Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription

Bing Li,¹ Madelaine Gogol,¹ Mike Carey,^{1,2} Samantha G. Pattenden,¹ Chris Seidel,¹ and Jerry L. Workman^{1,3}

¹Stowers Institute for Medical Research, Kansas City, Missouri 64110, USA; ²Department of Biological Chemistry, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California 90095, USA

The presence of Set2-mediated methylation of H3K36 (K36me) correlates with transcription frequency throughout the yeast genome. K36me targets the Rpd3S complex to deacetylate transcribed regions and suppress cryptic transcription initiation at certain genes. Here, using a genome-wide approach, we report that the Set2–Rpd3S pathway is generally required for controlling acetylation at coding regions. When using acetylation as a functional readout for this pathway, we discovered that longer genes and, surprisingly, genes transcribed at lower frequency exhibit a stronger dependency. Moreover, a systematic screen using high-resolution tiling microarrays allowed us to identify a group of genes that rely on Set2–Rpd3S to suppress spurious transcripts. Interestingly, most of these genes are within the group that depend on the same pathway to maintain a hypoacetylated state at coding regions. These data highlight the importance of using the functional readout of histone codes to define the roles of specific pathways.

[**Keywords:** Histone methylation; Rpd3; Set2; chromatin; cryptic transcripts; deacetylation]

Supplemental material is available at <http://www.genesdev.org>.

Received February 7, 2007; revised version accepted April 13, 2007.

Histones are subject to a wide variety of post-translational modifications (PTMs). The pattern of PTMs is an important indicator of the architecture and potential function of certain chromatin domains (Turner 2002). PTMs such as acetylation and ubiquitination may dramatically alter the biophysical properties of histones, thus exerting a direct influence on chromatin structure (Jason et al. 2002; Shogren-Knaak et al. 2006). Other PTMs such as methylation, perhaps due to their small size or lack of altering charge, likely function as a signal in certain pathways (Jenuwein and Allis 2001). In some cases, PTMs work through both of these mechanisms.

Histone modifications have been generally categorized based on their physical presence, with active marks referring to those at highly transcribed regions and repressive marks denoting those at silenced loci. This correlative definition, however, has very little implication for the direct function of these histone marks in biological processes. For instance, histone H3K9 methylation had long been thought to be a silencing mark due to its ability to recruit heterochromatin protein 1 (HP1) (Banister et al. 2001; Lachner et al. 2001). Nevertheless,

HP1/H3K9me interaction may also play a positive role in transcription regulation because H3K9me has been located at actively transcribed genes (for review, see Hediger and Gasser 2006). Therefore, defining the purpose of particular histone codes based merely on their physical presence is highly risky without rigorous functional investigation.

Two of the best characterized histone transcriptional marks are methylation of histone H3 at Lys 4 (K4me) and at Lys 36 (K36me). Both marks are enriched at coding regions of actively transcribed genes in all species examined and involve mono-, di-, and trimethylated states (Shilatifard 2006). A genome-wide study demonstrated that the occupancy of K4me3 and K36me3 correlates with transcription frequency of underlying genes (Pokholok et al. 2005). In addition, the histone methyltransferases Set1 and Set2 are responsible for all three states of K4me and K36me in yeast. Both Set1 (COMPASS) and Set2 have been shown to directly associate with the elongating forms of RNA Polymerase II (Pol II) (for review, see Hampsey and Reinberg 2003). These observations led to the view that K4me and K36me function in transcriptional activation. Deletion of *SET1* or *SET2* and mutations at histone K4 or K36 do not significantly affect transcription of the majority of yeast genome; however, both methylation states are dependent on active transcription (Li et al. 2007a). These

³Corresponding author.

E-MAIL jlw@stowers-institute.org; FAX (816) 926-4692.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1539307>.

observations imply that methylation is not required to facilitate the initial round of transcription. A recent finding that K4 methylation does not influence transcription elongation per se in a well-defined biochemical system (Pavri et al. 2006) supports this notion and suggests that methylation is not required for Pol II to transcribe through the nucleosomal barrier.

