12660–12674 *Nucleic Acids Research, 2020, Vol. 48, No. 22* *Published online 25 November 2020*
doi: 10.1093/nar/gkaa1091

HP1γ regulates H3K36 methylation and pluripotency in embryonic stem cells

Nur Zafirah Zaidan[1,2] and Rupa Sridharan [1,3,*]

[1]Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI 53715, USA, [2]Genetics Training Program, University of Wisconsin-Madison, Madison, WI 53715, USA and [3]Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI 53715, USA

Received April 14, 2020; Revised October 12, 2020; Editorial Decision October 24, 2020; Accepted October 27, 2020

ABSTRACT

The heterochromatin protein 1 (HP1) family members are canonical effectors and propagators of gene repression mediated by histone H3 lysine 9 (H3K9) methylation. HP1γ exhibits an increased interaction with active transcription elongation-associated factors in embryonic stem cells (ESCs) compared to somatic cells. However, whether this association has a functional consequence remains elusive. Here we find that genic HP1γ colocalizes and enhances enrichment of transcription elongation-associated H3K36me3 rather than H3K9me3. Unexpectedly, sustained H3K36me3 deposition is dependent on HP1γ. HP1γ-deleted ESCs display reduced H3K36me3 enrichment, concomitant with decreased expression at shared genes which function to maintain cellular homeostasis. Both the H3K9me3-binding chromodomain and histone binding ability of HP1γ are dispensable for maintaining H3K36me3 levels. Instead, the chromoshadow together with the hinge domain of HP1γ that confer protein and nucleic acid-binding ability are sufficient because they retain the ability to interact with NSD1, an H3K36 methyltransferase. HP1γ-deleted ESCs have a slower self-renewal rate and an impaired ability to differentiate towards cardiac mesoderm. Our findings reveal a requirement for HP1γ in faithful establishment of transcription elongation in ESCs, which regulates pluripotency.

INTRODUCTION

Pluripotent stem cells such as embryonic stem cells (ESCs) have the remarkable ability to self-renew indefinitely while retaining the potential to differentiate into multiple cell types. These properties partially derive from the plasticity of ESC chromatin, which is in a decondensed state with fewer regions of compacted heterochromatin as compared to differentiated cells (1,2). Chromatin-associated proteins and the linker histone H1 are dynamically exchanged at a higher rate in ESCs than in differentiated cells (3). ESCs also have a global decrease in repressive histone modifications that are associated with heterochromatin (4,5). A dramatic manifestation of the reduction in heterochromatin is depletion of heterochromatin protein 1α (HP1α) foci in ESCs as compared to somatic cells (3).

The HP1 family of proteins was originally described as key regulators of heterochromatin formation. They are required in stable gene silencing classically observed in position-effect variegation (PEV) in *Drosophila melanogaster* (6,7). They recognize and perpetuate histone H3 lysine 9 (H3K9) methylation, a modification associated with gene repression, by recruiting SU(VAR)3–9, a methyltransferase for H3K9, creating a feed-forward loop and the spreading of heterochromatin (8,9). In mammals, there are three members: HP1α (Cbx5), HP1β (Cbx1) and HP1γ (Cbx3), which share sequence similarities and functions in heterochromatin regulation (10). However, emerging evidence suggests that they can perform divergent and non-redundant functions in heterochromatin as well as active transcription (11–20).

We have previously found that in somatic cells, HP1γ is located in foci of pericentromeric heterochromatin similar to HP1α. However, while HP1α remains at such foci in ESCs, HP1γ is spread over the nucleoplasm (21). This dramatic difference in localization is concomitant with increased protein interactions of HP1γ with transcription elongation-associated factors in ESCs. Replication-independent histone variant H3.3, which is deposited on actively transcribing genes, is enriched in HP1γ protein complexes in ESCs, with a corresponding decrease in linker histone H1. The FACT complex, which facilitates the opening of chromatin by removing H2A-H2B dimers to allow passage of RNA polymerase II, and transcription elongation associated histone modifications H3 lysine 79 methylation (H3K79me) and H3 lysine 36 methylation (H3K36me) are more prevalent in HP1γ associated complexes in ESCs (21). While transcription elongation is correlated with enrichment of histone modifications defined by genome–wide chromatin association (ChIP-Seq) studies, it is unclear

*To whom correspondence should be addressed. Tel: +1 608 316 4422; Email: rsridharan2@wisc.edu

whether their deposition is interdependent. H3K79 methylation, which generally peaks at the beginning of expressed genes, is correlated with elongation rates of RNA Pol II (22,23). On the other hand, H3K36me3, which is enriched towards the end of genes, can be recruited by RNA Pol II during transcription (24–26). These associations suggest that HP1γ may have an underappreciated role in transcription elongation in ESCs.

Therefore, we investigated the functional role of HP1γ in active transcription in mouse ESCs. Surprisingly, HP1γ enrichment on gene bodies was correlated with H3K36me3, especially on larger multiexonic genes, and not the canonical H3K9me3. The knockout (KO) of HP1γ in mouse ESCs resulted in both molecular and cellular phenotypes. There was a decreased enrichment of H3K36me3 and reduced expression of targeted genes that functioned in cellular homeostasis, morphogenesis and ribonucleotide binding in HP1γ-KO cells as compared to wild type controls. Restoring expression of HP1γ fragments that only contained hinge and chromoshadow domains that interact with nucleic acids and proteins was sufficient in maintaining H3K36me3 enrichment. These domains recruited NSD1, an H3K36 methyltransferase.

The cellular phenotype was manifested in a slower rate of self-renewal in HP1γ-KO ESCs. In addition, HP1γ KO ESCs displayed impaired differentiation towards the mesoderm lineage. Taken together our results reveal an unknown role for HP1γ in faithful establishment of transcription elongation in ESCs and pluripotency homeostasis.

MATERIALS AND METHODS

Derivation and maintenance of cell lines

HP1γ-KO ESC line was generated by introducing dual gRNAs targeting intron 2 (5′-GTTCTAGGAGGTGCT TACAC-3′) and intron 3 (5′-GATGATAGCCTTGCCGA GCG-3′) into Tet-inducible V6.5 ESC line (27), resulting in a deletion of exon 3. gRNAs were cloned into gRNA cloning vector (Addgene, 41824) and transfected into ESCs together with hCas9 (Addgene, 41815) using Lipofectamine3000 (Thermo Fisher Scientific). Cells were plated at low density to isolate single clones and screened via PCR, immunofluorescence, and Western Blot. Rescue lines were generated by FLP-mediated recombination of N-terminal 3X-FLAG-tagged HP1γ full length or mutant constructs into the Col1A1 locus of HP1γ-KO tet-inducible V6.5 ESC line. FLAG-tagged protein expression was induced with 1ug/ml doxycycline. All cell lines were maintained on irradiated mouse embryonic fibroblasts (MEFs) in ESC media (KnockOut DMEM, 15% fetal bovine serum (FBS), L-glutamine, Penicillin/Streptomycin, non-essential amino acids, 2-mercaptoethanol and leukemia inhibitory factor). Cells were regularly tested for mycoplasma upon each thaw and at least once a month in culture.

Immunofluorescence

Immunofluorescence was performed as described (28). Briefly, cells on coverslips were fixed with 4% paraformaldehyde for 10 minutes, followed by permeabilization with PBS + 0.5% Triton-X-100 for 10 min. Cells were washed in PBS + 0.1% Tween, blocked (1× PBS with 5% normal donkey serum, 0.2% Tween-20 and 0.2% fish skin gelatin) for 30 min, followed by primary antibody staining in blocking buffer overnight. After washing cells, secondary antibody staining was performed for 30 min at room temperature. Coverslips were mounted onto slides using Aqua-Polymount. Antibodies used were HP1γ Clone 2MOD-1G6 (Active Motif 39981) at 1:100 dilution, and FLAG clone M2 (Sigma-Aldrich F3165) at 1:1000 dilution. Secondary antibodies used are DyLight488 anti-mouse IgG (ThermoFisher Scientific, 35502) at 1:1000 dilution. Imaging was performed on Nikon Eclipse Ti using NIS Elements software.

Western blot

Equivalent number of cells were lysed using a syringe in 1X Laemmli Sample Buffer. Samples were separated on SDS-PAGE gels and transferred onto nitrocellulose membranes at 250 mA for 2 h at 4°C. Membranes were blocked in blocking buffer (5% milk in 1XPBS) for 30 min at room temperature and blotted for primary antibodies in blocking buffer overnight at 4°C. Membranes were washed with 1XPBS and incubated with HRP-conjugated secondary antibodies in blocking buffer for 1 h at room temperature. Membranes were washed and developed with ECL reagent. Images were quantified using ImageStudioLite software. Antibodies used were FLAG at 1:5000 (Santa cruz), α-TUBULIN at 1:3000 (Cell Signaling, 3873) and NSD1 at 1:500 (Sigma-Aldrich, 04-1565).

RT-qPCR

Following MIQE guidelines we provide all essential information for RT-PCR. The experimental groups were the HP1γ KO ESCs and control were WT ESCs for Supplementary Figure S2G and post differentiation on the days indicated for Figure 5D. Three biological replicates are in each group. RNA from each sample was extracted using Isolate II RNA Kit (Bioline) according to manufacturer's instructions which has on-column digestion with DNase I for 15 min. 1 ug of RNA measured using Nanodrop with a purity of greater than 1.9 for A260/280 was used for cDNA conversion. No other evaluations of integrity were performed. 1ug RNA was converted in a 10 ul reaction volume reaction using qScript cDNA synthesis kit (QuantaBio) with 0.5 ul of qScript RT enzyme for 22°C for 5 min, 42°C for 30 min, and 85°C for 5 min according to manufacturer's instructions. Relative gene expression was assayed with gene-specific primers (specificity validated using UCSC in-silico PCR, https://genome.ucsc.edu/cgi-bin/hgPcr): Actin (5′-TGTTACCAACTGGGACGACA-3′, 5′-TCTCAGCTGTGGTGGTGAAG-3′); Gapdh (5′-TTCACCACCATGGAGAAGGC-3′, 5′-CCC TTTTGGCTCCACCCT-3′); Polr2a (RNA Polymerase II) (5′-CGCACCACGTCCAATGATA T-3′, 5′-GTGCTGCTGCTTCCATAAGG-3′); T (Brachyury) (5′-TCTCTCTCCCCTCCACACAC-3′, 5′-ACTGCAGCATGGACAGACAA-3′); Tbx5 (5′-GGAGCCTGATTCCAAAGACA-3′, 5′-TTCAGCC ACAGTTCACGTTC-3′); Nkx2.5 (5′-CCAAGTGCTCT

Table 1. Primer information for RT-qPCR

Gene name	Accession #	Amplicon length (bp)	Location	Splice variant
Actin	NM_007393.5	392	Exon 3–4	-
Gapdh	NM_001289726.1, NM_008084.3	52	Exon 3–4	All variants (1,2)
Polr2a (RNA Polymerase II)	NM_001291068.1	268	Exon 24–26	-
T (Brachyury)	NM_009309.2	144	Exon 7–8	-
Tbx5	NM_011537.3	153	Exon 2–3	-
Nkx2–5	NM_008700.2	136	Exon 1–2	-
Mest	NM_001252292.1, NM_001252293.1, NM_008590.2	131	Exon 6–7	All variants (1, 2, 3)
Abca1	NM_013454.3	150	Exon 30–31	-
Hes6	NM_019479.3, NM_001360900.1	145	Exon 2–4	Variant 1 & 3
Six1	NM_009189.3	111	Exon 1–2	-

CCTGCTTTC-3′, 5′-GGCTTTGTCCAGCTCCACT-3′); Mest (5′-CCGGCTCACCATAAAGAGTC-3′, 5′-AATTCATGAGCCTGGTGAGG-3′); Abca1 (5′-CCAGGAACTCCTGAATGCTC-3′, 5′-TTCTGGAAGAGGTCCACGAT-3′); Hes6 (5′-GGTACCGAGGTGCAGGC-3′, 5′-GGATGTAGCCAGCAGCGAA-3′); Six1 (5′-TTAAGAACCGGAGGCAAAGA-3′, 5′-CCCCTTCCAGAGGAGAGAGT-3′). Additional primer information is in Table 1.

