ChIP-DIP maps binding of hundreds of proteins to DNA simultaneously and identifies diverse gene regulatory elements

Andrew A. Perez[1,6], Isabel N. Goronzy[1,2,3,6], Mario R. Blanco[1], Benjamin T. Yeh [1,2], Jimmy K. Guo[1,4], Carolina S. Lopes[5], Olivia Ettlin [1], Alex Burr [1] & Mitchell Guttman [1] ✉

Gene expression is controlled by dynamic localization of thousands of regulatory proteins to precise genomic regions. Understanding this cell type-specific process has been a longstanding goal yet remains challenging because DNA–protein mapping methods generally study one protein at a time. Here, to address this, we developed chromatin immunoprecipitation done in parallel (ChIP-DIP) to generate genome-wide maps of hundreds of diverse regulatory proteins in a single experiment. ChIP-DIP produces highly accurate maps within large pools (>160 proteins) for all classes of DNA-associated proteins, including modified histones, chromatin regulators and transcription factors and across multiple conditions simultaneously. First, we used ChIP-DIP to measure temporal chromatin dynamics in primary dendritic cells following LPS stimulation. Next, we explored quantitative combinations of histone modifications that define distinct classes of regulatory elements and characterized their functional activity in human and mouse cell lines. Overall, ChIP-DIP generates context-specific protein localization maps at consortium scale within any molecular biology laboratory and experimental system.

Although every cell in the body inherits the same genomic DNA sequence, distinct cell types express different genes to enable specific functions. Cell type-specific gene regulation involves the coordinated activity of thousands of regulatory proteins that localize at precise DNA regions to activate, repress and quantitatively control transcription levels. Genomic DNA is organized around nucleosomes[1], which contain histone proteins that undergo extensive post-translational modifications[2,3] and together define cell type-specific chromatin states. Chromatin state is controlled by regulators that directly read, write and erase specific histone modifications[2,4] as well as control nucleosome positioning and DNA accessibility[5,6]. This determines which genomic

regions are accessible for binding by sequence-specific transcription factors (TFs)[7], enzymes that transcribe DNA into RNA (RNA polymerases)[8] and other general and specific regulatory proteins that promote or suppress transcriptional initiation[9,10]. Conversely, recruitment of these regulatory proteins to specific DNA regions, along with transcriptional changes, can facilitate changes in chromatin state and DNA accessibility[5,11].

Understanding how regulatory protein binding leads to cell type-specific gene expression has been a central goal of molecular biology for decades[2]. Over the past 20 years, important technical advances have enabled genome-wide mapping of regulatory proteins

[1]Division of Biology and Bioengineering, California Institute of Technology, Pasadena, CA, USA. [2]David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [3]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA. [4]Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [5]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA, USA. [6]These authors contributed equally: Andrew A. Perez, Isabel N. Goronzy. ✉e-mail: mguttman@caltech.edu

and histone modifications (for example, ChIP followed by sequencing (ChIP–seq))[12–15], improved binding site resolution (ChIP–exo)[16,17], increased sample throughput (for example, through automation and/or sample pooling)[18,19] and enabled mapping within limited numbers of cells (for example, cleavage under targets and release using nuclease (CUT&RUN) and cleavage under targets and tagmentation (CUT&Tag))[20–22]. Yet, while these innovations have uncovered critical insights into gene regulation, most work by studying a single protein at a time. The few exceptions are multiplexed versions of CUT&Tag, which can measure up to three proteins in a single experiment[23]. However, these approaches are not readily scalable to larger numbers of proteins[23–25] and are primarily limited to mapping modified histones and other highly abundant proteins but not most TFs and chromatin regulators[26]. In contrast to CUT&Tag methods, CUT&RUN can map many TFs and regulatory proteins, but it is not amenable to multiplexed mapping of more than one protein at a time[27]. Due to the large number of distinct regulatory proteins involved and the cell type-specific nature of their interactions, constructing a comprehensive map of regulatory factors to dissect gene regulation remains a challenge using existing approaches. Initial attempts to overcome this led to the formation of various international consortia that generated reference maps of hundreds of proteins within a small number of cell types (ENCODE[28], PsychENCODE[29], ImmGen[30], etc.). Although these efforts have provided many critical insights[31–33], it is not possible to study cell type-specific regulation using maps generated from reference cell lines because protein binding maps and gene expression programs are intrinsically cell type specific[34–36]. To date, most mammalian cell types, model organisms and experimental models remain uncharacterized because generating additional cell type-specific regulatory maps using current approaches requires thousands of individual experiments for each cell type. Accordingly, there is a clear need for a highly scalable, multiplexed protein profiling method that can increase throughput of protein mapping by orders of magnitude and profile the diverse categories of DNA-associated proteins, including classes that have been traditionally easier to map (for example, modified histones) and those that have been more challenging (for example, TFs)[37]. Such a method would allow any laboratory to generate comprehensive maps for any cell type of interest in a rapid and cost-effective manner and would enable exploration of key questions that is not currently possible.

Results

Chromatin immunoprecipitation done in parallel enables multiplexed mapping of DNA-associated proteins

To enable highly multiplexed, genome-wide mapping of hundreds of DNA-associated proteins in a single experiment, we developed chromatin immunoprecipitation done in parallel (ChIP-DIP) (Fig. 1a, Supplementary Notes 1 and 2 and related Extended Data Fig. 1, and Supplementary Figs. 1–3). ChIP-DIP works by (1) using a rapid, modular approach to couple individual antibodies to beads containing a unique oligonucleotide tag (Extended Data Fig. 1a), (2) combining

sets of different antibody–bead–oligonucleotide conjugates to create an antibody–bead pool, (3) performing ChIP, (4) barcoding chromatin–antibody–bead–oligonucleotide conjugates via split-and-pool ligation[38–40] and (5) sequencing DNA and computationally matching split-and-pool barcodes that are shared between genomic DNA and the antibody–oligonucleotide. We define all unique reads containing the same split-pool barcode as a cluster and combine reads from all clusters corresponding to the same antibody to generate a localization map for each protein. The output of ChIP-DIP is analogous to the data generated by ChIP–seq; however, instead of a single map, ChIP-DIP generates a map for each antibody used (Fig. 1b).

To ensure that chromatin–antibody–bead–oligonucleotide conjugates remain intact throughout the ChIP-DIP procedure, we designed a series of experiments to measure dissociation between oligonucleotide and bead, antibody and bead, or antibody and chromatin (Extended Data Fig. 1b and Supplementary Note 1).

(1) Oligonucleotide–bead dissociation. We found that most clusters (>95%) contained only a single oligonucleotide type (Extended Data Fig. 1c), indicating that oligonucleotide movement between beads is rare.
(2) Antibody–bead dissociation. We found that beads that were not coupled to any antibodies were associated with little chromatin (<0.5%; Extended Data Fig. 1d), indicating that antibody movement between beads is rare.
(3) Antibody–chromatin dissociation. We purified human and mouse chromatin using differentially labeled beads, mixed them together and observed minimal levels of chromatin assigned to the bead type of the incorrect species (4–6%; Extended Data Fig. 1e), indicating that the vast majority of antibody–chromatin interactions (>88–92%) remain intact throughout the ChIP-DIP procedure.

Together, these results demonstrate that chromatin–antibody–bead–oligonucleotide conjugates remain intact throughout the ChIP-DIP procedure, enabling accurate multiplexed protein–DNA assignment (we discuss additional technical validations of ChIP-DIP in Supplementary Note 2 and the related Supplementary Figs. 1–3).

ChIP-DIP maps protein–DNA interactions in diverse pools

To test whether ChIP-DIP can accurately map genome-wide protein localization, we performed ChIP-DIP in human K562 cells using four well-studied proteins: (1) the CTCF sequence-specific DNA binding protein that binds to insulator sequences[41], (2) the histone H3 lysine 4 (H3K4) trimethylation (H3K4me3) modification that localizes at the promoters of active genes[14,42], (3) the RNA polymerase (RNAP) II enzyme that transcribes RNA[43] and (4) the histone H3 lysine 27 (H3K27) trimethylation (H3K27me3) modification that accumulates over broad genomic regions that are associated with Polycomb-mediated transcriptional repression[14,42] (Supplementary Table 1). We observed


Fig. 1 | ChIP-DIP is a highly multiplexed method for mapping proteins to genomic DNA. a, Schematic of the ChIP-DIP method. (1) Beads are coupled with an antibody and labeled with the associated oligonucleotide (oligo) tag (antibody ID). (2) Sets of antibody–bead–oligonucleotide conjugates are then mixed (antibody–bead pool) and used to perform ChIP. (3) Multiple rounds of split-and-pool barcoding are performed to identify molecules associated with each chromatin–antibody–bead–oligonucleotide conjugate. (4) DNA is sequenced, and genomic DNA and antibody (Ab)–oligonucleotide containing the same split-and-pool barcode are grouped into a cluster, which are used to assign genomic DNA regions to their linked antibodies. (5) All DNA reads from all clusters corresponding to the same antibody are used to generate protein localization maps. **b**, Protein localization maps over a specific human genomic region (hg38, chromosome (chr)12:53,649,999–54,650,000) for four protein targets: CTCF, H3K4me3, RNAP II and H3K27me3. Left, protein localization generated by ChIP-

DIP in K562 cells. Top track shows read coverage before protein assignment, and the bottom four tracks correspond to read coverage after assignment to individual proteins. Right, ChIP–seq data generated by ENCODE in K562 cells for these same four proteins are shown for the same region. To enable direct comparison of scales between datasets, we normalized the scale to coverage per million aligned reads. Scale is shown from zero to maximum coverage within each region. **c**, Comparison of ChIP-DIP and ChIP–seq maps over specific regions corresponding to magnified views of the larger region shown in **b**. The locations presented are demarcated by colored bars above the gene track in **b**. Scale shown is like that in **b**. **d**, Genome-wide comparison (density plots of signal correlation) between the localization of each individual protein measured by ChIP-DIP (x axis) or ChIP–seq (y axis). Points are measured genome wide across 10-kb windows (CTCF, H3K27me3) or all promoter intervals (H3K4me3, RNAP II).

localization patterns that are highly comparable at specific genomic sites (Fig. 1b,c) and strongly correlated genome wide ($r = 0.837$–0.956; Fig. 1d) to ChIP–seq profiles generated by the ENCODE consortium[28,31,44] (Supplementary Table 2).

Because there are many hundreds of regulatory proteins, we explored whether ChIP-DIP could generate maps for large pools of distinct proteins. We considered two possibilities that might limit the scale of ChIP-DIP. (1) As the size of each pool increases, the background

levels of immunoprecipitated chromatin might increase and obscure our ability to generate high-quality binding maps for individual proteins ('pool size'). (2) If multiple proteins bind to similar DNA regions, this might deplete the associated chromatin and preclude our ability to accurately map each protein. In this way, the exact composition of the antibody pool used might impact the maps obtained for an individual protein ('pool composition'). To explore these possibilities, we analyzed the genome-wide profiles of the same four proteins (CTCF,

a Protein tagging | ChIP | >5 rounds split-pool tagging | Sequence + assign | Analyze

b ChIP-DIP chr12:53,649,999–54,650,000 | ChIP–seq (ENCODE) chr12:53,649,999–54,650,000

c 1 CTCF | 2 H3K4me3 | 3 RNAP II | 4 H3K27me3

**d** CTCF $r = 0.843$ | H3K4me3 $r = 0.956$ | RNAP II $r = 0.898$ | H3K27me3 $r = 0.837$
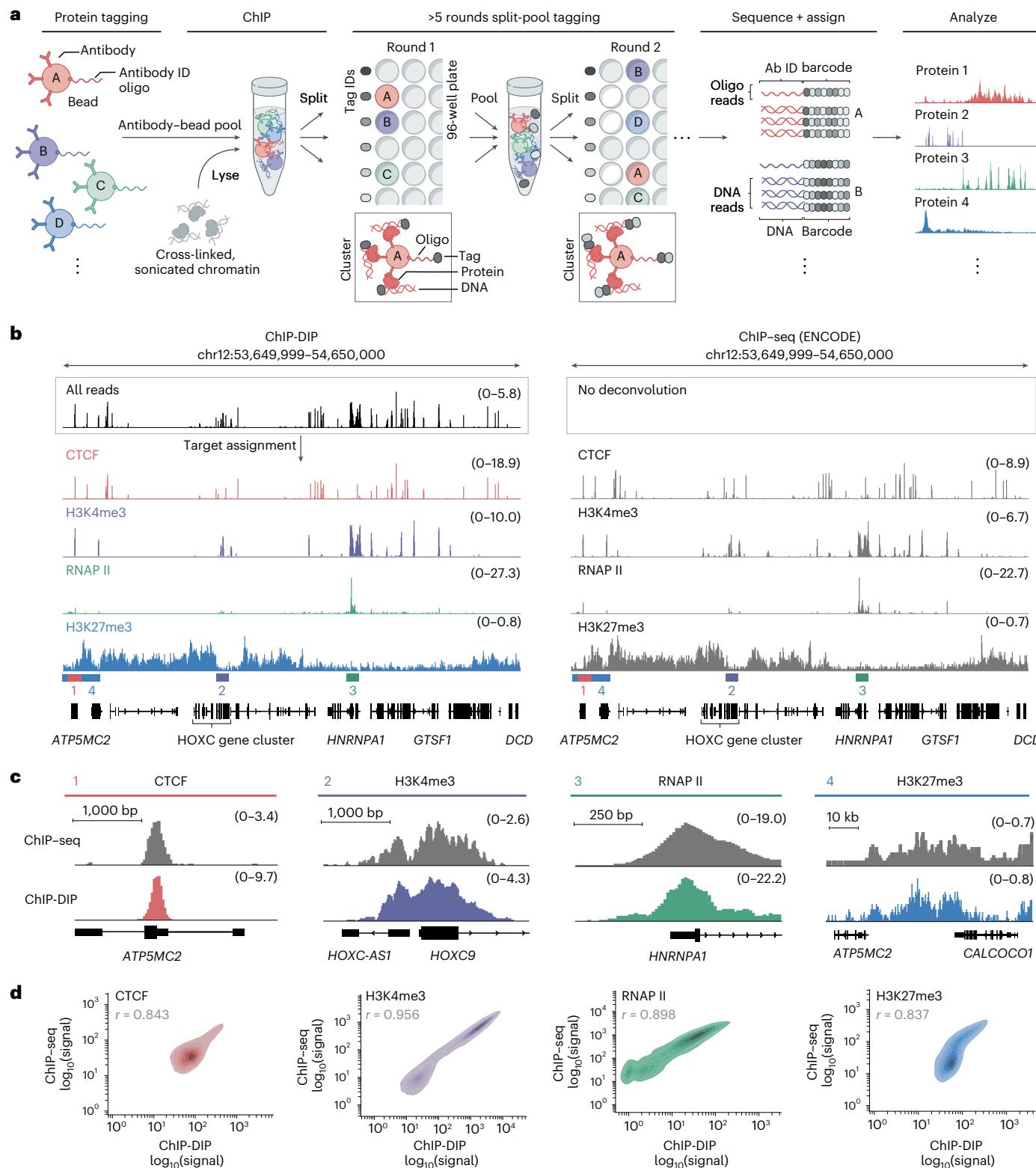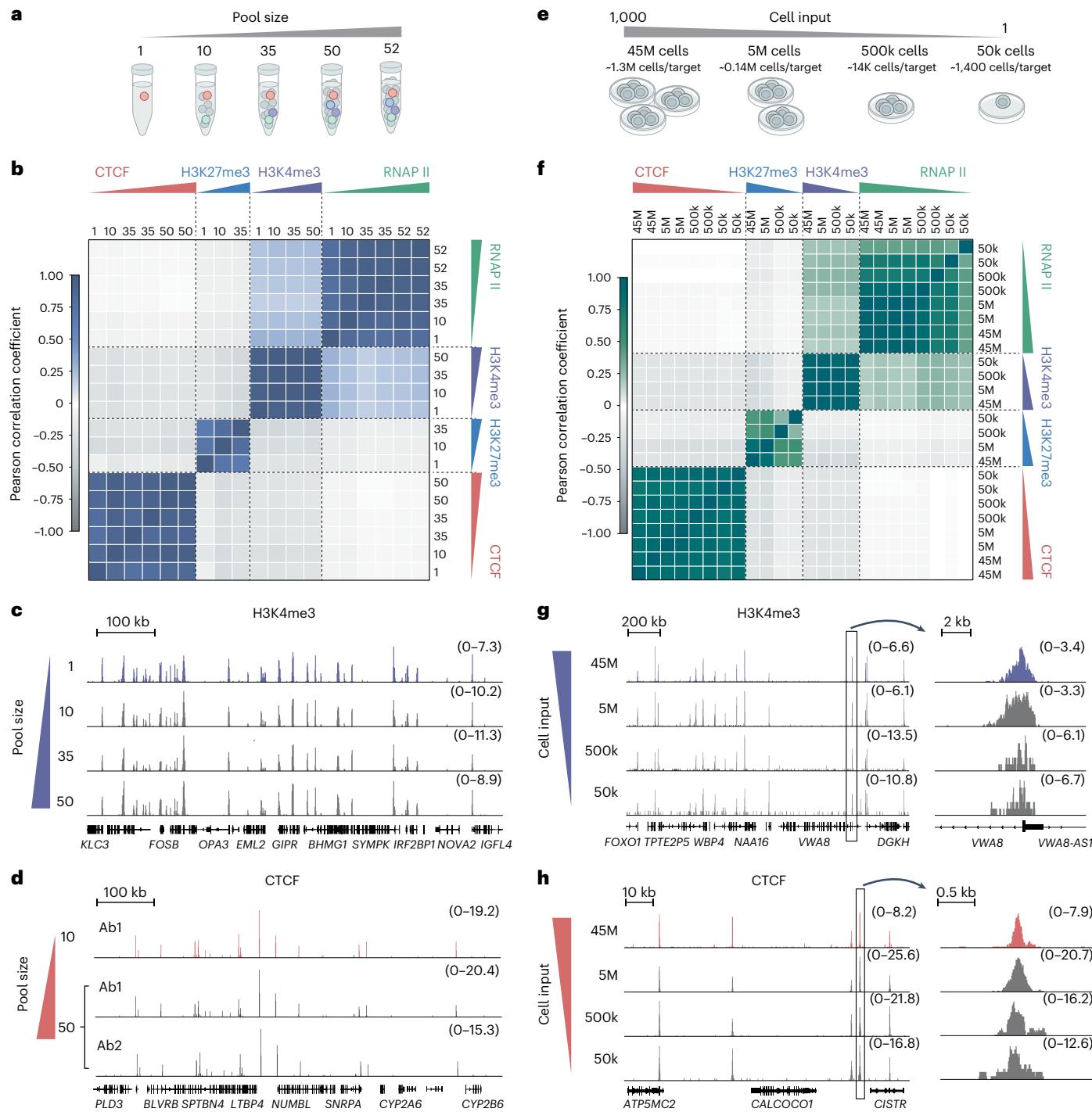
Fig. 2 | ChIP-DIP accurately maps known protein–DNA interactions across a range of multiplexed protein numbers, protein compositions and cell numbers. a, Schematic of the experimental design to test the scalability of antibody–bead pool size and composition. **b**, Correlation heatmap for protein localization maps of 4 proteins (CTCF, H3K4me3, RNAP II and H3K27me3) generated using antibody pools of 5 different sizes (1, 10, 35, 50 and 52 antibodies per pool) and compositions. Correlations were calculated over the set of regions corresponding to the union of all peaks called for any of the four targets in the K562 ten-antibody experiment and were calculated using the background-corrected ChIP-DIP signal for each sample (Methods). Pool sizes are listed along the top and left axes. Replicate proteins in the same pool indicate that a different antibody was used for that protein. Some proteins were not included in every pool. **c**, Comparison of H3K4me3 localization over a specific genomic region (hg38, chr19:45,345,500–46,045,500) when measured within various antibody pool sizes and compositions. Scale is normalized to coverage per million aligned reads. **d**, Comparison of CTCF localization over a specific genomic region

(hg38, chr19:40,349,999–41,050,000) when measured within a pool of 10 antibodies containing a single CTCF-targeting antibody (top) or within a pool of 52 antibodies containing 2 different CTCF-targeting antibodies (bottom). Scale is normalized to coverage per million aligned reads. **e**, Schematic of the experimental design to test the amount of cell input required for ChIP-DIP. k, thousand; M, million. **f**, Correlation heatmap for protein localization maps of four targets (CTCF, H3K4me3, RNAP II and H3K27me3) generated using various amounts of input cell lysate. Correlations were calculated over the same set of regions as **b** and using the background-corrected ChIP-DIP signal for each sample (Methods). Amounts of input cell lysate are listed along the top and left axes. **g**, Comparison of H3K4me3 localization over a specific genomic region (hg38, chr13:40,600,000–42,300,000) when measured using various amounts of input cell lysate. Scale is normalized to coverage per million aligned reads. **h**, Comparison of CTCF localization over a specific genomic region (hg38, chr12:53,664,000–53,764,000) when measured using various amounts of input cell lysate. Scale is normalized to coverage per million aligned reads.

H3K4me3, RNAP II, H3K27me3) measured across six distinct experiments containing increasing pool sizes (1, 10, 35, 50 and 52 antibodies per pool) and containing distinct pool compositions, including pools containing independent antibodies targeting the same protein (CTCF) or multiple proteins within the same complex (for example, multiple members of the Polycomb repressive complexes (PRC) 1 and 2; Extended Data Fig. 2). In all cases, we observed highly consistent genome-wide profiles generated from these distinct pools, regardless of pool size or protein composition (Fig. 2a–d, Supplementary Fig. 4 and Supplementary Table 3).

