



**HAL**  
open science

## **H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells.**

Krishanpal Karmodiya, Arnaud Krebs, Mustapha Oulad-Abdelghani, Hiroshi Kimura, Laszlo Tora

► **To cite this version:**

Krishanpal Karmodiya, Arnaud Krebs, Mustapha Oulad-Abdelghani, Hiroshi Kimura, Laszlo Tora. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells.. BMC Genomics, 2012, 13 (1), pp.424. 10.1186/1471-2164-13-424 . inserm-00742676

**HAL Id: inserm-00742676**

**<https://inserm.hal.science/inserm-00742676>**

Submitted on 16 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Open Access

# H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells

Krishanpal Karmodiya<sup>1</sup>, Arnaud R Krebs<sup>1,3</sup>, Mustapha Oulad-Abdelghani<sup>1</sup>, Hiroshi Kimura<sup>2</sup> and Laszlo Tora<sup>1\*</sup>

## Abstract

**Background:** Transcription regulation in pluripotent embryonic stem (ES) cells is a complex process that involves multitude of regulatory layers, one of which is post-translational modification of histones. Acetylation of specific lysine residues of histones plays a key role in regulating gene expression.

**Results:** Here we have investigated the genome-wide occurrence of two histone marks, acetylation of histone H3K9 and K14 (H3K9ac and H3K14ac), in mouse embryonic stem (mES) cells. Genome-wide H3K9ac and H3K14ac show very high correlation between each other as well as with other histone marks (such as H3K4me3) suggesting a coordinated regulation of active histone marks. Moreover, the levels of H3K9ac and H3K14ac directly correlate with the CpG content of the promoters attesting the importance of sequences underlying the specifically modified nucleosomes. Our data provide evidence that H3K9ac and H3K14ac are also present over the previously described bivalent promoters, along with H3K4me3 and H3K27me3. Furthermore, like H3K27ac, H3K9ac and H3K14ac can also differentiate active enhancers from inactive ones. Although, H3K9ac and H3K14ac, a hallmark of gene activation exhibit remarkable correlation over active and bivalent promoters as well as distal regulatory elements, a subset of inactive promoters is selectively enriched for H3K14ac.

**Conclusions:** Our study suggests that chromatin modifications, such as H3K9ac and H3K14ac, are part of the active promoter state, are present over bivalent promoters and active enhancers and that the extent of H3K9 and H3K14 acetylation could be driven by cis regulatory elements such as CpG content at promoters. Our study also suggests that a subset of inactive promoters is selectively and specifically enriched for H3K14ac. This observation suggests that histone acetyl transferases (HATs) prime inactive genes by H3K14ac for stimuli dependent activation. In conclusion our study demonstrates a wider role for H3K9ac and H3K14ac in gene regulation than originally thought.

**Keywords:** ChIP-seq, Histone acetylation, CpG islands, Embryonic stem cells, Gene regulation, Genome-wide mapping, Bivalent promoters, Epigenetics

\* Correspondence: laszlo@igbmc.fr

<sup>1</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 964, Université de Strasbourg, BP 10142-67404 ILLKIRCH Cedex, CU de Strasbourg, France

Full list of author information is available at the end of the article

## Background

Embryonic stem (ES) cells provide an important model system to study developmental regulation and hold significant potential for clinical therapeutics because of their unique capabilities to self re-new and differentiate into multiple lineages (reviewed in [1]). The chromatin of pluripotent ES cells have unique characteristics, including an open conformation, a hyper-dynamic organization of chromatin proteins, and less condensed heterochromatin domains, suggesting the plasticity of the genome in ES cells (reviewed in [2]).

Different modifications of chromatin are associated with variable functions. Histone modifications such as trimethylation of H3 lysine 4 (H3K4me3) and hyperacetylation of histone H3 and H4 are known as active marks and are often associated with ongoing transcription [3,4]. On the other hand, methylation of H3K9 and H3K27, are known as repressive marks and are associated with gene silencing [5]. Promoters of key regulatory genes have unique chromatin modification signatures, which contain both an active histone mark, H3K4me3, as well as a repressive histone mark, H3K27me3, also known as bivalent promoters, and are thought to be poised for gene activation during differentiation [6-8]. Distal regulatory regions, such as enhancers, are enriched in H3K4me1 (as compared to H3K4me3), histone acetyl transferase (HAT) co activators (i.e. p300 or ATAC) and have an open chromatin structure [4,9,10]. Not only promoters, but enhancer regions were also shown to be poised for gene activation during differentiation as only active enhancers are marked by the H3K27ac modification [11,12].

One of the most studied modifications of histones is acetylation of specific lysine (K) residues, which generally correlates with gene activation. The level of histone acetylation is regulated by the activity of both histone acetyl transferases (HATs) and histone deacetylases (HDACs), which acetylate and deacetylate lysine residues of the N terminal histone tails, respectively. Genetic and biochemical studies suggested that HATs have rather specific roles in gene activation, while genome-wide experiments rather suggested that HATs are often recruited simultaneously, and together are acetylating multiple lysine residues at a given loci [13,14]. Thus, the biological function of histone acetylation may be rather additive than specific. In ES cells, acetylation of H3K9 was shown to predict the pluripotency and reprogramming capacity [15] and its level reduces with ES cell differentiation [16]. Recent genome-wide studies shed light on various histone modifications in mES cells [6], however, the genome-wide role of histone H3 acetylation in mES cells is poorly understood. In this study, we have investigated two histone acetylation marks, H3K9 and H3K14. Acetylation of H3K9 is mainly performed by histone acetyl transferases GCN5/PCAF and/or Tip60,

whereas acetylation of H3K14 is mediated by GCN5/PCAF, p300/CBP and/or Myst3 [17-19].

The lack of in-depth studies concerning the role of H3K14ac was due to the absence of reliable and specific antibodies. Antibodies against modified histone tails are central research tools in studying chromatin biology at a genome-wide level. Thus, we have developed a new specific ChIP-grade antibody against the H3K14 acetylation mark. By using the antibody we have developed against H3K14ac and another published commercial antibody against H3K9ac, we made genome-wide location analysis of these two acetylation marks in mouse (m) ES cells, and compared the presence of two marks over various genomic regions. Our study suggests that these two marks are present not only over promoters of actively transcribed genes, but also on the developmentally regulated bivalent promoters, as well as over active enhancers in mES cells. Moreover, the degree of H3K9 and H3K14 acetylation correlates with the CpG content of the promoters and transcription level of the genes. Finally, we observed differential presence of these two acetylation marks over a subset of inactive genes, which is marked by low level of H3K14ac and thus seemed to be prepared for future activation.

## Results

### Genome-wide acetylation profiles of H3K9 and H3K14 correlate with each other

To understand the role of histone H3 acetylation at positions K9 and K14 in mouse ES cells, we have systematically tested the available antibodies raised against these modified histone tails. We and others have found that the anti-H3K9ac antibody from Abcam (ab4441) is specific for the corresponding modification in various applications including chromatin immunoprecipitation (ChIP) [20,21]. However, the anti-H3K14ac antibody from Upstate (07-353), which was used earlier for genome-wide localization of H3K14ac [3], was shown to cross-react with other histone modifications [20] and was found non-applicable for chromatin immunoprecipitation coupled high throughput sequencing (ChIP-seq) [21]. Our ELISA tests showed that it not only cross-reacts with the H4K5acK12ac peptide, it also recognizes the non-acetylated H3K14 peptide (Additional file 1: Figure S1). To overcome this limitation and to study the real genome-wide distribution of H3K14ac, we raised a specific mouse monoclonal antibody (mAb) against this modification and confirmed its specificity in several different tests (Additional file 2: Figure S2).

To gain insight in the genome-wide acetylation profile of H3K9 and H3K14 residues, ChIP-seq was performed using the commercially available antibody against H3K9ac and the new antibody developed against H3K14ac in this study. Peaks of local enrichment for H3K9ac and

H3K14ac were determined after sequence alignment and normalization to input DNA. Further to validate the peaks obtained in ChIP-seqs for H3K9ac (using ab4441 antibody) and H3K14ac (using 13HH3-1A5 antibody), we performed ChIP-qPCR on randomly selected genomic loci enriched for H3K9ac and H3K14ac. All the selected peaks of H3K9 and H3K14 acetylation from the ChIP-seq experiments were validated by ChIP-qPCR (Additional file 3: Figure S3 and Additional file 4: Table S1).

In order to test whether H3K9ac and H3K14ac modifications have differential preference over various chromatin regions, we compared their presence over promoters (2000 bp upstream of transcription start sites (TSSs)), coding exons, introns and distal intergenic regions (Figure 1A), which represent 2.5%, 1.6%, 38.7% and 57.2% of the total genome, respectively [22]. First, peaks of H3K9ac and H3K14ac local enrichment were determined after sequence alignment and normalization to input DNA using MACS [23]. Our analyses show that H3K9ac and H3K14ac peaks are distributed over all four genomic regions and the frequency of distribution over promoters is 13.2% and 12.4%, respectively (Figure 1A). However, approximately 85% of both the H3K9ac and H3K14ac marks are observed in distal intergenic and intronic regions with significant enrichments, comparable to promoters (Additional file 5: Figure S4), suggesting that these two modifications may have a role at distal intergenic and intronic regions. Further to compare H3K9 and H3K14ac marks, a combined list of binding sites at transcription start sites (TSSs) was established containing 15595 TSSs. The comparison of H3K9ac and H3K14ac over these TSSs shows a Pearson correlation coefficient of 0.73 (Figure 1B), suggesting that the studied two H3 acetylation marks are present simultaneously on promoters (Figure 1C). Moreover, H3K9 and H3K14 acetylations have a characteristic bimodal distribution around the TSSs, with one peak upstream of the TSS, another single peak (stronger in the case of H3K14ac) downstream of the TSS, and depletion of the signal right on the TSS (Figure 1C). To examine the distribution of these two histone marks over the gene body, composite profile of both marks spanning the entire gene body and extending 5 kb upstream from TSSs and 5 kb downstream of the 3' end of the genes for combined list of binding sites over TSSs (15595) was generated (Figure 1D). The H3K9ac and H3K14ac distribution profiles around TSSs suggest that both marks are predominantly located in regions surrounding the TSSs of genes. Further to confirm that the co-occupancy of H3K9 and H3K14 observed is not because of cellular heterogeneity, we performed sequential ChIP for H3K9ac followed by H3K14ac (Figure 1E). Sequential ChIP demonstrates that genomic loci are acetylated simultaneously both at H3K9 as well H3K14. Thus, our analyses suggest that on

