Open Access Full Text Article

ORIGINAL RESEARCH

Calculation of the ratio of the mononucleosome mapping number to the dinucleosome mapping number for each nucleotide position in the *Aspergillus fumigatus* genome

Hiromi Nishida

Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo, Japan

Abstract: The ratio of the mononucleosome mapping site-number to the dinucleosome mapping site-number calculated using nucleosome mapping data of *Aspergillus fumigatus* was employed as a marker for the sensitivity of nucleosomes to micrococcal nuclease (MNase). This ratio was obtained for four different regions, namely, active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters and compared. It was found that the nucleosomes located in the active gene promoters were more sensitive to MNase than those in the active gene bodies but less sensitive than those in the inactive gene bodies and promoters. In addition, compared to the gene bodies, gene promoters have a greater variety of nucleosomes that are sensitive to MNase, regardless of the transcriptional activity of these genes. This study showed that in addition to a mononucleosome map, a dinucleosome map is useful for estimating the sensitivity of nucleosomes to MNase.

Keywords: *Aspergillus fumigatus*, chromatin, dinucleosome stability, micrococcal nuclease sensitivity, nucleosome, transcriptional activity

Introduction

Eukaryotic genomic DNA is packaged with histones to form chromatin,[1] in which the most fundamental repeating unit is the nucleosome.[2] Nucleosomes consist of an octamer of histones around which genomic DNA is wrapped.[2] Adjacent nucleosomes are separated by unwrapped linker DNA. Analyses performed using genomic tiling arrays or massively parallel DNA sequencers have led to the high-resolution mapping of nucleosome positions.[3–13] The genome-wide nucleosome map has revealed the relationship between nucleosome density (particularly, the presence or absence of nucleosomes in gene promoter regions) and gene expression.

In the genome of the filamentous fungus *Aspergillus fumigatus*, 7,715,001 mononucleosomal and 8,565,279 dinucleosomal DNA fragments have already been sequenced and mapped using a massively parallel DNA sequencer.[13] In a previous study, a relationship was shown to exist between the distribution of nucleosomal DNA fragment lengths and gene expression levels. The distribution of mononucleosomal DNA fragment lengths showed two peaks: one at 135 nt and the other at 150 nt.[13] The gene bodies of active and inactive genes and inactive gene promoters showed these two peaks; however, the peak produced at 150 nt by the active gene promoters was lost.[13] This may be because the active promoters were more sensitive to micrococcal nuclease (MNase) than the other regions and therefore may have lost the farther peak.

Correspondence: Hiromi Nishida
Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan
Tel +81 3 5841 1297
Fax +81 3 5841 1136
Email hnishida@iu.a.u-tokyo.ac.jp

1

In order to ascertain this assumption, sensitivity to MNase should be examined.

Genomic DNA consists of regions with varied sensitivity to MNase.[14–16] MNase digests unwrapped linker DNA but does not digest nucleosome-bound DNA. Prolonged treatment with MNase (eg, a 15-min treatment of rye chromatin with MNase) results in nearly complete cleavage of the total DNA into mononucleosomes;[16] this finding suggests that MNase-digested dinucleosomes are precursors of mononucleosomes. Assuming that *A. fumigatus* dinucleosomes function as mononucleosome precursors, it may be possible to use the ratio of the mononucleosome mapping site-number to the dinucleosome mapping site-number as a marker for sensitivity of nucleosomes to MNase or dinucleosome stability. By comparing the ratios obtained in four regions, namely, active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters, the sensitivity of the nucleosomes to MNase or the stability of dinucleosomes can be estimated.

Methods

On the basis of each nucleosomal DNA fragment sequence obtained from mononucleosome and dinucleosome mapping data in *A. fumigatus*,[13] mononucleosome and dinucleosome mapping numbers were estimated for each nucleotide position

in the region from 1 kb upstream of the translational start site to the translational end site of a gene (Figure 1). Next, the ratio of the mononucleosome mapping number to the dinucleosome mapping number was calculated for each nucleotide. In this study, 50 genes with the highest expression levels (excluding rRNA genes) and 50 genes with the lowest expression levels were analyzed; the genes were selected on the basis of microarray data obtained in a previous study.[13] The gene promoter was defined as the region from 1 kb upstream of the translational start site, and the gene body was defined as the region from the translational start site to the translational end site.