One of the major challenges with mapping DNA binding proteins in primary cell types, disease models and rare cell populations is the large numbers of cells required to map each protein target. To explore the number of cells required to generate reliable genome-wide maps with ChIP-DIP, we used the same pool of 35 different antibodies across an ~1,000-fold range of input cell numbers (4.5×10^7, 5×10^6, 5×10^5 and 5×10^4 total cell equivalents). We observed strong genome-wide correlations and peak overlap across the range of cell numbers (Fig. 2e–h and Supplementary Figs. 5 and 6) and enrichment profiles similar to data generated by low-cell number CUT&Tag (Supplementary Figs. 7 and 8). Because ChIP-DIP generates many individual maps from the same lysate, this further reduces the effective number of cells required to map each protein. In this example, we mapped 35 different proteins using 5×10^4 total cell equivalents, which corresponds to using the chromatin yield from ~1×10^3 cells to map each individual protein with a traditional individual assay.

Together, these results demonstrate that ChIP-DIP generates data that are highly comparable to those generated by standard methods and is robust across different antibody pools and input cell numbers.

ChIP-DIP maps hundreds of diverse DNA-associated proteins

We next explored whether ChIP-DIP can simultaneously map proteins from distinct functional categories, some of which have been traditionally easier to map than others[45,46]. To do this, we performed ChIP-DIP on >60 distinct proteins in human K562 cells and >160 distinct proteins in mouse embryonic stem cells (mESCs) across six experiments (Supplementary Table 1 and Supplementary Note 2). These included 39 histone modifications, 67 chromatin regulators, 51 TFs and all three RNAPs and four of their post-translationally modified isoforms.

Histone modifications. Histone modifications define cell type-specific chromatin states and have proven incredibly useful for annotating cell type-specific regulatory elements[47]. We mapped 39 histone modifications, including 18 acetylation, 17 methylation, three ubiquitination and one phosphorylation marks, in either mESCs or K562 cells (Fig. 3a). We confirmed the specific localization of five histone modifications commonly used to demarcate functional states[7] as well as additional modifications associated with each state (Extended Data Fig. 3 and Supplementary Fig. 9): enhancer regions[48] (H3K4 monomethylation (H3K4me1), H3K4 dimethylation (H3K4me2), H3K27 acetylation (H3K27ac); Fig. 3b), transcribed regions[14,49,50] (H3K36me3, H3K79me1, H3K79me2; Fig. 3c), promoter regions[14,42,51] (H3K4me3, H3K9ac; Fig. 3d), Polycomb-repressed regions[52] (H3K27me3, histone H2A lysine 119 ubiquitination (H2AK119ub); Fig. 3e) and constitutive heterochromatin regions[53] (H3K9me3, H4K20me3; Fig. 3f). These data indicate that ChIP-DIP accurately maps histone modifications with distinct genome-wide patterns (broad and focal localization) that represent distinct activity states (active or repressive) and that localize at distinct functional elements (promoters, enhancers, gene bodies and intergenic regions).

Chromatin regulators. Chromatin regulators are responsible for reading, writing and erasing specific histone modifications and are critical for the establishment, maintenance and transition between chromatin states[11,54]. We measured 67 chromatin regulators associated

with various histone methylation, acetylation and ubiquitination marks as well as with DNA methylation in either mESCs or human K562 cells (Fig. 3a). As expected, we observed that an eraser (JARID1A)[55] and a writer (RBBP5-containing complex)[56] of H3K4me3 localize at H3K4me3-modified promoter sites (Fig. 3g and Extended Data Fig. 4a). Additionally, we observed that components of the PRC1 (RING1B, CBX8)[57] and PRC2 complexes (EED, SUZ12, EZH2)[58] colocalize and are enriched over genomic regions containing their respective histone modifications (H2AK119ub and H3K27me3; Fig. 3h and Extended Data Figs. 2 and 4b). Similarly, we observed colocalization of heterochromatin protein (HP)1α and HP1β at genomic DNA regions containing their associated heterochromatin marks, H3K9me3 and H4K20me3 (ref. 59) (Fig. 3i and Extended Data Fig. 4c). These data indicate that ChIP-DIP accurately maps chromatin regulators from diverse complexes and with distinct functions.

TFs. TFs bind *cis*-regulatory elements in combinatorial patterns to control gene expression. Generating comprehensive maps of TF localization has proven difficult because there are large numbers of distinct TFs, most are cell type specific and they are challenging to map by ChIP–seq because they tend to be lower in abundance and only transiently associated with DNA[60,61]. To explore whether ChIP-DIP can map large sets of TFs, we measured 15 TFs in K562 cells and 43 TFs in mESCs, including constitutive (for example, SP1 and USF2)[62,63], stimulus-dependent (for example, p53 and NRF1)[64–66] and developmental and/or cell type-specific (for example, NANOG and RFX1)[67,68] DNA binding proteins[69] (Fig. 4a). We obtained high-resolution binding maps for TFs in both cell types, with previously characterized TFs showing localization at their expected genomic DNA targets[62,66,70–73] and a median peak concordance of >90% for known binding sites (Fig. 4a,b and Supplementary Table 4). Using these genome-wide localization data, we accurately identified expected DNA binding motifs (Supplementary Fig. 10 and Supplementary Table 5), including the 20-bp dimer motif of p53 (ref. 74) and the 21-bp RE-1 consensus sequence of the TF REST[75] (Fig. 4c). Together, these data indicate that ChIP-DIP generates accurate, high-resolution binding maps of diverse TFs in multiple cell types.

RNAPs. Different classes of RNA are transcribed by distinct RNAPs: RNAP I transcribes the 45S ribosomal RNA (rRNA) encoding the 18S, 28S and 5.8S rRNAs; RNAP II transcribes messenger RNA and various noncoding RNAs, including small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and long noncoding RNA; and RNAP III transcribes diverse small RNAs, including transfer RNA (tRNA), 5S rRNA and 7SL, 7SK and U6 snRNA[8]. We leveraged ChIP-DIP to simultaneously map all three RNAPs and the post-translationally modified forms of RNAP II. We observed that each RNAP localizes with high selectivity to its corresponding classes of genes; RNAP I binds at ribosomal DNA (rDNA), RNAP II binds at mRNA and snRNA genes, and RNAP III binds at tRNA genes (Fig. 4d and Extended Data Fig. 5a). Moreover, we observed distinct localization patterns of different RNAP II phosphorylation states: serine 5-phosphorylated RNAP II localizes at promoters, while serine 2-phosphorylated RNAP II accumulates over the gene body and past the 3′ end of the gene (Extended Data Fig. 5b). These data indicate that ChIP-DIP accurately maps the localization of the three RNAPs, including multiple functional phosphorylation states of RNAP II, at distinct gene classes and gene features.

Together, these results establish ChIP-DIP as a modular, highly multiplexed method that generates high-quality maps for a wide range of DNA-associated proteins spanning diverse biological functions.

Multisample maps reveal chromatin changes in immune cells

Because gene regulation is highly dynamic in nature, we explored whether ChIP-DIP can be used to study changes in protein localization across multiple experimental conditions simultaneously. To explore
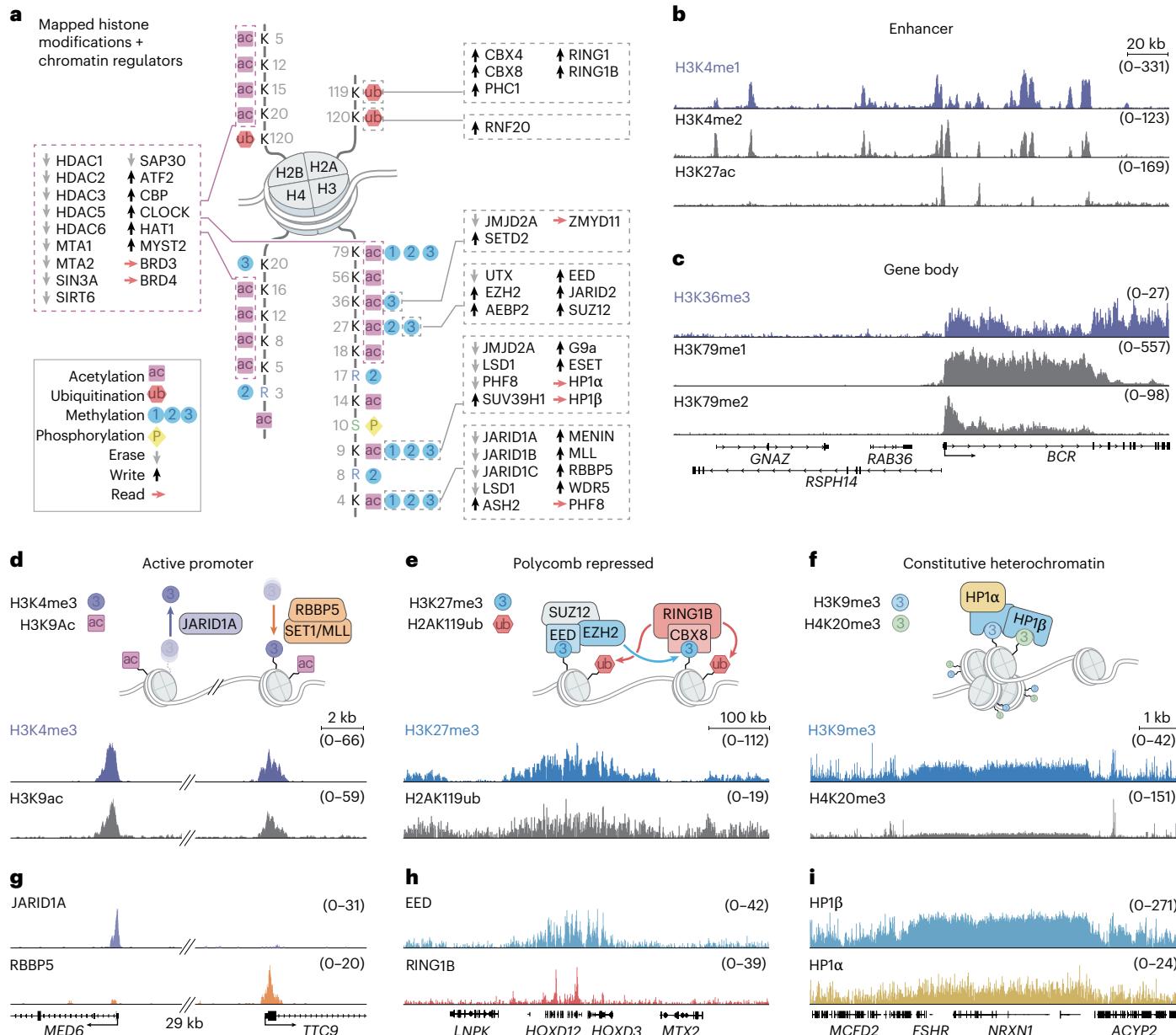
Fig. 3 | ChIP-DIP accurately maps dozens of functionally diverse histone modifications and chromatin regulators. a, Illustration of the diverse histone modifications and chromatin regulatory proteins mapped in K562 cells or mESCs using ChIP-DIP. **b,c**, Visualization of multiple histone modifications across a genomic region (hg38, chr22:23,050,000–23,290,000) in K562 cells corresponding to multiple histone modifications associated with enhancers (H3K4me1, H3K4me2 and H3K27ac) (**b**) and active gene bodies (H3K36me3, H3K79me1 and H3K79me2) (**c**). **d**, Top, schematic of histone modifications and chromatin regulators associated with active promoters. Bottom, visualization of multiple histone modifications associated with active promoters (H3K4me3 and H3K9ac) across a genomic region (mm10, chr12:81,590,000–81,636,000) in mESCs. Hash marks indicate an intervening 29-kb region that is not shown. **e**, Top, schematic of histone modifications and chromatin regulators associated with Polycomb-mediated repression. Bottom, visualization of multiple histone modifications associated with Polycomb-mediated repression (H3K27me3 and H2A119ub) across a genomic region (hg38, chr2:175,846,000–176,446,000) containing the silenced HOXD cluster in K562 cells. **f**, Top, schematic of histone modifications and chromatin regulators associated with constitutive heterochromatin. Bottom, visualization of multiple histone modifications associated with constitutive heterochromatin (H3K9me3 and H4K20me3) across a genomic region (hg38, chr2:46,200,000–55,700,000) in K562 cells. **g**, Visualization of an H3K4me3-associated eraser (JARID1A) and writer component (RBBP5) across the same genomic region as that in **d**. **h**, Visualization of PRC2 (EED) and PRC1 (RING1B) components across the same genomic region as that in **e**. **i**, Visualization of HP1β and HP1α across the same genomic region as that in **f**.

this, we treated primary mouse bone marrow-derived dendritic cells (mDCs) with lipopolysaccharide (LPS), which induces an anti-bacterial pathogen response that leads to changes in the expression of hundreds of genes[76], and collected cells at 0, 6 and 24 h after stimulation (Fig. 5a). In all, we used a pool of 25 antibodies to map 22 distinct chromatin modifications, including all five canonical active and repressed functional states, at all three time points.

We found that multiple chromatin modifications, including modifications demarcating active promoters, gene bodies and enhancers as well as insulation domains (CTCF) and repressive domains (H3K27me3), change substantially across the time course (Fig. 5b and Extended Data Fig. 6a), with the largest number of changes occurring within the first 6 h. For H3K27ac-enriched regions, these changes follow three temporal patterns: (1) one set stays the same across the entire time course
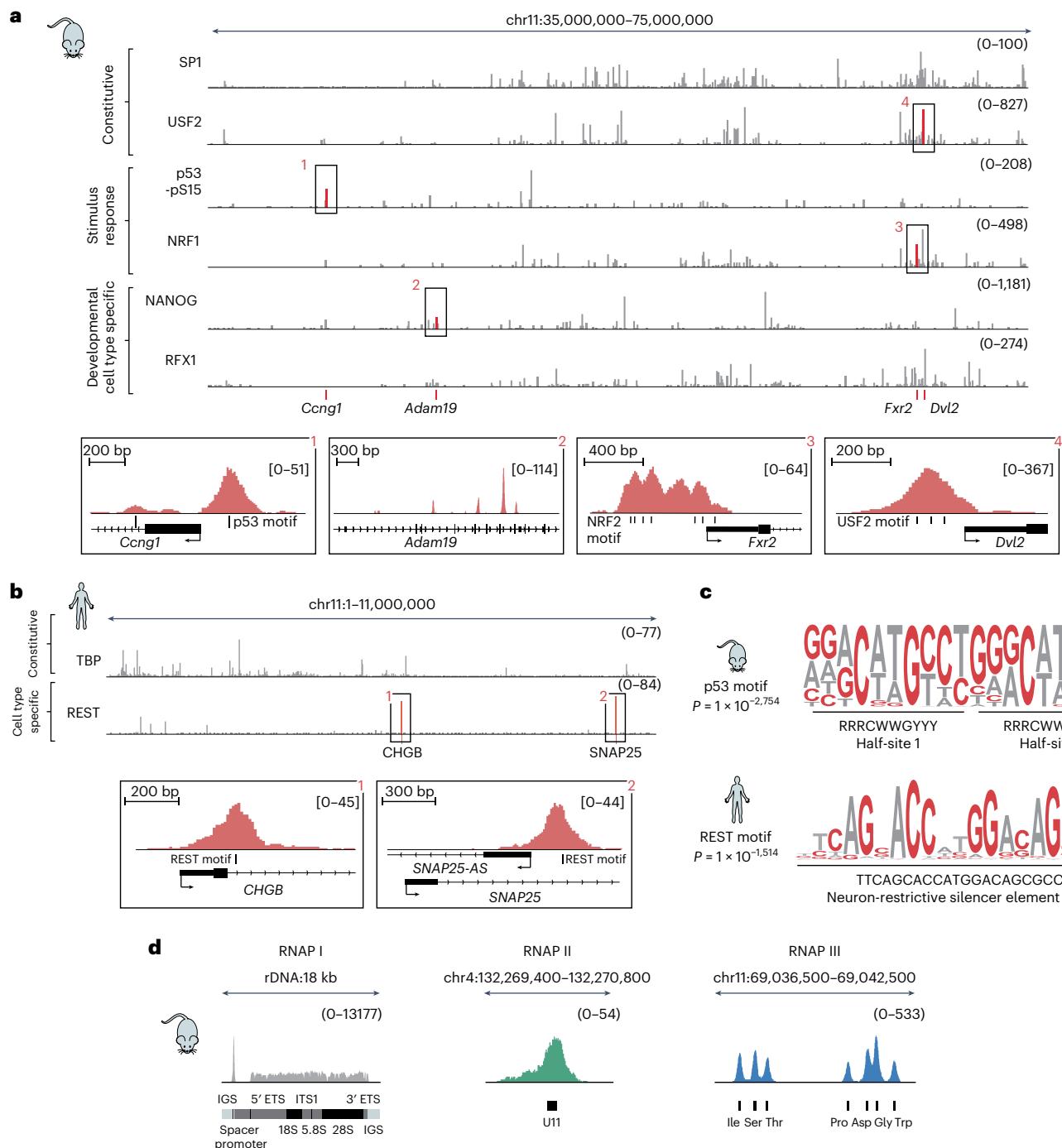
Fig. 4 | ChIP-DIP accurately maps dozens of TFs representing diverse functional classes and all three RNAPs. a, Top, visualization of six TFs (SP1, USF2, p53-pSer15, NRF1, NANOG, RFX1) representing three broad functional classes (constitutive, stimulus response, development–cell type specific) across a genomic region (mm10, chr11:35,000,000–75,000,000) in mESCs. Bottom, higher-resolution magnified views showing individual TF binding patterns at selected targets and motif sites. (1) p53 binding the p53 response element on the cyclin G1 gene (*Ccng1*) promoter. (2) NANOG binding a cluster of sites internal to the developmental gene *Adam19*. (3) NRF1 binding multiple copies of its motif at the *Fxr2* promoter. (4) The constitutively active USF2 binding its triplicate

E-box motif. **b**, Visualization of the TF TBP (constitutive) and REST (NRSF; cell type specific) across a genomic region (hg38, chr11:1–11,000,000) in K562 cells. Bottom, higher-resolution magnified views highlight two individual peaks of REST at motif sites near promoters of known neuronal genes *CHGB* and *SNAP25*. **c**, De novo generated motifs for p53 (top) in mESCs and REST (bottom) in K562 cells using binding sites identified using ChIP-DIP. **d**, Visualization of RNAP I at the promoter and along the gene body of rDNA (left), RNAP II at an snRNA gene (middle) and RNAP III at a cluster of tRNA genes (right) in mESCs. ITS1, internal transcript spacer 1; IGS, intergenic spacer; ETS, external transcript spacer.

('stable'), (2) one set increases upon stimulation ('activated') and (3) one set decreases upon stimulation ('repressed') (Fig. 5c). These three sets of regions correspond to genes for which transcription remains unchanged, increases and decreases, respectively (Extended Data

Fig. 6b). As an example, we observed that H3K27ac regions near inflammatory cytokine and chemokine genes, which increase in expression upon LPS stimulation, show a dramatic increase in acetylation (Fig. 5b and Extended Data Fig. 6c). We observed three temporal patterns of

acetylation at 'activated' regions: (1) H3K27ac regions that initially increase but then return to baseline by 24 h ('pulse'), (2) H3K27ac regions that only increase after 6 h of stimulation ('delayed') and (3) H3K27ac regions that continue to increase throughout the time course ('sustained') (Fig. 5d). At 'activated' regions, we find temporally matching patterns of gene expression, while, at 'repressed' regions, gene expression can recover without a corresponding return in acetylation. Notably, while both promoter and enhancer chromatin modifications change upon LPS stimulation, enhancer modifications show a stronger concordance with changes in transcriptional activity than changes at promoters (Fig. 5e,f and Extended Data Fig. 6d).

In sum, ChIP-DIP enables direct characterization of protein localization changes across distinct samples, time points or perturbations in various biological systems, including primary cells.

Protein localization analysis reveals distinct *cis*-regulators

Previous large-scale analyses have identified histone modifications that demarcate distinct genomic elements (for example, promoters, enhancers, transcribed regions, etc.)[77], their activity state (active, inactive, repressed) and regulatory potential (poised or primed for activation)[78]. However, because of the large number of histone modifications and regulatory proteins, many efforts have focused on mapping only five histone modifications (that is, H3K4me3, H3K4me1, H3K36me3, H3K9me3 and H3K27me3)[7]. Because ChIP-DIP can map large numbers of diverse proteins, we asked whether combinations of histone modifications and regulatory proteins can provide additional information about activity states and regulatory potentials of *cis*-regulatory elements beyond those captured by the five commonly studied histone modifications.

Promoter type and activity state are defined by combinations of histone modifications. H3K4me3 is generally thought to mark the promoters of actively transcribed RNAP II transcripts[14,42,79]. Consistent with this, we found H3K4me3 over the promoters of actively transcribed RNAP II genes but also near RNAP I promoters (rRNA) and active RNAP III genes (tRNA) (Fig. 6a). Similarly, we observed that other histone modifications associated with active RNAP II promoters, including H3K4me2, H3K9ac, H3K27ac and H3K56ac, were also enriched at RNAP I and III genes (Fig. 6a,b and Supplementary Fig. 11). For example, focusing on a genomic region containing neighboring RNAP II and RNAP III genes, we observe specific binding of the associated RNAP with shared TFs and chromatin modification patterns over these genes (Fig. 6b).

Although the presence of these histone modifications does not appear to distinguish between genes transcribed by different polymerases, we observed that their position relative to the transcriptional start site (TSS) varies with RNAP gene type: for RNAP I genes, these modifications localize before the TSS; for RNAP II, they flank the promoter and are enriched downstream of the TSS, both when considering all promoters and when excluding bidirectional promoters (Methods); and for RNAP III, they flank the gene body, localizing both upstream of the TSS and downstream of the transcriptional termination site (Fig. 6a and Supplementary Fig. 11). In addition, the three RNAPs have different relative levels of these histone modifications near their respective gene promoters. Specifically, we found that RNAP I and II promoters display stronger H3K56ac enrichment and RNAP I and III display stronger H3K4me2 enrichment relative to H3K4me3 (Fig. 6c). Although different antibodies have intrinsically different sensitivities that can confound direct comparisons, these distinct patterns correspond to differences in the relative signals of the same proteins within the same sample at distinct genomic regions (for example, RNAP I, II and III promoters). In this way, both quantitative combinations of histone modifications and their relative positions define distinct classes of promoters (Fig. 6d).

