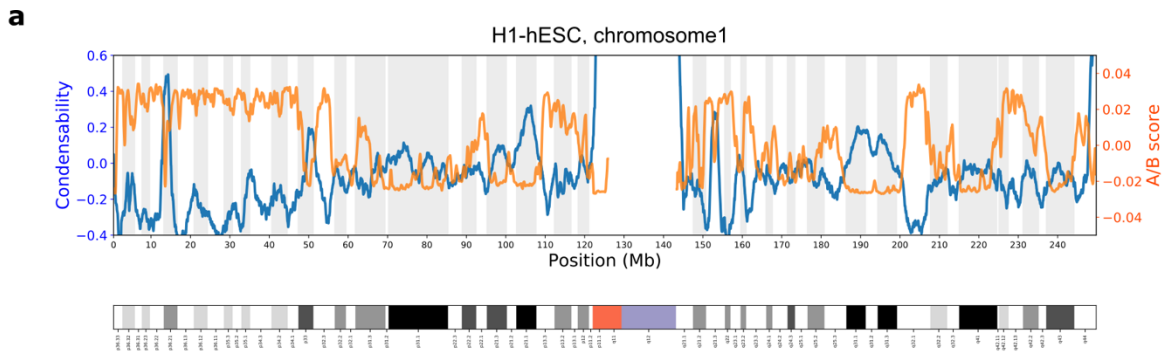
Extended Data Fig. 5: Spatial separation of molecules promoted by condensability difference to compartmentalize the genome.

H1-hESC condensability (blue) and A/B compartment scores based on Micro-C data (orange) in mega base-pair resolution of chromosome 1 (**a**) and finer resolution (**b**). **c**, PCR amplified AT-rich (Cy3 labeled) and GC-rich (Cy5 labeled) DNAs were equally mixed and condensed in various spermine concentrations. For each condition, DNA condensates were imaged using wide-field microscope. As spermine concentration increased, AT-rich DNAs formed a condensed core first, and GC-rich DNAs condensed over

54   the AT-rich core at higher spermine concentrations, promoting the spatial separation between AT-rich
55   versus GC-rich condensate. **d**, Chromosome polymer simulation with condense-seq data using spermine
56   as only input (GM12878, chr12) shows that highly condensable chromatin is compacted into the core and
57   the rest is excluded to generate spatially separate compartments (red = condensability score).

58

a

Condensability (A.U.) vs AT content (%)

b

Chromosome1

Nucleosome Counts vs Condensability (A.U.)

I II III IV V VI VII

c

AA/AT/TA/TT freqency · CC/CG/GC/GG freqency vs Super Helical Location

AT-rich · GC-rich

d

|	1	2	3	4	5	6	7	Biological functions
CHG density								pluripotency, neural development
CHH density								pluripotency, neural development
CpG density								gene silencing
H2AFZ								transcriptional activations
H2AK5ac								transcriptional activations
H2BK120ac								transcriptional activations
H2BK12ac								transcriptional activations
H2BK15ac								transcriptional activations
H2BK20ac								transcriptional activations
H2BK5ac								transcriptional activations
H3K14ac								transcriptional activations (promoters)
H3K18ac								transcriptional activations
H3K23ac								transcriptional activations
H3K23me2								gene silencing
H3K27ac								transcriptional activations (active/poised enhancers)
H3K27me3								gene silencing (polycomb repression, poised promoters)
H3K36me3								transcriptional elongations (gene bodies)
H3K4ac								transcriptional activations (promoters)
H3K4me1								transcriptional activations (primed enhancers)
H3K4me2								transcriptional activations
H3K4me3								transcriptional activations (promoters)
H3K56ac								transcriptional activations, DNA replication and repair
H3K79me1								transcriptional activations, DNA damage response
H3K79me2								transcriptional activations, DNA damage response
H3K9ac								transcriptional activations (active/poised enhancers)
H3K9me3								gene silencing (pericentric heterochromatin)
H4K20me1								transcriptional activations
H4K5ac								transcriptional activations
H4K8ac								transcriptional activations
H4K91ac								transcriptional activations

Z-score 1.5 / -1.5

e

UMAP: metric=cosine, n_neighbors=15, min...

f

Condensability by property class

Condensability (A.U.)