Quantitative PCR was performed in a 10 ul reaction of 1:5 diluted cDNA—i.e about 0.2 ug of converted cDNA per reaction using iTaq Universal SYBR Green Supermix (Bio-Rad, Cat#1725125) exactly according to manufacturer's specifications without any additive on a Bio-Rad c1000 thermocycler. The cycling parameters were 95°C for 20 s, 60°C for 20 s and 72°C for 30 s, for 40 cycles. Specificity of product was confirmed by melt curve. Since relative expression was used to analyze data, calibration curves were not performed. $2^{\Delta\Delta Ct}$ method was used for relative expression calculated against the geometric mean of three reference genes (Actin, Gapdh, and RNA Pol II) that were chosen to capture the dynamic range expected for target genes.

Chromatin immunoprecipitation and sequencing (ChIP-seq)

Following ENCODE and MINSEQE guidelines we provide all essential information for ChIP-seq. WT and HP1γ-KO ESCs were grown for four or six passages, respectively, and harvested at 80% confluency. Rescue HP1γ ESCs were grown for 5–6 passages, induced with 1ug/ml doxycycline for 72 h until 60–70% confluency. ChIP-seq for each sample was performed in duplicates. ESCs were harvested, MEF-depleted by plating onto a dish for 30 min and non-attached cells were collected. Between 10–15 million nuclei was collected from one 15 cm dish. Chromatin was prepared by cross-linking cells with 1% formaldehyde for 5 min, followed by adding glycine to 0.125 M for 5 min. Cells were washed and resuspended in ChIP Lysis Buffer (50mM Tris-HCl pH8.0, 20mM EDTA,1% SDS, protease inhibitor cocktail) and sonicated using Covaris S220 for 14 cycles of alternating 45s on/off with peak power = 170, duty factor = 5, cycles per burst = 200, 10–13 × 10^6 cells at a time. Soluble chromatin was collected by centrifuging at 12 000 × g for 10 min. 20ug of chromatin was used for each sample, together with a 0.4 ug spike-in of human chromatin from 293T cells (50:1 of sample to spike-in chromatin ratio). The combined chromatin was diluted with nine parts volume of Dilution Buffer (16.7 mM Tris pH8.0, 0.01% SDS, 1.1% Triton-X-100, 1.2 mM EDTA, 167 mM NaCl), and incubated with 3 ug of H3K36me3 antibody (Active Motif, 61101) overnight at 4°C. Equal portions of protein A and protein G magnetic beads (Thermo Fisher Scientific, 1002D and 1004D, respectively) were washed and added to each sample for 2 h at 4°C. Following IP, beads were washed twice with low salt buffer (50 mM HEPES, 0.1% SDS, 1% Triton X-100, 0.1% deoxycholate, 1 mM EDTA and 140 mM NaCl), high salt buffer (50 mM HEPES, 0.1% SDS, 1% Triton X-100, 0.1% deoxy-cholate, 1 mM EDTA and 500 mM NaCl), LiCl buffer (20 mM Tris, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA, and 250 mM LiCl) and TE. Elution of DNA was performed by incubating beads with elution buffer (50 mM Tris, 1 mM EDTA, and 0.1% SDS) followed by reverse cross-linking at 65°C overnight with 1% SDS. Samples were treated with RNase A for 30 min followed by 2 h of proteinase K digestion. Magnetic beads were removed, and DNA was purified by phenol:chloroform extraction followed by ethanol precipitation. ChIP-seq libraries were constructed using the Ovation Ultralow System V2 (NuGen, 0344) according to manufacturer's protocol. Libraries were sequenced on Illumina HiSeq, 50 or 75 bp single end, at NuCore of Northwestern University. There was >90% concordance between replicates based on overlap in genes.

Table 2. Datasets used in this study

HP1γ (ESCs)	GSM1081158
HP1γ (reprogramming intermediate, pre-iPSC)	GSM1081159
H3K9me3 (ESCs)	GSM2417112
H3K36me3 (ESCs)	GSM2417108
H3K79me2 (ESCs)	GSM2417104
H3K4me3 (ESCs)	GSM2417080
H3K27me3 (ESCs)	GSM2417100
H3K36me3 (ESCs) – parental	GSM3772689
H3K36me3 (ESCs) – Nsd1 KO	GSM3772690

ChIP-Seq processing

FASTQ files from published data were obtained from Gene Expression Omnibus (GEO) and associated GSM number can be found in 'Datasets used in study' (Table 2). HP1γ ChIP in mESCs and reprogramming interme-

diate (pre-iPSCs) were from Sridharan *et al.* (GSE44242) (5); H3K9me3, H3K36me3, H3K79me2, H3K4me3 and H3K27me3 ChIP in mESCs were from Chronis *et al.* (GSE90895) (29), and H3K36me3 from NSD1−/− ESCs was from Weinberg *et al.* (GSE118785) (30). FASTQ files were aligned to the mm9 genome using Bowtie aligner with default parameters. BAM and index files were generated using the SAMtools toolkit (31) with the 'view' and 'index' commands, respectively. Peak calling was performed using the MACS2 (32) 'callpeak' command between target and input samples with default narrowpeak settings, *P*-value cutoff of 1e−3 for H3K9me3 and H3K79me2 and *P*-value cutoff of 1e−4 for H3K4me3 and H3K36me3; or broadpeak settings with '–broad-cutoff 0.1' for HP1γ and H3K27me3 ChIP. To determine distribution of ChIP peaks to genomic region, the Cis-regulatory element annotation system (CEAS) (33) was performed with the mm9 gene annotation table file and default settings.

To find intersecting peaks between ChIP samples, called peaks were converted to GRanges and overlapped via R package ChIPpeakAnno (34) using the 'findOverlapsOfPeak' command which has a default of 1 bp. Peak annotation was done using the EASeq (35) 'Annotate' tool with the center of the peak set at nearest start or end of a gene. Metaplots of averaged ChIP density across gene sets in Figures 1, 2, Supplementary Figures S1 and S2 were generated using ngs.plot (36) with default settings, including normalization per million mapped reads. To categorize HP1γ-bound genes as in Figure 2, genes were first grouped based on the location of HP1γ peaks that is annotated to the nearest gene, either near the TSS (3 kb upstream and 0.5 kb downstream of the TSS), near the TES (0.5 kb upstream and 3 kb downstream of the TES), as well as any HP1γ peaks that falls within genes (from TSS to TES). Since many genes contain multiple HP1γ peaks annotated to various locations on the gene, we performed an overlap of genes from these categories (Figure 2A) and focused on the genes that had HP1γ annotated peaks unique to each category (colored sections in Figure 2A).

For quantitative ChIP normalization based on spike-in control (37), we first quantified the total reads aligning to the human genome, hg19. Since the amount of human spike-in chromatin is expected to be constant across samples, we determined a Scaling Factor such that the total human mapped reads would be the same for all samples. We then normalized ChIP samples by the scaling factor using Deeptools (38) 'bamCoverage' command with default settings of 50 bp bins. Metaplots and heatmaps of the normalized reads (as in Figures 3, 4 and Supplementary Figure S3) were generated using Deeptools 'computeMatrix', 'plotHeatmap', and 'plotProfile' commands. The average read counts for each genomic region were computed using Deeptools 'multiBigwigSummary' command in 'BED-file' mode. Nsd1 KO ChIP normalization was done by scaling to reads aligned to the Drosophila genome, dm6.

Enriched functional annotation was determined by DAVID Bioinformatics Resources (39) and Gene Set Enrichment Analysis (GSEA) (40). Venn diagram for overlapped genes was performed using web software Venny (http://bioinfogp.cnb.csic.es/tools/venny/).

RNA sequencing

Samples for RNA-seq were prepared as described (41). RNA was extracted using TriZol followed by purification on Qiagen RNeasy Mini Kit. 1 ug of RNA was used for each sample, combined with 20ng of RNA from 293T (human) cells, prepared as above, as a spike-in control for sequencing normalization. cDNA library was constructed using TruSeq RNA Sample Preparation Kit (Illumina, RS-122-2002) according to the manufacturer's instructions. Libraries were assessed by Qubit and Bioanalyzer3.0. Samples were multiplexed and sequenced paired end for 150bp on Illumina NovaSeq platform at Novogene. Experiments were performed in biological duplicates.

RNA sequencing analysis

RNA-seq analysis was performed as described (41). Sequenced reads were trimmed, cropped, filtered by quality, and adapter sequence clipped using Trimmomatic (42) paired-end mode, with the options 'LEADING:3 TRAILING:3 CROP:95 HEADCROP:13 SLIDINGWINDOW:4:15 MINLEN:36'. Accepted sequences were aligned to mouse reference genome mm9 and human reference genome hg18 constructed to include only annotated genes (i.e. NM_ RefSeqs) by Bowtie 2 alignment, via RSEM-1.2.4 (43) with a mismatch per seed set to 2 and seed length set to 28. A matrix of unnormalized reads mapping to the mouse or human genome was generated using R. Differentially Expressed (DE) genes was determined using EBSeq program which uses a Bayes approach to calculate the distribution of expression (44), using the human aligned reads matrix as median normalization (MedianNorm(human_matrix)). PostFC values from EBseq output was used to plot fold change in gene expression as in Figure 3H. Genes were considered differentially expressed with a posterior probability > 0.95 and FDR $= 0.05$. TPM values for all genes was calculated using RSEM.

Protein extraction and Immunoprecipitation

Cell pellets were extracted in Buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM DTT, phosphatase inhibitor and protease inhibitor cocktail) followed by nuclear pellet extraction in Buffer C (0.42 M KCl, 20 mM HEPES pH 7.9, 0.2 mM EDTA pH 8.0, 5% glycerol, 1 mM DTT, phosphatase and protease inhibitor cocktail); or in Micrococcal Nuclease digestion buffer (MCN; 50 mM Tris pH 7.6, 1 mM $CaCl_2$, 0.2% Triton-X-100, phosphatase inhibitor and protease inhibitor cocktail) with 20 units S7 Micrococcal Nuclease per 1×10^5 cells for 10 min at 37°C and stopped with 50 mM EDTA. NaCl was added to 0.3 M and incubated at 4°C for 1 h before centrifuging at 13 000 × g for 30 min to isolate soluble nuclear extracts.

