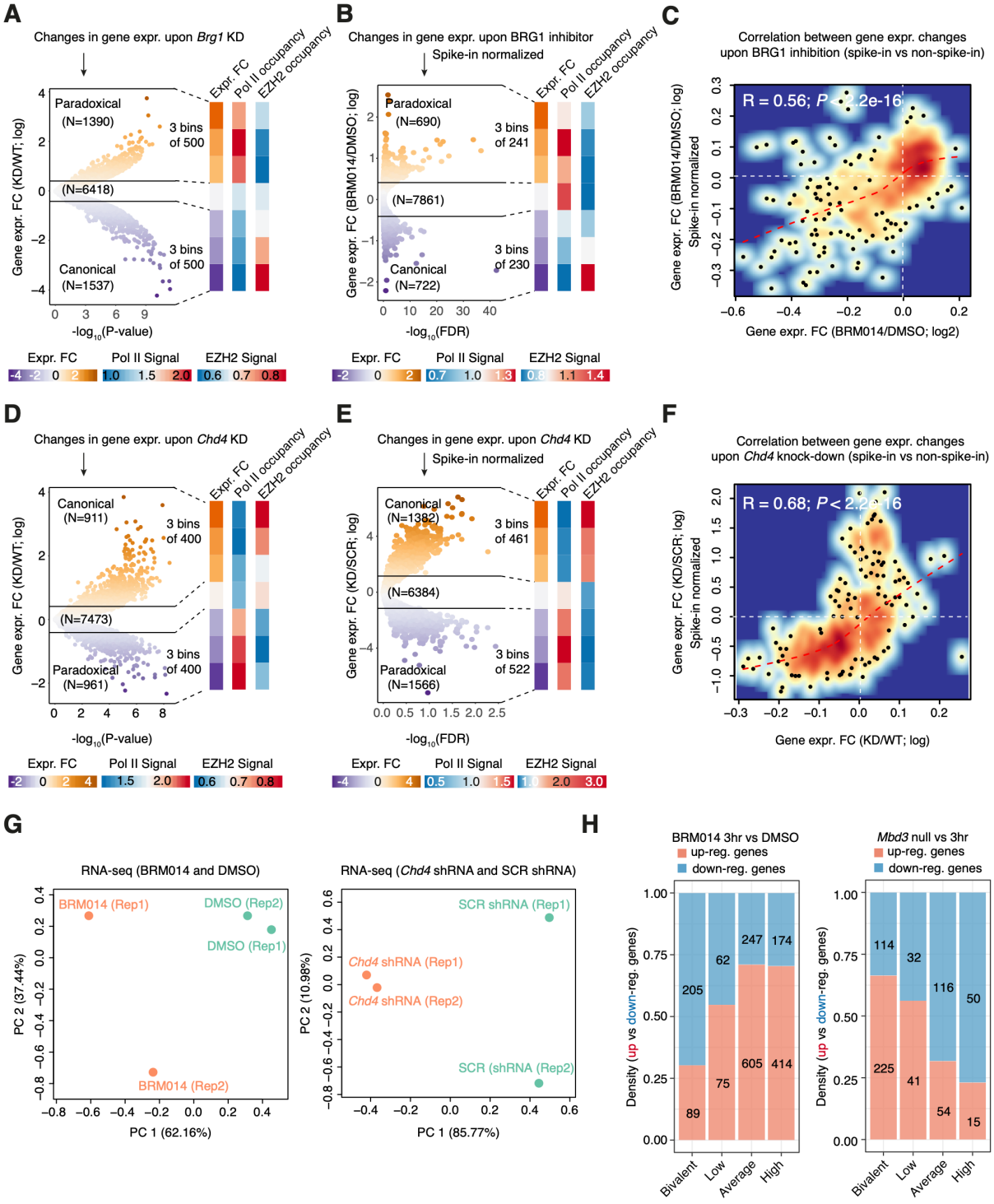
The impact of SWI/SNF and NuRD inactivation on gene expression is tightly coupled with levels of RNA Polymerase II occupancy at promoters

Sachin Pundhir[1,2,3#*], Jinyu Su[1,2,3#], Marta Tapia[1,2,3], Anne Meldgaard Hansen[1,2,3], James Seymour Haile[1,2,3], Klaus Hansen[1,2,3], Bo Torben Porse[1,2,3*]

