Genome-Wide Studies Reveal that H3K4me3 Modification in Bivalent Genes Is Dynamically Regulated during the Pluripotent Cell Cycle and Stabilized upon Differentiation

Rodrigo A. Grandy, a Troy W. Whitfield, b Hai Wu, a Mark P. Fitzgerald, a Jennifer J. VanOudenhove, a Sayed K. Zaidi, a Martin A. Montecino, c Jane B. Lian, a André J. van Wijnen, a Janet L. Stein, a Gary S. Stein a

Department of Biochemistry and University of Vermont Cancer Center, University of Vermont College of Medicine, Burlington, Vermont, USA; aDepartment of Cell and Developmental Biology and Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, Massachusetts, USA; aCenter for Biomedical Research and FONDAP Center for Genome Regulation, Universidad Andres Bello, Santiago, Chile; bDepartments of Orthopedic Surgery and Biochemistry and Molecular Biology, Mayo Clinic, Rochester, Minnesota, USA.

Stem cell phenotypes are reflected by posttranslational histone modifications, and this chromatin-related memory must be mitotically inherited to maintain cell identity through proliferative expansion. In human embryonic stem cells (hESCs), bivalent genes with both activating (H3K4me3) and repressive (H3K27me3) histone modifications are essential to sustain pluripotency. Yet, the molecular mechanisms by which this epigenetic landscape is transferred to progeny cells remain to be established. By mapping genomic enrichment of H3K4me3/H3K27me3 in pure populations of hESCs in G2, mitotic, and G1 phases of the cell cycle, we found striking variations in the levels of H3K4me3 through the G2-M-G1 transition. Analysis of a representative set of bivalent genes revealed that chromatin modifiers involved in H3K4 methylation/demethylation are recruited to bivalent gene promoters in a cell cycle-dependent fashion. Interestingly, bivalent genes enriched with H3K4me3 exclusively during mitosis undergo the strongest upregulation after induction of differentiation. Furthermore, the histone modification signature of genes that remain bivalent in differentiated cells resolves into a cell cycle-independent pattern after lineage commitment. These results establish a new dimension of chromatin regulation important in the maintenance of pluripotency.

H

uman embryonic stem cells (hESCs) are an increasingly powerful tool for regenerative medicine. They recapitulate, in vitro, the molecular phenomena that take place during the first stages of embryonic development. Like their in vivo counterparts, ESCs proliferate rapidly and are able to form the three embryonic germ layers (1). This highly self-renewing and pluripotent state is sustained by a unique epigenetic landscape, consisting of transcription factors, chromatin remodeling complexes, and histone modifications that provide the transcriptional plasticity required for rapid response to differentiation cues (2).

Histone H3 lysine 4 and 27 trimethylations (H3K4me3 and H3K27me3, respectively) are key histone modifications that are involved in transcriptional regulation (3, 4). H3K4me3 near transcriptional start sites (TSSs) marks regions of active transcription or transcriptional readiness (5). H3K27me3 modification, in contrast, is a well-established negative regulator of gene expression that repels transcriptional activators and attracts chromatin repressors that promote chromatin compaction (6). Genomic regions that host both histone marks, so-called bivalent domains, were first observed in ESCs, primarily near promoters of genes with developmental functions (7–9). Significant effort has gone into understanding the biological role of bivalency; the consensus is that, in ESCs, it represses transcription but poises genes for rapid expression during lineage commitment (10). Although this proposition is not yet supported with direct evidence, it has become clear that bivalent domains are essential for maintaining ESC pluripotency and self-renewing capacity (10). Despite the extensive availability of genome-wide maps of these histone marks in pluripotent and committed cells, it is not understood how they contribute to faithful reestablishment of transcriptional status after cell division. Compelling questions remain, including the detailed localization of H3K4me3 and H3K27me3 during mitosis, whether these histone marks are gained/lost exclusively during mitosis, and perhaps more importantly, whether they constitute bivalent domains that are retained after cells exit mitosis.

Here, we show that dynamic cell cycle control of H3K4 methylation/demethylation of bivalent genes represents a new dimension to chromatin regulation that advances understanding of how the pluripotent histone modification landscape contributes to maintenance of hESC identity. We developed a new method for isolating pure populations of hESCs at the G2, mitosis (M), and G1 phases of the cell cycle and used these phase-specific populations to map the genome-wide distribution of bivalent domains (H3K4me3/H3K27me3) throughout the pluripotent cell cycle. Consistent with a pivotal developmental function, we demonstrate that bivalent genes enriched with H3K4me3 during mitosis are maximally upregulated following induction of hESC differentiation, and subsequently, H3K4me3 on these genes becomes cell

Received 16 September 2015 Returned for modification 20 October 2015 Accepted 30 November 2015

Accepted manuscript posted online 7 December 2015


Address correspondence to Gary S. Stein, gary.stein@uvm.edu.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/MCB.00877-15.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.
cycle independent. Finally, we show that chromatin modifiers involved in H3K4 methylation/demethylation are recruited to bivalent gene promoters in a cell cycle-dependent fashion.

**MATERIALS AND METHODS**

**hESC culture and differentiation.** The H9 hESC line from WiCell Research Institute (Madison, WI) was maintained on hESC-qualified Matrigel (BD Bioscience; catalog no. 354277) in mTeSR-1 medium (Stemcell Technologies; catalog no. 05850) or essential E8 medium (Life Technologies; catalog no. A1517001), as recommended by the supplier. Cells were expanded every 5 to 6 days, using nonenzymatic passaging according to WiCell Research Institute standard protocols. To generate PAX6 cells, undifferentiated ESCs were incubated in mTeSR-1 medium supplemented with 10 nM retinoic acid (RA) (Sigma-Aldrich; catalog no. R2625-50MG) for 5 days. The treatment started 1 day after plating of the cells, and medium was changed every day. hESC research was approved by the Institutional Embryonic Stem Cell Research Oversight Committee at the University of Vermont.