	1	2	3	4	5	6	7	8	9	10
AT content	0.0	0.0	0.0	0.0	0.0	0.02	0.0	0.71	1.09	0.0
poly-G/C length	0.05	0.05	0.02	0.24	0.01	0.0	0.28	0.0	0.13	0.09
meCpG density	0.0	0.0	0.0	0.0	0.0	0.0	0.56	1.41	0.0	0.0
meCHG density	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.03	0.04	0.19
meCHH density	0.0	0.01	0.0	0.01	0.0	0.0	0.02	0.01	0.03	0.06
H2AFZ	0.0	0.0	0.0	0.0	0.0	0.25	0.0	0.0	0.0	0.0
H2AK5ac	0.0	0.32	0.0	0.0	0.01	0.0	0.0	0.0	0.0	0.0
H2BK120ac	0.0	0.47	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H2BK12ac	0.0	0.65	0.02	0.09	0.01	0.0	0.0	0.0	0.0	0.0
H2BK15ac	0.0	0.38	0.01	0.06	0.0	0.0	0.0	0.0	0.0	0.0
H2BK20ac	0.0	0.61	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H2BK5ac	0.02	0.21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K14ac	0.0	0.04	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K18ac	0.05	0.44	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K23ac	0.0	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K23me2	0.06	0.05	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
H3K27ac	0.04	0.09	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K27me3	0.0	0.0	0.0	2.66	0.0	0.0	0.0	0.0	0.0	0.0
H3K36me3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.38
H3K4ac	0.01	0.36	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K4me1	0.0	0.0	0.36	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K4me2	0.4	0.0	0.07	0.43	0.0	0.0	0.0	0.0	0.0	0.0
H3K4me3	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K56ac	0.05	0.14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H3K79me1	0.0	0.0	0.02	0.0	0.0	0.05	0.0	0.0	0.0	0.03
H3K79me2	0.0	0.0	0.0	0.0	0.0	0.45	0.0	0.0	0.0	0.0
H3K9ac	0.07	0.04	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.0
H3K9me3	0.0	0.0	0.0	0.0	0.0	0.0	0.02	0.0	0.01	0.0
H4K20me1	0.0	0.0	0.0	0.01	0.0	0.14	0.0	0.12	0.0	0.1
H4K5ac	0.03	0.15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H4K8ac	0.05	0.08	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.0
H4K91ac	0.01	0.52	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Property class

g

Model prediction (Boosting)

Prediction vs Experiment

h

10-fold cross validation

Pearson correlation

Linear reg · SVR · Boosting · Random Forest · Neural Network

i

Importance of variables (Boosting)

% of importance

AT content, poly-G/C length, meCpG density, meCHG density, meCHH density, H2AFZ, H2AK5ac, H2BK120ac, H2BK12ac, H2BK15ac, H2BK20ac, H2BK5ac, H3K14ac, H3K18ac, H3K23ac, H3K23me2, H3K27ac, H3K27me3, H3K36me3, H3K4ac, H3K4me1, H3K4me2, H3K4me3, H3K56ac, H3K79me1, H3K79me2, H3K9ac, H3K9me3, H4K20me1, H4K5ac, H4K8ac, H4K91ac

Extended Data Fig. 6: Deciphering the genetic and epigenetic determinants of genomic nucleosome condensation.

