Histone Variant H2A.Bbd Is Associated with Active Transcription and mRNA Processing in Human Cells

Michael Y. Tolstorukov,[1,5,6] Joseph A. Goldman,[2,6] Cristele Gilbert,[3] Vasily Ogryzko,[3] Robert E. Kingston,[2,*] and Peter J. Park[1,4,5,*]

[1]Center for Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA
[2]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA
[3]Université Paris-Sud 11 CNRS UMR 8126, Interactions Moléculaires et Cancer, Institut de Cancérologie Gustave-Roussy, F-94805 Villejuif Cedex, France
[4]Informatics Program, Children's Hospital, Boston, MA 02115, USA
[5]Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA
[6]These authors contributed equally to this work
*Correspondence: kingston@molbio.mgh.harvard.edu (R.E.K.), peter_park@harvard.edu (P.J.P.)
http://dx.doi.org/10.1016/j.molcel.2012.06.011

SUMMARY

Variation in chromatin composition and organization often reflects differences in genome function. Histone variants, for example, replace canonical histones to contribute to regulation of numerous nuclear processes including transcription, DNA repair, and chromosome segregation. Here we focus on H2A.Bbd, a rapidly evolving variant found in mammals but not in invertebrates. We report that in human cells, nucleosomes bearing H2A.Bbd form unconventional chromatin structures enriched within actively transcribed genes and characterized by shorter DNA protection and nucleosome spacing. Analysis of transcriptional profiles from cells depleted for H2A.Bbd demonstrated widespread changes in gene expression with a net downregulation of transcription and disruption of normal mRNA splicing patterns. In particular, we observed changes in exon inclusion rates and increased presence of intronic sequences in mRNA products upon H2A.Bbd depletion. Taken together, our results indicate that H2A.Bbd is involved in formation of a specific chromatin structure that facilitates both transcription and initial mRNA processing.

INTRODUCTION

The dynamic packaging of DNA into chromatin regulates gene expression by altering the accessibility of DNA to the transcriptional machinery. The basic building block of chromatin is the nucleosome, consisting of about 150 base pairs (bp) of DNA wrapped around a core of eight histones of four different types: H2A, H2B, H3, and H4 (Luger et al., 1997). Histones are subject to modifications including addition and removal of covalent marks and replacement of "major" histones with nonallelic variants. The exchange of an entire histone with a variant represents a drastic alteration in local chromatin organization and as such is expected to be significant to chromatin biology. Replacement histones are typically expressed throughout the cell cycle (Wu and Bonner, 1981) and play a role in various processes by affecting chromatin structure and dynamics during G1 and G2 beyond bulk packaging of the genome. These processes include chromosome segregation, DNA repair, transcriptional regulation, and mRNA processing (Talbert and Henikoff, 2010). However, the biological pathways and functions of histone variants are not fully understood.

In this study, we focus on a histone H2A family member, the mammalian-specific variant H2A.Bbd (Barr Body Deficient). This is one of the most rapidly evolving variants, which shares approximately 50% sequence identity with the major H2A histone (Chadwick and Willard, 2001; Gonzalez-Romero et al., 2008; Malik and Henikoff, 2003). The histone H2A family includes a number of replacement variants shown to have distinct biophysical properties and biological functions. For instance, H2A.Z is associated with transcriptional regulation and is enriched in nucleosomes positioned near transcriptional start sites (TSSs) (Barski et al., 2007; John et al., 2008). The macroH2A variant is associated with the inactivated X chromosome in female mammals, indicating a possible relationship with gene silencing. Recent studies showed depletion of this variant at the TSS of active genes (Buschbeck et al., 2009; Gamble et al., 2010). In contrast to macroH2A, H2A.Bbd is excluded from the inactive X chromosome and is enriched in active regions of genome (Chadwick and Willard, 2001). The genes encoding H2A.Bbd are expressed in a variety of tissues and cell types in both mouse and human, suggesting a widespread role in chromatin regulation (Gonzalez-Romero et al., 2008).

Biochemical analyses using recombinant H2A.Bbd support the association of this variant with active transcription. H2A.Bbd nucleosomes were reported to be less stable than those bearing the major H2A histone and to organize only ~120 bp of DNA, indicating that DNA in H2A.Bbd chromatin

may be more accessible and, thus, more amenable to transcription (Bao et al., 2004). An increased transcription rate was demonstrated in vitro on arrays of H2A.Bbd nucleosomes as compared to arrays containing nucleosomes with the major H2A histone (Zhou et al., 2007). Furthermore, the H2A.Bbd nucleosomes exhibit faster exchange rates in vivo, a hallmark of active transcription (Gautier et al., 2004).

To further study the role of H2A.Bbd in vivo, we profiled the genome-wide localization of nucleosomes containing H2A.Bbd, macroH2A1.2, H2A.Z, and major H2A using high-throughput sequencing. We report that H2A.Bbd is associated with altered chromatin organization, is significantly enriched in actively transcribed genes, and is associated with the elongating form of RNA polymerase. Analysis of the protein composition of H2A.Bbd-enriched chromatin and the effects of H2A.Bbd depletion on gene expression indicates that H2A.Bbd is involved with mRNA splicing. Therefore, our data imply that H2A.Bbd may link transcription and subsequent mRNA processing, revealing a role for nucleosomes in integrating these two critical gene expression pathways.

RESULTS

H2A.Bbd Nucleosomes Arrange Less DNA and Decrease Nucleosome Repeat Length

To characterize the function of nucleosomes containing H2A.Bbd, we compared the properties of variants from the H2A family using HeLa cells stably expressing FLAG-tagged versions of H2A.Bbd, macroH2A1.2, H2A.Z, and the major H2A histone. Each recombinant histone variant was found predominantly in the chromatin fraction (H2A, 99%; H2A.Z, 99%; H2A.Bbd, 84%; macroH2A, 99%), suggesting that each recombinant protein was efficiently incorporated. The recombinant histones H2A, H2A.Z, and macroH2A were expressed at levels lower than the endogenous histone but sufficiently high to constitute a significant fraction of the total histone content (17%, 33%, and 44%, respectively). The level of the recombinant H2A.Bbd protein relative to endogenous H2A.Bbd could not be calculated, since the amount of endogenous H2A.Bbd protein is unknown (see the Experimental Procedures). Western analysis using FLAG antisera showed that recombinant H2A.Bbd is 4-fold less abundant than recombinant H2A.Z and 12-fold less abundant than endogenous H2A.Z (data not shown). Nuclei were MNase digested, and chromatin was enriched for each variant by immunoprecipitation. Purified DNA was subjected to high-throughput sequencing (ChIP-seq) using the Illumina platform, resulting in at least 10 million uniquely mappable tags for each variant (see Table S1 available online).

We began characterization of the H2A.Bbd in vivo by determining whether the unusual biochemical properties of this variant (e.g., length of DNA protected from nuclease digestion by the histone octamer) were conserved in cultured cells. Structural features of chromatin containing this variant can be inferred from deep sequencing data (Figure 1A) (Tolstorukov et al., 2009). Correlation analysis of the tag frequencies on the positive and negative DNA strands reveals that the H2A.Bbd nucleosomes protect an average of 124 bp of DNA (Figure 1B), while other forms of histone H2A analyzed in this study show protection

lengths close to 147 bp, as expected for a canonical nucleosome (Figure S1A). This observation was corroborated by capillary gel electrophoresis analysis of the DNA fragments submitted for sequencing (Figure 1C, Figure S1B). Thus, these data are in agreement with previous reports based on in vitro-assembled H2A.Bbd nucleosomes (Bao et al., 2004).

We analyzed tag distributions on the same strands to estimate the average nucleosome repeat length, which represents the sum of the nucleosome-associated DNA and the intervening linker DNA. Thus, this analysis addresses the issue of whether the shortened amount of DNA associated with the H2A.Bbd nucleosomes is also reflected in H2A.Bbd nucleosome spacing in cells. We found that the average repeat length is 137 bp for H2A.Bbd nucleosomes, substantially shorter than the 170–180 bp reported for other nucleosomes in human chromatin and other variants analyzed in our study (Lohr et al., 1977) (Figure 1D, Figure S1C). This finding agrees with a previous study in which nucleosomes were assembled on a template in vitro (Bao et al., 2004). Assuming that H2A.Bbd nucleosomes are arranged adjacently, this indicates that the linker length in H2A.Bbd chromatin is 13 bp (137–124) rather than canonical 25–30 bp. We note that canonical nucleosomes, which organize 147 bp of DNA (Luger et al., 1997), could not be arranged in arrays with the repeat lengths of 137 bp observed for H2A.Bbd, due to likely steric clashes between adjacent nucleosomes. The shorter protection observed for the H2A.Bbd nucleosomes has been thought to reflect dynamic breathing of the DNA arms off the histone surface (Zhou et al., 2007). However, the shortened repeat length observed for H2A.Bbd nucleosomes argues for stable changes in the nucleosome particle upon H2A.Bbd incorporation (Doyen et al., 2006), as only nucleosomes with altered structural organization can be arranged with such short repeats.