**Cell sorting.** Pure populations of cells at the G2, mitosis, or G1 phase of the cell cycle were isolated by fluorescence-activated cell sorting (FACS), taking advantage of differences in DNA content to distinguish cells in G2/M from cells in G1 and the exclusive presence of histone H3 serine 28 phosphorylation (H3S28p) in mitosis to discriminate cells in G2 from those in M phase (Fig. 1A and C). As indicated in the figure legends, both nocodazole-synchronized and untreated cells were sorted using the procedure described here. After fixation, cells were permeabilized for 10 min using a mild permeabilization/wash buffer containing saponin (BD Bioscience; catalog no. 51-2091KZ). For ESC isolation, cells were incubated with labeled antibodies to OCT4 (phycoerythrin [PE] conjugated; BD Bioscience; catalog no. 561556) and H3S28p (Alexa Fluor 647 conjugated; BD Bioscience; catalog no. 558609) for 30 min. Labeled antibodies to PAX6, instead of OCT4 (PE conjugated; BD Bioscience; catalog no. 561552) were used to stain PAX6 cells. After staining, cells were washed with permeabilization/wash buffer, resuspended in 2% fetal bovine serum (FBS) in phosphate-buffered saline (PBS), and counterstained with 1 g/ml 4',6-diamidino-2-phenylindole (DAPI) (Life Technologies; catalog no. D1306) for 30 min. FACS was performed on a BD FACSAria II cell sorter using linear forward scatter (FSC) and side scatter (SSC) scaling, followed by height- and area-based doublet discrimination. Compensation was calculated using FACS Diva autocompensation algorithms and supplemented by manual compensation to correct for autofluorescence. Cells with subnormal DNA content were filtered out during the gating, and only samples with a purity of 95% or higher were used in the following experiments.

**Synchronization of hESCs.** Before every cell sorting experiment, cells were allowed to establish robust colonies. Normally, they were collected 4 to 5 days after plating. To improve the cell yield from sorting, in some experiments we synchronized hESCs at either the G2/M or G1 phase of the cell cycle 1 day before the collection, by incubation with nocodazole (Sigma-Aldrich; catalog no. M1404-2MG) (200 ng/ml for 16 h when cells were grown in mTeSR-1). Blocking of hESCs in G2/M phase was reversible, and cells resumed the cell cycle and progressed into G1 after nocodazole was withdrawn from the culture medium. For cells growing in E8 medium, 25 ng/ml of nocodazole for 16 h was enough to block most of the cells in G2/M and allowed progression into G1 after nocodazole was withdrawn from the medium (data not shown). After synchronization, cells were collected as single cells by treatment with Acutase (MP Biomedicals; catalog no. 1000449) and subsequently fixed with 1% formaldehyde for 10
min followed by 5 min of incubation with 0.125 M glycine (Sigma-Aldrich; catalog no. G8790-100G). Cells were counted and stored at −80°C.

**Immunofluorescence microscopy.** H9 cells were grown on coverslips previously coated with hESC-qualified Matrigel (BD Biosciences; catalog no. 354277). Normally, 3 days after plating, cells were fixed with 1% formaldehyde in PBS for 10 min and processed for immunofluorescence. For PAX6 cells, the fixation was performed at the end of the differentiation protocol. Subsequently, cells were permeabilized in 0.25% Triton X-100 in PBS and blocked with 0.5% bovine serum albumin (BSA) in PBS. Cells were stained with the following antibodies for 1 h at room temperature: anti-OCT4 (Abcam; catalog no. ab1857 [1:2,000]; Santa Cruz Biotechnology; catalog no. sc8628 [1:1,000]), anti-H3S28p (Millipore; catalog no. 07-145 [1:5,000]), or anti-PAX6 (Santa Cruz Biotechnology; catalog no. sc81649 [1:1,000]). Nuclear DNA was counterstained with DAPI (0.1 μg/ml). Secondary antibodies conjugated with Alexa Fluor 594 or Alexa Fluor 488 (1:1,000; Molecular Probes/Invitrogen) were incubated for 1 h in a humidity chamber and at room temperature. Finally, cells were mounted in Prolong-Gold (Invitrogen).

**Immunofluorescence analysis for cells in suspension.** After cell sorting experiments, the purity of the sorted samples was evaluated by immunofluorescence analysis. Briefly, a small aliquot of sorted cells was spun down on coverslips by using Cytospin4 (Thermo Shandon). Subsequently, cells were stained using the standard immunofluorescence protocol. Anti-Ki67 antibody was used in these experiments (Santa Cruz Biotechnology; catalog no. sc23900).

**Immunoblotting.** Whole-cell extracts from cell cycle-specific or unsynchronized cells were used to analyze protein expression by Western blotting. The following antibodies were used: anti-H3 (Abcam; catalog no. ab1791 [1:5,000]), anti-H3S10p (Upstate; catalog no. 05-598 [1:5,000]), anti-CDT1 (Abcam; catalog no. ab70829 [1:2,000]), antitubulin (Sigma; catalog no. T9026 [1:30,000]), anti-EZH2 (Cell Signaling; catalog no. AC22 [1:1,000]), anti-MLL-C (antiserum [1:1,000]), anti-KDM1 (Abcam; catalog no. ab17721 [1:1,000]), anti-KDM5A (Abcam; catalog no. ab70892 [1:1,000]), and anti-KDM5B (Abcam; catalog no. ab56759 [1:5,000]).

**Chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR) assays.** Cells at specific phases of the cell cycle, obtained from 3 to 4 different cell sorting experiments, were pooled and treated as a single sample (or biological replicate). Cells were washed twice with PBS and subjected to extraction of nuclei according to the method in reference 11 with modifications. Isolated nuclei were sonicated using a Misonix S-4000 ultrasonic processor (Qsonica) to obtain sheared chromatin ranging from 0.2 to 0.6 kb in size (see Fig. S2A in the supplemental material). Sheared chromatin (4 μg) was immunoprecipitated with either anti-H3K4me3 (Abcam; catalog no. ab1012) or anti-H3K27me3 (Millipore; catalog no. 07-449). Immunoprecipitated complexes were purified using protein G Dynabeads (Invitrogen), eluted, reverse cross-linked, and quantified, and subjected to library preparation. Immunoprecipitated DNA or input DNA for each cell cycle phase, cellular condition (ESC versus PAX6 cells), and biological replicate was end-repaired using the End-It DNA end repair kit (Epicon), extended using a Klenow fragment (3′-5′ exo) (Epicon), and ligated to sequencing adaptors (Illumina), according to the manufacturers’ recommendations. Each library was then subjected to 15 cycles of PCR amplification using PFU Ultra II Hotstart master mix (Agilent) and size selected in the range of 300 ± 50 bp. The final libraries were quantified using both Qubit fluorimeter (Life Technologies) and Bioanalyzer (Agilent Technologies) systems and then submitted for sequencing.

**ChIP-seq analysis.** Single-end 36-base sequencing was performed on an Illumina Genome Analyzer II. Base calls and sequence reads were generated using Illumina CASAVA software (version 1.6; Illumina). Reads were mapped to the human genome (GRCh37, hg19) using bowtie (version 0.12.8) (12) with up to two mismatches allowed. For each cellular condition (i.e., defined by cell cycle phase and differentiation state), genome-wide enrichment profiles were generated from pooled ChIP-seq replicates versus input for H3K27me3 and H3K4me3, respectively, using SPP (13). The resulting enrichment profiles are conservative statistical estimates of log fold enrichment across the genome, corresponding to a 99% confidence interval, α = 0.01. The UCSC known gene annotations (14) for canonical transcripts were used to analyze enrichments around transcriptional start sites (TSSs). In order to restrict our analysis to genes enriched for both H3K27me3 and H3K4me3, so-called bivalent genes, we required that for both histone Modifications during any pair of cell cycle phases, min[log2(ε′1), log2(ε′2), · · · , log2(ε′l)] > 1, where ε′i is the enrichment at the ith position around the TSS for gene j and N is chosen such that the window width is 500 bp. That is, bivalent genes were defined to show a 2-fold or greater enrichment everywhere in at least one 500-bp window of the spanned TSS neighborhood (within 2 kb upstream or downstream of the TSS). This restriction resulted in 5,240 bivalent genes. These bivalent genes were examined by direct inspection of the relative H3K4me3 enrichments between different phases of the cell cycle and also using k-means clustering. For the relative H3K4me3 enrichments, the mean ratio of M-phase to G1-phase or M-phase to G2-phase enrichments for gene j was calculated as, for example, [log2(ε′M/G2)] = [log2(M)] − [log2(G2)], where [log2(M)] = \[\frac{N\log^2_{N^2} \log^2_{N^2} \log^2_{N^2}}{N}\] and N was chosen to bracket the TSS by 1 kb upstream and downstream. Functional enrichment analysis for the genes grouped by ChIP-seq enrichment patterns used DAVID v. 6.7 (15, 16).

**RNA-seq sample preparation.** Total RNAs from undifferentiated H9 cells or PAX6 cells were isolated by using TRIzol (Invitrogen). Subsequently, we removed DNA contaminants using a DNA-free RNA extrac-
tion kit (Zymo Research). According to Illumina’s recommended protocol, 2 μg of DNA-free total RNA was used to construct each pair-end transcriptome sequencing (RNA-seq) library with the TruSeq RNA sample preparation kit v2 (Illumina). Libraries were subjected to 15 cycles of amplification with pair-end PCR primers (Illumina). Libraries were quantified using a Bioanalyzer system and sequenced on the HiSeq-1000 platform (Illumina) for the read length of 100 bases. To address the influence of culture medium on RNA expression, RNA-seq libraries were built from stocks of ESCs maintained in either mTeSR-1 or essential E8 medium. The RNA-seq libraries for PAX6 cells were generated from original cellular stocks maintained in either mTeSR-1 or E8 medium. No major differences in RNA expression were detected among different stocks. Therefore, libraries of similar cellular condition (control or differentiated) were considered biological replicates.

RNA-seq analysis. Raw sequences from RNA-seq libraries were mapped to genome assembly hg19 using TopHat (version 2.0.8b) (17) with parameters “-b2-sensitive -r 80 -g 10.” Mapped reads were analyzed using SeqMonk software (v0.27.0; Babraham Bioinformatics Institute). Expression profiles were calculated using the RNA-seq quantitation pipeline and normalized by globally adjusting count distributions at the 75th percentile. Differential expression was calculated using the intensity difference statistical test included in the software. Differential expression was called on log2-transformed counts by selecting transcripts that changed with a significance ($P$) of $<0.05$ after Benjamini and Hochberg correction.

Expression profiles were calculated using the RNA-seq quantitation pipeline and normalized by globally adjusting count distributions at the 75th percentile. Differential expression was calculated using the intensity difference statistical test included in the software. Differential expression was called on log2-transformed counts by selecting transcripts that changed with a significance ($P$) of $<0.05$ after Benjamini and Hochberg correction.

RESULTS

Genome-wide mapping of H3K4me3 and H3K27me3 across the cell cycle in hESCs reveals that the bivalent histone modification landscape is highly dynamic. To gain insights into the mechanisms that control the mitotic inheritance of bivalency in pluripotent cells, we globally mapped the genomic locations of the H3K4me3 and H3K27me3 modifications in pure populations of hESCs at the G2, M, or G1 phase of the cell cycle, by chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-seq). To this end, OCT4+ hESCs at G2, M, or G1 were isolated by fluorescence-activated cell sorting (FACS), based on both DNA content and H3S28p (histone H3 phosphorylated at serine 28), a mitosis-specific epigenetic mark (18) (Fig. 1A to C). We enriched cultures for mitotic cells by treating them with nocodazole (200 ng/ml, 16 h), a widely used synchronizing agent that has been successfully employed to synchronize ESCs (19–22). Cells in G2 were purified 1 h after release of the nocodazole block.

