HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2011 June 29.

Published in final edited form as:

Nat Genet. 2009 August ; 41(8): 941–945. doi:10.1038/ng.409.

H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions in the human genome

Chunyuan Jin[1,*], **Chongzhi Zang**[2,*], **Gang Wei**[3], **Kairong Cui**[3], **Weiqun Peng**[2], **Keji Zhao**[3,**], and **Gary Felsenfeld**[1,**]

[1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[2]Department of Physics, The George Washington University, Washington, DC 20052, USA

[3]Laboratory of Molecular Immunology, The National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA

Abstract

To understand how chromatin structure is organized by different histone variants, we have measured the genome-wide distribution of NCPs (nucleosome core particles) containing the histone variants H3.3 and H2A.Z. We find a special class of NCPs containing both variants, enriched at 'nucleosome-free regions' of active promoters, enhancers and insulator regions. We show that previous preparative methods resulted in loss of these unstable double variant NCPs. This instability should facilitate the accessibility of transcription factors to promoters and other regulatory sites *in vivo*. Other combinations of variants have different distributions, consistent with distinct roles for the histone variants in modulation of gene expression.

Precise global mapping of histone variants is indispensable for understanding the interaction between chromatin structure and gene expression. Genome-wide surveys of the distribution of individual histone variants H2A.Z or H3.3 have revealed that they are widely distributed in the genome1-11. However, no one has attempted a genome-wide study of the distribution of individual nucleosome core particles (NCPs) that contain both variants; it has therefore not been possible to distinguish NCPs carrying only one of these variants from those carrying both. This is particularly important because we have shown in a previous study 12 that H3.3/H2A.Z NCPs are unusually unstable under conditions normally used in

**To whom correspondence may be addressed. Laboratory of Molecular Immunology, NHLBI, NIH, Building 10, Room 7B06A , 10 Center Dr, Bethesda, MD 20892-0540, USA, zhaok@nhlbi.nih.gov, Telephone: 1-301-496-2098. . **To whom correspondence may be addressed. Laboratory of Molecular Biology, NIDDK, NIH, Building 5, Room 212, 9000 Rockville Pike, Bethesda, MD 20892-0540, USA, . gary.felsenfeld@nih.gov, Telephone: 1-301-496-4173. .

*These authors contribute equally to this work.

AUTHOR CONTRIBUTIONS C.J. and G.F. designed the experiments and C.J. carried them out, C.Z. and W.P. performed computational analyses, G.W. did template preparation for Solexa sequencing, K.C. contributed to the study, C.J., C.Z., W.P., K.Z. and G.F. analyzed the data, C.J., C.Z., W.P. and G.F. wrote the paper. K.Z. and G.F. directed the study.

ACCESSION CODES NCBI Short Read Archive: raw sequence tags for H2A.Z, H2A.Z (high salt), H3.3, *Double* (H3.3/H2A.Z), Input and Genomic DNA in HeLa cells have been deposited with accession code GSE13308.

preparations for such studies; it therefore seemed possible that preferential loss could account for earlier reports3,4,7,8,11,13 that H2A.Z-containing nucleosomes are absent from ‘nucleosome-free regions’13-22 at transcription start sites (TSSs) of active genes. With the use of isolation procedures that preserve the stability of the NCPs containing double variant, it becomes possible to carry out a genome-wide survey of each of these kinds of variant NCPs.

To determine the distribution of different combinations of histone variants, we prepared monomer NCPs from a HeLa cell line expressing FLAG-tagged human histone H3.3 23, and carried out individual or sequential immunopurification followed by high throughput Solexa sequencing4 to obtain genome-wide high resolution profiles for *total* H2A.Z, *total* H3.3 or *double* (H3.3/H2A.Z) NCPs (Supplementary Fig. 1 online). From these libraries, it was also possible using computational analysis to deduce the relative profiles of NCPs carrying H2A.Z *only* (in combination with H3.1 or H3.2 but not H3.3) and NCPs containing H3.3 *only* (not in combination with H2A.Z) (see Methods). H3.3/H2A.Z NCPs are disrupted by exposure to moderate salt concentrations12; we therefore carried out all of the purifications in low ionic strength solvents except as indicated. The mononucleosomes that we used as input in our study reflect the bulk of the genome (Supplementary Fig. 2 online).

We first investigated the distribution of histone variants around genomic TSSs. To correlate the distribution with gene expression, we created separate profiles containing 1000 genes each for highly expressed, intermediately expressed and silent genes. The data show that H3.3, H2A.Z, and H3.3/H2A.Z NCPs are selectively enriched at TSSs of active genes (Fig. 1a-c). Only a small fraction of H2A.Z *only* and almost none of H3.3 *only* NCPs are detected at such sites (Fig. 1d,e). The results for H3.3 and H2A.Z separately are apparently at variance with high resolution (mononucleosome level) studies, which have indicated that sites immediately upstream of the TSS of active genes tend to be generally depleted of H2A.Z NCPs and to a lesser extent of H3.3 NCPs1,3,4,7,8,13. Since H3.3/H2A.Z NCPs are easily disrupted 12, and these comprise a large fraction of *total* H2A.Z NCPs at TSS (compare Fig. 1a and 1d), it seemed possible that when isolated at higher salt concentrations they would be under-represented. As we anticipated, the second genome-wide screen, using NCPs prepared under conditions which exposed them to 150 mM NaCl, showed a relative minimum of H2A.Z abundance at the TSS, reproducing the earlier findings (Fig. 1f). We conclude that underrepresentation of H2A.Z-containing NCPs at TSS can arise from preferential disruption of H3.3/H2A.Z NCPs.