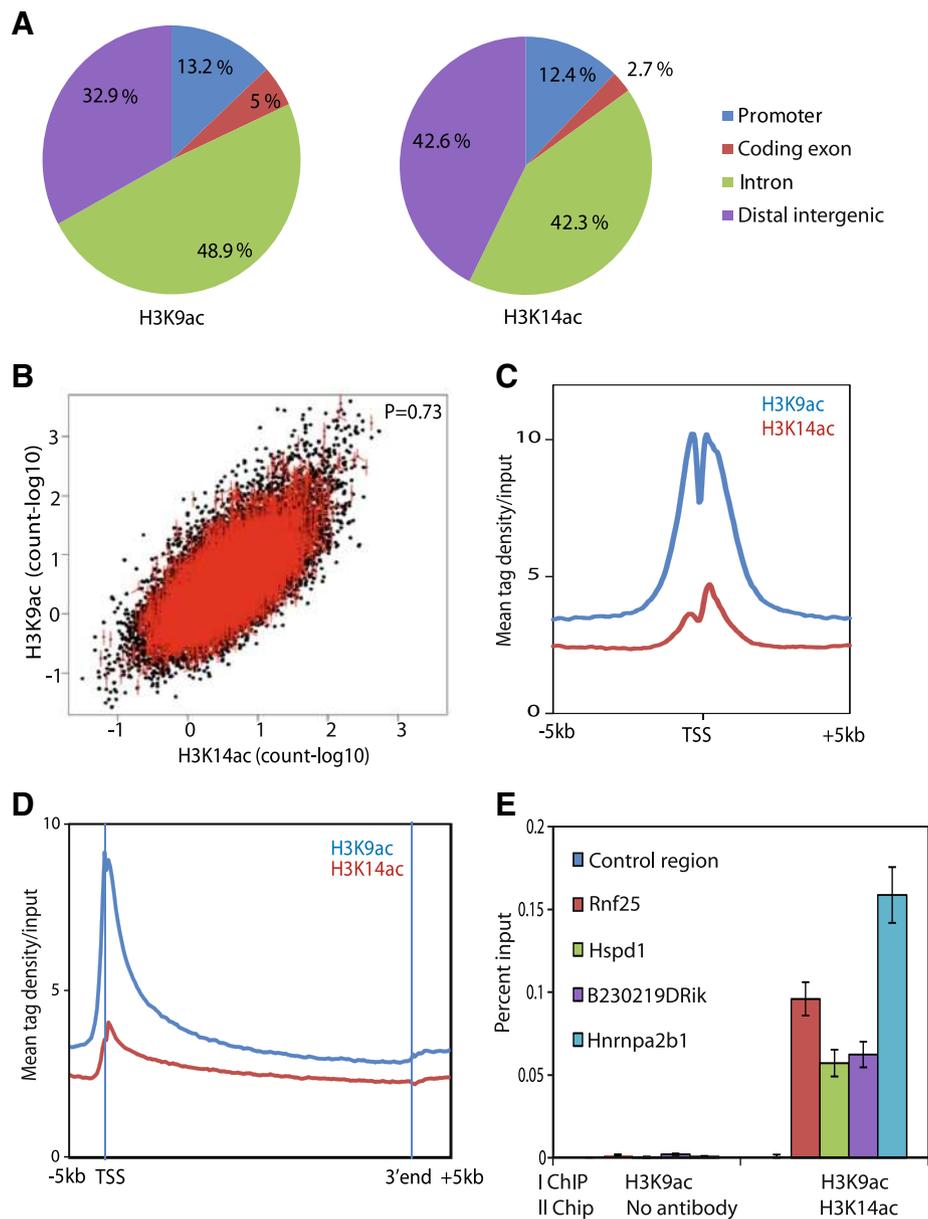
the genome H3K9ac and H3K14ac are mostly present at distal intergenic and intronic regions, specifically enriched at promoters when localized in the vicinity of genes and that the two acetylation marks co-occur at promoters as well as other locations.

#### **Levels of H3K9 and H3K14 acetylations correlate with the magnitude of gene expression**

Global transcription is a hallmark of pluripotent ES cells that contributes to plasticity and lineage specification [24]. Histone modifications are known to act in a combinatorial fashion to determine the overall outcome of the gene expression [3]. To explore the correlative relationship between various active histone marks detected at promoters and the transcription of the corresponding genes in mES cells, we compared the level of various active histone marks with the transcriptional level of the genes. Densities of active histone marks (H3K9ac, H3K14ac, H3K4me3 and H3K27ac), as well as total H3 and RNA polymerase II (Pol II), within a 3000 bp window flanking the TSSs of the expressed genes (12100) were collected. All expressed genes [8] were divided into ten categories ranked on the basis of their expression level. Presence of various active histone marks was analyzed over these categories. Analysis of histone H3 occupancy, histone modifications (H3K9ac, H3K14ac, H3K4me3 and H3K27ac) and Pol II around the TSSs suggest that depletion of the total histone H3 signal and enrichment of the active promoter marks (H3K9ac, H3K14ac, H3K4me3 and H3K27ac) at or around TSSs correlate with the increase in gene expression levels. Interestingly, H3K9ac is more spread than the other analyzed active histone marks around the TSSs (Figure 2). While Pol II is enriched at or slightly downstream of the TSSs, on these sites the nucleosomes (H3) are depleted (Figure 2). This genome-wide observation with various active histone marks is consistent with the notion that H3K9ac and H3K4me3 near the TSSs destabilize interaction between histones and DNA leading to nucleosome eviction [25,26]. Taken together, these results suggest that level of active histone marks (H3K4me3, H3K9ac, H3K14ac and H3K27ac) over the active promoter chromatin state correlates with the magnitude of gene expression.

#### **H3K9 and H3K14 acetylation levels correlate with the CpG content**

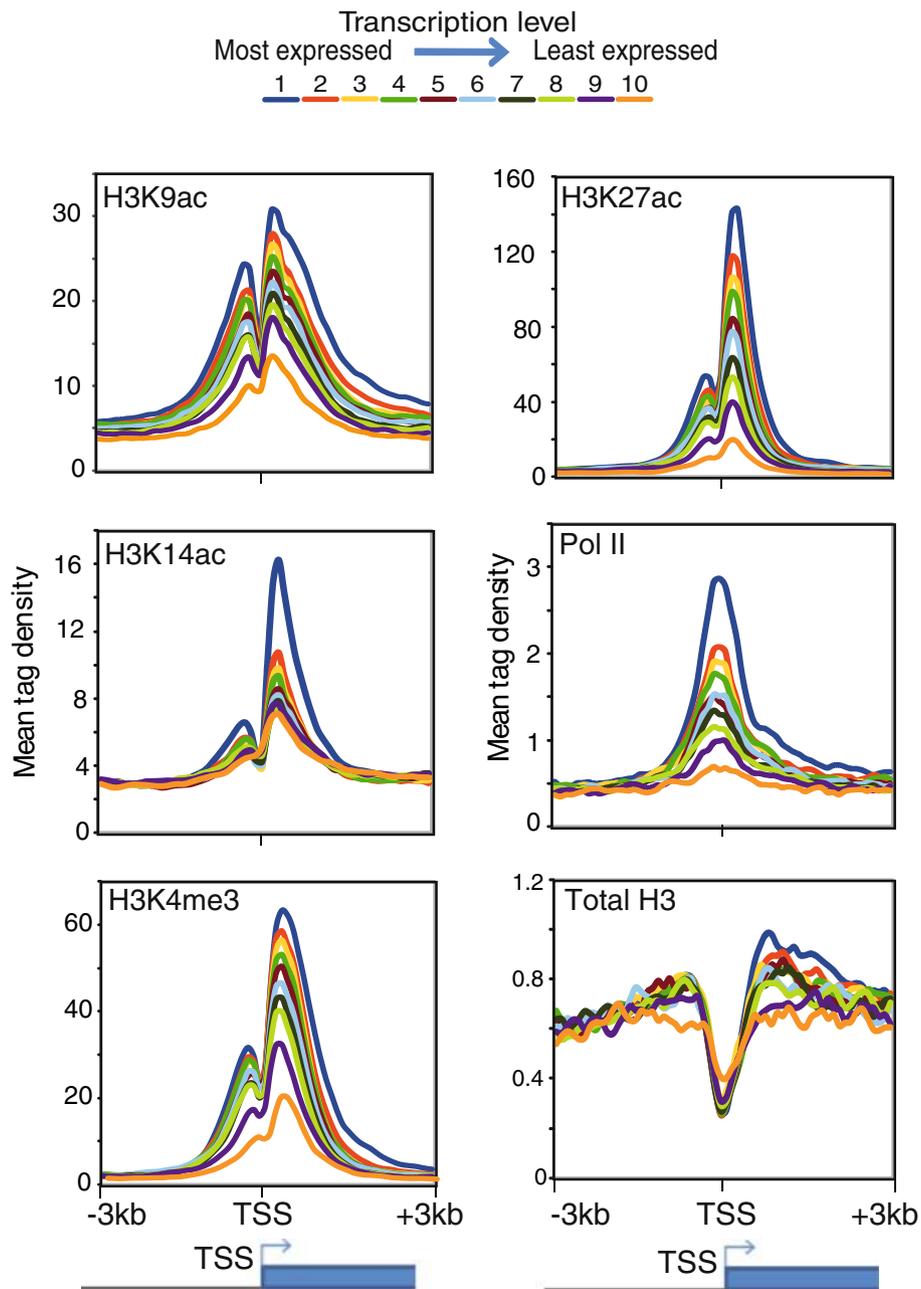
Cytosine-phosphate diester-guanine (CpG) islands are usually found at the 5' end of the regulatory regions of genes [27]. CpG islands are GC rich, predominantly non-methylated and their content correlates with H3K4me3 chromatin modification. To explore the relationship between the CpG content and the levels of H3K9ac and H3K14ac, we took all the CpG island sites (16026) from



**Figure 1 Genomic distribution of H3K9ac and H3K14ac in mES cells.** (A) Distribution of H3K9ac and H3K14ac peaks over the promoters (2000 bp upstream of TSS), coding exons, introns and distal intergenic regions. Many of the H3K9 and H3K14 acetylation peaks are at distal intergenic regions. (B) Dot plot representation of genome-wide co-localization analysis of the H3K9ac and H3K14ac modifications over the 15595 combined promoter list of H3K9 and H3K14ac suggests a strong correlation between these two modifications at promoters. (C) Average input normalized profile of 15595 combined promoter list of H3K9 and H3K14ac around the transcription starts sites shows bimodal distribution. (D) Average input normalized whole gene profiles for H3K9ac and H3K14ac modifications over 15595 combined promoter list of H3K9 and H3K14ac. (E) Sequential ChIP-qPCR quantification for co-occupancy of H3K9ac (primary ChIP) and H3K14ac (secondary ChIP) at randomly selected H3K9 and H3K14 acetylated loci suggest that these loci are co-marked with H3K9 as well H3K14 acetylation. Enrichment after first ChIP using H3K9ac followed by re-ChIP with no antibody was used as a control. Primer sequences used in ChIP-qPCR is provided in Additional file 4: Table S1. Error bars represent the standard deviation for three technical replicates.