The sorted cells exhibited the expected cell cycle-dependent morphological and molecular characteristics (see Fig. S1A in the supplemental material). M-phase cells displayed condensed chromosomes and relatively dispersed OCT4 distribution. G1 cells were the smallest and exhibited the typical micropunctate pattern for the Ki67 protein (19). G2 cells were larger than G1 cells and displayed one or two large Ki67 foci. Biochemical analysis (Fig. 1C) confirmed that H3S28p was exclusive to mitosis (18) and that the licensing factor CDT1 reaches its minimum protein levels in G2 and its maximum during mitosis (23). Importantly, the protein levels of the pluripotency markers OCT4 and SSEA4 were not affected by the nocodazole treatment (see Fig. S1B). These analyses demonstrate that our strategy yields highly pure populations of hESCs at the G2, mitosis, or G1 phase of the cell cycle that are suitable for genome-wide analysis.

Bivalent domains are typically located at gene promoters, and the vast majority of the H3K27me3 modifications observed in ESCs are found in these regions (7, 8). Therefore, we used the more than 15 million uniquely mapped reads generated for each histone mark and cellular condition, to calculate conservative genome-wide ChIP-seq enrichments within regions centered at the canonical TSS (kb ±2) of the UCSC-known genes (Spearman correlations between biological replicates, 0.84 to 0.98 [see Fig. S2E and F in the supplemental material]). We found that about one-third (5,340) of the 17,002 genes modified by H3K4me3 were also enriched in developmental functions and, on average, disodified by H3K27me3. Consistent with previous reports, this group was enriched in genes with developmental functions (Fig. 1E) (7–9). Importantly, the majority of these genes (~3,000) have been previously classified as bivalent in hESCs (24) (see Fig. S2C). The additional 2,154 genes classified as bivalent in our study are also enriched in developmental functions and, on average, displayed H3K4me3/H3K27me3 enrichment comparable to that observed in “established” bivalent genes (see Fig. S2B to D). In general, the enrichment profiles of bivalent genes at G2, M, and G1 phases of the cell cycle displayed the narrow peak-like distribution typical for H3K4me3 and the broader domain-like configuration associated with H3K27me3 (Fig. 1D) (8, 9).
Interestingly, H3K4me3 levels showed striking variations throughout the cell cycle, reaching maximum enrichment during mitosis and minimum enrichment in G1 (Fig. 1D): comparing these enrichments, normalized using an MA-norm-like procedure (see below) and averaged over a region within 2 kb upstream or downstream of each TSS, across these three cell cycle phases using Tukey’s range test showed that the means differed with a P value of $10^{-5}$. This result correlates well with immunofluorescence analysis showing that there are foci with enhanced H3K4me3 signal during mitosis, compared with cells in interphase (see Fig. S2G in the supplemental material).

In contrast, H3K27me3 levels remained relatively constant across the G2-M-G1 transition (Fig. 1D): comparing these enrichments, normalized using an MA-norm-like procedure (see below) and averaged over a region within 2 kb upstream or downstream of each TSS, across these three cell cycle phases using Tukey’s range test showed that the means differed relatively little, with a $P$ value of $0.09$. Consistently, immunofluorescence assays showed no differences in H3K27me3 levels between mitotic and interphase cells (see Fig. S2H in the supplemental material). Overall, these results indicate that H3K27me3 is stable across the cell cycle, suggesting that this histone mark differs from H3K4me3 not only in its relationship with transcription but also in its regulation throughout the cell cycle.

Cell cycle-dependent changes in H3K4me3 modification define new classes of bivalent genes. Motivated by repeated observations of genes enriched with H3K4me3 during mitosis (mitosis enriched [ME]), genes that exhibited constant levels of H3K4me3 in G2 and M followed by a pronounced drop during G1 (G1 depleted [G1-dep]), and genes that maintained relatively stable H3K4me3 levels across the cell cycle (ubiquitous [Ub]) (Fig. 2A; see also Fig. S3C in the supplemental material), we classified the bivalent genes into these three categories by defining threshold values for the relevant ratios of mean enrichments (see Materials and Methods and Fig. S3B for details). In this straightforward way, we identified 690 ME genes with levels of H3K4me3 modification at least 2-fold higher during mitosis than in G2 or G1, including 28 genes that were manually reassigned based on their histone modification profiles (see Table S1). We found 2,665 G1-dep genes with similar levels of H3K4me3 in mitosis and G2 and at least 2-fold-lower levels during G1 and 1,886 Ub genes with relatively constant levels of H3K4me3 across the cell cycle (Fig. 2B and C; see also Fig. S3C). As an alternative means of classifying cell cycle-dependent H3K4me3 enrichments, $k$-means clustering ($k=3$) was applied to the enrichment profiles, after normalizing the ChIP-seq data by either read counts (25) or an MA-norm-like scaling in which linear scaling is used to ensure that the average ChIP-seq enrichment across different cellular conditions is equal over the set of common peaks (26) (see below and Fig. 4D). It is important to note the consistency with which these three cell cycle-dependent signatures were observed, indicating that these histone marks are controlled in a cell cycle-dependent manner.

Cell cycle-dependent changes in H3K4me3 define new classes of bivalent genes. Motivated by repeated observations of genes enriched with H3K4me3 during mitosis (mitosis enriched [ME]), genes that exhibited constant levels of H3K4me3 in G2 and M followed by a pronounced drop during G1 (G1 depleted [G1-dep]), and genes that maintained relatively stable H3K4me3 levels across the cell cycle (ubiquitous [Ub]) (Fig. 2A; see also Fig. S3C in the supplemental material), we classified the bivalent genes into these three categories by defining threshold values for the relevant ratios of mean enrichments (see Materials and Methods and Fig. S3B for details). In this straightforward way, we identified 690 ME genes with levels of H3K4me3 modification at least 2-fold higher during mitosis than in G2 or G1, including 28 genes that were manually reassigned based on their histone modification profiles (see Table S1). We found 2,665 G1-dep genes with similar levels of H3K4me3 in mitosis and G2 and at least 2-fold-lower levels during G1 and 1,886 Ub genes with relatively constant levels of H3K4me3 across the cell cycle (Fig. 2B and C; see also Fig. S3C). As an alternative means of classifying cell cycle-dependent H3K4me3 enrichments, $k$-means clustering ($k=3$) was applied to the enrichment profiles, after normalizing the ChIP-seq data by either read counts (25) or an MA-norm-like scaling in which linear scaling is used to ensure that the average ChIP-seq enrichment across different cellular conditions is equal over the set of common peaks (26) (see below and Fig. 4D). It is important to note the consistency with which these three cell cycle-dependent signatures were observed, indicating that these histone marks are controlled in a cell cycle-dependent manner.