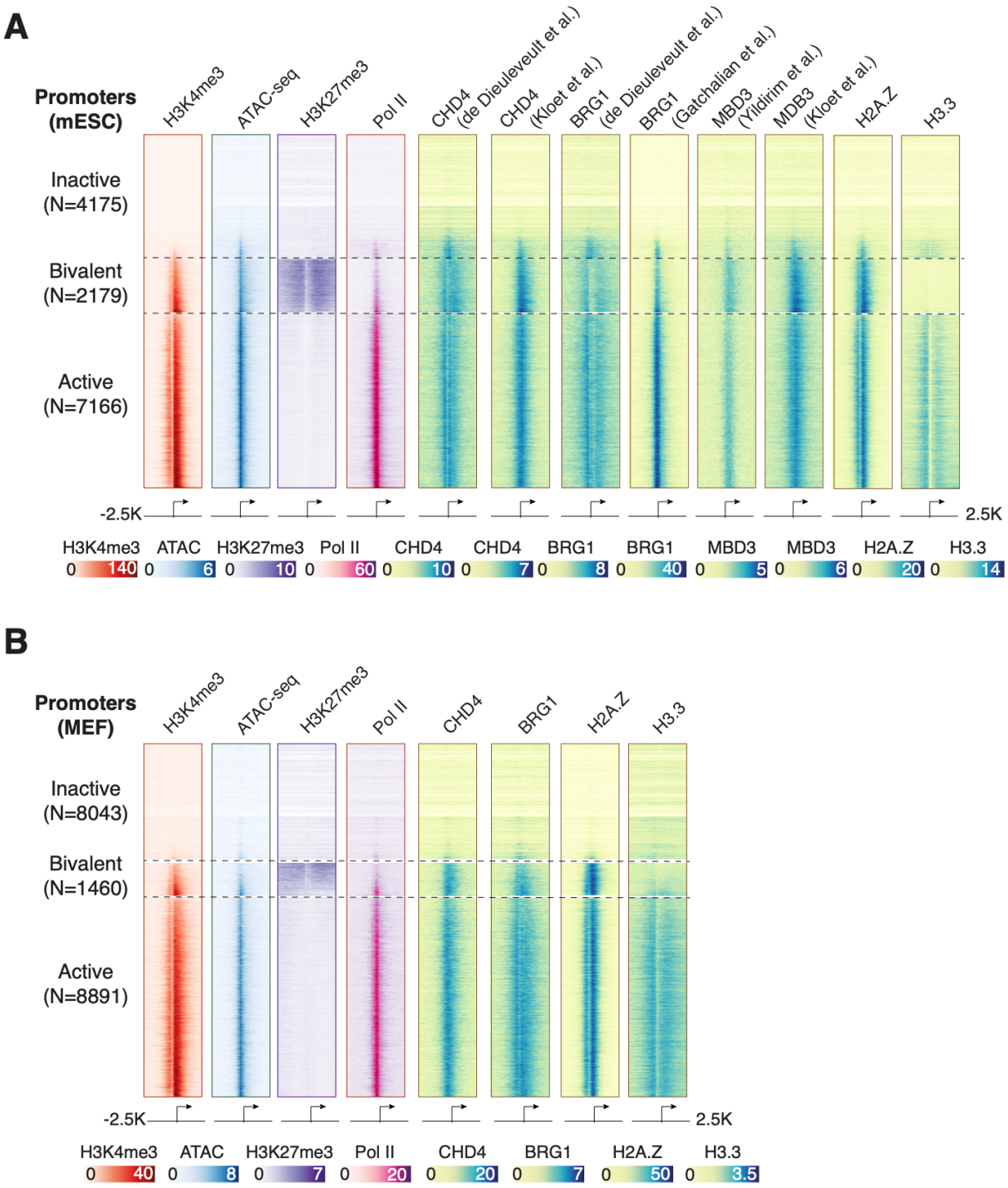
Supplementary Information

Supplementary Figures

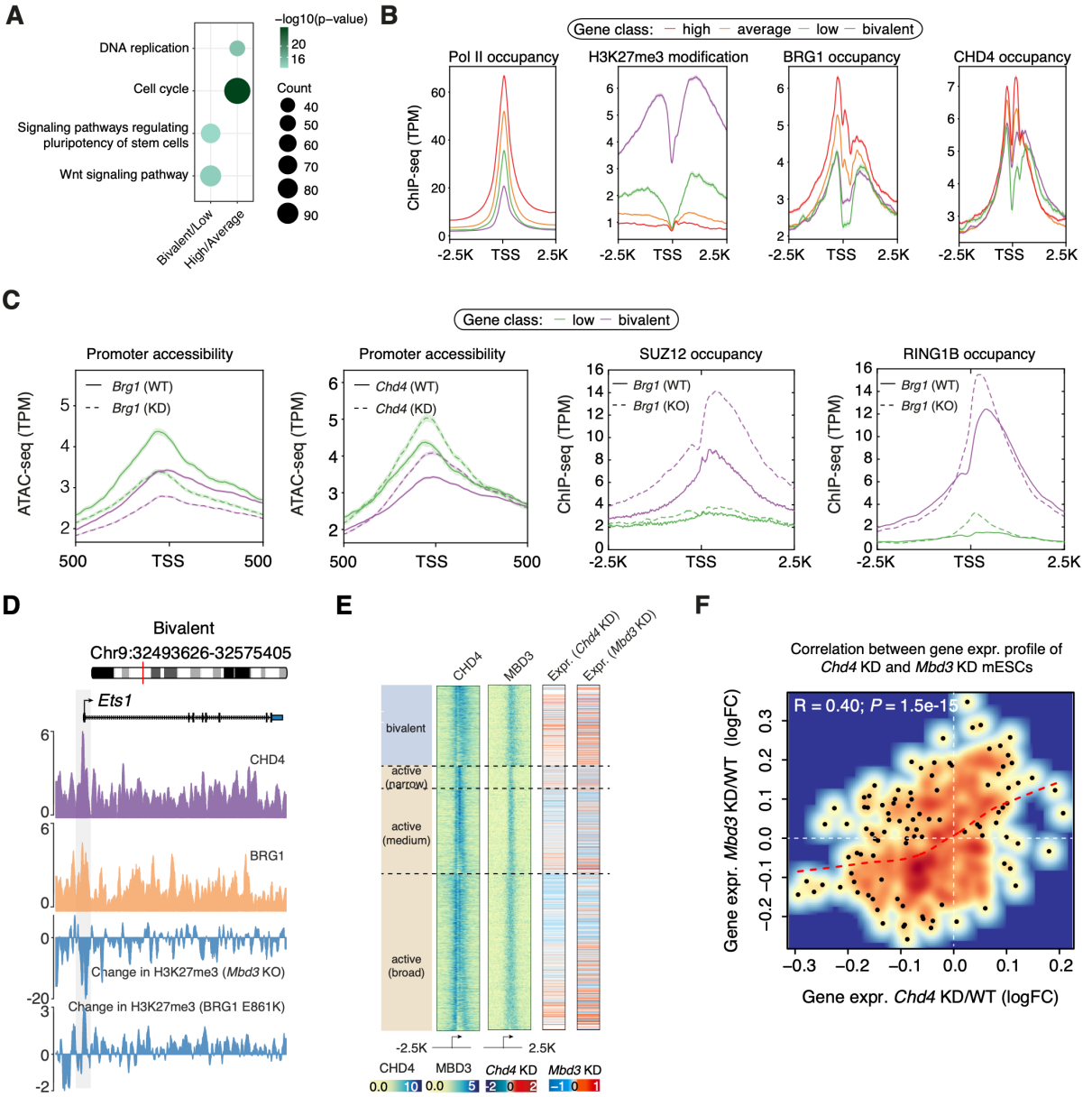


Supplementary Figure S1: Paradoxical changes in gene expression following loss in *Chd4* and *Brg1* expression is a feature of Pol II-bound promoters. (A) Volcano plot showing fold changes (log) in gene expression as a function of the significance level (p-value) following *Brg1* KD. Genes are binned into groups based on the fold changes in their expression, and for each group, the corresponding median fold change in expression as well as the RNA Polymerase II and EZH2 binding levels are shown. **(B)** Same as A but shows changes in spike-in