H2A.Bbd Nucleosomes Are Enriched in Active Genes

Deviation of the H2A.Bbd variant from major H2A histone in both primary sequence and nucleosome arrangement suggests that H2A.Bbd nucleosomes may serve a distinct function in the genome and are unlikely to be distributed randomly. We therefore analyzed the ChIP-seq data to investigate and compare genomic distribution of H2A.Bbd and other members of the H2A family (Figure 2). In agreement with previous reports, H2A.Z nucleosomes showed localized enrichment at TSS (Barski et al., 2007) and macroH2A occupied extended regions that often encompassed multiple genes (Gamble et al., 2010). In contrast, H2A.Bbd was enriched within individual genes, and the regions of its enrichment often corresponded to the regions of macroH2A depletion. Overall, 63% of H2A.Bbd enrichment clusters overlap with genes, and 36% of genes are associated with at least one H2A.Bbd locus, in comparison to 25% and 14%, respectively, for macroH2A (see the Experimental Procedures), indicating a possible role for H2A.Bbd within genes.

To further characterize the localization in transcription units, we calculated the average profile of each variant around the TSS and transcription termination sites (TTSs) (Figures 3A–3D). In this analysis, the genes were divided into five groups (quantiles) by expression levels, and the profiles were normalized by the tag densities in bulk chromatin. H2A.Bbd was enriched in
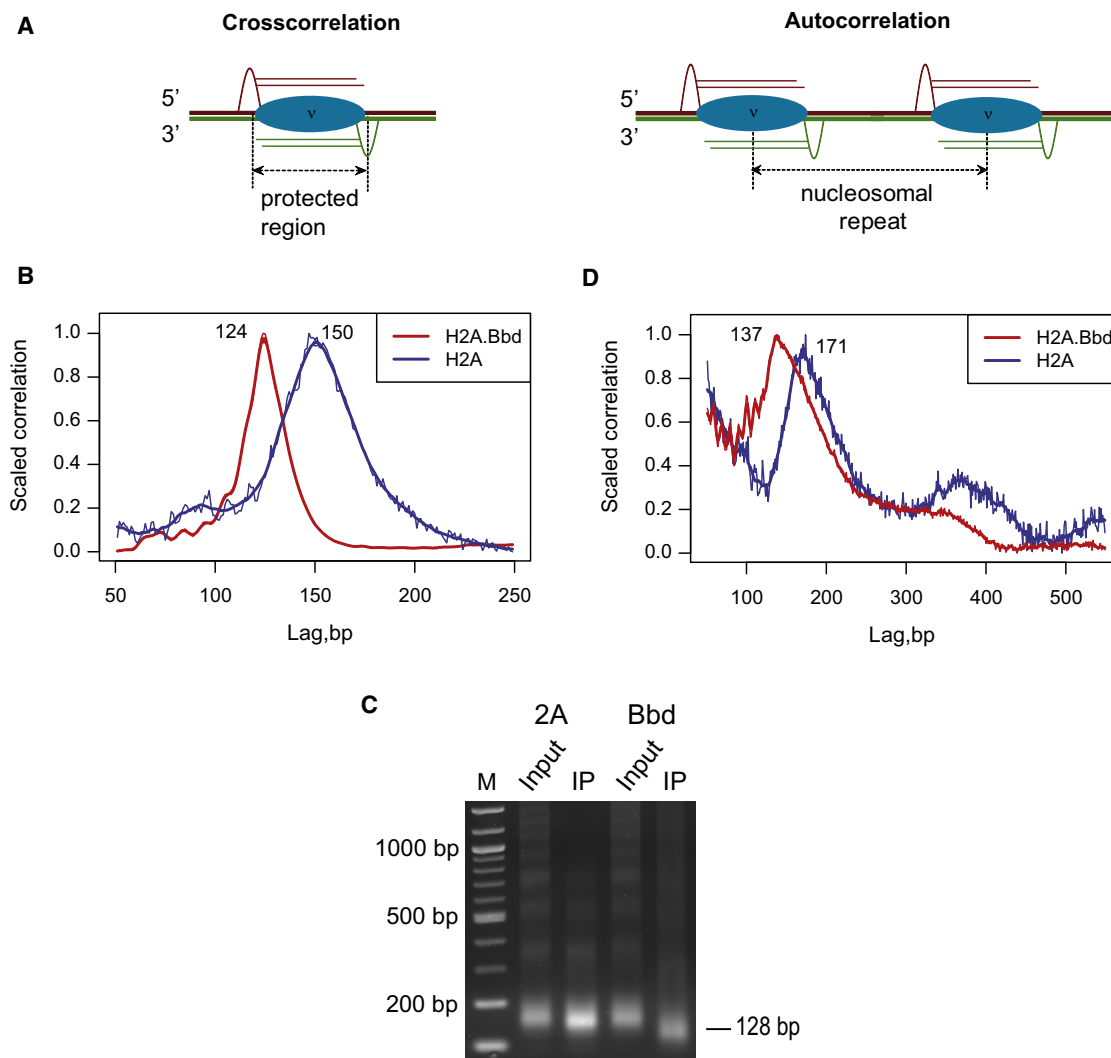
Figure 1. Properties of H2A.Bbd-Enriched Nucleosomes

(A) Schematic illustration of the algorithms used to compute cross- and autocorrelations of sequenced tag frequencies. Nucleosomes are represented with blue ovals. DNA-positive (red) and -negative (green) strands are shown with thick solid lines, and mononucleosomal fragments released by chromatin fragmentation are shown with thin solid lines. Distances between buildups in the frequencies of mapped tags (bell-shaped curves) were used to identify the length of DNA protected by the histone core or the nucleosomal repeat.

(B) Crosscorrelation of tag frequency distributions for H2A.Bbd (red) and major H2A variant (blue). Peaks in the tag frequency crosscorrelation plot provide information about the characteristic size of the nucleosome-protected fragment (indicated with numbers). Shorter DNA protection is a characteristic of H2A.Bbd-containing nucleosomes.

(C) Input chromatin DNA was compared to immunopurified chromatin DNA as indicated along the top by agarose gel electrophoresis.

(D) Autocorrelation of tag frequency distributions is labeled as in (B). The peaks in tag autocorrelation correspond to the nucleosomal repeat lengths, and a shorter repeat length is detected for H2A.Bbd nucleosomes as compared to nucleosomes containing the major H2A variant.

bodies of expressed genes, with strong enrichment throughout the entire gene and clear correlation with the expression level. H2A.Bbd was depleted in the immediate vicinity of TSS, with steep increase to the body of the genes, especially for highly expressed genes (Figure 3A). In contrast, macroH2A was depleted throughout expressed genes but showed elevated enrichment levels over both promoters and transcribed regions of silent genes (Figure 3B). In accordance with published results, H2A.Z was enriched while H2A was depleted at the starts of active

genes. Association of the H2A variants with gene expression status was statistically significant (Figures S2A–S2C).

