HISTONE VARIANTS DELINEATE THE TRANSCRIPTION ORIENTATION AT ENHANCERS

Inchan Choi[1,2], Banjamin A. Garcia[1] & Kyoung-Jae Won[1,2]*

The Institute for Diabetes[1], Obesity and Metabolism, Department of Genetics[2], Epigenetics Program[3], and Department of Biochemistry and Biophysics[4], University of Pennsylvania.
Speaker email: wonk@upenn.edu

Traditionally, enhancers have been defined as remote elements that increase transcription independently of their orientation. Recent genome-wide survey suggests that enhancers might have orientation. In this paper, we show that enhancers have transcriptional orientation and histone variants delineate the transcriptional orientation. Using chromatin immunoprecipitation followed by next generation sequencing (ChIP-seq) against the four histone variants (H3.1, H3.3, H2A.Z and macroH2A), we identified various combinations of histone variants (histone variants codes) at distal regulatory regions, suggesting that histone variants are deposited at regulatory regions to assist gene regulation. More importantly, we identified both symmetric and asymmetric patterns of histone variant (H3.3 and H2A.Z) occupancies at intergenic regulatory regions. Strikingly, these directional patterns were associated with RNA Polymerase II (PolII). These asymmetric patterns correlated with the enhancer activities measured by global run-on sequencing (GRO-seq) data. We also showed that enhancers with skewed histone variants patterns well facilitate enhancer activity. Our study indicates that H2A.Z and H3.3 delineate the orientation of transcription at enhancers as observed at promoters.

INTRODUCTION

The eukaryotic genome is packaged in the nucleus as chromatin, a dynamic arrangement which serves to compact the DNA. The fundamental unit of chromatin is the nucleosome. Nucleosomes are composed of an octamer of histone proteins comprised of two copies each of H2A, H2B, H3, and H4[1]. In addition to the canonical histones there are also protein variants encoded by separate genes[2]. These variants play further important roles in DNA packaging and controlling gene expression[3]. For instance, histone H2A.Z replaces canonical H2A at some 5' end of both active and inactive genes[4,5]. Histone H3.3 is specially enriched at transcriptionally active genes[6,7]. Repressive macroH2A mark is mainly associated with heterochromatic regions[8]. The presence of upstream H2A.Z nucleosomes only seen in some organisms with bidirectional transcription at promoters suggests that histone variants are associated with transcriptional direction at promoters[9]. Recent studies identified that histone variants are enriched in distal regulatory regions[6,7,10-12]. However, our understanding about the function of histone variants at distal regulatory regions is still limited. We investigated if histone variants are associated with the transcription orientation at enhancers.

RESULTS

Histone variant codes at regulatory regions

In order to understand the strategic deposition of histone variants and their functional roles, we used chromatin immunoprecipitation followed by massive sequencing (ChIP-seq) to map the FLAG-tagged histone variants (H2A.Z, macroH2A, H3.1, and H3.3) in the genome. First, we asked how the histone variants are positioned around genes. After sorting the annotated Refseq genes based on their expression levels, we observed sharp enrichment of H2A.Z marking the two nucleosomes flanking the nucleosome free regions (NFRs) at active promoters (Fig 1A)[13,14]. In gene body, H2A.Z was absent regardless of their expression levels (data not shown). Both H3.1 and H3.3 were depleted around the transcription start sites (TSSs) of active genes (Fig. 1A). At transcription termination sites (TTSs), we also found a depletion of H3.1 and H3.3[15]. Interestingly, both H3.1 and H3.3 levels correlated with gene expression levels in gene body.

