European Review for Medical and Pharmacological Sciences 2015; 19: 2651-2665

The determinations of nucleosome positioning

Y.-N. CAI[1,2], J.-Y. LIU[3], R. HU[1], Q.-H. ZHANG[1], Z.-M. DAI[1], Y.-J. QIAO[4,5], X.-H. DAI[1,2]

[1]School of Information Science and Technology, Sun Yat-sen University, Higher Education Mega Center, Guangzhou, China
[2]SYSU-CMU Shunde International Joint Research Institute (JRI), Shunde, Guangdong, China
[3]College of Chinese Language and Culture, Jinan University, Guangzhou, PR China
[4]School of Economics and Commerce, South China University of Technology, Guangzhou Higher Education Mega Center, Guangzhou, China
[5]Guangzhou Rural Commercial Bank, Zhujiang New Town, Guangzhou, China

Abstract. – OBJECTIVE: Nucleosomes are the basic packaging units of chromatin, determinants of nucleosome organization playing a major role in genome packaging. Although a wide variety of nucleosome organization factors have been considered separately across the whole or partial human genomic regions, it is unclarified that what the major determinants and their roles in scale are when being put all together. And it is also unknown that what the similarities and differences of determinants between different genomic features such as genes of different expression levels or genomic regions with different functions.

MATERIALS AND METHODS: We detected commonalities and characteristics of nucleosome positioning determinants in different genes and regions with 1591486 nucleosomes identified by ourselves in human CD4+ cell.

RESULTS: It was found that a distinct linear combination of about 20 nucleosome-positioning factors explained nucleosome occupancy for each genomic feature. In those linear combinations, 6 DNA sequence attributes (Roll stiffness and Twist stiffness, CT and AG, CG and shift stiffness) and a histone modification (H4R3me2) are shared. And other factors are varied. Roll stiffness and Twist stiffness are the most important features. They are dominant, alone explaining 96.61-98.45% of the positioning weight in each genomic feature. The characteristic factors in each combination are larger in number, but weaker in power. Numerous histone modifications play a subtle role for nucleosome positioning.

CONCLUSIONS: The present study provides a more accurate positioning nucleosome-map with higher resolution and a dramatically simplified means to predict and understand intrinsic nucleosome occupancy in different genomic features in human CD4+ cell. Roll stiffness and Twist stiffness are the two most important determinants in all genomic features. They may dominate because they both determine the degree of DNA bending and correlates with many other DNA structural characteristics. Histone modifications play a role of subtle allocation for nucleosome occupancy.

Key Words:
Nucleosome, Positioning, Chromatin.

Introduction

Nucleosome is the fundamental packing unit of chromatin[1] composed of DNA and histones, meaning a DNA superhelix of about 200bp, one histone octamer (composed respectively by 2 copies of the four fundamental histones: H2A, H2B, H3, and H4), and one monomer histone H1. Histone variants such as H2A.Z and H3.3also can be found, especially in active genes[2]. Nucleosome plays a crucial role in the fundamental molecular activities, such as DNA replication[3], DNA repair[4], transcription[5-9], and RNA alternative splicing[10]. The exact position of nucleosome in the genome sequence is recognized to be influenced by several factors, including DNA sequence preferences of nucleosomes themselves[11-15], chromatin remodeling factors[16-18], insulator[19], Alu elements[20,21], TF[22], DNA methylation[23-27], histone modification[28], histone variants[29,30], nucleotide content[31], physicochemical properties, and so on[32-34]. Researches on nucleosome position have been carried out on S. cerevisiae[35,36] and C. elegans genomes[37]. However, most studies on nucleosome positioning factors simply focused on one certain type respectively, such as DNA methylation[38], high G+C content (which correlates positively with intrinsic nucleosome occupancy), nucleotide content, histone modifi-

Corresponding Author: Yanning Cai, Ph.D and Xianhua Dai, Ph.D; e-mail: caiyanningmitl@sina.com

cation, or DNA physical properties[32]. Some were just confined to one certain human genomic region such as exon[39], promotor[40,41], or region near alternative splicing sites[42-44]. Although each of these factors respectively has a certain effect on nucleosome positioning, it's better to take them all into account.

This study aims to investigate the major determinants and their respective strength under the interaction of 3 types of factors across different human genomic features, including the genes of five expressed levels (active, median, middle, low and silent level) and four genomic regions (3' untranslated region (UTR), coding-area, promoter, and intron). Since there are more nucleosomes in intron than other regions, for a better clarification, intron is divided into 5 types (I-active, I-median, I-middle, I-low, and I-silent) according to the express level of the genes they belong to.

Methods

A feature selection step is performed to identify which sequence features influencing nucleosome occupancy or positioning are strongly associated with the *in vitro* nucleosome data of Kaplan et al (Nature 2009; 458: 362-366). The features included: (1) mononucleotide frequency (i.e. G+C content); (2) predicted DNA structural characteristics (each calculated from the dinucleotide content using a simple linear formula[17]); (3) nucleosome positioning and excluding sequences from the literature[10-14]; and (4) the frequency of 4-bp sequences over a 150-bp window. We used 4-mers instead of 5-mers (as in the Kaplan model) were used in order to limit the number of features, and to obtain inputs that correlate independently with nucleosome occupancy (since each 4-mer occurs more frequently than nucleosomes, on average). We identified 130 features that we deemed to be associated with *in vitro* nucleosome occupancy across the yeast genome, including representatives of all categories (1-4) mentioned above.

Based on available data, 97 factors (http://diprogb.fli-leibniz.de/) are selected for analyzing the nucleosome-positioning determinants in this study. They belong to 3 categories, the nucleotide content, the nucleotide physicochemical properties and the histone modification. And the details are as follows: contents of 4 types of mononucleotide, 16 types of dinucleotide and of G/C, purine (A/G), Y (C/T) and Keto (G/T), physical properties of 35 types of single strand dinucleotide, 21 types of histone methylation (H2A.Z included), and 18 types of histone acetylation.

Nucleosomes' Positioning-map and Occupancy scores

Nucleosome positioning is usually dynamic and unstable. The exact position of nucleosome may shift, vibrate or remove with rest or active changes of cell state or cellular processes of DNA replication, transcription, and recombination. Therefore, the certain position of nucleosome should be described by "positioning" and "occupancy". Positioning refers to the times (or probability) of the nucleosome appearing at the exactly same place on the DNA sequence, while occupancy refers to the times (or probability) of the nucleosome appearing at certain locus on the DNA sequence, as is shown in Figure 1[45].

We use available data from genome-wide studies to rebuild a new position map. In order to investigate intragenic nucleosome positioning, we analyzed available sequencing data from human CD4 + T-cells[46,47]. Furthermore, public gene-expression data used in this study are also from human CD4+ T-cells, Schones DE et al[46-48] provides the distribution counts of short sequence reading for nucleosomes at 10 bp resolution in genome-wide human CD4+ cell, which is nucleosome occupancy score (NOScore) at 10 bp resolution, but the exact nucleosome locations and NOScore at single-nucleotide resolution are not given. So, we have to rebuild the nucleosome map and NOScore at single-nucleotide resolution with the ChIP-seq reads data of 21 histone methylation (including histone variant H2A.Z)[49], as well as 18 histone acetylation[50] in human CD4+ cell.

DANPOS[51] is the most optimal algorithm by now designed for nucleosome calling and scoring at single-nucleotide resolution, which is written with python programmer language. It can detect three categories of nucleosome dynamics, such as position shift, fuzziness change, and occupancy change for nucleosomes associated with environmental changes. Besides, it can also be used to give positioning and NOScore for nucleosomes from a single environmental condition. The latter function was used in our studies. The algorithm is also applicable for analyzing histone modifications.