In principle, the average profile may be a composite of multiple classes of genes, each containing distinct enrichment patterns across the transcribed regions. To rule out this possibility, we grouped the genes using K-means clustering after averaging the profiles over the TSS-proximal, gene body, and the TTS-proximal regions. We then computed the average expression for each resulting cluster (Figures S2D–S2F). The result of this
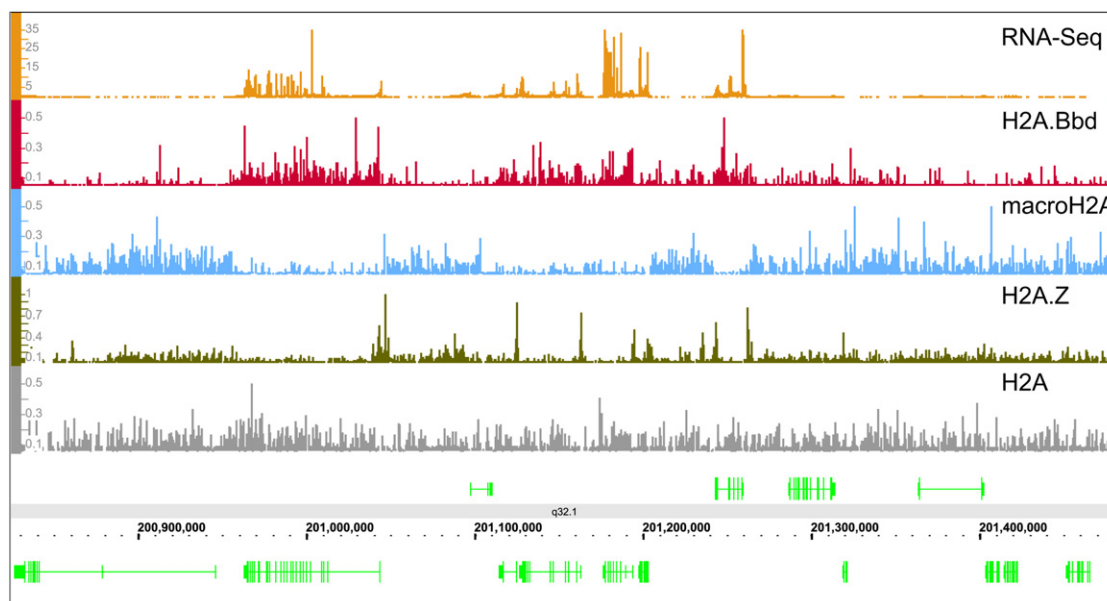
Figure 2. A Snapshot of the Genome-wide Distribution of H2A Variants
Distributions of the sequenced tag density for H2A.Bbd (red), macroH2A (blue), H2A.Z (green), and major H2A variant (gray) are shown within a locus on chromosome 1. The upper track (orange) shows RNA-seq coverage and provides reference of gene expression levels. Gene structure is indicated in green at the bottom of the plot. Note that H2A.Bbd is enriched in active genes, while macroH2A is enriched in silent genes and tends to cover domains spanning several genes.

analysis was consistent with the data presented in Figure 3, showing, for example, that the clusters characterized by high H2A.Bbd enrichment within gene bodies are comprised of expressed genes.

Since we used cells expressing the recombinant H2A.Bbd in addition to endogenous histone, we wondered if H2A.Bbd enrichment in transcribed genes could be explained merely by nonspecific association with regions of high nucleosome turnover. However, we observed depletion rather than enrichment of H2A.Bbd at DNase I hypersensitive sites (Figure S2G) and TSSs (Figure 3), which are often associated with high nucleosome turnover rates (Deal et al., 2010; Mito et al., 2007). Thus, our observation of H2A.Bbd enrichment in transcribed regions is unlikely to be a simple consequence of nonspecific association of H2A.Bbd with "open" chromatin regions.

The enrichment of H2A.Bbd within active genes may have resulted from strong association of H2A.Bbd with a group of active genes involved in specific biological processes or from ubiquitous enrichment of H2A.Bbd in all active genes. To distinguish between these alternatives, we performed gene ontology (GO) analysis for the top 10% of the H2A.Bbd-enriched genes (Table 1). This analysis showed that genes involved in intracellular metabolic processes, including protein and RNA metabolism, are strongly overrepresented among the H2A.Bbd-enriched genes. This is in contrast to the results for macroH2A-enriched genes, which showed enrichment for intercellular processes including development (Gamble et al., 2010). To distinguish H2A.Bbd-specific GO categories from those enriched in highly expressed genes, we repeated the analysis using only the top 40% of active genes as our background reference

set, and we identified similar GO categories (Table S2). We conclude that H2A.Bbd enrichment is a characteristic of genes that are involved in a broad range of cellular functions but is not pronounced in all active genes.

H2A.Bbd Is Associated with Markers of Active Chromatin

We expected that H2A variant-containing chromatin would be enriched for covalent marks appropriate to the regulatory characteristics of genes associated with the particular variant (Barski et al., 2007). To investigate whether the expected covalent marks on histone H3 colocalize with H2A variant chromatin, we analyzed bulk-purified H2A variant chromatin by western blot (Figure 4A). Marks correlated with gene silencing were depleted from H2A.Bbd chromatin (see results for H3K27me3 and H3K79me3 marks), while active marks were enriched (compare results for H3K4me2 and H3K4me3 marks). A covalent modification enriched in gene bodies without preference for particular regions of gene bodies (H3K27me1) was also slightly enriched with H2A.Bbd. Overall, the abundance of H3 modifications on bulk H2A.Bbd chromatin is in agreement with the localization observed previously (see Figure 3).

To further explore the possible involvement of H2A.Bbd in transcriptional regulation, we probed equal amounts of purified chromatin enriched in each of the H2A variants with an antibody against the N terminus of RNA Pol II by western blot (Figure 4B). As expected, relative to H2A, both H2A.Z and H2A.Bbd chromatin are enriched for total RNA Pol II. The Ser2-phosphorylated form of RNA Pol II, however, is only enriched with H2A.Bbd chromatin. The observed specificity of the elongating RNA Pol
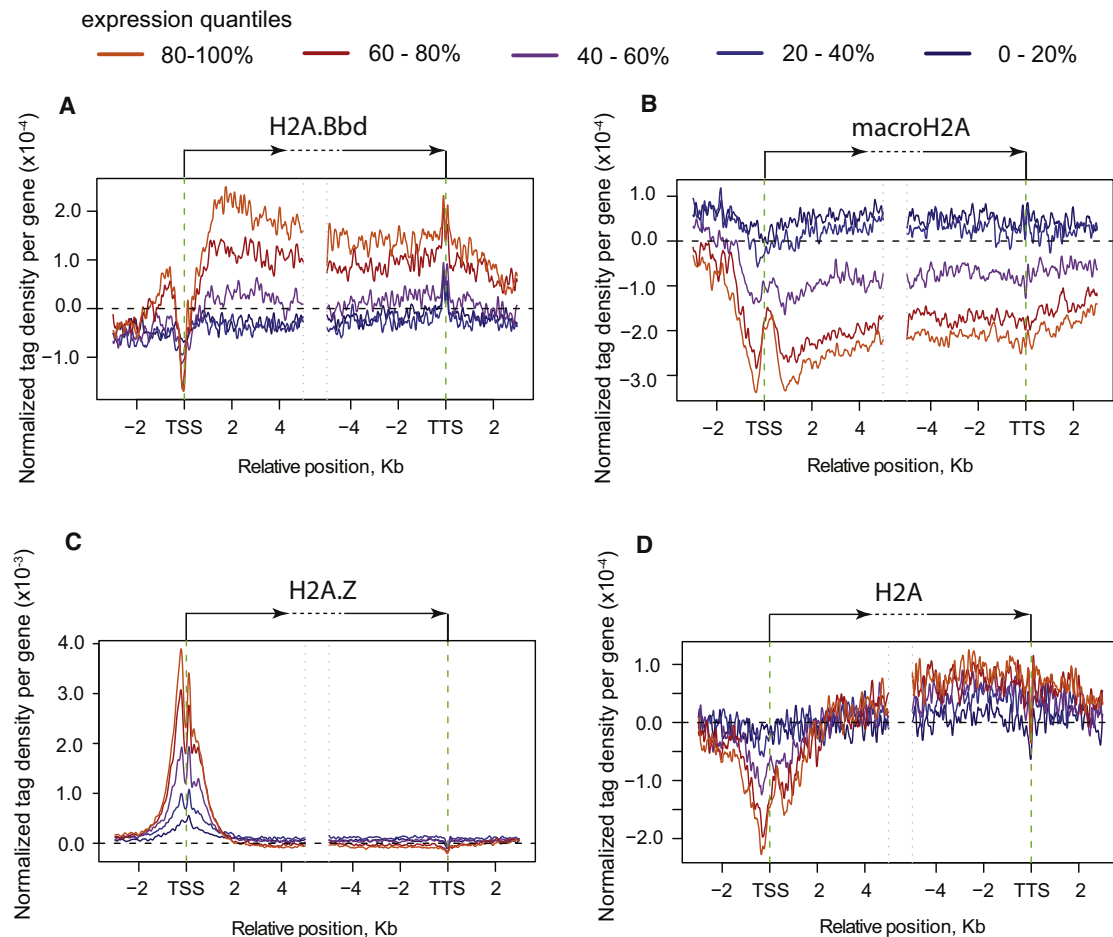
Figure 3. Average Enrichment Profiles around Transcription Start and Termination Sites Calculated for the Histone Variants H2A.Bbd, MacroH2A, H2A.Z, and Major H2A

The genes were placed into five groups of equal sizes according to their expression levels (expression quantiles, with "0–20%" corresponding to the lowest, and "80–100%" to the highest expression levels). Transcription start and termination sites are indicated with thin green lines. Arrows at the top of the plots indicate direction of the transcription. Overlapping genes and genes shorter than 500 bp were excluded from the analysis.