UCSC genome browser [28] and sorted them according to their CpG content. On these CpG islands, which are sorted on the basis of their increasing CpG content, we looked for the H3K9ac and H3K14ac profile. We found that indeed CpG content correlates with the level of

H3K9 and H3K14 acetylation, as acetylation over those sites increases in parallel with the CpG content (Figure 3). This in turn suggests that the levels of H3K9 and H3K14 acetylations on the nucleosomes positioning around the TSSs of the promoters correlate with the CpG content of



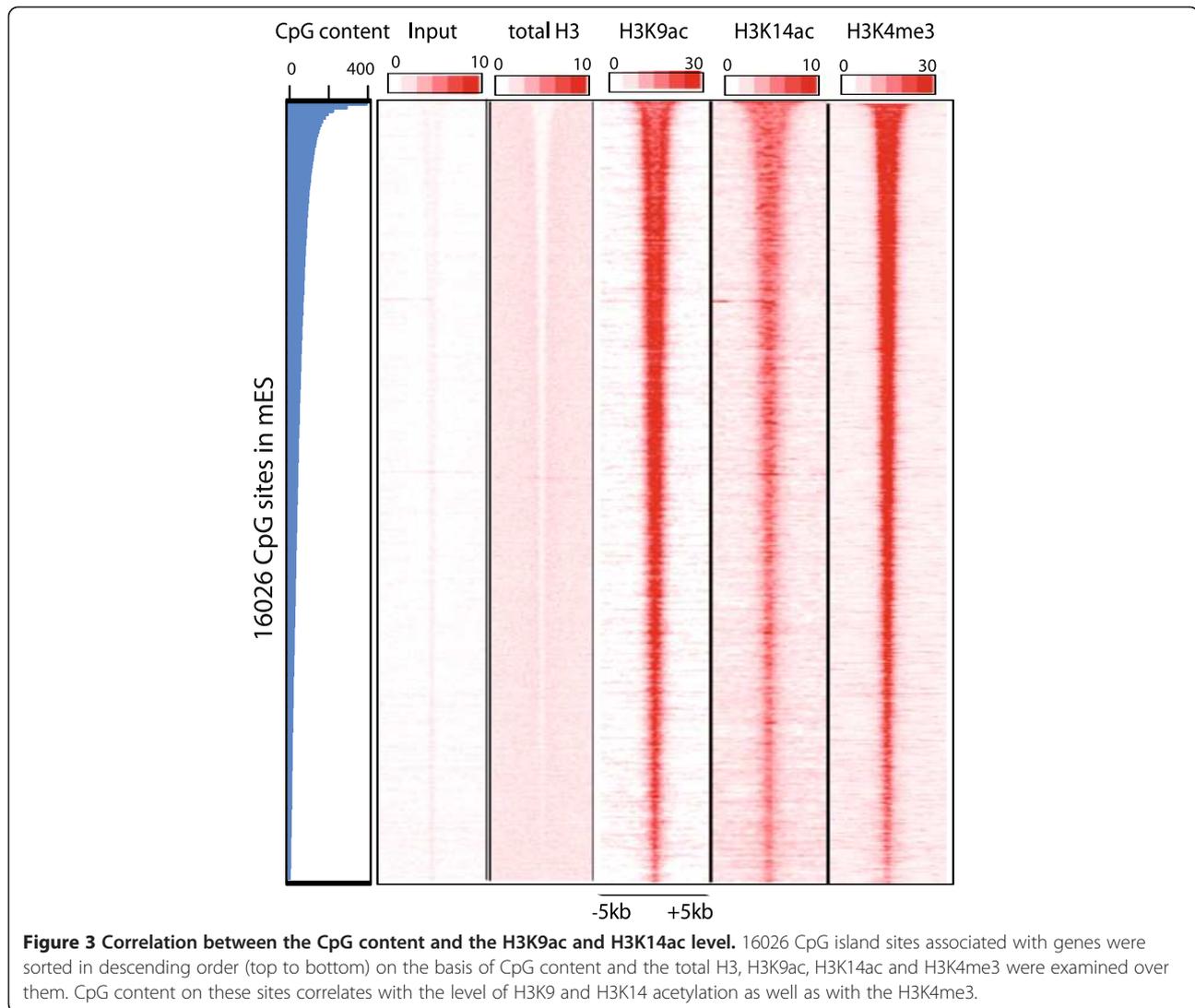
**Figure 2 Level of H3K9 and H3K14 acetylation correlates with magnitude of gene expression.** A total of 12000 expressed genes in mES cells were divided into ten groups based on their expression levels, from the top 10% (blue, group 1) to the lowest 10% (purple, group 10). Mean tag densities of active histone marks; H3K9ac, H3K14ac, H3K4me3, H3K27ac and Pol II within (-/+ ) 3 kb are positively correlated with the transcription level of the genes. On the other hand, total H3 densities in the same regions are negatively correlated with the transcription level.

the underlying DNA sequence. This is in accordance with the fact that CpG enriched genes are generally housekeeping and are widely expressed [8,29].

#### H3K9 and H3K14 acetylations occur at active enhancers

Transcription from enhancers resulting in enhancer RNAs (eRNAs) play important regulatory role in maintenance of gene expression programs [30-32]. Enhancers are key cis-

regulatory elements that can affect gene expression independent of their orientation or distance in a cell type specific manner [9]. Enhancers are marked by the presence of H3K4me1, DNase I hypersensitivity and histone acetyltransferases such as p300 [4,9] or the GCN5/PCAF-containing ATAC complex [10]. The important proportion of H3K9 and H3K14 acetylation sites in distal intergenic regions (Figure 1A) motivated us to further analyze



these sites for the presence of various histone modifications, Pol II and p300 that are indicative of enhancers. Identification of H3K9ac and H3K14ac peaks in intergenic regions is described in Materials and Methods. During this analysis we found a strong correlation between H3K9ac or H3K14ac intergenic sites with either, H3K4me1, H3K27ac, the presence of Pol II and p300 suggesting that H3K9ac and H3K14ac mark also enhancers (Additional file 6: Figure S5). To confirm the presence of H3K9ac and H3K14ac over enhancers, we took 25036 putative enhancers reported in mES cells [11] and subjected them to k-means clustering using seqMiner [33]. Further, H3K27ac was used to distinguish active and inactive/poised enhancers. Our analyses show that active enhancers are marked by the strong presence of H3K9ac and H3K14ac along with H3K27ac and H3K4me1 (Figure 4A and 4B) and those inactive/poised enhancers are marked by the presence of H3K4me1 together with relatively weak levels of H3K14ac (Figure 4A and 4C). Thus, our study

suggests that H3K9ac and H3K14ac mark enhancers and can further discriminate active enhancers from poised/inactive enhancers.

#### **Bivalent promoters are also marked by H3K9 and K14 acetylation in pluripotent mES cells**

Many promoters of developmentally regulated genes in mES cells are marked by H3K4me3 (active histone mark) as well as polycomb mediated repressive histone mark, H3K27me3 [6,7] and are known as bivalent promoters. Recent studies have shown that these bivalent promoters are also bound by Pol II and are transcribed at very low level [34]. As these promoters have H3K4me3 and show a very low level of active transcription we looked for the presence of H3K9 and H3K14 acetylation over the bivalent promoters to see if this low level of transcription would also associated with acetylation on these promoters.

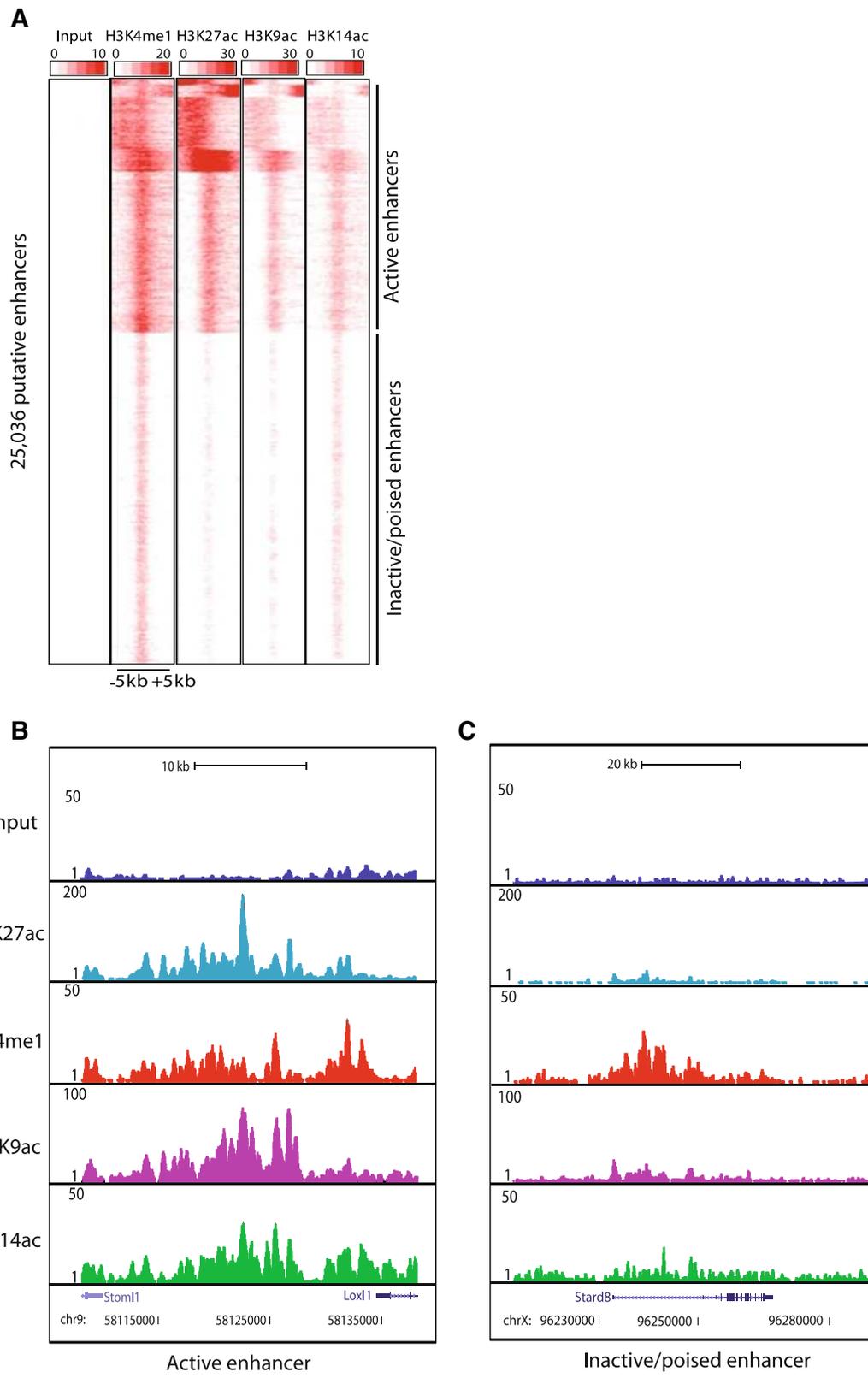


Figure 4 (See legend on next page.)

