

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Nucleosome Positioning and Its Role in Gene Regulation in Yeast

Hongde Liu, Weiheng Ma, Jiahao Xie, Huamei Li,
Kun Luo, Donghui Luo, Lei Liu and Xiao Sun

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.70935>

Abstract

Nucleosome, composed of a 147-bp segment of DNA helix wrapped around a histone protein octamer, serves as the basic unit of chromatin. Nucleosome positioning refers to the relative position of DNA double helix with respect to the histone octamer. The positioning has an important role in transcription, DNA replication and other DNA transactions since packing DNA into nucleosomes occludes the binding site of proteins. Moreover, the nucleosomes bear histone modifications thus having a profound effect in regulation. Nucleosome positioning and its roles are extensively studied in model organism yeast. In this chapter, nucleosome organization and its roles in gene regulation are reviewed. Typically, nucleosomes are depleted around transcription start sites (TSSs), resulting in a nucleosome-free region (NFR) that is flanked by two well-positioned H2A.Z-containing nucleosomes. The nucleosomes downstream of the TSS are equally spaced in a nucleosome array. DNA sequences, especially 10–11 bp periodicities of some specific dinucleotides, partly determine the nucleosome positioning. Nucleosome occupancy can be determined with high throughput sequencing techniques. Importantly, nucleosome positions are dynamic in different cell types and different environments. Histones depletions, histones mutations, heat shock and changes in carbon source will profoundly change nucleosome organization. In the yeast cells, upon mutating the histones, the nucleosomes change drastically at promoters and the highly expressed genes, such as ribosome genes, undergo more change. The changes of nucleosomes tightly associate the transcription initiation, elongation and termination. H2A.Z is contained in the +1 and -1 nucleosomes and thus in transcription. Chaperon Chz1 and elongation factor Spt16 function in H2A.Z deposition on chromatin. The chapter covers the basic concept of nucleosomes, nucleosome determinant, the techniques of mapping nucleosomes, nucleosome alteration upon stress and mutation, and Htz1 dynamics on chromatin.

Keywords: nucleosome, 10–11 bp periodicities, nucleosome-free region, MNase-sequencing, histone mutation, H2A.Z

1. Basic conceptions about chromatin and nucleosome

1.1. Chromatin of eukaryotic DNA, nucleosome, nucleosome compositions, and histone

1.1.1. Chromatin

Eukaryotic DNA exists as chromatin structure, which is composed of DNA and proteins in the nucleus (**Figure 1**). The proteins can divide into histone proteins (H1/H5, H2A, H2B, H3, and H4) and non-histone ones. The former acts as core which DNA winds. The histone winding with DNA acts as a ball that forms the basic structure. Non-histone proteins have three main functions: (1) enzyme used in different DNA activities, for example, DNA reparation, duplication, and translation, such as DNA polymerase and DNA ligase; (2) scaffold proteins. They play the role of skeleton; and (3) other motor proteins. All play essential roles in cell structure and regulatory functions that make life possible.

Since the package of DNA must be rapidly accessible so that protein machinery is able to interact with DNA in replication, transcription, DNA repair, and recombination, the chromatin is highly different in different cells and different periods. Chromatin can be divided into euchromatin and heterochromatin. Heterochromatin is characterized by its high compactness and its inhibitory effect on DNA transactions such as gene expression. However, according to Volpe et al. [1], many of them actually can transcribe but are silenced by RNA by RNA-induced transcriptional silencing (RITS). Euchromatin is the chromatin which is not packaged tightly like heterochromatin so it is more accessible. Most of chromatin are euchromatin (92% of the human genome [2]); it contains activating genes and changes its condensation during cell cycle.

1.2. Nucleosome and histones

Nucleosomes are the basic unit of chromatin. The nucleosome consists of 147 bp of DNA wrapped around an octamer of histones, with two copies of each H2A, H2B, H3, and H4, and about 1.65 superhelical turns arranged in a left-handed manner [3] (**Figure 2**). The nucleosome cores are connected by linker DNA, which typically ranges from 10 bp to 90 bp in length, to form a “beads-on-a-string” nucleosomal array with a diameter of 11 nm [4]. At the entry and the exit of the nucleosome, H1 binds the DNA to make the nucleosomes fixate in the space.

The “tails” of these histone proteins stick out, especially H3 and H4, where they can be modified in many ways. Modifications of the tail include methylation, acetylation, phosphorylation, ubiquitination, SUMOylation, citrullination, and ADP-ribosylation. Through these chemical modifications, histone can change its interaction with DNA. Interestingly, many of these modifications have fixed position and function (**Table 1**).

All histones have its variants (**Figure 3**) [5], and they have different biological function compared to canonical histones. Exchanging with canonical histones dynamically also plays an important role in regulation of gene expression.