Results and discussion

The mononucleosome and dinucleosome mapping numbers and their ratio were calculated for each nucleotide position in the region from 1 kb upstream of the translational start site to the translational end site of each of the 50 transcriptionally active and 50 transcriptionally inactive genes. The genomic nucleotide positions that exhibited a ratio of the mononucleosome mapping number to the dinucleosome mapping number with both numbers being ≥1 for each nucleotide position in the active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters were 71,873, 37,734, 60,478, and 42,386, respectively.
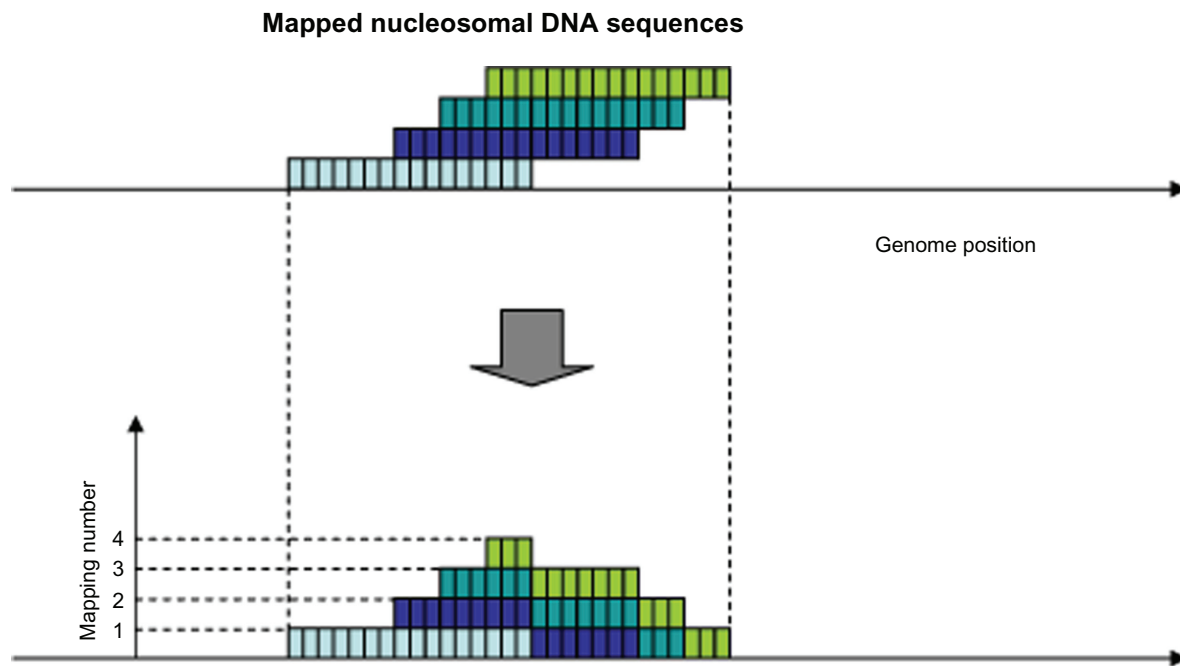

Figure 1 An example of mapping numbers for each nucleotide position. The four colors indicate the different nucleosomal DNA fragment sequences. Each box represents a genomic nucleotide. In this example, three nucleotide positions are shared by four different sequences (mapping number, 4).

The genes *Afu5g03760* and *Afu2g13025* had the highest and lowest expression, respectively, and are shown in Figure 2. Although the dinucleosome mapping numbers were higher than the mononucleosome mapping numbers, the profiles of the two types of numbers were similar. The Spearman's rank correlation coefficients between the mononucleosome mapping numbers and the dinucleosome mapping numbers were 0.68 (the 50 transcriptionally active gene promoters), 0.61 (the 50 transcriptionally active gene bodies), 0.65 (the 50 transcriptionally inactive gene promoters), and 0.71 (the 50 transcriptionally inactive gene bodies). These findings suggest that many of the dinucleosomes served as precursors for the mononucleosomes. If this assumption holds true, genomic regions with a high ratio of the mononucleosome mapping number to the dinucleosome mapping number will be more sensitive to MNase than other regions. In other words, the dinucleosome stability of such regions would be low.