(See figure on previous page.)

**Figure 4 H3K9ac and H3K14ac mark active enhancers along with H3K27ac.** (A) Heatmap of the signal density using k-means clustering observed on 25036 putative enhancers (-/+ 5 kb) from mES cells for H3K4me1 (mark of putative enhancers), H3K27ac, H3K9ac and H3K14ac. On the basis of presence or absence of H3K27ac enhancers are categorized as active or poised/inactive. UCSC genome browser track of representative examples of (B) active enhancer and (C) poised/inactive enhancer. Active enhancers have significant enrichment of H3K9ac and H3K14ac as compared to poised/inactive enhancers as was observed for H3K27ac.

In order to test the presence of H3K9 and H3K14 acetylations over the bivalent promoters, all the 27095 mouse promoters from UCSC genome browser were taken [28] and subjected to k-means clustering using seqMiner [33]. H3K4me3 and H3K27me3 (dual hallmarks of bivalent promoters) were used to get the bivalent loci. Pol II was included in the clustering to differentiate active and inactive genes. The resulting heatmap is shown in Figure 5A. In agreement with previous genome-wide studies, there are three distinct categories of loci. In respect to the studied acetylation (i) active loci marked by Pol II, H3K4me3, and devoid of strong H3K27me3 signals are highly enriched in H3K9ac and H3K14ac, (ii) bivalent loci, which are marked by H3K4me3, H3K27me3 and low Pol II signals contain both H3K9ac and H3K14ac signals, and (iii) inactive loci, on which either of the above signals is missing. Our analysis suggests that indeed on bivalent promoters H3K9 and H3K14 acetylations occur together with H3K4 and H3K27 trimethylations (Figure 5A).

To test the presence of H3K9ac and H3K14ac over bivalent promoters, we took a subset of randomly selected loci and successfully validated them by ChIP-qPCR (Figure 5B). Further to test if H3K9ac and H3K14ac are functional at these bivalent loci, a HDAC inhibitor and followed the level of H3K9ac and H3K14ac over time. First, Oct4 levels were monitored by Western blot to assess the pluripotent state of the ES cells at various time points after sodium butyrate treatment. We found that Oct4 levels are comparable in non-treated and sodium butyrate treated ES cells at various time points (Additional file 7: Figure S6), suggesting that the pluripotent state of the cells did not change during the treatment. Next we analyzed two selected loci for the presence of H3K9ac and H3K14ac during the sodium butyrate treatment. We found that inhibition of HDACs by sodium butyrate leads to increase in H3K9ac and H3K14ac suggesting that these marks are indeed functionally deposited and are actively maintained (Figure 5C). UCSC genome browser tracks of two representative examples on previously characterized bivalent promoters [8] harboring the H3K9ac and H3K14ac over the promoters along with H3K4me3 and H3K27me3 are shown in Figure 5D and 5E. Thus, our genome-wide analyses show that H3K9ac and H3K14ac mark bivalent promoters along with active (H3K4me3) and repressive (H3K27me3) histone marks over developmentally regulated genes in undifferentiated ES cells.

#### Differential correlation of H3K14ac with repressive histone marks as compared to H3K9ac

We next analyzed co-localization of the H3K9 and H3K14 acetylation marks with various active and repressive histone modifications. The Pearson correlation coefficient was calculated for these modifications in a window of 2 kb upstream and downstream the TSSs of all the mouse refseq genes (27095). Heatmap of various histone modifications suggest that active histone marks have high correlation and are grouped together for efficient gene expression (Figure 6A, blue square). Surprisingly, we observed higher correlation of H3K14ac with various repressive marks, such as H3K27me3 and H3K9me3, when compared to H3K9ac (Figure 6A, red square). To further test the association of H3K14ac with inactive marks, H3K27me3 and H3K9me3, we took 7924 inactive genes, which lack Pol II and H3K4me3 (Figure 5A) and calculated the ratio of H3K14ac/H3K9ac tag density. The ratio of H3K14ac/H3K9ac is significantly higher over inactive genes as compared to active genes suggesting that H3K14ac is specifically and significantly enriched at inactive genes (Figure 6B). In Figure 6B we used H3K4me3, Pol II signals distinguish active TSSs from inactive ones. To further strengthen our observations we took the 500 weakest (or not expressed) and the 500 highest expressed gene promoters, but this time based on their expression profiles as calculated from RNA-seq data (kindly provided by the D. Schübeler), and calculated the ratio of H3K9 and H3K14 acetylation tag densities over the least expressed and highly expressed promoters. This analysis again shows that H3K14ac is specifically enriched at the weakest (or inactive) promoters as compared to H3K9ac (Additional file 8: Figure S7). This in turn suggests that H3K14ac, which is generally considered as a mark of active promoters along with other acetylation marks, can also mark inactive promoters, although to a lesser extent.

#### A subset of inactive genes is specifically enriched for H3K14ac

To understand the functionality of H3K14ac over inactive promoters, we examined the level of this histone mark over active and inactive genes in presence of sodium butyrate, a HDAC inhibitor, and compared it to H3K9ac. In agreement with our above observations, we found that active genes exhibited a remarkable increase in H3K9ac and H3K14ac in presence of sodium butyrate (Figure 7A

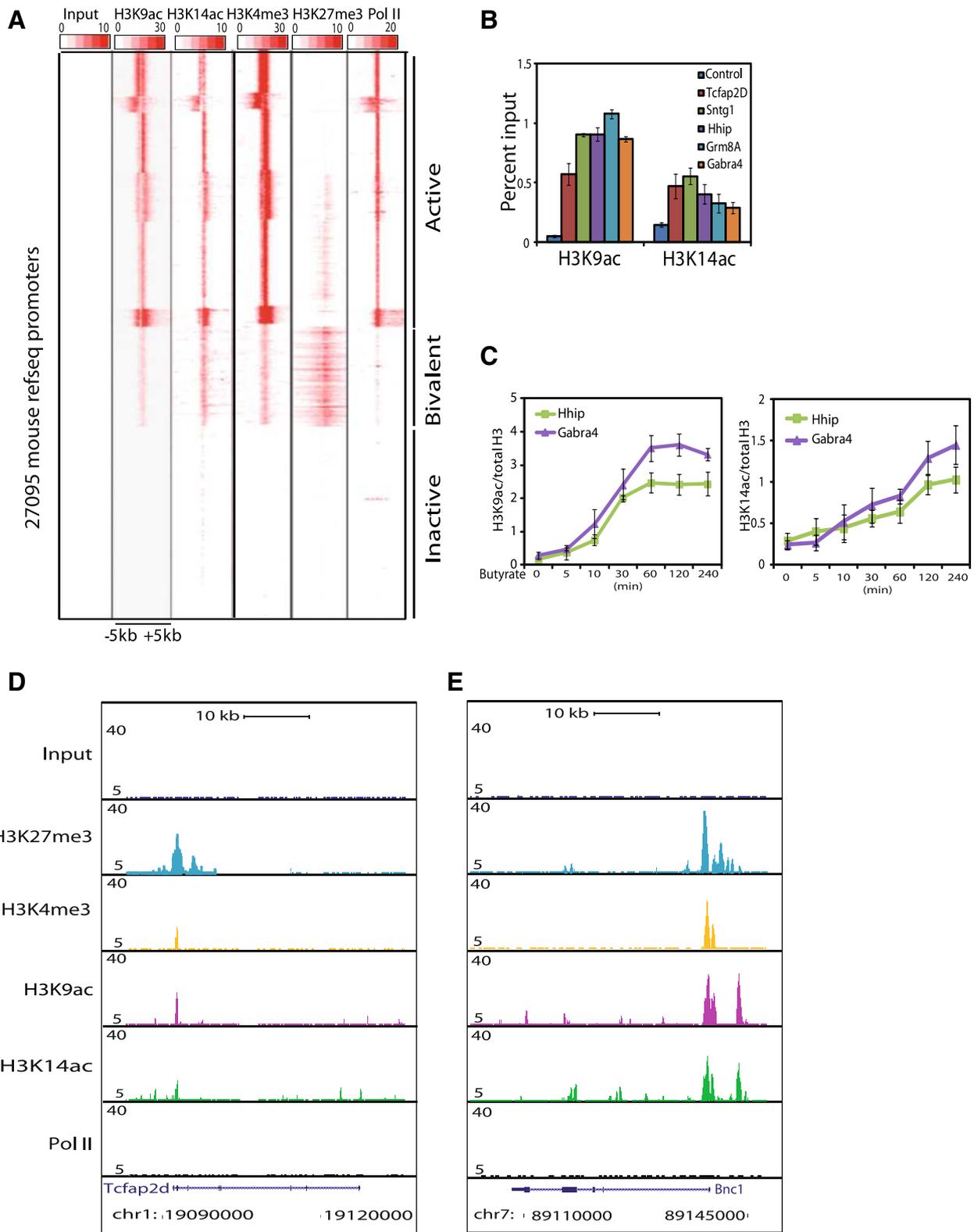


Figure 5 (See legend on next page.)

(See figure on previous page.)