We further carried out an analysis of positioning for all NCPs containing H2A.Z, making use of tags on both strands to determine accurately the boundaries of each NCP13. Consistent with published data, NCPs prepared in 150 mM NaCl show a 200 bp region depleted of H2A.Z NCPs immediately upstream of the TSS (−1 nucleosome), whereas in the surrounding region four phased nucleosomes are detected (from −2 to +3) (Fig. 1g and Supplementary Fig. 3 online). In contrast, the low salt preparation clearly reveals the enrichment of H2A.Z NCPs at the −1 position; the peaks in the region corresponding to −1 and −2 nucleosomes are not well ordered (Fig. 1h). The observed irregular patterns are entirely consistent with a population of sites in which one or two NCPs can occupy any of several positions in this ~400 bp region (Supplementary Fig. 4 online). Individual active

genes also displayed similar changes at TSS (Fig. 1i). It should be noted that these previously undetected NCPs carry both H3.3 and H2A.Z.

Next, we examined the distribution over other regulatory elements, including CTCF-binding sites, which typically represent regions with insulator activity 24, and DNase I hypersensitive sites, typically associated with the centers of regulatory activity25. *Total* H2A.Z is enriched at the center of the intergenic CTCF-binding sites26 (Fig. 2a). A small number of H2A.Z *only* NCPs (less than 20% of total) contribute to this enrichment. Interestingly, *total* H3.3 also had its highest peak at the sites, but again only one fifth of them are H3.3 *only* NCPs (Fig. 2a), suggesting that the majority of NCPs at the center of the binding sites are the H3.3/H2A.Z double variant. This is confirmed by the profile for *double* (H3.3/H2A.Z) NCPs (Fig. 2b). We next examined H2A.Z nucleosome positioning around CTCF-binding sites. Under the low salt conditions, the two highest peaks for both 5′ tags and 3′ tags are observed at the center of the binding sites (Fig. 2c and Supplementary Fig. 5 online). However, these two peaks are nearly missing (Fig. 2d) under higher salt conditions and the pattern is now quite similar in many respects to the one reported earlier 27, which showed a nucleosome-free gap at the binding sites surrounded by an ordered array of H2A.Z NCPs . These results reveal the presence of H2A.Z nucleosomes, largely H3.3/H2A.Z NCPs, at this “nucleosome-free” region. The distribution of nucleosome levels around CTCF-binding sites (Supplementary Fig. 6 online) in low salt condition indicates that a single H2A.Z NCP can bind in several different positions within the CTCF-binding region, a pattern resembling that seen at TSS sites. A survey of intergenic ENCODE DNase I hypersensitive sites28,29 reveals high concentrations of *total* H3.3 and *total* H2A.Z (Fig. 2e), whereas there is only a small enrichment of NCPs containing H3.3 or H2A.Z alone; the *double* variant NCP predominates. H3.3/H2A.Z NCPs are not detectable in HeLa cells at sites that are DNase I hypersensitive in CD4+ T cells but not in HeLa (Fig. 2f). This shows that the presence of the unstable NCPs reflects the activity of the hypersensitive sites, which also carry histone modifications correlated with enhancer activity (K.C., C.Z., D.S., W.P. and K.Z., unpublished data). Taken together, H3.3/H2A.Z NCPs mark ‘nucleosome-free regions’ of active promoters as well as enhancers and insulator regions.

We then examined patterns of distributions of histone variants at the transcription termination sites (TTSs). The abundance of *total* H2A.Z near the TTS is low, nearly uniform and almost independent of gene activity (Fig. 3a). In accordance with previous observations in the *Drosophila* genome1, over the most active genes H3.3 abundance reaches a broad peak around TTS and then decreases on either side (Fig. 3b). *Double* variant NCPs rise slightly in abundance 3′ of the TTS of the more active genes (Fig 3c). These may function in transcriptional termination, antisense transcription or antisilencing. There is a narrow local minimum at the TTS in the H3.3, H2A.Z and H3.3/H2A.Z distributions (Fig. 3b and Supplementary Fig. 7a-f online). Similar patterns are seen with the input sample of NCPs (before immunoprecipitation) and total genomic DNA (Supplementary Fig. 7g,h), suggesting that these very low level signals are an artifact associated with TTS sequences, and should be taken into account in analyses of this kind.