a, Scatter plot of the condensability of mono-nucleosomes in chromosome 1 and the AT contents of corresponding nucleosomal DNA. **b**, The nucleosome population was partitioned into seven partitions, from low to high condensability. The periodicity of AT-rich versus GC-rich dinucleotides was analyzed. The results for each partition (**c**, left), and its amplitude and phase were computed using Fourier transformation and represented in polar plot (**c**, right, radius: amplitude, angle: phase). **d**, The enrichment analysis of all DNA methylation and histone ChIP-seq data available in ENCODE over different condensability partitions from low to high (1-7 partitions in **b**). **e**, After randomly selecting 1% of the mono-nucleosomes in chromosome 1, their genetic and epigenetic features were projected into two-dimensional scatter plots through UMAP embedding. Each color represents one of the property classes of the NMF analysis shown in panel **f**. **f**, The genetic and epigenetic features of all mono-nucleosomes in chromosome 1 were linearly decomposed into 10 property classes by NMF algorithms. Each property class has a specific combination of features, as shown in the matrix (lower panel). All nucleosomes were clustered into each representative property class with the biggest contributions. After clustering, nucleosome condensabilities were plotted as boxplot for each class (upper panel). **g**–**l**, Multivariate linear regression (linear reg), Supported Vector Machine regression (SVM), gradient boosting regression (Boosting), random forest regression (Random Forest), and neural networks were used to predict nucleosome condensability. All showed similar correlations between experimental values and predictions in a 10-fold cross-validation (**g**,**h**). The importance of genetic–epigenetic features in prediction was computed using the boosting method (**i**).

81
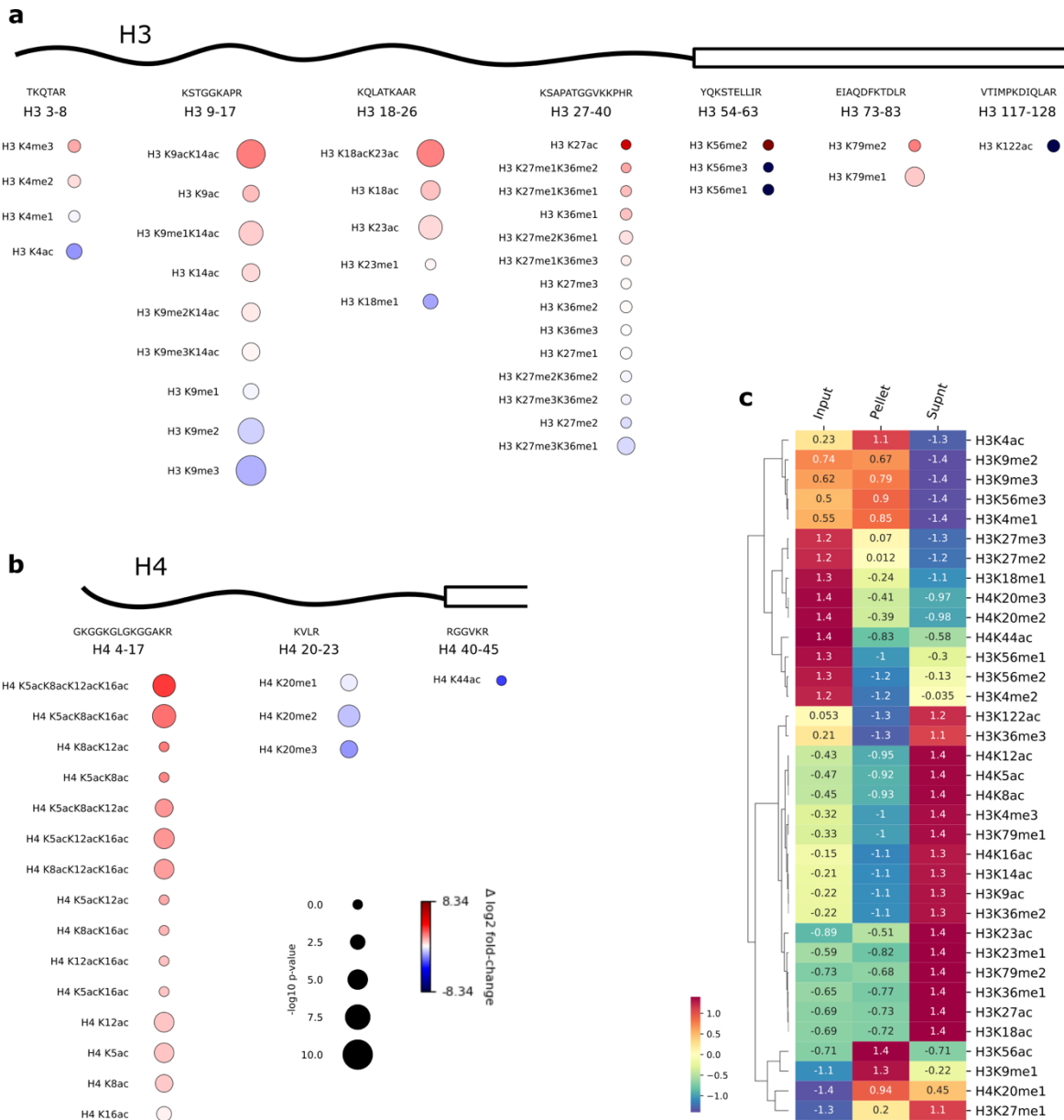
Extended Data Fig. 7: Mass spectrometry identification of histone PTM marks with biased enrichment during native mononucleosome condensation experiments.