The identification of various protein domains that recognize K4me and K36me provided a clue to the importance of these methylation marks in vivo. The chromodomain within the SAGA histone acetyltransferase (Pray-Grant et al. 2005) and the hCHD1 chromatin remodeling factor (Sims et al. 2005), the PHD domain of NURF (for review, see Zhang 2006), and yeast NuA3 (Taverna et al. 2006) can specifically bind to the methylated H3K4 residue and potentially target these factors to regions that are rich in H3K4 methylation for transcription activation. The details of this targeting mechanism are poorly defined. A vital role for histone methylation marks in targeting chromatin-modifying complexes in vivo was recently discovered. Methylated K36 is recognized by the chromodomain of Eaf3 in the Rpd3S histone deacetylase complex, which leads to recruitment of Rpd3S to the coding regions of transcribed genes. Histone deacetylation by locally recruited Rpd3S is an important event that is key to the suppression of spurious internal transcription initiation (Carrozza et al. 2005; Joshi and Struhl 2005; Keogh et al. 2005). Although this mechanism has been demonstrated at a few individual genes, several key questions about the role of Set2 signaling and its mechanism of action remain unanswered. Does K36me regulate ORF acetylation genome wide? Does K36me control cryptic transcription events across the genome? Which classes of genes are dependent on Set2 to control ORF acetylation? Which classes of genes rely on Set2 to prevent internal initiation? Is there any overlap between these two groups?

Here we provide genome-wide evidence that Set2-mediated K36me is required for the control of global acetylation at coding regions. Surprisingly, using acetylation changes in Δ *set2* as a functional readout, we discovered that genes that are longer and transcribed less often display a stronger dependence on the Set2–Rpd3S pathway to suppress acetylation of their coding regions. This trend is consistent with our high-resolution transcription profiling experiment, illustrating that the occurrence of cryptic transcripts correlates more closely with genes on which histone acetylation levels increase in Set2 mutants than with genes with high H3K36me3 levels in wild-type cells.

Results

The Set2–Rpd3S pathway controls global histone acetylation at ORFs

At many individual loci, defects in Set2-mediated methylation and Rpd3S complex function lead to an increase in the level of acetylation within the coding region (Reid et al. 2004; Carrozza et al. 2005; Joshi and Struhl 2005;

Keogh et al. 2005). To examine this phenomenon on a genome-wide level, we performed chromatin immunoprecipitation coupled with microarray analysis (ChIP–chip) experiments using an antibody against acetylated histone H4 (AcH4). Immunoprecipitated DNA from both wild-type and Δ *set2* strains were labeled with fluorescent dye and competitively hybridized with the respective input DNA on high-density tiling microarrays. Consistent with previous results (Pokholok et al. 2005), we found that enrichment for AcH4 in wild-type cells peaked at promoter regions (data not shown). When the level of AcH4 in Δ *set2* and wild-type strains was directly compared, we found that coding regions also frequently had increased levels of acetylation (data not shown) in Δ *set2*. A statistical assessment of changes in the acetylation pattern was performed using a modified average gene analysis that was originally developed by the Young laboratory (see Materials and Methods). The average distribution profile of AcH4 changes between Δ *set2* and wild-type strains revealed a peak at the 3' portion of gene coding regions (Li et al. 2007b), resembling that of K36 trimethylation in wild-type cells (Pokholok et al. 2005). Thus, reduction of K36 trimethylation marks led to a corresponding increase in H4 acetylation marks. When we normalized the AcH4 level according to nucleosome density (histone H4) before comparing the values obtained from wild-type and Δ *set2* strains, we obtained almost identical results (data not shown). This observation is consistent with the previous finding that deletion of *SET2* does not alter histone density genome wide (Rao et al. 2005).