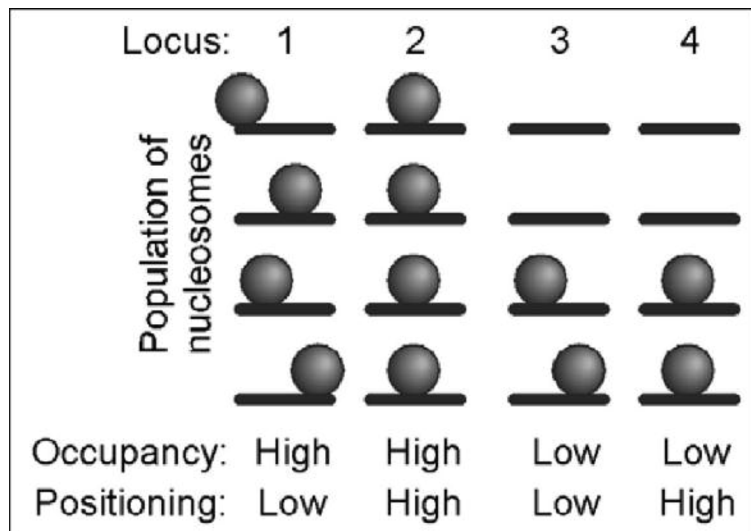


Figure 1. Nucleosome occupancy and positioning. Illustration that distinguishes nucleosome occupancy level and positioning. Spheres represent nucleosomes resting on DNA "lines". Occupancy is high when all lines contain a sphere. Positioning is high when all spheres that are present are aligned vertically.

The Histone Modification Score of Nucleosomes

Similarly, the histone modification (including histone variants) scores of mononucleotide discrimination rate was calculated by the software program danpos[51] and, then, the average nucleotide score of the nucleosome will be the modification score of the nucleosome.

The Nucleotide Content of Nucleosome and its Physical Property Score

After the nucleosome position sequences have been got, the nucleosome DNA sequence (hg18) can be downloaded from the UCSC genome browser (http://genome.ucsc.edu/)[52]. Next, the data of nucleotide contents and physical properties for nucleosome were calculated by the software DiProGB (The Dinucleotide Properties Genome Browser provided on the website, http://diprogb.fli-leibniz.de/)[53].

The Gene Expression Level and Gene Function Regions

The gene expression level is obtained from the Su et al[48]. Based on expression level, the 12000 ref genes are divided into five groups: the active group, the medium group, the middle group, the low group, and the silent group. Downloaded from the UCSC genome browse, the different functional regions of human genome hg18 are 3'UTR, CDS, intron, and the 1000bp region in the upper TSS of each gene are read as promoter.

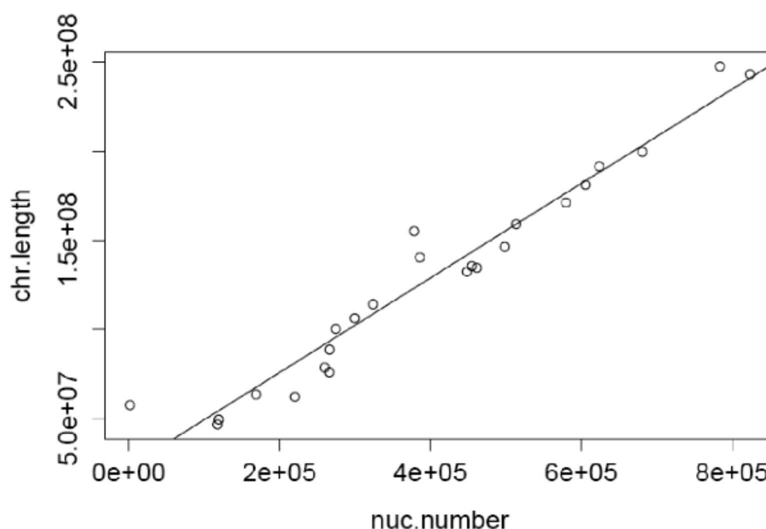


Figure 2. The proportional relation between nucleosome numbers and lengths of chromosomes. The regression parameters of Nu.num = -6.217*104+3.586*10-3 Chr. len both pass t-tests (p <0.05), and their regression formula passes f-test (p =1.529*10-15), Adjusted R2=0.9448.

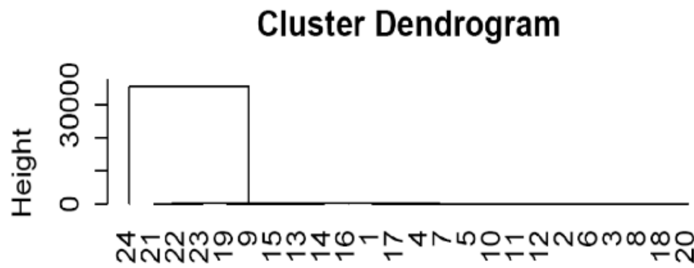


Figure 3. The clustering analysis of nucleosome density of chromosomes.

All the nucleosomes respond their sequence positions to different gene expression levels and different functional regions. The nucleosomes are proportionally shown in Figure 3. The nuleosomes accounts for only 0.006% of all nucleosomes in the 5'UTR region, which is abandoned because of the small proportion.

Statistical Methods and Software

The possible positioning factors for nucleosomes are quite huge in number and complex in interrelationship. In order to clearly find out the main positioning factors and their interrelationship, Lasso (The Least Absolute Shrinkage and Selectionator, Tibshirani (1996)) is utilized here to make linear regression models which is simplified as much as possible (to make the p value of each final positioning factor less than 0.05). Lasso is a kind of shrinkage estimation, and its fundamental theory is that under the restrained condition, that the sum of the regression coefficient absolute value is smaller than a constant, the residual sum of squares is minimized, and consequently some coefficient shrinkages and some models whose coefficients exactly equal to zero will be produced. In this way the weights of some relevant factors can be cut to a maximum extent and therefore allocate greater weight to the most contributing factors, and the over-fitting can be avoided as well.

By the package glmnet of the software R, the Lasso regression analysis of the above various positioning factors is performed. After each regession analysis, the obtained model and its positioning factors will be f-tested ($p<0.01$) and t-tested ($p<0.05$). If there is any positioning factor in the model that cannot be t tested, that factor will be deleted from the model. The rest factors are analyzed by lasso regression analyzed again. The process is repeated until all factors in the model can be tested. Through Lasso regression of nucleosome of different expression levels and different gene regions, finally the nucleosome positioning factor regression models of corresponding expression levels and different gene regions are obtained. The regression coefficient of each factor in the model (hereinafter referred to as positioning weight) stands for their contribution to the nucleosome positioning for the group. After standardization, the above positioning factors of each equation are detected and the final weights of positioning factors are comparable directly. The heavier the factor is, the greater contribution it will make to the nucleosome positioning. The plus or minus before the weight shows its positive or negative role to the nucleosome positioning: plus indicates the induction of nucleosome positioning and positive factors while minus shows rejection of nucleosome positioning and negative factors.

To make a contrast of the specificities of nucleosome positioning factor combinations of each group, layering sampling of random selection of the same number of nucleosomes respectively from the 24 chromosomes are analyzed by the same way of lasso regression and correspondingly the regression factors of each group and their influence degrees (the specified data is omitted here). The regression factors of contrasting groups indicate that the regression factors are special.