a,**b**, Histone PTM marks detected in each histone H3/H4 peptides are shown. Its relative enrichment difference compared with the unmodified peptide is represented by color (red: more enriched in supernatant, blue: more depleted in supernatant) and its signification is represented by the size of bubble (-log p-value). **c**, Combinatorial histone PTM enrichment data was aggregated into single PTM modifications, and the relative enrichment in each phase of condensation (input/pellet/supernatant) is shown in the z-score heat map.
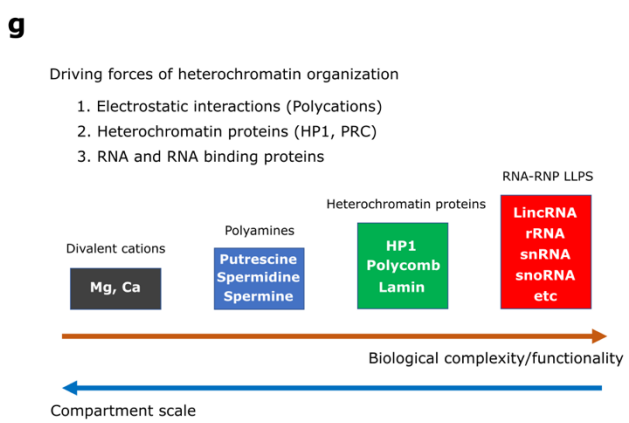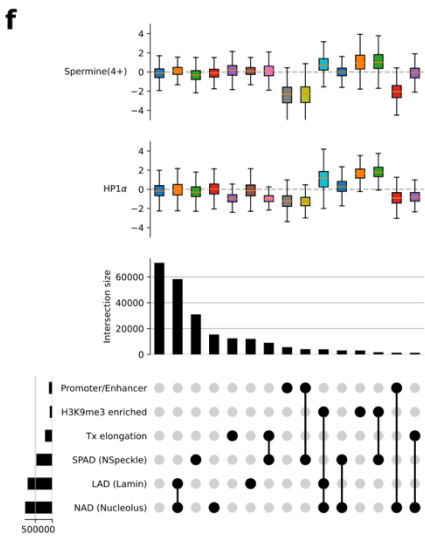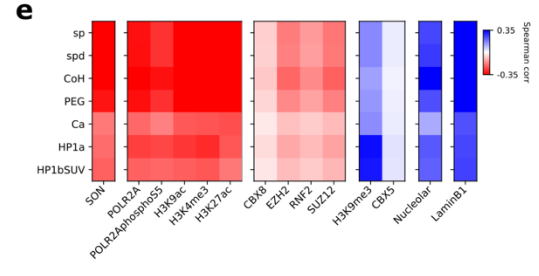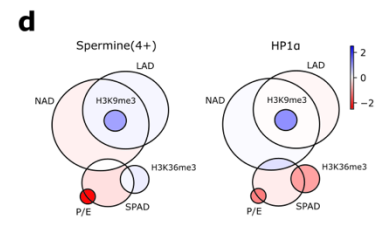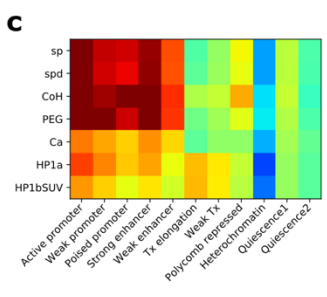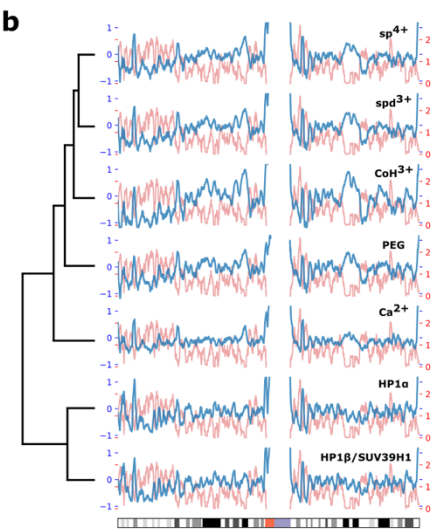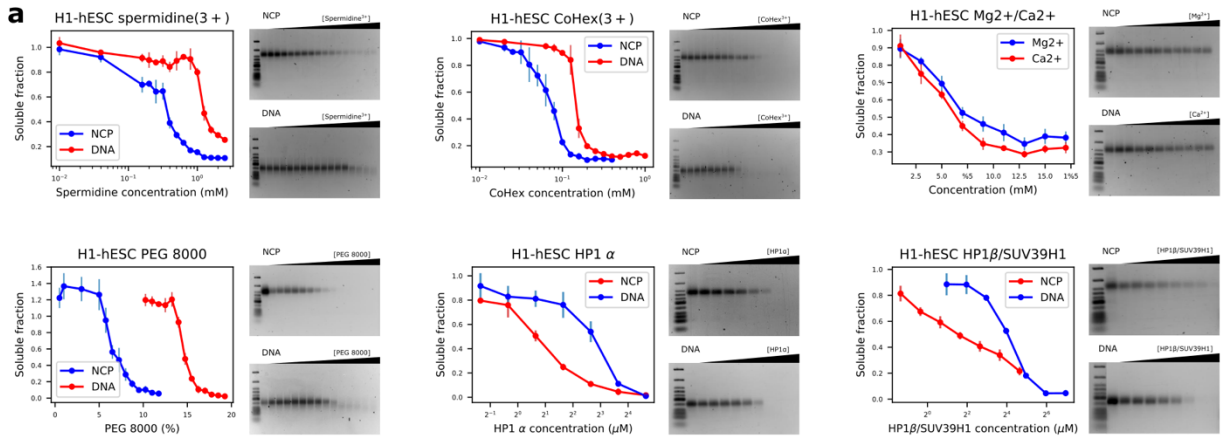
Extended Data Fig. 8: Condense-seq of H1-hESC native mono-nucleosomes using various condensing agents.