Next, we considered whether other histone modifications may distinguish activity states of RNAP II promoters. Previously, co-occurring H3K4me3 and H3K27me3 modifications ('bivalent domains') have been shown to associate with a poised transcriptional state[13,78], an effect we also observe in our data (Supplementary Fig. 12). To explore the spectrum of co-occurring modifications at promoters, we quantified the levels of ten histone modifications at H3K4me3-enriched regions and identified five clusters; four are enriched with other histone modifications (clusters 1–4), and one is not (cluster 5). The four co-occurring clusters correspond to H3K4me3 along with H3K27me3–H2AK119ub (cluster 1), H3K36me3–H3K79me2–H3K79me3 (cluster 2), H3K9me3–H4K20me3 (cluster 3) or H3K4me1–H3K27ac (cluster 4) (Fig. 7a). These clusters correspond to promoters that exhibit distinct transcriptional activity (Fig. 7b and Extended Data Fig. 7) and are enriched for distinct gene classes, such as ribosomal protein and cell cycle genes (cluster 2), zinc finger (ZNF) protein (cluster 3) and long intergenic noncoding RNA genes (sets 3 and 4)[13,80] (Fig. 7c–g). Consistent with the fact that H3K4me3 localization associates with functionally distinct classes of promoters, we observed different combinations of H3K4me3-associated readers, writers and erasers at distinct promoters (Extended Data Fig. 4a).

In sum, these results demonstrate that combinations of histone modifications can distinguish promoter features including polymerase (Fig. 6d), gene type and activity level (Fig. 7h).

Enhancer type, activity and potential are defined by combinations of histone modifications. There are >40 different histone acetylation marks[3], many of which have been associated with enhancers and active transcription. We mapped 15 acetylation marks on all four histone proteins and observed that they colocalize at similar sites genome wide (Pearson $r = 0.86$–0.97)[81] (Extended Data Fig. 8). We considered whether these strong correlations indicate redundancy or whether there is additional regulatory information encoded by the relative levels of each acetylation mark at specific genomic sites. To explore this, we used a matrix factorization algorithm to define five weighted combinations at highly acetylated regions (Methods, Fig. 8a,b, Supplementary Note 3 and Supplementary Fig. 13). These quantitative combinations

---

Fig. 5 | ChIP-DIP reveals dynamics changes in the chromatin landscape following LPS stimulation of primary mDCs. a, Schematic of the experimental design to profile chromatin changes in primary cells following LPS stimulation. **b**, Visualization of H3K27ac, H3K9ac, H3K36ac and transcription levels at 0 h, 6 h and 24 h across a genomic region (mm10, chr2:129,298,000–129,420,000) containing the LPS-stimulated interleukin genes *Il1a* and *Il1b*. To enable direct comparison of time points, we normalized the scale to coverage per million aligned reads, and, for each target, scale is shown from zero to maximum coverage for all three time points. **c**, *k*-means clustered heatmap of H3K27ac coverage at individual enriched genomic regions (*y* axis) across time points (*x* axis). Three distinct sets of regions showing differential temporal patterns are labeled along the left side. Regions associated with example inflammatory genes are labeled on the right side. **d**, Line plots of relative H3K27ac coverage of regions from **c** (left) and expression of associated genes (right) versus time. Subsets of enhancer regions that are newly acetylated after stimulus ('activated') are shown

above the dashed line, and subsets of enhancer regions that are deacetylated after stimulus ('repressed') are shown below the dashed line (Supplementary Methods). Mean levels are shown as solid lines with surrounding 95% confidence interval bands. **e**, Visualization of H3K27ac, H3K79me1 and transcription levels at 0 h, 6 h and 24 h across a genomic region (mm10, chr9:25,440,000–25,640,000) containing regions belonging to the 'repressed' set from **c**. A masked region of ~6 kb within the gene has been removed and is indicated by hash marks. **f**, Visualization of H3K27ac, H3K79me1 and transcription levels at 0 h, 6 h and 24 h across a genomic region (mm10, chr5:92,320,000–92,380,000) containing regions belonging to the 'activated' set from **c**. For **e** and **f**, scale per histone target is shown from zero to maximum coverage across both regions; scale for transcription is shown from zero to maximum coverage across a single region. Schematics showing relative quantification of levels across a region are shown on the right of each track.
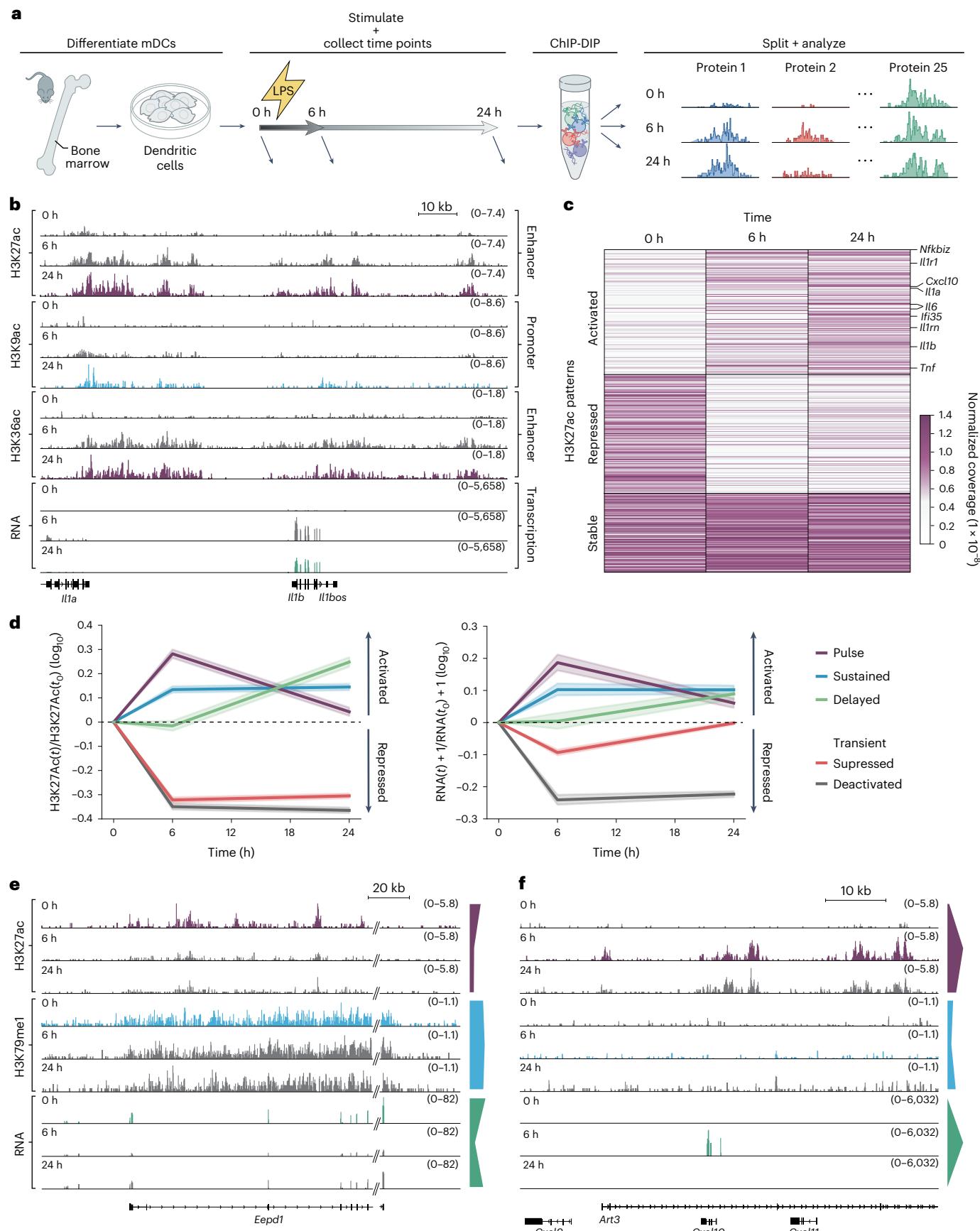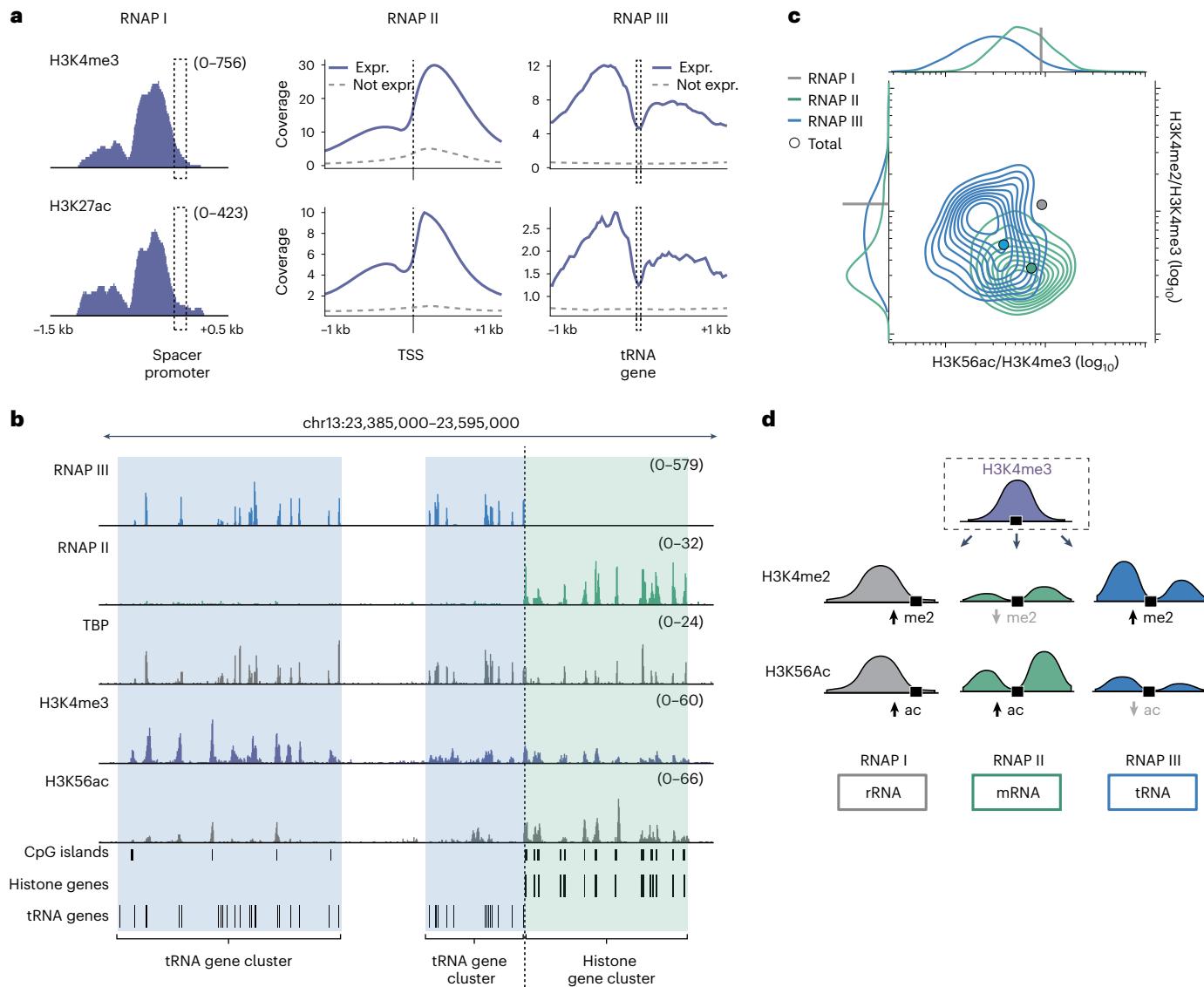
Fig. 6 | Distinct chromatin signatures define the promoters of each RNAP.
a, Comparison of H3K4me3 and H3K27ac profiles at the promoters of RNAP I, II and III genes. The profile over RNAP I genes is displayed over the rDNA spacer promoter (left), while profiles over RNAP II and III genes are displayed as metaplots across active (blue) and inactive (dashed gray) promoters. Expr., expressed. **b**, Visualization of RNAP II and RNAP III along with the shared TF TBP and histone modifications H3K4me3 and H3K56ac across a genomic region (mm10, chr13:23,385,000–23,595,000) containing a tRNA gene cluster (RNAP III-transcribed genes) adjacent to a histone gene cluster (RNAP II-transcribed genes), separated by a dashed line. **c**, Density distribution of H3K4me2/H3K4me3 versus H3K56ac/H3K4me3 ratios at RNAP I, active RNAP II and active RNAP III promoters. Points show ratios when computed using the total sum of histone coverage over all respective promoters. Marginal distributions are shown for RNAP II and III along x and y axes. Axes are $\log_{10}$ scaled. This plot compares the relative signals of the same antibodies within the same sample across distinct genomic regions corresponding to known promoters of RNAP I, II and III genes. **d**, Schematic showing relative levels of histone modifications H3K4me2 and H3K56ac at H3K4me3-enriched regions and the relative position of the associated RNAP promoter.

correspond to genomic regions that contain distinct TF and chromatin regulator binding profiles (Fig. 8c–f and Extended Data Fig. 9).

Active promoter-proximal elements. The first group (C1) is defined by H3K9ac and several other H3 acetylation marks (H3K14ac, H3K18ac, H3K36ac, H3K56ac and H3K79ac) (Fig. 8b). Genomic regions containing this signature tend to be localized near the promoter region of transcribed genes and are enriched for RNAP II, TFIIB and CpG island-associated factors (for example, E2F1, CXX1) (Fig. 8c,e,f).

Poised promoter-proximal elements. The second group (C2) contains high levels of H3K9ac and H2AZac (Fig. 8b). Genomic regions

containing this signature tend to have lower levels of RNAP II relative to C1 and are strongly enriched for Polycomb (JARID2, SUZ12, RING1B) and other repressive chromatin regulators (KDM2B, HDAC2) (Fig. 8e,f).

Stress and signaling response elements. The third group (C3) contains high levels of H2AZac and H4ac (Fig. 8b), is enriched for RNAP II and bound by p53, and contains other stress response motifs (for example, BACH1, NRF2) or signaling response motifs (for example, CRE) (Fig. 8c,e,f). Consistent with these observations, H2AZ has been proposed as a facilitator of inducible transcription (for example, signaling pathway responses and p53 regulation)[82–85].
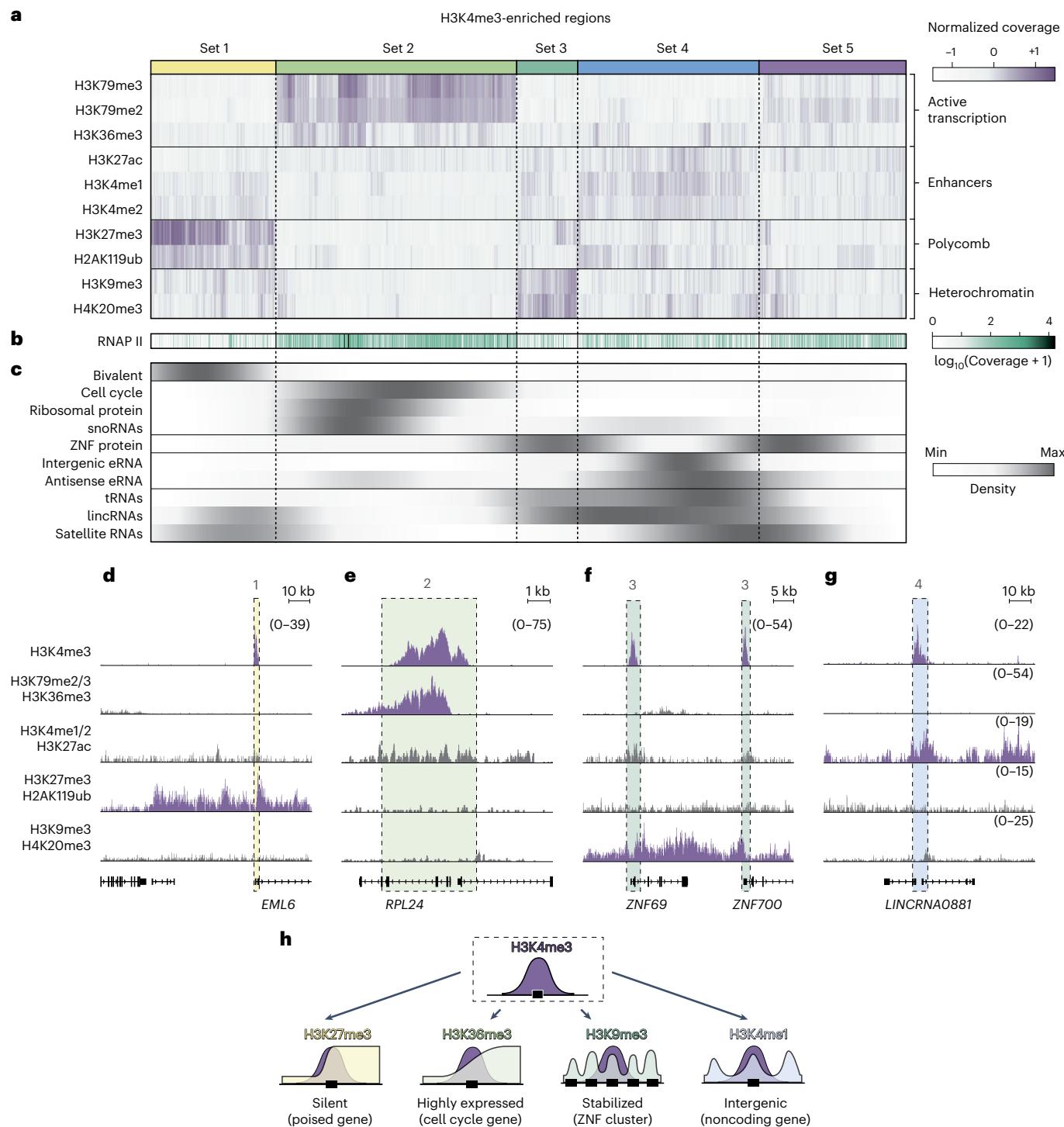
Fig. 7 | Combinations of histone modifications distinguish RNAP II promoter type, activity and potential. a, Hierarchically clustered heatmap of coverage levels of ten different histone modifications (*y* axis) at individual H3K4me3-enriched genomic regions (*x* axis). Five distinct clusters of regions are indicated by colored bars along the top axis. **b**, RNAP II coverage at H3K4me3-enriched regions, as sorted in **a**. **c**, Gene density of ten different gene classes at H3K4me3-enriched regions, as sorted in **a**. eRNA, enhancer RNA; lincRNA, long intergenic noncoding RNA. **d**, Visualization of H3K4me3 and H3K27me3–H2AK119ub (associated with cluster 1) across the *EML5* gene in K562 cells. **e**, Visualization of H3K4me3 and H3K79me2–H3K79me3–H3K36me3 colocalization (associated with cluster 2) across the ribosomal

protein gene *RPL24* in K562 cells. **f**, Visualization of H3K4me3 and H4K20me3–H3K9me3 colocalization (associated with cluster 3) across neighboring ZNF genes *ZNF69* and *ZNF700* in K562 cells. **g**, Visualization of H3K4me3 and H3K4me1–H3K4me2–H3K27ac (associated with cluster 4) across the long intergenic noncoding RNA gene *LNCRNA0881* in K562 cells. For tracks in **d**–**g**, the non-H3K4me3 tracks represent the sum of histone tracks associated with each set and are scaled to the maximum value across all panels. H3K4me3 tracks are scaled to the maximum for each panel. **h**, Schematic summarizing the co-occurring histone modifications at H3K4me3-enriched regions and their associated gene groups.
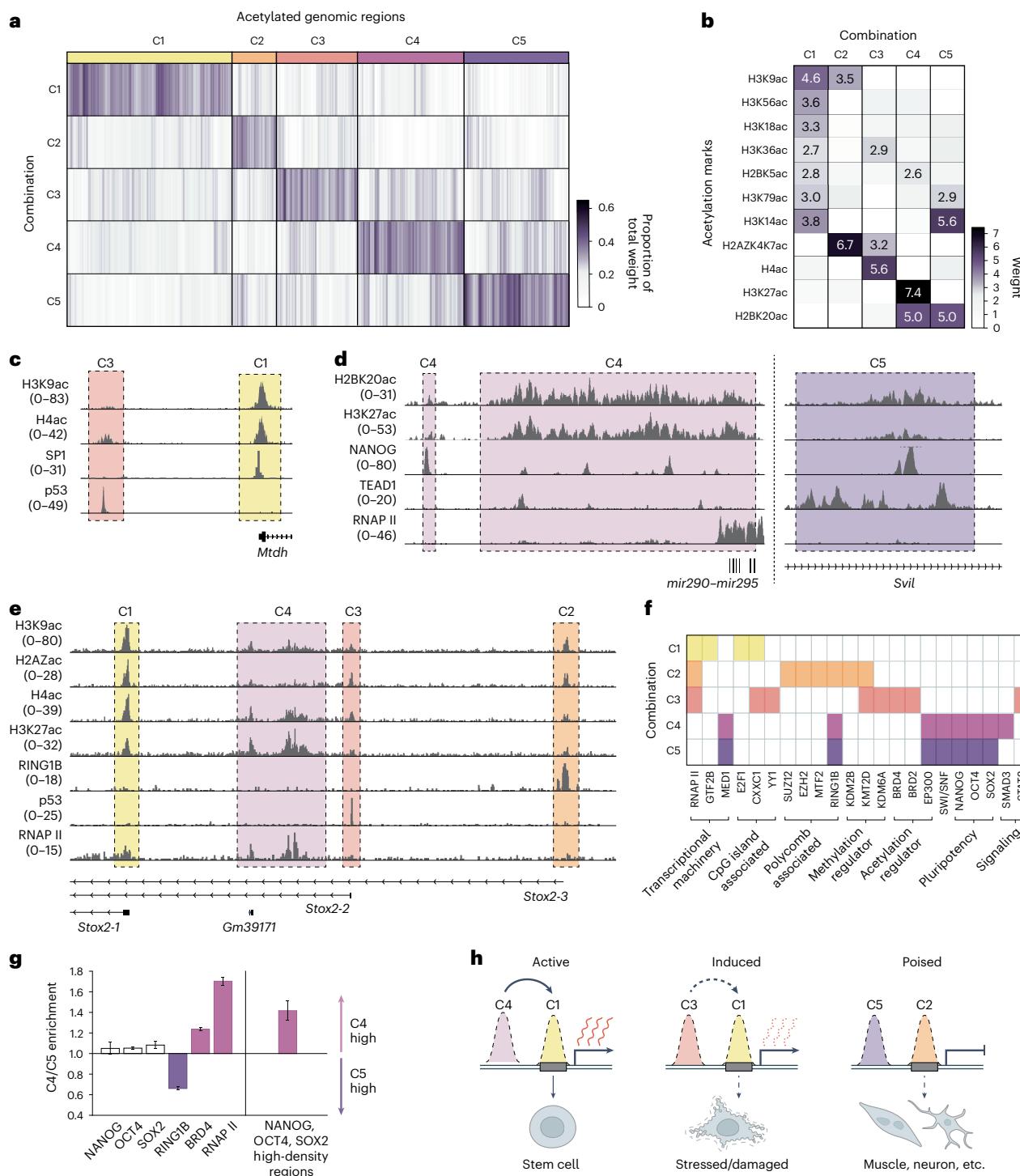
Fig. 8 | Distinct combinations of histone acetylation marks define unique enhancer types that differ in their activity and developmental potential. a, The relative weights of five different combinations of histone acetylation marks (C1–C5, *y* axis) for each acetylated genomic region (*x* axis). Regions are grouped according to the combination that received the greatest weight, and groups are indicated along the top axis. **b**, The relative weights of each histone acetylation mark (*y* axis) within each combination (*x* axis). Only weights greater than 2.5 are labeled. **c**, Visualization of H3K9ac and H4ac along with SP1 and p53 across a genomic region (mm10, chr15:34,065,000–34,086,000) containing enhancers assigned to the C1 (yellow) and C3 (red) states. **d**, Visualization of H2BK20ac and H3K27ac along with NANOG, TEAD1 and RNAP II across two genomic regions (left, mm10, chr7:3,191,500–3,221,500; right, mm10, chr18:5,006,500–5,016,500) containing enhancers assigned to C4 (left) and to