Among the 50 transcriptionally active genes, 22 genes had the region(s) with high ratio (>2 in $\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number (Supplementary Data 1, Table 1; see http://www.dovepress.com/submission_7797_files.php for all supplementary data). Among the 50 transcriptionally inactive genes, 27 genes had the region(s) with high ratio (>2 in $\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number (Supplementary Data 2, Table 1). Most of those regions were AT rich, regardless of their transcriptional activation status (Supplementary Data 3 and 4, Table 1). This result is consistent with the yeast nucleosome positioning.[17]

The median values of the ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number in the active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters were -1.37, -1.00, -0.72, and -0.76, respectively (Figure 3). Comparison of the four median values strongly indicated that the dinucleosomes are more stable in active genes (both promoters and bodies) than in inactive genes.
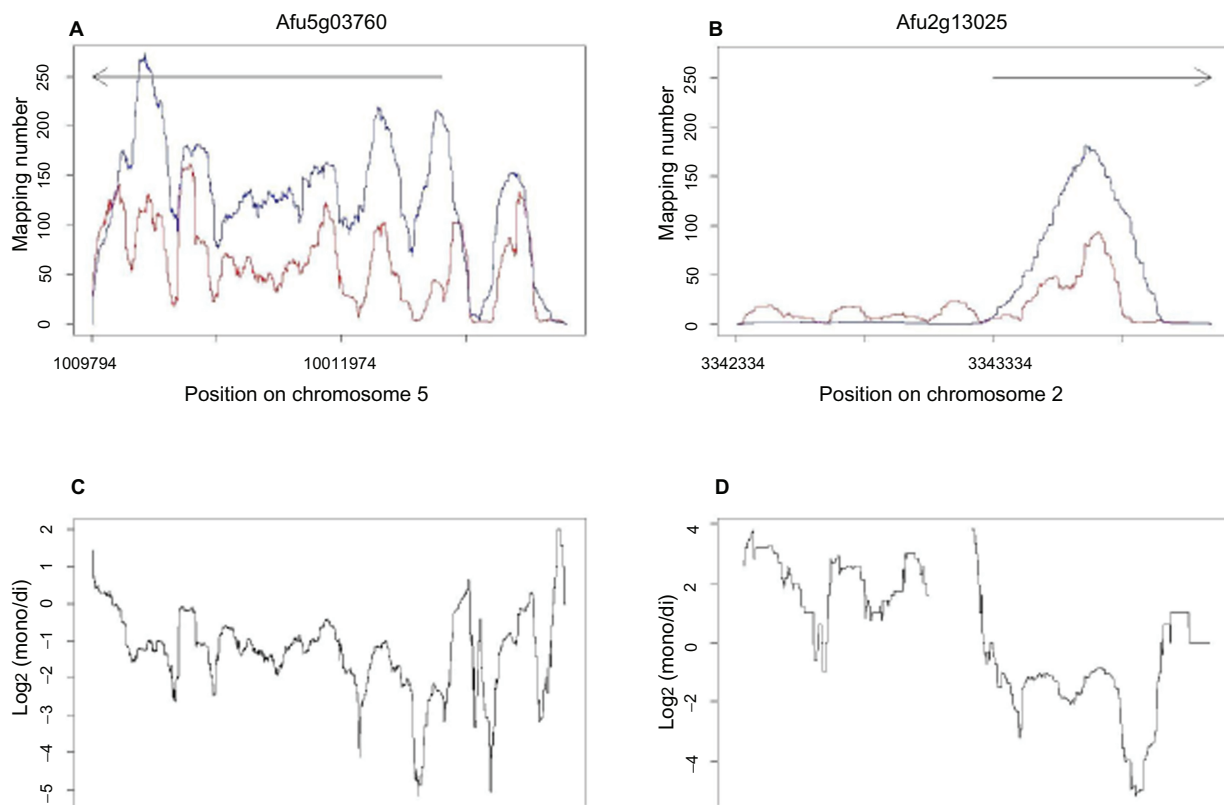

Figure 2 Mapping numbers and ratios of the mononucleosome mapping number to the dinucleosome mapping number for the genes *Afu5g03760* and *Afu2g13025*, which had the highest and lowest expression, respectively. **A**) Red and blue have been used to indicate the mononucleosome and dinucleosome mapping numbers, respectively, in the region from 1 kb upstream of the translational start to the translational end site in *Afu5g03760*. The arrow indicates the region from the translational start site to the translational end site. **B**) Red and blue indicate the mononucleosome and dinucleosome mapping numbers, respectively, in the region from 1 kb upstream of the translational start site to the translational end site in *Afu2g13025*. The arrow indicates the region from the translational start site to the end site. **C**) Ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number for *Afu5g03760*. **D**) Ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number for *Afu2g13025*.