94   **a**, The soluble fraction of nucleosomes was measured by titrating the various condensing agents other than
95   spermine, including spermidine, cobalt-hexamine, magnesium/calcium, PEG 8000, and HP1α, HP1β with
96   SUV39H1 complex. **b**, Condensability scores were computed in 10kb resolution and plotted over
97   chromosome 1 (blue) and compared with the gene expression level (red). Hierarchical clustering of the
98   condensability profile shows that all ionic condensing agents (spermine/spermidine/cobalt-
99   hexamine/PEG/calcium) are clustered together but other protein-based condensing agents (HP1α and
100   HP1β) are clustered in a separate group. Comparison of condensability scores across various chromatin
101   states (**c**) and various nuclear compartments (**d**, LAD: lamin-associated domain, NAD: nucleolar-associated
102   domain, SPAD: nuclear speckle-associated domain, P/E: promoter or enhancer) for different condensing
103   agents. **e**, Correlations between nucleosome condensability and the enrichment of other nuclear
104   compartment hall mark proteins are represented as heat map. **f**, Upset plot of intersections between various
105   nuclear compartments and the corresponding nucleosome condensabilities by spermine and HP1α shown
106   as bar plots. **g**, Hypothetical hierarchal model of the biophysical driving force of chromatin organization: At
107   a large scale, chromatin is compartmentalized via ubiquitous charge–charge interactions, but other proteins
108   and RNAs are involved to generate sub-compartments that are smaller in scale but more specific function
109   directed.