II to H2A.Bbd chromatin is consistent with the distinct distribution patterns of H2A variants along genes shown in Figure 3. Thus, our results indicate that H2A.Bbd is a reliable marker of expressed genes and is associated with other chromatin features of transcribed genes.

H2A.Bbd Depletion Affects Gene Expression Patterns

The genomic distribution and association of H2A.Bbd with elongating RNA Pol II led us to consider that H2A.Bbd might be involved in the regulation of transcription or processes related to transcriptional elongation. To investigate how deficiency in H2A.Bbd can affect gene expression, we depleted H2A.Bbd using short-hairpin RNA (shRNA) construct and used high-throughput sequencing of nuclear RNA (RNA-seq) to characterize changes in transcript abundance. To rule out possible artifacts due to overexpression of H2A.Bbd in the transgenic cell line, we used such cells only to check efficiency of the shRNA constructs, while the actual RNA-seq experiment was performed in nontransgenic cells. Transcription profiles for cells treated with shRNA constructs targeting the H2A.Bbd transcript were compared with those obtained using a control shRNA with no homology to the human genome.

Depletion of H2A.Bbd results in substantial changes in gene expression levels, with genes being both up- and downregulated (Figure 4C). With two biological replicates in each experiment, we identified ∼1,200 differentially expressed genes (false discovery rate of 0.05 and greater than two fold change for a conservative estimate; see the Experimental Procedures), with more genes downregulated than upregulated (716 and 522 genes, respectively, $p = 6 \times 10^{-9}$).

To evaluate if H2A.Bbd nucleosomes directly regulate transcription, we examined how the amount of H2A.Bbd in a gene is correlated with the extent of change in gene expression following H2A.Bbd depletion. The expression fold change is mildly correlated with the enrichment of the gene in H2A.Bbd (Pearson's correlation coefficient, R = 0.2, Figure S3). Also, the

Table 1. Enrichment of H2A.Bbd- and MacroH2A-Associated Genes in Gene Ontology Terms

Gene Ontology Term	Frequency in Gene Test Set	Corrected p Value
H2A.Bbd (2,475 Genes)		
Metabolic process	1,014 of 2,475 genes, 41.0%	6.01E-60
Cellular process	1,534 of 2,475 genes, 62.0%	7.02E-56
Cellular protein metabolic process	471 of 2,475 genes, 19.0%	6.74E-45
Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	382 of 2,475 genes, 15.4%	1.00E-43
RNA metabolic process	237 of 2,475 genes, 9.6%	4.08E-37
Nitrogen compound metabolic process	417 of 2,475 genes, 16.8%	2.42E-35
Gene expression	276 of 2,475 genes, 11.2%	5.57E-33
MacroH2A (2,483 Genes)		
Multicellular organismal process	584 of 2,483 genes, 23.5%	2.01E-10
Ion transport	136 of 2,483 genes, 5.5%	8.73E-07
Multicellular organismal development	386 of 2,483 genes, 15.5%	9.92E-06
Cell-cell signaling	105 of 2,483 genes, 4.2%	4.01E-05
Cell adhesion	121 of 2,483 genes, 4.9%	6.45E-05
System development	317 of 2,483 genes, 12.8%	6.99E-05
Anatomical structure development	334 of 2,483 genes, 13.5%	0.00014

For each variant, 10% of the genes associated with highest tag density were taken for the analysis, and all human genes were used as a background. Top seven GO categories are reported for each variant (if several categories showed significant overlap, only one of them was retained).

range of the expression fold change is considerably smaller for the genes with lower levels of H2A.Bbd enrichment. These findings, combined with the association of H2A.Bbd with the marks of active transcription (Figures 4A and 4B), suggest the importance of this variant for positive regulation of transcription. We note, however, that the skew in the number of affected genes toward upregulation is only modest (less than 2-fold), which can be indicative of a complex regulatory role of this variant. We therefore considered other regulatory mechanisms that may involve H2A.Bbd.

Proteomic Analysis Reveals Association of H2A.Bbd with RNA Processing

To further investigate possible functions of H2A.Bbd, we undertook a proteomic approach to determine the factors that are associated with H2A.Bbd chromatin using mass spectrometry. We isolated chromatin enriched for H2A.Bbd using the

epitope-tagged H2A.Bbd cell line and compared it to similarly purified macroH2A-enriched chromatin (Figure 5A). The mass spectrometry analysis revealed that 57 proteins were enriched only with H2A.Bbd chromatin and that an additional nine proteins were at least 2-fold enriched over macroH2A (Table S3). Some proteins were also enriched with macroH2A (data not shown), most notably the previously described PARP1, which was present in roughly stoichiometric amounts with the tagged macroH2A (Figure 5A) (Ouararhni et al., 2006). The majority of proteins specific to H2A.Bbd chromatin were RNA processing factors including 15 proteins identified as components of the spliceosome, leading us to the surprising hypothesis that H2A.Bbd regulates interactions with RNA splicing machinery. In particular, five of the seven members of the Sf3b complex required for branchpoint recognition and U2 snRNP assembly were found specifically enriched in the H2A.Bbd sample (Corrionero et al., 2011; Folco et al., 2011; Gozani et al., 1996). These data indicate that H2A.Bbd chromatin associates with proteins important for assembly and function of the spliceosome.

The results obtained with mass spectrometry were confirmed by western blot, comparing similar amounts of purified chromatin-containing H2A variants. To ensure that Sf3b association with H2A.Bbd is not a common feature of chromatin present in active gene bodies, the histone variant H3.3 was included in this analysis, since H3.3 is also found inside active genes (Jin et al., 2009; Schwartz and Ahmad, 2005). Again, proteins from the Sf3b complex as well as other spliceosome factors were found enriched with H2A.Bbd chromatin (Figure 5B). Enzymes involved in incorporation of H2A.Z and H3.3 (EP400 and DAXX) were found enriched with chromatin only from the purified variant, validating the purification of H2A.Z and H3.3 chromatin. Since MNase used to solubilize chromatin prior to purification is a potent single-stranded nuclease, little or no RNA remained in the purification, indicating that the specific interaction of Sf3b with H2A.Bbd chromatin likely occurs through nucleosomes and not RNA (Figure S4A). We conclude that H2A.bbd nucleosomes associate with proteins linked to the spliceosome, particularly with the Sf3b subcomplex required for assembly of the U2 snRNP.

Furthermore, analysis of the H2A.Bbd density around splice sites reveals a distinct pattern in nucleosome positioning around the intron-exon junction, which extends to the branchpoint recognized by U2 snRNP (Figure 5C). Comparison of the H2A.Bbd and H2A.Z densities shows that this pattern is specific to H2A.Bbd nucleosomes. In particular, the peak in the tag density located on the intron side of the junction is pronounced for H2A.Bbd and not for H2A.Z. Interestingly, the distance between two stable H2A.Bbd nucleosome positions flanking the junction is about 135 bp, which is less than the DNA protection length in canonical nucleosomes but is consistent with the size of DNA fragments associated with H2A.Bbd nucleosomes. This observation suggests that the specific structural properties of H2A.Bbd nucleosomes, described above, can be operational for splice site definition.