For FLAG IP, 3mg of nuclear extract was incubated with 150 μl of FLAG-M2 agarose beads (Sigma, 50% slurry) for 2 h at 4°C, washed thrice with MCN buffer with 0.3M NaCl, followed by washing twice with the same buffer but without Triton-X-100. IP complexes were eluted in Laemmli Sample Buffer. Three independent immunoprecipitation exper-

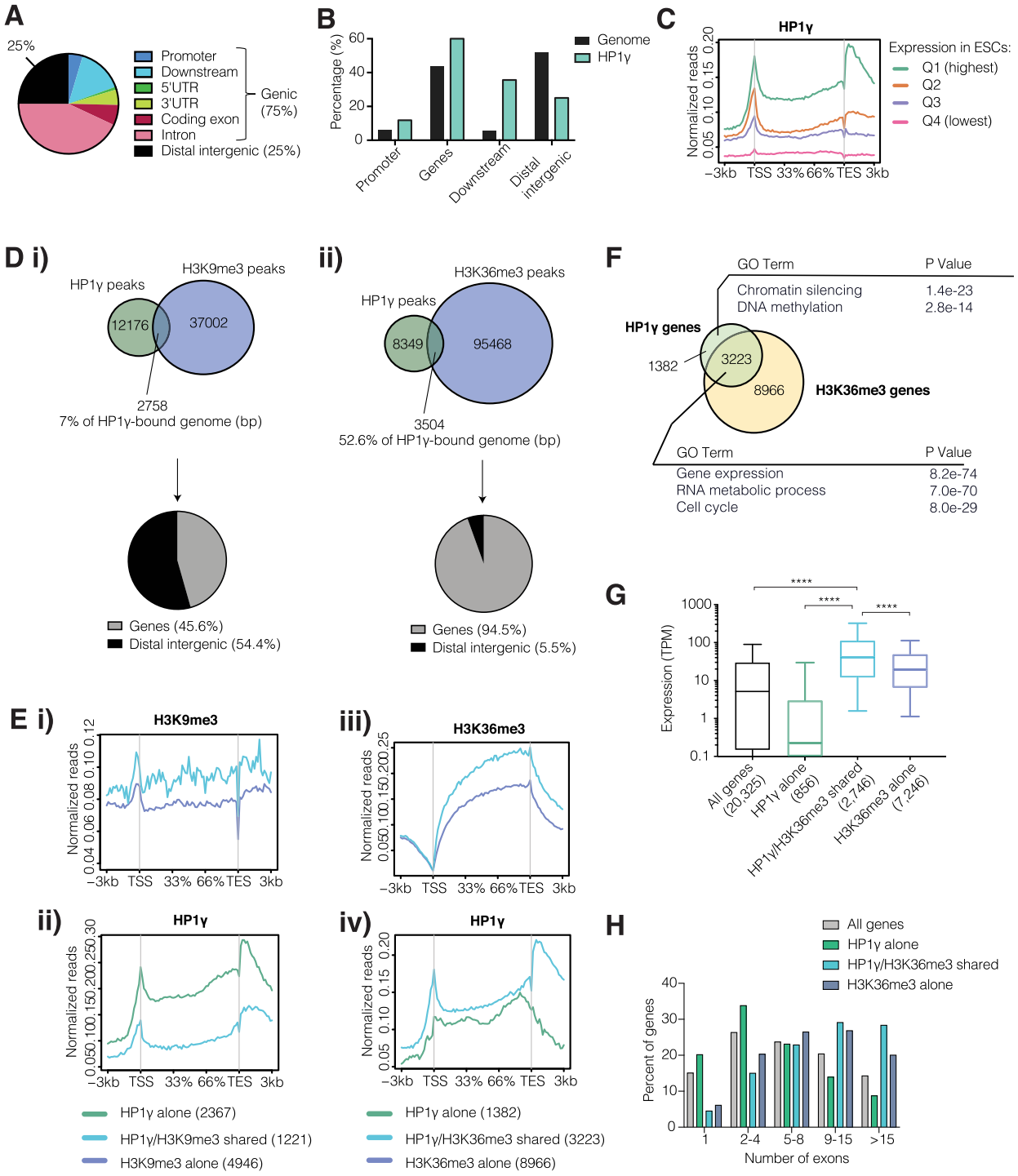

Figure 1. HP1γ and H3K36 methylation are reciprocally enriched on shared genes. (**A**) *Cis*-regulatory element annotation system (CEAS) analysis reveal the distribution of HP1γ binding sites according to different genomic regions. (**B**) Percentage of HP1γ-bound loci that are found in coding or non-coding regions as compared to genomic distribution. (**C**) Metaplot of averaged HP1γ ChIP enrichment (normalized per million mapped reads) on genes categorized into quartiles based on expression in ESCs. TSS: Transcription Start Site; TES: Transcription End Site. (**D**) Venn diagram (not to scale) of HP1γ ChIP peaks overlapped with (i) H3K9me3 or (ii) H3K36me3. Numbers indicate the number of regions within each category. Pie chart below each Venn show the distribution of shared peaks within or outside of genes. (**E**) (i-ii) Metaplot of H3K9me3 (i) and HP1γ (ii) ChIP enrichment (normalized per million mapped reads) on ±3 kb of genes that were bound by both HP1γ and H3K9me3 (shared), or by either HP1γ or H3K9me3 without the other (alone). (iii-iv) As above but for H3K36me3 (iii) and HP1γ (iv). Number of genes within each category are in parentheses. TSS: Transcription Start Site; TES: Transcription End Site. (**F**) Venn diagram of HP1γ-bound genes overlapped with H3K36me3-bound genes. Gene ontology (GO) enrichment categories are listed for genes bound by HP1γ alone, or both HP1γ and H3K36me3. Numbers indicate the number of genes within each category. (**G**) Boxplot of absolute expression (TPM) in ESCs for genes bound by HP1γ alone, H3K36me3 alone, or both HP1γ and H3K36me3. Boxplot hinges correspond to 10th and 90th percentile, with median center line. ****$P < 0.001$ assessed by unpaired Welch's *t*-test. (**H**) Bar graph shows the percent of genes within each category, binned by number of exons. The distribution of all genes is included as comparison.

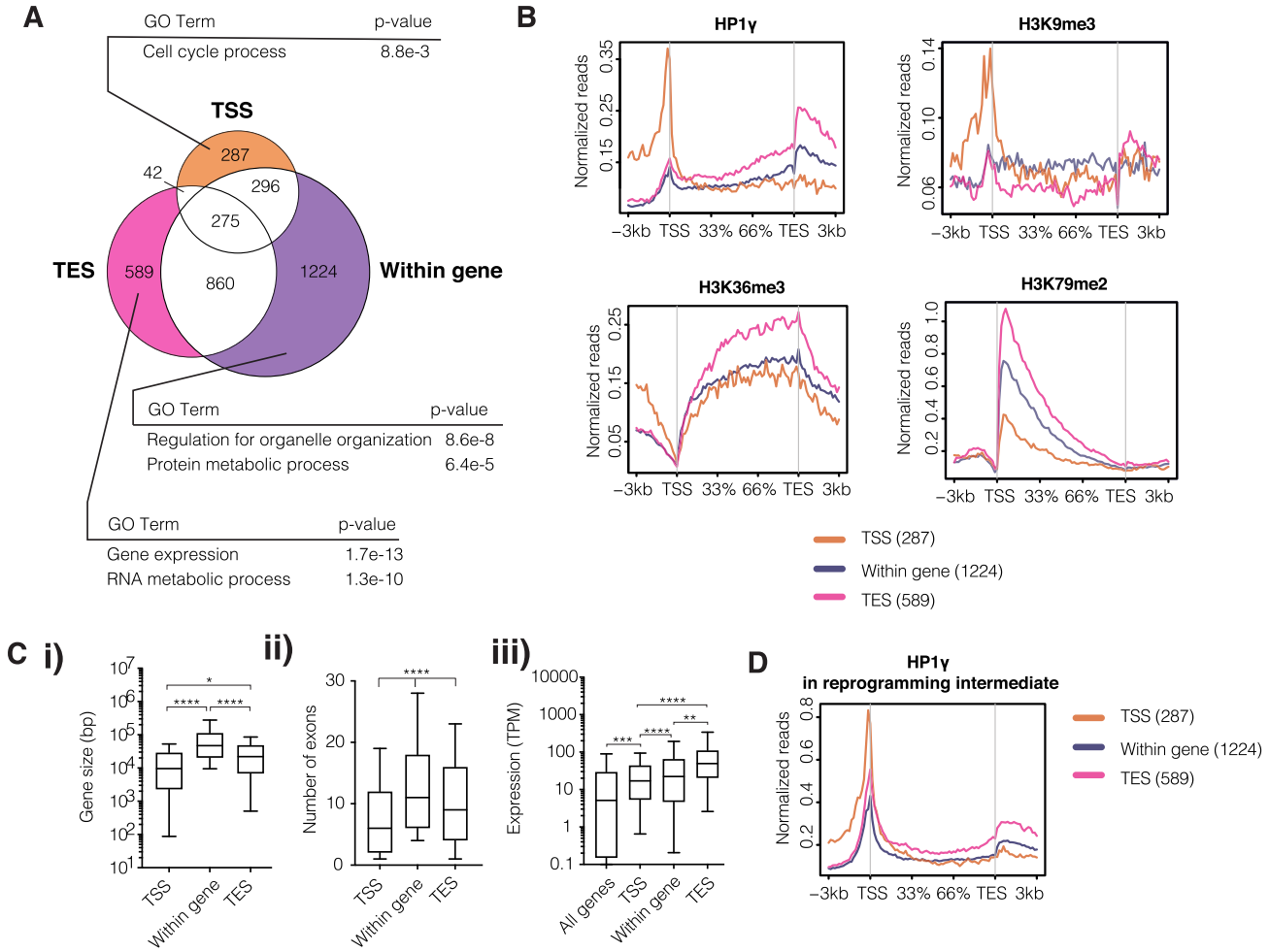

Figure 2. The presence of HP1γ in the gene body and TES, rather than TSS, correlate with higher H3K36me3 enrichment. (**A**) Venn diagram of HP1γ-bound genes in which HP1γ enrichment were found in the Transcription Start Site (TSS; −3kb to 500bp of start of a gene), Transcription End Site (TES; −500bp to +3kb of the end of a gene) or within gene body. Numbers indicate the number of genes that are found in each category. Gene ontology (GO) enrichment categories are listed for genes that are in the colored sections of the diagram: HP1γ-TSS, HP1γ-TES, or HP1γ-within gene. (**B**) Metaplot of HP1γ (top left), H3K9me3 (top right), H3K36me3 (bottom left) and H3K79me2 (bottom right) ChIP read counts (normalized per million mapped reads) on ±3 kb of genes that are HP1γ-TSS, HP1γ-TES, or HP1γ-within gene. TSS: Transcription Start Site; TES: Transcription End Site. (**C**) Boxplot shows the (i) gene size, (ii) number of exons and (iii) absolute expression in ESCs (TPM) of HP1γ-bound genes in different enrichment pattern. Boxplot hinges correspond to 10th and 90th percentile, with median center line. *$P < 0.05$, **$P < 0.01$, ***$P < 0.005$, ****$P < 0.001$ assessed by unpaired Welch's t-test. (**D**) Metaplot of HP1γ ChIP read counts (normalized per million mapped reads) on ±3 kb of genes that are HP1γ-TSS, HP1γ-TES, or HP1γ-within gene in reprogramming intermediate cells. TSS: Transcription Start Site; TES: Transcription End Site.

iments were performed. 75% of IP elute was used for immunoblotting for NSD1, while 15% of IP elute was used for immunoblotting for HP1γ.