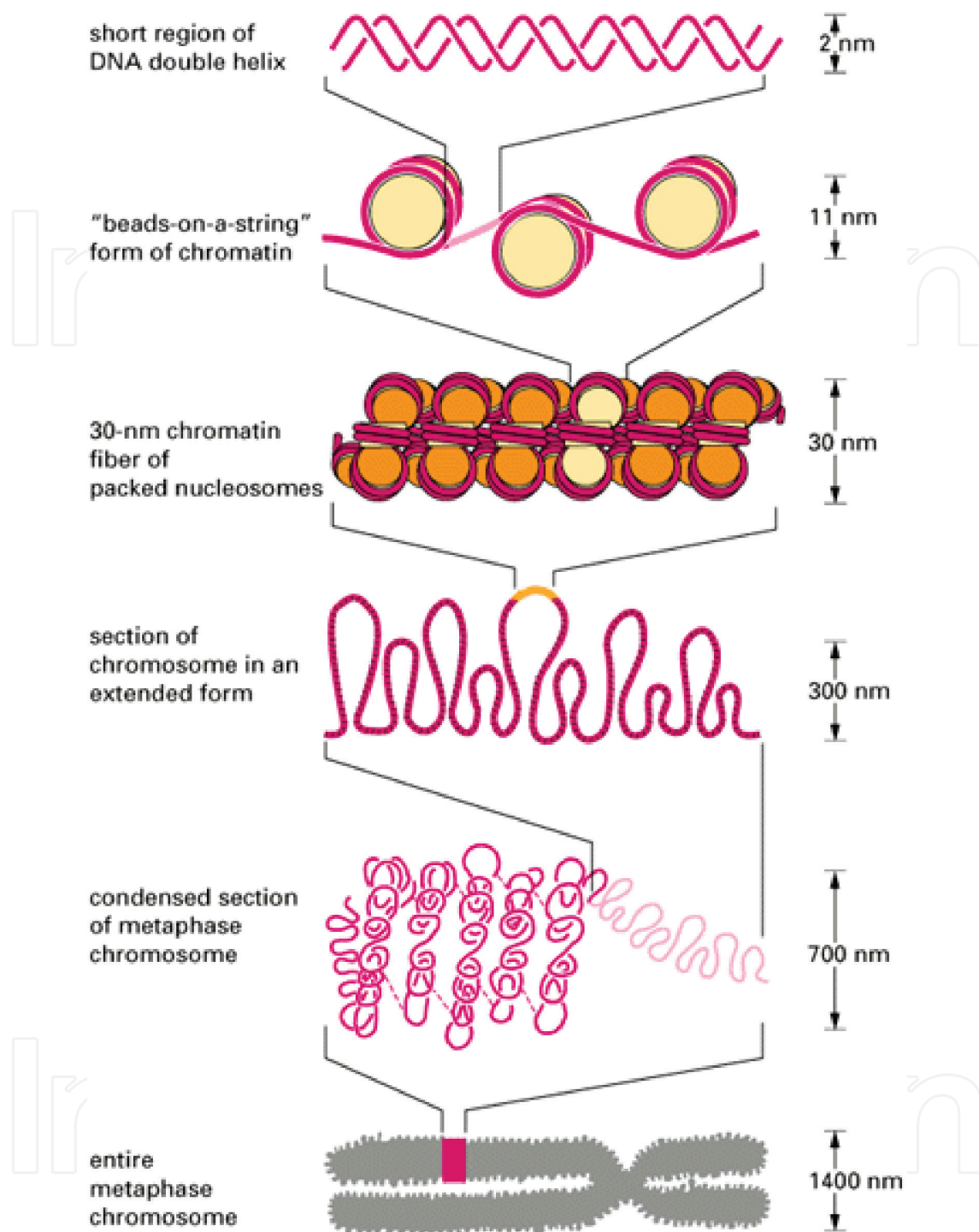


Figure 1. Chromatin structure (<http://csma31.csm.jmu.edu/chemistry/faculty/mohler/chromatin.htm>).

1.3. Research history of nucleosomes, especially in yeast

Clark and Felsenfeld first used staphylococcal nuclease to digest chromatin in 1971 and found that some regions were sensitive to nuclease while some were insensitive; insensitive regions were homogeneous, suggesting it contains subunits. Then Hewish and Burgoyne Researchers

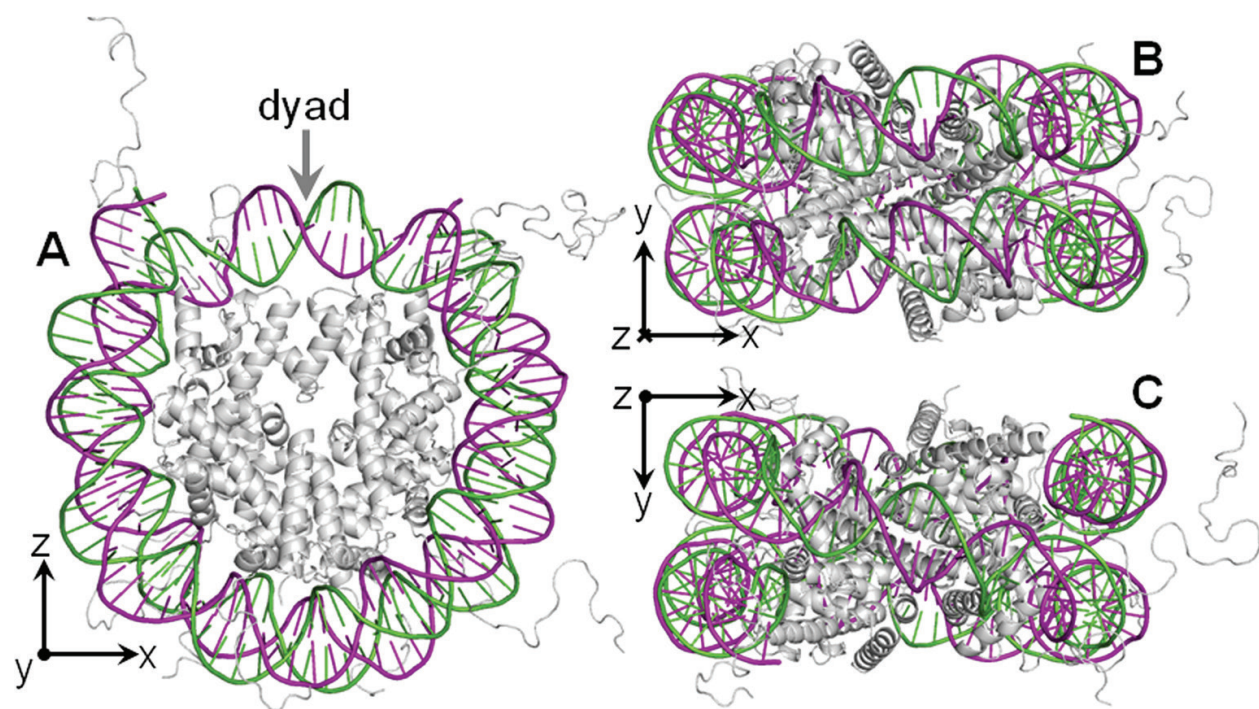


Figure 2. These views of NCP147 (147-bp nucleosome core particle), at Å resolution, show the two strands of the double-helix in purple and green, with the protein core in gray. (A) The curvature of DNA around the histone core, with the dyad at the top, center; (B) a rotation of the particle, showing the adjacent segments of DNA, opposite the dyad; and (C) a rotation in the opposite direction, showing the DNA crossing over the dyad. As indicated by the coordinate system axes, in (A) the y-axis is pointing out of the page, in (B) the z-axis is pointing into the page, and in (C) the z-axis is pointing out of the page [6].

Type of modification	Histone								
	H3K4	H3K9	H3K14	H3K27	H3K79	H3K36	H4K20	H2BK5	H2BK20
Mono-Me	A	A		A	A		A	A	
Di-Me		R		R	A				
Tri-Me	A	R		R	A & R	A		R	
Ac		A	A	A					A

Note: Me, methylation; Ac, acetylation; A, activation; R, repression.

Table 1. Different modifications in transcription regulation (from) [7–10].

in previous study digested the nuclei with endogenous nuclease and isolated DNA from the nucleus. As a result, a series of DNA fragments were found, which corresponded to a basic unit of about 200 bp, indicating that histones bind to DNA in a regular manner which results in only certain restricted regions are sensitive to nuclease.