To characterize the distributions of histone variants across entire genes, we displayed our data on a normalized distance scale with the TSS set at 0 and the TTS at 1, and with a

compressed scale for the regions around the TSS and TTS. Of all the H2A.Z containing NCPs near the TSS of active genes, the majority carry both H3.3 and H2A.Z (Fig. 4a,b). There is a slight but consistent elevation of H2A.Z *only* particles over the gene bodies and downstream of TTS of the silent gene population (Supplementary Fig. 8 online). *Total* H3.3 NCPs shows a gradient of increasing abundance from 5′ to 3′ over the entire transcribed regions of active genes (Fig. 4c), reminiscent of the distribution of histone H3 lysine 36 trimethylation in active genes4,19. Interestingly, NCPs containing *only* H3.3 are almost completely absent from TSS (Fig. 4d), showing that in this region H3.3, when present, is almost always partnered with H2A.Z. In contrast, the pattern and the density of NCPs containing *only* H3.3 within the gene bodies are very close to those of total H3.3, indicating that the majority of H3.3 NCPs over transcribed regions carry the single variant H3.3 but not H2A.Z. We note that NCPs containing the single variant H3.3 are still relatively unstable compared to canonical NCPs12 and might accommodate the passage of RNA polymerase. *Double* variant NCPs are enriched over the TSS and at a relatively low abundance near TTS, both correlated with transcriptional level (Fig. 4e). Some of these particles are also present in gene bodies, but at quite low concentrations (see Supplementary Fig. 9,10 online). The presence of these NCPs over transcribed regions might facilitate chain elongation of Pol II and/or the rapid loss of nucleosomes over some gene bodies, perhaps of immediately inducible genes, a phenomenon seen at *Hsp70* loci 30.

As we show here, the distribution of the unstable double variants is distinct and quite different from the distributions of NCPs carrying either H3.3 or H2A.Z alone. The ‘nucleosome-free region’ of active promoters is likely to be occupied to a considerable extent by the labile H3.3/H2A.Z NCPs (Supplementary Note online). These unstable NCPs could serve as ‘place holders’ to prevent the region from being covered by adjacent quite stable (canonical) NCPs and/or nonspecific factors, as might occur if the region was completely free of nucleosomes. At the same time, because of their relative instability the H3.3/H2A.Z NCPs could more easily be displaced by transcription factors. Our results suggest a new model for the chromatin structure at vertebrate promoters and other regulatory sites, in which the site is dynamically cycling between occupancy by these unstable nucleosomes, or by transcription factors, or perhaps by some canonical nucleosomes if the site is temporarily silent or has not yet been replaced by variant histone after replication. For some small fraction of the time the site may also be vacant during the period in which these components are exchanging places (Fig. 5). Which of these states is detected will depend on the measurement method, but they are all part of the promoter structure, in which the double variant H3.3/H2A.Z NCP appears to play an important role.

Each combination of histone variants gives rise to a distinct and characteristic nucleosome stability31. It is not yet clear which of these differences in stability arise from differences in amino acid composition, and which are caused by a combination of histone modifications unique to each variant. Our present results clearly show, however, that each variant or combination of variants has a highly specific pattern of distribution *in vivo*, suggesting that these differences in stability are elaborately exploited in the regulation of gene expression.

METHODS

Cell lines

H3.3 HeLa S3 cell lines, stably expressing the H3.3 fused with C-terminal FLAG- and HA-epitope tags, were grown as described23. The fusion gene is driven by an MMLVLTR promoter23,32. The tagged H3.3 is about one third of total histone H3 in expression level (Supplementary Fig. 11). There has been extensive experience with the use of H3.3-FLAG and H3.1-FLAG12,23,33-35. The FLAG tag does not appreciably perturb nucleosome behavior: nucleosomes containing these tagged histones show distinctive properties that resemble those of the corresponding endogenous nucleosomes12,23,34,35.

Isolation of mononucleosomes

Nuclei were isolated with a cell lysis buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, and 0.4% NP-40. All buffers were supplemented with 10 mM Nabutyrate, 0.5 μg/mL aprotinin, 0.5 μg/mL leupeptin, and 1 μg/mL aprotinin. We previously showed that the instability of the variant histones seemed unaffected by acetylation12, but butyrate is used during preparation to mimic the *in vivo* state as closely as possible. Nuclei were pelleted and resuspended in the same buffer plus 1 mM $CaCl_2$. The A_{260} was adjusted to 1.25, and the resuspended nuclei were digested with 12×10^{-2} U/μL MNase (Worthington) for 10 min at 37°C to generate mostly mononucleosomes. The reaction was stopped by adding EDTA (pH 8.0) to a final concentration of 10 mM, and the suspension was centrifuged at 2500 rpm for 5 min, retaining supernatant S1. The pellet was resuspended in lysis buffer plus 0.25 mM EDTA, incubated on ice for 15 min, and recentrifuged at 10,000 rpm for 10 min after passing four times through a 20-gauge needle followed by four passes through a 25-gauge needle. The supernatant S2 was combined with S1. Mononucleosomes were then purified on a 5%–30% sucrose gradient containing 10 mM NaCl, 10 mM Tris-HCl (pH 7.4), and 0.2 mM EDTA. To prepare the mononucleosomes exposed to a high salt, S1 and S2 were incubated with 150 mM NaCl for 20 min at 4°C before loading them on the sucrose gradient centrifugation.