FIG 2 Cell cycle-dependent changes in H3K4me3 define new classes of bivalent genes. (A) ChIP-seq enrichment at select bivalent genes from each of the three cell cycle-dependent bivalency profiles. Shaded areas indicate bivalent domains. ChIP-seq enrichments/input in a log2 scale are displayed. (B) Heat map of H3K27me3 and H3K4me3 profiles over 4,000-bp regions spanning the TSSs of UCSC canonical genes, grouped according to the divisions illustrated in Fig. 3B in the supplemental material: mitosis enriched (top), G1 depleted (middle), and ubiquitous (bottom). The total number of genes clustered into each group is displayed on the right. (C) Aggregation plots of H3K4me3 and H3K27me3 ChIP-seq enrichments for bivalent genes clustered according to panel B. (D) GO term analyses of the three groups of bivalent genes clustered by H3K4me3 enrichment profiling. Tables display randomly selected genes from each cluster. (E) ChIP-qPCR analysis of H3K4me3 in hESCs sorted at specific phases of the cell cycle without prior chemical synchronization. Results ($n=3$) were normalized by total histone H3 levels and expressed as percent input. Ordinary one-way analysis of variance followed by Tukey’s test for multiple comparisons was performed. *, $P<10^{-2}$; ***, $P<10^{-3}$; ****, $P<10^{-5}$; ns, not significant.
tone modification patterns are robust; the same pattern emerges regardless of the data processing method. Throughout the remainder of this paper, the gene classification from Fig. S3B is used.

A more in-depth examination revealed that the ME category includes bivalent genes with the highest levels of H3K27me3 modification (Fig. 2C; see also Fig. S4B in the supplemental material), and gene ontology (GO) analysis revealed that developmentally relevant transcription factor genes are overrepresented in this group (Fig. 2D). Of note, many of these transcription factor genes were marked with H3K4me3 only during mitosis, suggesting stringent cell cycle control of this histone mark (Fig. 2A, D, and E). The Ub and G1-dep categories displayed comparable average levels of H3K27me3 across the cell cycle (Fig. 2C; see also Fig. S4B), and both include many genes involved in membrane biology and cell signaling; however, the G1-depleted group uniquely includes genes implicated in phosphorylation processes (Fig. 2D). Overall, these observations suggest a functional separation of genes associated with these three bivalency profiles. Interestingly, we observed genomic regions in which two different cell cycle-dependent bivalency patterns are separated by only a few kilobases (Fig. 2A; see also Fig. S3A). This finding further supports the notion that bivalency is strictly regulated and molecularly compartmentalized during the cell cycle in hESCs.

To validate these results, we carried out ChIP-qPCR analysis from G2-, M-, and G1-phase cells isolated from asynchronous hESCs (not nododazole treated) using FACS. Because only limited numbers of mitotic cells are present in asynchronous cultures and H3K27me3 does not vary significantly across the cell cycle, we evaluated H3K4me3 levels for selected genes representative of each of the three bivalency categories. Total levels of histone H3 were assessed to control for possible local variations in nuclear somal density. These experiments corroborated that, in asynchronous cultures, the ME TBX1 and TLX2 genes are modified with H3K4me3 exclusively during mitosis, and H3K4me3 levels in the Ub bivalent TBC1D20 and PPP1R1A1 genes do not vary significantly across the cell cycle. Moreover, genes representative of the G1-dep category, the COL1A1 and FOXJ1 genes, show somewhat reduced levels of H3K4me3 in G1 (Fig. 2A and E; see also Fig. S3A in the supplemental material).

Taken together, these results reveal unexpected cell cycle-dependent variations in the histone modification landscape that expose previously unknown categories of bivalent genes.

**Bivalent genes enriched with H3K4me3 during mitosis experience the strongest upregulation at the onset of hESC differentiation.** To understand how the dynamic regulation of H3K4me3 during the pluripotent cell cycle influences transcription of bivalent genes during lineage commitment, we used high-throughput RNA sequencing (RNA-seq). RNA expression levels were analyzed in hESCs before or after induction of retinoic acid (RA) differentiation. As in previous reports (7–9), we found that in undifferentiated hESCs, bivalent genes are either repressed or expressed at very low levels (Fig. 3A). Interestingly, genes in the ME class were the most repressed, whereas those in the G1-dep group were the least repressed (Fig. 3A). Consistently, genes in the ME class displayed the highest levels of the repressive H3K27me3 mark, and G1-dep genes showed the highest average levels of H3K4me3 (Fig. 2C). As expected, after inducing differentiation with RA (10 μM, 5 days), we observed downregulation of the pluripotency marker OCT4 and upregulation of the early ectodermal marker PAX6 (Fig. 4A and B) (8). Accordingly, the epigenetic status of the OCT4 promoter changed to a repressed form, with no H3K4me3 and high H3K27me3 (see Fig. S4D in the supplemental material). Likewise, the PAX6 promoter switched from a bivalent, ME state to an active configuration rich in H3K4me3, with very little H3K27me3 (see Fig. S4D). Compared with undifferentiated hESCs, the PAX6-expressing cells (PAX6 cells) displayed significant differential expression of 546 genes (P < 0.05) (Fig. 3B). Among these, 138 upregulated genes were bivalent (Fig. 3D). The ME category had the highest proportion of upregulated bivalent genes after RA induction. More importantly, most of these upregulated genes encode transcription factors with developmental functions (Fig. 3D), and they displayed the largest changes in RNA levels in response to RA treatment (Fig. 3C). In contrast, although many genes in the G1-dep and Ub categories were significantly upregulated, changes in their expression were smaller (Pearson’s correlation coefficient R = 0.661 for ME genes, R = 0.876 for G1-dep genes, and R = 0.825 for Ub genes) (Fig. 3E). These changes are consistent with the H3K4me3 and H3K27me3 ChIP-seq profiles seen in bivalent genes before and after induction of hESC differentiation (see Fig. S4C).