Kornberg and Thomas then digested the chromatin with a small cellulase in 1974 and centrifuged it to obtain monomers, dimers, trimers, and tetramers. Using electron microscopy, the monomer was observed as a 10 nm body, and the dimer was two associated bodies. The same

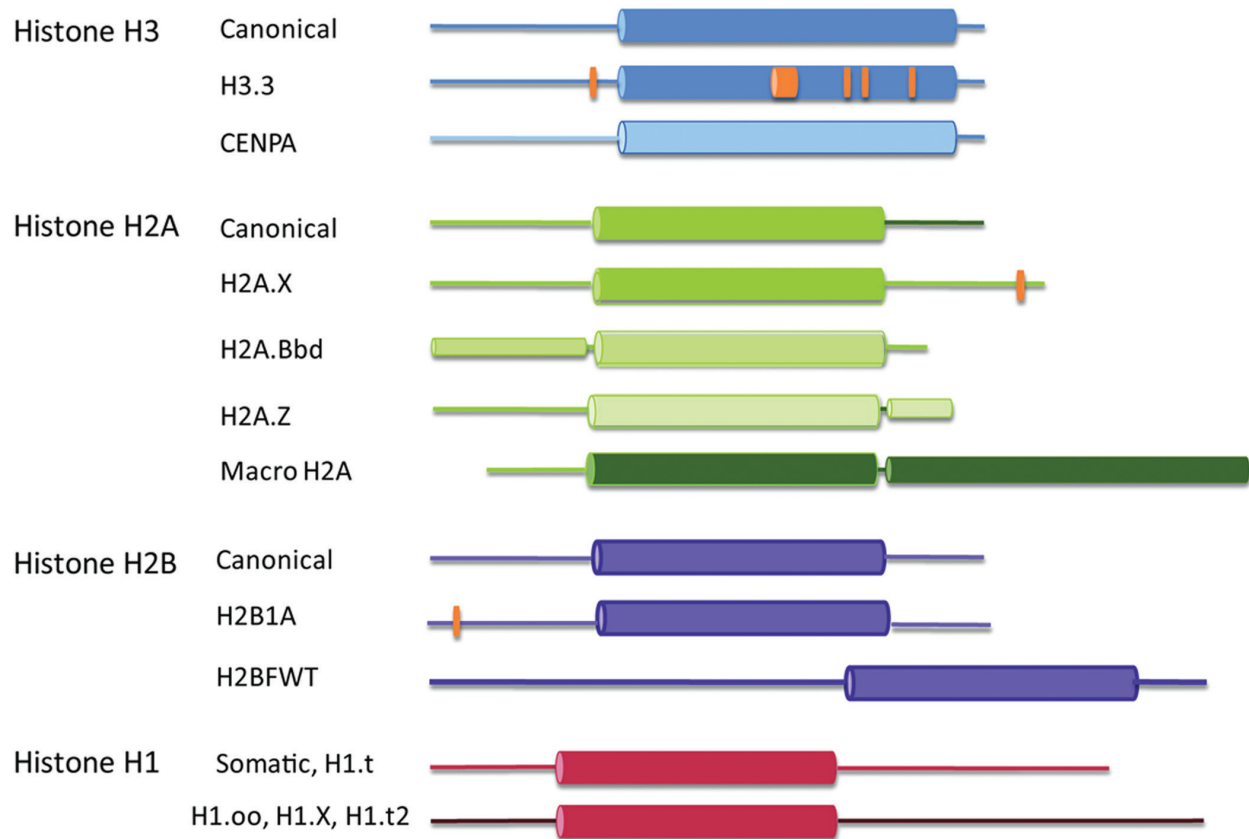


Figure 3. All of the histone variants contain a highly conserved histone fold domain and vary mainly in their C- and N-terminal sequences. The above shown is a schematic comparing histone variant sequences. Boxes represent the histone fold domain and orange lines represent site-specific sequence variations. Histones that are in different shades of the same color are from the same histone family but have large differences in sequence [5].

trimer and tetramer consisted of three bodies and four bodies, respectively, indicating that the structure consisting of 200 bp DNA was “rope beads” units, which are called nucleosomes.

Through all kinds of experiments, it was found that the structure of the nucleosome core is relatively invariant from yeast to metazoans [11, 12] containing a 147 bp DNA wrapped around a histone protein octamer. In 2005, Yuan et al. developed a tiled microarray approach to identify at high resolution the translational positions of 2278 nucleosomes over 482 kb of *Saccharomyces cerevisiae* DNA, including almost all of chromosome III and 223 additional regulatory regions [13]. However, the study of the location of nucleosomes is quite time-consuming and costly if using experiments alone, so the researchers began to build nucleosome positioning prediction model based on the existing experimental data [14]. In yeast genome, Segal et al. found that DNA sequence contains ~10-bp period pattern of AA-TT-TA/GC dinucleotides [15]. Nucleosomal DNA sharp bending occurs at every DNA helical repeat (~10 bp), when the major groove of the DNA faces inward toward the histone octamer, and again ~5 bp away, with opposite direction, when the major groove faces outward. The property of the ~10-bp periodicity is called “a genome code” for nucleosomes. Since that, many nucleosome prediction models were developed.

2. Nucleosome positioning and its determinant

2.1. Concepts of nucleosome positioning and nucleosome occupancy

The term “nucleosome positioning” is used to indicate where nucleosomes are located with respect to the genomic DNA sequence [16]. Generally, nucleosome positioning can divide into two parts: rotational positioning and translational positioning. The first one is to describe the side of the DNA helix that faces the histones and the next one is to determine the nucleosome midpoint with regard to the DNA sequence.

By doing statistical analysis, “nucleosome occupancy” tries to identify the possibilities of a base pair whether it is in a nucleosome region [16]. It is possible to calculate average nucleosome positioning levels on a given region of DNA in a population of cells. In ideal conditions, nucleosome is “shaking” in the perfect position. By counting the time of sequenced reads that are overlapped by nucleosome center in a ~147 bp window, it gives the most conservative locus which means that it is most possible to have a nucleosome there (Figure 4).