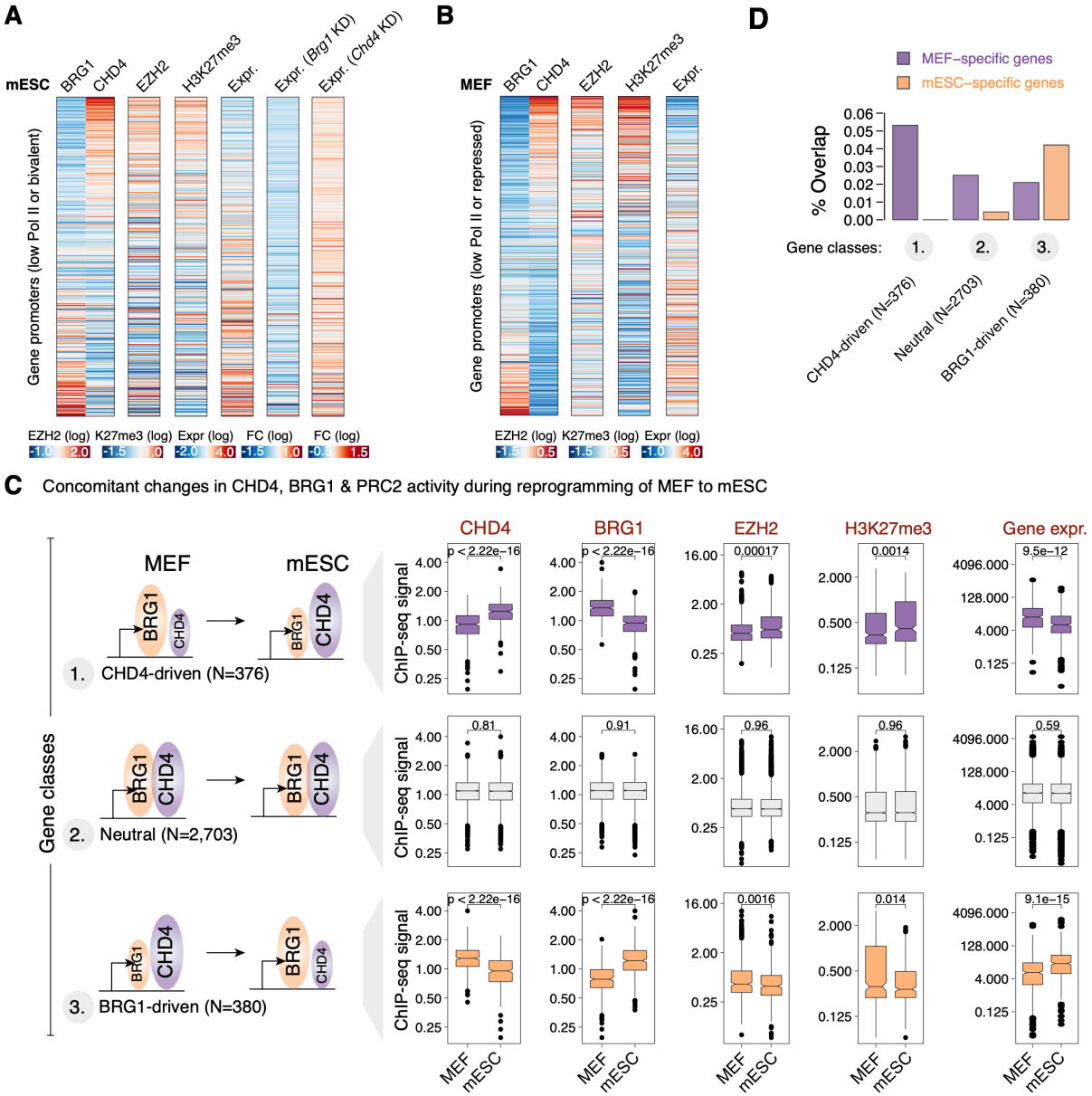
normalized nascent RNA levels upon 3 hours BRG1 inhibitor (BRM014) treatment. **(C)** Spearman's rank correlation between changes in mature RNA (x-axis) and spike-in normalized nascent RNA (y-axis) levels upon 3 hours of BRG1 inhibitor treatment. Genes binned into groups of 25 based on Pol II occupancy levels. **(D)** Same as A but shows changes in gene expression upon *Chd4* knock-down (KD). **(E)** Same as A but shows changes in spike-in normalized nascent RNA levels upon *Chd4* KD**. (F)** Spearman's rank correlation between changes in mature (x-axis) and nascent RNA (y-axis) levels upon *Chd4* KD. Genes binned into groups of 25 based on Pol II occupancy levels. **(G)** PCA plots showing spatial clustering of nascent RNA-seq samples (two biological replicates) subjected to BRG1 inhibitor treatment and WT (left panel) or Chd4 KD and WT (right panel). **(H)** Number of genes up- and down-regulated upon BRG1 inhibitor treatment or upon *Mbd3* KD that overlap with our four group of genes defined in based on Pol II occupancy levels (high, average, low and bivalent).