Table 1 Genes with high ratio (>2 in $\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number

| Gene	Region	AT-rich
Transcriptionally active genes (22 genes)		
Afu1g02550	Promoter	No
Afu1g03570	Promoter/Body	Yes/Yes
Afu1g03630	Promoter	Yes
Afu1g14220	Promoter	Yes
Afu2g11150	Promoter/Body	Yes/No
Afu2g11520	Promoter	Yes
Afu2g13530	Promoter	No
Afu2g16060	Promoter	Yes
Afu3g08010	Promoter	No
Afu3g08160	Promoter	Yes
Afu3g10660	Promoter	Yes
Afu3g10960	Promoter	Yes
Afu4g06820	Promoter	No
Afu4g13120	Promoter	Yes
Afu5g03020	Promoter	Yes
Afu5g04210	Promoter	Yes
Afu5g04230	Promoter	No
Afu5g08830	Promoter	Yes
Afu6g04740	Promoter	Yes
Afu8g03930	Promoter	Yes
Afu8g05320	Promoter	Yes
Afu8g05610	Promoter	Yes
Transcriptionally inactive genes (27 genes)		
Afu1g00140	Body	Yes
Afu1g09340	Promoter/Body	Yes/Yes
Afu1g13810	Body	Yes
Afu1g15410	Promoter/Body	Yes/Yes
Afu1g17300	Promoter/Body	Yes/Yes
Afu2g04180	Promoter	Yes
Afu2g04420	Promoter	Yes
Afu2g09970	Promoter/Body	Yes/Yes
Afu2g13025	Promoter	Yes
Afu2g13300	Promoter	No
Afu2g17550	Promoter	Yes
Afu3g01470	Promoter	Yes
Afu3g03270	Promoter	Yes
Afu3g14580	Promoter	Yes
Afu4g06650	Promoter/Body	Yes/Yes
Afu4g14530	Promoter	No
Afu4g14690	Promoter	Yes
Afu5g11060	Promoter	Yes
Afu5g12690	Promoter	Yes
Afu5g14100	Promoter	Yes
Afu5g14840	Body	Yes
Afu5g14990	Promoter	Yes
Afu6g11280	Promoter	Yes
Afu7g04530	Promoter	No
Afu8g00820	Promoter	Yes
Afu8g02230	Promoter	Yes
Afu8g02270	Promoter	Yes
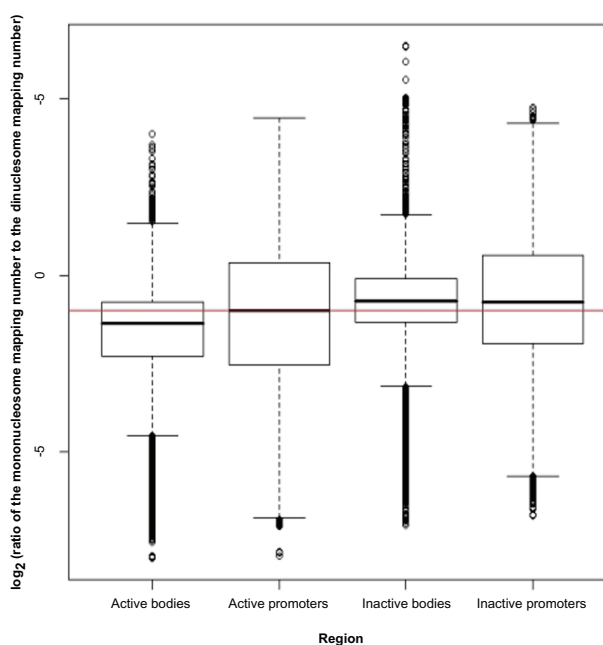
Figure 3 Boxplots of the ratios ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number in transcriptionally active and inactive gene bodies and promoters. The boxes are composed of the medians with the first and third quartiles. The dots indicate outliers. The genomic nucleotide positions with a significant ratio of the mononucleosome mapping number to the dinucleosome mapping number (with both numbers being ≥1 for a single nucleotide position) in the active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters were 71,873, 37,734, 60,478, and 42,386, respectively. The median values for the ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number in the active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters were −1.37, −1.00, −0.72, and −0.76, respectively. The red line indicates the median for the ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number in the active gene promoters.