Embryoid body differentiation

In vitro embryoid body differentiation from ESCs was performed as previously described (45). ESCs were MEF-depleted twice and washed with 1X DPBS before plated as hanging droplets of 500 cells per 20ul in differentiation media (ESC media but without LIF). After two days, embryoid bodies in droplets were transferred into each well of 96-well ultralow attachment plate filled with 180 ul of differentiation media. On day 8, embryoid bodies were transferred into each well of 48-well plate on gelatin with an additional of 300 ul differentiation media. Spontaneous cardiomyocyte contractions were scored on day 10, and cells were harvested for RT-qPCR.

Alkaline Phosphatase (AP) staining

Cells were fixed with 4% paraformaldehyde for 20 min, washed with water, followed by staining with 1 mg/ml Fast Red TR Salt Hemi (Sigma, F8764) in water with 40ul/ml Naphthol AS-MS phosphate (Sigma, 855–20 ml).

Statistical analysis

Significance was calculated using the two-tailed t-test function in Graphpad Prism.

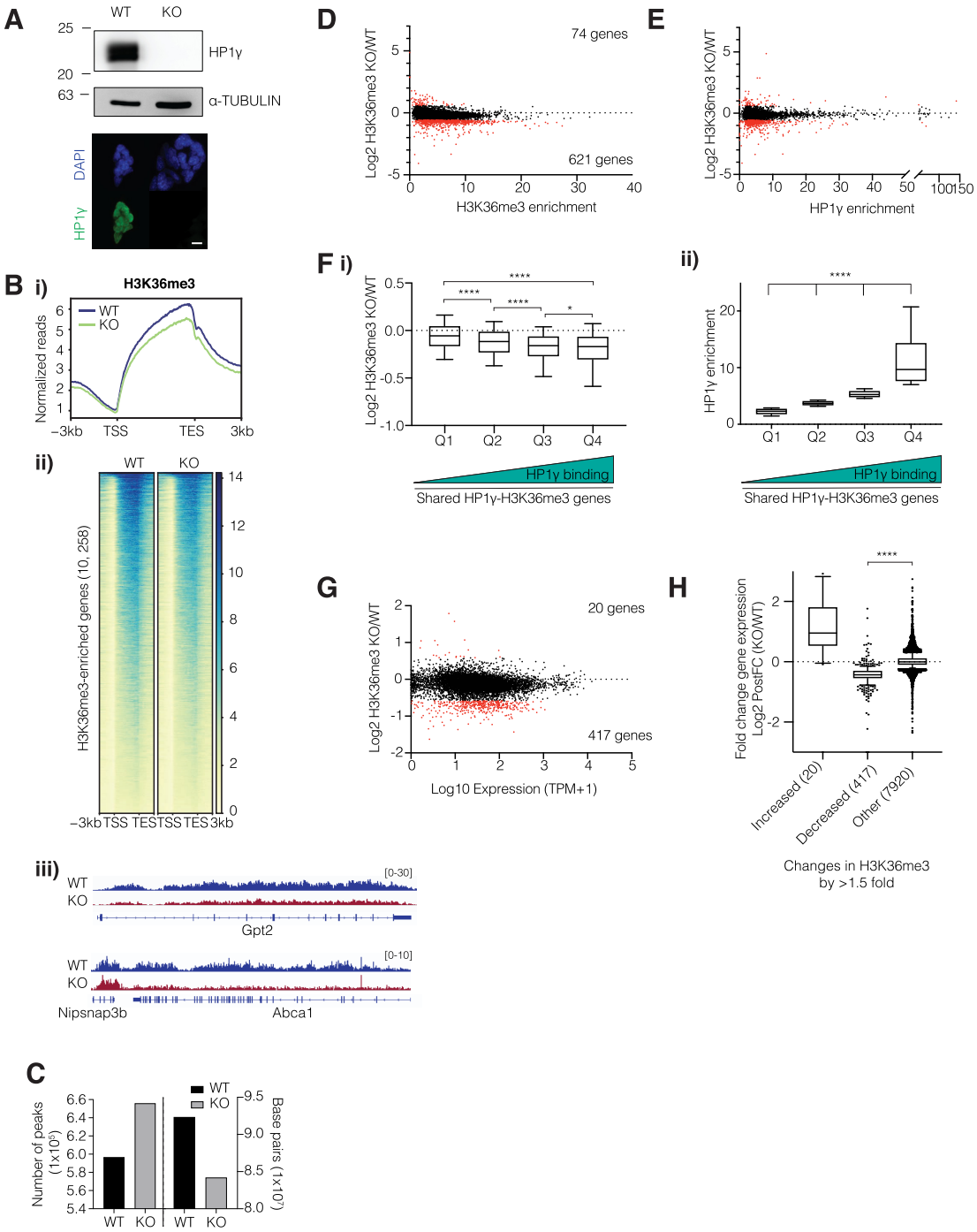

Figure 3. Loss of HP1γ impacts H3K36me3 enrichment. (**A**) HP1γ protein levels as shown by western blot (top) and immunofluorescence (bottom) of WT and HP1γ KO ESCs. Scale bar = 10 μm. (**B**) Metaplot (i) and heatmap (ii) of H3K36me3 ChIP enrichment (normalized to human mapped reads) on ±3 kb of H3K36me3-bound genes in WT and HP1γ KO ESCs. TSS: Transcription Start Site; TES: Transcription End Site. (iii) Representative H3K36me3 enrichment profile in WT and HP1γ KO ESCs. (**C**) Total number of H3K36me3-enriched peaks (left) and base pairs of H3K36me3-enriched genome coverage (right) for shared HP1γ/H3K36me3 genes in WT and HP1γ KO ESCs. (**D**) Scatterplot of fold change of H3K36me3 enrichment (KO over WT) on genes plotted against absolute H3K36me3 enrichment (reads) in WT. Numbers and highlighted in red are genes that changed in H3K36me3 enrichment by at least 1.5 fold. (**E**) Scatterplot of fold change of H3K36me3 enrichment (KO over WT) on genes plotted against absolute HP1γ enrichment (reads) in WT. Highlighted in red are genes that changed in H3K36me3 enrichment by at least 1.5-fold. (**F**) (i) Boxplot of fold change in H3K36me3 enrichment (KO over WT) of genes from each quartile of HP1γ-H3K36me3 shared genes. Quartiles are determined based on HP1γ enrichment (ii), where Q1 = lowest, and Q4 = highest. Boxplot hinges correspond to 10th and 90th percentile, with median center line. $*P < 0.05$, $**P < 0.01$, $***P < 0.005$, $****P < 0.001$ assessed by unpaired Welch's *t*-test. (**G**) Scatterplot of fold change of H3K36me3 enrichment (KO over WT) on genes plotted against absolute expression in WT ESCs. Highlighted in red are genes that changed in H3K36me3 enrichment by at least 1.5-fold. (**H**) Boxplot showing fold change in gene expression (KO over WT) in genes that have an increase or decrease in H3K36me3 enrichment by at least 1.5-fold. Genes that are in between this threshold are also shown (Other). Boxplot hinges correspond to 10th and 90th percentile, with median center line. $*P < 0.05$, $**P < 0.01$, $***P < 0.005$, $****P < 0.001$ assessed by unpaired Welch's *t*-test.

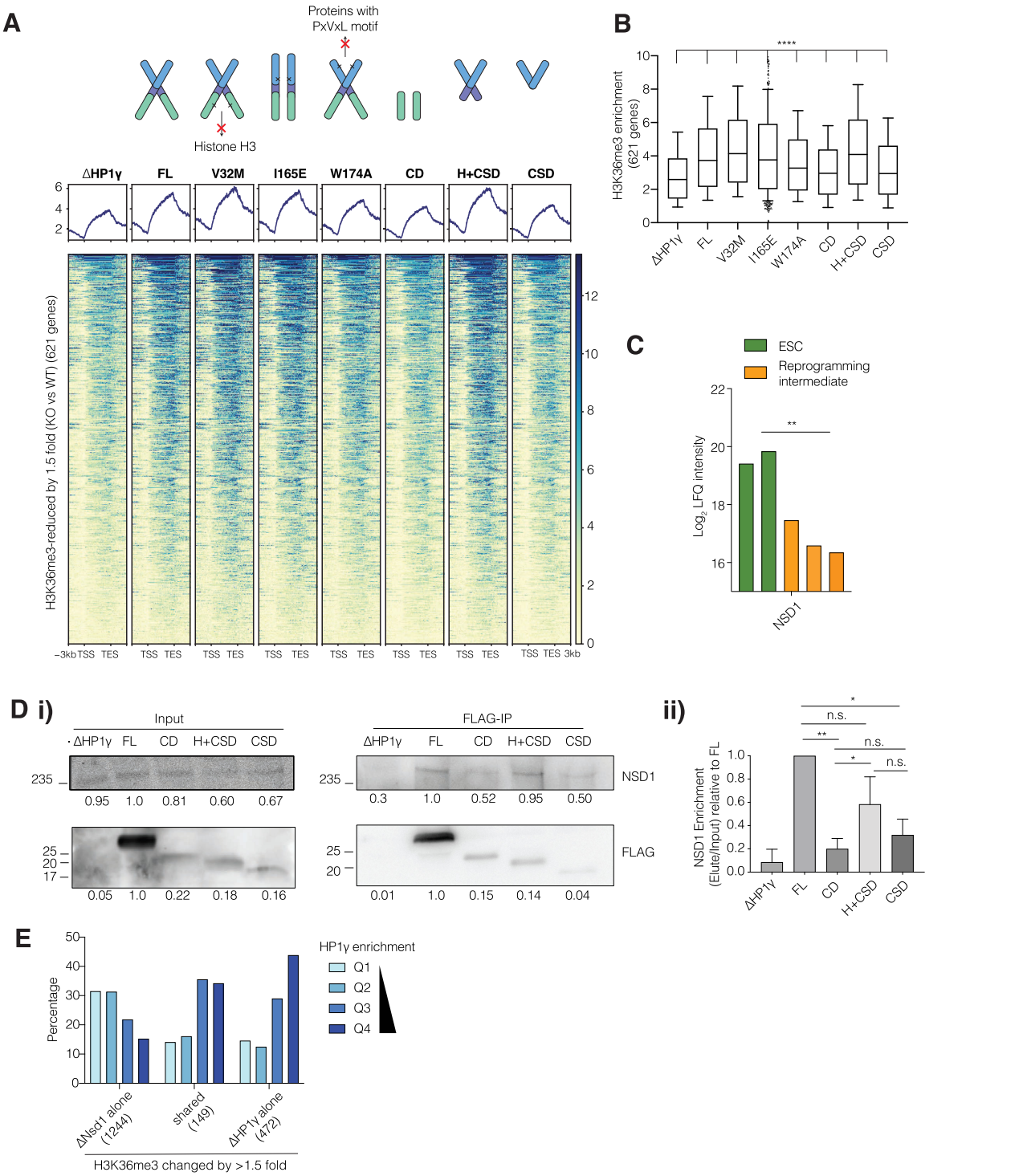

Figure 4. The protein and DNA/RNA-binding ability of HP1γ is important in maintaining H3K36me3 enrichment. (**A**) Top: Cartoon depicting each HP1γ rescue constructs. Bottom: Heatmap of H3K36me3 ChIP enrichment (normalized to human mapped reads) in either uninduced control (ΔHP1γ) or rescue HP1γ constructs. Regions shown are ±3 kb of genes that had reduced H3K36me3 by at least 1.5-fold in HP1γ KO over WT ESCs. Metaplots of averaged enrichment are shown above each heatmap. TSS: Transcription Start Site; TES: Transcription End Site. (**B**) Boxplot of averaged read counts of H3K36me3 ChIP reads from (**A**). Boxplot hinges correspond to 10th and 90th percentile, with median center line. $^{****}P < 0.001$ assessed by paired two-sided Wilcoxon test. (**C**) Bar graph showing differential enrichment of NSD1 by FLAG-HP1γ in ESCs and reprogramming intermediate. Data from two or three independent immunoprecipitation experiments for ESC and reprogramming intermediates, respectively, are presented. $^{**}P < 0.01$ assessed by Student's *t*-test. (**D**) (i) Immunoprecipitation (IP) with FLAG antibody with immunoblot for NSD1 and FLAG in either uninduced control (ΔHP1γ) or rescue HP1γ constructs. (ii) Quantitation of NSD1 enrichment (elute/input) relative to full-length (FL) in either uninduced control (ΔHP1γ) or rescue HP1γ constructs. Error bars represent standard deviation of the mean from three independent immunoprecipitation experiments. $^{*}P < 0.05$, $^{**}P < 0.01$ assessed by Student's *t*-test. n.s., not significant. (**E**) Percentage of genes of those that had reduced H3K36me3 by at least 1.5 fold in NSD1 KO ESCs, HP1γ KO ESCs, or both, in quartiles based on HP1γ enrichment, where Q1 = lowest, and Q4 = highest.