To establish the extent of the association between H2A.Bbd and splicing, we compared the enrichment of H2A.Bbd in single- and multiple-exon genes (Figures 5D and 5E). We hypothesized
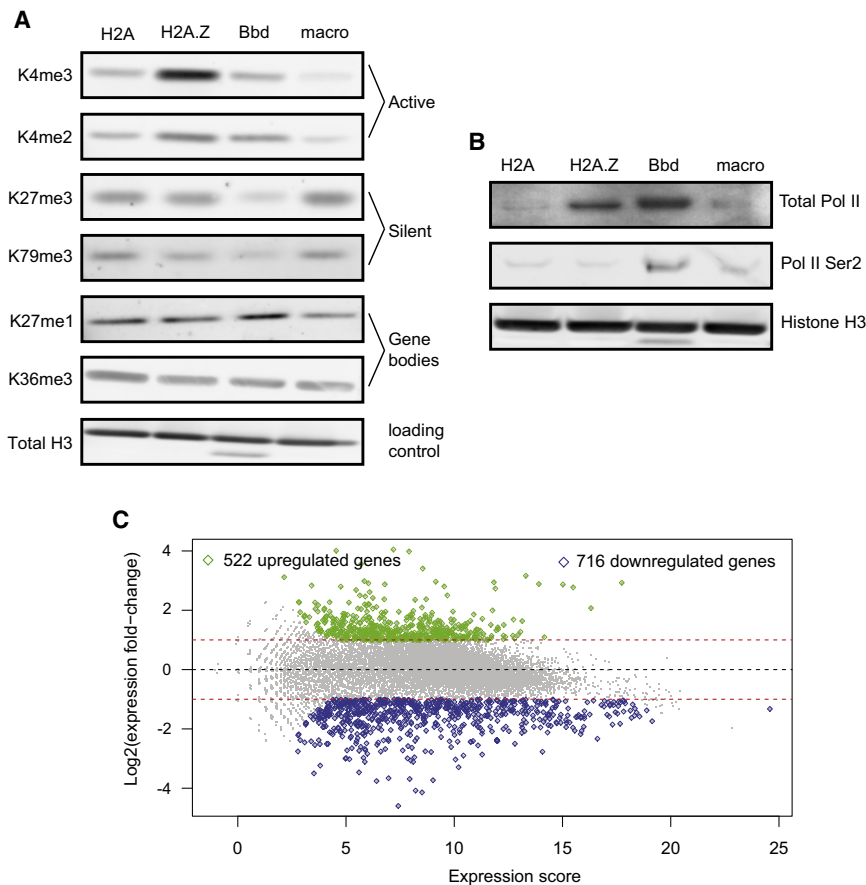
Figure 4. Role of H2A.Bbd Variant in Transcription

(A) Association of H2A.Bbd-enriched chromatin with different histone marks. Similar amounts of affinity-purified chromatin from the cell lines indicated at the top were analyzed by western blot using antibodies against different histone H3 methylation marks as indicated on the left. Total H3 was used as a chromatin loading control. Marks were grouped according to their relationship with gene location or expression as indicated on the right.

(B) Western blot analysis as in (A) but performed with antibodies for RNA polymerase and histone H3. Enrichment with a particular histone variant is determined relative to the amount of RNA polymerase detected with H2A-purified chromatin which is more representative of bulk chromatin.

(C) A scatterplot illustrating the changes in gene expression scores upon H2A.Bbd depletion. For each gene with nonzero RNA-seq fragment density in all considered samples, the longest annotated transcript was included in the analysis (~18,000). The fold changes of the expression values (log2) in the shRNA-treated sample as compared to the control sample are indicated on the y axis. The larger of the expression scores between the shRNA and control samples for each gene is indicated on the x axis. Green and blue squares mark transcripts that were significantly up- or downregulated by H2A.Bbd depletion (absolute value of the expression log2 fold change $\geq$1 and FDR $\leq$0.05).

that if H2A.Bbd were generally important to RNA splicing, multiple-exon genes (average length, ~50 Kb) would be more enriched with H2A.Bbd, since single-exon genes (average length, ~2 Kb) are not likely to require splicing. As expected, we observed significant enrichment along the entire length of multiple-exon genes, in contrast to nonuniform and weak enrichment in single-exon genes. As another control, we used a set of randomly selected genes with the same number of genes as in the single-exon group (Figure S4B), which confirmed that the difference between the single- and multiple-exon groups is not a function of group size. We also confirmed that the fraction of single-exon genes in a group decreases with the increase in H2A.Bbd density (Figure S4C). Based on these unexpected observations, we hypothesized that the H2A.Bbd variant is associated with mRNA processing. Alterations in splicing efficiency are also likely to lead to changes in steady-state amounts of affected RNAs, consistent with the conclusion of the expression analysis described above.

H2A.Bbd Depletion Results in Alteration of mRNA Splicing

To investigate the role of histone H2A.Bbd in regulating RNA processing, we further characterized the RNA-seq data from H2A.Bbd-depleted cells. As in the analysis of the transcription effects of tagged H2A.Bbd, the cells treated with shRNA with no homology to the human genome were used as control. To determine whether the effects observed upon H2A.Bbd depletion were specific to this variant, the cells depleted for another H2A variant associated with active chromatin, H2A.Z, were used for comparison. Analysis of RNA-seq data revealed changes in mRNA splicing patterns upon H2A.Bbd depletion, with the characteristic feature of increased read density within introns (see example in Figure 6A). The frequency of paired-end reads that indicate retention of intronic sequences (those with one end mapping to an exon and the other mapping to an intron) also increases upon H2A.Bbd depletion. These changes are likely to result from a decrease in efficiency of splicing upon H2A.Bbd depletion. We note that the density of intronic reads was nonzero in the control sample in which no variant was depleted; the density increased upon depletion of H2A.Z, but it was much less than the increase in the H2A.Bbd-depleted sample.

To assess how widespread this phenomenon is, we performed a number of genome-wide analyses. First, we confirmed that on average the fraction of intronic reads is significantly higher in the H2A.Bbd-depleted sample compared to both controls (Figure 6B). A more detailed analysis reveals that 1,122 transcripts have significantly increased intronic read density upon H2A.Bbd depletion, which is almost twice as many as 594 transcripts in the case of H2A.Z depletion (Figure S5A). The change in the intronic read density shows moderate correlation with the density of H2A.Bbd tags inside the genes even though
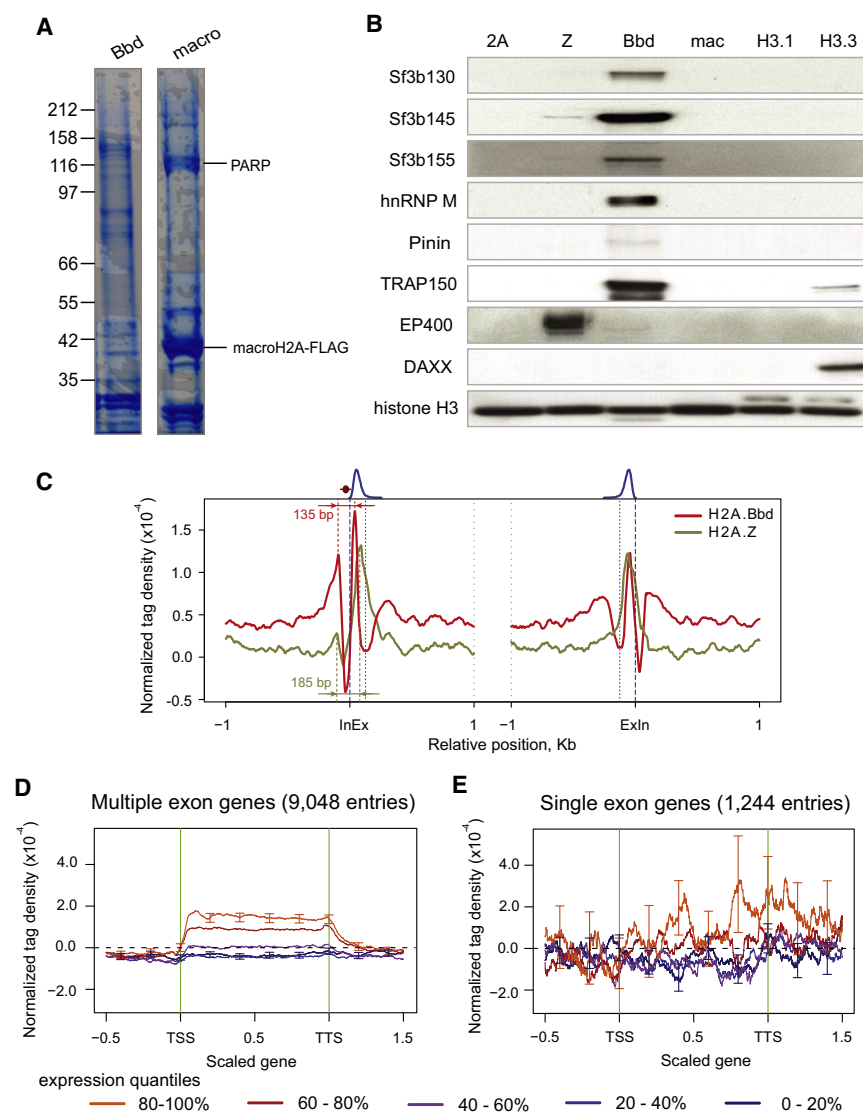
Figure 5. Association of H2A.Bbd with mRNA Processing

(A) Equal amounts of affinity-purified H2A.Bbd and macroH2A chromatin were resolved by SDS-PAGE and stained with Coomassie blue.