Results

9555851 nucleosomes of mononucleotide resolution inside CD4+ cell genome are obtained ($p < 10^{-5}$). The number of nucleosome in each chromosome is shown in Table I. The numbers of each chromosome are different and so is their proportion in all nucleosomes of human genomes. The nucleosomes of chr1-chr12 all respectively exceed 4% of all nucleosomes of genomes and their total sum exceeds 70% of the whole. However, taking the length of each chromosome into consideration, it can be found that the correlation coefficient of the nucleosome numbers of each

Table I. Length of 24 human chromosomes, nucleosome number, their percentage and density.

Chr	Chr Length (bp)	Nucleosome number (Nn)	percentage	Density (L/Nn)
1	247249719	783879	8.20	315
2	242951149	823684	8.62	295
3	199501827	680950	7.13	293
4	191273063	623827	6.53	307
5	180857866	605946	6.34	298
6	170899992	579708	6.07	295
7	158821424	513830	5.34	309
8	146274826	498448	5.22	293
9	140273252	385858	4.04	364
10	135374737	454847	4.76	298
11	134452384	461601	4.83	291
12	132349534	448252	4.69	295
13	114142980	323990	3.39	352
14	106368585	299672	3.14	355
15	100338915	274370	2.87	366
16	88827254	266105	2.78	334
17	78774742	259372	2.71	304
18	76117153	265819	2.78	286
19	63811651	168518	1.76	379
20	62435964	220165	2.30	284
21	46944323	117620	1.23	399
22	49691432	119100	1.25	417
X	154913754	378681	3.96	409
Y	57772954	1609	0.02	35906

chromosome and their respective base pair length is 0.97, and is statistically significant (a=0.01). Software R is used to make linear regression analysis and detect that they are proportional in relation, which is shown in Figure 2.

Although the nucleosome numbers of chromosome are all proportional to their respective lengths, the nucleosome densities of each chromosome, that is the average base pair number of nucleosomes, are different. The nucleosome density of each chromosome is shown in Table I. We can found that the density of chrY is only 1/126 to 1/ 62 that of other chromosomes and the difference is quite remarkable. Hierarchical clustering (hclust) analysis by software R indicates that the nucleosome densities of 24 chromosomes are divided into two types: the type of chrY, the rest chromosomes belong to the other type, as is shown in Figure 3. In the same way, the variance analysis by software R is made to check the nu-

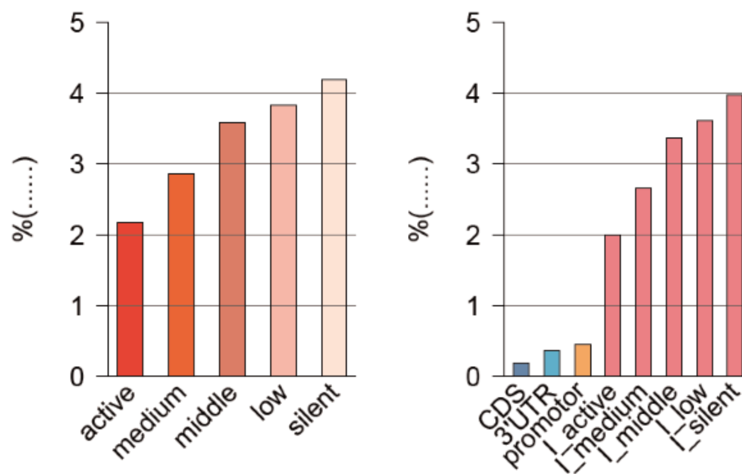


Figure 4. The nucleosome percentage of different gene expression levels (left) and different functional regions (right).

Table II. 10 most important positioning factors for nucleosomes of each level.

active	Roll stiffness, Twist stiffness,H4R3me2, H3K9me3, Slide(Dpc) CG, shift stiffness, H4K20me3, H4K20me1, H2BK5ac
medium	Roll stiffness, Twist stiffness, H4R3me2, 64Shift, H3K9me3 H2AK5ac, CG, Shift stiffness, H4K20me1, H4K91ac
middle	Roll stiffness, Twist stiffness, H4R3me2, H3K27me3, 64Shift CG, Shift stiffness, H4K91ac, H4K20me1, H4K20me3
low	Roll stiffness, Twist stiffness, 64Shift, H4R3me2, H3R2me2 CG, Shift stiffness, H4K91ac, H4K20me1, H3K4me3
silent	Roll stiffness, Twist stiffness, H3K9me3, 64Shift, H4R3me2 H3K36me1, CG, H3K27me2, Shift stiffness, H4K20me1

For nucleosomes of each expression level, 5 most important positive factors and 5 most important negative ones are in the upper and lower lines, respectively.

cleosome density of 24 chromosomes and we can find that the density of chrY is notably different from those of other chromosomes at a=0.01 (p = 0), indicating that as the unique chromosome for male, the nucleosome positioning system of it is possibly different from that of other chromosomes. Consequently, its internal gene function and expression may also be strikingly different from those of other chromosomes.

Distribution on Genes

Through gene projection, 1591486 nucleosomes for positioning factor analysis are obtained. All these nucleosomes are divided into groups based on their gene expression power and gene functional regions and the proportion of nucleosome of each group accounting for all genome nucleosomes are obtained as shown in Figure 4. With the decrease of gene expression power and functional region expression level, the nucleosome proportion increases. As for the complete no expression power intron, its nucleosome percentage also rises with the decrease of its gene expression level.

To get the exact influential factors of nucleosome positioning of human CD4+ cell genes with different expression levels or different gene regions, the whole explaining ability of each model is sacrificed and the lasso regression has been used repeatedly on each nucleosome group to get the final p value of each positioning factor smaller than 0.05.

To better analyze the positioning factor weights of each model group, we divide the positioning weights into 4 classes according to their quantity: **Overwhelming main forces**: weight ≥ 10. The overwhelming main forces exert huge effect on nucleosome positioning, and each factor weight accounts for about 20-80% of the whole group weight. **Main forces**: 1 ≤ weight < 10. The single main force power lies between the overwhelming main forces and the common factor forces. **Common forces**: 0.1 ≤ weight < 1. Each common factor force weight accounts for about 1% of the whole group weight. **Tiny adjusting forces**: weight < 0.1. The factors of this group exert very tiny influence on nucleosome positioning. Each tiny adjusting force weight accounts for about 0.1% or even less of the whole group weight.

Analyzed by Different Gene Expression Levels

The five groups of nucleosome positioning factors of different gene expression power and their weights are shown in Figure 5. The five groups of positioning factors vary in number and type, ranging from 19 to 23. Among the largest weights of each positioning model group, the first five positive factors and the first negative factors are shown in Table II.

Analyzed by different gene functional regions the eight groups of nucleosome positioning factors of different gene functional regions and their weights are shown in Figure 6.

Among the largest weights of each positioning model group, the first five positive factors and the first negative factors are shown in Table III.