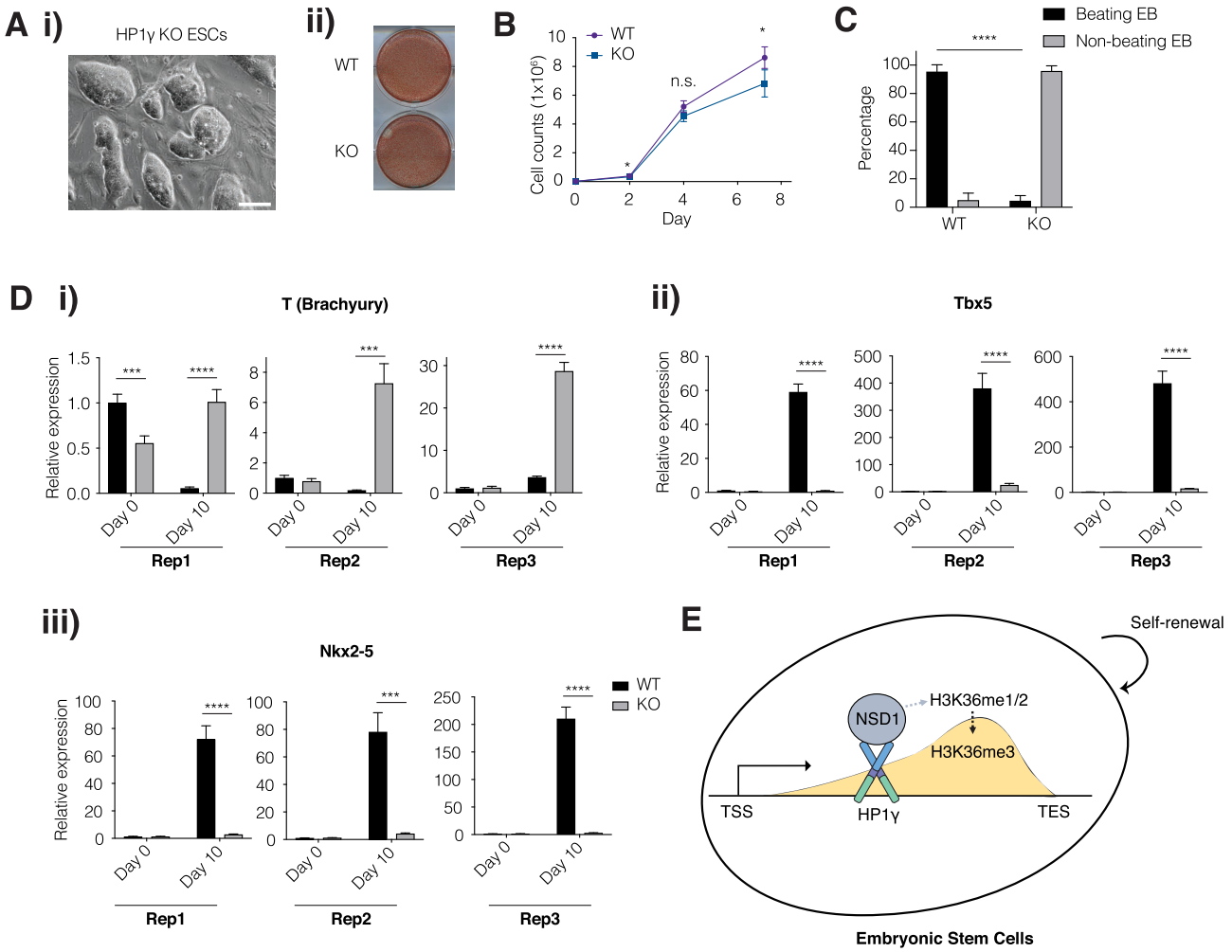

Figure 5. HP1γ KO ESCs exhibit impaired self-renewal and differentiation. (**A**) (i) Bright field image of HP1γ KO ESCs. Scale bar, 100 μm. (ii) AP staining of WT and HP1γ KO ESCs. (**B**) Cell counts of WT and HP1γ KO ESCs. Error bars indicate standard deviation of the mean from counts of three replicates. $*P < 0.05$ assessed by Student's *t*-test. n.s., not significant. (**C**) Percentage of beating and non-beating cardiomyocytes visualized at day 10 of embryoid body differentiation in WT and HP1γ KO cells. Error bars represent standard deviation of the mean from four independent differentiation experiments. $****P < 0.001$ assessed by Student's *t*-test. (**D**) Relative expression levels of mesoderm lineage markers Brachyury (i), Tbx5 (ii) and Nkx2-5 (iii) on day 0 and day 10 of embryoid body differentiation in WT and HP1γ KO cells. Error bars represent standard deviation of the mean from three technical replicates. Shown are results from three independent embryoid body differentiation experiments. $*P < 0.05$, $**P < 0.01$, $***P < 0.005$, $****P < 0.001$ assessed by Student's *t*-test. (**E**) Proposed model depicting regulation of H3K36me3 by HP1γ. HP1γ binding is prevalent in gene bodies of actively transcribing genes, which correlated with H3K36me3 enrichment towards the end of genes (yellow shadow). HP1γ interacts with NSD1 through its Hinge+Chromoshadow domain. NSD1 mediates mono- and di-methylation of H3K36, which serve as potential substrate for tri-methylation.

RESULTS

HP1γ enhances enrichment of H3K36me3 and H3K79me2 at shared genes

The HP1 proteins are known as 'readers' of H3K9 methylation, a histone modification associated with gene repression in numerous model systems (46–49). However, our previous work suggests that HP1γ interacts with transcription elongation-associated factors such as the FACT complex and histone H3.3 more than HP1α or HP1β (21). To investigate how HP1γ is associated with transcription elongation, we first analyzed the genomic distribution of HP1γ in ESCs (5). HP1γ binding is predominantly in genic (75% of peaks) as compared to distal intergenic regions (25%) (Figure 1A). In the genic regions, HP1γ peaks were ∼6 fold (35.6% versus 5.4%) more prevalent downstream of genes as compared to the genome average (Figure 1B). In a metaplot of averaged ChIP enrichment across gene bodies, HP1γ binding profile showed a peak at the Transcription Start Site (TSS), which gradually increases within the gene body and even further downstream of genes, after the Transcription End Site (TES) (Figure 1C). Greater HP1γ enrichment is correlated with higher gene expression since most HP1γ binding occurs in genes that belong to the top two quartiles (Q1 and Q2) as compared to genes that are lowly or not expressed (Q4) (Figure 1C).

We next asked whether HP1γ binding in coding regions correlates with particular histone modifications. Despite the fact that HP1γ is known to be a reader of H3K9me3, only 7% of the HP1γ-bound genome (in base pairs) is shared with H3K9me3 (Figure 1D, i, Supplementary Figure S1A). In contrast, H3K36me3 and H3K79me2, both of

which are implicated with transcriptional elongation and are enriched on highly expressed genes (22–24,50), have a greater co-occurrence with HP1γ (Supplementary Figure S1A). About 52.6% of HP1γ-bound genome (in base pairs) is shared with H3K36me3, whereas 37% is shared with H3K79me2 (Figure 1D, ii, Supplementary Figure S1A, S1B, i). As expected, HP1γ-bound regions do not correlate with enrichment for the polycomb-mediated repressive H3K27me3 histone modification but have shared regions with H3K4me3, a mark associated with active transcription (Supplementary Figure S1A). Despite some TSS enrichment of HP1γ, the co-occurrence with H3K4me3 is small because they are targeted to different genes. The percentage of HP1γ-bound regions shared with H3K36me3 is the highest compared to other histone modifications analyzed (Supplementary Figure S1A). Of the HP1γ/H3K9me3 shared regions, only 45.6% are within genes (Figure 1D, i) while >90% of the HP1γ/H3K36me3 and HP1γ/H3K79me2 shared regions are predominantly in genes (Figure 1D, ii, Supplementary Figure S1B, i). Thus, contrary to the prevalent view (18), it is unlikely that the large proportion of genic HP1γ is recruited in response to the presence of H3K9me3, the canonical histone modification associated with HP1 proteins.

To further investigate the influence of HP1γ on the histone marks, we separated genes based on the pairwise presence of HP1γ and each modification. HP1γ and H3K9me3 have previously been found on genes and implicated in RNA splicing (18,51). Comparing genes bound by HP1γ with or without H3K9me3, the presence of HP1γ does correlate with higher H3K9me3 enrichment, although the signal is low and noisy within genes (Figure 1E, i, Supplementary Table S1). However, the presence of H3K9me3 does not correlate with high HP1γ enrichment, as HP1γ binding is drastically higher on HP1γ-bound genes without H3K9me3 (Figure 1E, ii, Supplementary Table S1). The HP1γ bound genes that do not contain H3K9me3 are higher expressed than shared genes (Supplementary Figure S1D, Table S2).

The enrichment of H3K36me3 is higher on genes that are also bound by HP1γ, suggesting that the presence of HP1γ enhances H3K36me3 enrichment (Figure 1E, iii, Supplementary Table S1). Reciprocally, HP1γ is higher enriched on genes whose histones are modified for H3K36me3 (Figure 1E, iv, Supplementary Table S1). Similar analysis performed with H3K79me2 revealed the same correlation pattern (Supplementary Figure S1B, ii, Table S1). Genes with both HP1γ/H3K36me3 were enriched with functional categories such as gene expression, RNA and DNA metabolic processes and cell cycle (Figure 1F). The genes that were bound by HP1γ and enriched for H3K36me3 also had the highest expression levels (Figure 1G).

In contrast, the 1382 genes that were bound by HP1γ alone did not display increased density after the TES (Figure 1E, iv, Supplementary Table S1). Gene Ontology (GO) and GSEA analysis of such HP1 genes, that were without H3K36me3, revealed categories such as chromatin silencing and DNA methylation, and mainly consisted of histones and small RNAs (Figure 1F, Supplementary Figure S1C, i). These genes were largely repressed (Figure 1G).