Taken together, these results indicate that bivalent genes which accumulate H3K4me3 modifications during mitosis in pluripotent cells experience more pronounced upregulation at the onset of differentiation. This suggests that these genes are controlled by a distinct mechanism of mitotic chromatin regulation that allows transcriptional activation in response to differentiation cues.

**The pluripotent H3K4me3 landscape becomes cell cycle independent upon cell fate commitment.** To address whether the cell cycle-dependent, histone modification patterns observed in hESCs are retained in nonpluripotent cells, we generated ChIP-seq libraries for H3K4me3 and H3K27me3 histone modifications from RA-differentiated hESCs (PAX6 cells) synchronized and isolated by FACS at the G2, M, or G1 phase of the cell cycle (see Fig. S4A in the supplemental material). Comparison of the H3K4me3/H3K27me3 ChIP-seq profiles obtained in hESCs and PAX6 cells revealed that, although numerous genes lost bivalency after induction of differentiation, many others remained bivalent (Fig. 4C and E; see also Fig. S4C and D). This result was consistent with upregulation of only a fraction of bivalent genes after induction of differentiation (Fig. 3B) and also with reports that the gain or loss of bivalency depends on whether or not genes are activated during a specific lineage-determination process (27, 28).

Importantly, as in ESCs, the average level of H3K27me3 in PAX6 cells remained relatively constant through the G2-M-G1 transition, which suggests that the cell cycle-related regulation of H3K27me3 is similar in hESCs before and after the onset of differentiation (Fig. 4C and E; see also Fig. S4B in the supplemental material). In contrast, the H3K4me3 enrichments became cell cycle independent when cells began to differentiate—in constant levels of H3K4me3 were observed across the different categories of bivalent genes, during the G2-M-G1 transition in PAX6 cells (Fig. 4C to E). Of note, although our initial analysis showed higher levels of the H3K4me3 mark during G2, these enrichments were abrogated when different types of data normalization were applied (read number versus MA-norm-like scaling normalization) (Fig. 4D).

Together, these results indicate that regulation of H3K4me3 at genes that remain bivalent after the onset of differentiation (PAX6 cells) is cell cycle independent and, furthermore, that the cell cycle-dependent regulation of H3K4me3 modification within bivalent domains is unique to the pluripotent state. This finding is
consistent with a model in which active accumulation of H3K4me3 during mitosis prepares repressed bivalent genes for activation in response to differentiation cues at the start of the G1 phase. If those cues do not materialize, H3K4me3 is removed during G1.

Cell cycle-dependent variation of H3K4me3 correlates with sequential recruitment of KDM5A, MLL1/2, and WDR5. To understand the dynamic regulation of H3K4me3 levels in bivalent genes, we assessed the chromatin occupancy of proteins with H3K4me3 regulatory activity during the pluripotent cell cycle. First, because MLL has H3K4 methyltransferase activity and remains associated with chromatin during mitosis (20), we used ChIP-qPCR to evaluate its interaction with a representative set of bivalent gene promoters during the G2-M-G1 transition in unsynchronized hESCs (Fig. 5A). We detected significant MLL binding on promoters of ME genes only during mitosis. Promoters of G1-dep genes showed modest levels of MLL binding that remained constant during G2 and mitosis and decreased significantly in G1. Consistent with expectations, the MLL binding pattern for Ub gene promoters was constant during the G2-M-G1 transition. Furthermore, because it was recently shown that MLL2 is the main lysine methyltransferase responsible for H3K4me3 on bivalent genes in mouse ESCs (29, 30), we also assessed the chromatin occupancy of MLL2. Although the binding of MLL2 was modest among the genes evaluated here, the trends were similar to those of MLL cell cycle-dependent binding. Likewise, WDR5, a subunit of the MLL methyltransferase complex that is essential for MLL activity and critical for maintaining ESC self-renewal (3, 31), occupied ME and G1-dep gene promoters during G2 and mitosis only, becoming completely dissociated during G1. Interestingly, however, no significant cell cycle variation in WDR5 binding to Ub gene promoters was detected.

These results suggest that MLL/WDR5 proteins associate with bivalent gene promoters in a cell cycle-dependent manner that...
correlates with H3K4me3 enrichment in G2 and mitosis (Fig. 2). However, they do not explain the reduced H3K4me3 levels observed in G1 phase. We therefore evaluated KDM5A, a member of the H3K4 di-/tridemethylase family involved in repression of many developmental genes in ESCs (32). Using asynchronous hESC cultures, we found that levels of KDM5A were modestly but significantly enriched in G1 near ME and G1-dep gene promoters, which is consistent with the variations in H3K4me3 observed among these genes during the cell cycle (Fig. 5A). Strikingly, the chromatin occupancy of KDM5A on Ub category gene promoters was also significantly increased during G1 (Fig. 5A). This finding was unexpected because of the constant H3K4me3 levels observed among Ub gene promoters through the cell cycle. However, it is consistent with reports that H3K27me3 facilitates KDM5A recruitment to chromatin (4, 32) and could explain the slight increase in global H3K27me3 levels that we observed in hESCs after mitosis (Fig. 1D). This mechanism would reinforce repression of bivalent genes during G1 phase in the absence of differentiation signals and would be relevant when enzymes with H3K4 methyltransferase activity are not present.