Discussion

The above model groups, share the following similarities:

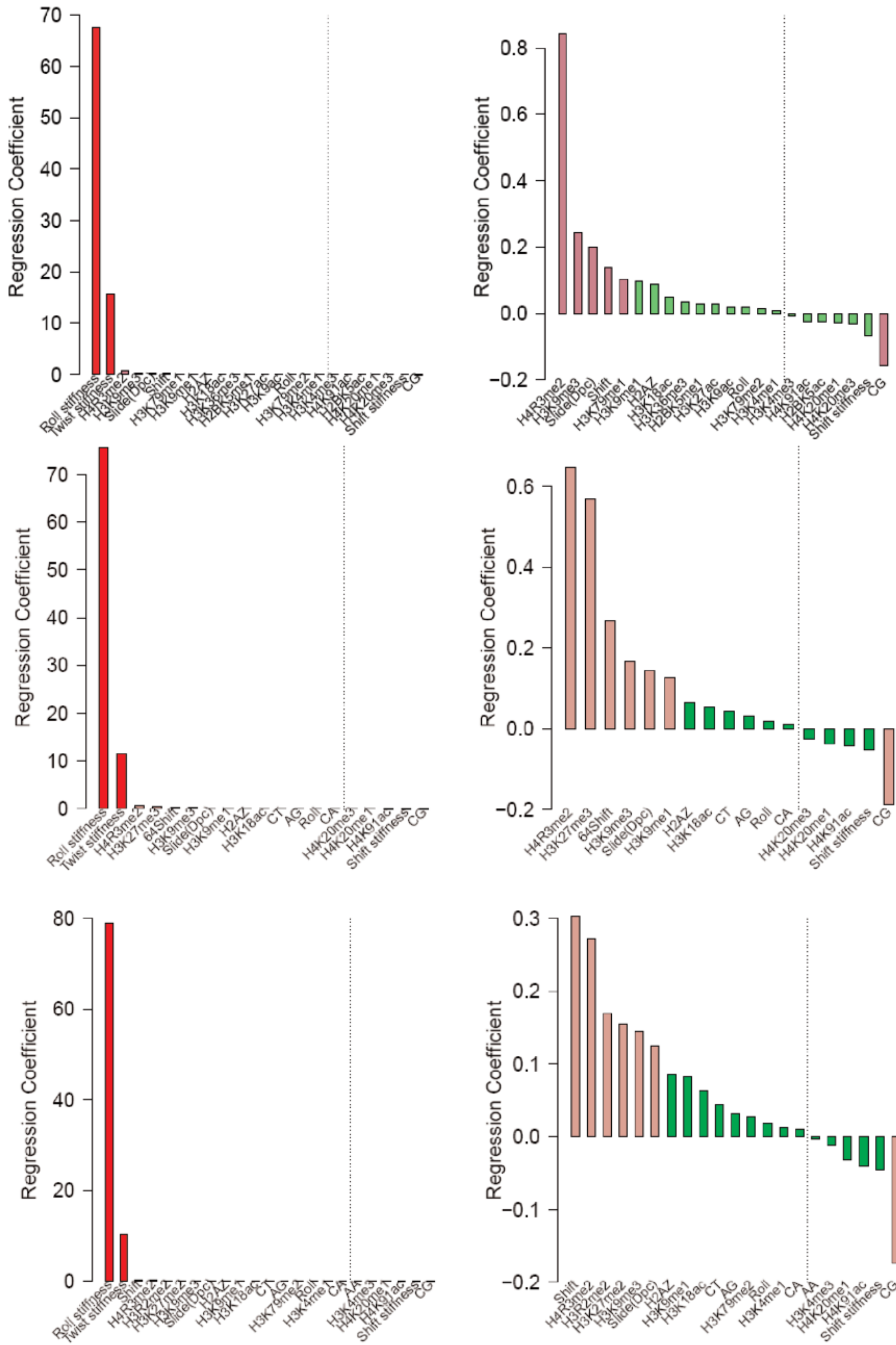
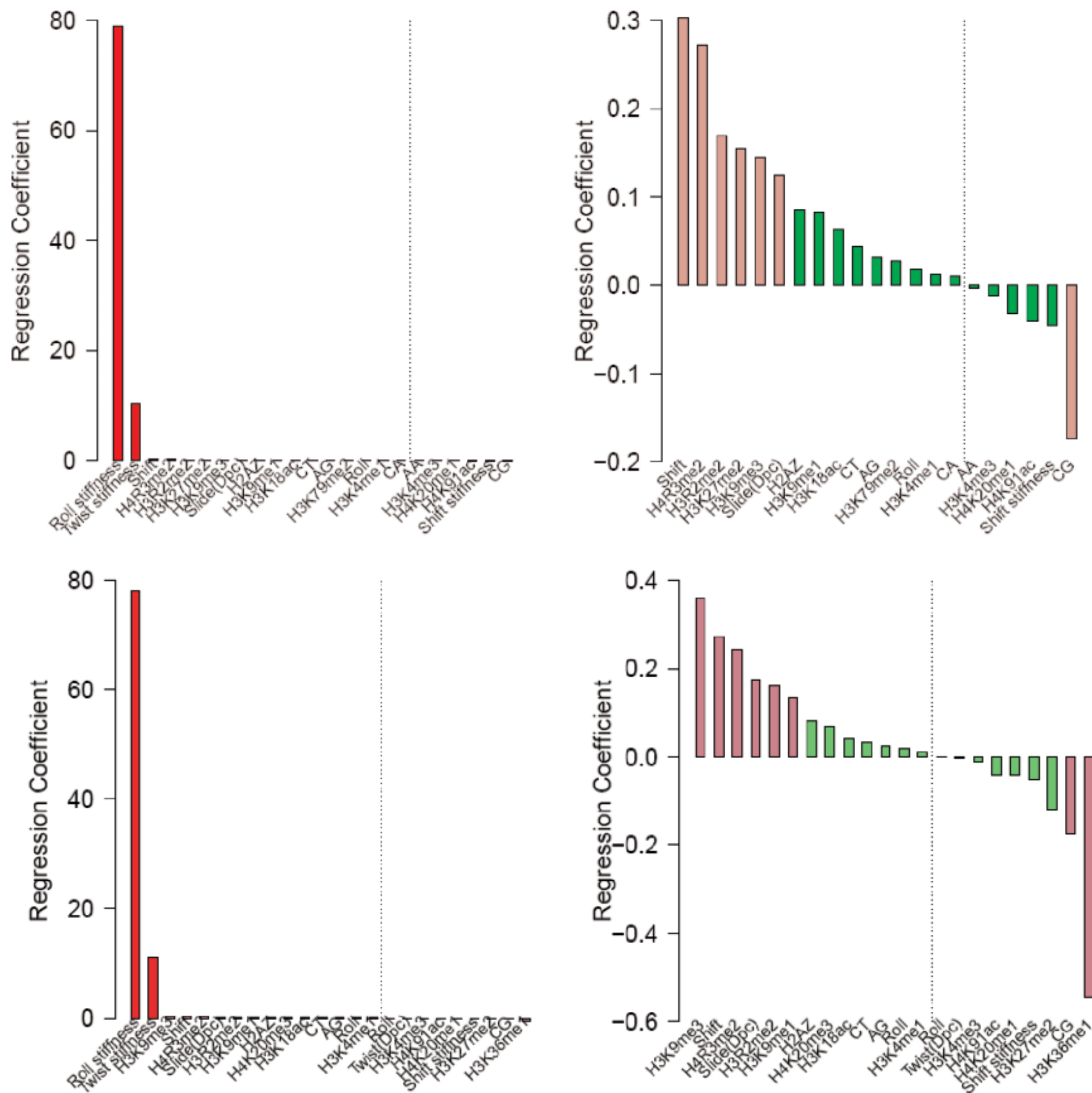


Figure 5. The histogram of nucleosome positioning factors of different gene expression levels.

The figure continued



Nucleosome positioning factors for active, medium, middle, low, and silent gene, respectively. The right is zoom in of the left, respectively.