The pattern of H3K36me3 enrichment always peaked at the TES irrespective of HP1γ binding (Figure 1E, iii). The H3K36me3 genes that did not have HP1γ were less expressed (Figure 1G) and enriched for MYC targets by GSEA analysis (Supplementary Figure S1C, ii). While H3K36me3 is generally enriched on multi-exonic genes (Supplementary Figure S1E), genes with the most exons (>15) had both HP1γ and H3K36me3 enrichments (Figure 1H). HP1γ and H3K36me3 enriched genes functioned in housekeeping and were associated with c-MYC control (Supplementary Figure S1C, iii). Notably when normalized to gene length, the highest HP1γ enrichment was found on genes of mid-length (Supplementary Figure S1F). Since c-MYC is known to improve elongation rates (52,53), together these characteristics suggest that HP1γ may enhance elongation and high expression of large multi-exonic genes that are controlled by c-MYC.

H3K36me3 is highly enriched when HP1γ is bound at TES rather than TSS

Given that the pattern of HP1γ binding in genes (Figure 1C) could vary with H3K36me3 enrichment (Figure 1E), we asked whether we could categorize the genes according to specific HP1γ-enrichment pattern. The composite pattern of binding could arise either by each gene having the same profile that is then averaged in a metaplot or genes having different patterns of binding that can be sorted based on the location of highest enrichment. We categorized HP1γ ChIP peaks annotated to the TSS, within gene, or the TES, and grouped the HP1γ-bound genes based on the location of HP1γ enrichment (Figure 2A, Materials and Methods). We focused on three major patterns of HP1γ: genes that had HP1γ only around the TSS (TSS), genes with HP1γ throughout the gene body and onwards (within gene) and genes with HP1γ enrichment after the TES (TES) (Figure 2A and B).

We next asked how the different HP1γ binding patterns correlate to the enrichment of elongation histone marks. Interestingly, both H3K36me3 and H3K79me2 enrichment are highest on genes with HP1γ bound at the TES, rather than the TSS (Figure 2B). In stark contrast, H3K9me3 enrichment correlates with HP1γ at TSS-bound genes (Figure 2B). Moreover, as compared to HP1γ-TSS genes, genes with HP1γ at TES and within gene tend to be larger in size and have higher number of exons (Figure 2C), concomitant with the observation of HP1γ/H3K36me3 shared genes (Supplementary Figure S1E). Interestingly, each pattern correlated with differential expression levels. Genes with HP1γ-TSS are generally expressed lower than genes with HP1γ-TES (Figure 2C) and are enriched in GO categories such as cell cycle process (Figure 2A). On the other hand, genes with HP1γ at TES or throughout the gene are enriched in categories such as DNA and RNA metabolic process (Figure 2A).

Somatic cells can be reprogrammed to induced pluripotent stem cells (iPSCs) that are functionally identical to ESCs. In intermediates of this reprogramming process, HP1γ binding does not conform the patterns found in ESCs; instead, it is concentrated mostly at the TSS and less

so in gene bodies (Figure 2D). A similar TSS enrichment of HP1γ observed in neural progenitor cells (54). Therefore, the non-TSS HP1γ binding profiles in ESCs may be unique to pluripotent stem cells.

Loss of HP1γ impacts H3K36me3 enrichment and expression of H3K36me3-bound genes

The correlation of HP1γ with H3K36me3 suggests that HP1γ may be important for the deposition of this elongation-associated modification. We therefore generated ESC cell lines in which endogenous HP1γ was deleted via CRISPR-Cas9 technology and validated the loss of HP1γ protein by western blot and immunofluorescence (Figure 3A). We next performed ChIP-seq of H3K36me3 in wild type (WT) and HP1γ KO ESCs. Interestingly, there was a reduced global enrichment of H3K36me3 in HP1γ-KO ESCs compared to WT cells (Supplementary Figure S2A, Methods). The reduced H3K36me3 enrichment is particularly apparent on HP1γ-TES genes (Supplementary Figure S2B). To precisely quantitate this loss of H3K36me3, we normalized the reads to a human spike in control (Methods, Figure 3B). We quantified H3K36me3 enrichment on genes in both WT and HP1γ KO and examined the profiles of genes that differed by at least 1.5-fold. There were almost 10-fold more genes that exhibited a significant decrease in H3K36me3 (621 genes) compared to an increase (74 genes). Among genes that retained H3K36me3 in the HP1γ KO, the number of peaks associated with H3K36me3 increased. However, the size of the peaks occupied by the modification as reflected in the base pairs of genome coverage was reduced (Figure 3C). Hence HP1γ promotes sustained enrichment of H3K36me3.

This difference in H3K36me3 levels was not due to altered expression of the H3K36 methyltransferases or demethylases as measured by RNA-Seq (Supplementary Figure S2C). There was no detectable change in levels of the H3K36me3 modification itself by western blot upon HP1γ loss (Supplementary Figure S2D). Thus, in the HP1γ KO there is a specific loss of H3K36me3 on a subset of genes.

We next sought to determine whether there was specificity in the genes that were affected in their H3K36me3 enrichment in the HP1γ KO. H3K36me3 loss in HP1γ KO cells was found in genes with a wide range of H3K36me3 enrichment in WT ESCs, suggesting that the effect is not specific to high H3K36me3-enriched genes (Figure 3D). Similarly, both low and high HP1γ-enriched genes showed a change in H3K36me3 enrichment (Figure 3E).

To better quantify the genes most vulnerable to H3K36me3 loss, we divided H3K36me3-enriched genes into four quartiles, with Q1 having little to no HP1γ enrichment and Q4 having the highest. We found a greater proportion of genes with high HP1γ (Q4) that are reduced in H3K36me3 by 1.5-fold as compared to genes in the lower quartiles (Q1 and Q2) (Supplementary Figure S2E). Since the range of HP1γ binding strength on H3K36me3-enriched genes is very broad (Figure 3E) and genes within the lower quartiles could include those that are not bound by HP1γ, we further narrowed our analysis to genes that are bound both by HP1γ and H3K36me3 (Figures 1F and 3F). Remarkably, high HP1γ enrichment (genes in Q4) was correlated with the most pronounced reduction in H3K36me3 levels upon HP1γ loss (Figure 3F).

Given that both HP1γ and H3K36me3 are associated with highly expressed genes, we examined the absolute expression of genes in WT ESCs with the change in H3K36me3 enrichment upon HP1γ loss. Interestingly, the change in H3K36me3 does not seem to be specific to highly-expressed transcripts, as even those with low absolute expression were affected (Figure 3G). Next, we sought to determine whether the loss of HP1γ would affect expression of genes. We therefore performed RNA-seq in HP1γ KO and WT ESCs while including a spike-in human RNA for normalization control. As previously reported (5,55,56), the depletion of HP1γ does not elicit a great change in the number or magnitude of gene expression, but there are more differentially upregulated genes compared to down (601 vs. 293) (Supplementary Figure S2F). Despite the lack in expression changes globally, genes that had reduced H3K36me3 upon HP1γ loss are specifically downregulated in expression (Figure 3H, Supplementary Figure S2G).

Both HP1γ and H3K36me3 can influence the incorporation of exons (16,20,55,57,58) and we have also found an interaction of HP1γ with some splicing components (21). However, we did not find any significant differences in the number of retained introns or skipped exons (data not shown).

Taken together, these results suggest that HP1γ is important in maintaining H3K36me3 levels on specific genes, and that loss of HP1γ elicits a reduction in H3K36me3 enrichment as well as steady-state expression.

HP1γ control of H3K36me3 enrichment is dependent on its protein and DNA/RNA binding ability

The HP1 proteins have a conserved domain structure. Besides the chromodomain that binds H3K9me3 (9,59), there is a chromoshadow domain involved in heterodimerization and protein binding (60), as well as an unstructured hinge domain for nucleic acid interaction (61–63). We therefore wondered which of the domains or functions of HP1γ that is important in maintaining H3K36me3 enrichment on genes. We generated clones of doxycycline-inducible Flag-tagged HP1γ integrated at a single locus in the HP1γ KO ESCs. The rescue ESC lines included full-length (FL) or single amino acid mutations which abolish binding to histone H3 (V32M) (8), dimerization (I165E) (64) or interaction to proteins with PxVxL motifs (where x = any amino acid) (W174A) (64) in the context of the FL. In addition, we generated domain mutants consisting of Chromodomain (CD), Chromoshadow domain (CSD) and Hinge plus Chromoshadow domain (H+CSD). While all the full-length proteins, including the single amino acid mutants, were expressed at equivalent levels, the domain mutants were about 2.5–10-fold less expressed than full length (Supplementary Figure S3A and B) with equivalent levels of induction. This indicates that the domain mutants alone might be unstable without the remaining structure of the full-length protein.

We performed ChIP-seq of H3K36me3 upon induction of each of the rescue constructs, as well as an uninduced control (ΔHP1γ). On genes that were affected in

H3K36me3 upon HP1γ loss, the full-length HP1γ construct showed an increase in H3K36me3 enrichment compared to uninduced control (Figure 4A and B). Interestingly, single amino acid mutants V32M as well as I165E are also able to maintain H3K36me3 enrichment at a similar level to that of full length HP1γ (Figure 4A and B). This indicates that HP1γ binding to histone H3, as well as its dimerization properties are not necessary to maintain H3K36me3. On the other hand, the W174A mutant maintains H3K36me3, albeit to a lesser extent than full length HP1γ, indicating that the interaction of HP1γ to other proteins through their PxVxL motif is partially required (Figure 4A and B). This result also raises an interesting conundrum because it is thought that dimerization is required for binding to PxVxL motif containing proteins. Therefore in the context of maintenance of H3K36me3 levels, the PxVxL interaction surface may be able to function independent of dimerization.

More drastic differences were observed when we analyzed the ESC lines rescued with the domain mutant constructs. Neither the CD alone nor CSD alone were able to rescue H3K36me3 enrichment on these genes (Figure 4A and B). Interestingly, the H+CSD showed a robust restoration of H3K36me3 similar to the FL (Figure 4A and B) despite the lower levels of protein expression as compared to the full-length wild type clone (Supplementary Figure S3A and B). This restoration was most significant in the group of genes that were direct targets of HP1γ (Supplementary Figure S3C). We could not derive clones that only contained the hinge region. Since the CSD includes the region for binding PxVxL motif proteins and the W174A mutant can partially mimic FL function, together these results indicate that binding of HP1γ to both nucleic acid and proteins is sufficient to maintain H3K36me3 enrichment.

We next sought to determine the mediator of HP1γ effects on H3K36me3 enrichment. Re-examining our proteomics analyses of HP1γ, we found that HP1γ interacts with NSD1 (21), a H3K36me2 methyltransferase (65–67). We did not recover any peptides from the H3K36me3 enzyme SETD2 (21). The interaction of NSD1 with HP1γ is significantly enriched in ESCs, but not in reprogramming intermediates (Figure 4C), concomitant with the gene body enrichment of HP1γ that is found in the former but not the latter (Figure 2B and D). We validated this interaction in ESCs by performing FLAG immunoprecipitation (FLAG-IP) of the full-length rescue construct of HP1γ followed by western blot. NSD1 co-immunoprecipitated with FLAG-FL, but not in the uninduced control (Figure 4D). Given the dramatic differences in H3K36me3 enrichment abilities of the domain mutants, we wondered whether there might also be different binding ability to NSD1. We therefore performed FLAG-IPs in the domain rescue ESC lines. Interestingly, despite having lower protein levels as compared to FL, the H+CSD domain mutant is able to co-immunoprecipitate NSD1, comparable to that of FL (Figure 4D). Importantly, the CD domain only showed a weak interaction to NSD1, despite having similar levels of FLAG-HP1γ immunoprecipitated as compared to H+CSD (Figure 4D, Supplementary Figure S3D). On the other hand, CSD alone also co-immunoprecipitated NSD1 at lower levels, although this could be due to the low enrichment of the bait itself (Figure 4D, Supplementary Figure S3D). Thus, minimally the CSD domain is sufficient to interact with NSD1, with an additional contribution of the Hinge to the interaction or the stability of the truncated protein. Taken together, these results implicate the interaction of the Chromoshadow domain of HP1γ with NSD1 in maintaining H3K36me3 levels in ESCs. Since NSD1 does not have a PxVxL motif, this interaction is likely to represent the non-PxVxL portion of HP1γ mediation of H3K36me3 enrichment.