To further analyze the trend of acetylation changes following deletion of *SET2*, we performed a class-average analysis in which only genes within certain categories were averaged and the resulting average distribution of all classes were plotted on the same graph (Fig. 1A). Initially, we grouped the genes based on their ORF length. As shown in Figure 1A, deletion of *SET2* led to a more dramatic increase in acetylation at genes with longer ORFs, suggesting that Set2 dependence was proportional to gene length. This trend was reminiscent of the positive correlation observed between ORF length and the presence of K36me2 (Rao et al. 2005) and K36me3 (Fig. 1B). We next subjected the matrix generated from the average gene analysis to a K-mean cluster analysis, which categorized genes according to similarity in distribution profiles. We found that two clusters (Fig. 1C, bottom two panels), which constituted ~25% of the total genome, displayed a significant increase of AcH4 following deletion of *SET2*. Interestingly, when we compared ORF length of this group of genes to that of the entire genome in a histogram analysis (Fig. 1D), we found that increase of AcH4 upon deletion of *SET2* tends to occur at longer genes. This result agreed with those of the class-average analysis discussed above (Fig. 1A).

The enrichment of K36me3 at coding regions has been shown to be proportional to the transcription frequency of underlying genes (Pokholok et al. 2005). K36me2, however, only correlates with the ON/OFF state of tran-

Figure 1. The Set2-Rpd3 pathway controls global histone acetylation at ORFs. (A) Longer genes depend on the Set2-Rpd3S pathway to maintain the acetylation status at ORFs. ChIP-chip was performed using high-resolution tiling arrays manufactured by Agilent Technologies. The log₂ ratio of acetylation of H4 (AcH4) in $\Delta set2$ over AcH4 in wild type were subjected to a modified average gene analysis originated by the Young laboratory (see Materials and Methods for details). All genes were divided into eight subclasses based on the length of their coding regions. The averages of each subclass were plotted. The number of genes in each class is indicated within the parentheses behind the gene categories in the white box. (B) K36 trimethylation is preferentially enriched at longer ORFs. The similar class-average analysis was used to show the distribution pattern of K36me3 on average genes based on Pokholok et al. (2005). (C) Identification of genes that require Set2 to suppress hyperacetylation at ORFs. The matrix generated from the average gene analysis that covers the ORF region was subjected to the K-mean cluster analysis using TM4 microarray suite software (Saeed et al. 2006). The bottom two clusters (~25% of genome) displayed significant AcH4 increase upon deletion of *SET2*. (D) The ORF length of the genes that are within the bottom two clusters from the above analysis (in C) is directly compared with that of the entire genome in a histogram analysis (Microsoft Excel).

scription, and not transcription frequency (Rao et al. 2005). To test the relationship of the acetylation changes following deletion of *SET2* with frequency of gene transcription, we used the class-average analysis described earlier, except that genes were categorized based on their transcription frequency in wild-type cells. Surprisingly, we found that genes transcribed less often (particularly for those that were less than five copies per hour) showed strong dependence on Set2 to control ORF acetylation (Fig. 2). Therefore, we concluded that although K36me3 is enriched at highly active genes, its functional role may be more important at genes with a low transcription frequency.

Rpd3S is required for controlling genome-wide acetylation at ORFs

Having established the role of Set2 in regulating global acetylation at ORFs, we asked if defects in the Rpd3S complex resulted in a similar genome-wide phenotype. In a parallel study, we learned that the two critical domains of Rpd3S—the chromodomain of Eaf3 (CHD_{Eaf3}) and the PHD domain of Rco1 (PHD_{Rco1})—were essential for full Rpd3S function (Li et al. 2007b). Therefore, we took advantage of two mutants, *eaf3 Δ chd* (deletion of CHD_{Eaf3}) and *rco1 Δ phd* (deletion of PHD_{Rco1}), and performed ChIP-chip experiments as described above. We

Figure 2. Infrequently transcribed genes are sensitive to disruption of the Set2–Rpd3S pathway. Average gene analysis was performed essentially as described in Figure 1, except that genes were divided into five subgroups based on their transcription rates (Holstege et al. 1998). The averages of each subclass are plotted.

discovered that both mutations result in acetylation changes strikingly similar to those noted for deletion of *SET2* (Fig. 3). This result supports the notion that Set2 and Rpd3S function primarily in the same pathway, and that this pathway, as a whole, exerts a stronger effect on longer and less frequently transcribed genes.