Figure 5. *(Continued).*

In each group, there must be the overwhelming main positioning factors of Roll stiffness and Twist stiffness, and always there are CT and AG, the two tiny adjusting positive factors and CG and Shift stiffness, the two common and tiny adjusting negative factors. 2. Every group is short of main positioning factors. 3. In each group, there must be the common positioning factors of H4R3me2. The common positioning factors mainly appear as positive factors, seldom as negative factors. 4. In each group, the number of the tiny adjusting positioning factors is the largest, accounting for nearly 50% of the total.

The above similarities indicate that nucleosome positioning gets great support: the two overwhelming main forces appear only as positive factors, small in number, but huge in power, and their total weight accounts for 80% to 95% of each group weight. The common forces mainly appear as positive factors, seldom as negative factors; the tiny adjusting factors are great in number, but tiny in function and in total sum, and they are the main composing force of negative force. Therefore, the positive force composed by overwhelming main forces, common forces and tiny adjusting forces far outweighs the negative

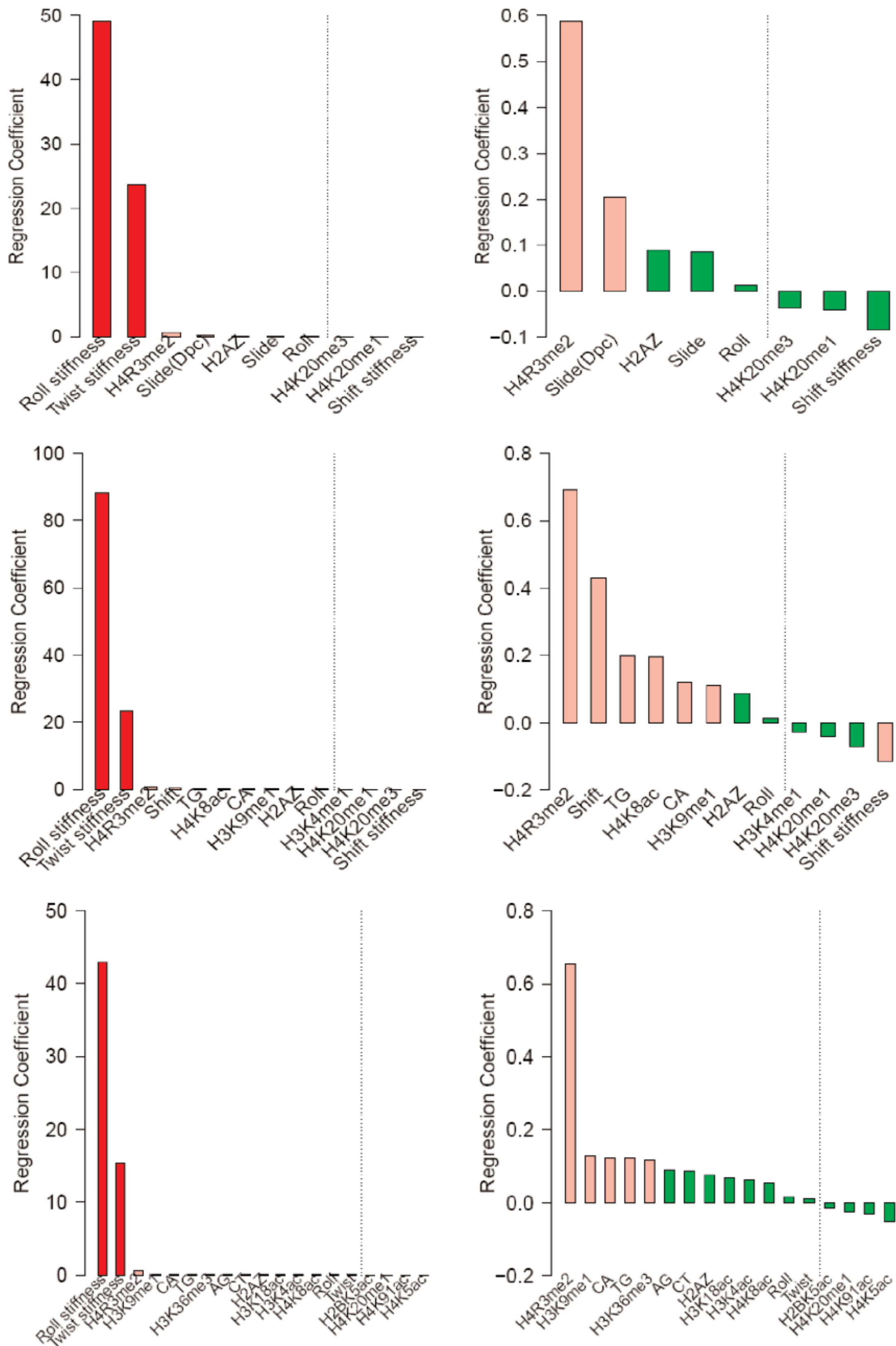


Figure 6. The histogram of nucleosome positioning factors of different genomic regions.