110

111
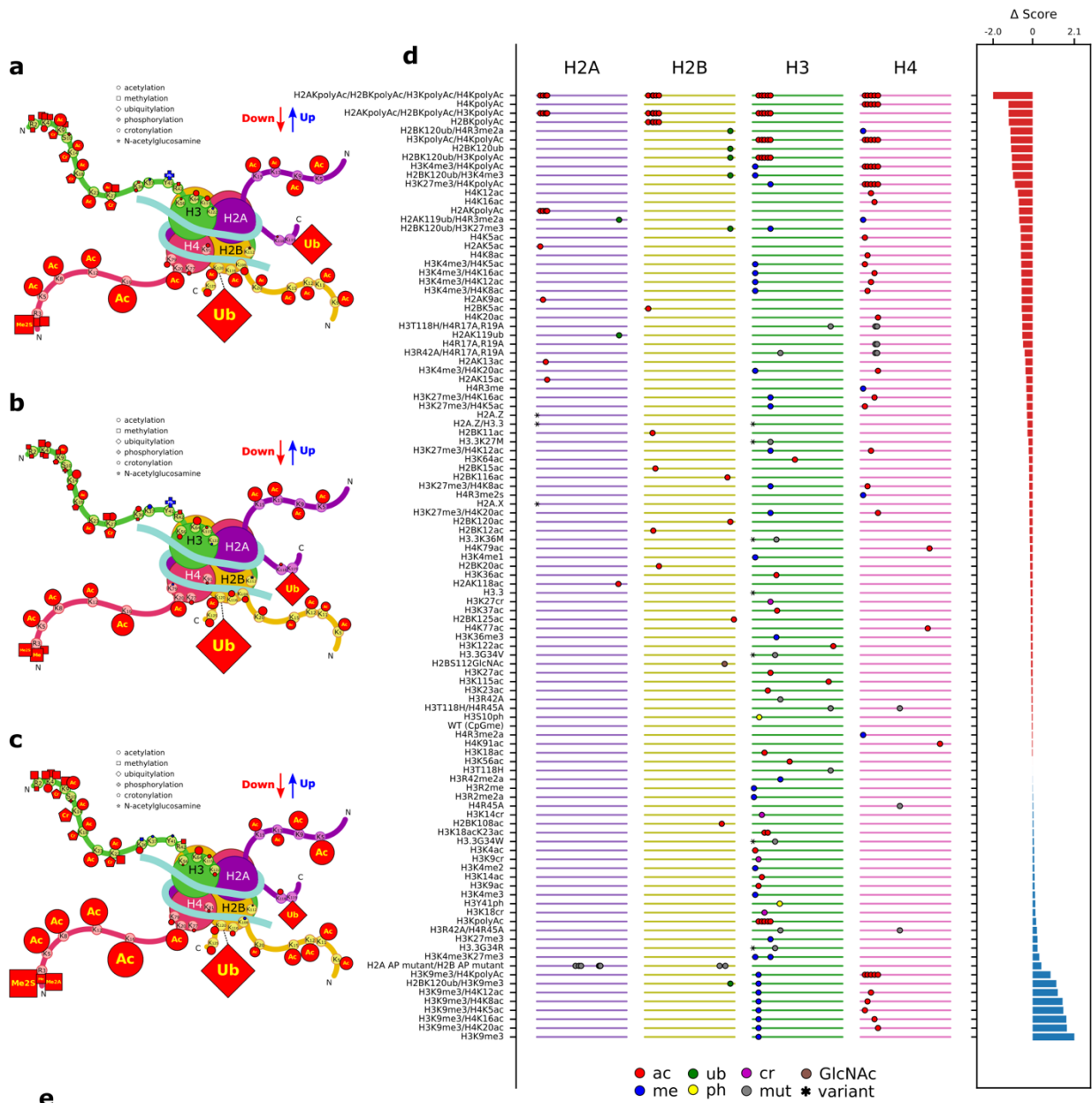
Extended Data Fig. 9: Summary of histone PTM effects on the nucleosome condensation by various condensing agents on the synthetic nucleosome library with PTM marks.