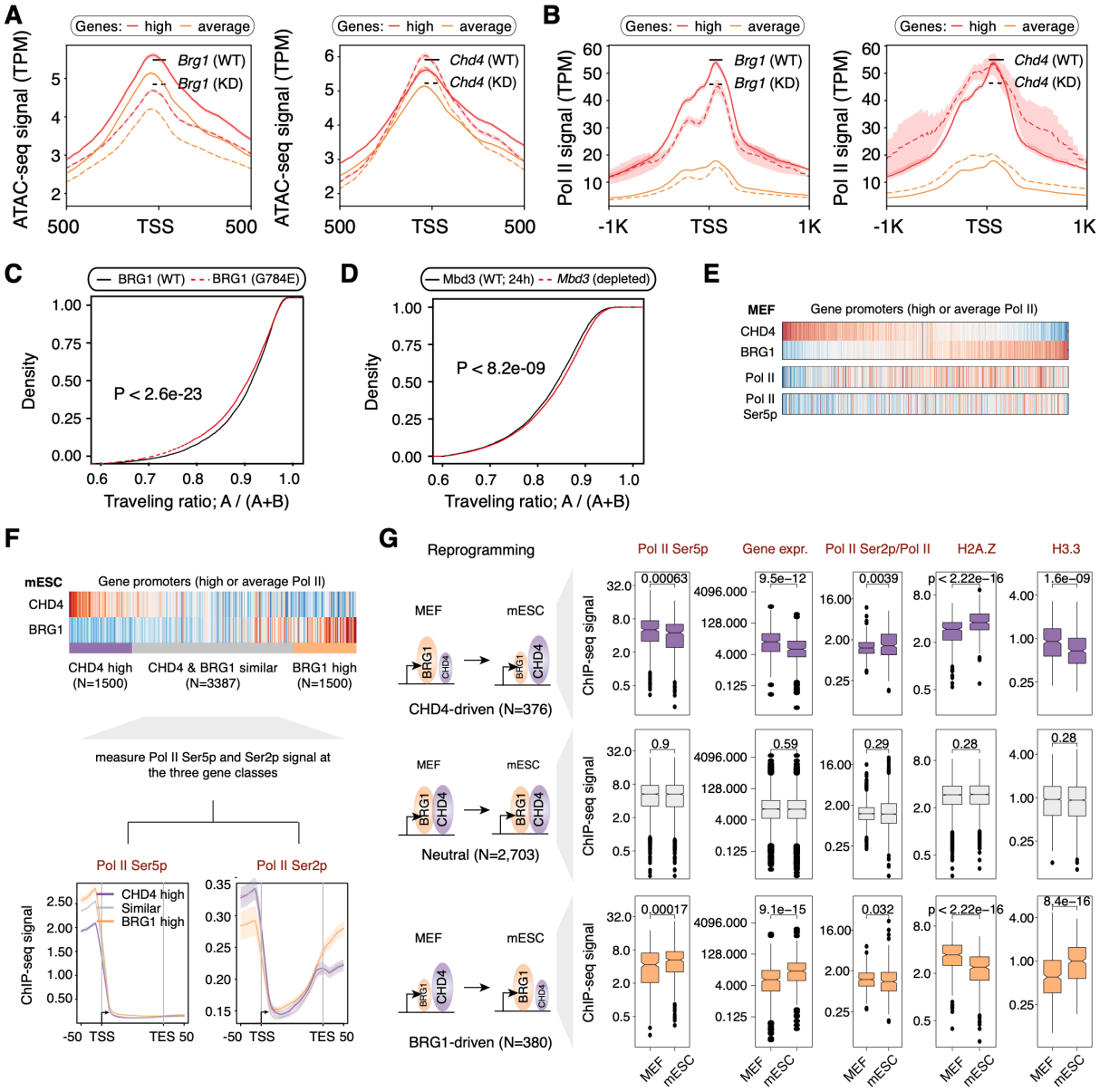
Supplementary Figure S2: Gene promoter classification based on histone modifications in mESCs and MEFs. **(A)** Classification of genes into bivalent (H3K4me3 and H3K27me3-enriched), active (H3K4me3-enriched) and inactive (remaining) gene classes in mESCs (de Dieuleveult et al. 2016; Kloet et al. 2018; Gatchalian et al. 2018; Yildirim et al. 2011). (**B)** Classification of genes into bivalent (H3K4me3 and H3K27me3 enriched), active (H3K4me3 enriched) and inactive (remaining) gene classes in MEFs.



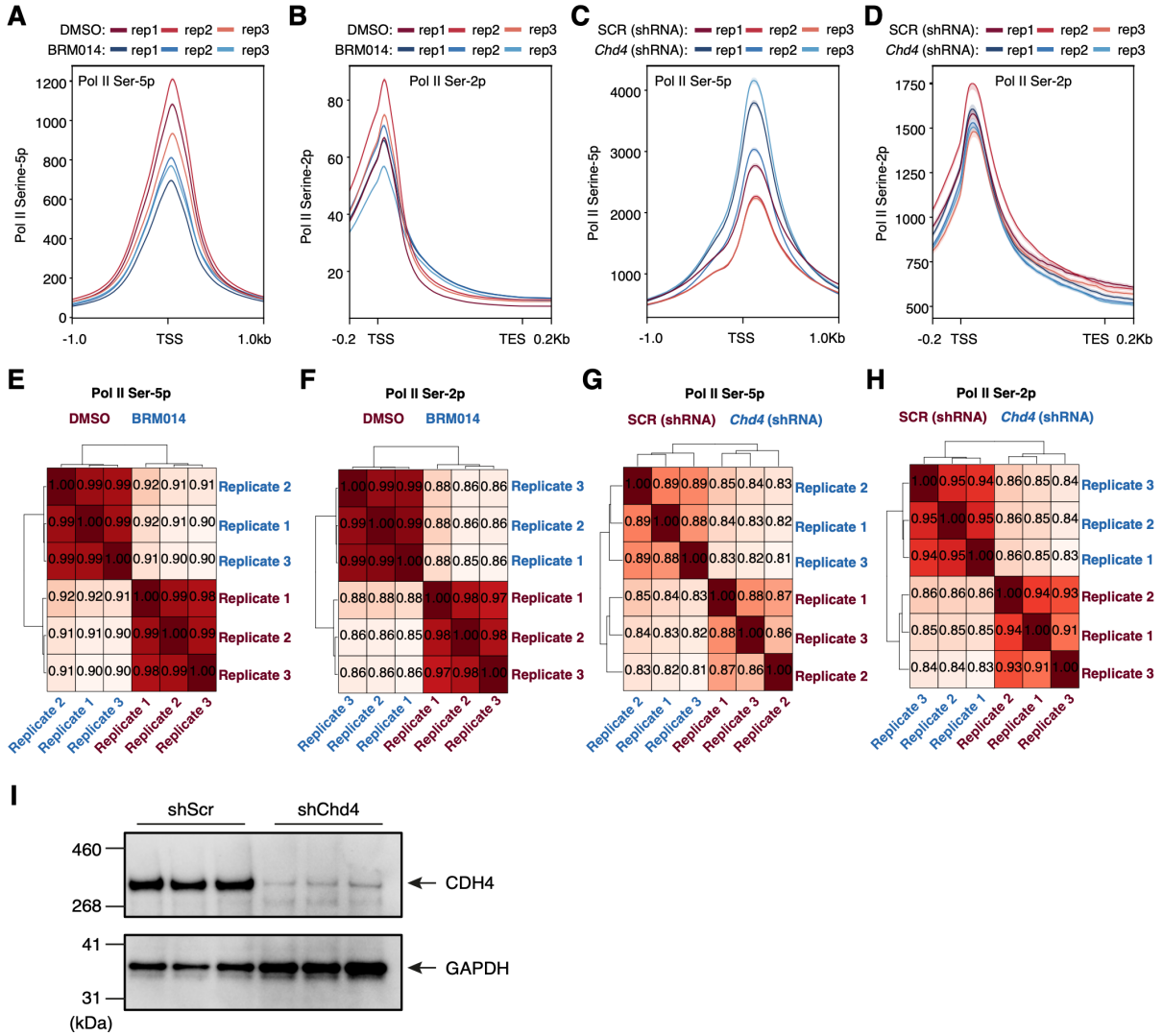
Supplementary Figure S3: The two NuRD components, CHD4 and MBD3, exhibit similar activity profiles at target gene promoters. (A) Differential biological pathways (KEGG) regulated by highly/average expressed (Pol II-bound; N=6387) and bivalent/lowly expressed (PRC2-bound; N=2958) genes. **(B)** Pol II, H3K27me3, CHD4 and BRG1 ChIP-seq signals at promoters of genes (N=9,345) classified based on their activity (high, N=2294; average, N=4093; low, N=779; and bivalent, N=2179) in mESCs. **(C)** Changes in DNA accessibility at promoters of lowly expressed/bivalent genes in mESCs following *Brg1* KD (first panel) or *Chd4* KD (second panel). Increase in SUZ12 (third panel) and RING1B (fourth panel) binding following *Brg1* KO in mESCs. **(D)** Genome browser view for *Ets1* (bivalent gene in Figure 2A) where loss in BRG1 activity leads to an increase in H3K27me3 levels. The opposite is observed following *Mbd3* KO. **(E)** ChIP-seq binding levels of CHD4 and MBD3 at gene promoters (bivalent, N=2179 and active, N=7166) in mESCs. Genes are binned into groups of 25 each and rank-ordered based on the Pol II binding levels at their promoters (same as in Figure 1D). For each group, the median gene expression fold changes following *Chd4* or *Brg1* knock-down (KD) are shown. **(F)** Spearman's rank correlation between the expression profiles of *Chd4* and *Mbd3* depleted ES cells.