Genome-wide identification of genes that rely on the Set2–Rpd3S pathway to suppress cryptic internal initiation

We wanted to determine if the mechanism of internal initiation suppression by the Set2–Rpd3S pathway was a global phenomenon, so we conducted a genome-wide survey that directly compared RNA from wild-type and $\Delta set2$ strains on a high-resolution tiling array with multiple probes for each ORF region. We defined target genes to be genes that had similar transcript levels at the 5' end of the ORF (ratio of $\Delta set2$ over wild type = 1), but higher levels of mRNA at the 3' ORF in $\Delta set2$, since the presence of cryptic transcripts would increase the amount of RNA from one end of the gene. Genes that showed increased or decreased transcripts across the entire ORF region were not considered.

RNA for microarray analysis is commonly labeled using RT-based PCR amplification with an oligo-dT primer. These reactions, however, are inevitably biased toward regions close to the 3' end of the ORF due to the processivity nature of RT. To circumvent this potential problem and preserve the signals from longer transcripts, we adapted a direct, nonenzymatic labeling approach (Heetebrij et al. 2003). This method involved the use of a platinum-based compound to directly couple a fluorescent dye to RNA molecules. Initially, we examined the

transcripts corresponding to the 5' end of *STE11*. The microarray results showed no difference between $\Delta set2$ and wild type (Log2 ratio = 0). Signals from the middle of the gene to the 3' end of ORF in $\Delta set2$, however, were clearly increased (Fig. 4A). The start of each stepwise increase signal matched the internal start sites mapped by Northern blot for the *STE11* transcript. The results for the *STE11* message validate our microarray approach for detecting genes that have cryptic transcripts in $\Delta set2$ mutants.

We next sought to identify all genes exhibiting similar expression pattern changes upon deletion of *SET2*. First, we generated an expression matrix using the data points in the ORF regions through the average gene analysis method as described in the Materials and Methods. We then subjected this matrix to a K-mean cluster analysis. Two of the clusters that met the standard for target genes contained a total of 621 genes. An example of one such gene, *AZR1*, is shown in Figure 4B and confirmed by Northern blot (Fig. 4C). Thus, we concluded that Set2 was required for suppression of internal transcription initiation at many sites genome-wide.

To further extend our transcription profiling analysis, we applied the labeled mRNA to another set of microarrays on which DNA oligos were chosen from exactly the same regions as described above, except that the sense strand was used. These microarrays were designed for identifying internal transcripts that initiated from antisense templates. Indeed, we found 494 genes that fell within this category. One example is demonstrated in

Figure 3. Patterns of global acetylation increase at coding regions caused by specific mutations in the Rpd3S complex are similar to those seen in *SET2* deletion mutants. ChIP–chip experiments were carried out using yeast strains bearing either *eaf3Δchd* (YBL619) or *rco1Δphd* (YBL632) mutations. Microarray analyses were performed exactly as described in Figures 1 and 2.

Figure 5. Functional readout of the histone K36 methylation code. (A) Acetylation changes were used as the readout for K36 methylation. The K-mean cluster analysis revealed that in the $\Delta set2$ mutant, 621 genes possess internally initiated mRNA transcripts that are transcribed in the same direction as full-length transcripts, and 494 genes have cryptic transcripts that use the antisense strand as a template. The Venn diagram analysis (<http://www.pangloss.com/seidel/Protocols/venn.cgi>) of these genes and genes that are dependent on Set2–Rpd3S to maintain acetylation status (1685 genes) demonstrates an obvious overlap. (B) The enrichment level of K36 methylation was used as the readout for K36 methylation. The data set of the relative enrichment of K36me3 at coding regions (Pokholok et al. 2005) was subjected to K-mean Cluster analysis as described in Figure 1C. The two clusters that displayed the highest K36 trimethylation levels were chosen for subsequent Venn diagram analysis. The level of K36me3 at ORFs does not reflect well the functional dependence of genes on the Set2–Rpd3S pathway.