**Figure 5 H3K9ac and H3K14ac associate with active as well as bivalent promoters.** (A) Heatmap of the signal density using k-means clustering observed on 27095 mouse refseq promoters (-/+5 kb) for H3K9ac and H3K14ac along with H3K4me3 and H3K27me3 (hallmark of bivalent promoters) and Pol II. The clustering of density map shows three different categories of genes. Active promoters having H3K9ac, H3K14ac, H3K4me3 and Pol II, bivalent promoters showing H3K4me3 and H3K27me3 along with H3K9ac and H3K14ac, and inactive promoters lacking all above marks along with Pol II. (B) Presence of H3K9ac and H3K14ac over randomly chosen bivalent loci was validated by ChIP-qPCR. (C) Increase in the H3K9 and H3K14 acetylation over bivalent promoters (Hhip and Gabra4) following HDAC inhibition by sodium butyrate. The presence of H3K9ac and H3K14ac over these bivalent loci at the indicated time points after the sodium butyrate treatment was measured by ChIP-qPCR. ChIP signals for H3K9ac and H3K14ac were normalized to total H3. Primer sequences used in ChIP-qPCR is provided in Additional file 4: Table S1. Error bars represent the standard deviation for three technical replicates. (D and E) UCSC genome browser track of two representative examples of loci showing H3K9ac and H3K14ac over the bivalent promoters.

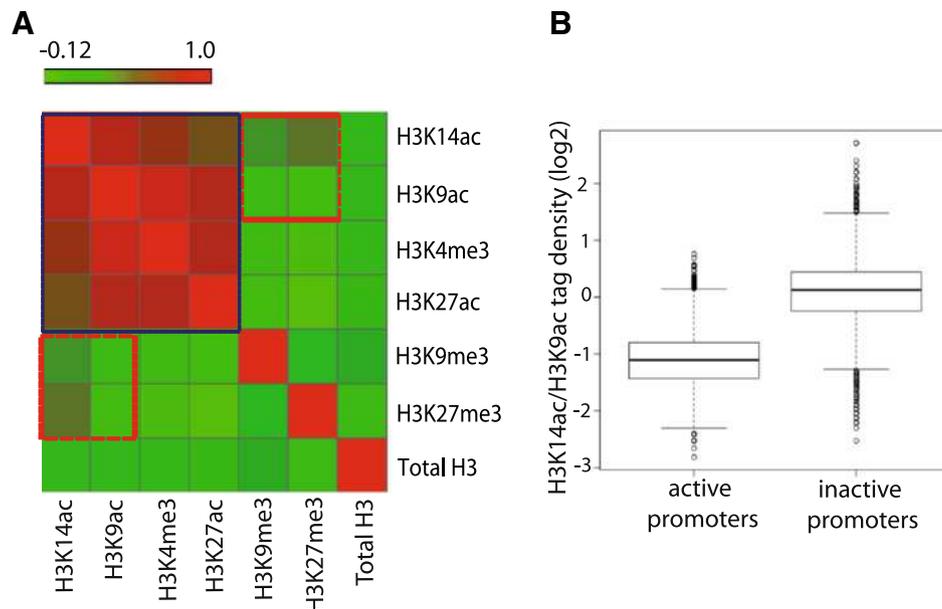
and 7B). However, examination of H3K9ac and H3K14ac level at inactive genes in presence of sodium butyrate revealed a slow and selective increase in H3K14ac, while at these sites H3K9ac did not change (Figure 7C and 7D). The selective increase in H3K14ac at inactive genes in presence of sodium butyrate suggests that inactive genes are subjected to constant H3K14 acetylation and deacetylation. This dynamic H3K14ac at inactive genes may poise them for future activation. To test this hypothesis, we selected 500 inactive genes having higher H3K14ac as compared to H3K9ac (Figure 8A) and subjected them to gene ontology (Figure 8B). Our analysis suggests that inactive genes, which are having significant H3K14ac over H3K9ac belong to various pathways that are induced by various stimuli such as sensory perception, olfaction and chemosensory perception. Other pathways include receptor activities, which are induced upon ligand binding to

activate the signal transduction and ultimately cellular responses (Figure 8B). Thus, our data indicate that H3K14ac is selectively present over a subset of non-expressed genes and HDACs frequently remove this mark to keep the genes inactive. The coordinated action of HATs and HDACs may poise these genes for stimuli dependent activation.

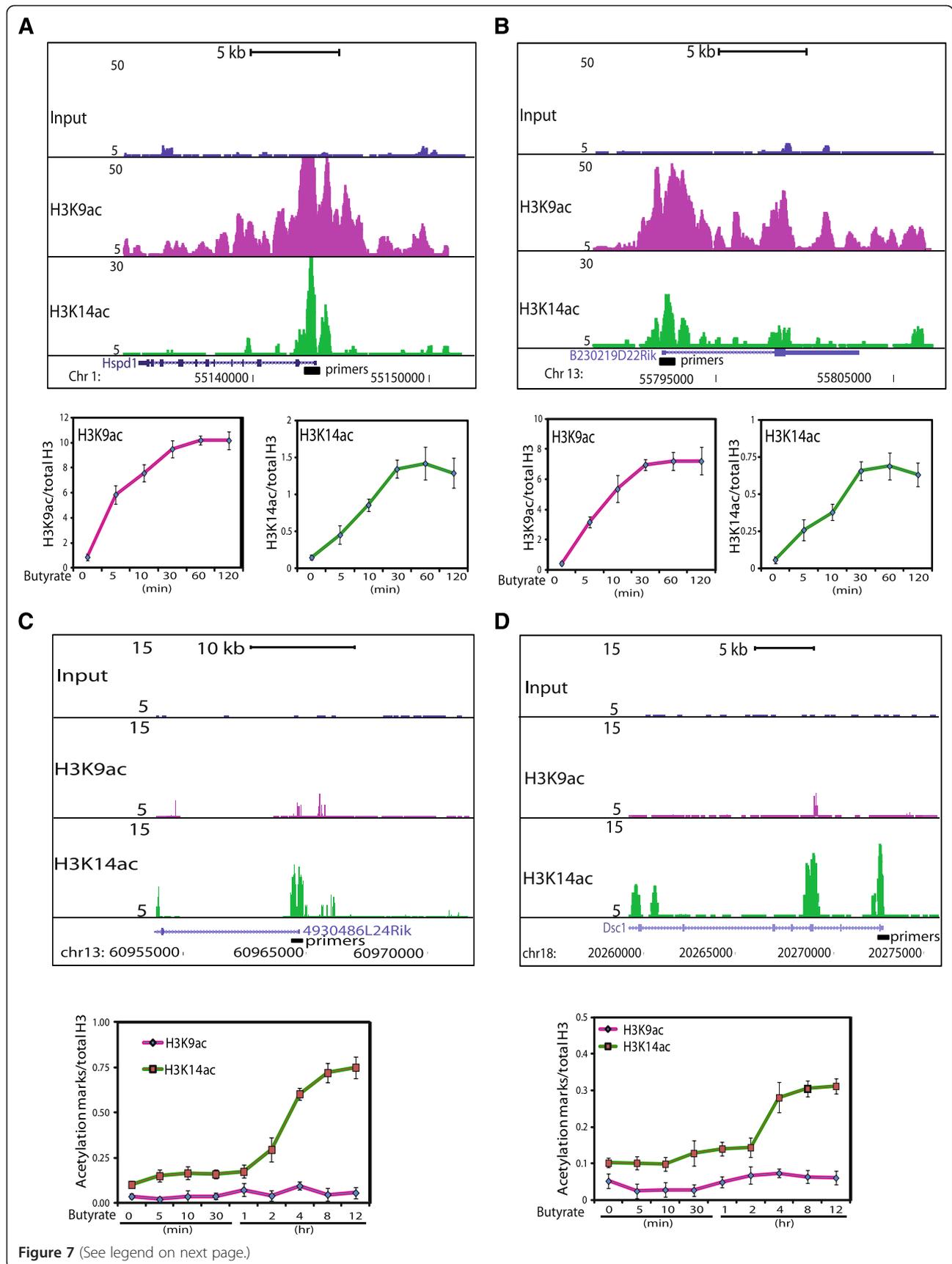
## Discussion

### H3K9ac and H3K14ac co-occur with other "active" histone modifications establishing a chromatin conformation that is compatible with transcription

Recent genome-wide studies have generated comprehensive chromatin landscapes for various transcription factors and histone modifications [3,14,35]. In our study, by using a new specific H3K14ac antibody, we show that H3K9ac and H3K14ac co-occur with other active histone



**Figure 6 Differential H3K14ac over inactive promoters as compared to H3K9ac.** (A) Heatmap representing the correlation between total H3, H3K9ac, H3K14ac, H3K4me3, H3K27ac, H3K27me3 and H3K9me3 around the TSS (-/+ 2 kb). Active promoter marks; H3K14ac, H3K9ac, H3K4me3 and H3K27ac clustered separately (blue square) to form an active promoter chromatin state. Inactive marks (H3K9me3 and H3K27me3) have higher co-occurrence with H3K14ac as compared to H3K9ac (red square) (B) Ratio of H3K14ac/H3K9ac ChIP-seq tag density plotted for active promoters and inactive promoters. Ratio is significantly higher for inactive promoters suggesting that level of H3K14ac is higher over inactive promoters as compared to H3K9ac.



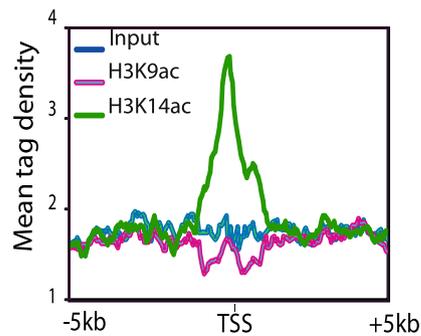
(See figure on previous page.)

**Figure 7 Selective increase in H3K14ac over inactive genes in presence of sodium butyrate.** (A, B) Rapid increase in the H3K9 and H3K14 acetylation over active promoters (Hspd1 and B230219D22Rik) caused by HDAC inhibitor. The upper panels show the UCSC genome browser tracks for the genes analysed. (C, D) Selective increase in H3K14ac over the inactive promoters (4930486L24Rik and Dsc1) as compared to H3K9ac following the treatment with HDAC inhibitor. The upper panels show the UCSC genome browser track for the genes analysed. The presence of H3K9ac and H3K14ac over these loci at the indicated time points after the sodium butyrate treatment was measured by ChIP-qPCR. ChIP signals for H3K9ac and H3K14ac were normalized to total H3. Positions of the primers used for ChIP-qPCR is shown in the UCSC genome browser track and primer sequences are provided in Additional file 4: Table S1. Error bars represent the standard deviation for three technical replicates.