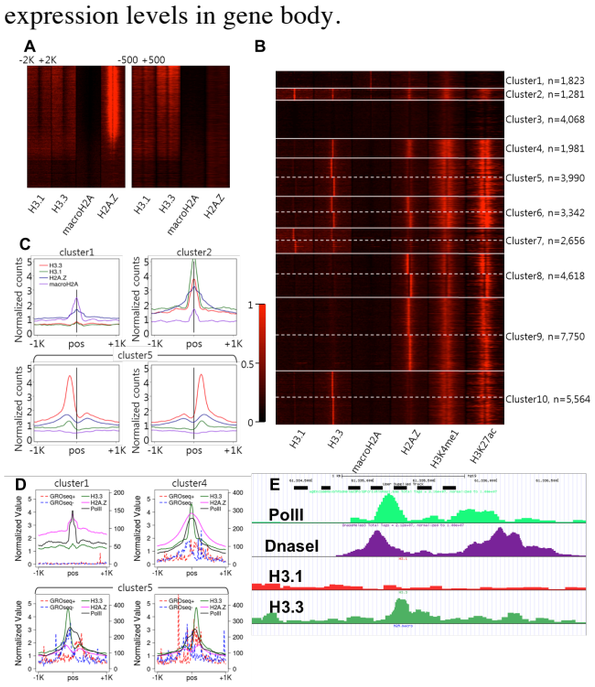

***FIGURE 1**. Genomic profiles of histone variants*

(A) Distribution of ChIP-seq reads at annotated TSSs (+/- 2K) and TTSs (+/- 500) and *(B)* at distal regulatory regions. We clustered DHSs located in the intergenic region. We identified 16 groups and rearranged them to 10 clusters based on their profiles. Various compositions of histone variants were found. Clusters 5-10 are composed of 2 mirroring groups. After clustering based on histone variants, we aligned histone modification. Histone variants are off-centered for the mirroring clusters (cluster 5-10), suggesting orientation at regulatory *(C)* Symmetric and asymmetric profiles of histone variants. Clusters 1 and 4 show symmetric profiles with various compositions of histone variants. Cluster 5 shows mirroring asymmetric profile. *(D)*

The profiles of H3.3 and H2A.Z are associated with the pattern of PolII and enhancer transcripts. For symmetric clusters PolII is located at the center. For asymmetric clusters, PolII is skewed to the direction of the peaks of histone variants. Transcripts at enhancers show bidirectional patterns. Strand-specific transcripts are stronger in the asymmetric clusters. *(E)* The screenshot of histone variant around DNaseI. H3.3 is enriched more to the right side of the peak of DNaseI. PolII has its peak to the right side to DNaseI.

We then investigated if histone variants are enriched at promoter distal (>2Kbp from the annotated TSSs) regulatory regions. Using 37,073 DNaseI hypersensitive sites (DHSs) located at intergenic regions, we defined 16 groups (Fig. 1B). Majority of the distal DHSs were enriched for H3.3 and/or H2A.Z as well as H3K27ac indicating that these histone variants are important for enhancer function. Interestingly, we also found clusters marked by H3.1 (cluster 2 and 7) or even with repressive macroH2A mark (cluster 1).

Histone variants have symmetric and asymmetric patterns at distal regulatory regions

Histone H3.3 and other histone variants were observed to have asymmetric profiles (Fig. 1B and C). The clusters 5-7 showed that H3.3 and H2A.Z were skewed to one side. The H3K27ac profiles were enriched on the side where the H3.3 peak was located even though the skewness was less dramatic as compared to histone variant profiles. The histone acetyltransferase p300 as well as other co-factors profiles were centered at the DHSs (symmetric), indicating that the skewness in the histone variant profiles was independent from transcription factors and their co-factors.

Strikingly, we found that PolII occupancy was skewed towards the peak of H3.3 and H2A.Z (Fig. 1D (cluster 5) and E). H3.3 is located on a side of the DHS and PolII peak was observed between the DHS and H3.3 peak, as shown in the profile. We validated this observation by performing ChIP-qPCR of the 5 regions around the DHSs. The qPCR experiment confirmed that PolII and H3.3 were not symmetric at a potential enhancer and skewed towards the same direction (data not shown).

Role of chromatin domains containing studied histone variants – We investigated the enhancer activities of these potential regulatory using the global run-on sequencing (GRO-seq) data in HeLa cells[16]. We observed modest but significant bias of the strand-specific transcripts that matched with the PolII orientation in the asymmetric clusters (Fig. 1D). The enhancer activity and the PolII levels of the cluster with macroH2A (cluster 1) were very weak. We also observed stronger eRNA levels for cluster 8 than cluster 9, where skewness was more strongly observed (p-value= 6.4e-11). Collectively, these suggest that skewed histone variants facilitate enhancer activity.

H3.3 and H2A.Z are with activating histone modification mark – As we identified enhancers enriched for histone variants, we further investigated the composition of histone variant containing nucleosome using chromatin immunoprecipitation coupled to mass spectrometry (ChIP-MS) (Fig. 2). Activating enhancer marks (H3K4me2 and H3K27ac) were highly enriched in nucleosomes ChIPed with H3.3 or H2A.Z (11.5 fold in H3.3 and 19.8 fold in H2A.Z as compared with the genomic chromatin levels). H3K4me3 was found to be enriched almost 30-fold in H2A.Z purified nucleosomes as compared to global input (Fig. 2), which was consistent with previous observations that investigated active promoters[14]. A similar trend was observed for H3K36me3 in H3.3 and H2A.Z purified nucleosomes, which was about 7% and 15% of the total histone H3, respectively. In genomic chromatin and macroH2A containing nucleosomes H3K36me3 was only 4% and 1%, respectively (Fig. 2).