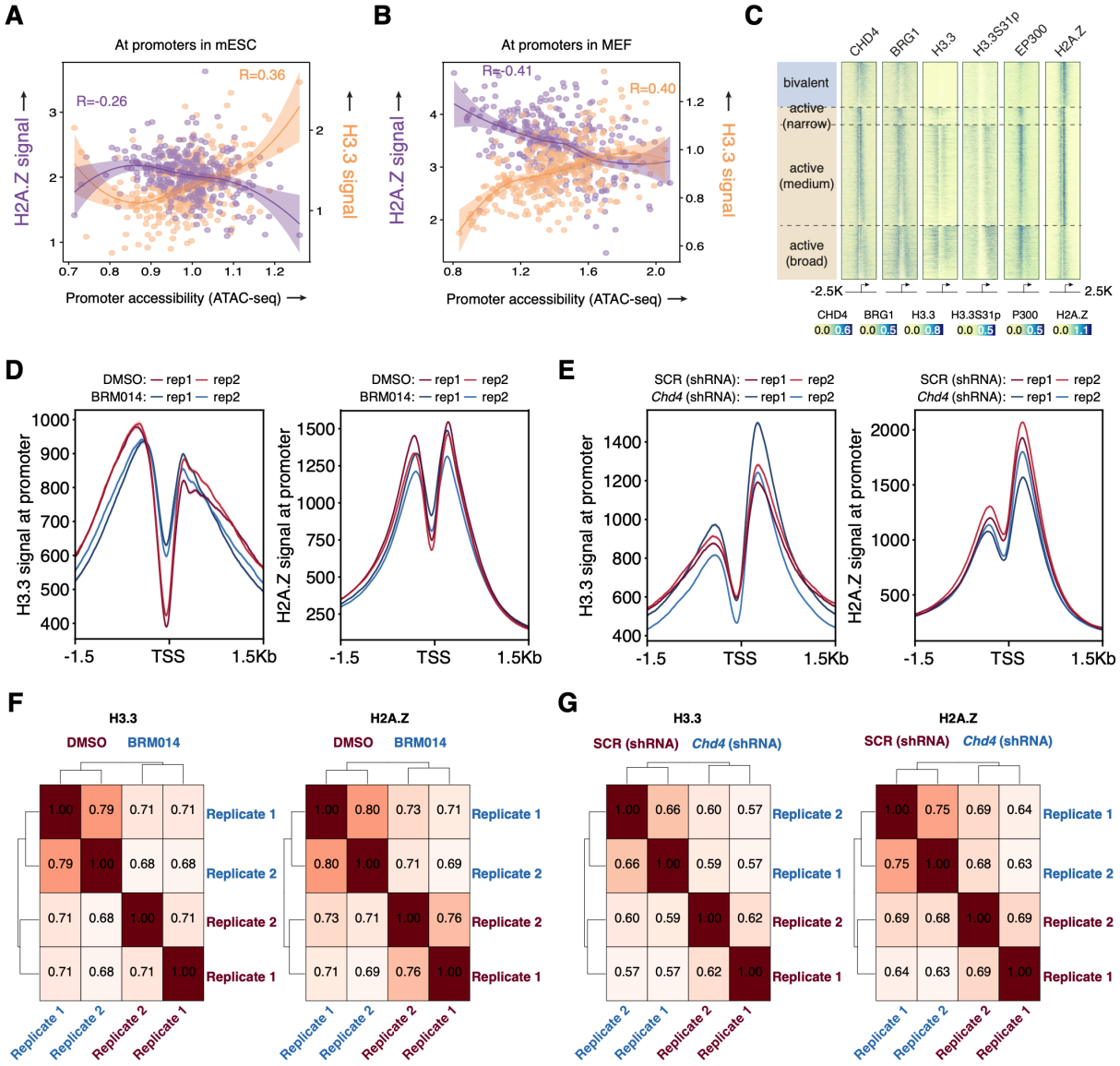
Supplementary Figure S4: The role of SWI/SNF and NuRD at low Pol II occupancy/PRC2-bound gene promoters in mESCs and MEFs. (A) Low Pol II occupancy/bivalent genes in mESCs (N=2,958) were binned into groups of 25 each based on the CHD4 and BRG1 binding ratio at their promoters in mESCs. For each group, the median promoter ChIP-seq signals of CHD4, BRG1, EZH2 and H3K27me3 are shown. Also shown is the median expression of these genes in mESCs along with the median fold changes in their expression following KD of *Brg1* and *Chd4*. **(B)** Low Pol II occupancy/repressed genes (N=3,117) were binned into groups of 25 each based on the CHD4 and BRG1 binding ratio at their promoters in MEFs. For each group, the median promoter ChIP-seq signals of CHD4, BRG1, EZH2 and H3K27me3 are shown. Also shown is the median gene expression in MEFs. **(C)** Classification of gene promoters based on increases in CHD4 activity (CHD4-driven), BRG1 activity (BRG1-driven) or no activity changes (neutral) during MEF to mESC reprogramming. Shown are the changes in PRC2 activity and gene expression levels for the three gene promoter classes during reprogramming (see Methods). **(D)** Percentage overlap of previously defined MEF- and mESC-specific genes (Chronis et al. 2017) with our CHD4-driven, BRG1-driven and CHD4 & BRG1-driven gene classes.