C5 (right), respectively (the scale of the NANOG track is capped to the maximum of the left region; TEAD1 data are from published ChIP-seq data from fetal cardiomyocytes[86]). **e**, Visualization of H3K9ac, H2AZac and H4ac along with RING1B, p53 and RNAP II over a genomic region (mm10, chr8:47,272,800–47,427,000) containing multiple isoforms of the gene *STOX2* and enhancers assigned to states C1–C4. **f**, DNA-associated proteins (*x* axis, ordered by function) with significant binding at genomic regions defined by each combination (*y* axis) are indicated in color (Methods). **g**, Bars show the enrichment value of selected transcription-associated factors or regions with a high density of pluripotency TFs (Supplementary Methods) in C4- versus C5-associated regions. Whiskers indicate the 5th and 95th percentiles from permutation-based resampling (*n* = 200 permutations) in which each permutation retained three-quarters of the C4 or C5 region. **h**, Schematic of C1–C5-associated regions and their corresponding functions.

Active pluripotency distal regulatory elements. The fourth group (C4) is defined by H2BK20ac and H3K27ac (Fig. 8b). These regions tend to be promoter distal (Extended Data Fig. 9b) and are associated with actively transcribed embryonic- and stem cell-specific genes (Fig. 8d). These regions are enriched for binding of pluripotency TFs, including NANOG, OCT4 and SOX2, as well as the p300 acetyltransferase and components of mediator (Fig. 8f).

Poised differentiation distal regulatory elements. The fifth group (C5) is defined by H2BK20ac and H3K14ac (Fig. 8b). These regions displayed TF and chromatin regulator occupancy similar to that of C4 regions[86] (Fig. 8f,g). However, in contrast to C4 regions, C5 regions bound by pluripotency factors correspond to enhancers of genes involved in post-embryonic development (Extended Data Fig. 10) and are enriched for sequence motifs of TFs involved in lineage specification and morphogenesis (for example, TEA domain TF (TEAD) family)[87] (Extended Data Fig. 9c). This suggests that C5 enhancers might be important in establishing the gene expression program needed upon differentiation (regulatory potential). Interestingly, we identified a third set of genomic regions that also contain a high density of pluripotency TFs but lack the C4 or C5 acetylation signatures; these are associated with genes involved in later stages of organogenesis (for example, kidney and sensory systems) (Extended Data Fig. 10).

These analyses indicate that histone acetylation is not a redundant marker of enhancers, but that combinations of acetylation modifications can define unique classes of *cis*-regulatory elements (promoter-proximal versus -distal enhancers) that act in distinct ways (stimulus responsive versus developmentally regulated) and that exhibit different activity (for example, active gene expression versus poised for activation upon differentiation) (Fig. 8h).

Discussion

We demonstrated that ChIP-DIP enables highly multiplexed mapping of hundreds of regulatory proteins to genomic DNA in a single experiment. Although the largest ChIP-DIP experiment in this study contained >225 distinct antibodies, this number was primarily limited by the availability of high-quality antibodies, and we expect that ChIP-DIP could profile larger pools of antibodies. Because this approach employs standard molecular biology techniques, we expect that it will be readily accessible to any laboratory without the need for specialized training or equipment. As such, we anticipate that ChIP-DIP will enable a fundamental shift from large consortia generating reference maps for a limited number of cell types to individual laboratories generating cell type-specific maps within any specific experimental system of interest.

In recent years, several methods, including multi-CUT&Tag[24], MulTI-Tag[88], MAbID[23], NTT-seq[89], Nano-CT[90] and uCoTarget[91], have made it possible to simultaneously profile multiple proteins. Yet, these multiplexed methods have two limitations: (1) they are limited in scale (two to six proteins simultaneously), and (2) they primarily map histone modifications and other abundant proteins but cannot map most chromatin regulators and TFs (protein diversity)[22,24,26]. ChIP-DIP overcomes both limitations by (1) generating high-quality datasets within pools containing >160 distinct antibodies and (2) mapping distinct protein types including histone modifications, chromatin regulators and TFs.

Given the important information encoded within quantitative combinations of histone modifications, chromatin regulators and TFs, comprehensively mapping these factors across cell types will be critical for studying gene regulation and for defining the putative effects of genetic variants associated with human disease. For instance, while specific regulatory states have been shown to be encoded by combinations of histone modifications (for example, bivalent domains), the number and diversity of such states have remained largely unexplored. The large number of chromatin proteins has necessitated a tradeoff between mapping many marks in a few cell types or a few marks in many cell types. ChIP-DIP overcomes this by mapping hundreds of proteins in

a single experiment. Moreover, due to the nature of split-pool barcoding used in ChIP-DIP and because there is negligible antibody–bead–chromatin dissociation during the procedure, ChIP-DIP can also be used to map protein binding within multiple samples simultaneously using distinct sets of antibody–oligonucleotide-labeled beads. In addition to the increase in scale provided by mapping multiple proteins and samples simultaneously, ChIP-DIP multiplexing also reduces many sources of technical and biological variability associated with processing individual proteins and samples. This ability will enable large-scale mapping of dynamic protein localization across distinct cell types and time points.

Beyond the applications highlighted in this work, ChIP-DIP can be directly integrated into existing split-pool approaches to create additional capabilities. For example, we previously showed that we can map the 3D genome structure surrounding individual protein binding sites (SIP)[92]; integrating this with ChIP-DIP will enable mapping of 3D structure at hundreds of distinct binding sites simultaneously. Moreover, we previously developed a method to map 3D genome contacts within thousands of individual single cells using this same split-pool approach[93]. Integrating this approach with ChIP-DIP will enable comprehensive mapping of hundreds of regulatory binding sites within thousands of individual cells. Finally, we previously used split-and-pool barcoding to simultaneously map the spatial proximity of DNA and RNA and measure noncoding RNA localization and the levels of nascent RNA transcription at individual DNA sites[94]. Integrating this with ChIP-DIP will enable simultaneous measurement of protein binding and transcriptional activity at individual genomic locations, providing a direct link between binding events and the associated transcription activity. For these reasons, we expect that ChIP-DIP will represent a transformative tool for dissecting gene regulation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-024-02000-5.

References

1. Bednar, J. et al. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl Acad. Sci. USA* **95**, 14173–14178 (1998).
2. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
3. Huang, H., Sabari, B. R., Garcia, B. A., Allis, C. D. & Zhao, Y. SnapShot: histone modifications. *Cell* **159**, 458 (2014).
4. Tekel, S. J. & Haynes, K. A. Molecular structures guide the engineering of chromatin. *Nucleic Acids Res.* **45**, 7555–7570 (2017).
5. Mashtalir, N. et al. Chromatin landscape signals differentially dictate the activities of mSWI/SNF family complexes. *Science* **373**, 306–315 (2021).
6. He, S. et al. Structure of nucleosome-bound human BAF complex. *Science* **367**, 875–881 (2020).
7. Kundaje, A. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
8. Barba-Aliaga, M., Alepuz, P. & Pérez-Ortín, J. E. Eukaryotic RNA polymerases: the many ways to transcribe a gene. *Front. Mol. Biosci.* **8**, 663209 (2021).
9. Roeder, R. G. Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb. Symp. Quant. Biol.* **63**, 201–218 (1998).
10. Malik, S. & Roeder, R. G. Regulation of the RNA polymerase II pre-initiation complex by its associated coactivators. *Nat. Rev. Genet.* **24**, 767–782 (2023).

11. Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463**, 474–484 (2010).

12. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of in vivo protein–DNA interactions. *Science* **316**, 1497–1502 (2007).

13. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).

14. Barski, A. et al. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).

15. Robertson, G. et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657 (2007).

16. He, Q., Johnston, J. & Zeitlinger, J. ChIP-nexus enables improved detection of in vivo transcription factor binding footprints. *Nat. Biotechnol.* **33**, 395–401 (2015).

17. Serandour, A. A., Brown, G. D., Cohen, J. D. & Carroll, J. S. Development of an Illumina-based ChIP-exonuclease method provides insight into FoxA1–DNA binding properties. *Genome Biol.* **14**, R147 (2013).

18. Tehranchi, A. K. et al. Pooled ChIP–seq links variation in transcription factor binding to complex disease risk. *Cell* **165**, 730–741 (2016).

19. Aldridge, S. et al. AHT-ChIP–seq: a completely automated robotic protocol for high-throughput chromatin immunoprecipitation. *Genome Biol.* **14**, R124 (2013).

20. Janssens, D. H. et al. Automated CUT&Tag profiling of chromatin heterogeneity in mixed-lineage leukemia. *Nat. Genet.* **53**, 1586–1596 (2021).

21. Kaya-Okur, H. S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).

22. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).

23. Lochs, S. J. A. et al. Combinatorial single-cell profiling of major chromatin types with MAbID. *Nat. Methods* **21**, 72–82 (2024).

24. Gopalan, S., Wang, Y., Harper, N. W., Garber, M. & Fazzio, T. G. Simultaneous profiling of multiple chromatin proteins in the same cells. *Mol. Cell* **81**, 4736–4746 (2021).

25. Gopalan, S. & Fazzio, T. G. Multi-CUT&Tag to simultaneously profile multiple chromatin factors. *STAR Protoc.* **3**, 101100 (2022).

26. 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 (2020).

27. Kong, N. R., Chai, L., Tenen, D. G. & Bassal, M. A. A modified CUT&RUN protocol and analysis pipeline to identify transcription factor binding sites in human cell lines. *STAR Protoc.* **2**, 100750 (2021).

28. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

29. PsychENCODE Consortium et al. The PsychENCODE project. *Nat. Neurosci.* **18**, 1707–1712 (2015).

30. The Immunological Genome Project Consortium et al. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).

31. Partridge, E. C. et al. Occupancy maps of 208 chromatin-associated proteins in one human cell type. *Nature* **583**, 720–728 (2020).

32. He, Y. et al. Spatiotemporal DNA methylome dynamics of the developing mouse fetus. *Nature* **583**, 752–759 (2020).

33. Sisu, C. et al. Transcriptional activity and strain-specific history of mouse pseudogenes. *Nat. Commun.* **11**, 3695 (2020).

34. Chasman, D. & Roy, S. Inference of cell type specific regulatory networks on mammalian lineages. *Curr. Opin. Syst. Biol.* **2**, 130–139 (2017).

35. Ota, M. et al. Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. *Cell* **184**, 3006–3021 (2021).

36. Madhani, H. D. et al. Epigenomics: a roadmap, but to where? *Science* **322**, 43–44 (2008).

37. Kidder, B. L., Hu, G. & Zhao, K. ChIP–seq: technical considerations for obtaining high-quality data. *Nat. Immunol.* **12**, 918–922 (2011).

38. Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757 (2018).

39. Quinodoz, S. A. et al. RNA promotes the formation of spatial compartments in the nucleus. *Cell* **184**, 5775–5790 (2021).

40. Quinodoz, S. A. et al. SPRITE: a genome-wide method for mapping higher-order 3D interactions in the nucleus using combinatorial split-and-pool barcoding. *Nat. Protoc.* **17**, 36–75 (2022).

41. Kim, S., Yu, N.-K. & Kaang, B.-K. CTCF as a multifunctional protein in genome regulation and gene expression. *Exp. Mol. Med.* **47**, e166 (2015).

42. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).

43. Girbig, M., Misiaszek, A. D. & Müller, C. W. Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nat. Rev. Mol. Cell Biol.* **23**, 603–622 (2022).

44. Abascal, F. et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).

45. Adli, M., Zhu, J. & Bernstein, B. E. Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors. *Nat. Methods* **7**, 615–618 (2010).

46. Karimzadeh, M. & Hoffman, M. M. Virtual ChIP–seq: predicting transcription factor binding by learning from the transcriptome. *Genome Biol.* **23**, 126 (2022).

47. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).

48. Spicuglia, S. & Vanhille, L. Chromatin signatures of active enhancers. *Nucleus* **3**, 126–131 (2012).

49. Steger, D. J. et al. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol. Cell. Biol.* **28**, 2825–2839 (2008).

50. Gates, L. A., Foulds, C. E. & O'Malley, B. W. Histone marks in the 'driver's seat': functional roles in steering the transcription cycle. *Trends Biochem. Sci.* **42**, 977–989 (2017).

51. Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H. & Tora, L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* **13**, 424 (2012).

52. Chen, Z., Djekidel, M. N. & Zhang, Y. Distinct dynamics and functions of H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. *Nat. Genet.* **53**, 551–563 (2021).

53. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3 (2015).

54. Chen, T. & Dent, S. Y. R. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat. Rev. Genet.* **15**, 93–106 (2014).

55. Kirtana, R., Manna, S. & Patra, S. K. Molecular mechanisms of KDM5A in cellular functions: facets during development and disease. *Exp. Cell Res.* **396**, 112314 (2020).

56. Shilatifard, A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20**, 341–348 (2008).

57. Geng, Z. & Gao, Z. Mammalian PRC1 complexes: compositional complexity and diverse molecular mechanisms. *Int. J. Mol. Sci.* **21**, 8594 (2020).

58. Mierlo, G., van, Veenstra, G. J. C., Vermeulen, M. & Marks, H. The complexity of PRC2 subcomplexes. *Trends Cell Biol.* **29**, 660–671 (2019).

59. Bosch-Presegué, L. et al. Mammalian HP1 isoforms have specific roles in heterochromatin structure and organization. *Cell Rep.* **21**, 2048–2057 (2017).

60. Mazzocca, M., Colombo, E., Callegari, A. & Mazza, D. Transcription factor binding kinetics and transcriptional bursting: what do we really know? *Curr. Opin. Struct. Biol.* **71**, 239–248 (2021).

61. Bartman, C. R. et al. Transcriptional burst initiation and polymerase pause release are key control points of transcriptional regulation. *Mol. Cell* **73**, 519–532 (2019).

62. Rada-Iglesias, A. et al. Whole-genome maps of USF1 and USF2 binding and histone H3 acetylation reveal new aspects of promoter structure and candidate genes for common human disorders. *Genome Res.* **18**, 380–392 (2008).

63. O'Connor, L., Gilmour, J. & Bonifer, C. The role of the ubiquitously expressed transcription factor Sp1 in tissue-specific transcriptional regulation and in disease. *Yale J. Biol. Med.* **89**, 513–525 (2016).

64. Li, Z., Cogswell, M., Hixson, K., Brooks-Kayal, A. R. & Russek, S. J. Nuclear respiratory factor 1 (NRF-1) controls the activity dependent transcription of the GABA-A receptor β1 subunit gene in neurons. *Front. Mol. Neurosci.* **11**, 285 (2018).

65. Horn, H. F. & Vousden, K. H. Coping with stress: multiple ways to activate p53. *Oncogene* **26**, 1306–1316 (2007).

66. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36**, 3943–3956 (2017).

67. Akberdin, I. R. et al. Pluripotency gene network dynamics: system views from parametric analysis. *PLoS ONE* **13**, e0194464 (2018).

68. Reith, W. et al. MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain. *Genes Dev.* **4**, 1528–1540 (1990).

69. Brivanlou, A. H. & Darnell, J. E. Signal transduction and the control of gene expression. *Science* **295**, 813–818 (2002).

70. Satoh, J., Kawana, N. & Yamamoto, Y. Pathway analysis of ChIP–seq-based NRF1 target genes suggests a logical hypothesis of their involvement in the pathogenesis of neurodegenerative diseases. *Gene Regul. Syst. Biol.* **7**, GRSB.S13204 (2013).

71. Qi, B., Newcomer, R. & Sang, Q.-X. ADAM19/adamalysin 19 structure, function, and role as a putative target in tumors and inflammatory diseases. *Curr. Pharm. Des.* **15**, 2336–2348 (2009).

72. Schoch, S., Cibelli, G. & Thiel, G. Neuron-specific gene expression of synapsin I. Major role of a negative regulatory mechanism. *J. Biol. Chem.* **271**, 3317–3323 (1996).

73. Martin, D. & Grapin-Botton, A. The importance of REST for development and function of beta cells. *Front. Cell Dev. Biol.* **5**, 12 (2017).

74. Bao, F., LoVerso, P. R., Fisk, J. N., Zhurkin, V. B. & Cui, F. p53 binding sites in normal and cancer cells are characterized by distinct chromatin context. *Cell Cycle* **16**, 2073–2085 (2017).

75. Otto, S. J. et al. A new binding motif for the transcriptional repressor REST uncovers large gene networks devoted to neuronal functions. *J. Neurosci.* **27**, 6729–6739 (2007).

76. Garber, M. et al. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* **47**, 810–822 (2012).

77. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* **28**, 817–825 (2010).

78. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).

79. Wang, H. et al. H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature* **615**, 339–348 (2023).

80. Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* **23**, 2484–2489 (2009).

81. Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* **20**, 259–266 (2013).

82. Giaimo, B. D. et al. Histone variant H2A.Z deposition and acetylation directs the canonical Notch signaling response. *Nucleic Acids Res.* **46**, 8197–8215 (2018).

83. Giaimo, B. D., Ferrante, F., Herchenröther, A., Hake, S. B. & Borggrefe, T. The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* **12**, 37 (2019).

84. Gévry, N., Chan, H. M., Laflamme, L., Livingston, D. M. & Gaudreau, L. p21 transcription is regulated by differential localization of histone H2A.Z. *Genes Dev.* **21**, 1869–1881 (2007).

85. Gévry, N. et al. Histone H2A.Z is essential for estrogen receptor signaling. *Genes Dev.* **23**, 1522–1533 (2009).

86. Akerberg, B. N. et al. A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. *Nat. Commun.* **10**, 4907 (2019).

87. Currey, L., Thor, S. & Piper, M. TEAD family transcription factors in development and disease. *Development* **148**, dev196675 (2021).

88. Meers, M. P., Llagas, G., Janssens, D. H., Codomo, C. A. & Henikoff, S. Multifactorial profiling of epigenetic landscapes at single-cell resolution using MulTI-Tag. *Nat. Biotechnol.* **41**, 708–716 (2023).

89. Stuart, T. et al. Nanobody-tethered transposition enables multifactorial chromatin profiling at single-cell resolution. *Nat. Biotechnol.* **41**, 806–812 (2023).

90. Bartosovic, M., Kabbe, M. & Castelo-Branco, G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat. Biotechnol.* **39**, 825–835 (2021).

91. Xiong, H., Wang, Q., Li, C. C. & He, A. Single-cell joint profiling of multiple epigenetic proteins and gene transcription. *Sci. Adv.* **10**, eadi3664 (2024).

92. Vangala, P. et al. High-resolution mapping of multiway enhancer–promoter interactions regulating pathogen detection. *Mol. Cell* **80**, 359–373 (2020).

93. Arrastia, M. V. et al. Single-cell measurement of higher-order 3D genome organization with scSPRITE. *Nat. Biotechnol.* **40**, 64–73 (2022).

94. Goronzy, I. N. et al. Simultaneous mapping of 3D structure and nascent RNAs argues against nuclear compartments that preclude transcription. *Cell Rep.* **41**, 111730 (2022).

Methods

Cells, cell culture and cross-linking

Cell lines. Two cell lines were used: (1) female mESCs (pSM44 mESC line) derived from a 129 × castaneous F_1 mouse cross and (2) K562, a female human lymphoblastic cell line (ATCC, CCL-243).