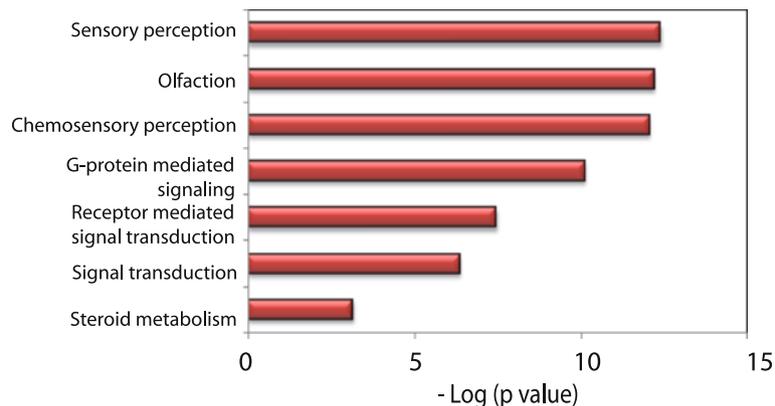
modifications, H3K4me3 and H3K27ac in mES cells. Thus, it seems that at these sites all these “active” histone modifications are deposited by the corresponding activities sequentially, or at the same time, to act in a coordinated way. A possible mechanistic cross-talk between H3K4me3 and H3 acetylation could occur in the following steps: (i) histone lysine methyltransferases (HMT), Set1/COMPASS associates with the early initiating RNA Pol II through the PAF1 complex to methylated histone H3K4 around promoters [36]; (ii) various HAT complexes (such as p300/CBP, NuA3, NuA4) would then recognize this H3K4me3 indirectly or directly [37] and acetylate the chromatin; (iii) in parallel or subsequently other HATs,

such as GCN5/PCAF-containing SAGA and ATAC complexes, may recognize the H3K4me3 by the Tudor domain of their SGF29 subunit [38,39]. This cross-talk between the H3K4me3 and H3 acetylation is supported by the observations indicating that knockdown of WDR5 (subunit of several chromatin modifying complexes including HMTs and HATs) or Set1 (catalytic subunit) of H3K4 methyltransferases complex not only decrease the H3K4me2/3, but they also decrease the acetylation levels at given promoters [40-42]. Thus, H3K4me3 seems to provide a binding platform for HATs, which are specific for the H3 tails, but may not be specific for the given lysine residue. This in turn would explain the fact that we

**A**



**B**



**Figure 8 A subset of inactive promoters having high H3K14ac is poised for stimuli dependent activation.** (A) Average ChIP-seq profile of 500 inactive genes around the transcription starts sites (-/+ 5 kb) shows specific enrichment of H3K14ac as opposed to H3K9ac. (B) Gene ontology (using David Bioinformatics, <http://david.abcc.ncifcrf.gov/>) analysis of 500 inactive genes having high H3K14ac suggest that these genes are activated in a stimuli dependent manner. The gene ontology term is on the y axis, and the negative log of p value indicating significance of enrichment is on the x axis.

see the acetylation of various H3 residues together (such as H3K9, H3K14 and H3K27) at the 5' end of the genes, which follow the presence of H3K4me3 and the underlying CpG islands (see below). This would suggest that some marks (i.e. H3K9ac and H3K14ac) are establishing more the openness of the chromatin, while others (i.e. H3K4me3) may serve more as a docking site [43,44]. Thus, all these marks together may be required to regulate Pol II transcription initiation positively.

#### **The presence of H3K9ac and H3K14ac at promoters correlates with their CpG content**

Approximately 70% of all the annotated gene promoters are associated with a CpG islands [27], which have distinct patterns of chromatin configuration. Interestingly, in mES cells CXXC finger protein 1 (Cfp1), a component of Setd1 histone methyltransferases complex has a preference for CpG clusters and can be recruited to artificial promoterless CpG clusters that subsequently lead to trimethylation of H3K4, a CpG island promoter signature [45,46]. The presence of H3K4me3 might be followed by a cascade of events and consequent acetylation of different H3 and/or H4 residues (as discussed above). Acetylation (H3K9 and H3K14) of the nucleosomes associated with the CpG islands might be acting as a physical barrier to inhibit the H3K9me3 of the chromatin [47]. Housekeeping gene promoters have generally high CpG content and are often protected from DNA methylation and methylation of H3K9, enabling constitutive expression of the associated genes. Thus, our findings which show a very good correlation between CpG content of promoter regions and acetylation of the associated nucleosomes on H3K9 and H3K14, along with H3K4me3 are in a good agreement with the model in which all these "active" histone marks influence the local chromatin structure to simplify the regulation of gene activity (reviewed in [27]). Our results also suggest the dependence of histone modifications such as H3K4me3, H3K9ac and H3K14ac on the DNA sequences underneath.

#### **H3K9 and H3K14 acetylation label active enhancers together with H3K27 acetylation**

Histone modification profiles have been used for identifying enhancer elements [4]. Enhancers are characterized by high H3K4me1/H3K4me3 ratio, open chromatin, low Pol II and presence of HAT co-activators such as p300 [9] and ATAC [10]. Interestingly, chromatin modification patterns at enhancers are much more variable and cell type specific, than chromatin patterns at promoters that are much more conserved [9,10]. Our study establishes that H3K9ac and H3K14ac also mark enhancers. Furthermore, like H3K27ac [11], these two marks can differentiate between the active and inactive/poised enhancers (Figure 4). RNA Pol II from these active enhancers produces bidirectional short (<2 kb)

non-coding eRNAs, whose expression level correlate with the proximal gene activity. [30]. Thus, along with H3K27ac, H3K9ac and H3K14ac can be used to discriminate the active enhancers from the inactive/poised enhancers at global and gene specific level.

#### **Genome-wide bivalent promoters are also marked by H3K9 and H3K14 acetylation**

Here we also show that bivalent promoters are not only marked by the H3K4me3 and H3K27me3, but also by the other active histone marks, such as H3K9ac and H3K14ac. There are two possibilities for the presence of H3K9ac and H3K14ac over the bivalent promoters. First, both acetylation marks may be required in addition to H3K4me3 for preparing these promoters of developmental genes for rapid induction at later stages. Second, the presence of H3K9ac and H3K14ac over bivalent promoters might act as a physical barrier and thereby inhibiting the trimethylation of H3K9, which eventually can lead to a stable silencing of developmentally regulated genes [48]. Thus acetylation of H3K9 and H3K14 may be required as a balancing mechanism to inhibit the permanent transcriptional silencing of the developmentally regulated genes [47]. Our genome-wide study showing presence of H3K9ac at bivalent promoters is in accordance with earlier studies describing the presence of H3K9ac, along with the active (H3K4me3) and repressive (H3K27me3) chromatin marks, over bivalent promoters in undifferentiated ES cells [49]. Note however that Hezroni et al. [50] observed only very low H3K9ac levels over bivalent gene promoters in ES cells. The difference observed between H3K9ac enrichment over bivalent promoters in our data and that of Hezroni et al. [50] could be attributed to the fact that our ChIP-seq data shows an increased depth in the coverage of H3K9ac sites compared to the one obtained previously. Indeed, we obtained 23 million reads for our H3K9ac ChIP seq data versus 14 million reads in the former study. Higher sequencing depth will result in a more complete coverage of the identified H3K9ac signals and also identify smaller peaks. Thus, it is possible that we detect more peaks that would allow us a more thorough analysis. Alternatively, differential enrichment of H3K9ac over bivalent promoters in our study and that of Hezroni et al. [50] is also possible because of the different antibodies used in the two studies, which can have distinct specificities.

#### **A subset of inactive promoters are specifically enriched for H3K14ac**

H3K9ac and H3K14ac are associated with active promoters and are considered as hallmarks of active transcription. However, we observed that in contrast to H3K9ac, H3K14ac is selectively and specifically associated with

repressive marks such as H3K9me3 and H3K27me3 on a subset of inactive promoters. Presence of H3K14ac on inactive promoters along with H3K9me3 and H3K27me3 is in agreement with previous observation where high throughput characterization of combinatorial histone marks using proteomic approach shows highest abundance for H3K14ac peptide with activation and repressive marks [51]. Our data suggest that HATs may bind transiently to a subset of inactive promoters and acetylated H3K14, which is then removed by the transient binding of HDACs. Inhibition of HDACs leads to substantial increase in H3K14ac of these inactive promoters, suggesting that these promoters are subjected to dynamic regulation by HATs and HDACs. Transient acetylation and deacetylation of these promoters may keep these promoters unexpressed but at the same time primed for future activation upon receiving the external stimuli. It is thus possible that the presence of H3K14ac over a subset of inactive promoters may serve as a landing platform for further HATs, which have bromodomains, such as GCN5/PCAF, and propagate the acetylation of the given loci at a later stage. Thus, presence of H3K14ac and the consequent coordinated action of HATs and HDACs over these inactive genes might poise them for stimuli dependent future activation.

## Conclusions

In this study, we characterized the H3K9 and H3K14 acetylations over various genomic regions in mES cells. Taken together our data suggest that genome-wide coordinated modifications of histone acetylation (i.e. H3K9, H3K14 and H3K27) and methylation (H3K4me3) provide a general signature for facilitating gene expression over active promoters. On the other hand we showed differential enrichment of H3K14ac over inactive promoters. Furthermore our analyses suggest correlation between acetylation of nucleosomes over the promoters and the underlying sequences and demonstrate the presence of H3K9ac and H3K14ac over enhancers, which can be used in the future studies to discriminate active enhancers from inactive/poised ones. In conclusion, our comprehensive study of H3K9ac and H3K14ac demonstrates a wider role for these to marks in gene regulation than originally thought.

## Methods

### Cell culture

Wild type embryonic stem cells were derived from blastocysts (3.5 PC) and cultivated on feeder cells (37°C, 5% CO<sub>2</sub>) in DMEM (4.5 gm/lit), 15% FCS, leukemia inhibiting factor, penicillin/streptomycin, L-glutamine, and non-essential amino acids. At least three passages under feeder free conditions on 0.1% gelatin were used to exclude feeders.

### Antibodies

Mouse monoclonal anti-H3K14ac antibody (13HH3-1A5) was produced against the peptide "STGGK(ac)APRKC". The characterization of the antibody is shown in Additional file 2: Figure S2. The H3K9ac antibody is from Abcam -ab4441.

### Sodium butyrate treatment of mES cells

mES cells were treated with 5 mg/ml sodium butyrate for 0 min, 5 min, 10 min, 30 min, 1 hr, 2 hr, 4 hr, 8 hr and 12 hr. Cells were cross linked with 1% formaldehyde and ChIP was performed as described below. Level of Oct4 was monitored by Western blot to see the pluripotent state of the ES cells at various time points after sodium butyrate treatment (Additional file 7: Figure S6). Similar level of Oct4 at all time points suggest that the time points we have used for sodium butyrate treatment have no effect on pluripotency of the ES cells. Tubulin was used as a loading control. Anti-Oct4 antibody (611202) used in the Western blot is from BD labs.