Supplementary Figure S5: The role of SWI/SNF and NuRD at average/high Pol II occupancy gene promoters in mESCs and MEFs. (A) Changes in DNA accessibility following *Brg1* KD (left) and *Chd4* KD (right) at genes with average/high Pol II promoter occupancy (N=6,387). **(B)** Same as A but depicting the changes in Pol II occupancy following *Brg1* KD (left) and *Chd4* KD (right). **(C)** Changes in traveling ratio upon loss of BRG1 activity. **(D)** Changes in traveling ratio upon loss of Mbd3 activity. **(E)** Genes with average/high Pol II promoter occupancy in MEFs (N=7,234) binned into groups of 25 each based on the CHD4 and BRG1 binding ratio at their promoters. For each bin, the levels of Pol II and Pol II (Ser5p) at promoters are shown. **(F)** Genes with average/high Pol II promoter occupancy in mESCs (N=6,387) binned into three groups based on the CHD4 and BRG1 binding ratio at their promoters. For each group, the levels of initiating (Ser5p) and released (Ser2p) Pol II at promoters are shown. **(G)** Differential changes in Pol II (Ser5p), gene expression, Pol II (Ser2p), H2A.Z and H3.3 levels following changes in the CHD4 to BRG1 binding ratios at gene promoters during MEF to mESC reprogramming (refer to Supplementary Figure S3C for class definition).



Supplementary Figure S6: SWI/SNF and NuRD antagonistically modulate Pol II release kinetics at genes having average/high Pol II promoter occupancy levels. (A) Changes in Pol II initiation (Ser-5p) levels at gene promoters (N=6,387) in mESCs upon inhibiting BRG1 activity by BRM014, compared to control treated with DMSO. Shown are the spike-in normalized levels for three individual replicates of DMSO-treated and BRM014 treated conditions. **(B)** Same as A but shows changes in Pol II release (Ser-2p) levels at gene promoters. **(C)** Changes in Pol II initiation (Ser-5p) levels at gene promoters (N=6,387) upon *Chd4* KD compared to the control. Shown are the spike-in normalized levels for three individual replicates of the two conditions. **(D)** Same as C but shows changes in Pol II release (Ser-2p) levels at gene promoters. **(E-F)** Hierarchical clustering of individual biological replicates of Pol II Ser5p (E) and Pol II Ser2p (F) samples (DMSO and BRM014) based on similarity in their ChIP-seq signal profiles. **(G-H)** Hierarchical clustering of individual biological replicates of Pol II Ser5p (E) and Pol II Ser2p (F) samples (Scrambled and Chd4 shRNA) based on similarity in their ChIP-seq signal profiles. **(I)** CHD4 protein levels measured by western blotting. Lanes 1-3 are triplicates of control mESCs (shScr) and lanes 4-6 are triplicates of *Chd4* KD mESCs (ShChd4). GAPDH is used for loading control.



Supplementary Figure S7: SWI/SNF modulate H3.3 and NuRD modulate H2A.Z levels at gene promoters. (**A**) Genes (N=6,387) are binned into groups of 25 each based on their DNA accessibility levels. For each bin, shown are the median H2A.Z and H3.3 levels (Loess regression lines are included). **(B)** Genes (N=7,234) are binned into groups of 25 each based on their DNA accessibility levels. For each bin, shown are the median H2A.Z and H3.3 levels (Loess regression lines are included). **(C)** CHD4, BRG1, H3.3, H3.3S31p, P300 and H2A.Z ChIP-seq signals at promotors of active and bivalent genes in mESCs. Genes are rank-ordered based on the Pol II binding levels at their promoters (same as in Figure 1D). **(D)** Changes in H3.3 (left panel) and H2A.Z (right panel) signal (spike-in normalized) at gene promoters in mESCs (N=6,387; average/high Pol II occupancy) upon inhibiting BRG1 activity using BRM014, compared to DMSO treatment. **(E)** Changes in H3.3 (left panel) and H2A.Z (right panel) signal (spike-in normalized) at gene promoters in mESCs (N=6,387; average/high Pol II occupancy) in *Chd4* KD mESCs compared to the scrambled control. **(F)** Hierarchical clustering of individual biological replicates of H3.3 (left) and H2A.Z (F) samples (DMSO and BRM014) based on similarity in their ChIP-seq signal profiles. **(G)** Hierarchical clustering of individual biological replicates of H3.3 (left) and H2A.Z (right) samples (Scrambled and Chd4 shRNA) based on similarity in their ChIP-seq signal profiles.

Supplementary Methods

Cell culture

Mouse embryonic stem cells (mESCs) were cultured on 0.2% gelatin-(Sigma) coated dishes in serum-free 2i medium comprised of 50% DMEM/F-12 (Gibco) and 50% Neurobasal medium (Gibco), supplemented with N2 supplement (Gibco), B27 serum-free supplement (Gibco), non-essential amino acids (Gibco), penicillin-streptomycin (Gibco), Sodium pyruvate (Gibco), Glutamax (Gibco), 10 μM β-mercaptoethanol (Gibico), 1 μM non-ATP-competitive MEK inhibitor (PD0325901, MedChemExpress), 3 μM glycogen synthase kinase 3 (GSK-3) inhibitor (CHIR-99021, MedChemExpress) and leukemia inhibitory factor (LIF, produced in the lab). Cells were passaged every 2-3 days by Accutase release (Stem Cell Technologies) and reseeded in fresh 2i medium.

Chd4 shRNA knockdown

HEK293 cells were co-transfected with PLKO.1 shRNA constructs, PAX8 packaging plasmid and VSVG envelope plasmid to produce shRNA lentivirus as previously described (Radzisheuskaya et al. 2016). Supernatants containing shRNA-encoding lentivirus were collected after 48 hours and lentivirus particles were immobilized onto RetroNectin (T100B, TaKaRa)-coated cell culture plates. mESCs were subsequently added to plates and transduced cells were selected by addition of 1.5 μg/mL puromycin 48 hours post-infection. After two days of selection cells harvested for western blotting and chromatin immunoprecipitation.