Overall, these results indicate that the methylase/demethylase machinery involved in regulating H3K4me3 levels in hESCs is recruited differentially to bivalent genes in a cell cycle-dependent manner (Fig. 5B). Furthermore, they suggest that similar, related mechanisms may be responsible for the H3K4me3 profiles observed on all bivalent genes.

Disruption of MLL1/2 trimethylase activity abrogates H3K4me3 mitotic enrichment and delays expression of bivalent genes upon induction of cellular differentiation. To further understand the role of H3K4me3 enrichment in bivalent genes during mitosis in pluripotent cells, we assessed the biological consequences of disrupting the activity of the H3K4 trimethylases MLL1/2. Since the genetic ablation of components of the COMPASS complex, including MLL, DPY30, or DWR5 proteins, not only reduces H3K4me3 levels on target genes but also affects the self-renewal capabilities of pluripotent cells (20, 31, 33), we conducted our experiments using less disruptive conditions. We took advantage of two recently developed chemical agents that inhibit MLL1/2 activity, Mi-2 and MM-102. These small nonpeptide molecules decrease H3K4me3 levels on genes by disrupting the interaction between MLL-menin and MLL-WDR5, respectively (34, 35) (Fig. 6A). Our results show that in hESCs, a 24-h treatment with these inhibitors drastically reduces the levels of H3K4me3 on bivalent genes in mitotic cell-enriched (nocodazole-treated) cultures (Fig. 6B and D), without causing major morphological alterations.
(DMSO versus Mi-2/MM-102, 20 μM for 24 h) (see Fig. S5 in the supplemental material). These results further confirmed that H3K4me3 levels are increased during mitosis on certain bivalent genes, such as the PAX6 and HOXA4 genes (from the ME category), and that this phenomenon is dependent on the activity of MLL1/2 complexes (Fig. 6D). As expected, genes from the G1-dep (COL1A1 gene) and ubiquitous (PPP1R1A gene) categories did not display H3K4me3 enrichment after synchronization of the cells in mitosis. However, after treatment with MLL1/2 inhibitors, H3K4me3 levels on these genes were reduced (Fig. 6D). Consistent with the distribution of H3K4me3 near the TSSs of genes, the effect of the inhibitors was observed in regions proximal to the TSSs of all the genes analyzed (PAX6, HOXA4, COL1A1, PPP1R1A, and CDC25A genes) but not on a region distal to the TSS (distal CDC25A gene [Fig. 6D]).

Once we established that the acute treatment of human ESCs with a combination of inhibitors successfully inhibited MLL1/2 activity and partially or completely reduced the levels of H3K4me3 on bivalent genes, we assessed the effect on expression of bivalent genes during differentiation. To do this, we incubated cells for 24 h with Mi-2 or MM-102 (20 μM). Subsequently, the MLL1/2 inhibitors were withdrawn; cells were washed with fresh medium, resuspended, and plated as single cells in medium containing RA (1 μM); and then cells were incubated for 24 h (Fig. 6C).
Disruption of MLL1/2 trimethylase activity abrogates H3K4me3 mitotic enrichment and delays expression of bivalent genes upon induction of cellular differentiation. (A) MLL1/2 inhibitors MM-102 and Mi-2 disrupt the interaction of MLL proteins with WDR5 and menin, respectively. (B) Summary of the strategy used to analyze the effect of MLL1/2 inhibitors on H3K4me3 levels in H9 hESCs. Briefly, 2 days after passaging, cells were incubated in mTeSR-1 medium containing Mi-2/MM-102 (20 μM each) for 8 h, followed by 16 h of culture with nocodazole (200 ng/ml) plus Mi-2/MM-102 (20 μM each). Cells were then collected, fixed, and subjected to the ChIP protocol. (C) Strategy used to address the effect of blocking MLL1/2 activity before induction of differentiation in H9 hESCs. Briefly, 2 days after passaging, cells were preincubated in E8 medium containing 20 μM (each) Mi-2 and MM-102 for 24 h. Subsequently, single-cell suspensions were prepared, and 8 × 10⁶ cells were seeded in E6 medium containing 1 μM retinoic acid (RA); 24 h later, total RNA was extracted. (D) ChIP-qPCR analysis of the H3K4me3 levels in hESCs synchronized in mitosis with or without MLL1/2 inhibitors. DMSO, cells treated with DMSO only; Noc, cells blocked in mitosis with nocodazole; Noc + MLL Inh, cells blocked in mitosis with nocodazole in the presence of Mi-2 and MM-102 (20 μM each). Values for percent input are normalized to the IgG control (n = 2). One-way analysis of variance followed by Tukey’s test for multiple comparisons was performed. *, P < 10⁻²; ***, P < 10⁻⁴; ns, not significant. (E) Reverse transcription-qPCR analysis of RNA expression levels for bivalent genes after induction of differentiation in unsynchronized H9 ES cells pretreated with MLL inhibitors. Differentiation was induced using E6 medium containing 1 μM RA per 24 h. DMSO, cells pretreated with DMSO only; MLL Inh, cells pretreated with Mi-2 and MM-102 (20 μM each) for 24 h. Unpaired parametric t test with Welch’s corrections was performed. Two-tailed P values were calculated. *, P < 10⁻²; ***, P < 10⁻⁴; ns, not significant. Log₂ RNA levels were normalized against HPRT1 RNA levels and expressed as fold change between undifferentiated and differentiated cells. (F) Summary of the transcriptional outcome of decreasing the mitotic H3K4me3 enrichment in hESCs prior to induction of cellular differentiation. Sizes of the green circles and ovals denote levels of histone marks or protein enrichments around the TSSs (arrows) of bivalent genes. The thickness and number of arrows represent levels of gene expression.
Using qPCR, we found that, in response to RA induction, the expression of ME-category bivalent genes, the PAX6, HOXA4, GATA6, and MEIS2 genes, was reduced or delayed in cells that had been preincubated with MLL1/2 inhibitors, compared to differentiated cells that were not exposed to the MLL1/2 inhibitors (Fig. 6E). Interestingly, inhibition of MLL1/2 prior to the induction of differentiation also delayed the differentiation-dependent repression of genes that are expressed in the pluripotent state (NANOG and MIXL1 genes) (Fig. 6E). Importantly, treatment with MLL1/2 inhibitors did not affect the expression of control genes, the SIP1, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and HRPT1 genes, indicating that these inhibitors do not alter gene expression globally. This is consistent with the fact that in ESCs, MLL1/2 predominantly catalyze trimethylation of H3K4 on bivalent genes (29, 30).