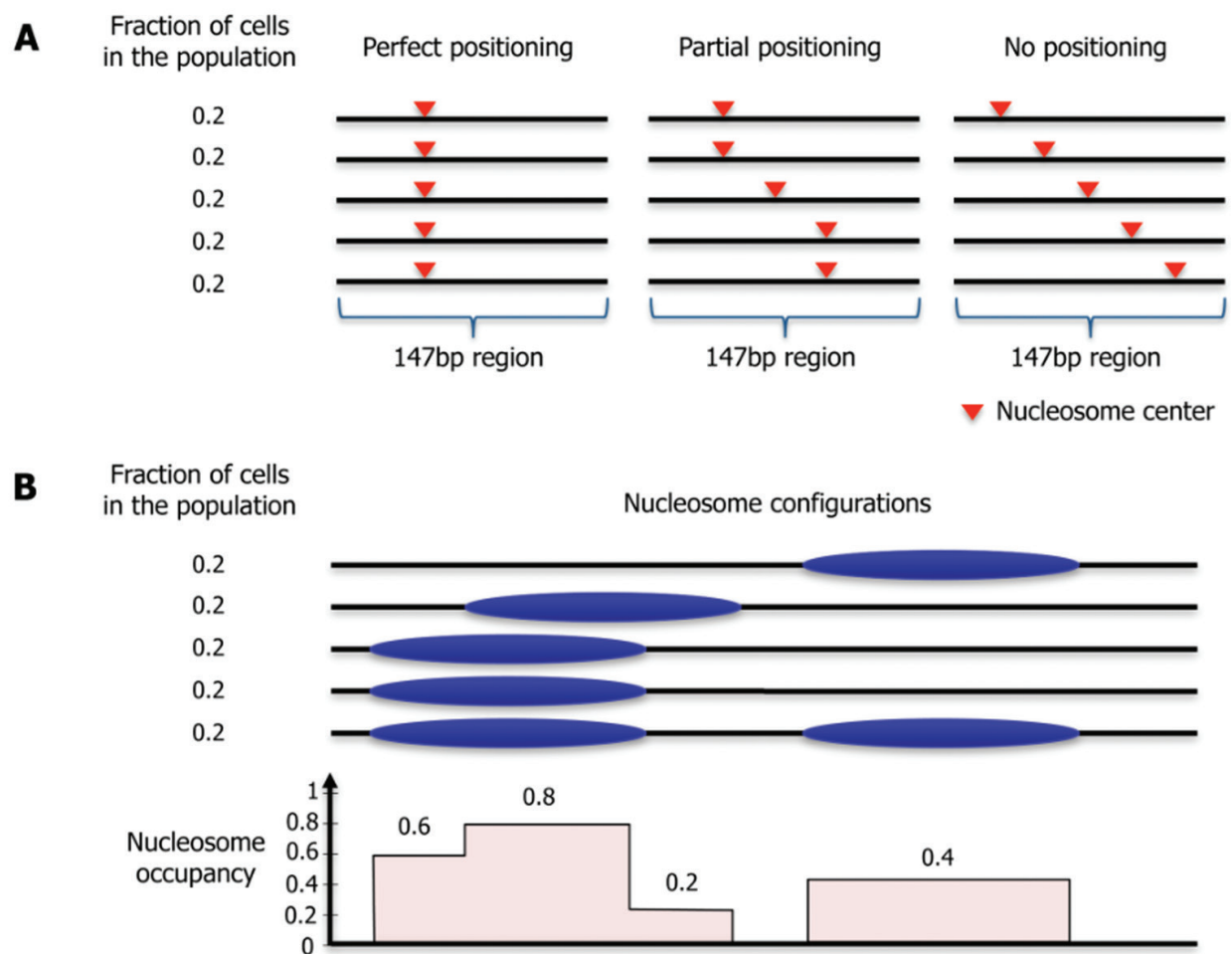


Figure 4. Illustration of the concepts of nucleosome positioning and nucleosome occupancy. (A) We use fraction of cells from the population in which that basepair is in the middle of a 157bp nucleosome representing the nucleosome positioning along every basepair in the genome. The left figure exhibit perfect-positioning region, where the nucleosome center is located at the same basepair all over population cells; the other two showed partial-positioning and no-positioning region. (B) [16].

2.2. The association between nucleosome positioning and 10–11 bp periodicities in DNA sequence in yeast

Early in 1990, the 10–11 bp periodicities were reported [17]. In addition to 3-bp periodicity, which is due to the fact that three consecutive bases encode one type of amino acids, the genomic DNA exhibits 10–11 bp periodicities. The 10–11 bp periodicities in complete genomes reflect protein structure and DNA folding [17].

In alpha helices structure, the hydrophobic amino acids (aa) occur with a ~ 3.5 aa, and all five hydrophobic amino acids L, I, V, F, and M have a base T (thymine) at middle position of their codons. This leads $\sim 3.5 \times 3 = \sim 10.5$ bp periodicity in protein coding DNA sequences, called protein-induced periodicity.

On the other hand, the 10–11 bp periodicities have an intimate association with nucleosome positioning. To sharply bent and tightly wrapped around a histone protein octamer, DNA sequence has intrinsic bias. The position of certain dinucleotides, such as AA, TA, and TT in minor grooves facing toward (every 10 bp) and GC in minor grooves facing away from the histone octamer favors these (Figure 5) distortions [15]. Moreover, when digesting DNA using DNase I (Deoxyribonuclease I), it was observed that the cleavage pattern in nucleosome position shows a ~ 10.3 bp period, which is equal to a minor groove. For the naked DNA, which is entirely devoid of nucleosomes, the oscillatory pattern in cleavage profile was disappeared in digesting [18]. All of these strongly suggested the role of the 10–11 bp periodicities of the specific dinucleotides in positioning nucleosomes. Based on the features of DNA sequences, many models are developed to predict nucleosomes (Table 2).

We further found that within frequency domains, weakly bound dinucleotides (AA, AT, and the combinations AA-TT-TA and AA-TT-TA-AT) present doublet peaks in a periodicity range