Western blotting

Samples for western blotting were prepared by addition of Laemmli sample buffer (BioRad) directly to cell pellets and boiled for 5 min. Proteins were separated on NuPAGE 3-8% Tris-acetate protein gels (Invitrogen) and blotted onto PDVF (polyvinylidene difluoride) membranes (Millipore) using a wet-transfer system at 30V overnight. Nonspecific binding sites were blocked with 5% milk in TBST (20 mM Tris pH 7.5, 150 mM NaCl and 0.05% Tween 20) for 1 hour at room temperature. The membrane was incubated with the anti-CHD4 antibody (ab264417, Abcam) overnight followed by incubation with the secondary antibody, swine anti-rabbit immunoglobulins/HRP (P021702-2, Agilent), and subsequently visualized using the SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific) in a ChemiDoc Imaging System (BioRad). In order to ensure similar loading, the membrane was subsequently reprobed with a monoclonal anti-GAPDH-peroxidase antibody (G9295, Sigma-Aldrich).

Promoter identification in mESCs and MEFs

We downloaded Ensembl-defined gene coordinates (release 97, mm10) (Cunningham et al. 2022). The TSS position corresponding to the longest isoform for each protein-coding gene was identified (N=21,827) and used to check for overlap with H3K4me3 and H3K27me3 peaks in both mESCs and MEFs. Gene TSSs overlapping with H3K4me3 peaks are marked as active, and those overlapping with both H3K4me3 and H3K27me3 peaks are marked as bivalent, and if none of the two histone modifications are enriched, the gene is marked as inactive. Furthermore, genes for which we have no expression measurements in the analyzed cell types (mESCs or MEFs) were discarded, giving us a final set of 13,520 genes in mESCs (inactive: 4,175; bivalent: 2,179 and active: 7,166) (Supplementary Figure S2A) and 18,394 genes in MEFs (inactive: 8,043; bivalent: 1,460; and active: 8,891) (Supplementary Figure S2B). A gene

can have multiple TSSs spread across its promoter region (canonical Nucleosome Free Region, NFR), the magnitude of which (promoter width) can be determined by measuring the extent of H3K4me3-modified non-canonical (fragile) nucleosomes (de Dieuleveult et al. 2016) (Supplementary Figure S2A). We therefore used the width of H3K4me3 peak at active and bivalent promoters to define promoter boundaries and observed it to contain multiple Pol II binding sites (TSS) flanked by up- (-1) and down-stream (+1) canonical nucleosomes targeted by NuRD (CHD4) and SWI/SNF (BRG1) chromatin remodelers (Supplementary Figure S2A-B). ChIP-seq signals for CHD4, BRG1, Pol II (ser2p or ser5p), EZH2, H3K4me3, H3K27me3, H3K27ac, H2A.Z and H3.3 at gene promoters were measured as tags per million and were further normalized by the width of the promoter region for unbiased comparison.

ChIP-seq data and analysis

Read mapping: We downloaded raw ChIP-seq data from GEO (Barrett et al. 2013) and ArrayExpress (Parkinson et al. 2011) to measure histone modifications, chromatin remodelers, histone variants and Pol II occupancy at gene promoters (source identifiers are in Supplementary Table S1). Reads were mapped to the mouse (mm10) genome assembly using Bowtie2 (Langmead et al. 2009). For Pol II Serine-5p, Serine-2p, H3.3 and H2A.Z ChIP-seq data in WT, BRG1 inhibitor (BRM014) treated and *Chd4* KD treated mESCs, reads were mapped to a combined *Mouse* (mm10) and *Drosophila* (dm6) genome, and read counts were spike-in normalized (Orlando et al. 2014) followed by coverage normalization using total number of reads mapping in genomic regions outside peaks (background). Only properly mapped paired reads were used for analysis. Furthermore, we only used uniquely mapped and PCR duplicates (exact copies) were collapsed into single reads, that were extended to their fragment length by determining the read extension size using MASC2 (predictd

parameter) (Zhang et al. 2008). Raw read counts were normalized to TPM using deeptools (bamCoverage) (Ramírez et al. 2014).

Genome bowser tracks: Normalized reads (bigwig files) for genomic regions of interest were visualized using Gviz (Hahne and Ivanek 2016). Changes in Pol II and H3K27me3 levels following BRG1 and CHD4 loss are computed by subtracting normalized read signals in relevant conditions using deeptools (bigwigCompare --ratio subtract) (Ramírez et al. 2014).

Histone modification enrichment analysis: Genomic regions enriched for H3K4me3 and H3K27me3 histone modifications were determined using a custom script which determines both the signal strength (peaks) as well as the width of a particular histone modification at a region (https://github.com/spundhir/findNFR, bam2peaks).

Gene expression data and analysis

Gene expression measurements corresponding to *Chd4* KD, *Brg1* KD, *Mbd3* KD and WT ES cells were retrieved from GEO using GEO2R utility (source identifiers are in Supplementary Table S1) (Barrett et al. 2013). Genes are designated as up-regulated (>1.5 FC upon KD at an adjusted P-value < 0.05), down-regulated (<1.5 FC upon KD at an adjusted P-value <0.05) or neutral for each chromatin remodeler. Gene expression measurements corresponding to mESCs and MEFs were retrieved from (Chronis et al. 2017). Raw gene expression measurements corresponding to H3.3 deficient ES cells were retrieved via email correspondence with Drs Maja Gehre and Kyung Min Noh (Gehre et al. 2020). Similarly gene expression measurements in mESCs treated with BRG1 molecular inhibitor or following re-expression of MBD3 in *Mbd3 null* mESCs were retrieved from (Bornelov et al. 2018; Iurlaro et al. 2021). Fold changes in gene expression were measured using DESeq2 (Love et al. 2014).

Gene expression measurements corresponding to H2A.Z deficient mESCs were retrieved from (Subramanian et al. 2013).

Functional annotation analysis: We analyzed 9,345 genes in mESCs (bivalent/low: 2,958; average/high: 6387) for functional annotation analysis in Figure 1D and Supplementary Figure S3A. The ClusterProfiler package (Yu et al. 2012) was used to determine the top biological pathways (KEGG) differentially enriched between the two gene groups.