In NSD1 KO ESCs, the major effect is on H3K36me2 enrichment and far fewer genes are affected in H3K36me3 levels (30). To further evaluate the NSD1 contribution to the maintenance of H3K36me3 we analyzed the enrichment of this mark in NSD1-KO ESCs (30). In NSD1-KO ESCs, about 3-fold more genes were decreased for H3K36me3 as compared to increased and this change was located largely at the end of the gene (Supplementary Figure S3E and F). We found that about 10% of genes showing a decrease in H3K36me3 were shared with HP1γ (Supplementary Figure S3G). The genes that were commonly reduced in H3K36me3 in both the NSD1-KO and HP1γ KO were highly enriched for HP1γ (Figure 4E). Together these results implicate a co-operative interaction between NSD1 and HP1γ to maintain H3K36me3 at genes where there is high enrichment of HP1γ.

HP1γ KO ESCs exhibit impaired ESC self-renewal

Given that loss of HP1γ has a molecular phenotype in dysregulating transcription elongation marks such as H3K36me3, we wondered whether there would also be a cellular phenotype. HP1γ KO ESCs are morphologically indistinguishable to WT ESCs (Figure 5A). However, self-renewal is compromised with HP1γ KO ESCs dividing slower than WT (Figure 5B). This is likely due to the decreased expression of housekeeping genes that compromises pluripotency homeostasis, such as Cdk10 and Mcm5.

Next, we sought to investigate the role of HP1γ during differentiation. We performed an undirected embryoid body differentiation, that generates cells from all three germ layers. Phenotypically, both WT and HP1γ KO ESCs are able to form comparable embryoid bodies in culture. Surprisingly, we observed a difference in the presence of beating cardiomyocytes which usually emerges around day 10 of embryoid body differentiation. While 95% of our WT-derived embryoid bodies produced beating cardiomyocytes, in stark contrast, almost no beating was observed in embryoid bodies derived from HP1γ KO ESCs (Figure 5C). Thus, loss of HP1γ impaired differentiation to cardiac mesodermal fate. We therefore quantitated the expression of mesodermal lineage marker genes. Intriguingly, Brachyury, an early mesodermal marker which is usually expressed early during differentiation and is reduced in expression shortly after, continues to be maintained in HP1γ KO cells (Figure 5D). Concomitantly, expression of cardiac mesoderm markers (Tbx5, Nkx2.5) was absent in HP1γ KO derived EBs (Figure 5D). These observations suggest that the absence of HP1γ compromises differentiation toward the mesodermal lineage.

DISCUSSION

The HP1 family of proteins play instrumental roles in a myriad of functions, including DNA damage repair, RNA processing and splicing (10). While HP1γ binds to promoters of repressed genes (54,56), they can also be found in gene bodies of actively transcribed genes (16,18,19). In this study, we have demonstrated that the binding of HP1γ in the gene body correlated to high enrichment of H3K36me3 and H3K79me2. HP1γ enrichment pattern in genes overlapped more with H3K36me3, which also peaked towards the end of genes. This high HP1γ enrichment downstream of genes may be explained by its role in transcription termination especially of highly transcribed genes that are prone to R-loop formation (68). Corroborating this notion of function in robust transcription elongation, our ChIP-seq analyses revealed that loss of HP1γ affected both H3K36me3 enrichment and expression of shared genes.

In heterochromatin, HP1γ binds to H3K9me3 via the highly conserved amino-terminal chromodomain (9). On the other hand, euchromatic HP1γ and H3K9me3 have been found to be present in gene bodies of active genes in HeLa cells and implicated with a role in splicing (51). In contrast, our analyses in mouse ESCs show that HP1γ correlates with H3K9me3 enrichment when present at the TSS, rather than in gene bodies. Therefore, HP1γ may be recruited to gene body locations with a novel mechanism. Our finding that the Hinge plus Chromoshadow domain is sufficient to maintain H3K36me3 enrichment further demonstrate a role of HP1γ that is independent of its H3K9me3-recognizing chromodomain.

The chromoshadow domain of HP1 is important for its dimerization and interaction with other proteins, many of which contains a PxVxL motif (60). The Hinge region, which is less conserved, is not well studied but has been shown to interact with DNA and/or RNA (62,63). HP1γ does not directly interact with H3K36me3 in in-vitro pull-down assays with H3 peptides (16,17). This would suggest that the effect of HP1γ on H3K36me3 enrichment may be due to interaction with other proteins. We have identified NSD1, an H3K36 methyltransferase as an important mediator of this interaction that can be maintained with just the H+CSD domains of HP1γ. Although NSD1 is a mono- and di-methylase of H3K36, these modifications serve as a potential substrate for tri-methylation (65–67). In fact, depletion of NSD1 reduces the levels of all three methylation states of H3K36 (66). HP1γ also interacts with RNAPII (17–19) and it is known that RNAPII interacts with SETD2, a H3K36me3 methyltransferase (24–26,69). Given that only 10% of the genes that lose H3K36me3 in the NSD1-KO are common with the HP1γ-KO, this suggests that these genes are part of a specific regulatory context that may be in play at long highly transcribed genes. NSD1 was not recovered in the interactome of HP1α or HP1β. Furthermore the hinge and chromoshadow domain of HP1γ that are important for its NSD1 interaction have less conservation to HP1β (73% identity, 85% similarity) as compared to chromodomain (82% identity, 88% similarity). Interestingly in Drosophila, HP1α has an interaction with KDM4A, a H3K36me3 demethylase through its chromoshadow domain, and loss of the HP1α leads to an increase in H3K36me3 (70). Thus, the HP1 proteins may have opposing functions in maintaining H3K36me3 at transcribed genes.

Our finding that loss of HP1γ impairs differentiation and growth of ESCs seems to be contradictory to the mouse KO of HP1γ that is viable although infertile. One explanation for this could be that the viable HP1γ KO mouse was a result of a cross between heterozygote parents (11,12). Therefore, there is a possibility that the wild-type maternal transcript is present and required during early development of the homozygous null embryos. The levels of HP1γ also seem to be important for differentiation potential since a previous study demonstrated that overexpression of HP1γ in P19 embryonal carcinoma cells exhibited significant morphological differentiation, which included beating cardiomyocytes (71).

CONCLUSION

Taken together, our study reveals a role for HP1γ in maintaining genic H3K36me3 levels and self-renewal capability in ESCs, that is independent of its H3K9me3-recognizing ability (Figure 5E).

DATA AVAILABILITY

The datasets supporting the conclusions of this article are available in the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) repository with the accession code GSE144863.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Dr Coral Wille for alternative splicing analysis, Amulya Suresh and Nikita Patel for technical assistance; members of the Sridharan lab for discussion and critical reading of the manuscript.
Author contributions: N.Z.Z. participated in the conception of the project and design of the study, conducted all experiments and bioinformatic analyses, analyzed and interpreted data, prepared figures and wrote the paper. R.S. conceived the project, procured funding, interpreted data and wrote the paper.

FUNDING

NIH [R01GM113033 to R.S.]; N.Z.Z. was supported by the UW-Madison Stem Cell and Regenerative Medicine Center fellowship; American Heart Association [18PRE34080337]. Funding for open access charge: NIH [R01GM113033].
Conflict of interest statement. None declared.

REFERENCES

1. Fussner,E., Djuric,U., Strauss,M., Hotta,A., Perez-Iratxeta,C., Lanner,F., Dilworth,F.J., Ellis,J. and Bazett-Jones,D.P. (2011) Constitutive heterochromatin reorganization during somatic cell reprogramming. *EMBO J.*, **30**, 1778–1789.