The effects of single PTMs on nucleosome condensation are depicted by the cartoons (**a**: spermidine, **b**: cobalt-hexamine, **c**: PEG 8000 as the condensing agent). Each symbol represents different types of PTMs as shown in the legend, and the size is proportional to the strength of effects. The colors of the marks indicate the direction of the effect (red: decreases condensation, blue: increases condensation) compared with the unmodified control. **d**, All condensability scores of the PTM library using HP1α as a condensing agent are summarized in the ladder bar plot. The library members are sorted from the lowest to the highest condensability scores from top to bottom. On the left panel, the ladder-like lines represent each histone subunit peptide from N-terminal (left) to the C-terminal (right). Each mark on the line indicates the location of the PTMs and the shape of the marks represents the PTM type (ac: acetylation, me: methylation, cr: crotonylation, ub: ubiquitylation, ph: phosphorylation, GlcNAc: GlcNAcylation, mut: amino acid mutation, var: histone variant). On the right panel, differences in the condensability score compared with the unmodified control are shown as bar plots for each member of the library. **e**, PCA analysis was conducted by combining the condensability scores of all five condensing agents (spermine/spermidine/cobalt-hexamine/PEG 8000/HP1α) into the five-dimensional state vector. In the PCA plot, each member of the library is represented by a symbols according to categories such as canonical wild-type nucleosome (WT), wild type with CpG methylation (WT+CpGme), mutations on wild type (WT+mut), nucleosome with histone variants (Var), mutations on histone variants (Var+mut), Acidic patch mutants (AP mutants), and nucleosomes with acetylation on H2A/B dimer (H2A/Bac), acetylation on H3 (H3ac), acetylation on (H4ac), having poly-acetylation (KpolyAC), methylation on H3 (H3me), methylation on H4 (H4me), acetylation on H4 and methylation on H3 (H4ac + H3me), crotonylation on H3 (H3cr), GlcNAcylation (GlcNAc), phosphorylation on H3 (H3ph), and ubiquitylation (+ub), all of which are shown in the figure legend. **f**, Comparison of condensability scores across different condensing agents. Scatter plots of condensability across different condensing agents are shown in the lower triangle, and the corresponding Spearman's correlations are shown in the upper triangle of the matrix.