Primary cells. mDCs were derived from bone marrow collected from 6–8-week-old female C57BL6 mice[95] (Supplementary Methods). mDCs were stimulated with 100 ng ml^{-1} LPS (rough, ultra-pure *Escherichia coli* K12 strain, Invitrogen) and collected at 0 h, 6 h and 24 h after treatment, as previously described[95].

Cell cross-linking. Cells were cross-linked in suspension with 1% formaldehyde for 10 min at room temperature.

ChIP-DIP: bead preparation

Protein G bead biotinylation. Protein G Dynabeads (Invitrogen, 10003D) were incubated with EZ-Link Sulfo-NHS-Biotin (Thermo Scientific, 21217), and the NHS reaction was quenched with 1 M Tris, pH 7.4.

Preparation of streptavidin-coupled oligonucleotides. Biotinylated antibody ID oligonucleotides were coupled to streptavidin (BioLegend, 280302) in a 96-well PCR plate.

Preparation of oligonucleotide-labeled protein G beads. Ten microliters of biotinylated beads were aliquoted into individual wells of a deep-well 96-well plate (Nunc 96-Well DeepWell Plates with Shared-Wall Technology, Thermo Scientific, 260251), and 14 µl of 5.675 nM streptavidin-coupled oligonucleotide was added.

Antibody coupling. Antibody (2.5 µg) was added to each well of the 96-well plate.

Preparation of the bead pool. Beads were pooled using equal amounts of prepared beads for each antibody (10 µl beads per antibody) or titrated based on the determined chromatin pulldown efficiency measured in QC experiments (Supplementary Methods).

ChIP-DIP: immunoprecipitation, split-and-pool and library preparation

The pool of labeled beads was added to lysate, incubated for 1 h at room temperature and then washed. To blunt end and phosphorylate double-stranded DNA, the NEBNext End Repair Module (NEB, E6050L; containing T4 DNA polymerase and T4 PNK) was used. Split-and-pool barcoding was performed as previously described[40], with modifications described in Supplementary Methods. After split-and-pool barcoding was complete, beads were resuspended in 1 ml proteinase K buffer, digested with proteinase K (NEB) and reverse cross-linked at 65 °C overnight. DNA from each reverse cross-linked aliquot was isolated and amplified for 9–12 cycles using the Q5 Hot-Start High-Fidelity 2× Mastermix (NEB, M0294L) and primers that added the full Illumina adaptor sequences. Sequencing was performed on the Illumina NovaSeq S4, NextSeq or AVITI (Element Biosciences).

Data-processing pipeline

Reads were split into two files, one for antibody ID reads and one for DNA reads, based on the presence of 'BPM' (bead tag) or 'DPM' (DNA tag), respectively, in read 1. For DNA reads, the DPM sequence was trimmed and aligned to mm10 or hg38 using Bowtie 2 (version 2.3.5)[96] with default parameters. For antibody ID reads, the BPM sequence was trimmed, and the UMI was extracted from the remaining sequence. A 'cluster file' was generated by aggregating all reads that share the same split-and-pool barcode sequence. Individual clusters in the 'cluster file' were assigned to a specific antibody based on antibody ID reads within the cluster (see Supplementary Methods for assignment details).

Genomic DNA alignments were split into separate BAM files, one per antibody, based on cluster assignment.

Visualization and peak calling

BigWig files were generated from each antibody-specific BAM file using the 'bamCoverage' function from deepTools version 3.1.3 (ref. 97) and were visualized with IGV[98]. For normalization, a background model was generated for each individual antibody using the total pool of assigned sequencing reads (Supplementary Methods). Background-normalized tracks were generated using the scaled background distribution. Track visualizations are scaled to the maximum over the region, and scales indicate reads per bin, unless indicated otherwise. Peaks were called using the HOMER version 4.11 (ref. 99) program 'findPeaks' on tag directories generated for target datasets. Background-normalized peaks were generated using the scaled background distribution as input. TF motifs were predicted using the HOMER program 'findMotifsGenome' with the parameters '-s 200 –mask –len 10' on peaks generated as described above.

Heatmaps, summary plots and other graphical visualizations

Genome-wide metaplots, pairwise scatterplots and correlation heatmaps were generated and visualized using deepTools version 3.1.3, a suite of Python tools designed for efficient analysis of high-throughput sequencing data. We used the following functions: 'multiBamSummary', 'multiBigWigSummary', 'plotCorrelation', 'plotCoverage', 'multiBamCoverage' and 'computeMatrix'. All other graphical visualizations (for example, line plots, violin plots, etc.) were generated using seaborn version 0.13.2, a Python data visualization library.

ChIP-DIP experiments

We performed 11 ChIP-DIP experiments in this paper, each of which, along with the associated antibodies, proteins and statistics, is described in Supplementary Table 1. All ChIP-DIP experiments were performed using the same general protocol with a few experiment-specific modifications described in detail in Supplementary Methods.

Comparison with ENCODE data

ChIP-DIP comparisons with ENCODE-generated ChIP–seq data in Fig. 1 were made using the K562 ten-antibody pool experiment. Genome-wide coverage comparisons were calculated across all RefSeq TSSs for H3K4me3 and POLR2A or across 10-kb bins for CTCF and H3K27me3. Calculations were performed using 'multiBamSummary' and plotted as 2D kernel density plots.

For all ChIP-DIP K562 datasets, comparisons with ENCODE-generated ChIP–seq were made for all targets for which ENCODE datasets were available. ENCODE accession numbers are listed in Supplementary Methods. The Pearson correlation coefficients of genome-wide coverage comparisons were calculated at 1,000 bp using 'multiBigwigSummary', and the fraction of overlapping peaks is reported in Supplementary Table 2.

Comparison with CUT&Tag data

ChIP-DIP data from the K562 35-antibody experiment were compared with data from ChIPmentation[100] and high-throughput CUT&Tag performed using various starting cell numbers (10,000, 100,000, 500,000 or 10 million)[21]. SRA accession numbers are listed in Supplementary Methods. The estimated library complexity was calculated using the 'lc_extrap' function from preseq version 3.2.0 (ref. 101). Fraction of reads in peaks (FRIP) scores were calculated for all samples with at least 100 called peaks using the intersect function from BEDTools version 2.29.2 (ref. 102).

Transcription factor peak comparison

Peak sites from ChIP-DIP TF data were compared to reference binding sites retrieved from ReMap2022, a database of transcriptional

regulator peaks derived from curated ChIP–seq, ChIP-exo and DAP-seq experiments[103]. To ensure that only high-quality datasets are included, ReMap2022 implements four different ENCODE-defined quality metrics. Only TFs with a minimum FRIP score of 0.5% and for which reference data were available were analyzed. Results are reported in Supplementary Table 4.

Pool size comparison analysis
To measure the influence of the number of antibodies contained within an individual pool, read coverage profiles of four targets (H3K4me3, H3K27me3, CTCF and RNAP II) generated in four different ChIP-DIP experiments in K562 cells (ten-, 35-, 50- and 52-antibody pools) or generated by ENCODE (one-antibody pool) were compared. For both RNAP II and CTCF, two different antibodies were included. Coverage of normalized BigWig files across the set of all peak regions from the ten-antibody pool experiment was calculated using 'multiBigwigSummary'. Pearson correlation coefficients for all pairs were calculated and plotted as a heatmap using 'plotCorrelation', manually ordering the rows and columns from smallest to largest pool size for each target.

Peak overlaps were calculated for each target between experiments of different pool sizes as the (number of peaks in experiment 1 intersecting peaks in experiment 2)/(total number of peaks in experiment 1). The numbers of intersecting peaks were calculated using BEDTools version 2.29.2 (ref. [102]) and are reported in Supplementary Table 3.

To ensure that the results were robust to read coverage, the analysis was repeated after downsampling. Specifically, the target-separated BAM files from each pool size were downsampled to an equal number of reads for each target: H3K4me3, 9 million reads; CTCF, 2 million reads; H3K27me3, 15 million reads; RNAP II, 1 million reads. Genome-wide correlation of BigWig profiles was calculated as described above.

Cell lysate amount comparison analysis
To measure the amount of cell input material required for ChIP-DIP, we performed a series of ChIP-DIP experiments using the same antibody pool and differing amounts of cell lysate (50,000, 500,000, 5,000,000 or 45,000,000 cells) (Supplementary Methods). Read coverage profiles of four targets (H3K4me3, H3K27me3, CTCF and RNAP II) were compared across different levels of input cell lysate. For both RNAP II and CTCF, two different antibodies were included, coverage comparison was performed identical to that described for 'Pool size comparison analysis'. Peak overlaps were calculated for each antibody between pairs of experiments as described above. For target and condition pairs with sufficient read depth, the estimated library complexity was calculated using the 'lc_extrap' function from preseq version 3.2.0.

Histone modification diversity analysis
Chromatin state. Genome-wide coverage for 10-kb windows for 12 histone marks (H3K27me3, H2AK119ub, H3K9me3, H4K20me3 and H3K9me3 from the 5 million (5M) condition in the K562 35-antibody pool experiment; H3K79me2, H3K79me1, H3K4me3, H3K4me2, H3K4me1, H3K9ac and H3K27ac from the K562 50-antibody pool experiment) was calculated using 'multiBamCoverage'. These values were standardized for each mark by transforming into z-score values. The UMAP reduction was generated using the UMAP[104] Python package and parameters n_components = 2 and n_neighbors = 3.

Heterochromatin-associated histone modifications. Validation of heterochromatin-associated histone modifications used the 5M condition in the K562 35-antibody pool experiment. Read coverage of H3K9me3, H4K20me3 and H3 was computed over annotation groups (ZNFs, LTRs, LINES, SINES, TSS ± 2 kb) using the 'depth' function from SAMtools version 1.9 (ref. [105]). An enrichment score was calculated by normalizing for feature and target abundance (Supplementary Methods).

Promoter-associated histone modifications. Validation of promoter-associated histone modifications used the mESC 67-antibody pool experiment. Promoter coverage correlations were calculated across promoters from EPDNew[106], a database of non-redundant eukaryotic RNAP II promoters, ±500 bp using 'multiBamSummary' and 'plotCorrelations'.

Gene body-associated histone modifications. Validation of gene body-associated histone modifications used the 5M condition in the K562 35-antibody pool experiment and the K562 50-antibody pool experiment. Values in coverage metaplots over the gene bodies of all protein-coding genes from the GENCODE[107] version 38 basic annotation were calculated using 'computeMatrix' and normalized to the maximum and the minimum for each target.

Chromatin regulator diversity analysis
Polycomb-associated chromatin regulators. Validation of Polycomb-associated chromatin regulators used the K562 50-antibody pool experiment. Metaplots respective to RING1B peak sites were calculated using 'computeMatrix'.

Heterochromatin-associated chromatin regulators. Validation of heterochromatin-associated chromatin regulators used the K562 50-antibody pool experiment. Genome-wide coverage for 10-kb windows and Pearson correlation coefficients were calculated using 'multiBigwigSummary' and 'plotCorrelation'.

H3K4me3-associated chromatin regulators. Analysis of H3K4me3-associated chromatin regulators used the mESC 165-antibody pool experiments. Binding profiles of JARID1A, RBBP5 and PHF8 were measured ±1 kb around the TSS of all representative promoters from EPDNew and were clustered using k-means clustering with $k = 4$ by 'plotCoverage'.

Polymerase diversity analysis
RNAP I, II and III comparison. Validation of the various RNAPs used the mESC 165-antibody pool experiment. First, read coverage within a ±100-bp window surrounding the promoters and TSSs of various gene groups was calculated. Next, for each polymerase, coverage was normalized to the total reads aligned with any gene group. Finally, an enrichment score of the relative coverage compared to an IgG isotype control was calculated and plotted as a bar graph.

RNAP II phosphorylation state comparison. Validation of the various RNAPs used the K562 52-antibody pool experiment. Values in the metaplots over the gene bodies of all protein-coding genes from the GENCODE version 38 basic annotation were calculated using 'computeMatrix'.

Mouse dendritic cell LPS stimulation time course analysis
Temporal pattern analysis. For genome-wide time course analysis of individual histone modifications, the read coverage per 100-kb bin genome wide for each target at each time point was calculated using 'multiBigwigSummary'. Next, for each target, for each time point pair (for example, 6 h and 0 h, 24 h and 6 h), the enriched bins were determined by finding the knee point of the summed coverage per bin-versus-rank graph using the Python package 'kneed'[108]. The scaled coverage per bin was calculated for each time point as $(x - \text{min})/(\text{max} - \text{min})$. The difference in coverage per bin was computed for enriched bins by subtracting the scaled coverage between time points.

For time course analysis of H3K27ac, enriched acetylated regions were determined by peak calling on the merged read file of histone acetylation at all time points using the 'findPeaks' function from HOMER with '−minDist 100000' and removing peaks with size <1,500 bp. Coverage of H3K27ac at each enriched region was computed using

'multiBamSummary' and normalized for read depth by dividing by the total reads for H3K27ac per respective time point. To focus on regions with higher coverage, regions with normalized read depth <0.05 CPM for all time points were removed from analysis. Coverage per time point was then rescaled to a minimum of 0 and a maximum of 1 using $(x - \text{min})/(\text{max} - \text{min})$ per region. Rescaled regions were clustered into three clusters using the 'kmeans' function from scipy.cluster version 1.13.0. See Supplementary Methods for subset criteria of 'activated' and 'repressed' regions.

Relationship with gene expression. RNA-seq data for LPS-stimulated mDCs at 0 h, 6 h and 24 h[109] were processed using the RNA-seq pipeline from Nextflow[110]. Transcript abundances across different time samples were normalized using quantile normalization and filtered to remove low- or non-expressed genes. H3K27ac-enriched regions were paired with the gene that had the maximum cosine similarity between H3K27ac coverage and gene expression versus time of all genes within a 100-kb window. Transcription levels of genes assigned to each of the three H3K27ac clusters were plotted as a violin plot using seaborn version 0.13.2 after removing duplicate genes and outlier values, defined as $x < Q_1 - 1.5 \times IQR$ or $x > Q_3 + 1.5 \times IQR$.

For comparison of enhancer versus promoter versus gene body, the coverage of promoter-associated histone marks in the window ±1,000 bp surrounding the TSS of expressed genes and the coverage of gene body-associated histone marks across the gene body of expressed genes were calculated using 'multiBamSummary'. Coverage was scaled by the length of the region and, for each histone mark, normalized by sequencing depth and rescaled to a minimum of 0 and a maximum of 1 using $(x - \text{min})/(\text{max} - \text{min})$. For each histone mark and each pair of time points $((t_0, t_6), (t_6, t_{24}), (t_0, t_{24}))$, the Spearman rank correlation coefficient between time points of histone coverage and transcription for each gene was calculated using the 'spearmanr' function from scipy.stats version 1.13.0.

Histone combinatorial analyses
Polymerase-associated histone profiles. For RNAP I, track coverage profiles of various histone modifications 1.5 kb upstream to 0.5 kb downstream of the spacer promoter were visualized using IGV.

For RNAP II, metaplots of coverage profiles for various histone modifications were generated around active and inactive RNAP II promoters using 'computeMatrix' (reference-point -a 1000 -b 1000) and 'plotProfile'. See Supplementary Methods for analysis involving bidirectional versus unidirectional promoters.

For RNAP III, metaplots of coverage profiles for various histone modifications were generated around active and inactive tRNA genes using 'computeMatrix' (scale-regions -a 1000 -b 1000 -m 75 -bs 25) and 'plotProfile'. tRNA genes were grouped into active or inactive categories based on the read coverage of RNAP III.

For comparison of relative histone levels, total coverage for each histone mark was calculated in the −1.5-kb to +0.5-kb window surrounding the spacer promoter for rDNA, the −0.5-kb to +0.5-kb window around active RNAP II promoters and the −0.5-kb to +0.5-kb window around active RNAP III tRNA gene promoters. To account for differences in window size, the coverage of H3K56ac and H3K4me2 was normalized to the level of H3K4me3.

H3K4me3-enriched region clustering. Combinatorial histone modification analysis for H3K4me3 regions used the 5M condition of the K562 35-antibody pool experiment. Read coverage of ten histone targets (H3K79me3, H3K79me2, H3K36me3, H3K4me1, H3K4me2, H3K27ac, H3K27me3, H2AK119ub, H3K9me3 and H4K20me3) was calculated over all H3K4me3 peak regions using the 'multicov' function of BEDTools version 2.29.2. The resulting region-versus-histone data matrix (A) was normalized using log normalization[111] (Supplementary Methods). The regions of the normalized data matrix were clustered using the cluster.

hierarchy.linkage function from SciPy version 1.6.2 (ref. [112]) with a Euclidean distance metric and a complete linkage method.

Gene annotation of H3K4me3 regions was performed using the 'annotatePeaks.pl' function from HOMER version 4.11. Definitions for each annotation group (ZNF genes, RP genes, lincRNA genes, snoRNA genes, satellite RNA genes, tRNA genes, cell cycle genes, bivalent genes and enhancer RNA regions) are provided in Supplementary Methods. To visualize enrichments of gene annotations in sets and subsets of the hierarchically clustered heatmap, the kernel density estimate (KDE) was calculated for each annotation group based on their clustering-defined order.

ChromHMM model of acetylation. The ChromHMM genome segmentation model was built using 15 different histone acetylation modifications measured in the mESC 67-antibody pool experiment. BAM files were binarized using the BinarizeBam function from ChromHMM with a Poisson threshold of 0.000001 and other default parameters. The signal threshold was increased from the default to remove spurious noise. State models with 5–20 states were built using the LearnModel function with default parameters. States were manually reordered and grouped based on transition probabilities between states. Nineteen states were selected for the final model to retain state 17, a state with a distinctive enrichment and transition profile.

Non-negative matrix factorization of acetylated regions. Non-negative matrix factorization analysis used the histone acetylation mark data from the mESC 67-antibody pool experiment. A normalized read coverage matrix of acetylation-enriched genomic regions (N) versus histone acetylation marks (M) was generated (Supplementary Methods). NMF was performed on this data matrix using 'NIMFA'[113], a Python library for non-negative matrix factorization, with the nndsvd initialization method. The rank k was selected empirically, taking into account the biological assignability of the resulting states, the complexity of the model and the stability of the factorization (the number of iterations the algorithm required to coverage). After factorization, the resulting basis matrix ($N \times k$) contained the coefficient of each combination i for each genomic region. A sorted heatmap of the basis matrix was generated by grouping the regions according to the combination that contributed the greatest coefficient for each region. For visualization, this heatmap was normalized by dividing the coefficients for each region by the total coefficient sum of the region. To profile and assign a biological interpretation to individual combinations, each region was assigned to the combination with the maximum coefficient. Identification of TFs with significant binding overlap with regions assigned to a single combination was performed using the Cistrome Data Browser, an interactive database of public ChIP–seq data[114]. Motif enrichment was calculated using the HOMER function 'findMotifs' on all genomic regions assigned to each combination. For comparison of enrichment levels in C4 versus C5, enrichments were calculated using bedgraphs from the mESC 165-antibody pool experiment and the ChromHMM program 'OverlapEnrichment' (java -jar ChromHMM.jar OverlapEnrichment -binres 1 -signal). Interval bars were generated by permutation-based resampling; enrichments were recalculated for 200 independent draws of 75% of the regions assigned to C4 or C5.

Statistics and reproducibility
Pearson correlation coefficients were calculated using the pearsonr function of scipy.stats version 1.13.0 (ref. [112]) or generated using the 'plotCorrelation' function from deepTools version 3.1.3 (ref. [97]). Spearman rank correlation coefficients were calculated using the 'spearmanr' function from scipy.stats version 1.13.0. The Mann–Whitney U-test was used to compare gene expression fold change following LPS stimulation between sets (Extended Data Fig. 6b) and was calculated using the 'mannwhitneyu' function from scipy.stats version 1.13.0. Statistical tests and distribution assumptions for peak calling are

intrinsic to the HOMER version 4.11 (ref. 99) peak-calling algorithm and commonly used and accepted for ChIP–seq data. *P* values for TF motifs were generated using HOMER version 4.11. Other statistical tests were performed using permutation or random sampling and make no implicit assumptions about distributions. Experimental details needed to reproduce individual ChIP-DIP experiments are provided in Supplementary Methods. Key proteins were mapped in multiple different experimental replicates and show comparable results.

Reporting summary

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

Data availability

All ChIP-DIP datasets generated in this study are available at GEO: GSE227773. Accession numbers for publicly available datasets used in this study are listed in Supplementary Methods.

Code availability

Publicly available software and packages were used in this study as indicated in Methods and Supplementary Methods. The original code for the ChIP-DIP pipeline is available on GitHub at https://github.com/GuttmanLab/chipdip-pipeline/tree/Paper (https://doi.org/10.5281/zenodo.13952458) (ref. 115).

References

95. Donnard, E. et al. Comparative analysis of immune cells reveals a conserved regulatory lexicon. *Cell Syst.* **6**, 381–394 (2018).
96. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
97. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
98. Robinson, J. T. et al. Integrative Genomics Viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
99. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
100. Schmidl, C., Rendeiro, A. F., Sheffield, N. C. & Bock, C. ChIPmentation: fast, robust, low-input ChIP–seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 (2015).
101. Daley, T. & Smith, A. D. Predicting the molecular complexity of sequencing libraries. *Nat. Methods* **10**, 325–327 (2013).
102. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
103. Hammal, F., Langen, P., de Bergon, A., Lopez, F. & Ballester, B. ReMap 2022: a database of human, mouse, *Drosophila and Arabidopsis* regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.* **50**, D316–D325 (2021).
104. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018).
105. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
106. Dreos, R., Ambrosini, G., Groux, R., Cavin Périer, R. & Bucher, P. The Eukaryotic Promoter Database in its 30th year: focus on non-vertebrate organisms. *Nucleic Acids Res.* **45**, D51–D55 (2017).
107. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
108. Satopää, V., Albrecht, J., Irwin, D. & Raghavan, B. Finding a 'kneedle' in a haystack: detecting knee points in system behavior. In *2011 31st International Conference on Distributed Computing Systems Workshops* 166–171 (IEEE, 2011).
109. Liang, K., Patil, A. & Nakai, K. Discovery of intermediary genes between pathways using sparse regression. *PLoS ONE* **10**, e0137222 (2015).
110. Tommaso, P. D. et al. Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* **35**, 316–319 (2017).
111. Kluger, Y., Basri, R., Chang, J. T. & Gerstein, M. Spectral biclustering of microarray data: coclustering genes and conditions. *Genome Res.* **13**, 703–716 (2003).
112. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
113. Zitnik, M. & Zupan, B. NIMFA: a Python library for nonnegative matrix factorization. *J. Mach. Learn. Res.* **13**, 849–853 (2012).
114. Zheng, R. et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res.* **47**, D729–D735 (2019).
115. Yeh, B. & Goronzy, I. GuttmanLab/chipdip-pipeline: *Nature Genetics* (2024) paper release (v1.0_publication). *Zenodo* https://doi.org/10.5281/zenodo.13952458 (2024).