In addition, the difference (0.37) between the medians of the ratios for the active gene bodies and promoters was higher than the difference (0.04) between the medians of the ratios for the inactive gene bodies and promoters. Thus, the nucleosomes in the active gene promoters are certainly more sensitive to MNase than the nucleosomes in the active gene bodies but are less sensitive than those in the inactive gene bodies and promoters. This result indicates that the sensitivity of the active promoters to MNase is not the sole reason for the loss of the farther peak in the distribution of the mononucleosomal DNA fragment lengths in the active gene promoters.[13]

The histograms for the ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number in the 4 different regions were compared (Figure 4); it was found that the distributions in the gene bodies (both active and inactive) produced a larger peak than those in the corresponding gene promoters. This indicates that compared to gene bodies, gene promoters have a greater variety of
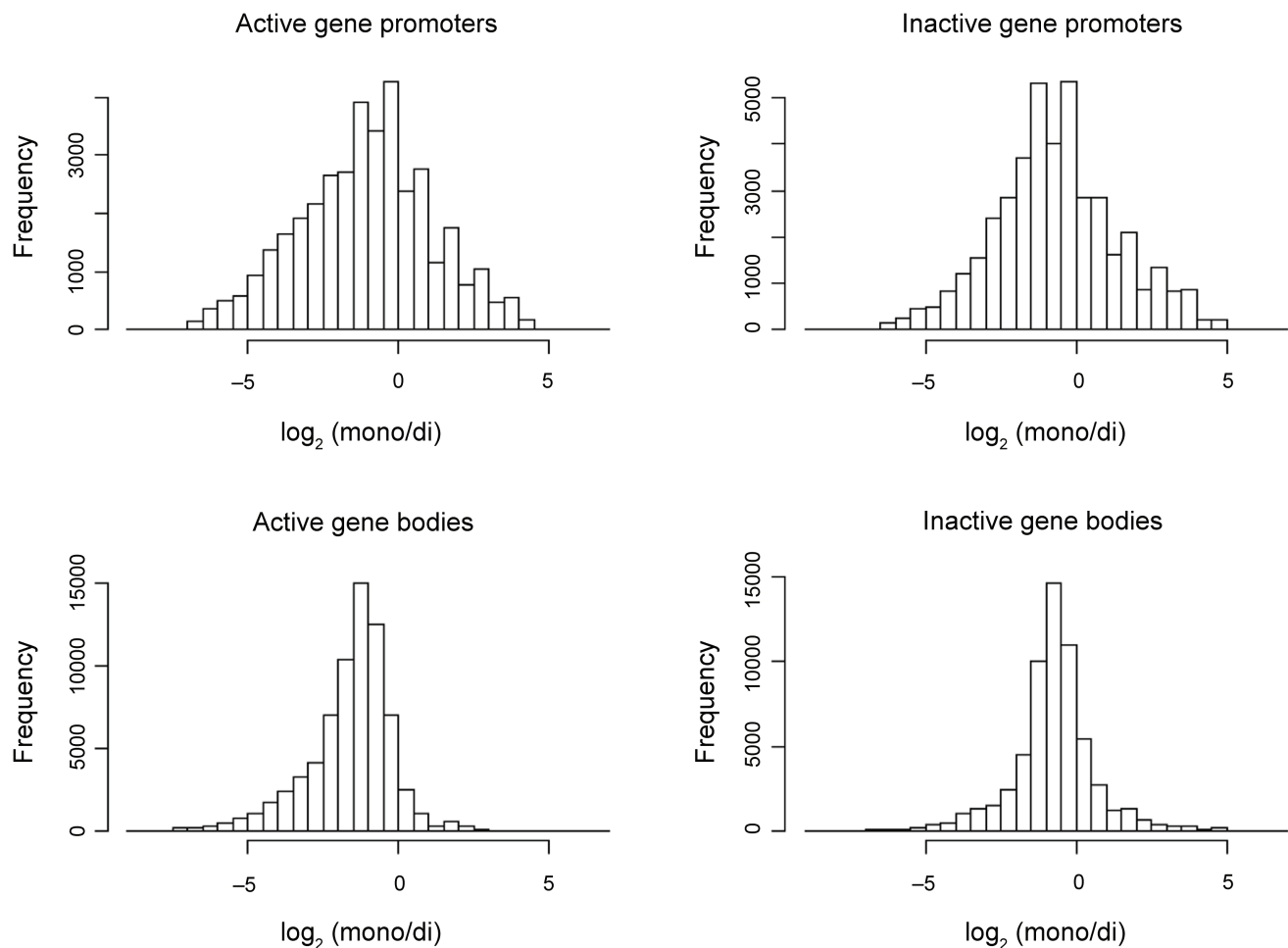
Active gene promoters