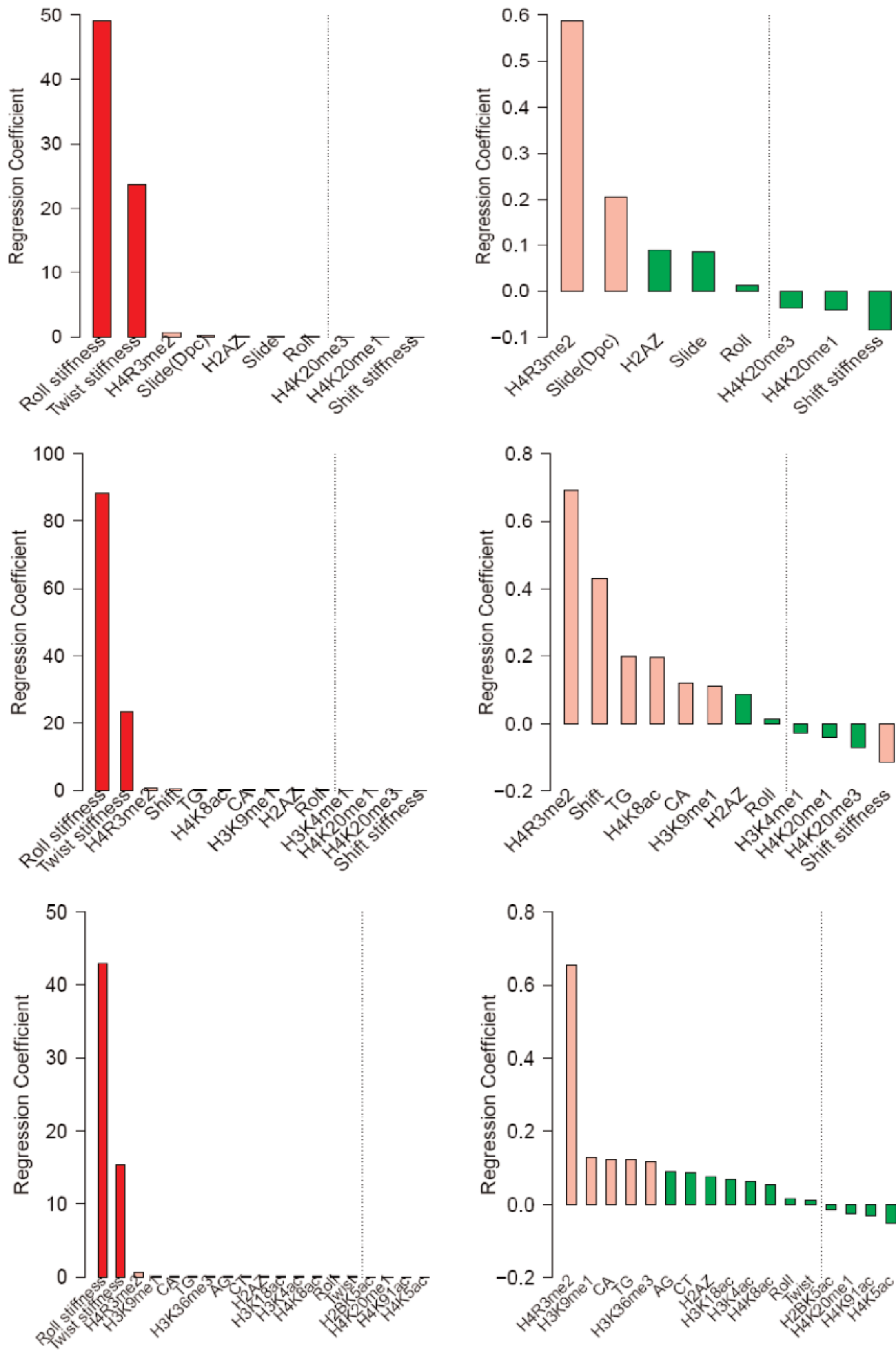


Figure 6. The histogram of nucleosome positioning factors of different genomic regions.

Figure continued

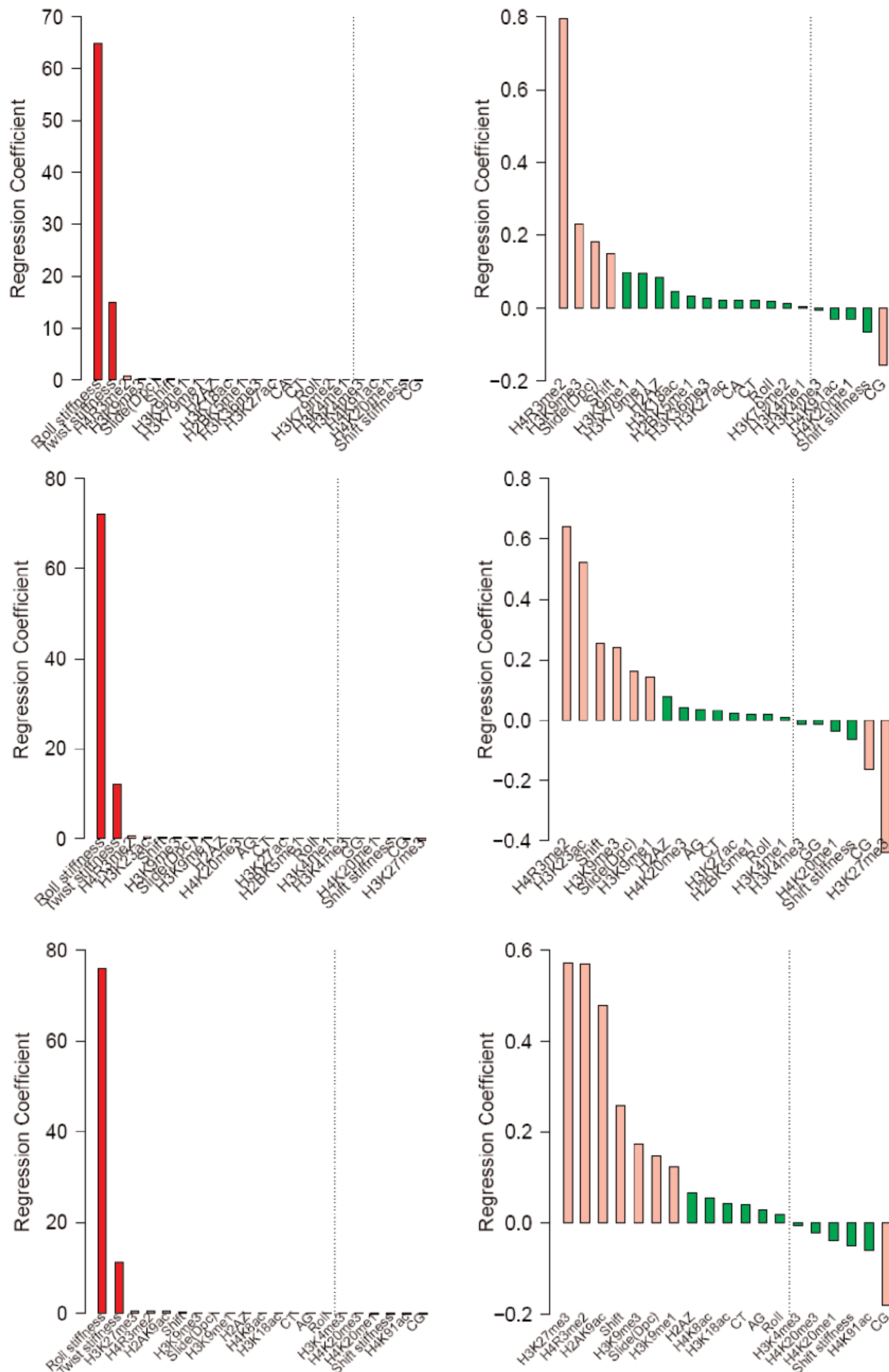
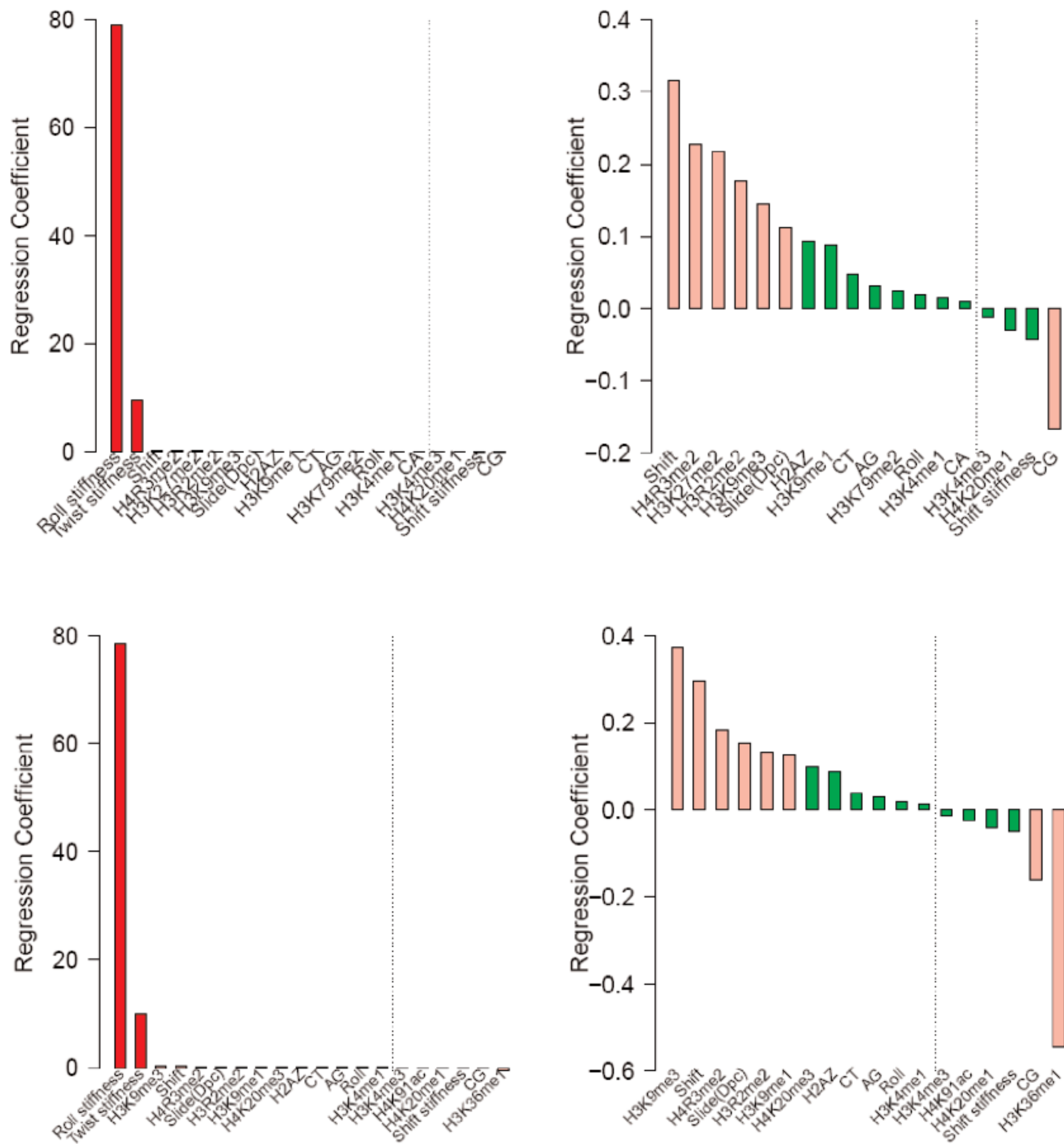