Author contributions

A.A.P., M.R.B. and M.G. conceived ChIP-DIP; A.A.P. and M.R.B. developed ChIP-DIP; A.A.P., I.N.G. and J.K.G. optimized ChIP-DIP; A.A.P. and I.N.G. generated the data presented in this paper; C.S.L., O.E. and A.B. cultured, collected and treated cells; I.N.G. developed the computational pipeline; B.T.Y. generated the GitHub repository for the pipeline; I.N.G. performed data analysis and visualization; A.A.P., I.N.G. and M.G. generated figures and wrote the paper.

Competing interests

M.G., A.A.P., M.R.B., I.N.G. and J.K.G. are inventors of a submitted patent covering the ChIP-DIP method. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-024-02000-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-024-02000-5.

Correspondence and requests for materials should be addressed to Mitchell Guttman.

Peer review information *Nature Genetics* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at www.nature.com/reprints.
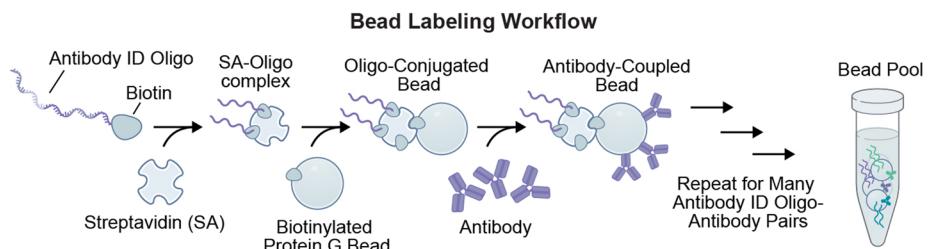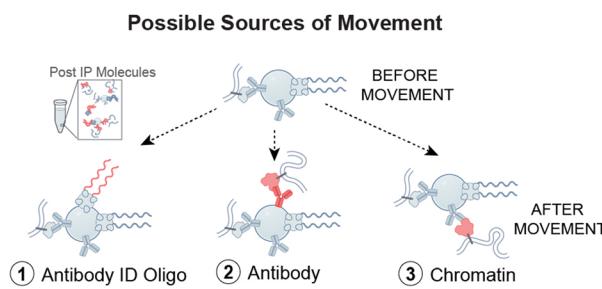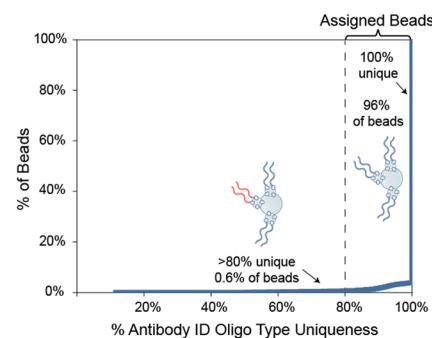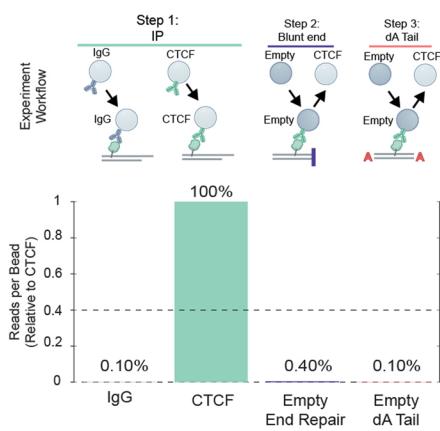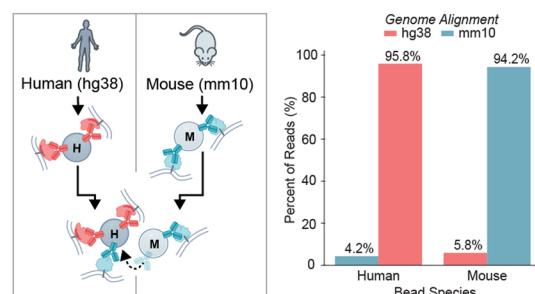
Extended Data Fig. 1 | Potential sources of mixing in ChIP-DIP. (a) Schematic of labeling strategy to generate Protein G beads coupled with a unique antibody-identifying oligonucleotide and a matched antibody. (i) Protein G beads are covalently modified with a biotin, (ii) oligonucleotides containing a 3' biotin are conjugated to streptavidin, (iii) oligo-streptavidin complexes are mixed with biotinylated protein G beads and (iv) protein G beads are mixed with antibodies. This process is repeated for each unique oligonucleotide-antibody pair and then all bead-antibody conjugates are pooled together. **(b)** Schematic of three potential sources of dissociation of chromatin-antibody-bead-oligo conjugates that could lead to mixing during ChIP-DIP: dissociation 1) between oligo and bead, 2) between antibody and bead, or 3) between antibody and chromatin. **(c)** If oligos dissociate from their original beads and bind to distinct beads (oligo-bead dissociation), we would expect multiple distinct oligo types on the same bead. To quantify this, we computed the percent uniqueness of oligo-types within each

split-pool cluster. The cumulative distribution of the uniqueness of antibody-ID oligos type (x-axis) within individual clusters is shown. **(d)** If antibodies dissociate from their original bead and reassociate with a different bead (antibody-bead dissociation), we expect that chromatin would associate with empty beads present in the experiment. We show a schematic of the experimental design to test for antibody movement between beads (top) and the quantification of reads per bead assigned to true targets (CTCF) or empty beads added during experimental processing steps (bottom). **(e)** If proteins (and their crosslinked chromatin) dissociate and reassociate to other beads containing the same epitope-specific antibodies (antibody-chromatin dissociation), we would expect that chromatin purified independently from human and mouse lysates would mix during the procedure. We show a schematic of the human-mouse mixing experimental design to test for chromatin movement (left) and quantification of species-specific reads assigned to human or mouse beads (right).
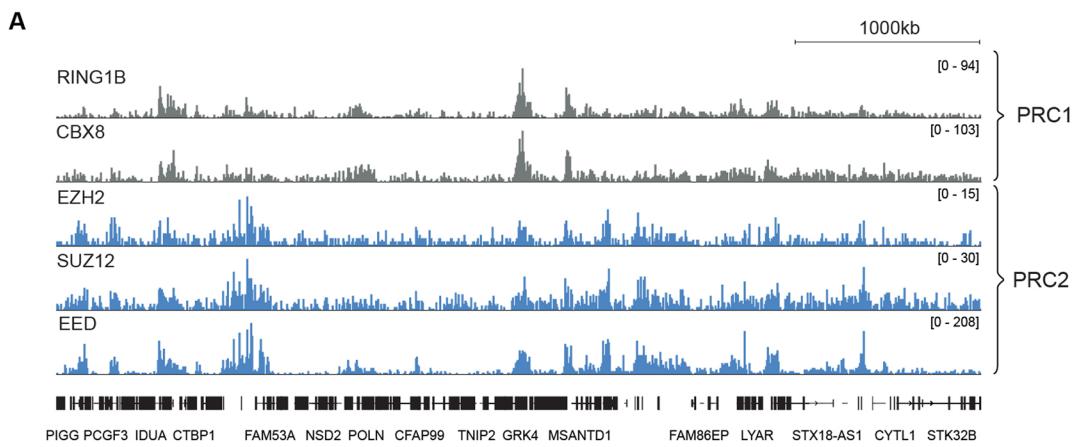
Extended Data Fig. 2 | Mapping multiple components of the same regulator complex within a single experiment. (a) Visualization of various components of the PRC1 (RING1B, CBX8) and PRC2 (EZH2, SUZ12, EED) complexes that were mapped within the same ChIP-DIP pool (K562 52 Antibody Pool) along a genomic region (hg38, chr4:500,000-5,500,000).
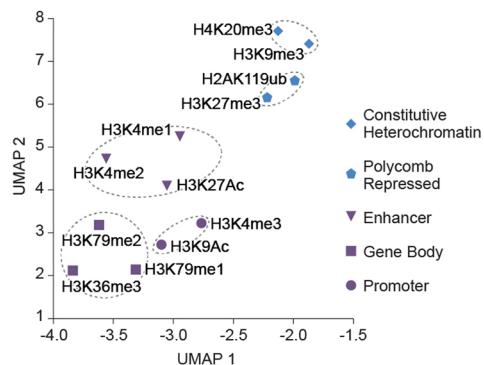
A Five Chromatin States   **B** Gene Body   **C** Active Promoter
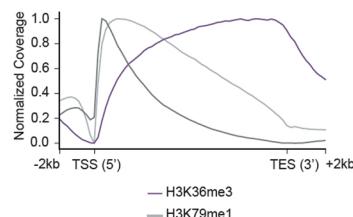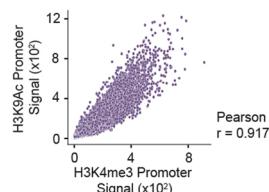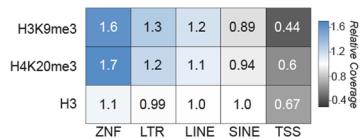
Extended Data Fig. 3 | Histone modifications associated with five chromatin states. (a) UMAP embedding of 12 histone modifications measured in K562 correspond to five chromatin states. **(b)** Metaplot of signal distribution of H3K36me3, H3K79me1 and H3K79me2 across the gene body of protein coding genes in K562. **(c)** Correlation scatterplot of H3K9Ac and H3K4me3 signals at promoter sites in mESC. **(d)** Enrichment heatmap of H3K9me3 and H4K20me3 at various associated (ZNF genes, LTRs, LINES) and unassociated (SINES, TSS) genomic elements in K562. H3 is shown as reference. For A-D, see Methods for details on ChIP-DIP experiments used for each analysis.

A H3K4me3 Readers/Writers/Erasers
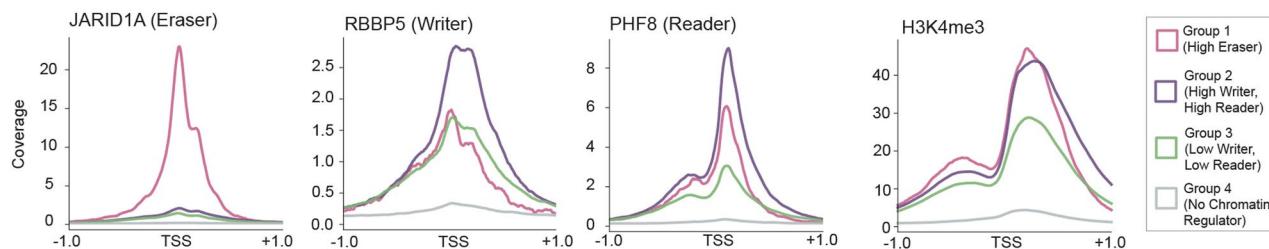
B H3K27me3 and PRC1/2
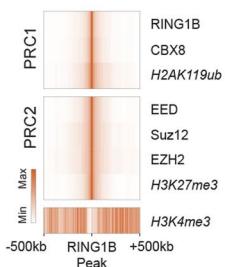
C H3K9me3 and Heterochromatin Readers
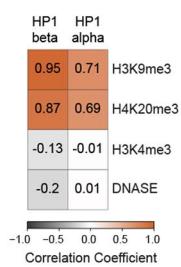
Extended Data Fig. 4 | Chromatin regulators co-localizing with known histone targets. (a) Metaplots of read coverage for three H3K4me3-associated chromatin regulators (JARID1A, RBBP5, PHF8) and H3K4me3 at four promoter groups in mESC. Promoter groups were identified using k-means clustering of CR signal. **(b)** Metaplot showing colocalization of multiple PRC1 and PRC2 members and their respective histone modifications at RING1B sites in K562. **(c)** Genome-wide correlation matrix of multiple HP1 proteins versus heterochromatin and euchromatin markers in K562. For A-C, see Methods for details on ChIP-DIP experiments used for each analysis.

A
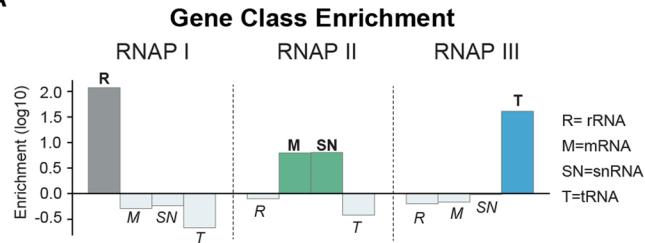
B
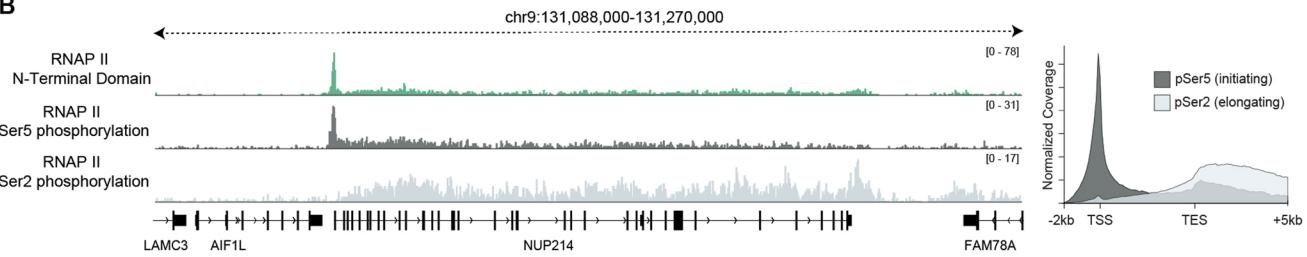
Extended Data Fig. 5 | Simultaneous mapping of distinct RNA polymerases and their isoforms. (a) Bar graph showing enrichment of gene class coverage (rRNA, mRNA, snRNA or tRNA) for RNAP I, II and III in mESC. For each RNAP, the bar of its associated class (or classes) is highlighted. **(b)** Visualization of RNAP II phosphorylation isoforms across the *NUP214* gene in K562 (left). Metaplot of signal distribution of RNAP II phosphorylation isoforms across the gene body of protein coding genes in K562 (right).
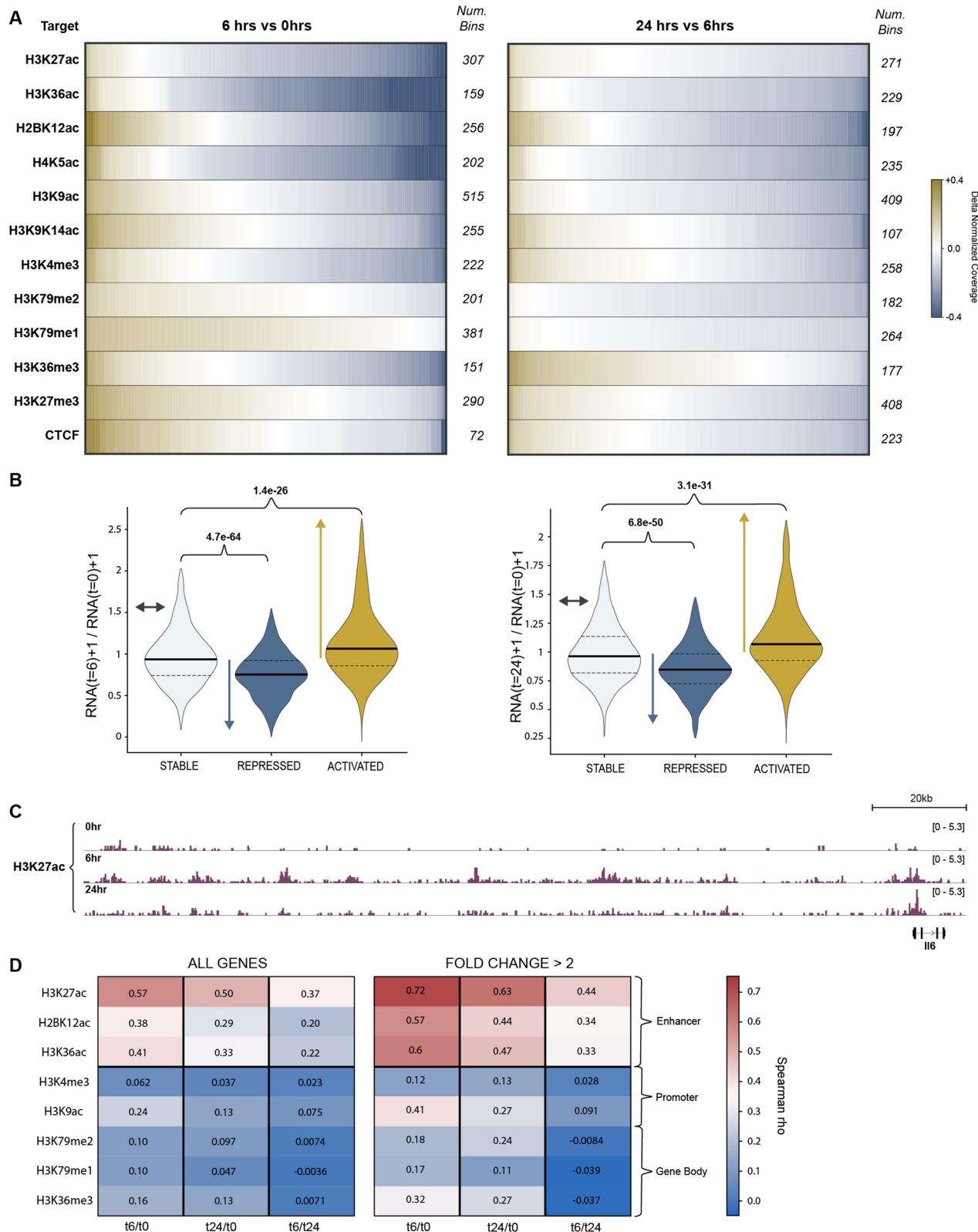
Extended Data Fig. 6 | Chromatin dynamics and the relationship to gene expression following LPS stimulation in mDCs. (a) Heatmap of change in normalized coverage per 100 kb bin for various mapped factors. For each factor, only enriched bins are shown and bins are sorted left-to-right by magnitude of change. **(b)** Violin plot of gene expression fold change for 6hrs vs 0hrs (left) and 24hrs vs 0hrs (right) grouped by sets of genes corresponding to sets of regions from Fig. 5C (see Methods). Shown are Mann-Whitney U test p-values. **(c)** Track visualization of H3K27ac at 0hrs, 6hrs and 24hrs across a genomic region (mm10, chr5:29,838,000-30,024,000) upstream of the inflammatory gene *IL6* and containing regions belonging to the 'activated' set from Fig. 5B. **(d)** Heatmap of spearman correlation coefficients between histone coverage change and gene expression change between time points. Change is defined as the ratio between the two time points. All genes were included in the correlation heatmap on the left; only genes with a fold change of >2 in gene expression were included in the correlation heatmap on the right (see Methods and Supplemental Methods).
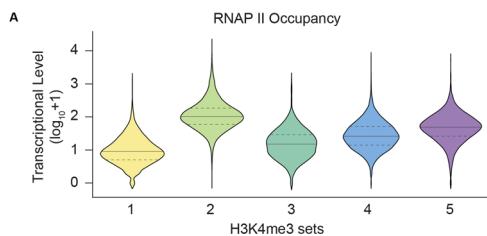
Extended Data Fig. 7 | Transcription levels of specific clusters of H3K4me3 enriched regions. (a) Violin plot of the transcriptional levels, measured by the RNAP II occupancy, of the five major clusters of H3K4me3 regions identified in Fig. 7.
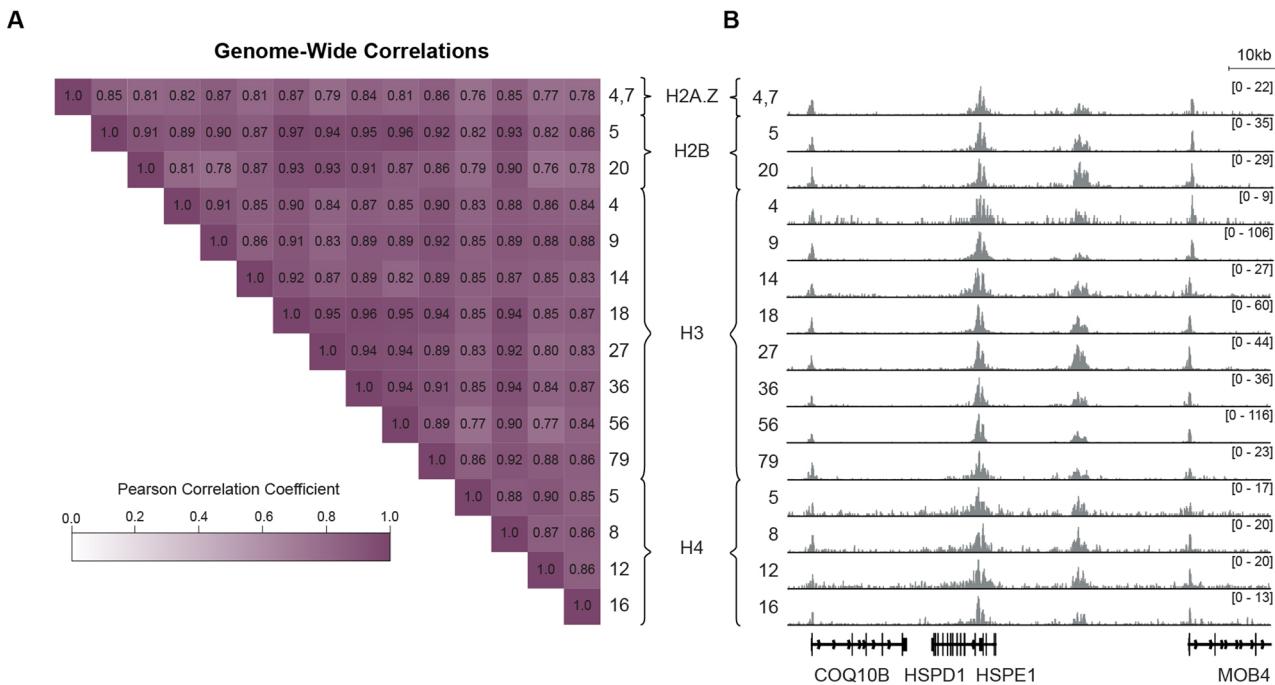
Extended Data Fig. 8 | Histone acetylation marks are highly correlated genome-wide. (a) Genome-wide pearson correlation coefficients of 15 different histone acetylation marks in mESC. Correlations are based on coverage computed in 10 kb windows. **(b)** Comparison of 15 different histone acetylation marks across a genomic region (mm10, chr1:55,048,000-55,148,000) in mESC.