### Chromatin Immunoprecipitation (ChIP) and sequential ChIP

All ChIP experiments were carried out on  $2 \times 10^7$  cells per antibody. Cells were cross-linked with 1% formaldehyde, lyses and syndicated in sonication buffer (10 mM Tris-HCl pH 7.5, 200 mM NaCl, 1% SDS, 4% NP-40, 1 mM PMSF) to obtain an average chromatin size of 200-500 bp. Chromatin was pre-cleared using 50  $\mu$ l of a 50% protein A sepharose (GE healthcare) slurry for 1 h at 4°C with gentle inverting. Immunoprecipitation were carried out in 10 ml of IP buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton-X 100). 15  $\mu$ l of ZZZ3 and SPT20 serum antibodies were used. Input chromatin was obtained after pre-clearing, by de-cross linking and purifying input DNA using a Qiaquick column (Qiagen) according to manufacturer's instructions. Immunoprecipitation were carried out with inverting at 4°C for 14-16 h. The samples were then incubated with 50  $\mu$ l of a 50% Protein A sepharose slurry for 3 h at 4°C with gentle inverting. IP samples were reverse-cross linked and the DNA was purified using a Qiaquick column (Qiagen). Q-PCR using SYBR green was used to validate known target sites before and after sequencing. For sequential ChIP, at least four ChIP assays of  $2 \times 10^7$  cells were used for the first IP (H3K9ac). Following standard washing, elution was performed with 10 mM DTT (30 min, 37°C). The eluates from four ChIPs were combined, diluted at least 30 times with ChIP dilution buffer and secondary antibody (H3K14ac) was incubated overnight. The subsequent steps were performed as for regular ChIPs. Validation of ChIP-seq was performed by the assessment of the individual enrichment over the control genomic region by

ChIP-qPCR in triplicate with primers specific for these regions using SYBR Green master mix (Roche).

### Chromatin Immunoprecipitation – High throughput-sequencing (ChIP-seq)

#### *Library preparation and sequencing*

We followed the manufacturer's (Solexa) protocol for creating genomic DNA libraries as previously described in [52]. ChIP DNA and input DNA were first band-isolated on a 2% agarose to obtain fragments between 150 and 350 base pairs and DNA was extracted using the QIAquick gel extraction kit (Qiagen) and eluted in 34  $\mu$ L. After end-repair and addition of a single adenosine ("A") nucleotide, adapters were ligated to samples for 15 min at room temperature in the following fashion: the samples eluted from the MinElute column in 10  $\mu$ L were ligated to 1  $\mu$ L of adapters using 1.3  $\mu$ L of LigaFast T4 DNA Ligase (3 Units/ $\mu$ L; Promega) and 12.3  $\mu$ L Rapid Ligation Buffer (Promega). For DNA libraries, the Illumina genomic DNA adapters were diluted 1:10. After 15 min, samples were purified with the MinElute PCR purification kit (Qiagen).

Adapters in excess were eliminated by using gel purification on a 2% agarose E-Gel (Invitrogen) for 20 min, together with Track-It 50 bp DNA ladder (Invitrogen). DNA fragments ranging from 150 base pairs to 500 base pairs were extracted and recovered in 28  $\mu$ L EB with a QIAquick gel extraction kit (Qiagen). To amplify the library, PCR was performed using Illumina genomic DNA primer "1.1" and Illumina genomic DNA primer "2.1" with 15 cycles (Input DNA) or 17 cycles (ChIP DNA) of amplification. A final size selection was performed using a 2% agarose E-Gel to obtain a library with a median length of ~230 bp which is within the recommended size range for cluster generation on Illumina's flow cell. The library was recovered in 20  $\mu$ L EB using MinElute Gel Extraction kit (Qiagen). Finally, DNA concentrations and purities (A260/280 nm ratios) were measured on a Nanodrop spectrophotometer. Sequencing was carried out on the Illumina (Solexa) platform at a sequencing depth of 1 lane averaging 10 million reads, read length of 27 + bp, single-end reads and mapped to mouse genome build (mm9).

#### *Establishment of list of reference loci*

ChIP-seq data were mapped using the ELAND software (Illumina) allowing one mismatch. Mapped read data were used as an input and to establish list of loci using MACS (Using default parameters except: mfold 12; tag size according to platform; band width 100) [23]. Input DNA file was used as a control in all peak detection analyses. To calculate a single enrichment value for a binding site, tag density is defined as the number of tags present or overlapping in user-defined

window around the reference site. The algorithm for the signal enrichment calculation is described in details in [33].

#### *Average profile calculations*

We extracted the tag density in a 5 kb window surrounding the TSSs and gene body using the program seqMINER which generates heatmap as well as the profiles [33]. The sequenced ChIP-seq reads represent only the end of each immunoprecipitated fragments instead of the precise protein-DNA binding sites. To illustrate the entire DNA fragment, basically before analysis, 3' end of each ChIP-seq read was extended to 200 bp in the direction of the reads. For average gene profiles, genes (+/-5000 bp from binding site) were divided in 100 bins of length relative to the gene length. Moreover 10 equally sized (50 bp) bins were created on the 5' and 3' of the gene and ChIP-seq densities were collected for each dataset in each bin. The average of enrichments over input in each bin was plotted for each dataset. For average TSS profile, the genes as above were aligned on their TSSs and densities were collected in 100 equally sized bins around the TSS (-/+5000 bp).

#### *Identification of intergenic regions*

The list of refseq genes was obtained from UCSC table browser [28]. Co-ordinates for the intergenic regions were extracted from refseq genes using complement intervals of a dataset tools from Galaxy web server (<https://main.g2.bx.psu.edu/>). All those regions, which do not correspond to the beforehand defined promoter regions, that are not gene bodies, and are not regions that are situated 2 kb downstream of the end of annotated genes, are considered as intergenic regions. Overlapping piece of intervals (from Galaxy web tools) with an overlap of at least 500 bps was used to distinguish peaks, which are intergenic for H3K9 and H3K14 acetylation.

#### *Data source*

ChIP-seq datasets were downloaded from the public data bank Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/gds>) under the accession number: GSM307618 (H3K4me3-mES); GSM307619 (H3K27me3-mES); GSM307621 (H3K9me3-mES); GSM594577 (H3K4me1-mES); GSM594578 (H3K27ac-mES); GSM307623 (Pol II-mES); GSM307624 (total H3-mES); GSM307625 (input-mES); GSM699164 (p300-mES). The sequencing data we have generated is deposited in the GEO database under the accession number GSM775313 (H3K9ac-mES) and GSM775314 (H3K14ac-13HH3-1A5-mES). CpG island definition and locations of known genes were downloaded from UCSC genome browser [28]. Number of mapped reads for each data set is provided in the Additional file 9: Table S2.

Link to allow review of record:

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=nhclxooaewkmaq&acc=GSE31284>.

### Creation of density files for genome browser data visualization

Raw BED files are used as input for ad hoc (WIG) density file creation script as described in [33]. Reads are directionally extended of their theoretical length (200 bp), and 25 bp bins are created. In each bin, the maximal number of overlapping reads is computed. Tracks were uploaded and displayed using fixed scale representation in the UCSC genome browser [28].

### Data analysis

The scatter plots, k-means clustering and average gene profiles are created using seqMiner [33]. Box plots and correlation analysis are produced using R software (<http://r-project.org/>). The distribution of H3K9ac and H3K14ac in various genomic regions is calculated using the CEAS software [22]. Briefly, CEAS estimates the relative enrichment level of ChIP regions in each gene feature with respect to the whole genome. In our analysis promoters correspond to the 2 kb upstream regions from the transcription start site (TSS) of refseq genes. All those regions, which do not correspond to the beforehand defined promoter regions, that are not gene bodies, and are not regions that are situated 2 kb downstream of the end of annotated genes, are considered as intergenic regions (more details are available at <http://liulab.dfci.harvard.edu/CEAS/usermanual.html> and [22]). Gene ontology was performed using David Bioinformatics tools (<http://david.abcc.ncifcrf.gov/home.jsp>).

### Additional files

**Additional file 1: Figure S1.** The commercially available anti-H3K14ac antibody (Upstate (07-353)) cross-reacts with other peptides. Enzyme linked immunosorbent assay (ELISA) using various peptides such as H3K14 acetylated (H3K14ac) and non-acetylated (H3K14) and non specific histone H4 acetylated peptide (H4K5ac-K12ac) suggest that anti-H3K14ac antibody from Upstate (07-353) cross-reacts with non-acetylated H3K14 peptide as well as with H4K5andK12ac peptide.

**Additional file 2: Figure S2.** Characterization of mouse monoclonal anti-H3K14ac (13HH3-1A5) antibody. (A) Western blot analysis using the 13HH3-1A5 (anti H3K14ac antibody) on the recombinant E. coli expressed histone H3 and histones extracted by and acidic extraction protocol from human HeLa cells (right panel). Coomassie blue stained SDS-PAGE of the proteins used for the western blot analysis in a 100-fold dilution. (B) Enzyme linked immunosorbent assay (ELISA) using various peptides such as H3K14 acetylated (H3K14ac) and non-acetylated (H3K14), H3K14 dimethylated, H3S10 phosphorylated and K14 dimethylated (H3pS10K14dimethyl), H3S10 phosphorylated (H3pS10), H3K9 dimethylated (H3K9dimethyl), H3S10 phosphorylated and K9 dimethylated (H3pS10K9dimethyl) and histone H4 acetylated peptide (H4K5, 8, 12 and 16 ac).

**Additional file 3: Figure S3.** Validation of (A) H3K9ac and (B) H3K14ac ChIP-seq in mouse embryonic stem cells by ChIP-qPCR. Peaks of local

enrichment of H3K9 and H3K14 acetylation were determined after sequence alignment and normalization to input DNA. Tag density per peak (shown on the right of the graph) is plotted as fold enrichment over an arbitrarily chosen control genomic region. The list of primer is provided as Supplementary Table S1. These ChIP-qPCR results confirmed the specificity of the predicted peaks for H3K9 and H3K14 acetylation.

**Additional file 4: Table S1.** List of primers used for ChIP-qPCR validation.

**Additional file 5: Figure S4.** Enrichment of H3K9ac and H3K14ac over various genomic regions. Average enrichment profile of H3K9 and H3K14ac over the promoters, coding exons, introns and distal intergenic regions. Enrichment over distal intergenic and intronic sites is comparable to promoters.