(>25%) (Fig. 1). Only a small fraction of genes that contain embedded cryptic promoters, however, might display the spurious transcription phenotype. When we used Venn diagram analysis to compare the list of genes showing either acetylation changes or appearance of cryptic transcripts upon deletion of *SET2*, we observed apparent overlap between the genes that depend on Set2–Rpd3S to suppress acetylation at ORFs and those that rely on the same pathway to prevent internal initiation (Fig. 5A). In contrast, very little similarity was found when comparing the genes that were highly enriched with K36me3 to the genes that require Set2 to repress spurious transcripts (Fig. 5B). Therefore, both functional readouts of K36 methylation (acetylation changes and cryptic initiation phenotype) suggested that the Set2–Rpd3S pathway played an important role at genes that normally were transcribed less often.

Discussion

Earlier studies showed that Set2-mediated K36 methylation targeted Rpd3S, which maintained a hypoacetylated state within the coding regions of certain genes.

This mechanism was essential to suppress the initiation of intragenic transcription (Carrozza et al. 2005; Joshi and Struhl 2005; Keogh et al. 2005). Here, using a genome-wide localization approach, we demonstrate that the Set2–Rpd3S pathway is generally responsible for controlling acetylation levels at coding regions (Fig. 1). Hyperacetylation caused by defects in this pathway triggers chromatin structure changes (data not shown) and exposes cryptic promoters to the transcriptional machinery, which results in the generation of spurious transcripts within a subset of genes (total 1032 genes) (Fig. 5). In addition, we discovered that genes with a low transcription frequency and longer ORF are strongly dependent on the functional integrity of the Set2–Rpd3S pathway.

We demonstrate that the Set2–Rpd3S pathway controls global acetylation at coding regions. Further analysis of acetylation changes in $\Delta set2$ mutants or Rpd3S mutant yeast strains suggests that longer genes are more dependent on this pathway (Fig. 1). When we examined transcription frequency, however, it was obvious that the less frequently transcribed genes are dependent on the Set2–Rpd3S pathway (Figs. 2, 3). This result is an apparent paradox, since previously published work demonstrates that K36me2 is associated with the ON/OFF state of transcription (Rao et al. 2005), and the level of K36me3 is indicative of transcription frequency (Pokholok et al. 2005). Thus, it seems that K36 methylation should instead be more important at highly transcribed genes. In order to resolve these seemingly conflicting results, we compared the list of genes that show an increase in AcH4 following *SET2* deletion to those that are dependent on Set2 to repress internal initiation. This Venn diagram analysis suggested that predictions made using acetylation changes as a readout (Fig. 5A) are much more accurate than those generated from examining the overall enrichment level of K36 methylation (Fig. 5B). Therefore, these data resolve the apparent paradox and demonstrate that less frequently transcribed genes are more dependent on the Set2–Rpd3S pathway. These results provide an example that the presence of a histone modification alone may not provide an accurate prediction of its direct *in vivo* function.

Why are infrequently transcribed genes so exquisitely sensitive to disruption of the Set2–Rpd3S pathway? The passage of Pol II is believed to alter the configuration of transcribed nucleosomes, presumably through histone eviction and/or redeposition (Workman 2006). The principle function of the Set2–Rpd3S pathway is to restore the acetylation status behind transcribing Pol II (Workman 2006). Genes that are not frequently transcribed have a low density of Pol II, and therefore histones that are deposited must be deacetylated by the Set2–Rpd3S pathway to prevent general transcription factor (GTF) binding and the initiation of aberrant cryptic transcription. For highly transcribed genes, the density of elongating Pol II at coding regions may sterically inhibit the binding of GTFs and the assembly of preinitiation transcription complexes at cryptic promoters in the absence of Set2.