A Weighted Combination Model Factor Enrichments

B Weighted Combination Model Position Relative to TSS Enrichments

C Weighted Combination Model Motif Enrichments

C1	E2F4, ELF1,ELF4, ETS, ETS1, ETV4, ELK1, ELK4, FLI1, GABPA
C2	BACH1, ELF1, ETS, ETV1, ELK1, ELK4, FLI1, GABPA, p53, p73
C3	ELF1, ETS, ELK4, ELK1, GFY, GFY-STAF, HINFP, NRF1, RONIN, ZFP57
C4	BACH1, KLF1, KLF3, KLF5, KLF6, NF-E2, NRF2, OCT4-SOX2-TCF-NANOG, SOX2, SP2
C5	BACH1, BACH2, ERRG, NF-E2, NRF2, TEAD1, TEAD4, TEAD3, TEAD, TEAD2

Extended Data Fig. 9 | Enrichment profiles for NMF generated combinations (C1-C5) of histone acetylation marks. (a) RNAP II, TF and CR enrichment matrix for regions assigned to combinations (C1-C5) from NMF decomposition of highly acetylated regions using histone acetylation marks, shown in Fig. 8. **(b)** Heatmap of genome position enrichments relative to TSS for regions assigned to combinations. **(c)** Transcription factors of top 10 most significant sequence motifs for regions assigned to each combination are listed.
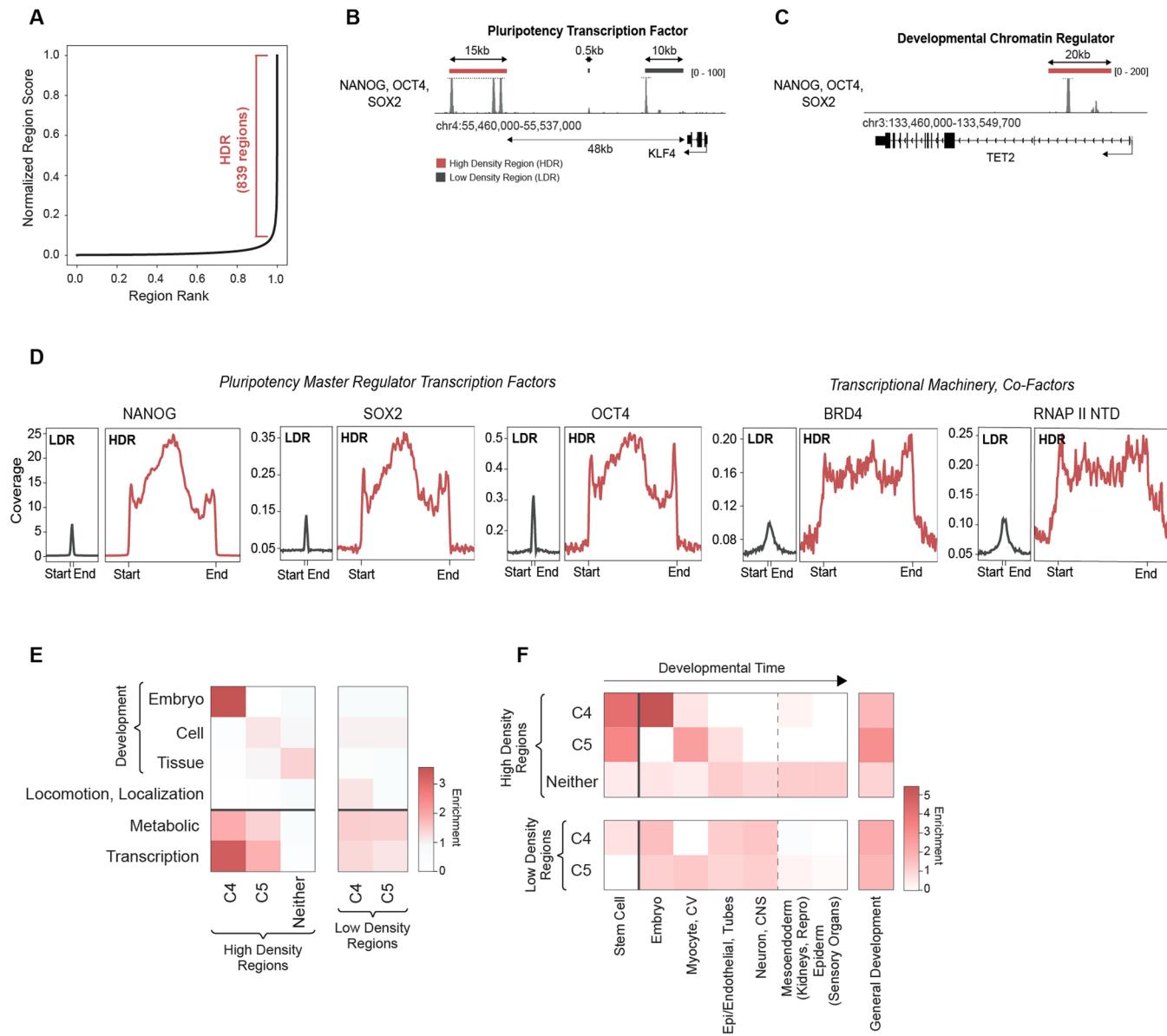
Extended Data Fig. 10 | Profiles for high density regions of NANOG-OCT4-SOX2. (a) Plot showing normalized region scores (x-axis) for peak regions of NANOG-OCT4-SOX2, ordered by rank (y-axis). High density regions are defined as regions past the point where the slope = 1. **(b)** Track visualization of NANOG-OCT4-SOX2 upstream of the gene for the pluripotency transcription factor KLF4 in mESC. A high density region is indicated with a red bar; low density regions are indicated with grey bars. **(c)** Visualization of NANOG-OCT4-SOX2 near the *TET2* gene, a developmentally associated chromatin regulator, in mESC. A high density region internal to the gene is indicated with a red bar. **(d)** Coverage metaplots over low density regions (LDR) vs high density regions (HDR) for pluripotency transcription factors and other transcriptional-related factors. Metagenes are centered on the region and the lengths represent the approximate difference in mean lengths (500 bps for LDRs and 14,500 bps for HDRs). An additional 4 kb surrounding each region is shown. **(e)** Enrichment heatmap for GO terms of genes associated with HDRs or LDRs containing C4, C5 or neither C4/C5 chromatin signatures. **(f)** Enrichment heatmap for development-associated GO terms of genes associated with HDRs or LDRs containing C4, C5 or neither C4/C5 chromatin signatures.

nature portfolio

Corresponding author(s): Mitchell Guttman

Last updated by author(s): Sep 30, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	N/A
Data analysis	- bedtools=2.29.2 - bowtie2=2.3.5 - chromhmm=1.24 - cutadapt=3.4 - deeptools=3.1.3 - fastqc=0.11.8 - homer=4.11 - multiqc=1.6 - nimfa=1.3.1 - numpy=1.17.2 - pandas=1.1.3 - pigz=2.6 - pysam=0.16.0.1 - python=3.6 - samtools=1.9 - scipy=1.6.2 - seaborn=0.12.2 - tqdm=4.38.0 - trim-galore=0.6.2

Online software:
http://go.cistrome.org/
https://igv.org/app/

Custom Code:
https://github.com/GuttmanLab/chipdip-pipeline

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All ChIP-DIP datasets generated in this study have been deposited at GEO: GSE227773 and will be publicly available as of the date of publication.

Accession numbers for publicly available ENCODE datasets used in this study are as follows: ENCFF121RHF, ENCFF508LLH, ENCFF035SOZ, ENCFF272JVI, ENCFF816ECC, ENCFF880HKV, ENCFF352HXD, ENCFF446FUS, ENCFF656DMV, ENCFF465UWC, ENCFF149MXA, ENCFF155UQU, ENCFF187HIQ, ENCFF702HIC, ENCFF800GVR, ENCFF982AFE, ENCFF178ARN, ENCFF096FWU, ENCFF844WTT, ENCFF801TEZ, ENCFF174BEG, ENCFF014HSG, ENCFF656FDC, ENCFF457PGP, ENCFF108EMO, ENCFF617EYG, ENCFF986KSB, ENCFF572IXE, ENCFF893LSE, ENCFF785OCU, ENCFF816VGU, ENCFF179XDZ, ENCFF863MYY, ENCRR698RKX, ENCRR736TRL, ENCFF222OPH, ENCFF731WDM, ENCFF479VOH, ENCFF514SHW, ENCFF643TQX, ENCFF586UHP, ENCFF456SZC, ENCFF108JNI, ENCFF632MQY, ENCFF924JXT, ENCFF834RTA, ENCFF660HBV, ENCFF567BOX, ENCFF563CWK, ENCFF394JNY, ENCFF307DMW, ENCFF181AVH, ENCFF169ZQQ, ENCFF040TWS, ENCFF990EZP, ENCFF755KWM, ENCFF269JZL, ENCFF106GOL, ENCFF827OVS, ENCFF789NDS, ENCFF567HEH, ENCFF519QMV, ENCFF008XKX, ENCFF362DZS, ENCFF158GBZ, ENCFF993IIG, ENCFF668UBL, ENCFF925BSH
Previously published ChIP-Seq data for TEAD1 was retrieved from GEO: GSE124008.
SRA accession numbers for RNAseq for LPS stimulated mDCs are as follows: DRR013943, DRR013949, DRR013952.
Previously published Gro-seq for mESC was retrieved from GEO: GSE4889528.

Accession numbers for publicly available CUTnTAG/ChIPmentation datasets used in this study are as follows: SRX1080059, SRX1080058, SRX1080056, SRX1080055, SRX1080050, SRX1080049, SRX1080048, SRX1080047, SRX1080026, SRX1080025, SRX1080024, SRX1080023, SRX1080022, SRX1080021, SRX1080028, SRX1080029, SRX1080018, SRX1080017, SRX1079997, SRX1079998, SRX1079996, SRX1079995, SRX1079994, SRX1079993, SRX5193376, SRX5193377, SRX5193378, SRX5193379, SRX5193380, SRX5193382, SRX5193384, SRX5193385, SRX5242669, SRX5545333

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	N/A
Data exclusions	None
Replication	Key proteins were mapped in multiple different experiments and show comparable results.
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Product Name,Vendor,Catalog Number
CTCF (D31H2) XP(R) Rabbit mAb,Cell Signaling Technologies ,3418S
Empty (no antibody),NA,NA
InVivoMAb recombinant human IgG1 Fc,Bio X Cell ,BE0096
Rabbit IgG, monoclonal [EPR25A],Abcam,ab172730
Rpb1 NTD (D8L4Y) Rabbit mAb,Cell Signaling Technologies ,14958S
SP1 (D4C3) Rabbit mAb,Cell Signaling Technologies ,9389S
TCF4/TCF7L2 (C48H11) Rabbit mAb,Cell Signaling Technologies ,2569S
Tri-Methyl-Histone H3 (Lys27) (C36B11) Rabbit mAb,Cell Signaling Technologies ,9733S
Tri-Methyl-Histone H3 (Lys4) (C42D8) Rabbit mAb,Cell Signaling Technologies ,9751S
V5 Tag Polyclonal Antibody,Bethyl,A190-120A
ZFX (L28B6) Mouse mAb,Cell Signaling Technologies ,5419S
AbFlex HDAC1,Active Motif,91216
AbFlex HDAC2,Active Motif,91198
AbFlex KDM5A antibody,Active Motif,91212
AbFlex BRD4 antibody ,Active Motif,91301
AbFlex EED antibody ,Active Motif,91135
AbFlex RNA Pol II antibody ,Active Motif,91151
AbFlex RNA pol II CTD phospho Ser2 antibody ,Active Motif,91115
AbFlex RNA pol II CTD phospho Ser5 antibody ,Active Motif,91119
Anti-KAT3B / p300 antibody [EPR23495-268] ,Abcam,ab275378
Anti-Rabbit IgG, monoclonal [EPR25A] ,Abcam,ab172730
BACH2 (D3T3G) Rabbit mAb,Cell Signaling Technologies,80775S
Biotinylated Anti-Rabbit IgG,Cell Signaling Technologies,14708S
BRD4 (E2A7X) Rabbit mAb,Cell Signaling Technologies,13440S
CBP (D6C5) Rabbit mAb,Cell Signaling Technologies,7389S
CBX4 (E6L7X) Rabbit mAb,Cell Signaling Technologies,30559S
CBX8 (D2O8C) Rabbit mAb,Cell Signaling Technologies,14696S
c-Fos (9F6) Rabbit mAb,Cell Signaling Technologies,2250T
c-Fos (E7L5L) Mouse mAb,Cell Signaling Technologies,74620S
c-Jun (60A8) Rabbit mAb,Cell Signaling Technologies,9165T
CTBP2 antibody (pAb),Active Motif,61261
EED (E4L6E) XP(R) Rabbit mAb,Cell Signaling Technologies,85322T
Ezh2 (D2C9) XP(R) Rabbit mAb,Cell Signaling Technologies,5246T

FOXG1 antibody (pAb),Active Motif,61211
FOXP1 antibody (mAb),Active Motif,61309
GATA-1 antibody (pAb),Active Motif,61535
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody,Life Technologies,A21037
HDAC2 (D6S5P) Rabbit mAb,Cell Signaling Technologies,57156S
HP1-alpha Antibody,Cell Signaling Technologies,2616S
HP1-beta (D2F2) XP(R) Rabbit mAb,Cell Signaling Technologies,8676T
JARID2 (D6M9X) Rabbit mAb,Cell Signaling Technologies,13594T
Non-phospho (Active) beta-Catenin (Ser33/37/Thr41) (D13A1) Rabbit mAb,Cell Signaling Technologies,8814S
p300 (D2X6N) Rabbit mAb,Cell Signaling Technologies,54062S
RB X-BACH1 ANTIBODY-TRIAL,Bethyl,A303-057A-T
RB X-REST ANTIBODY-TRIAL,Bethyl,A300-539A-T
RB X-SIRT6 ANTIBODY-TRIAL,Bethyl,A302-451A-T
RING1B (D22F2) XP(R) Rabbit mAb,Cell Signaling Technologies,5694T
Rpb1 NTD (D8L4Y) Rabbit mAb,Cell Signaling Technologies,14958S
SET1A (D3V9S) Rabbit mAb,Cell Signaling Technologies,61702S
SET1A (E3E2S) Rabbit mAb,Cell Signaling Technologies,50805S
SETD2 (E4W8Q) Rabbit mAb,Cell Signaling Technologies,80290S
SETD2 (E9H7P) Rabbit mAb,Cell Signaling Technologies,84384S
SP1 (D4C3) Rabbit mAb,Cell Signaling Technologies,9389S
SUV39H1 mAb (Clone MG44), sample,Active Motif,39786
SUZ12 (D39F6) XP(R) Rabbit mAb,Cell Signaling Technologies,3737T
Suz12 mAb (Clone 2A09), sample,Active Motif,39878
TBP (D5C9H) XP(R) Rabbit mAb,Cell Signaling Technologies,44059T
TBP (D5G7Y) Rabbit mAb,Cell Signaling Technologies,12578S
TCF4/TCF7L2 (C9B9) Rabbit mAb,Cell Signaling Technologies,2565S
TEAD1 antibody (pAb),Active Motif,61643
WDR5 antibody (pAb),Active Motif,61485
YY1 (D3D4Q) Rabbit mAb,Cell Signaling Technologies,63227S
Acetyl-Histone H3 (Lys27) (D5E4) XP(R) Rabbit mAb,Cell Signaling Technologies,8173S
Acetyl-Histone H3 (Lys9) (C5B11) Rabbit mAb,Cell Signaling Technologies,9649S
Anti-Actin antibody [ACTN05 (C4)] ,Abcam,ab3280
Anti-beta Catenin antibody [E247] ,Abcam,ab32572
Anti-CTCF antibody [EPR7314(B)],Abcam,ab128873
Anti-EED antibody [EPR23043-5],Abcam,ab240650
Anti-ERG antibody [EPR3864],Abcam,ab92513
Anti-HDAC2 antibody [EPR5001],Abcam,ab124974
Anti-Histone H3 (acetyl K27) antibody [EP16602],Abcam,ab177178
Anti-Histone H3 (acetyl K4) antibody [EPR16596],Abcam,ab176799
Anti-Histone H3 (di methyl K36) antibody [EPR16994(2)],Abcam,ab176921
Anti-Histone H3 (di methyl K4) antibody [Y47],Abcam,ab32356
Anti-Histone H3 (mono methyl K4) antibody [ERP16597] ,Abcam,ab176877
Anti-Histone H3 (mono+di+tri methyl K79) antibody [EPR17468],Abcam,ab177185
Anti-Histone H3 (tri methyl K27) antibody [EPR18607] ,Abcam,ab192985
Anti-hnRNP U/p120 antibody [EPR12278(2)(B)] ,Abcam,ab180952
Anti-KAT1 / HAT1 antibody [EPR18775],Abcam,ab194296
Anti-KAT3B / p300 antibody [EPR23495-268],Abcam,ab275378
Anti-KDM1/LSD1 antibody [EPR6825],Abcam,ab129195
Anti-KDM5A / Jarid1A / RBBP2 antibody [EPR18651],Abcam,ab194286
Anti-KDM5B / PLU1 / Jarid1B antibody [EPR12704] ,Abcam,ab181089
Anti-KDM5C / Jarid1C / SMCX + KDM5D / Jarid1D / SMCY antibody [EPR18653],Abcam,ab194288
Anti-KMT6 / EZH2 antibody [EPR20108],Abcam,ab191250
Anti-Lamin A + Lamin B1 + Lamin C antibody [EPR4068] ,Abcam,ab108922
Anti-Lamin B Receptor/LBR antibody [E398L] ,Abcam,ab32535
Anti-POLR3A antibody,Abcam,ab247007
Anti-RNA polymerase II CTD repeat YSPTSPS (phospho S2) antibody [EPR18855-87],Abcam,ab238146
Anti-SUZ12 antibody [EPR5234(N)],Abcam,ab175187
CSDE1 Antibody,Novus,NBP1-71914
CTCF (D31H2) XP(R) Rabbit mAb,Cell Signaling Technologies,3418S
Di-Methyl-Histone H3 (Lys4) (C64G9) Rabbit mAb,Cell Signaling Technologies,9725S
Di-Methyl-Histone H3 (Lys79) (D15E8) XP(R) Rabbit mAb,Cell Signaling Technologies,5427S
eIF3B/EIF3S9 Polyclonal Antibody,Bethyl,A301-761A
IgG Fraction Monoclonal Mouse Anti-Fluorescein (FITC),Jackson ImmunoResearch,200-002-037
InVivoMAb recombinant human IgG1 Fc,BioX Cell,BE0096
Living Colors® GFP Monoclonal Antibody,Living Colors,l632331
Mono-Methyl-Histone H3 (Lys4) (D1A9) XP(R) Rabbit mAb,Cell Signaling Technologies,5326S
Mono-Methyl-Histone H3 (Lys79) (D5X1S) Rabbit mAb,Cell Signaling Technologies,12522S
Phospho-Rpb1 CTD (Ser2) (E1Z3G) Rabbit mAb,Cell Signaling Technologies,13499S
Phospho-Rpb1 CTD (Ser2/Ser5) (D1G3K) Rabbit mAb,Cell Signaling Technologies,13546S
Phospho-Rpb1 CTD (Ser5) (D9N5I) Rabbit mAb,Cell Signaling Technologies,13523S
Rabbit IgG, monoclonal [EPR25A] ,Abcam,ab172730
Tri-Methyl-Histone H3 (Lys27) (C36B11) Rabbit mAb,Cell Signaling Technologies,9733S
Tri-Methyl-Histone H3 (Lys36) (D5A7) XP(R) Rabbit mAb,Cell Signaling Technologies,4909S
Tri-Methyl-Histone H3 (Lys4) (C42D8) Rabbit mAb,Cell Signaling Technologies,9751S
Tri-Methyl-Histone H3 (Lys79) (E8B3M) Rabbit mAb,Cell Signaling Technologies,74073S
Tri-Methyl-Histone H3 (Lys9) (D4W1U) Rabbit mAb,Cell Signaling Technologies,13969S