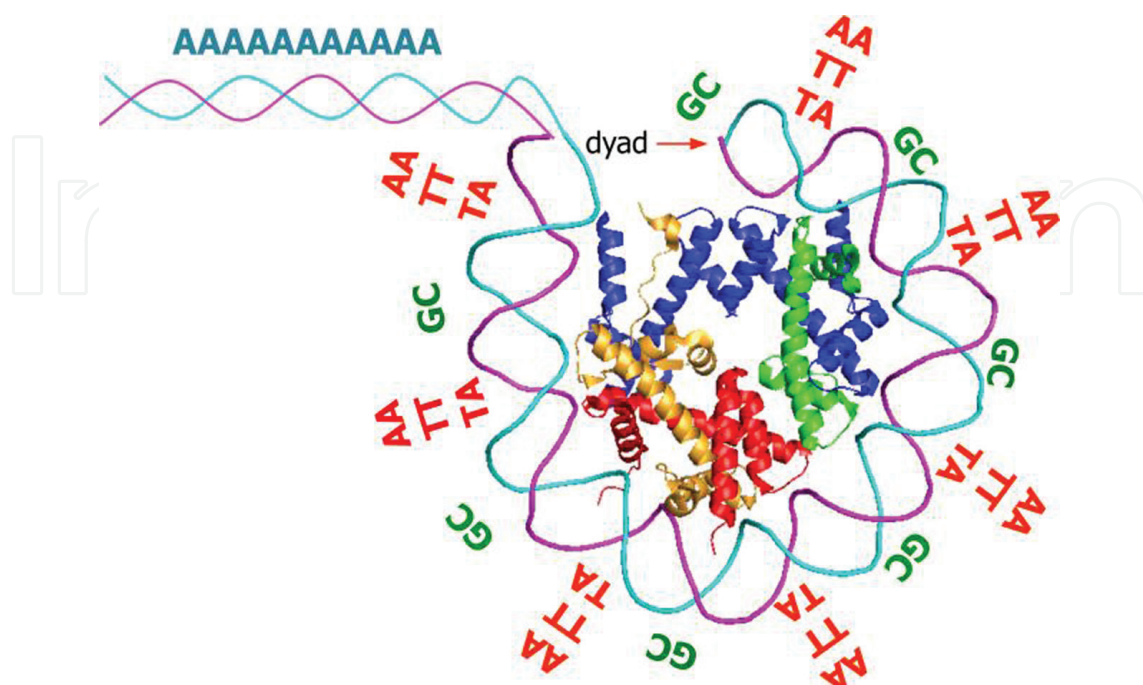


Figure 5. Illustration of nucleosome sequence preferences [16].

Software	Description	References
Segal Lab: Online Nucleosomes Prediction	This tool allows you to submit a genomic sequence and to receive a prediction of the nucleosomes positions on it, based on the nucleosome-DNA interaction model	[15, 21, 22]
iNuc-PseKNC	A predictor for predicting nucleosome positioning in <i>Homo sapiens</i> , <i>Caenorhabditis elegans</i> , and <i>Drosophila melanogaster</i> genomes	[23]
NuCMap	NuCMap is based on chemical modification of engineered histones. The tool reveals novel aspects of the in vivo nucleosome organization that are linked to transcription factor (TF) binding, RNA polymerase pausing, and the higher order structure of the chromatin fiber	[24]
NuPoP	Predicts nucleosome position by explicitly modeling the linker DNA length. NuPoP is based on a duration hidden Markov model (HMM)	[25]
Epidaurus	Epidaurus is a bioinformatics tool used to effectively reveal inter-dataset relevance and differences through data aggregation, integration, and visualization	[26]
Multi-Layer Model	Analyses nucleosome position data obtained with microarray-based approach. MLM is a classifier to distinguish between several kinds of patterns	[27]
NucEnerGen	Predicts nucleosome energetics by using high throughput sequencing. It establishes that nucleosome occupancies can be explained by systematic differences in mono- and dinucleotide content between nucleosomal and linker DNA sequences	[28]
nuMap	Implements the YR and W/S schemes to predict nucleosome positioning at high resolution. This methodology is based on the sequence-dependent anisotropic bending	[29, 30]
NPRD	Compiles the available experimental data on locations and characteristics of nucleosome formation sites (NFSs). The object of the database is a single NFS described in an individual entry	[31]
AWNFR	An algorithm based on down-sampling operation and footprint in wavelet	[32]
ICM	Allows users to assess nucleosome stability and fold sequences of DNA into putative chromatin templates. It uses an elastic model to place nucleosomes	[33]
SymCurv	The tool is able to capture sequence constraints, which are related to structure in genomic regions	
FineStr	Allows users to upload genomic sequences in FASTA format and to perform a single-base-resolution nucleosome mapping on them	[34]
iNuc-PhysChem	Identifies nucleosomal sequences by incorporating physicochemical properties into a 1788-dimensional feature vector. iNuc-PhysChem was able to identify nucleosome positioning for an independent DNA segment extracted from the <i>Saccharomyces cerevisiae</i> genome	[35]
TemplateFilter	High-resolution nucleosome mapping reveals transcription-dependent promoter packaging	[36]
DANPOS	A comprehensive bioinformatics pipeline explicitly designed for dynamic nucleosome analysis at single-nucleotide resolution. DANPOS is also robust in defining functional dynamic nucleosomes	[37]
BINOCh	A package that allows biologists to carry out an analysis of nucleosome occupancy data to discover stimulus-induced transcription factor binding	[38]

Software	Description	References
PING	A package for nucleosome positioning using MNase-seq data or MNase- or sonicated ChIP-seq data. PING uses a model-based approach, which enables nucleosome predictions even in the presence of low read counts	[39]
ChIPseqR	A package based on an algorithm for the analysis of nucleosome positioning and histone modification ChIP-seq experiments	[40]
NUCwave	A bioinformatic tool that generates nucleosome occupation maps from chromatin digestion with micrococcal nuclease (MNase-seq), chemical cleavage (CC-seq), chromatin immunoprecipitation (ChIP-seq) and fragmentation by sonication	[41]
NucPosSimulator	A simulation tool to identify positions of nucleosomes from next generation sequencing data	[42]
NucHunter	Inferring nucleosome positions with their histone mark annotation from ChIP data	[43]
DiNuP	A systematic approach to identify regions of differential nucleosome positioning	[44]
NucTools	Allows calculations of nucleosome occupancy profiles averaged over several replicates, comparisons of nucleosome occupancy landscapes between different experimental conditions, and the estimation of the changes of integral chromatin properties. NucTools facilitates the annotation of nucleosome occupancy with other chromatin features like binding of transcription factors (TF) or architectural proteins, and epigenetic marks like histone modifications or DNA methylation	[45]
Dimnp	Identifies differential nucleosome regions (DNRs) in multiple samples. Dimnp is able to identify all the DNRs that are identified by two-sample method Danpos. It shows a good capacity (area under the curve >0.87) compared with the manually identified DNRs	[46]
ArchAlign	ArchAlign identifies shared chromatin structural patterns from high-resolution chromatin structural datasets derived from next-generation sequencing or tiled microarray approaches for user defined regions of interest	[47]
SANEFALCON	A tool developed to calculate the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles, based on single end sequencing of cell-free DNA	[48]
NucDe	An R package mapping nucleosome-linker boundaries from both MNase-Chip and MNase-seq data using a non-homogeneous hidden-state model based on first-order differences of experimental data along genomic coordinates	
Nu-OSCAR	A program that can be used to identify binding sites of known transcription factors	
NSeq	A multithreaded Java application for finding positioned nucleosomes from sequencing data	[49]
ArchTEx	The extension of mapped sequence tags is a common step in the analysis of single-end next-generation sequencing (NGS) data from protein localization and chromatin studies. ArchTEx identifies the optimal extension of sequence tags based on the maximum correlation between forward and reverse tags and extracts and visualizes sites of interest using the predicted extension	[50]

Software	Description	References
PuFFIN	Builds genome-wide nucleosome maps specifically designed to take advantage of paired-end reads. This method can accurately determine a genome-wide set of nonoverlapping nucleosomes without any user-defined parameters	[51]
NPS	A python software package that can identify nucleosome positions given histone-modification ChIP-seq or nucleosome sequencing at the nucleosome level	[52]

Table 2. Nucleosome prediction models and nucleosome sequenced data-processed models.

of 10–11 bp, and strongly bound dinucleotides present a single peak [19]. A time-frequency analysis, based on wavelet transformation, indicated that weakly bound dinucleotides of nucleosomal DNA sequences were spaced smaller (~10.3 bp) at the two ends, with larger (~11.1 bp) spacing in the middle section. The finding was supported by DNA curvature and was prevalent in all core DNA sequences.