Together, these results reveal that in hESCs, the inhibition of MLL1/2 activity is sufficient to reduce the levels of H3Kme3 on bivalent genes, which correlates with a delay in the expression of bivalent genes after induction of cellular differentiation (Fig. 6F).

**DISCUSSION**

The current consensus is that most bivalent genes are largely repressed in pluripotent cells (7, 28, 33) but can rapidly respond to initiators of differentiation by turning on expression in the G₁ phase of the cell cycle (36). Because the G₁ phase in pluripotent hESCs is considerably shorter than that in differentiating cells, it is reasonable to assume that a unique mechanism, active during the M-G₁ transition, is required for cells to mount such a rapid response. In this context, a highly dynamic bivalent landscape may be critical for generating local chromatin conformations that facilitate rapid unfolding of mitotic chromosomes and expedite transcriptional activation in response to differentiation cues during G₁. In this model, genes required for early differentiation are poised for expression only during a very brief window of time, in the context of a transcriptionally unfavorable mitotic environment that prevents unnecessary gene expression. Accordingly, in the absence of differentiation cues, the drastic reduction in H3K4me3 levels observed during G₁ phase would prevent leaky transcription during this phase of the cell cycle when chromatin becomes more accessible. The fact that these histone modification patterns become cell cycle independent after induction of differentiation further supports this scenario and implies that a cell cycle-dependent epigenetic landscape is instrumental in maintaining chromatin architecture plasticity in pluripotent cells but not in committed cells. This is consistent with the finding of Blobel et al. that H3K4me3 remains constant during the G₂-M-G₁ transition in HeLa cells (20). The lengthening of the cell cycle associated with lineage commitment (21, 37) would provide adequate time for cell cycle-independent, expression-regulation mechanisms.

This interpretation is consistent with a recent report from the Vallier lab, documenting that developmental genes not expressed in undifferentiated hESCs become actively transcribed during G₁ phase in response to differentiation cues (36). There is, however, a conflicting recent study reporting that developmental genes are expressed during the G₁ phase of undifferentiated hESCs (38). Importantly, though, this study did not include a selection step to isolate highly pluripotent cells.

Our results reveal the existence of new categories of bivalent domains in ESCs that are associated with genes that have specific cellular functions and expression patterns before and after the onset of differentiation. Importantly, by using a representative set of bivalent genes, we showed that similar cell cycle-dependent histone modification profiles are seen in hESCs purified at G₂, mitosis, and G₁ with or without prior chemical synchronization and that the epigenetic status of these genes correlates with the sequential recruitment of the H3K4 methylase/demethylase enzymatic machinery to chromatin. This is consistent with the fact that the inhibition of MLL1/2 activity decreases the levels of H3K4me3, implying that, after disruption of trimethylase activity, histone demethylases are ready to take over. This mechanism would be important to maintain the dynamic regulation of the H3K4me3 histone mark during the cell cycle.

Of note, hESCs treated to inhibit MLL1/2 activity proliferated at a lower rate than control cells (data not shown); this is consistent with a study reporting that human leukemia cells treated with Mi-2 are arrested in the G₁ phase of the cell cycle (35). This observation suggests that a properly functioning MLL1/2 complex is critical to sustain self-renewal and that the deteriorated self-renewing capability observed in ESCs after knockdown of components of the COMPASS complex may be due more to a lack of H3K4 methylase activity than to the physical absence of these proteins (20, 31, 33). Although inhibition of MLL1/2 trimethylase activity may alter the expression profile of genes necessary for self-renewal, it is also possible that normal patterns of H3K4me3 on bivalent chromatin are required for cells to properly progress through the cell cycle. This is important because it would provide a plausible explanation for the existence of bivalent chromatin domains that do not lose H3K4me3 during the pluripotent G₁ phase (genes in the ubiquitously bivalent category). Whether these bivalent domains are important for cell cycle progression, perhaps by establishing the right conditions for proper DNA replication, is something that should be evaluated in the future.

Importantly, as Pauklin and Vallier showed (36), hESCs respond to differentiation cues during the G₁ phase of the cell cycle. Therefore, it should not be surprising that delaying or arresting G₁-phase progression could have consequences for the timing of cellular differentiation. In this regard, we have shown that decreasing the levels of H3K4me3 on bivalent genes prior to induction of differentiation delayed the expression of critical developmental bivalent genes. This reduced expression seems to be a consequence of impeding the differentiation process, as suggested by the delayed repression of the NANOG gene observed in hESCs induced to differentiate after preinhibition of the MLL1/2 complex.

Overall, our findings strikingly reveal that, at least in pluripotent cells, chromatin structure associated with bivalent genes is not static but dynamically regulated during the cell cycle. The cell cycle-dependent control of epigenetic modification in bivalent domains represents a novel dimension to chromatin regulation that advances understanding of how hESC identity is maintained.

**ACKNOWLEDGMENTS**

We thank Eli Canaani in the Department of Molecular Cell Biology at the Weizmann Institute of Science, Israel, for donating the anti-MLL473 antibody, Lori Martin-Buley for manuscript editing, Prachi N. Ghule for valuable discussion, the Deep Sequencing Core Facility and Core Flow Cytometry Lab at the University of Massachusetts Medical School, Worcester, MA, the Advanced Genome Technologies Core, and the Flow Cytometry and Cell Sorting Facility at the University of Vermont, Burlingame, CA, and the Cytometry and Cell Sorting Facility at the University of Vermont, Burlington, VT.
REFERENCES