Inactive gene promoters

Active gene bodies

Inactive gene bodies

Figure 4 Histograms of the ratio ($\log_2$ scale) of the mononucleosome mapping number to the dinucleosome mapping number in transcriptionally active and inactive gene bodies and promoters. The genomic nucleotide positions with a significant ratio of the mononucleosome mapping number to the dinucleosome mapping number (with both numbers being ≥ 1 for a single nucleotide position) in the active gene bodies, active gene promoters, inactive gene bodies, and inactive gene promoters were 71,873, 37,734, 60,478, and 42,386, respectively. The differences of the distributions between the gene promoters and the gene bodies were evaluated by the Kolmogorov-Smirnov test. The results indicated that those differences in both the transcriptionally active and inactive genes were statistically significant (P-value $< 2.2 \times 10^{-16}$).

nucleosomes that are sensitive to MNase, regardless of the transcriptional activity of these genes. It is possible that a variety of modified histones are related to the different types of nucleosomes in gene promoters that are sensitive to MNase. Fungi possess various histone modification systems.[18] In human cells, the relative nucleosome depletion in the vicinity of the transcription start site may not necessarily be associated with active transcription but with histone modification in gene promoters.[19] Hence, the sensitivity of nucleosomes to MNase may be related to histone modification, regardless of the transcriptional activity of the genes. Generally the nucleosome map studies are based on the mononucleosome map data. This study showed that in addition to a mononucleosome map, a dinucleosome map is useful for determining the sensitivity of nucleosomes to MNase.

References

1. Igo-Kemenes T, Hörz W, Zachau HG. Chromatin. *Annu Rev Biochem*. 1982;51:89–121.
2. Luger K. Dynamic nucleosome. *Chromosome Res*. 2006;14:5–16.
3. Dennis JH, Fan HY, Reynold SM, et al. Independent and complementary methods for large-scale structural analysis of mammalian chromatin. *Genome Res*. 2007;17:928–939.
4. Kaplan N, Moore IK, Fondufe-Mittendorf Y, et al. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*. 2009;458:362–366.
5. Lee W, Tillo D, Bray N, et al. A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet*. 2007;39:1235–1244.
6. Mavrich TN, Jiang C, Ioshikhes IP, et al. Nucleosome organization in the *Drosophila* genome. *Nature*. 2008;453:358–362.

7. Schones DE, Cui K, Cuddapah S, et al. Dynamic regulation of nucleosome positioning in the human genome. *Cell.* 2008;132:887–898.

8. Segal E, Fondufe-Mittendorf Y, Chen L, et al. A genomic code for nucleosome positioning. *Nature.* 2006;442:772–778.

9. Shivaswamy S, Bhinge A, Zhao Y, Jones S, Hirst M, Lyer VR. Dynamic remodeling nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* 2008;6:e65.

10. Song JS, Liu X, Liu XS, He X. A high-resolution map of nucleosome positioning on a fission yeast centromere. *Genome Res.* 2008;18: 1064–1072.

11. Valouev A, Ichikawa J, Tonthat T, et al. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* 2008;18:1051–1063.

12. Yuan GC, Liu YJ, Dion MF, et al. Genome-scale indentification of nucleosome positions in *S. cerevisiae. Science.* 2005;309:626–630.

13. Nishida H, Motoyama T, Yamamoto S, Aburatani H, Osada H. Genome-wide maps of mono- and di-nucleosomes of *Aspergillus fumigatus. Bioinformatics.* 2009;25:2295–2297.

14. Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res.* 2009;19:460–469.

15. Tolstorukov MY, Kharchenko PV, Goldman JA, Kingston RE, Park PJ. Comparative analysis of H2A.Z nucleosome organization in the human and yeast genomes. *Genome Res.* 2009;19:967–977.

16. Vershinin AV, Heslop-Harrison JS. Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Mol Biol.* 1998;36:149–161.

17. Ioshikhes IP, Albert I, Zanton SJ, Pugh BF. Nucleosome positions predicted through comparative genomics. *Nat Genet.* 2006;38:1210–1215.

18. Nishida H. Evolutionary conservation levels of subunits of histone-modifying protein complexes in fungi. *Comp Funct Genomics.* 2009;2009:379317.

19. Nishida H, Suzuki T, Kondo S, Miura H, Fujimura Y, Hayashizaki Y. Histone H3 acetylated at lysine 9 in promoter is associated with low nucleosome density in the vicinity of transcription start site in human cell. *Chromosome Res.* 2006;14:203–211.