We assessed the roles of the 10–11 bp periodicities for different kinds of dinucleotides [20]. Near the transcription start site, the signals reveal a similar feature that the nucleosome organization exhibits (**Figure 6**). But, it seems that the species do not share the same dinucleotides patterns. Furthermore, the dinucleotides patterns are dominant at the specific region of genome, indicating their diverse roles in forming and organizing nucleosomes.

2.3. Nucleosome prediction models for yeast

In **Table 2**, the models for both nucleosome prediction and nucleosome sequencing data processing are listed.

2.4. The chromatin remodeling complex and its roles in altering nucleosomes

Chromatin remodeling complex helps cell to establish the access of genomic DNA for transcription factors. The complexes have two major groups, namely covalent histone-modifying complexes and ATP-dependent chromatin remodeling complexes [53]. They work in a different way.

ATP-dependent chromatin-remodeling enzymes are helicase which use ATP’s energy to reposition (slide, twist or loop) nucleosomes along the DNA, expel histones away from DNA or facilitate exchange of histone variants, and thus creating nucleosome-free regions of DNA for gene activation [54]. All known ATP-dependent chromatin complex can be organized into SWI/SNF, ISWI, CHD, and INO80 families. Each family of ATPase has distinct remodeling activities, including incremental nucleosome sliding on DNA in *cis*; the creation of DNA loops on the surface of the nucleosome; eviction of histone H2A/H2B dimers; eviction of the histone octamer; or the exchange of histone octamer subunits within the nucleosome to change its composition [55].

Covalent histone-modifying complexes modify the histone including acetylation, methylation, and phosphorylation which can change the interaction between histone and DNA; for

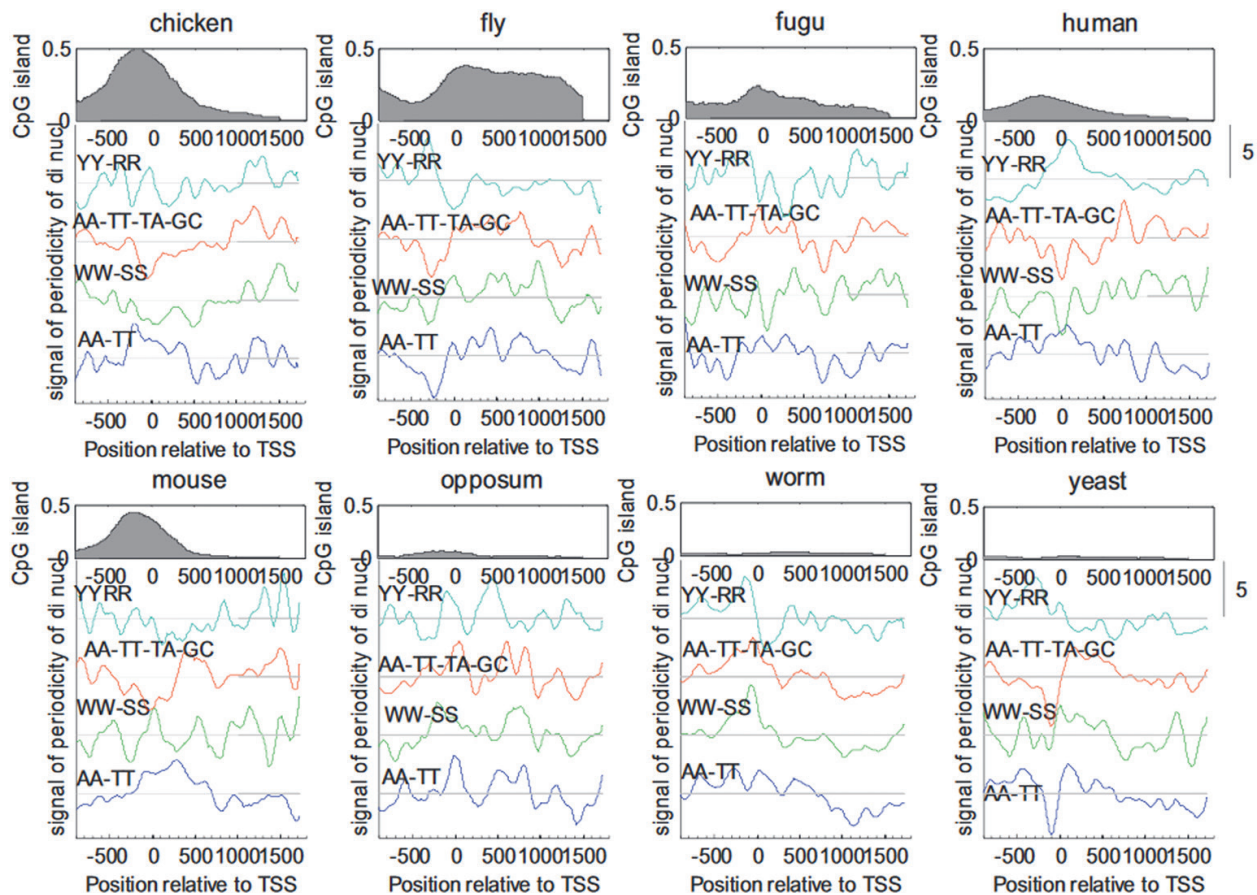


Figure 6. The 10–11 bp periodicities signals of the dinucleotides patterns around TSSs of eight species (human, mouse, chicken, worm, fly, fugu, lancelet, and yeast) [20].

example, methylation of specific lysine residues in H3 and H4 causes further condensation of DNA around histones, making it hard to bind transcription factor or other proteins.