We also observed a dramatic enrichment of H4K16ac in nucleosomes purified with H3.3 (~50%) and H2A.Z (~40%) as compared to this modification in genomic chromatin (~20%). The repressive mark H3K27me3 was enriched approximately three fold changes in macroH2A purified nucleosomes as compared to genomic chromatin (~18% in macroH2A vs. ~6% in genomic or ~5% in H2A.1 purified nucleosomes).

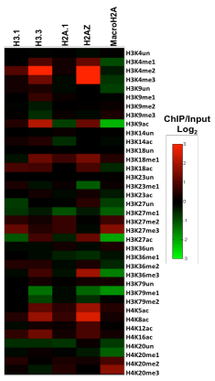

Figure 2. *relative ratio of histone post-translational modifications in FLAG-IP experiments as compared to the global HeLa extract.*

Relative abundance of single histone PTMs were calculated. Log2 ratio was calculated between each FLAG-IP sample (listed on top of the heatmap) and the HeLa input. Single PTMs were sorted by common regulation into a hierarchical tree.

CONCLUSIONS

We identified diverse compositions of four histone variants at potential regulatory regions. Besides H3.3 and H2A.Z, we found regulatory regions enriched for H3.1 or even repressive macroH2A mark. More importantly, we observed asymmetric patterns of histone variants. The asymmetric patterns of histone variants were associated with PolII occupancy as well as transcriptions at enhancers. Taken together, this demonstrated that skewed histone variants were not just noise, but such deposition dictates the direction of PolII movement.

Previous study already identified a number of enhancer groups with asymmetric histone modification patterns[18]. More directly, nascent RNAs at enhancers (or enhancer RNAs (eRNAs)) show both bidirectional and unidirectional transcripts[21]. 5' ends of capped RNAs detected by Cap Analysis of Gene Expression (CAGE) in HeLa cells confirmed unidirectional transcripts at enhancers[22]. The asymmetric patterns of eRNA as well as the skewed PolII occupancy in our study suggest that enhancers have directional information.

PolII orientation is clearly defined at promoters. A remarkable observation was made for bidirectional as well as unidirectional promoters in association with histone variants. H2A.Z at active promoters show strong up- as well as downstream peaks in human and yeast[13,14,26], but not in flies[27] or Arabidopis[28]. The presence of upstream H2A.Z nucleosomes seen in some organisms correlates with bidirectional transcription in yeast and mammals[9,29]. This suggests that histone variants are associated with transcriptional direction. At promoters, the nucleosome located to the transcriptional direction blocks the movement of PolII[30]. Depletion of H2A.Z from a nucleosome

position resulted in a higher barrier to PolII[30]. Enhancers bare similar characteristics with promoters. Besides eRNAs, some enhancers are even with TATA box sites[22].

In conclusion, why do enhancers have orientation? A DNA looping model has been suggested where promoter-enhancer interactions facilitate gene transcription[31-33]. These observations questioned the transcriptional mechanism associated with histone variants and PolII. By forming a looping structure, PolII needs to have a preference for its movement as it needs to move towards the transcription orientation of the associated gene (Fig. 3). The nucleosome on one side of the regulatory region may block the movement of PolII at distal regulatory regions. H3.3 and H2A.Z may help the movement of PolII by destabilizing the nucleosome that blocks the movement of PolII.

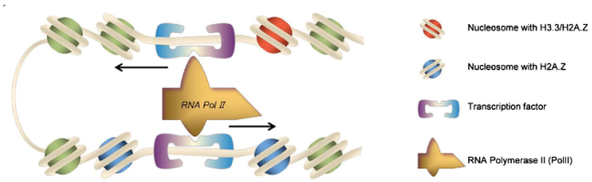

***FIGURE 3*. A possible model for gene regulation associated with H3.3**. The nucleosome at enhancer located to the transcription direction is with histone variants to facilitate the movement of PolII.

This work has been published in *Chromatin and Epigenetics*.