ChIP and double ChIP

ChIP and Double ChIP were carried out as described elsewhere with minor modifications 12. For double ChIP, FLAG-H3.3-containing mononucleosomes were isolated first by anti-FLAG antibody immuno-affinity gel purification and then subjected to the second immunoprecipitation with anti-H2A.Z antibodies. Anti-FLAG M2 affinity gel (A2220, Sigma) and anti-H2A.Z antibodies (07-594, Millipore) were used for ChIP and Double ChIP.

Template preparation for Solexa sequencing analysis

Genomic DNA was purified from H3.3 HeLa S3 cells and fragmented by sonication to an average size of ~200-300 bp. The genomic DNA, mononucleosomal DNA (Input) and the ChIP DNA ends were repaired using PNK and Klenow enzyme (ER0720, Epicenter Biotechnology), followed by the treatment with Taq polymerase to generate a protruding 3′ A base used for the adaptor ligation. Following the ligation of a pair of Solexa adaptors to

the repaired ends, the adapter-ligated DNAs were amplified using the adaptor primers for 18 cycles and the corresponding fragments were isolated from the agarose gel. The purified DNA was used directly for the cluster generation and the sequencing analysis using the Solexa 1G Genome Analyzer following the manufacturer's protocols.

Solexa pipeline analysis

Sequence tags of mostly 25 bp were obtained using the Solexa Analysis Pipeline. All tags were mapped to the human genome (hg18) and only uniquely matching reads were retained. Unique tag numbers for each sample are listed in Supplementary Table 1. The output of the Solexa Analysis Pipeline was converted to browser extensible data (BED) files detailing the genomic coordinates of each tag. In all analysis, tags with multiple identical copies were trimmed to be below five copies to reduce potential PCR amplification bias.

Identification of ChIP enriched regions

In preference to measuring local enrichment, we employed SICER (C.Z., D.S., C.Z., K.C., K.Z. and W.P., unpublished data), an algorithm that identify ChIP enriched regions by looking for clusters of windows occupied by tags unlikely to appear by chance. In this method, the genome was partitioned into non-overlapping mono-nucleosomal summary windows of 200 bp. The number of tags in each 200 bp summary window was counted, with the location of each Watson (Crick) tag shifted by +75 bp (−75 bp) from its 5′ start to represent the center of the DNA fragment associated with the tag. The windows exhibiting ChIP enrichment (*p*-value of 0.2 based on a Poisson background model) were then identified. Islands were defined as clusters of enriched windows allowing gaps of at most two unenriched windows. To ensure high confidence, ChIP-enriched regions were identified as islands whose tag-counts were above a threshold determined by a very stringent *E*-value (i.e., the expected number of islands whose tag-counts are above the threshold under a background model of random tags) requirement of 0.1.

Identification of regions enriched with specific combination of variants

The summary windows on the significant islands identified using the method described above for the libraries for H2A.Z, H3.3 and *Double* (H3.3/H2A.Z) were compared. The 'H2A.Z (H3.3) *only*' contained tags in the island-filtered summary windows on H2A.Z (H3.3) that do not overlap with any significant islands from H3.3 (H2A.Z) and *Double*. Results for 'H2A.Z *only*' and 'H3.3 *only*' obtained with our approach are further supported by the following evidence: For the 'H3.3 *only*' profile (Fig.4d) produced by this method, the salient feature is the increasing enrichment of 'H3.3 *only*' toward the 3′ end. This feature can already be seen by direct comparison of the profile of H3.3 (Fig.4c) and *Double* (Fig.4e) at the gene body region, where H3.3 abundance increases whereas *Double* remains flat. For 'H2A.Z *only*', the SICER result implies that there are genes that are enriched with only H2A.Z around TSSs. Indeed, there are 924 genes 1) whose TSSs overlap with H2A.Z islands but do not overlap with H3.3 and *Double* islands, and 2) whose gene bodies and TTSs have no enrichment of any variants. Supplementary Figure 12 provides such an example.

The gene set

UCSC old known genes36 were obtained from UCSC genome browser. They were mapped to Affymetrix U133P2 probe IDs using the table provided in the UCSC genome browser. Genes without corresponding U133P2 ID were ignored. If multiple genes map to the same U133P2 ID, only one was retained. A total of 20,444 genes results after further removal of two genes in chromosome M. These genes were ranked according to their expression values in HeLa S3 cells obtained with the Affymetrix U133P2 microarray. The three sets of genes, 1-1,000, 9,001-10,000 and 19,445-20,444, were chosen as the highly expressed, intermediately expressed and silent genes.