Figure 6. *(Continued).*



Nucleosome positioning factors for 3'UTR, CDS, promotor, I_active, I_medium, I_middle, I_low and I_silent regions, respectively. The right is zoom in of the left, respectively.

Figure 6. *(Continued).*

force composed by common forces and tiny adjusting forces, which makes it easy for nucleosomes to be positioned. The contrast analysis of nucleosome positioning factors also demonstrates in an opposite way that exactness of uncleosome positioning identified by DANPOS. As for the lack of main positioning forces in each group, it may be relevant to the limited ranges of positioning factors in this study. Further studies need to be carried out when conditions are permitted.

Besides the above similarities, there are also some differences in each group of positioning factors.

1. Each group of positioning factors is different in number, ranging from 10 to 24. 2. Except the above mentioned a few stable positioning factors, the rest positioning factors in each group do not always remain the same. 3. The weights of the same positioning factors in each group are different, even opposite in function. Take H4K20me3 as an example, a factor is positive in

Table III. 10 most important positioning factors for nucleosomes of each region.

3' UTR	Roll stiffness, Twist stiffness, H4R3me2, Slide(Dpc), H2AZ Shift stiffness, H4K20me1, H4K20me3 (*)
CDS	Roll stiffness, Twist stiffness, H4R3me2, Shift,TG Shift stiffness, H4K20me3, H4K20me1, H3K4me1(*)
Promotor	Roll stiffness, Twist stiffness, H4R3me2, H3K9me1, CA H4K5ac, H4K91ac, H4K20me1, H2BK5ac(*)
I_active	Roll stiffness, Twist stiffness, H4R3me2, H3K9me3, Slide(Dpc) CG, Shift stiffness, H4K20me1, H4K91ac, H3K4me3
I_medium	Roll stiffness, Twist stiffness, H4R3me2, H3K23ac, Shift H3K27me3, CG, Shift stiffness, H4K20me1, GG
I_middle	Roll stiffness, Twist stiffness, H3K27me3, H4R3me2, H2AK9ac CG, H4K91ac, Shift stiffness, H4K20me1, H4K20me3
I_low	Roll stiffness, Twist stiffness, Shift, H4R3me2, H3K27me2 CG, Shift stiffness, H4K20me1, H3K4me3(*)
I_silent	Roll stiffness, Twist stiffness, H3K9me3, Shift, H4R3me2 H3K36me1, CG, Shift stiffness, H4K20me1, H4K91ac

(*): the negetive factors are less than 5.
For nucleosomes of each region, 5 most important positive factors and 5 most important negative ones are in the upper and lower lines, respectively.

the medium group but negative in the middle group. 4. In each group, the quantities of nucleotide, nucleotide properties and histone modification are changeable in number.

Although each group of positioning factors varies in number, weight and even in function detection, the total positive factor sum and their weights are far from being challenged by that of the negative factors, which is finally able to determine the nucleosome positioning. The differences of each positioning factor group indicate the flexibility of nucleosome positioning, which can meet the needs of genes to have different gene expression levels and different gene functions.

The analysis of the similarities and differences of nucleosome positioning well reflects the complexity of nucleosome positioning system, as well as its principle and flexibility.

DANPOS, the optimal method at present is utilized to identify 9555852 nucleosomes of human CD4+ cell genomes. These nucleosomes are uniform distributed into chromosomes according to length of chromosomes; however, the nucleosome densities of each chromosome are not even, especially the density of chrY is far lower than those of other chromosomes, which indicates that there must be a certain hidden nucleosome positioning system in chrY, which cuts down the nucleosome density in chrY. The nucleosomes will influence the gene function, which reminds us to investigate the different functions between the nucleosomes of chrY and those of other chromosomes.

Through the analysis of nucleosome distribution of genes, it can be found that the higher the gene expression level is, the lower the nucleosome proportion will be, and vice versa. This phenomenon is coherent with our current viewpoint, because the higher the gene expression level is, the more frequent the internal molecular activities of transcription and translation that nucleosome involves will be; the less bindings nucleosomes will get, and thus nucleosome as the DNA binding compositions will not appear so frequently. On the other side, the lower the gene expression level is, the more bindings from nucleosome DNA will get.

Through analysis of nucleosome positioning of genes, we found that the leading forces for nucleosome positioning of human CD4+ cell genes are Roll stiffness and Twist stiffness, however, there is not enough main force grade between the overwhelming main forces and the common forces. It might be because that the factors lacking force grade are not covered in this study and, thus, is the result of all groups of positioning factors lacking in the positioning factors this grade.

To meet the needs of different gene expression levels and different gene functions, nucleosomes of human CD4+ cell genomes have evolved various nucleosome positioning strategies. In these strategies, the physical properties of dinucleotide are the most important ones, always with Roll stiffness and Twist stiffness as the overwhelming leading positive factors. The negative positioning factors are always taking dinucleotide CG quantity and shift stiffness as leading factors. The histone modification varies with the change of gene expression levels and function regions, playing a tiny role of adjusting for nucleosome positioning.

Conclusions

This study provides a higher resolution and more accurate positioning nucleosome- map and a dramatically simplified means to predict and understand intrinsic nucleosome occupancy in different genomic features in human CD4+ cell. Roll stiffness and Twist stiffness are the two most important determinants in all genomic features. They may dominate because they both determine the degree of DNA bending and correlate with many other DNA structural characteristics. Histone modifications play a role of subtle allocation for nucleosome occupancy.

Acknowledgements

This paper is supported by National Nature Science Fund of China, Grant: 61174163.

Conflict of Interest

The Authors declare that they have no conflict of interests.