Methylation of H3K36 is tightly associated with active

Table 1. *Yeast strains used in this study*

Name	Parental strain	Genotype	Source
YBL202	S288C	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Rsc2-TAP::HIS3MX6</i>	Open Biosystems
YBL250	BY4741	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Δset2::KAN</i>	Open Biosystems
YBL360	BY4741	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Rpb9-TAP</i>	Open Biosystems
YBL534	YJW594	<i>MATa, his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Δrco1::KANMX6</i>	Open Biosystems
YBL555	YJW677	<i>MATα his3Δ1 leu2Δ0 lys2Δ 0 ura3Δ0 eaf3-chd77-113Δ Flag:Leu2</i>	Carrozza et al. 2005
YBL591	w303a	<i>MATa ura3-1 lys2Δ::hisG trp1-1 his3-11,15 leu2-3,112 can1-100 proGAL-HA STE11::KAN</i>	This study
YBL597	YBL139	<i>MATa ura3-1 lys2Δ::hisG trp1-1 his3-11,15 leu2-3,112 can1-100 proGAL-HA STE11::KAN/Δset2:URA3</i>	This study
YBL609	BY4741	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Δjhd1::KAN</i>	Open Biosystems
YBL632	YBL631	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Rco1-ΔPh.D.</i>	This study
YBL633	BY4741	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Δmac1::KAN</i>	Open Biosystems

transcription, with K36me2 indicating the ON/OFF state and K36me3 reflecting the transcription frequency (Pokholok et al. 2005; Rao et al. 2005). However, this finding does not necessarily mean that methylation directly facilitates transcription. In fact, the higher level of methylation at more active genes could result from an increased rate of cotranscriptional methylation relative to the rate of demethylation or methyl K36 H3 loss. Therefore, increased transcription should increase the frequency of elongating Pol II passage down the ORF, resulting in a higher steady-state level of K36me3.

Despite its association with actively transcribed genes, K36me actually plays a repressive role in transcription (Strahl et al. 2002). Its primary function is to direct the repression of unwanted intragenic transcription at actively transcribed genomic loci (Li et al. 2007a), thereby facilitating productive full-length transcription. Mistargeting of Set2 to the promoter region has been shown to cause transcriptional inhibition (Strahl et al. 2002). Interestingly, cells have evolved multiple mechanisms to potentially prevent Set2-mediated K36 methylation at the promoter proximal region. First, Set2 is specifically associated with elongating Pol II (Li et al. 2002, 2003; Krogan et al. 2003; Xiao et al. 2003), which is only present throughout coding regions. Second, when genes are inactive, the histone variant Htz1 is localized to promoter regions, which inhibits the histone methyltransferase activity of Set2 (Li et al. 2005). Finally, at active promoters, Lys 36 is acetylated by Gcn5 (Morris et al. 2007), thereby preventing it from being methylated by Set2. This tight control of methylation by Set2 demonstrates the importance of this mechanism for productive, full-length transcription.

The yeast genome is generally considered to be euchromatic, since this organism has a relatively high rate of transcription. In contrast, only a small portion of higher eukaryotic genomes are actually engaged in active transcription, and the average length of genes in these organisms is much longer than in yeast. We discovered in yeast that longer, less frequently transcribed genes are more sensitive to disruption of the Set2–Rpd3S pathway. Thus, it is conceivable that this pathway would play a much more critical role in higher eukaryotes. The human Set2 homolog, HYPB, contains a region that could

interact with both phosphorylated RNA Pol II and the Huntington disease protein, huntingtin (Sun et al. 2005). Therefore, future studies into the role of the Set2–Rpd3S equivalent pathway in higher eukaryotes is of certain interest.

Materials and methods

Yeast strains

All *Saccharomyces cerevisiae* strains are derived from either S288C or w303a background and are listed in Table 1. YBL632 (*rco1Δphd*) was generated by genetic manipulation at endogenous loci using a modified delitto perfetto method (Storici and Resnick 2006; Li et al. 2007b). All mutations were confirmed either by PCR or genomic sequencing.

ChIP–chip and transcription profiling

ChIP was performed essentially as described in two previous publications (Strahl-Bolsinger et al. 1997; Li and Reese 2001) with anti-acetylated H4 antibody (Upstate Biotechnology). DNA microarrays used for ChIP–chip were designed using previously published templates (Pokholok et al. 2005) and were manufactured by Agilent Technologies. Our probe amplification and labeling procedures were adopted from earlier publications (Iyer et al. 2001; Li et al. 2005). Briefly, in Round A amplification, a degenerate primer with a universal oligo tag at the 5' end was used in a two-round primer extension reaction catalyzed by T7 sequenase version 2.0 (Upstate Biotechnology); one-sixth of Round A products was then subjected to Round B amplification with ExTaq (Takara) using the universal primer. DNA was subsequently purified using Qiagen PCR purification kits and was chemically labeled using a platinum-based coordination compound as described previously (Heetebrij et al. 2003).