(B) Western blots were performed on chromatin affinity purified from the stable cell lines expressing the FLAG-tagged histone indicated at the top (see the Experimental Procedures). Equivalent amounts of chromatin were determined by histone H3 content. Antibodies against spliceosome components are specified on the left for the top five panels. TRAP150 is an H2A.Bbd-enriched protein with roles in both transcription and mRNA processing; EP400 and DAXX westerns indicate specific enrichment of other histone variant purifications.

(C) Average enrichment profiles around intron-exon (InEx) and exon-intron (ExIn) junctions calculated for the histone variants H2A.Bbd (red) and H2A.Z (green). Since H2A.Z shows strong enrichment around TSS, the first three exons of each gene were excluded from the analysis to avoid possible artifacts. Also the last exons were excluded to eliminate effects associated with TTS. The dashed blue lines indicate splice junctions, and dotted blue lines give reference of the median size of the exons included in the analysis. The solid blue lines at the top of the plot represent the distributions of the exon centers. The dark red oval schematically depicts the branchpoint. The distances between the stable nucleosome positions flanking the intron-exon junction are indicated for each histone variant.

(D and E) Average enrichment profiles calculated for the H2A.Bbd variant within genes containing multiple (D) and single (E) exons. Similar to the analysis for all genes (Figure 3), the genes were placed into five groups of equal sizes according to their expression levels (expression quantiles) and filtered for short and overlapping genes. All genes were scaled to the same length and oriented in the same direction. Thin green lines indicate gene starts and ends, positions zero and one, respectively. The numbers of single and multiple exon genes included in the analysis are indicated above the plots. The 95% confidence intervals are shown for the first and fifth expression qualtiles.

the RNA-seq and ChIP experiments were performed in different cell lines (see the Experimental Procedures and discussion above), indicating a direct nature of the effect (Figure S5B).

Second, we characterized each exon by estimating a "spilling" score, defined as the fraction of RNA-seq reads that map partly to this exon and partly to an intron (Figure 6C). Using replicate data, we identified exons for which spilling score significantly changed between the control sample and the samples where either H2A.Bbd or H2A.Z was depleted. Comparison of the scores for these exons showed that depletion of either H2A variant resulted in an increased fraction of exon-intron reads; however, the magnitude of the increase is significantly larger in the case of H2A.Bbd depletion ($p = 10^{-13}$).

Third, we computed the numbers of exon-exon junctions supported by the RNA-seq fragments spanning over the junction

from one exon to another, which represent instances of normal splicing (Figure 6D). Rather than using a threshold for the minimal number of RNA-seq fragments supporting a junction, we compared the results for several threshold values and observed that the numbers of supported junctions significantly decreased upon H2A.Bbd depletion. As before, the effect of H2A.Bbd depletion is significantly more pronounced than that of H2A.Z depletion ($p = 10^{-5}$). Results of all these analyses indicate that histone H2A.Bbd is involved in splicing of introns from precursor mRNAs.

RNA-seq experiment also allows estimation of the exon inclusion rates, which measure the frequency with which exons are included in the mature transcript. Here, we defined the inclusion rate for each exon as the ratio of the number of the tags mapped to that exon relative to the maximal number of tags mapped to
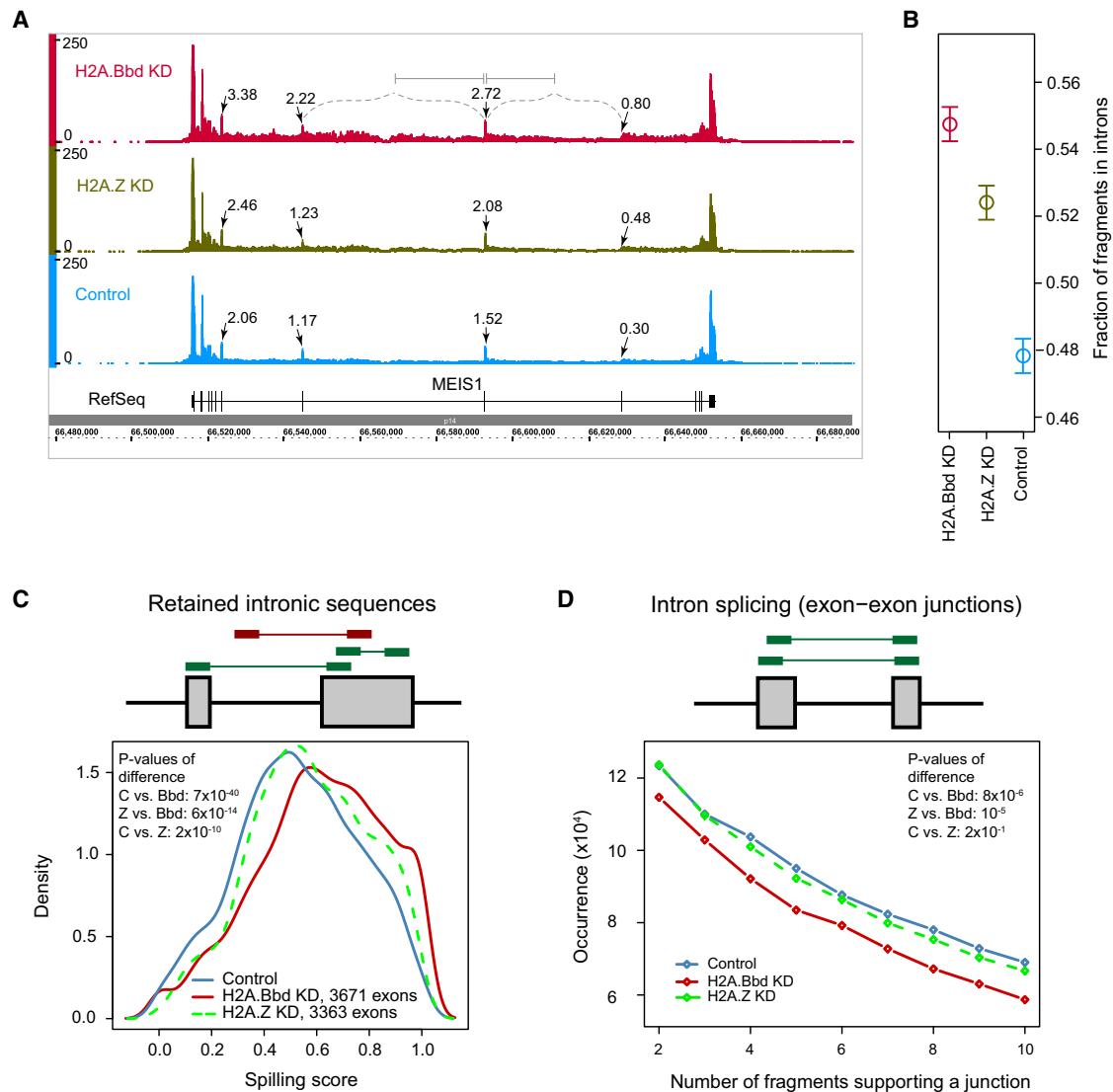
Figure 6. Effect of H2A.Bbd Depletion on mRNA Processing

(A) RNA-seq fragment coverage is shown for the locus on chromosome 2 encompassing gene MEIS1 as an example. The coverage for H2A.Bbd-depleted (red), H2A.Z-depleted (green), and control (blue) samples is shown. Normalized counts of the paired-end RNA-seq fragments (number per million of fragments mapped) spanning from exons to adjacent introns, as illustrated for an exon in the middle, are shown for several internal exons.