Reference

1) Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 1997; 389: 251-260.
2) Malik HS, Henikoff S. Phylogenomics of the nucleosome. Nat Struct Biol 2003; 10: 882-891.
3) Eaton ML, Galani K, Kang S, Bell SP, MacAlpine DM. Conserved nucleosome positioning defines replication origins. Genes Dev 2010; 24: 748-753.
4) Ying H, Epps J, Williams R, Huttley G. Evidence that localized variation in primate sequence divergence arises from an influence of nucleosome placement on DNA repair. Mol Biol Evol 2010; 27: 637-649.
5) He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q, Zhang Y, Xu K, Ni M, Lupien M, Mieczkowski P, Lieb JD, Zhao K, Brown M, Liu XS. Nucleosome dynamics define transcriptional enhancers. Nat Genet 2010; 42: 343-347.
6) Jin J, Bai L, Johnson DS, Fulbright RM, Kireeva ML, Kashlev M, Wang MD. Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. Nat Struct Mol Biol 2010; 17: 745-752.
7) Svaren J, Horz W. Regulation of gene expression by nucleosomes. Curr Opin Genet Dev 1996; 6: 164-170.
8) Jiang C, Pugh BF. Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet 2009; 10: 161-172.
9) Tirosh I, Barkai N. Two strategies for gene regulation by promoter nucleosomes. Genome Res 2008; 18: 1084-1091.
10) Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcarcel J, Guigo R. Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol 2009; 16: 996-1001.
11) Widom J. Role of DNA sequence in nucleosome stability and dynamics. Q Rev Biophys 2001; 34: 269-324.
12) Fernandez AG, Anderson JN. Nucleosome positioning determinants. J Mol Biol 2007; 371: 649-668.
13) Chung HR, Vingron M. Sequence-dependent nucleosome positioning. J Mol Biol 2009; 386: 1411-1422.
14) Fraser RM, Keszenman-Pereyra D, Simmen MW, Allan J. High-resolution mapping of sequence-directed nucleosome positioning on genomic DNA. J Mol Biol 2009; 390: 292-305.
15) Travers A, Hiriart E, Churcher M, Caserta M, Di Mauro E. The DNA sequence-dependence of nucleosome positioning in vivo and in vitro. J Biomol Struct Dyn 2010; 27: 713-724.
16) Mayes K, Qiu Z, Alhazmi A, Landry JW. ATP-dependent chromatin remodeling complexes as novel targets for cancer therapy. Adv Cancer Res 2014; 121: 183-233.
17) Guyon JR, Narlikar GJ, Sullivan EK, Kingston RE. Stability of a human SWI-SNF remodeled nucleosomal array. Mol Cell Biol 2001; 21: 1132-1144.
18) Becker PB, Horz W. ATP-dependent nucleosome remodeling. Annu Rev Biochem 2002; 71: 247-273.
19) 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.
20) Englander EW, Howard BH. Nucleosome positioning by human Alu elements in chromatin. J Biol Chem 1995; 270: 10091-10096.
21) Tanaka Y, Yamashita R, Suzuki Y, Nakai K. Effects of Alu elements on global nucleosome positioning in the human genome. BMC Genomics 2010; 11: 309.
22) Kim HD, O'Shea EK. A quantitative model of transcription factor-activated gene expression. Nat Struct Mol Biol 2008; 15: 1192-1198.

23) Pennings S, Allan J, Davey CS. DNA methylation, nucleosome formation and positioning. Brief Funct Genomic Proteomic 2005; 3: 351-361.

24) Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, Casero D, Bernal M, Huijser P, Clark AT, Kramer U, Merchant SS, Zhang X, Jacobsen SE, Pellegrini M. Relationship between nucleosome positioning and DNA methylation. Nature 2010; 466: 388-392.

25) Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee TH. DNA methylation increases nucleosome compaction and rigidity. J Am Chem Soc 2010; 132: 1782-1783.

26) Lee JY, Lee TH. Effects of DNA methylation on the structure of nucleosomes. J Am Chem Soc 2012; 134: 173-175.

27) Jimenez-Useche I, Ke J, Tian Y, Shim D, Howell SC, Qiu X, Yuan C. DNA methylation regulated nucleosome dynamics. Sci Rep 2013; 3: 2121.

28) Hayes JJ, Clark DJ, Wolffe AP. Histone contributions to the structure of DNA in the nucleosome. Proc Natl Acad Sci U S A 1991; 88: 6829-6833.

29) Guillemette B, Bataille AR, Gevry N, Adam M, Blanchette M, Robert F, Gaudreau L. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. PLoS Biol 2005; 3: e384.

30) Henikoff S. Labile H3.3+H2A.Z nucleosomes mark 'nucleosome-free regions'. Nat Genet 2009; 41: 865-866.

31) Tillo D, Hughes TR. G+C content dominates intrinsic nucleosome occupancy. BMC Bioinformatics 2009; 10: 442.

32) Miele V, Vaillant C, d'Aubenton-Carafa Y, Thermes C, Grange T. DNA physical properties determine nucleosome occupancy from yeast to fly. Nucleic Acids Res 2008; 36: 3746-3756.

33) Deniz O, Flores O, Battistini F, Perez A, Soler-Lopez M, Orozco M. Physical properties of naked DNA influence nucleosome positioning and correlate with transcription start and termination sites in yeast. BMC Genomics 2011; 12: 489.

34) Gan Y, Guan J, Zhou S, Zhang W. Structural features based genome-wide characterization and prediction of nucleosome organization. BMC Bioinformatics 2012; 13: 49.

35) Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. Genome-scale identification of nucleosome positions in S. cerevisiae. Science 2005; 309: 626-630.

36) Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. A high-resolution atlas of nucleosome occupancy in yeast. Nat Genet 2007; 39: 1235-1244.

37) Johnson SM, Tan FJ, McCullough HL, Riordan DP, Fire AZ. Flexibility and constraint in the nucleosome core landscape of Caenorhabditis elegans chromatin. Genome Res 2006; 16: 1505-1516.

38) Choi JK, Bae JB, Lyu J, Kim TY, Kim YJ. Nucleosome deposition and DNA methylation at coding region boundaries. Genome Biol 2009; 10: R89.

39) Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res 2009; 19: 1732-1741.

40) Schmid CD, Bucher P. ChIP-Seq data reveal nucleosome architecture of human promoters. Cell 2007; 131: 831-832; author reply 832-833.

41) Dennis JH, Fan HY, Reynolds SM, Yuan G, Meldrim JC, Richter DJ, Peterson DG, Rando OJ, Noble WS, Kingston RE. Independent and complementary methods for large-scale structural analysis of mammalian chromatin. Genome Res 2007; 17: 928-939.

42) Chen W, Luo L, Zhang L. The organization of nucleosomes around splice sites. Nucleic Acids Res 2010; 38: 2788-2798.

43) Tillo D, Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Field Y, Lieb JD, Widom J, Segal E, Hughes TR. High nucleosome occupancy is encoded at human regulatory sequences. PLoS One 2010; 5: e9129.

44) Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A. Determinants of nucleosome organization in primary human cells. Nature 2011; 474: 516-520.

45) Zhang Z, Wippo CJ, Wal M, Ward E, Korber P, Pugh BF. A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. Science 2011; 332: 977-980.

46) Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. Dynamic regulation of nucleosome positioning in the human genome. Cell 2008; 132: 887-898.

47) Schones DE, Zhao K. Genome-wide approaches to studying chromatin modifications. Nat Rev Genet 2008; 9: 179-191.

48) Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 2004; 101: 6062-6067.

49) Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. Cell 2007; 129: 823-837.

50) Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 2008; 40: 897-903.

51) Chen K, Xi Y, Pan X, Li Z, Kaestner K, Tyler J, Dent S, He X, Li W. DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res 2013; 23: 341-351.

52) Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res 2002; 12: 996-1006.

53) Friedel M, Nikolajewa S, Suhnel J, Wilhelm T. DiProDB: a database for dinucleotide properties. Nucleic Acids Res 2009; 37: D37-40.