Tri-Methyl-Histone H4 (Lys20) (D84D2) Rabbit mAb,Cell Signaling Technologies,5737S
AbFlex® RNA Pol II antibody (rAb),Active Motif,91151
AbFlex® RNA pol II CTD phospho Ser2 antibody (rAb),Active Motif,91115
AbFlex® RNA pol II CTD phospho Ser5 antibody (rAb),Active Motif,91119
Di-Methyl-Histone H3 (Lys4) (C64G9) Rabbit mAb,Cell Signaling Technologies,9725T
Di-Methyl-Histone H3 (Lys79) (D15E8) XP(R) Rabbit mAb,Cell Signaling Technologies,5427T
EED (E4L6E) XP(R) Rabbit mAb,Cell Signaling Technologies,85322S
Histone H3 (D2B12) XP(R) Rabbit mAb (ChIP Formulated),Cell Signaling Technologies,4620S
HP1-beta (D2F2) XP(R) Rabbit mAb,Cell Signaling Technologies,8676S
Rabbit (DA1E) mAb IgG XP® Isotype Control ,Cell Signaling Technologies,3900S
RING1B (D22F2) XP(R) Rabbit mAb,Cell Signaling Technologies,5694S
SUZ12 (D39F6) XP(R) Rabbit mAb,Cell Signaling Technologies,3737S
TBP (D5C9H) XP(R) Rabbit mAb,Cell Signaling Technologies,44059S
Ubiquityl-Histone H2A (Lys119) (D27C4) XP(R) Rabbit mAb,Cell Signaling Technologies,8240T
5-Hydroxymethylcytosine (5-hmC) antibody (pAb),Active Motif,39770
5-Hydroxymethylcytosine (5-hmC) antibody (pAb),Active Motif,39792
5-Hydroxymethylcytosine mAb (Clone 59.1), sample,Active Motif,40000
AbFlex® 5-methylcytosine antibody (rAb),Active Motif,91188
AbFlex® HIF-1 alpha antibody (rAb),Active Motif,91140
AbFlex® Histone H3K27ac antibody (rAb),Active Motif,91194
AbFlex® Histone H3K36me3 antibody (rAb),Active Motif,91266
AbFlex® Histone H3K4me2 antibody (rAb),Active Motif,91322
AbFlex® Histone H3K4me3 antibody (rAb),Active Motif,91264
AbFlex® Histone H3K9ac antibody (rAb),Active Motif,91104
AbFlex® Histone H4K20me3 antibody (rAb),Active Motif,91108
Acetyl-Histone H2AZ (Lys4/Lys7) (D3V1I) Rabbit mAb,Cell Signaling Technologies,75336S
Acetyl-Histone H2B (Lys12) (D7H4) Rabbit mAb (ChIP Formulated),Cell Signaling Technologies,9072S
Acetyl-Histone H2B (Lys15) (D8H1) XP(R) Rabbit mAb,Cell Signaling Technologies,9083S
Acetyl-Histone H2B (Lys20) (D7O9W) Rabbit mAb,Cell Signaling Technologies,34156S
Acetyl-Histone H2B (Lys5) (D5H1S) XP(R) Rabbit mAb,Cell Signaling Technologies,3418S
Acetyl-Histone H2B (Lys5) (D5H1S) XP(R) Rabbit mAb,Cell Signaling Technologies,12799T
Acetyl-Histone H3 (Lys14) (D4B9) Rabbit mAb,Cell Signaling Technologies,7627T
Acetyl-Histone H3 (Lys18) (D8Z5H) Rabbit mAb,Cell Signaling Technologies,13998S
Acetyl-Histone H3 (Lys36) (D9T5Q) Rabbit mAb,Cell Signaling Technologies,27683S
Acetyl-Histone H3 (Lys9/Lys14) Antibody,Cell Signaling Technologies,9677S
Acetyl-Histone H4 (Lys12) (D2W6O) Rabbit mAb,Cell Signaling Technologies,13944S
Acetyl-Histone H4 (Lys16) (E2B8W) Rabbit mAb,Cell Signaling Technologies,13534S
Acetyl-Histone H4 (Lys5) (D12B3) Rabbit mAb,Cell Signaling Technologies,8647S
Anti-Histone H3 (acetyl K27) antibody [EP16602] - ChIP Grade,Abcam,ab177178
beta-Catenin (D10A8) XP(R) Rabbit mAb,Cell Signaling Technologies,8480T
BRD3 antibody (pAb),Active Motif,61490
BRD4 (pan) (E1Y1P) Rabbit mAb,Cell Signaling Technologies,83375S
CDK8 antibody (pAb),Active Motif,61482
c-Fos (9F6) Rabbit mAb,Cell Signaling Technologies,2250S
c-Jun (60A8) Rabbit mAb,Cell Signaling Technologies,9165S
COBRA1 (D6K9A) Rabbit mAb (nelf-b),Cell Signaling Technologies,14894S
Cyclin T1 (D1B6G) Rabbit mAb,Cell Signaling Technologies,81464S
Di/Tri-Methyl-Histone H3 (Lys9) (6F12) Mouse mAb,Cell Signaling Technologies,5327T
Di-Methyl-Histone H3 (Lys27) (D18C8) XP(R) Rabbit mAb,Cell Signaling Technologies,9728T
Di-Methyl-Histone H3 (Lys9) (D85B4) XP(R) Rabbit mAb,Cell Signaling Technologies,4658T
DNMT3A (E9P2F) Rabbit mAb,Cell Signaling Technologies,49768S
DNMT3B (E4I4O) Rabbit mAb (Mouse Specific),Cell Signaling Technologies,48488S
ELL (D7N6U) Rabbit mAb,Cell Signaling Technologies,14468S
ESET (D4M8R) XP(R) Rabbit mAb (SETDB1),Cell Signaling Technologies,93212S
EZH2 antibody (mAb),Active Motif,39876
EZH2 antibody (pAb),Active Motif,61642
G9a/EHMT2 (D5R4R) XP(R) Rabbit mAb,Cell Signaling Technologies,68851S
GT X-SP1 ANTIBODY-TRIAL,Bethyl,A303-944A-T
HDAC1 antibody (mAb),Active Motif,39532
HDAC3 (D2O1K) Rabbit mAb,Cell Signaling Technologies,85057S
Histone H2BK120ub1 antibody (mAb),Active Motif,39624
Histone H3K14ac antibody (pAb),Active Motif,39698
Histone H3K18ac antibody (pAb),Active Motif,39756
Histone H3K27ac antibody (pAb),Active Motif,39134
Histone H3K27me3 antibody (pAb),Active Motif,39157
Histone H3K36me3 antibody (pAb),Active Motif,61102
Histone H3K4ac antibody (pAb),Active Motif,61810
Histone H3K4me1 antibody (pAb),Active Motif,61634
Histone H3K4me2 antibody (pAb),Active Motif,39914
Histone H3K4me3 antibody (pAb),Active Motif,39916
Histone H3K56ac antibody (pAb),Active Motif,39282
Histone H3K79ac antibody (pAb),Active Motif,39566
Histone H3K79me2 antibody (pAb),Active Motif,39924
Histone H3K9ac antibody (mAb),Active Motif,61252
Histone H3K9ac antibody (mAb),Active Motif,61664
Histone H3K9ac antibody (pAb),Active Motif,39918

Histone H3K9me1 antibody (pAb),Active Motif,39888
Histone H3K9me2 antibody (pAb),Active Motif,39754
Histone H3K9me3 antibody (pAb),Active Motif,39766
Histone H3R17me2a (asymmetric) antibody (pAb),Active Motif,39710
Histone H3R17me2aK18ac antibody (pAb),Active Motif,61614
Histone H3R8me2a (asymmetric) antibody (pAb),Active Motif,39652
Histone H4ac (pan-acetyl) antibody (pAb),Active Motif,39926
Histone H4K12ac antibody (pAb),Active Motif,39928
Histone H4K16ac antibody (pAb),Active Motif,39930
Histone H4K20me3 antibody (mAb),Active Motif,39672
Histone H4K5ac antibody (pAb),Active Motif,39700
Histone H4K8ac antibody (pAb),Active Motif,61104
Histone H4R3me2a (asymmetric) antibody (pAb),Active Motif,39706
Histone H4R3me2s (symmetric) antibody (pAb),Active Motif,61188
INTS9 Antibody,Cell Signaling Technologies,13945S
JARID1B (E2X6N) Rabbit mAb,Cell Signaling Technologies,15327S
JARID2 (D6M9X) Rabbit mAb,Cell Signaling Technologies,13594S
KLF5 antibody (pAb),Active Motif,61100
Lamin A/C antibody (mAb),Active Motif,39288
LSD1 (C69G12) Rabbit mAb (KDM1),Cell Signaling Technologies,2184S
MBD3 (D1B8F) Rabbit mAb,Cell Signaling Technologies,99169S
MeCP2 antibody (mAb),Active Motif,61292
Menin antibody (pAb),Active Motif,61006
MLL / HRX antibody (pAb),Active Motif,61296
Mono-Methyl-Histone H3 (Lys9) (D1P5R) Rabbit mAb,Cell Signaling Technologies,14186S
MYST2 (D4N3F) Rabbit mAb (KAT7),Cell Signaling Technologies,58418S
Nanog (D1G10) Rabbit mAb (Mouse Specific; ChIP Formulated),Cell Signaling Technologies,8785S
Nanog antibody (pAb),Active Motif,61420
NRF1 (D9K6P) Rabbit mAb,Cell Signaling Technologies,46743S
Oct-4 antibody (pAb),Active Motif,39812
Oct-4A (C30A3C1) Rabbit mAb (ChIP Formulated),Cell Signaling Technologies,5677S
Oct-4A (D6C8T) Rabbit mAb (Mouse Specific),Cell Signaling Technologies,83932T
p53 (D2H9O) Rabbit mAb (Rodent Specific) (trp53),Cell Signaling Technologies,32532S
PCAF (C14G9) Rabbit mAb (KAT2b),Cell Signaling Technologies,3378T
PHC1 (1F3F3) Mouse mAb,Cell Signaling Technologies,13768S
PHF8 (E6K3Y) Rabbit mAb,Cell Signaling Technologies,93801S
Phospho-c-Fos (Ser32) (D82C12) XP(R) Rabbit mAb,Cell Signaling Technologies,5348T
Phospho-c-Jun (Ser63) (E6I7P) XP(R) Rabbit mAb,Cell Signaling Technologies,91952T
Phospho-c-Jun (Ser73) (D47G9) XP(R) Rabbit mAb,Cell Signaling Technologies,3270T
Phospho-Histone H3 (Ser10) (D7N8E) XP(R) Rabbit mAb,Cell Signaling Technologies,53348T
Phospho-Histone H3 (Thr3) (D5G1I) Rabbit mAb,Cell Signaling Technologies,13576S
Phospho-p53 (Ser15) Antibody,Cell Signaling Technologies,9284T
Phospho-Rpb1 CTD (Ser7) (E2B6W) Rabbit mAb,Cell Signaling Technologies,13780S
Phospho-Rpb1 CTD (Thr4) (D7L9W) Rabbit mAb,Cell Signaling Technologies,26319S
POLR1A (D6S6S) Rabbit mAb,Cell Signaling Technologies,24799S
RB X-ASH2 ANTIBODY-TRIAL,Bethyl,A300-112A-T
RB X-ATF2 ANTIBODY-TRIAL,Bethyl,A301-649A-T
RB X-BAF57/SMARCE1-TRIAL,Bethyl,A300-810A-T
RB X-BHC110/LSD1 ANTIBODY-TRIAL,Bethyl,A300-216A-T
RB X-BRG1, BLR106H MAB-TRIAL (SMARCA4),Bethyl,A700-106-T
RB X-CLOCK ANTIBODY-TRIAL,Bethyl,A302-617A-T
RB X-CREB ANTIBODY-TRIAL,Bethyl,A301-669A-T
RB X-CREB IHC ANTIBODY-TRIAL,Bethyl,IHC-00332-T
RB X-E2F1 ANTIBODY-TRIAL,Bethyl,A300-766A-T
RB X-E2F4 ANTIBODY-TRIAL,Bethyl,A302-133A-T
RB X-E4F1 ANTIBODY-TRIAL,Bethyl,A300-832A-T
RB X-EGR1 ANTIBODY-TRIAL,Bethyl,A303-390A-T
RB X-ELK1 ANTIBODY-TRIAL,Bethyl,A303-530A-T
RB X-EZH2 ANTIBODY-TRIAL,Bethyl,A304-197A-T
RB X-HCF1 ANTIBODY-TRIAL,Bethyl,A301-400A-T
RB X-HDAC1 ANTIBODY-TRIAL,Bethyl,A300-713A-T
RB X-HDAC2 ANTIBODY-TRIAL,Bethyl,A300-704A-T
RB X-HDAC2 ANTIBODY-TRIAL,Bethyl,A300-705A-T
RB X-HDAC2 IHC ANTIBODY-TRIAL,Bethyl,IHC-00057-T
RB X-HDAC3 ANTIBODY-TRIAL,Bethyl,A300-464A-T
RB X-HDAC6 ANTIBODY-TRIAL,Bethyl,A301-342A-T
RB X-JARID1A/RBP2 ANTIBODY-TRIAL,Bethyl,A300-897A-T
RB X-JMJD2A ANTIBODY-TRIAL (KDM4A),Bethyl,A300-861A-T
RB X-LAMIN-A ANTIBODY-TRIAL,Bethyl,A303-432A-T
RB X-LAMIN-A/C ANTIBODY-TRIAL,Bethyl,A303-430A-T
RB X-LAP2 ALPHA/TMPO ANTIBODY-TRIAL,Bethyl,A304-839A-T
RB X-LAP2 BETA/TMPO ANTIBODY-TRIAL,Bethyl,A304-840A-T
RB X-MAX ANTIBODY-TRIAL,Bethyl,A302-866A-T
RB X-MAZ/SAF-1 ANTIBODY-TRIAL,Bethyl,A301-652A-T
RB X-MBD2 ANTIBODY-TRIAL,Bethyl,A301-633A-T
RB X-MNT ANTIBODY-TRIAL,Bethyl,A303-626A-T

RB X-MNT ANTIBODY-TRIAL,Bethyl,A303-627A-T
RB X-MTA1 ANTIBODY-TRIAL,Bethyl,A300-280A-T
RB X-MTA1 ANTIBODY-TRIAL,Bethyl,A300-911A-T
RB X-MTA2 ANTIBODY-TRIAL,Bethyl,A300-395A-T
RB X-NFRKB ANTIBODY-TRIAL,Bethyl,A301-460A-T
RB X-NONO ANTIBODY-TRIAL,Bethyl,A300-582A-T
RB X-NONO ANTIBODY-TRIAL,Bethyl,A300-587A-T
RB X-NONO IHC ANTIBODY-TRIAL,Bethyl,IHC-00191-T
RB X-OCT-4 ANTIBODY-TRIAL,Bethyl,A304-591A-T
RB X-OCT-4 ANTIBODY-TRIAL,Bethyl,A304-592A-T
RB X-PHF8 ANTIBODY-TRIAL,Bethyl,A301-772A-T
RB X-PHF8 IHC ANTIBODY-TRIAL,Bethyl,IHC-00343-T
RB X-PHOSPHO RNA POLY II (S2)-TRIAL,Bethyl,A300-654A-T
RB X-POLR3D ANTIBODY-TRIAL,Bethyl,A302-295A-T
RB X-POLR3D ANTIBODY-TRIAL,Bethyl,A302-296A-T
RB X-POLR3E ANTIBODY-TRIAL,Bethyl,A303-707A-T
RB X-POLYCYSTIN-2 ANTIBODY-TRIAL,Bethyl,A302-470A-T
RB X-RAD21 ANTIBODY-TRIAL,Bethyl,A300-080A-T
RB X-Rad21, BLR052F MAB-TRIAL,Bethyl,A700-052-T
RB X-RAP74 ANTIBODY-TRIAL (Gtf2f1),Bethyl,A302-757A-T
RB X-RBBP5 ANTIBODY-TRIAL,Bethyl,A300-109A-T
RB X-RBBP5 IHC ANTIBODY-TRIAL,Bethyl,IHC-00172-T
RB X-RFX1 ANTIBODY-TRIAL,Bethyl,A303-043A-T
RB X-RNA POLYMERASE II ANTIBODY-TRIAL,Bethyl,A304-408A-T
RB X-RNF2 ANTIBODY-TRIAL,Bethyl,A302-869A-T
RB X-SAP30 ANTIBODY-TRIAL,Bethyl,A303-551A-T
RB X-SIN3A ANTIBODY-TRIAL,Bethyl,A300-724A-T
RB X-SIN3A ANTIBODY-TRIAL,Bethyl,A300-725A-T
RB X-SIX4 ANTIBODY-TRIAL,Bethyl,A303-220A-T
RB X-SP1 ANTIBODY-TRIAL,Bethyl,A300-133A-T
RB X-SP1 ANTIBODY-TRIAL,Bethyl,A300-134A-T
RB X-TAF1 ANTIBODY-TRIAL,Bethyl,A303-505A-T
RB X-TBP ANTIBODY-TRIAL,Bethyl,A301-229A-T
RB X-USF2 ANTIBODY-TRIAL,Bethyl,A303-540A-T
RB X-USF2 ANTIBODY-TRIAL,Bethyl,A303-541A-T
RB X-UTX ANTIBODY-TRIAL,Bethyl,A302-374A-T
RB X-YB1 ANTIBODY-TRIAL,Bethyl,A303-230A-T
RB X-YB1 ANTIBODY-TRIAL,Bethyl,A303-231A-T
RB X-YY1 ANTIBODY-TRIAL,Bethyl,A302-778A-T
RB X-YY1 ANTIBODY-TRIAL,Bethyl,A302-779A-T
RB X-ZBTB1 ANTIBODY-TRIAL,Bethyl,A303-242A-T
RB X-ZNF24 ANTIBODY-TRIAL,Bethyl,A303-092A-T
RB X-ZNF687 ANTIBODY-TRIAL,Bethyl,A303-278A-T
RBPJ antibody (mAb),Active Motif,61506
RING1B antibody (mAb),Active Motif,39664
RNA pol II CTD phospho Ser5 antibody (pAb),Active Motif,39750
RNA Pol II CTD phospho Thr4 antibody (mAb),Active Motif,61362
RNF20 (D6E10) XP(R) Rabbit mAb,Cell Signaling Technologies,11974S
SIN3A (D1B7) Rabbit mAb,Cell Signaling Technologies,7691S
SMARCA5 antibody (mAb),Active Motif,39544
Sox2 (D9B8N) Rabbit mAb,Cell Signaling Technologies,23064S
Sox2 antibody (pAb),Active Motif,39844
SPT16 (D7I2K) Rabbit mAb (Supt16),Cell Signaling Technologies,12191S
SPT4 (D3P2W) Rabbit mAb (SUPT4A),Cell Signaling Technologies,64828S
SPT6 (D6J9H) Rabbit mAb,Cell Signaling Technologies,15616S
SRC-3 (5E11) Rabbit mAb,Cell Signaling Technologies,2126S
SRF antibody (mAb),Active Motif,61386
SSRP1 (E1Y8D) Rabbit mAb,Cell Signaling Technologies,13421S
Suz12 antibody (pAb),Active Motif,39358
Tet1 antibody (pAb),Active Motif,61444
TET2 (D9K3E) Rabbit mAb (Mouse Specific),Cell Signaling Technologies,92529S
TH1L (D5G6W) Rabbit mAb (NELFCD),Cell Signaling Technologies,12265S
Ubiquityl-Histone H2A (Lys119) (D27C4) XP(R) Rabbit mAb,Cell Signaling Technologies,8240S
Ubiquityl-Histone H2B (Lys120) (D11) XP(R) Rabbit mAb,Cell Signaling Technologies,5546T
WDR5 (D9E1I) Rabbit mAb,Cell Signaling Technologies,13105S
YY1 (D5D9Z) Rabbit mAb,Cell Signaling Technologies,46395S
ZMYND11/BS69 (E7P9O) Rabbit mAb,Cell Signaling Technologies,67713S
AEBP2 (D7C6X) Rabbit mAb,Cell Signaling Technologies,14129S
BCL6 (D65C10) Rabbit mAb,Cell Signaling Technologies,5650S
CBF-beta (D4N2N) Rabbit mAb,Cell Signaling Technologies,62184S
CBP (D9B6) Rabbit mAb,Cell Signaling Technologies,7425S
FoxP1 (D35D10) XP(R) Rabbit mAb,Cell Signaling Technologies,4402S
HDAC1 (D5C6U) XP(R) Rabbit mAb,Cell Signaling Technologies,34589S
HDAC5 (E6G3N) Mouse mAb,Cell Signaling Technologies,98329S
MITF (D3B4T) Rabbit mAb,Cell Signaling Technologies,97800S
MITF (D5G7V) Rabbit mAb,Cell Signaling Technologies,12590S

Nanog (D2A3) XP(R) Rabbit mAb (Mouse Specific),Cell Signaling Technologies,8822T
RB X-LAP2 ALPHA/TMPO ANTIBODY-TRIAL ,Bethyl,A304-839A-T
TCF4/TCF7L2 (C48H11) Rabbit mAb,Cell Signaling Technologies,2569S

Validation	Antibodies were validated by the manufactur as described online. No further in-house validation was performed. All CST antibodies are ChIP validated and validation can be found at http://cellsignal.com/. Validation of antibodies from Active Motif when available can be found at: https://www.activemotif.com/. Validation of antibodies from Bethyl laboratories when available can be found at: https://www.fortislife.com/

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	(i) pSM44 mES cell line derived from a 129 × castaneous F1 mouse cross (ii) K562, a female human lymphoblastic cell line (ATCC, Cat # CCL-243)
Authentication	(i) The pSM44 mES cell line is identical to that used in previously published work (doi: 10.1016/j.cell.2021.10.014). This line was validated by gDNA sequencing by amplifying a region of modification. The amplified region included a SNP to confirm integration of dox promoter on 129 allele. Dox induction of Xist was checked by qPCR. The line was passaged for a maximum of 28 passages and qPCR checks for X chromosome retention were performed every passage. (ii) The K562 line was purchased from ATCC and directly used. The line was validated by the producer. No further in-house authentication was performed.
Mycoplasma contamination	Cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	N/A

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A