The distribution of histone variants at specific sets of genomic landmarks

To examine the profile of histone variant around a set of transcription start sites (TSSs), the TSSs were aligned. Tags in non-overlapping windows of 20 bp were tallied in the set of TSSs. The total tag counts were normalized by the numbers of genes in each set and by the window size. An identical method was applied to transcription end sites (TTSs) and CTCF-binding sites. HeLa S3 intergenic CTCF-binding sites (CTCF sites away from promoters and gene bodies) were obtained from published data26. In Figure 1a-f, Figure 2a,b,e,f and Figure 3a-c, island-filtered 5′ tags were used. The profile of H2A.Z (‘H2A.Z *only*’ also) and H3.3 (‘H3.3 *only*’) were further normalized by the total tag numbers of island-filtered tags in H2A.Z and H3.3 library, respectively, while the profile of *Double* (H3.3/H2A.Z) was normalized by total island-filtered tags in *Double* library.

The distribution of histone variants at Encode DNase I hypersensitive sites

The Encode DNase I HS sites for HeLa S3 cells were downloaded from UCSC genome browser28,37 . All DNase I HS in the intergenic regions (away 1kb from TSS) were aligned and normalized to the same length, and partitioned into twenty blocks. Island-filtered tags in each block were tallied and normalized by the total number of base pairs in each block. Outside the DNase I HS region, island-filtered tags were tallied in the 2 kb upstream and downstream in 50 bp windows and normalized similarly. In the end, the profile was also normalized by the total number of island-filtered tags in each sample.

Nucleosome positioning analysis

All tags were used for nucleosome positioning analysis. For TSS (Fig. 1g,h), the number of 5′ tags and 3′ tags were separately counted in 20 bp windows surrounding each site. The counts from all sites were added up and normalized by the total number of sites, the window size and the total number of tags in the library. Similar analysis was applied to TTS (Supplementary Fig. 7d-h) as well as intergenic CTCF-binding sites (Fig. 2c,d).

Profiles of tag density in and around genes

For each gene, island-filtered tags were summed according to their shifted positions in 1 kb windows for the regions from 5 kb upstream of the transcription start site (txStart) to the txStart and from the transcription end site (txEnd) to 5 kb downstream. Within the gene bodies, island-filtered tags were summed according to their shifted positions in windows equal to 5% of the gene length. In Figure 4, all window tag-count were normalized by the

total number of bases in the windows, and the total number of island-filtered tags in the corresponding sample (Fig. 4a,b by H2A.Z; Fig. 4c,d by H3.3; Fig. 4e by itself) to obtain normalized tag density.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank H. Tagami, Y. Nakatani and G. Almouzni for Flag/HA H3.3 cells, Tae-young Roh, Dustin E. Schones, Pavel Khil for Solexa pipeline analysis, and Shaila Sharmeen, George Poy for Solexa sequencing. We also acknowledge members of the Felsenfeld laboratory for criticism of the manuscript. This research was supported by the Intramural Research Programs of the National Heart, Lung, and Blood Institute and the National Institute of Diabetes and Digestive and Kidney Diseases.

REFERENCES

1. Mito Y, Henikoff JG, Henikoff S. Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet. 2005; 37:1090–7. [PubMed: 16155569]
2. Mito Y, Henikoff JG, Henikoff S. Histone replacement marks the boundaries of cis-regulatory domains. Science. 2007; 315:1408–11. [PubMed: 17347439]
3. Albert I, et al. Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. Nature. 2007; 446:572–6. [PubMed: 17392789]
4. Barski A, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007; 129:823–37. [PubMed: 17512414]
5. Creyghton MP, et al. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell. 2008; 135:649–61. [PubMed: 18992931]
6. Li B, et al. Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. Proc Natl Acad Sci U S A. 2005; 102:18385–90. [PubMed: 16344463]
7. Mavrich TN, et al. Nucleosome organization in the Drosophila genome. Nature. 2008; 453:358–62. [PubMed: 18408708]
8. Raisner RM, et al. Histone variant H2A.Z marks the 5′ ends of both active and inactive genes in euchromatin. Cell. 2005; 123:233–48. [PubMed: 16239142]
9. Zhang H, Roberts DN, Cairns BR. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. Cell. 2005; 123:219–31. [PubMed: 16239141]
10. Whittle CM, et al. The genomic distribution and function of histone variant HTZ-1 during C. elegans embryogenesis. PLoS Genet. 2008; 4:e1000187. [PubMed: 18787694]
11. Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res. 2008
12. Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev. 2007; 21:1519–29. [PubMed: 17575053]
13. Schones DE, et al. Dynamic regulation of nucleosome positioning in the human genome. Cell. 2008; 132:887–98. [PubMed: 18329373]
14. Boeger H, Griesenbeck J, Strattan JS, Kornberg RD. Nucleosomes unfold completely at a transcriptionally active promoter. Mol Cell. 2003; 11:1587–98. [PubMed: 12820971]
15. Reinke H, Horz W. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. Mol Cell. 2003; 11:1599–607. [PubMed: 12820972]
16. Bernstein BE, Liu CL, Humphrey EL, Perlstein EO, Schreiber SL. Global nucleosome occupancy in yeast. Genome Biol. 2004; 5:R62. [PubMed: 15345046]

17. Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD. Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat Genet. 2004; 36:900–5. [PubMed: 15247917]
18. Roh TY, Ngau WC, Cui K, Landsman D, Zhao K. High-resolution genome-wide mapping of histone modifications. Nat Biotechnol. 2004; 22:1013–6. [PubMed: 15235610]
19. Pokholok DK, et al. Genome-wide map of nucleosome acetylation and methylation in yeast. Cell. 2005; 122:517–27. [PubMed: 16122420]
20. Yuan GC, et al. Genome-scale identification of nucleosome positions in S. cerevisiae. Science. 2005; 309:626–30. [PubMed: 15961632]
21. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. Cell. 2007; 130:77–88. [PubMed: 17632057]
22. Ozsolak F, Song JS, Liu XS, Fisher DE. High-throughput mapping of the chromatin structure of human promoters. Nat Biotechnol. 2007; 25:244–8. [PubMed: 17220878]
23. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell. 2004; 116:51–61. [PubMed: 14718166]
24. West AG, Gaszner M, Felsenfeld G. Insulators: many functions, many mechanisms. Genes Dev. 2002; 16:271–88. [PubMed: 11825869]
25. Felsenfeld G, Groudine M. Controlling the double helix. Nature. 2003; 421:448–53. [PubMed: 12540921]
26. Cuddapah S, et al. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. 2009; 19:24–32. [PubMed: 19056695]
27. Fu Y, Sinha M, Peterson CL, Weng Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. PLoS Genet. 2008; 4:e1000138. [PubMed: 18654629]
28. Crawford GE, et al. DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. Nat Methods. 2006; 3:503–9. [PubMed: 16791207]
29. Birney E, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007; 447:799–816. [PubMed: 17571346]
30. Petesch SJ, Lis JT. Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. Cell. 2008; 134:74–84. [PubMed: 18614012]
31. Henikoff S. Nucleosome destabilization in the epigenetic regulation of gene expression. Nat Rev Genet. 2008; 9:15–26. [PubMed: 18059368]
32. Nakatani Y, Ogryzko V. Immunoaffinity purification of mammalian protein complexes. Methods Enzymol. 2003; 370:430–44. [PubMed: 14712665]
33. Polo SE, Roche D, Almouzni G. New histone incorporation marks sites of UV repair in human cells. Cell. 2006; 127:481–93. [PubMed: 17081972]
34. Jin C, Felsenfeld G. Distribution of histone H3.3 in hematopoietic cell lineages. Proc Natl Acad Sci U S A. 2006; 103:574–9. [PubMed: 16407103]
35. Loyola A, Bonaldi T, Roche D, Imhof A, Almouzni G. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell. 2006; 24:309–16. [PubMed: 17052464]
36. Karolchik D, et al. The UCSC Table Browser data retrieval tool. Nucleic Acids Res. 2004; 32:D493–6. [PubMed: 14681465]
37. Brenner S, et al. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000; 18:630–4. [PubMed: 10835600]

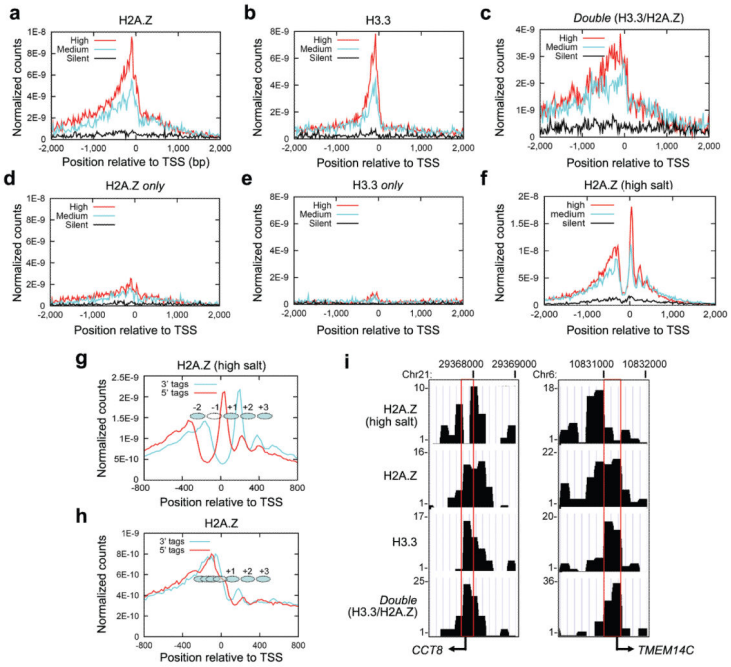

Figure 1.

H3.3/H2A.Z NCPs mark ‘nucleosome-free regions’ of active promoters.

Tags in non-overlapping 20 bp windows relative to the aligned transcription start sites (TSSs) were tallied in the gene set. The tag counts were normalized by the total numbers of bases (i.e. 20 multiplied by the number of genes in the gene set). Island-filtered 5′ tags were used in (a-f) and the profiles were further normalized by the total number of island-filtered tags in the library. All tags were used in (g,h) and the profiles were further normalized by the total number of tags in the library.