2.5. The statistical model for nucleosomes distribution

A typical nucleosome distribution around TSS is shown in **Figure 7** [56]. Nucleosomes are depleted around TSSs, resulting in a nucleosome-free region (NFR) that is flanked by two well-positioned nucleosomes whereas the nucleosomes downstream of the TSS are equally spaced in a nucleosome array. Of all nucleosomes around the gene, the +1 nucleosome often contains histone variants (H2A.Z and H3.3) and modification by acetyltransferases and methyltransferases. These may help to the nucleosome eviction when transcription is needed. The +2 nucleosome follows the +1 nucleosome immediately and shares the some properties but contains less H2A.Z and less methylation and acetylation. In a barrier model for nucleosome organization, the nucleosome distribution is largely a consequence of statistical packing principles. The genomic sequence specifies the location of the –1 and +1 nucleosomes. The +1 nucleosome forms a barrier against which nucleosomes are packed, resulting in uniform positioning, which decays at farther distances from the barrier [57].

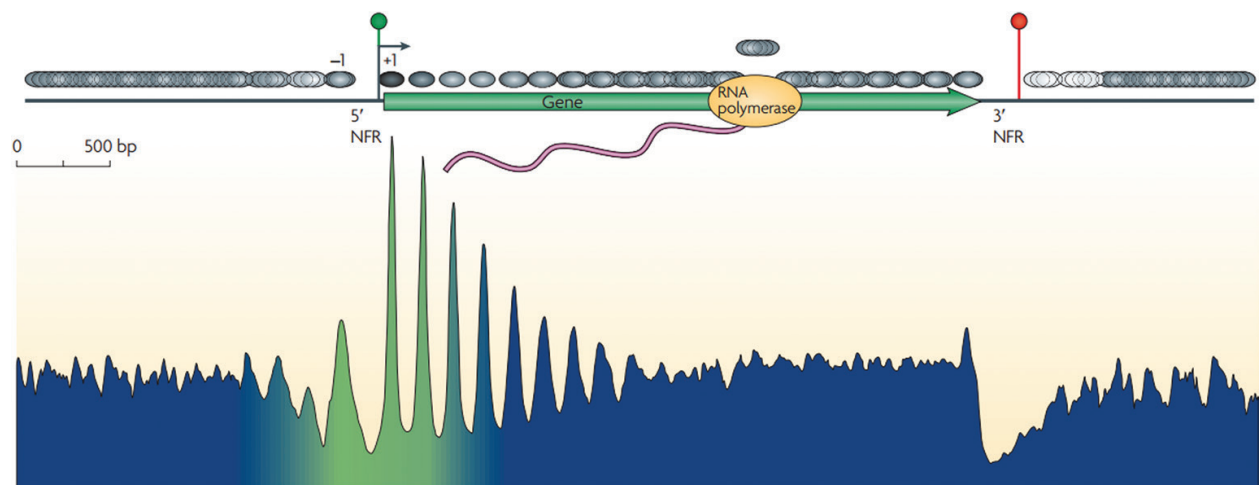


Figure 7. The consensus distribution of nucleosomes (gray ovals) around all yeast genes is shown, aligned by the beginning and end of every gene. The resulting two plots were fused in the genic region. The peaks and valleys represent similar positioning relative to the transcription start site (TSS). The arrow under the green circle near the 5' nucleosome-free region (NFR) represents the TSS. The green-blue shading in the plot represents the transitions observed in nucleosome composition and phasing (green represents high H2A.Z levels, acetylation, H3K4 methylation and phasing, whereas blue represents low levels of these modifications). The red circle indicates transcriptional termination within the 3' NFR. Figure is reproduced from REF(2008) Cold Spring Harbor Laboratory Press [56].

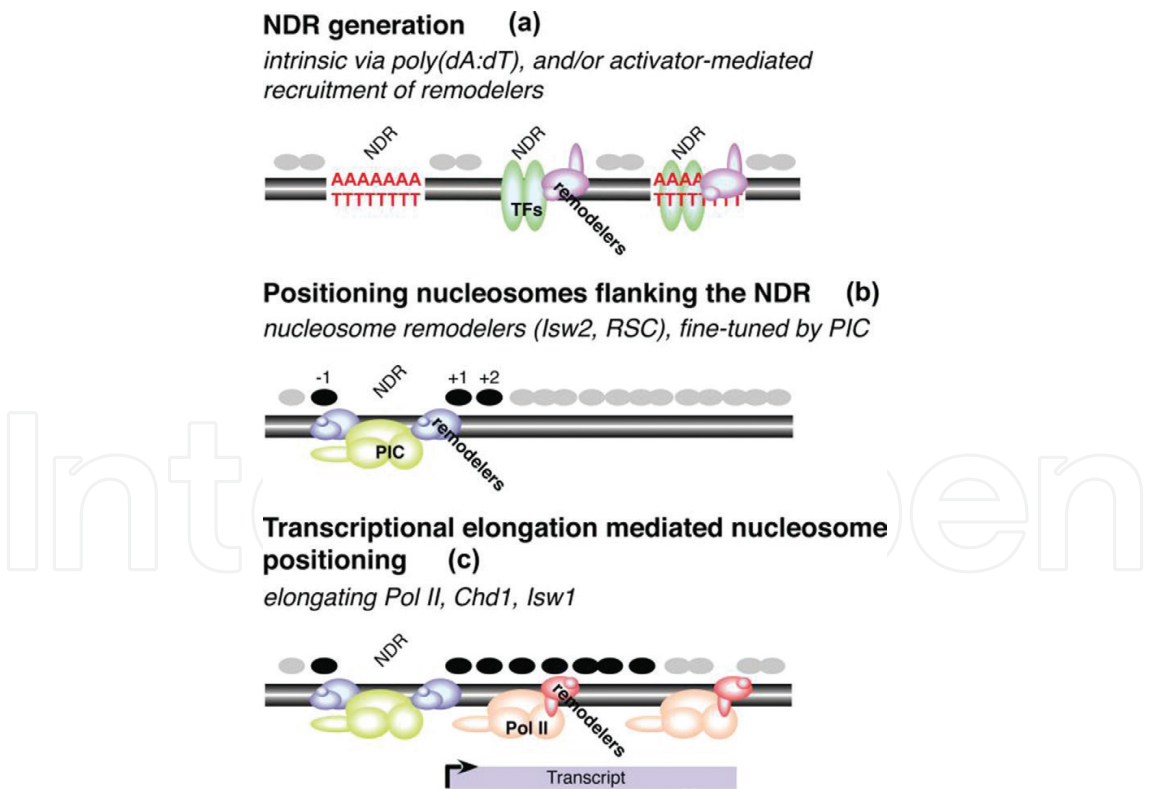


Figure 8. Determinants of nucleosome positioning. (a) Nucleosome-depleted regions (NDRs) are generated either by poly (dA:dT) tracts and/or by transcription factors and their recruited nucleosome remodeling complexes. Gray circles indicate nucleosomes. (b) Nucleosomes located at highly preferred positions (black circles) flanking the NDR are generated by nucleosome-remodeling complexes (for example, Isw2 and RSC, likely in a transcription-independent manner), and fine-tuned by the Pol II preinitiation complex (PIC) and associated factors. (c) Positioning of the more downstream nucleosomes depends on transcriptional elongation, and the recruitment of nucleosome-remodeling activities (for example, Chd1 and Isw1) and histone chaperones by the elongating Pol II machinery [16].