2. Ahmed,K., Dehghani,H., Rugg-Gunn,P., Fussner,E., Rossant,J. and Bazett-Jones,D.P. (2010) Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One*, **5**, e10531.
3. Meshorer,E., Yellajoshula,D., George,E., Scambler,P.J., Brown,D.T. and Misteli,T. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell*, **10**, 105–116.
4. Mattout,A., Biran,A. and Meshorer,E. (2011) Global epigenetic changes during somatic cell reprogramming to iPS cells. *J. Mol. Cell Biol.*, **3**, 341–350.
5. Sridharan,R., Gonzales-Cope,M., Chronis,C., Bonora,G., McKee,R., Huang,C., Patel,S., Lopez,D., Mishra,N., Pellegrini,M. *et al.* (2013) Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. *Nat. Cell Biol.*, **15**, 872–882.
6. James,T.C. and Elgin,S.C. (1986) Identification of a nonhistone chromosomal protein associated with heterochromatin in Drosophila melanogaster and its gene. *Mol. Cell. Biol.*, **6**, 3862–3872.
7. Eissenberg,J.C., James,T.C., Foster-Hartnett,D.M., Hartnett,T., Ngan,V. and Elgin,S.C. (1990) Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 9923–9927.
8. Nielsen,A.L., Oulad-Abdelghani,M., Ortiz,J.A., Remboutsika,E., Chambon,P. and Losson,R. (2001) Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol. Cell*, **7**, 729–739.
9. Lachner,M., O'Carroll,D., Rea,S., Mechtler,K. and Jenuwein,T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1proteins. *Nature*, **410**, 116–120.
10. Canzio,D., Larson,A. and Narlikar,G.J. (2014) Mechanisms of functional promiscuity by HP1 proteins. *Trends Cell Biol.*, **24**, 377–386.
11. Brown,J.P., Bullwinkel,J., Baron-Lühr,B., Billur,M., Schneider,P., Winking,H. and Singh,P.B. (2010) HP1gamma function is required for male germ cell survival and spermatogenesis. *Epigenet. Chromatin*, **3**, 9.
12. Takada,Y., Naruse,C., Costa,Y., Shirakawa,T., Tachibana,M., Sharif,J., Kezuka-Shiotani,F., Kakiuchi,D., Masumoto,H., Shinkai,Y.I. *et al.* (2011) HP1γ links histone methylation marks to meiotic synapsis in mice. *Development*, **138**, 4207–4217.
13. Aucott,R., Bullwinkel,J., Yu,Y., Shi,W., Billur,M., Brown,J.P., Menzel,U., Kioussis,D., Wang,G., Reisert,I. *et al.* (2008) HP1b is required for development of the cerebral neocortex and neuromuscular junctions. *J. Cell Biol.*, **183**, 597–606.
14. Maksakova,I.A., Goyal,P., Bullwinkel,J., Brown,J.P., Bilenky,M., Mager,D.L., Singh,P.B. and Lorincz,M.C. (2011) H3K9me3-binding proteins are dispensable for SETDB1/H3K9me3-dependent retroviral silencing. *Epigenet. Chromatin*, **4**, 12.
15. Bosch-Presegué,L., Raurell-Vila,H., Thackray,J.K., González,J., Casal,C., Kane-Goldsmith,N., Vizoso,M., Brown,J.P., Gómez,A., Ausió,J. *et al.* (2017) Mammalian HP1 isoforms have specific roles in heterochromatin structure and organization. *Cell Rep.*, **21**, 2048–2057.
16. Smallwood,A., Hon,G.C., Jin,F., Henry,R.E., Espinosa,J.M. and Ren,B. (2012) CBX3 regulates efficient RNA processing genome-wide. *Genome Res.*, **22**, 1426–1436.
17. Kwon,S.H., Florens,L., Swanson,S.K., Washburn,M.P., Abmayr,S.M. and Workman,J.L. (2010) Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II. *Genes Dev.*, **24**, 2133–2145.
18. Vakoc,C.R., Mandat,S.A., Olenchock,B.A. and Blobel,G.A. (2005) Histone H3 Lysine 9 Methylation and HP1γ Are Associated with Transcription Elongation through Mammalian Chromatin. *Mol. Cell*, **19**, 381–391.
19. Lomberk,G., Bensi,D., Fernandez-Zapico,M.E. and Urrutia,R. (2006) Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat. Cell Biol.*, **8**, 407–415.
20. Saint-André,V., Batsché,E., Rachez,C. and Muchardt,C. (2011) Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. *Nat. Struct. Mol. Biol.*, **18**, 337–344.
21. Zaidan,N.Z., Walker,K.J., Brown,J.E., Schaffer,L.V., Scalf,M., Shortreed,M.R., Iyer,G., Smith,L.M. and Sridharan,R. (2018) Compartmentalization of HP1 Proteins in Pluripotency Acquisition and Maintenance. *Stem Cell Rep.*, **10**, 627–641.
22. Veloso,A., Kirkconnell,K.S., Magnuson,B., Biewen,B., Paulsen,M.T., Wilson,T.E. and Ljungman,M. (2014) Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res.*, **24**, 896–905.
23. Duffy,E.E., Canzio,D., Maniatis,T. and Simon,M.D. (2018) Solid phase chemistry to covalently and reversibly capture thiolated RNA. *Nucleic Acids Res.*, **46**, 6996–7005.
24. Carvalho,S., Raposo,A.C., Martins,F.B., Grosso,A.R., Sridhara,S.C., Rino,J., Carmo-Fonseca,M. and de Almeida,S.F. (2013) Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. *Nucleic Acids Res.*, **41**, 2881–2893.
25. Kizer,K.O., Phatnani,H.P., Shibata,Y., Hall,H., Greenleaf,A.L. and Strahl,B.D. (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol. Cell. Biol.*, **25**, 3305.
26. Li,J., Moazed,D. and Gygi,S.P. (2002) Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J. Biol. Chem.*, **277**, 49383–49388.
27. Beard,C., Hochedlinger,K., Plath,K., Wutz,A. and Jaenisch,R. (2006) Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis*, **44**, 23–28.
28. Sridharan,R., Tchieu,J., Mason,M.J., Yachechko,R., Kuoy,E., Horvath,S., Zhou,Q. and Plath,K. (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell*, **136**, 364–377.
29. Chronis,C., Fiziev,P., Papp,B., Butz,S., Bonora,G., Sabri,S., Ernst,J. and Plath,K. (2017) Cooperative binding of transcription factors orchestrates reprogramming. *Cell*, **168**, 442–459.
30. Weinberg,D.N., Papillon-Cavanagh,S., Chen,H., Yue,Y., Chen,X., Rajagopalan,K.N., Horth,C., McGuire,J.T., Xu,X., Nikbakht,H. *et al.* (2019) The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature*, **573**, 281–286.
31. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
32. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
33. Shin,H., Liu,T., Manrai,A.K. and Liu,X.S. (2009) CEAS: cis-regulatory element annotation system. *Bioinformatics*, **25**, 2605–2606.
34. Zhu,L.J., Gazin,C., Lawson,N.D., Pages,H., Lin,S.M., Lapointe,D.S. and Green,M.R. (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics*, **11**, 237.
35. Lerdrup,M., Johansen,J.V., Agrawal-Singh,S. and Hansen,K. (2016) An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat. Struct. Mol. Biol.*, **23**, 349–357.
36. Shen,L., Shao,N., Liu,X. and Nestler,E. (2014) ngs.plot: quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*, **15**, 284.
37. Orlando,D.A., Chen,M.W., Brown,V.E., Solanki,S., Choi,Y.J., Olson,E.R., Fritz,C.C., Bradner,J.E. and Guenther,M.G. (2014) Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.*, **9**, 1163–1170.
38. Ramirez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
39. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
40. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S. *et al.* (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15545.
41. Wille,C.K. and Sridharan,R. (2019) Depletion of H3K79 methylation specifically enhances reprogramming to pluripotency but not

transdifferentiation. bioRxiv doi: https://doi.org/10.1101/835413, 07 November 2019, preprint: not peer reviewed.
42. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
43. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.
44. Leng,N., Dawson,J.A., Thomson,J.A., Ruotti,V., Rissman,A.I., Smits,B.M.G., Haag,J.D., Gould,M.N., Stewart,R.M. and Kendziorski,C. (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics*, **29**, 1035–1043.
45. Wang,X. and Yang,P. (2008) In vitro differentiation of mouse embryonic stem (mES) cells using the hanging drop method. *J. Vis. Exp.*, **17**, 825.
46. Lehnertz,B., Ueda,Y., Derijck,A.A.H.A., Braunschweig,U., Perez-Burgos,L., Kubicek,S., Chen,T., Li,E., Jenuwein,T. and Peters,A.H.F.M. (2003) Suv39h-Mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, **13**, 1192–1200.
47. Martens,J.H.A., O'Sullivan,R.J., Braunschweig,U., Opravil,S., Radolf,M., Steinlein,P. and Jenuwein,T. (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.*, **24**, 800–812.
48. Peters,A.H., O'Carroll,D., Scherthan,H., Mechtler,K., Sauer,S., Schofer,C., Weipoltshammer,K., Pagani,M., Lachner,M., Kohlmaier,A. *et al.* (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*, **107**, 323–337.
49. Schotta,G., Lachner,M., Sarma,K., Ebert,A., Sengupta,R., Reuter,G., Reinberg,D. and Jenuwein,T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.*, **18**, 1251–1262.
50. Cermakova,K., Smith,E.A., Veverka,V. and Hodges,H.C. (2019) Dynamics of transcription-dependent H3K36me3 marking by the SETD2:IWS1:SPT6 ternary complex. bioRxiv doi: https://doi.org/10.1101/636084, 14 May 2019, preprint: not peer reviewed.
51. Saint-André,V., Batsché,E., Rachez,C. and Muchardt,C. (2011) Histone H3 lysine 9 trimethylation and HP1. *Nat. Methods*, **18**, 337–344.
52. Rahl,P.B., Lin,C.Y., Seila,A.C., Flynn,R.A., McCuine,S., Burge,C.B., Sharp,P.A. and Young,R.A. (2010) c-Myc regulates transcriptional pause release. *Cell*, **141**, 432–445.
53. Baluapuri,A., Hofstetter,J., Dudvarski Stankovic,N., Endres,T., Bhandare,P., Vos,S.M., Adhikari,B., Schwarz,J.D., Narain,A., Vogt,M. *et al.* (2019) MYC recruits SPT5 to RNA polymerase II to promote processive transcription elongation. *Mol. Cell*, **74**, 674–687.
54. Huang,C., Su,T., Xue,Y., Cheng,C., Lay,F.D., McKee,R.A., Li,M., Vashisht,A., Wohlschlegel,J., Novitch,B.G. *et al.* (2017) Cbx3 maintains lineage specificity during neural differentiation. *Genes Dev.*, **31**, 241–246.
55. Yearim,A., Gelfman,S., Shayevitch,R., Melcer,S., Glaich,O., Mallm,J.-P., Nissim-Rafinia,M., Cohen,A.-H.S., Rippe,K., Meshorer,E. *et al.* (2015) HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep.*, **10**, 1122–1134.
56. Ostapcuk,V., Mohn,F., Carl,S.H., Basters,A., Hess,D., Iesmantavicius,V., Lampersberger,L., Flemr,M., Pandey,A., Thoma,N.H. *et al.* (2018) Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature*, **557**, 739–743.
57. Pradeepa,M.M., Sutherland,H.G., Ule,J., Grimes,G.R. and Bickmore,W.A. (2012) Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLos Genet.*, **8**, e1002717.
58. Spies,N., Nielsen,C.B., Padgett,R.A. and Burge,C.B. (2009) Biased chromatin signatures around polyadenylation sites and exons. *Mol. Cell*, **36**, 245–254.
59. Bannister,A., Zegerman,P., Partridge,J., Miska,E., Thomas,J. and Allshire,R. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, **410**, 120–124.
60. Thiru,A., Nietlispach,D., Mott,H.R., Okuwaki,M., Lyon,D., Nielsen,P.R., Hirshberg,M., Verreault,A., Murzina,N.V. and Laue,E.D. (2004) Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J.*, **23**, 489–499.
61. Smothers,J.F. and Henikoff,S. (2001) The hinge and chromo shadow domain impart distinct targeting of HP1-Like proteins. *Mol. Cell. Biol.*, **21**, 2555–2569.
62. Muchardt,C., Yaniv,M., Guilleme,M., Seeler,J.-S., Trouche,D. and Dejean,A. (2002) Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1. *EMBO Rep.*, **3**, 975–981.
63. Mishima,Y., Watanabe,M., Kawakami,T., Jayasinghe,C.D., Otani,J., Kikugawa,Y., Shirakawa,M., Kimura,H., Nishimura,O., Aimoto,S. *et al.* (2013) Hinge and chromoshadow of HP1α participate in recognition of K9 methylated histone H3in nucleosomes. *J. Mol. Biol.*, **425**, 54–70.
64. Brasher,S.V., Smith,B.O., Fogh,R.H., Nietlispach,D., Thiru,A., Nielsen,P.R., Broadhurst,R.W., Ball,L.J., Murzina,N.V. and Laue,E.D. (2000) The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *EMBO J.*, **19**, 1587–1597.
65. Li,Y., Trojer,P., Xu,C.-F., Cheung,P., Kuo,A., Drury,W.J. III, Qiao,Q., Neubert,T.A., Xu,R.-M., Gozani,O. *et al.* (2009) The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J. Biol. Chem.*, **284**, 34283–34295.
66. Lucio-Eterovic,A.K., Singh,M.M., Gardner,J.E., Veerappan,C.S., Rice,J.C. and Carpenter,P.B. (2010) Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 16952.
67. Qiao,Q., Li,Y., Chen,Z., Wang,M., Reinberg,D. and Xu,R.-M. (2011) The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *J. Biol. Chem.*, **286**, 8361–8368.
68. Skourti-Stathaki,K., Kamieniarz-Gdula,K. and Proudfoot,N.J. (2014) R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature*, **516**, 436–439.
69. Li,M., Phatnani,H.P., Guan,Z., Sage,H., Greenleaf,A.L. and Zhou,P. (2005) Solution structure of the Set2–Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 17636.
70. Lin,C.-H., Li,B., Swanson,S., Zhang,Y., Florens,L., Washburn,M.P., Abmayr,S.M. and Workman,J.L. (2008) Heterochromatin protein 1a stimulates histone H3 lysine 36 demethylation by the Drosophila KDM4A demethylase. *Mol. Cell*, **32**, 696–706.
71. Morikawa,K., Ikeda,N., Hisatome,I. and Shirayoshi,Y. (2013) Biochemical and biophysical research communications. *Biochem. Biophys. Res. Commun.*, **431**, 225–231.