(B) Median values (circles) and 95% confidence intervals (arrows) for the fractions of RNA-seq fragments mapped to the introns for the control and histone variant-depleted samples. Fractions were computed for all multiexon genes with nonzero tag density in both intronic and exonic sequences (~15,000 genes). Results obtained for all genes are essentially the same (data not shown). The differences between the distributions for all possible sample pairs are significant ($p < 10^{-7}$).

(C) Distributions of "spilling" scores characterizing aberrant splicing. The score is computed for each exon as the ratio between the number of paired reads that have some part of the read within the exon and at least some part outside of any exon of the gene (red read in the schematic illustration where a gray box indicates an exon) to the total number of reads spanning the exon in question (sum of red and green reads). The score of zero means that all RNA-seq fragments belong only to the exonic region, and the score of one means that all fragments "spill over" from the annotated exon. Only those exons that have statistically significant changes in the spilling score between variant-depleted and control samples were considered (see the Experimental Procedures).

(D) Occurrences of exon-exon junctions that are supported by the presence of RNA-seq fragments (per 50 million read pairs mapped) spanning between exons of the same gene (green reads), which is indicative of intron splicing. Significance of the difference between the results for control and variant-depleted samples was estimated using ANOVA test, and corresponding p values are indicated.

any single exon in the same gene. Using this approach, we detected significant changes in exon inclusion rates between the H2A.Bbd-depleted and control samples and observed that the mean inclusion rate was significantly increased in the H2A.Bbd-depleted sample (Figure S5C). On average, the changes in exon inclusion rates were positively correlated with changes in transcript levels (Figure S5F); however, the transcript level changes were relatively small (less than 2-fold) in most of

the transcripts exhibiting changes in exon inclusion. We also addressed the relationship between presence of H2A.Bbd and changes in exon inclusion rates by comparing the inclusion rates in control and H2A.Bbd-depleted samples for exons grouped according to H2A.Bbd enrichment levels. We observed that the change in inclusion rate upon H2A.Bbd depletion is more pronounced for the exons with higher normal enrichment levels of H2A.Bbd (Figures S5D and S5E). We conclude that the presence of H2A.Bbd affects alternative splicing with a net decrease in exon inclusion.

Taken together, our results provide evidence for functional association of H2A.Bbd chromatin with mRNA processing and, specifically, splicing. In particular, our observations show that H2A.Bbd depletion results in increased read density in intronic sequences and elevated exon inclusion rates. These changes are consistent with a decrease in splicing efficiency, indicating that H2A.Bbd facilitates splicing in unperturbed cells.

DISCUSSION

We report that the properties of the H2A.Bbd chromatin differ from those of bulk chromatin in physical parameters and in an increased association with transcribed genes. The changes we observe in splicing patterns upon H2A.Bbd depletion imply a role for this variant in modulating splicing efficiency. Examples of the impact of H2A.Bbd on chromatin function include the shorter length of nucleosome-protected DNA and the related nucleosomal repeat, corresponding to a noncanonical arrangement in the primary structure of chromatin. The enrichment of H2A.Bbd chromatin in spliceosome components and in the elongating Pol II suggests that H2A.Bbd facilitates a connection between transcription elongation and mRNA processing. Our analysis demonstrated that H2A.Bbd chromatin is associated with the Sf3b complex, a component of the U2 snRNP required for spliceosome function and that H2A.Bbd nucleosomes are arranged in a specific pattern at the intron-exon junction, which may facilitate splice site recognition. Consistent with a role in stimulating spliceosome assembly, depletion of H2A.Bbd resulted in fewer spliced junctions and increased abundance of intronic sequences in the transcripts.

We performed a number of controls and checks to ensure that the observed effects are specific to H2A.Bbd. Since transcription and splicing are known to be linked (Luco et al., 2011; Moore and Proudfoot, 2009), assessing the impact of another histone H2A variant involved in gene activation, H2A.Z, on mRNA processing was important for this purpose. While we observed some association of H2A.Z with Sf3b complex, it was considerably less pronounced than that of H2A.Bbd (Figure 5B). Likewise, the magnitude of the changes in splicing upon H2A.Z depletion was significantly lower than the magnitude of those observed for H2A.Bbd (Figure 6, Figure S5), while H2A.Z and H2A.Bbd depletions were equally strong and H2A.Z depletion resulted in a pronounced downregulation of a set of genes (see the Supplemental Experimental Procedures). Furthermore, the number of splicing junctions directly supported by RNA-seq reads did not change significantly upon H2A.Z depletion (Figure 6D), suggesting that the association of this variant with splicing is relatively weak. This indicates that the

impact of H2A.Bbd on splicing goes beyond what might be expected for a histone variant involved generally in active transcription.

Two hypotheses are consistent with H2A.Bbd-mediated integration of transcription and mRNA processing. First, H2A.Bbd may arrange chromatin structure that facilitates a transcriptionally favorable "open" chromatin state with indirect effects on downstream mRNA processing. In this "nonspecific" model, splicing regulation could be, e.g., a consequence of transit time of RNA Pol II with resultant changes in availability of splice sites (Keren et al., 2010; Kornblihtt et al., 2004; Tilgner et al., 2009). Second, the H2A.Bbd variant may specifically interact with the transcriptional machinery to coordinate both transcription elongation and mRNA processing factors (Luco et al., 2011). The two models are not mutually exclusive, and the changes in splicing that we observe upon H2A.Bbd depletion are consistent with either model. At the same time, the enrichment of H2A.Bbd chromatin with splicing factors favors the latter mechanism. Other studies have demonstrated a connection between chromatin factors and mRNA processing by the association between proteins that bind both nucleosomes and RNA-splicing machinery (Batsché et al., 2006; Luco et al., 2010; Sims et al., 2007). Here, we elaborate upon this idea by showing the involvement of the specific H2A variant in integration of transcription and splicing.

A possible role for nucleosomes in mRNA processing was hypothesized from the observation that typical exon length correlates well with nucleosome size in metazoans (Keren et al., 2010). One interpretation is that nucleosomes may be involved in exon definition, helping the splicing machinery to identify shorter exons from the many kilobases of introns in a gene. This is supported by the observation that nucleosomes display higher density in the exons separated by long rather than short introns (Spies et al., 2009). We note that the average human exon length of ~126 bp is shorter than both the exon length in other metazoans and canonical nucleosome protection size (Schwartz et al., 2009) but agrees with the average protection size of H2A.Bbd nucleosomes reported here. Furthermore, a decrease in splicing events and increased prevalence of intronic sequences in mRNA, as observed in the absence of H2A.Bbd, are phenotypes that are reconcilable with defects in splice site definition. However, the mechanism of such an association is yet to be identified.

The landscape of H2A variants in the genome reflects the local regulation of gene expression, with each variant being associated with different stages of transcriptional regulation. Active chromatin consists of promoters enriched with H2A.Z and gene bodies enriched with H2A.Bbd, both increasing with gene activity. Repressed genes are largely devoid of H2A.Z and H2A.Bbd but may have some association with macroH2A. Concordantly, H2A.Bbd and macroH2A generally display mutual exclusivity along the genome as they do on the inactive X. However, regions demonstrating inclusion of both variants are also apparent, and the presence of macroH2A was detected by mass spectrometry of H2A.Bbd chromatin (see Table S3). Therefore, alternative inclusion of H2A.Bbd and macroH2A is likely a consequence of their association with different regulation pathways rather than of possible physical limitation on

having a chromatin fiber with adjacent H2A.Bbd and macroH2A nucleosomes.

It is intriguing why H2A.Bbd is observed only in mammals and not in invertebrates, especially in light of our findings that H2A.Bbd is associated with transcription and splicing, which are processes occurring in all eukaryotes. While several histone variants are enriched in active genes throughout eukaryotes (Talbert and Henikoff, 2010; Weber et al., 2010), the hypervariabilty of quickly evolving H2A.Bbd might be instrumental during evolution of higher organisms (Gonzalez-Romero et al., 2008). Alternative splicing of mRNA, which increases protein product diversity without increasing the number of genes (Nilsen and Graveley, 2010), has been associated with development of organism complexity. Therefore H2A.Bbd may have coevolved with other pathways to further regulate mRNA splicing important during mammalian evolution. Additional studies will be required to test this hypothesis.