(a-e) Profiles of histone variants indicated above each panel across the TSS for 1000 highly active (red), intermediately active (cyan) and silent genes (black) are shown (see Methods).

(f) Profile of H2A.Z-containing NCPs isolated in high salt across the TSS for 1000 highly active (red), intermediately active (cyan) and silent genes (black) are shown.

(g,h) The H2A.Z nucleosome positioning near the TSS at high (g) or low salt (h). The y axis shows the normalized counts of sequenced tags from the upper strand and the lower strand of the DNA at each position, represent 5′ and 3′ boundaries of each NCP. ‘Open oval’ represents depleted NCP; ‘filled oval’ indicates phased NCP.

(i) Two typical examples of histone variants patterns at high resolution at TSSs of two active genes, shown as custom tracks on the UCSC genome browser. Both active genes *CCT8* and *TMEME14C* have high levels of H3.3/2A.Z NCPs at the TSS (lower three panels). The loss of these NCPs after exposure to high salt (top panels) is evident (red rectangles).

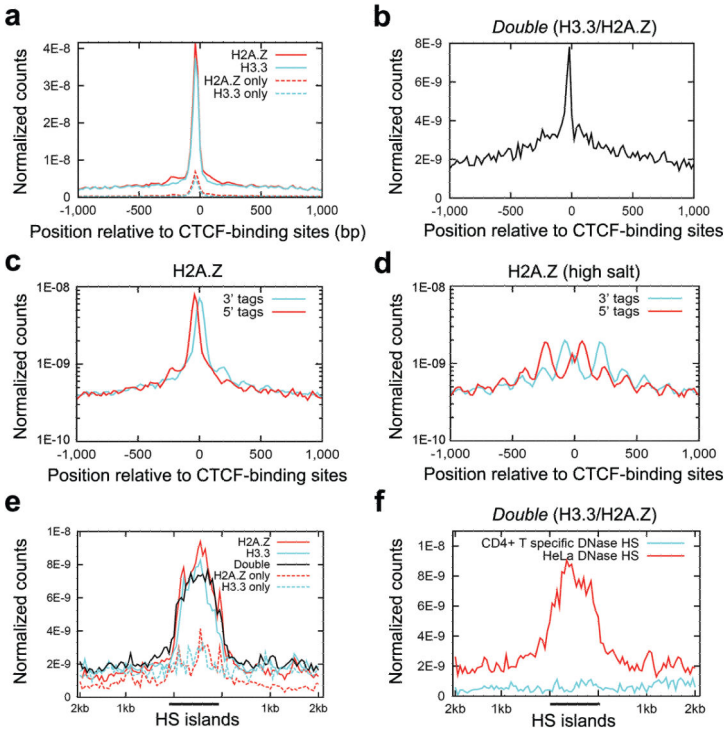

Figure 2.

H3.3/H2A.Z NCPs enriched at other regulatory elements.

(a,b) Histone variants at intergenic CTCF-binding sites. Methods used here were similar to Fig 1a-f. The H2A.Z (as well as 'H2A.Z *only*') and H3.3 (as well as 'H3.3 *only*') were normalized by the total tag numbers of island-filtered tags in H2A.Z and H3.3 libraries, respectively, while the profile of *Double* (H3.3/H2A.Z) was normalized by total island-filtered tags in the *Double* library.

(c,d) Averaged H2A.Z nucleosome positioning near the CTCF-binding sites at low (c) or high (d) salt shown by the sequenced 5′ (red) and 3′ (cyan) tags, representing the 5′ and 3′ boundary of each NCP. Method was similar to that described in Fig 1g,h, except the data is plotted as logarithmic curve.

(e) Histone variants at ENCODE DNase I hypersensitive sites. All DNase I HS sites were aligned and normalized to the same length, and partitioned into twenty blocks. Island-filtered tags in each block were tallied and normalized by the total number of bases in each block. Outside the DNase I HS sites, island-filtered tags were tallied in 50 bp windows in the 2 kb upstream and downstream regions and normalized similarly. At the end, the profile was also normalized by the total number of island-filtered tags in each sample.

(f) In HeLa cells, H3.3/H2A.Z NCPs are only enriched at HeLa DNase I hypersensitive sites (red) but not at sites (cyan) that are DNase I hypersensitive in CD4+ T cells but not in HeLa.

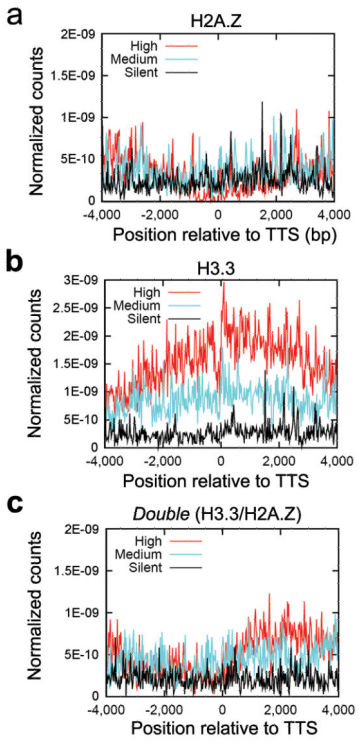

Figure 3.