**Additional file 6: Figure S5.** Correlation of H3K9ac (A) and H3K14ac (B) intergenic peaks with H3K4me1, H3K27ac, Pol II and p300. Heatmaps of the signal density using k-means clustering over H3K9ac and H3K14ac distal intergenic sites (-/+ 5 kb). Strong correlation with H3K4me1 and H3K27ac and presence of Pol II and p300 over H3K9ac and H3K14ac intergenic sites suggest that they may act as enhancers.

**Additional file 7: Figure S6.** Measurement of Oct4 level at various time points after the treatment with 5 mg/ml sodium butyrate in ES cells. Level of Oct4 was monitored by Western blot to see the pluripotent state of the ES cells at various time points after sodium butyrate treatment. The similar levels of Oct4 observed at different experimental time points suggest that the time points we have used for sodium butyrate treatment have no effect on pluripotency of the ES cells. Tubulin was used as a loading control.

**Additional file 8: Figure S7.** Ratio of tag density of least expressed to highest expressed gene promoters for H3K9 and H3K14 acetylation. H3K9 and H3K14 acetylation tag density were calculated over 500 least and highest expressed gene promoters based upon their RNA-sequencing expression profile. The ratio of tag density of least expressed to highest expressed plotted for H3K9 and H3K14 acetylation, which shows that H3K14ac is specifically enriched at inactive promoters as compared to H3K9ac.

**Additional file 9: Table S2.** Number of mapped reads for the data set used in this study.

### Abbreviations

mES: Mouse embryonic stem cells; HAT: Histone acetyl transferase; HDAC: Histone deacetylases; ChIP: Chromatin immunoprecipitation; ChIP-seq: Chromatin immunoprecipitation coupled high throughput sequencing; mAbs: Monoclonal antibodies; TSS: Transcription start sites; CpG: Cytosine-phosphate diester-guanine; PHD: Plant homeodomain; eRNA: Enhancer RNA; Pol II: RNA Polymerase II.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

KK generated ChIP-seq data and performed experiments. KK, ARK and LT analyzed data. MOA and HK generated antibodies. KK and LT wrote the manuscript. LT planned, coordinated and supervised the project. All authors read and approved the manuscript.

### Acknowledgements

The authors are very grateful to the labs of B. E. Bernstein, R. Jaenisch and D. Schübeler from which released data were used in the analyses and to D. Schübeler lab for providing unpublished RNA-seq data. The authors are grateful to T. Ye, and S. Le Gras for bioinformatics support; the IGBMC NGS platform for data generation; P. Eberling for generation of peptides and the IGBMC cell culture service. We thank D. Devys for critical reading of the manuscript. ARK was a recipient of a fellowship from INSERM-Region Alsace and the Association pour la Recherche sur le Cancer. This work was funded by grants from ANR (GenomATAC; ANR-09-BLAN-0266), the EU (EUTRACC and EPIDIACAN) and CNRS (LEA-SkinChroma). None of the authors have a financial interest related to this work.

#### Author details

<sup>1</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR 7104, INSERM U 964, Université de Strasbourg, BP 10142-67404 ILLKIRCH Cedex, CU de Strasbourg, France. <sup>2</sup>Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan. <sup>3</sup>Present address: Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland.

Received: 30 March 2012 Accepted: 22 August 2012

Published: 24 August 2012

#### References

1. Hochdinger K, Jaenisch R: Nuclear reprogramming and pluripotency. *Nature* 2006, **441**(7097):1061–1067.
2. Mattout A, Meshorer E: Chromatin plasticity and genome organization in pluripotent embryonic stem cells. *Curr Opin Cell Biol* 2010, **22**(3):334–341.
3. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, et al: Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 2008, **40**(7):897–903.
4. Heintzman ND, Ren B: Finding distal regulatory elements in the human genome. *Curr Opin Genet Dev* 2009, **19**(6):541–549.
5. Li B, Carey M, Workman JL: The role of chromatin during transcription. *Cell* 2007, **128**(4):707–719.
6. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007, **448**(7153):553–560.
7. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al: A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006, **125**(2):315–326.
8. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D: Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 2008, **30**(6):755–766.
9. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, et al: Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 2009, **459**(7243):108–112.
10. Krebs A, Karmodiya K, Lindahl-Allen M, Struhl K, Tora L: SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol Cell* 2011, **44**(3):410–423.
11. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al: Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 2010, **107**(50):21931–21936.
12. Rada-Iglesias A, Bajpai R, Swigut T, Bruggmann SA, Flynn RA, Wysocka J: A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 2011, **470**(7333):279–283.
13. Anamika K, Krebs AR, Thompson J, Poch O, Devys D, Tora L: Lessons from genome-wide studies: an integrated definition of the coactivator function of histone acetyl transferases. *Epigenetics Chromatin* 2010, **3**(1):18.
14. Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K: Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* 2009, **138**(5):1019–1031.
15. Hezroni H, Tzchori I, Davidi A, Mattout A, Biran A, Nissim-Rafinia M, Westphal H, Meshorer E: H3K9 histone acetylation predicts pluripotency and reprogramming capacity of ES cells. *Nucleus* 2011, **2**(4):300–309.
16. Krejčí J, Uhlířová R, Galiová G, Kozubek S, Smigová J, Bártová E: Genome-wide reduction in H3K9 acetylation during human embryonic stem cell differentiation. *J Cell Physiol* 2009, **219**(3):677–687.
17. Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, Wang C, Brindle PK, Dent SY, Ge K: Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27 ac in nuclear receptor transactivation. *EMBO J* 2011, **30**(2):249–262.
18. Nagy Z, Tora L: Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* 2007, **26**(37):5341–5357.
19. Lee KK, Workman JL: Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* 2007, **8**(4):284–295.
20. Fuchs SM, Krajewski K, Baker RW, Miller VL, Strahl BD: Influence of combinatorial histone modifications on antibody and effector protein recognition. *Curr Biol* 2011, **21**(1):53–58.
21. Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwiercz P, Alekseyenko AA, Cheung MS, Day DS, Gadel S, Gorchakov AA, et al: An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* 2011, **18**(1):91–93.
22. Shin H, Liu T, Manrai AK, Liu XS: CEAS: cis-regulatory element annotation system. *Bioinformatics* 2009, **25**(19):2605–2606.
23. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al: Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008, **9**(9):R137.
24. Efroni S, Duttgupta R, Cheng J, Dehghani H, Hoepfner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH, Gingeras TR, Misteli T, Meshorer E: Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* 2008, **2**(5):437–447.
25. Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, Lin MF, et al: Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* 2010, **330**(6012):1787–1797.
26. Boeger H, Griesenbeck J, Strattan JS, Kornberg RD: Nucleosomes unfold completely at a transcriptionally active promoter. *Mol Cell* 2003, **11**(6):1587–1598.
27. Deaton AM, Bird A: CpG islands and the regulation of transcription. *Genes Dev* 2011, **25**(10):1010–1022.
28. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D: The human genome browser at UCSC. *Genome Res* 2002, **12**(6):996–1006.
29. Antequera F, Bird A: Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A* 1993, **90**(24):11995–11999.
30. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, et al: Widespread transcription at neuronal activity-regulated enhancers. *Nature* 2010, **465**(7295):182–187.
31. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G: A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol* 2010, **8**(5):e1000384.
32. Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, Qiu J, Liu W, Kaikkonen MU, Ohgi KA, et al: Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* 2011, **474**(7351):390–394.
33. Ye T, Krebs AR, Choukrallah MA, Keime C, Plewniak F, Davidson I, Tora L: seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* 2011, **39**(6):e35.
34. Herz HM, Nakanishi S, Shilatifard A: The curious case of bivalent marks. *Dev Cell* 2009, **17**(3):301–303.
35. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al: Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* 2010, **471**(7339):480–485.
36. Shilatifard A: Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* 2006, **75**:243–269.
37. Zhang Y: It takes a PHD to interpret histone methylation. *Nat Struct Mol Biol* 2006, **13**(7):572–574.
38. Vermeulen M, Eberl HC, Matarese F, Marks H, Denisov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, et al: Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 2010, **142**(6):967–980.
39. Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Baryste D, Li J, Wu M, Zhou BO, et al: Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J* 2011, **30**(14):2829–2842.
40. Crump NT, Hazzalin CA, Bowers EM, Alani RM, Cole PA, Mahadevan LC: Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP. *Proc Natl Acad Sci U S A* 2011, **108**(19):7814–7819.
41. Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD: WDR5 associates with histone H3 methylated at

- K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 2005, **121**(6):859–872.
42. Li X, Wang S, Li Y, Deng C, Steiner LA, Xiao H, Wu C, Bungert J, Gallagher PG, Felsenfeld G, *et al*: **Chromatin boundaries require functional collaboration between the hSET1 and NURF complexes.** *Blood* 2011, **118**(5):1386–1394.
  43. Roh TY, Cuddapah S, Cui K, Zhao K: **The genomic landscape of histone modifications in human T cells.** *Proc Natl Acad Sci U S A* 2006, **103**(43):15782–15787.
  44. Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT: **Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4.** *Cell* 2007, **131**(1):58–69.
  45. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD, *et al*: **CpG islands influence chromatin structure via the CpG-binding protein Cfp1.** *Nature* 2010, **464**(7291):1082–1086.
  46. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, *et al*: **DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.** *Nature* 2007, **448**(7154):714–717.
  47. Wang H, Cao R, Xia L, Erdjument-Bromage H, Borchers C, Tempst P, Zhang Y: **Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase.** *Mol Cell* 2001, **8**(6):1207–1217.
  48. Margueron R, Trojer P, Reinberg D: **The key to development: interpreting the histone code?** *Curr Opin Genet Dev* 2005, **15**(2):163–176.
  49. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, *et al*: **Chromatin signatures of pluripotent cell lines.** *Nat Cell Biol* 2006, **8**(5):532–538.
  50. Hezroni H, Sailaja BS, Meshorer E: **Pluripotency-related, valproic acid (VPA)-induced genome-wide histone H3 lysine 9 (H3K9) acetylation patterns in embryonic stem cells.** *J Biol Chem* 2011, **286**(41):35977–35988.
  51. Young NL, DiMaggio PA, Plazas-Mayorca MD, Baliban RC, Floudas CA, Garcia BA: **High throughput characterization of combinatorial histone codes.** *Mol Cell Proteomics* 2009, **18**(10):2266–2284.
  52. Auerbach RK, Euskirchen G, Rozowsky J, Lamarre-Vincent N, Moqtaderi Z, Lefrancois P, Struhl K, Gerstein M, Snyder M: **Mapping accessible chromatin regions using Sono-Seq.** *Proc Natl Acad Sci U S A* 2009, **106**(35):14926–14931.

doi:10.1186/1471-2164-13-424

**Cite this article as:** Karmodiya *et al*: H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 2012 **13**:424.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