EXPERIMENTAL PROCEDURES

Cell Lines and Chromatin Immunoprecipitation

Stable HeLa cell lines expressing H2A.Bbd-FLAG, macroH2A1.2-FLAG, H2A.Z-FLAG, or H2A-FLAG were constructed as previously described (Nakatani and Ogryzko, 2003; Viens et al., 2006). The cell lines showed no obvious growth defects. Immuno-affinity-purified chromatin was isolated as described (Goldman et al., 2010, see the Supplemental Information for more details, including description of the control for relative expression of the recombinant protein and list of antibodies used). The purified DNA was cloned using standard Illumina sample kit and sequenced on Genome Analyzer I instrument.

ShRNA-Mediated Depletion and RNA-seq Library Preparation

A panel of five different H2A.Bbd (NM_080720) targeted constructs from Sigma-Aldrich (Mission shRNA clones) was assayed for the ability to deplete H2A.Bbd in the HeLa stable cell line after transduction of lentivirus expressing the respective shRNA. Construct TRC106827, which displayed the most potent H2A.Bbd depletion estimated at >90% after 7 days of puromycin selection following transduction, was selected. Construct TRC72585 (NM_002106) was used for H2A.Z with >95% depletion, and nontarget shRNA lentivirus (SHC002) was used as a control. Briefly, after transduction, nuclei were isolated to enrich for immature transcripts using an NP-40 lysis buffer (Pandya-Jones and Black, 2009). Total RNA was purified using the Trizol reagent (Invitrogen) and subsequently depleted for rRNA using the Ribo-zero kit (Epicenter #RZH1046). Purified RNA was then cloned using the Tru-Seq RNA sample kit without any poly(A) enrichment (Illumina).

Data Analysis

A detailed description of the computational procedures is provided in the Supplemental Information. In brief, sequenced tags were mapped to the human genome (hg18) using Bowtie aligner (Langmead et al., 2009) in the case of ChIP-seq data and Tophat in the case of RNA-seq data (Trapnell et al., 2009). Tags were filtered for possible artifacts prior to further analyses. Differential gene expression analysis was performed with bioconductor package DESeq (Anders and Huber, 2010), using RNA-seq data for two biological replicates for each sample (there were at least 85 million paired reads per replicate). Gene Ontology Term Finder web server (http://go.princeton.edu/cgi-bin/GOTermFinder/GOTermFinder) was used for gene ontology analysis. Exon inclusion and intron incorporation rates were estimated based on the normalized numbers of fragments mapped within each interrogated locus. The "spilling" score was computed for each exon as the occurrence of RNA-seq reads that have some part within the exon and at least some part of the read outside of all exons of the corresponding gene, normalized by total number of reads associated with this exon.

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Bao, Y., Konesky, K., Park, Y., Rosu, S., Dyer, P., Rangasamy, D., Tremethick, D., Laybourn, P., and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. EMBO J. 23, 3314–3324.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837.

Batsché, E., Yaniv, M., and Muchardt, C. (2006). The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat. Struct. Mol. Biol. 13, 22–29.

Buschbeck, M., Uribesalgo, I., Wibowo, I., Rue, P., Martin, D., Gutierrez, A., Morey, L., Guigo, R., Lopez-Schier, H., and Di Croce, L. (2009). The histone variant macroH2A is an epigenetic regulator of key developmental genes. Nat. Struct. Mol. Biol. 16, 1074–1079.

Chadwick, B.P., and Willard, H.F. (2001). A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. J. Cell Biol. 152, 375–384.

Corrionero, A., Minana, B., and Valcarcel, J. (2011). Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. Genes Dev. 25, 445–459.

Deal, R.B., Henikoff, J.G., and Henikoff, S. (2010). Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. Science 328, 1161–1164.

Doyen, C.M., Montel, F., Gautier, T., Menoni, H., Claudet, C., Delacour-Larose, M., Angelov, D., Hamiche, A., Bednar, J., Faivre-Moskalenko, C., et al. (2006). Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. EMBO J. 25, 4234–4244.

Folco, E.G., Coil, K.E., and Reed, R. (2011). The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. Genes Dev. 25, 440–444.

Gamble, M.J., Frizzell, K.M., Yang, C., Krishnakumar, R., and Kraus, W.L. (2010). The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. Genes Dev. 24, 21–32.

Gautier, T., Abbott, D.W., Molla, A., Verdel, A., Ausio, J., and Dimitrov, S. (2004). Histone variant H2ABbd confers lower stability to the nucleosome. EMBO Rep. 5, 715–720.

Goldman, J.A., Garlick, J.D., and Kingston, R.E. (2010). Chromatin remodeling by imitation switch (ISWI) class ATP-dependent remodelers is stimulated by histone variant H2A.Z. J. Biol. Chem. *285*, 4645–4651.

Gonzalez-Romero, R., Mendez, J., Ausio, J., and Eirin-Lopez, J.M. (2008). Quickly evolving histones, nucleosome stability and chromatin folding: all about histone H2A.Bbd. Gene *413*, 1–7.

Gozani, O., Feld, R., and Reed, R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev. *10*, 233–243.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat. Genet. *41*, 941–945.

John, S., Sabo, P.J., Johnson, T.A., Sung, M.-H., Biddie, S.C., Lightman, S.L., Voss, T.C., Davis, S.R., Meltzer, P.S., Stamatoyannopoulos, J.A., et al. (2008). Interaction of the glucocorticoid receptor with the chromatin landscape. Mol. Cell *29*, 611–624.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nat. Rev. Genet. *11*, 345–355.

Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004). Multiple links between transcription and splicing. RNA *10*, 1489–1498.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lohr, D., Corden, J., Tatchell, K., Kovacic, R.T., and Van Holde, K.E. (1977). Comparative subunit structure of HeLa, yeast, and chicken erythrocyte chromatin. Proc. Natl. Acad. Sci. USA *74*, 79–83.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science *327*, 996–1000.

Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., and Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. Cell *144*, 16–26.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Malik, H.S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat. Struct. Biol. *10*, 882–891.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2007). Histone replacement marks the boundaries of cis-regulatory domains. Science *315*, 1408–1411.

Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell *136*, 688–700.

Nakatani, Y., and Ogryzko, V. (2003). Immunoaffinity purification of mammalian protein complexes. Methods Enzymol. *370*, 430–444.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature *463*, 457–463.

Ouararhni, K., Hadj-Slimane, R., Ait-Si-Ali, S., Robin, P., Mietton, F., Harel-Bellan, A., Dimitrov, S., and Hamiche, A. (2006). The histone variant mH2A1.1 interferes with transcription by down-regulating PARP-1 enzymatic activity. Genes Dev. *20*, 3324–3336.

Pandya-Jones, A., and Black, D.L. (2009). Co-transcriptional splicing of constitutive and alternative exons. RNA *15*, 1896–1908.

Schwartz, B.E., and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev. *19*, 804–814.

Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. Nat. Struct. Mol. Biol. *16*, 990–995.

Sims, R.J., Millhouse, S., Chen, C.-F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. Mol. Cell *28*, 665–676.

Spies, N., Nielsen, C.B., Padgett, R.A., and Burge, C.B. (2009). Biased chromatin signatures around polyadenylation sites and exons. Mol. Cell *36*, 245–254.

Talbert, P.B., and Henikoff, S. (2010). Histone variants—ancient wrap artists of the epigenome. Nat. Rev. Mol. Cell Biol. *11*, 264–275.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., and Guigó, R. (2009). Nucleosome positioning as a determinant of exon recognition. Nat. Struct. Mol. Biol. *16*, 996–1001.

Tolstorukov, M.Y., Kharchenko, P.V., Goldman, J.A., Kingston, R.E., and Park, P.J. (2009). Comparative analysis of H2A.Z nucleosome organization in human and yeast genome. Genome Res. *19*, 967–977.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Viens, A., Mechold, U., Brouillard, F., Gilbert, C., Leclerc, P., and Ogryzko, V. (2006). Analysis of human histone H2AZ deposition in vivo argues against its direct role in epigenetic templating mechanisms. Mol. Cell. Biol. *26*, 5325–5335.

Weber, C.M., Henikoff, J.G., and Henikoff, S. (2010). H2A.Z nucleosomes enriched over active genes are homotypic. Nat. Struct. Mol. Biol. *17*, 1500–1507.

Wu, R.S., and Bonner, W.M. (1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. Cell *27*, 321–330.

Zhou, J., Fan, J., Rangasamy, D., and Tremethick, D. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. Nat. Struct. Mol. Biol. *14*, 1070–1076.