Histone variants near transcription termination sites (TTSs). Method was the same as used for Fig 1a-f.

(a-c) Profiles of histone variants indicated above each panel across the TTS for 1000 highly active (red), intermediately active (cyan) and silent genes (black) are shown.

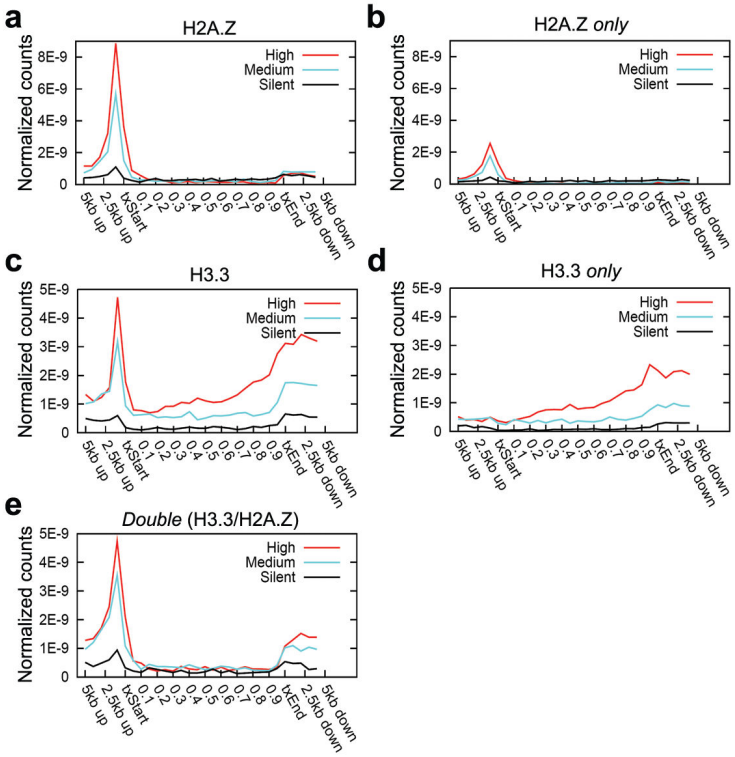

Figure 4.

Different combinations of histone variants have distinctive distribution patterns across genes.

(a-e) Profiles of histone variants indicated above each panel in and around gene bodies for 1000 highly active (red), intermediately active (cyan) or silent (black) genes are presented. For each gene, island-filtered tags were summed according to their shifted positions in 1 kb windows from 5 kb upstream of the transcription start site (txStart) to the txStart and from the transcription end site (txEnd) to 5 kb downstream. Within the gene bodies, island-filtered tags were summed according to their shifted positions in windows equal to 5% of the gene length. All window tag counts were normalized by the total number of bases in the windows. The profiles of H2A.Z and 'H2A.Z *only*' (H3.3 and 'H3.3 *only*') were further normalized by the total tag numbers of island-filtered tags in the H2A.Z (H3.3) library, respectively. Using the same normalization for H2A.Z and 'H2A.Z *only*' (H3.3 and 'H3.3 *only*') allows quantitative comparison of modification levels, while the profile of *Double* (H3.3/H2A.Z) was further normalized by total island-filtered tags in the *Double* library.

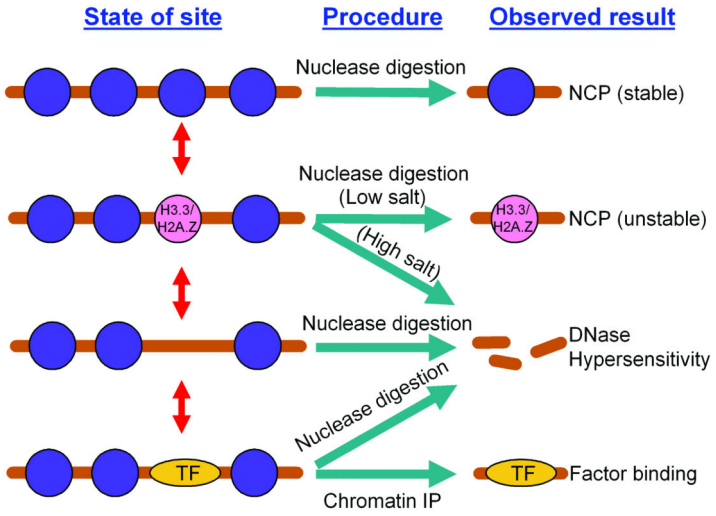

Figure 5.

Schematic representation of the dynamic exchange of factors at a transcriptionally active TSS or other regulatory elements.

These sites can be occupied transiently by H3/H2A-containing NCPs (blue circle), by H3.3/H2A.Z NCPs (pink circle), by transcription factors (TFs, yellow oval), or they can be free of proteins. The mechanisms of exchange among these different states are not clear. Each of these states will be detected by the assay method designed for that purpose, but over time each site will occupy all four states.