Traveling ratio analysis

We analyzed Ensembl-defined Transcription Start Sites (TSS) (release 97, mm10) (Cunningham et al. 2022) to identify the most active TSS for each of the 9,345 genes analyzed in mESCs. The most active TSS is identified as the one for which we observed the highest Pol II ChIP-seq signal in mESCs. Next, we measured raw Pol II ChIP-seq signal at TSSR (Transcription Start Site Region; -50 bp to +300 bp around TSS) and the gene body (region 300 bp down-stream of TSS and 3,000 bp beyond Transcription End Site (TES)) as described in (Day et al. 2016). This analysis is done for each of the two replicates of Pol II ChIP-seq experiments performed in *Mbd3* deficient, BRG1$^{G784E/+}$ and WT ES cells. Next, we used DESeq2 (Love et al. 2014) to measure the fold changes in Pol II occupancy levels at TSSRs and gene bodies following loss of MBD3 and BRG1 activity. To compute Pol II traveling ratio at gene promoters, we normalized the raw Pol II occupancy levels at TSSRs and gene bodies to the total mapped reads and TSSR (fixed 350 bp) and gene body lengths (+300 bp past TSS to 3 kb past TES of the gene). The normalized Pol II occupancy levels were used to compute the traveling ratio, which gives a measure of Pol II levels at promoter versus gene body.

$$\text{Traveling ratio} = \frac{\text{Pol II occupancy (TSSR)}}{\text{Pol II occupancy (TSSR)} + \text{Pol II occupancy (gene body)}}$$

We used a similar approach to compute the Traveling index using GRO-seq data in mESCs and MEFs, except that reads mapping in the same orientation as the host transcript were used for analysis. We use a distance of 3kb post TES to measure Pol II release into the gene body. To ensure that our measurements were not affected by TSS of genes in close proximity (<3 kb), we repeated the Pol II traveling ratio analysis as shown in Supplementary Figure 5C-D, Figure 6B-C using only those genes which are >3kb away from the closest gene, and observed similar results.

Patient data analysis

Mutational information and RNA-seq datasets for each cancer cohort (the gene-level transcription estimates, as in log2(x+1)-transformed RSEM estimated counts) were downloaded from UCSC Xena (https://xenabrowser.net/) (Goldman et al. 2020). We analyzed cancer cohorts containing at least ten patients having mutations in *SMARCA4* (encoding BRG1) or *CHD4* yielding a total of 11 cancer cohorts for both *SMARCA4* and *CHD4* (Supplementary Table S2). Specifically, for each of the 11 cancer cohorts, we compared the expression levels of protein-coding genes (N=19,292) from patients having non-synonymous and damaging single nucleotide variants (SNVs) in *SMARCA4* or *CHD4* to those of the remaining patients of that cohort. We did this by fitting a generalized linear model using the glm function in R giving us a regression coefficient for each gene, where positive regression coefficients reflect a higher expression in patients harboring *SMARCA4* or *CHD4* lesions relative to the remaining patients, and *vice versa* for negative regression coefficients (refer to Supplementary Table S2 for the number of patients in each group). For each cancer cohort, we also defined the base line expression of each gene by determining its median expression across all the patients, irrespective of their mutation status. To compute the statistical

significance of the observed regression coefficients across the 11 cancer cohorts for both *SMARCA4* and *CHD4* lesions, we repeated the regression fit a hundred times by randomly selecting a set of patients as having SNVs (same number as in real case) and compared their gene expression levels with the rest of the patients. Next, we assessed the significance level (p=0.05) by computing the frequency by which we observed regression coefficients that were higher (*SMARCA4*) or lower (C*HD4*) than those observed in the real data sets. Significance levels were computed using Kolmogorov-Smirnov test and were corrected for multiple testing using the Benjamini-Hochberg procedure.

References

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al. 2013. NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Research* **41**: D991-995.

Bornelov S, Reynolds N, Xenophontos M, Dietmann S, Bertone P, Reynolds N, Xenophontos M, Gharbi S, Johnstone E, Floyd R. 2018. The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Molecular Cell* 1–17.

Chronis C, Fiziev P, Papp B, Sabri S, Ernst J, Plath K. 2017. Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* 1–18.

Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov AG, Barnes I, Bennett R, et al. 2022. Ensembl 2022. *Nucleic Acids Research* **50**: D988–D995.

Day DS, Zhang B, Stevens SM, Ferrari F, Larschan EN, Park PJ, Pu WT. 2016. Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. *Genome Biology* **17**: 1–17.

de Dieuleveult M, Yen K, Hmitou I, Depaux A, Boussouar F, Dargham DB, Jounier S, Humbertclaude H, Ribierre F, Baulard C, et al. 2016. Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature* **530**: 113–116.

Gatchalian J, Malik S, Ho J, Lee D-S, Kelso TWR, Shokhirev MN, Dixon JR, Hargreaves DC. 2018. A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. *Nat Commun* **9**: 5139.

Gehre M, Bunina D, Sidoli S, Lübke MJ, Diaz N, Trovato M, Garcia BA, Zaugg JB, Noh K. 2020. Lysine 4 of histone H3.3 is required for embryonic stem cell differentiation, histone enrichment at regulatory regions and transcription accuracy. *Nature Genetics* **52**: 273–282.

Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, et al. 2020. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* **38**: 675–678.

Hahne F, Ivanek R. 2016. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* **1418**: 335–351.

Iurlaro M, Stadler MB, Masoni F, Jagani Z, Galli GG, Schübeler D. 2021. Mammalian SWI/SNF continuously restores local accessibility to chromatin. *Nat Genet* **53**: 279–287.

Kloet SL, Karemaker ID, van Voorthuijsen L, Lindeboom RGH, Baltissen MP, Edupuganti RR, Poramba-Liyanage DW, Jansen PWTC, Vermeulen M. 2018. NuRD-interacting protein

ZFP296 regulates genome-wide NuRD localization and differentiation of mouse embryonic stem cells. *Nat Commun* **9**: 4588.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short {DNA} sequences to the human genome. *Genome Biol* **10**: R25.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, Guenther MG. 2014. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep* **9**: 1163–1170.

Parkinson H, Sarkans U, Kolesnikov N, Abeygunawardena N, Burdett T, Dylag M, Emam I, Farne A, Hastings E, Holloway E, et al. 2011. ArrayExpress update--an archive of microarray and high-throughput sequencing-based functional genomics experiments. *Nucleic Acids Res* **39**: D1002-1004.

Radzisheuskaya A, Shlyueva D, Müller I, Helin K. 2016. Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. *Nucleic Acids Res* **44**: e141.

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187-191.

Subramanian V, Mazumder A, Surface LE, Butty VL, Fields PA, Alwan A, Torrey L, Thai KK, Levine SS, Bathe M, et al. 2013. H2A.Z acidic patch couples chromatin dynamics to regulation of gene expression programs during ESC differentiation. *PLoS Genet* **9**: e1003725.

Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS, Weng Z, Rando OJ, Fazzio TG. 2011. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147**: 1498–1510.

Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16**: 284–287.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**: R137.