RNA preparation was essentially as described by Carrozza et al. (2005). Briefly, total RNA was isolated by the acid-phenol extraction method from yeast cells exponentially grown in rich medium (YPD). Poly(A)⁺ mRNA was subsequently extracted from 3 mg of total RNA using oligo d(T) cellulose affinity chromatography. The eluted poly(A) RNA samples were directly labeled with Cy dyes using a platinum-based coordination compound (Heetebrij et al. 2003). Labeled RNA from wild type and *Δset2* were competitively hybridized to customized high-resolution oligonucleotide arrays (Agilent Technologies). The probe designs are available on request. In our case, the anti-sense design was intended to detect RNA transcribed from the

sense strand, while the sense design (by the Young Laboratory) was used to identify antisense RNA. The raw data were normalized so that the median ratio of intensities from two channels equals 1.

All microarray-based experiments were carried out at least in duplicate. The final data sets were generated from the median values of all repetitions with Pearson correlation exceeding 0.85.

Microarray data analysis

Final data sets from both ChIP-chip and transcription profiling experiments were pipelined into the average gene analysis based on the framework originated by the Young laboratory (Pokholok et al. 2005). Basically, the coding region of each gene in the yeast genome was divided into 40 equally sized bins. The upstream and downstream intergenic regions for the gene were each divided into 20 equally sized bins. The microarray expression value for each probe was assigned to the closest bin (whichever bin midpoint was closest to the probe midpoint). Empty bin values were calculated using linear interpolation from the last and next nonzero expression values or assigned to zero for transcriptional profiling. In this way, an expression matrix was generated with 80 columns (corresponding to the intergenic and ORF bins) and a row for each gene in the yeast genome. Each column was then averaged throughout the entire matrix or within certain subgroups to create the average expression values for each bin position. Microarray data are available at ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) under accession numbers E-TABM-260 and E-TABM-265.

Acknowledgments

We are grateful to Dr. Thomas Kusch and other members of the Workman laboratory for useful discussions and technical suggestions. This work was supported by a grant from the NIGMS to J.L.W and funding from the Stowers Institute.