Won KJ*, Choi I*, LeRoy G*, Zee BM, Sidoli S, Gonzales-Cope M, Garcia BA (2015) Proteogenomics analysis reveals specific genomic orientations of distal regulatory regions composed by non-canonical histone variants. ***Chromatin & Epigenetics***, 10;8:13

REFERENCES

1 Luger, K. Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev* **13**, 127-135 (2003).
2 Kamakaka, R. T. & Biggins, S. Histone variants: deviants? *Genes Dev* **19**, 295-310, doi:10.1101/gad.1272805 (2005).
3 Weber, C. M. & Henikoff, S. Histone variants: dynamic punctuation in transcription. *Genes Dev* **28**, 672-682, doi:10.1101/gad.238873.114 (2014).
4 Adam, M., Robert, F., Larochelle, M. & Gaudreau, L. H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. *Mol Cell Biol* **21**, 6270-6279 (2001).
5 Raisner, R. M. *et al.* Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**, 233-248, doi:10.1016/j.cell.2005.10.002 (2005).
6 Mito, Y., Henikoff, J. G. & Henikoff, S. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37**, 1090-1097, doi:10.1038/ng1637 (2005).
7 Ray-Gallet, D. *et al.* Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. *Mol Cell* **44**, 928-941, doi:10.1016/j.molcel.2011.12.006 (2011).
8 Changolkar, L. N. & Pehrson, J. R. macroH2A1 histone variants are depleted on active genes but concentrated on the inactive X chromosome. *Mol Cell Biol* **26**, 4410-4420, doi:10.1128/MCB.02258-05 (2006).
9 Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845-1848, doi:10.1126/science.1162228 (2008).
10 Mito, Y., Henikoff, J. G. & Henikoff, S. Histone replacement marks the boundaries of cis-regulatory domains. *Science* **315**, 1408-1411, doi:10.1126/science.1134004 (2007).
11 Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**, 1191-1200 (2002).
12 Goldberg, A. D. *et al.* Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678-691, doi:10.1016/j.cell.2010.01.003 (2010).
13 Tolstorukov, M. Y., Kharchenko, P. V., Goldman, J. A., Kingston, R. E. & Park, P. J. Comparative analysis of H2A.Z nucleosome organization in the human and yeast genomes. *Genome Res* **19**, 967-977, doi:10.1101/gr.084830.108 (2009).
14 Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).
15 Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362-366, doi:10.1038/nature07667 (2009).
16 Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-461, doi:10.1038/nature12787 (2014).
17 Ku, M. *et al.* H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. *Genome biology* **13**, R85, doi:10.1186/gb-2012-13-10-r85 (2012).
18 Kundaje, A. *et al.* Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements. *Genome Res* **22**, 1735-1747, doi:10.1101/gr.136366.111 (2012).
19 Lai, W. K. & Buck, M. J. ArchAlign: coordinate-free chromatin alignment reveals novel architectures. *Genome Biol* **11**, R126, doi:10.1186/gb-2010-11-12-r126 (2010).
20 Nielsen, F. G. *et al.* CATCHprofiles: clustering and alignment tool for ChIP profiles. *PLoS One* **7**, e28272, doi:10.1371/journal.pone.0028272 (2012).
21 Li, W. *et al.* Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516-520, doi:10.1038/nature12210 (2013).
22 Andersson, R. *et al.* Nuclear stability and transcriptional directionality separate functionally distinct RNA species. *Nature communications* **5**, 5336, doi:10.1038/ncomms6336 (2014).
23 Min, I. M. *et al.* Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev* **25**, 742-754, doi:10.1101/gad.2005511 (2011).
24 Lam, M. T. *et al.* Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**, 511-515, doi:10.1038/nature12209 (2013).
25 Hu, G. *et al.* H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* **12**, 180-192, doi:10.1016/j.stem.2012.11.003 (2013).
26 Bargaje, R. *et al.* Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Res* **40**, 8965-8978, doi:10.1093/nar/gks665 (2012).
27 Mavrich, T. N. *et al.* Nucleosome organization in the Drosophila genome. *Nature* **453**, 358-362, doi:10.1038/nature06929 (2008).
28 Zilberman, D., Coleman-Derr, D., Ballinger, T. & Henikoff, S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**, 125-129, doi:10.1038/nature07324 (2008).
29 Xu, Z. *et al.* Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457**, 1033-1037, doi:10.1038/nature07728 (2009).
30 Weber, C. M., Ramachandran, S. & Henikoff, S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell* **53**, 819-830, doi:10.1016/j.molcel.2014.02.014 (2014).
31 Blackwood, E. M. & Kadonaga, J. T. Going the distance: a current view of enhancer action. *Science* **281**, 60-63 (1998).
32 Krivega, I. & Dean, A. Enhancer and promoter interactions-long distance calls. *Curr Opin Genet Dev* **22**, 79-85, doi:10.1016/j.gde.2011.11.001 (2012).
33 Zhang, Y. *et al.* Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature* **504**, 306-310, doi:10.1038/nature12716 (2013).