139
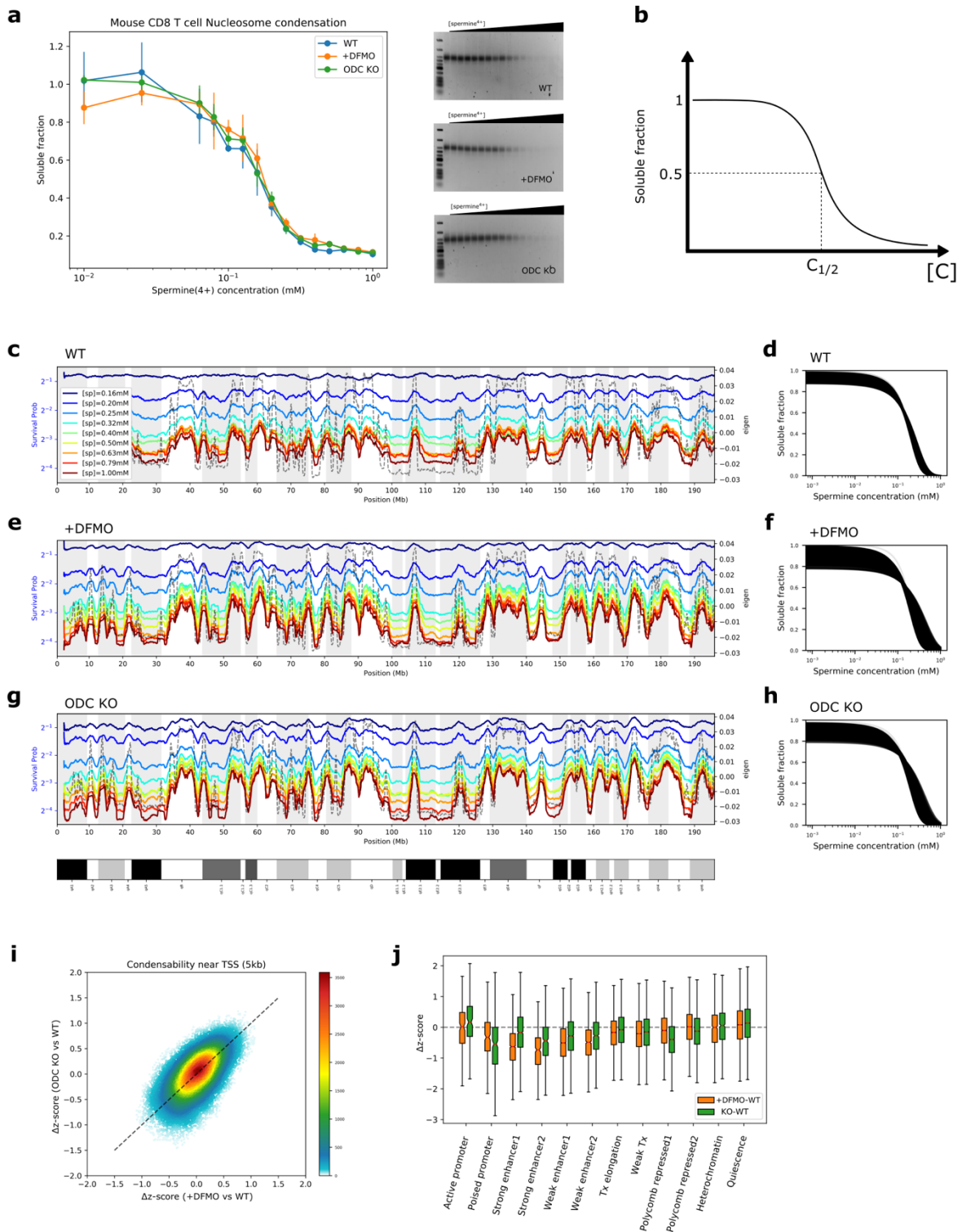
Extended Data Fig. 10: Condense-seq measurements of nucleosomes purified from mouse CD8+ T cells.

a, Soluble fractions were measured via UV-VIS spectroscopy and run in 2% agarose gel after condensation in various spermine concentrations. **b**, Condensation point ($c_{1/2}$) is defined by the concentration of condensing agent when the soluble fraction is half the input, so it is reversely correlated with condensability

145   score. **c–h**, Soluble fractions of nucleosomes in various spermine concentrations were calculated and
146   plotted over chromosome 1 in 10kb resolution. $C_{1/2}$ was computed for each bin after fitting the soluble
147   fraction change with a logistic function as shown fitting curves of all bins (**d**, **f**, **h**), and polyamine deficient
148   conditions show broader distribution of condensation points. (**c**, **d**: wild type control, **e**, **f**: +DFMO, **g**, **h**:
149   ODC KO) **i**, The scatter plot of Δ z-score of condensability near TSS shows a high correlation between
150   +DFMO and ODC KO. **j**, The Δ z-score of condensabilities is computed as the difference between the
151   standardized condensability of +DFMO or ODC KO conditions and the wild type control and then
152   categorized into the corresponding ChromHMM chromatin states.