References

- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120–124.
- Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581–592.
- Hampsey, M. and Reinberg, D. 2003. Tails of intrigue: Phosphorylation of RNA polymerase II mediates histone methylation. *Cell* **113**: 429–432.
- Hediger, F. and Gasser, S.M. 2006. Heterochromatin protein 1: Don't judge the book by its cover! *Curr. Opin. Genet. Dev.* **16**: 143–150.
- Heetebrij, R.J., Talman, E.G., van Velzen, M.A., van Gijlswijk, R.P., Snoeijs, S.S., Schalk, M., Wiegant, J., van de Rijke, F., Kerkhoven, R.M., Raap, A.K., et al. 2003. Platinum(II)-based coordination compounds as nucleic acid labeling reagents: Synthesis, reactivity, and applications in hybridization assays. *ChemBioChem* **4**: 573–583.
- Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717–728.
- Iyer, V.R., Horak, C.E., Scafe, C.S., Botstein, D., Snyder, M., and Brown, P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**: 533–538.
- Jason, L.J., Moore, S.C., Lewis, J.D., Lindsey, G., and Ausio, J. 2002. Histone ubiquitination: A tagging tail unfolds? *Bioessays* **24**: 166–174.
- Jenuwein, T. and Allis, C.D. 2001. Translating the histone code. *Science* **293**: 1074–1080.
- Joshi, A.A. and Struhl, K. 2005. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol. Cell* **20**: 971–978.
- Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., et al. 2005. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**: 593–605.
- Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., et al. 2003. Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **23**: 4207–4218.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120.
- Li, B. and Reese, J.C. 2001. Ssn6-Tup1 regulates RNR3 by positioning nucleosomes and affecting the chromatin structure at the upstream repression sequence. *J. Biol. Chem.* **276**: 33788–33797.
- Li, J., Moazed, D., and Gygi, S.P. 2002. Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J. Biol. Chem.* **277**: 49383–49388.
- Li, B., Howe, L., Anderson, S., Yates III, J.R., and Workman, J.L. 2003. The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **278**: 8897–8903.
- Li, B., Pattenden, S.G., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J., and Workman, J.L. 2005. Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. *Proc. Natl. Acad. Sci.* **102**: 18385–18390.
- Li, B., Carey, M., and Workman, J.L. 2007a. The role of chromatin during transcription. *Cell* **128**: 707–719.
- Li, B., Gogol, M., Carey, M., Lee, D., Seidel, C., and Workman, J.L. 2007b. Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science* **316**: 1050–1054.
- Morris, S.A., Rao, B., Garcia, B.A., Hake, S.B., Diaz, R.L., Shabanowitz, J., Hunt, D.F., Allis, C.D., Lieb, J.D., and Strahl, B.D. 2007. Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *J. Biol. Chem.* **282**: 7632–7640.
- Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A., and Reinberg, D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**: 703–717.
- Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolzheimer, E., et al. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**: 517–527.
- Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates III, J.R., and Grant, P.A. 2005. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433**: 434–438.
- Rao, B., Shibata, Y., Strahl, B.D., and Lieb, J.D. 2005. Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. *Mol. Cell. Biol.* **25**: 9447–9459.

Li et al.

- Reid, J.L., Moqtaderi, Z., and Struhl, K. 2004. Eaf3 regulates the global pattern of histone acetylation in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **24**: 757–764.
- Saeed, A.I., Bhagabati, N.K., Braisted, J.C., Liang, W., Sharov, V., Howe, E.A., Li, J., Thiagarajan, M., White, J.A., and Quackenbush, J. 2006. TM4 microarray software suite. *Methods Enzymol.* **411**: 134–193.
- Shilatifard, A. 2006. Chromatin modifications by methylation and ubiquitination: Implications in the regulation of gene expression. *Annu. Rev. Biochem.* **75**: 243–269.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**: 844–847.
- Sims III, R.J., Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. 2005. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem.* **280**: 41789–41792.
- Storici, F. and Resnick, M.A. 2006. The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol.* **409**: 329–345.
- Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., et al. 2002. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol. Cell. Biol.* **22**: 1298–1306.
- Strahl-Bolsinger, S., Hecht, A., Luo, K., and Grunstein, M. 1997. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes & Dev.* **11**: 83–93.
- Sun, X.J., Wei, J., Wu, X.Y., Hu, M., Wang, L., Wang, H.H., Zhang, Q.H., Chen, S.J., Huang, Q.H., and Chen, Z. 2005. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J. Biol. Chem.* **280**: 35261–35271.
- Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T., et al. 2006. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol. Cell* **24**: 785–796.
- Turner, B.M. 2002. Cellular memory and the histone code. *Cell* **111**: 285–291.
- Workman, J.L. 2006. Nucleosome displacement in transcription. *Genes & Dev.* **20**: 2009–2017.
- Xiao, T., Hall, H., Kizer, K.O., Shibata, Y., Hall, M.C., Borchers, C.H., and Strahl, B.D. 2003. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes & Dev.* **17**: 654–663.
- Zhang, Y. 2006. It takes a PHD to interpret histone methylation. *Nat. Struct. Mol. Biol.* **13**: 572–574.

Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription

Bing Li, Madelaine Gogol, Mike Carey, et al.

Genes Dev. 2007, **21**:

Access the most recent version at doi:[10.1101/gad.1539307](https://doi.org/10.1101/gad.1539307)

Supplemental Material

<http://genesdev.cshlp.org/content/suppl/2007/05/18/21.11.1422.DC1>

References

This article cites 38 articles, 17 of which can be accessed free at:
<http://genesdev.cshlp.org/content/21/11/1422.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).