2.6. The nucleosome determinant [16]

A variety of factors determine the location of nucleosomes including DNA sequence, nucleosome remodelers, transcription factors (TFs), and elongating Pol II (**Figure 8**). Each of these components has different contribution in nucleosome positioning. Interestingly, these components can affect each other thus resulting in different positioning pattern in a more complex way. The DNA sequence is critical for rotational positioning along the DNA helix, and it is also an important determinant for nucleosome occupancy. In particular, poly (dA:dT) and poly (dG:dC) tracts are intrinsically inhibitory to nucleosome formation, whereas non-homopolymeric GC-rich regions favor nucleosome formation.

3. The experiment methods of determining nucleosome occupancy and the bioinformatics analysis for the data

3.1. The techniques of determining nucleosomes positions

3.1.1. *Mnase-seq*

Micrococcal nuclease (MNase), one kind of glycolprotein of *Staphylococcus aureus*, has capacity of digesting the naked DNA. MNase, firstly, induces single-strand breaks, and then cleaves the complementary strand near the first break [58, 59]. Nucleosomal DNA is protected by wrapping on histone octamer in digesting with MNase, thus being remained as DNA fragments after the digestion. Taking this advantage, a high throughput sequencing technique MNase-seq is developed to probe nucleosome positions in a genome-wide manner. MNase cleavage favors AT-rich region in limiting enzyme concentrations.

3.1.2. *Dnase-seq*

DNase I, one kind of endonuclease, can cut the chromatin-accessible DNA, namely DNase I hypersensitive sites (DHSs), and thus is used in mapping opening chromatin regions (**Figure 9**) [60]. The opening chromatin region is mainly the regulatory sites in gene transcription. Thus, the opening region may alter in different cells types. This can be reflected in DHSs. The change of DHSs often associates one or more nucleosomes loss or formation [60].

DNase-seq means the DNase I digestion followed by DNA sequencing [60]. DNase-seq has been widely used in probing cell-specific chromatin accessibility. The rotational localization of individual nucleosomes is based on the inherent preference of DNA enzyme I cleavage of DNA at about 10 bp per nucleosome [61]. By coupling bioinformatics analysis, DNase-seq can be used in studying TF occupancy at nucleotide resolution in a qualitative and quantitative manner [62]. In DNase-seq, many cells and many sample preparations and enzyme titration steps are required [63].

3.1.3. *ATAC-seq*

ATAC-seq is an assay for transposase-accessible chromatin with high throughput sequencing [64]. The technique is based on Tn5 transposase's "cutting and pasting" function to probe

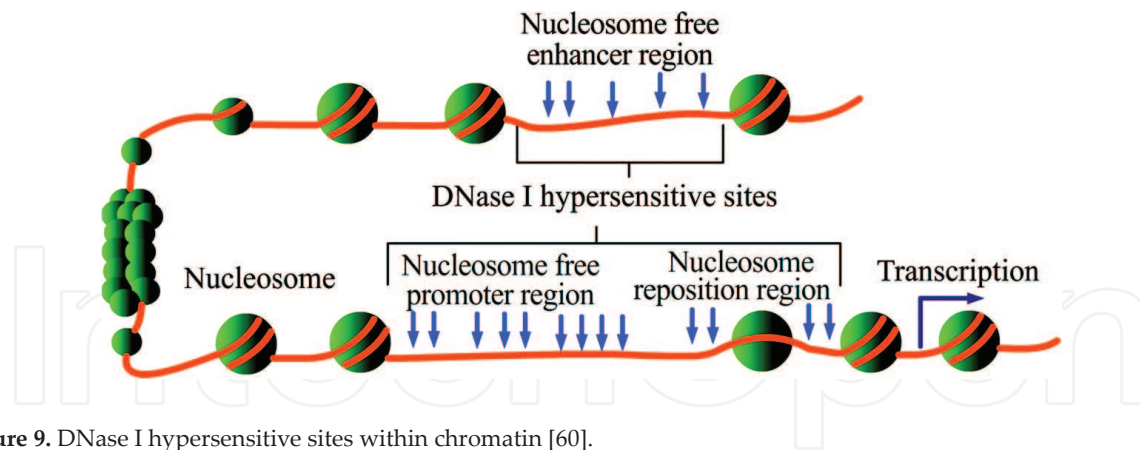


Figure 9. DNase I hypersensitive sites within chromatin [60].

the active regulatory regions [65]. ATAC-seq only needs a small number of cells, ~500–50,000 unfixed nuclei. Moreover, its procedure only involves two steps. Therefore, it is able to study multiple aspects of chromatin architecture simultaneously at high resolution, including nucleosomes, chromatin accessibility [64].

3.1.4. ChIP-seq

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) sequences the interest DNA fragments that are separated and collected from the immunoprecipitation [66]. The main area of ChIP-seq is in precisely mapping for transcription factor-binding sites (TFBSs). **Figure 10** shows a general procedure of a ChIP experiment [66]. This procedure includes the DNA-protein crosslinking with formaldehyde, sonication, immunoprecipitation, reversed crosslinking, and sequencing [66]. Using antibody of the histones, such as histone H3, ChIP-seq is immediately able to determine nucleosome positions.

3.1.5. Other techniques

In addition to the techniques mentioned above, there are other techniques often used, such as Formaldehyde-assisted isolation of regulatory elements (FAIRE-seq) and ChIP-exo. FAIRE-seq is based on the differences in crosslinking efficiencies between DNA and nucleosomes or sequence-specific DNA-binding proteins. Sequencing provides information for regions of DNA that are not occupied by histones [67]. ChIP-exo employs the use of exonucleases to degrade strands of the protein-bound DNA in the 5′–3′ direction to within a small number of nucleotides of the protein binding site [68]. The nucleotides of the exonuclease-treated ends are determined using DNA sequencing.

3.2. Procedures of dealing with the nucleosome DNA sequenced dataset

At the present, nucleosome sequencing dataset are mainly from MNase-seq. In some studies, dataset from ATAC-seq, DNase-seq, and ChIP-seq are used to infer nucleosome positions. A general analysis workflow includes data quality control, mapping, making nucleosome profile, determining nucleosome position, comparing between cell types, and associating with other omics-data (expression data) to find biological meanings.

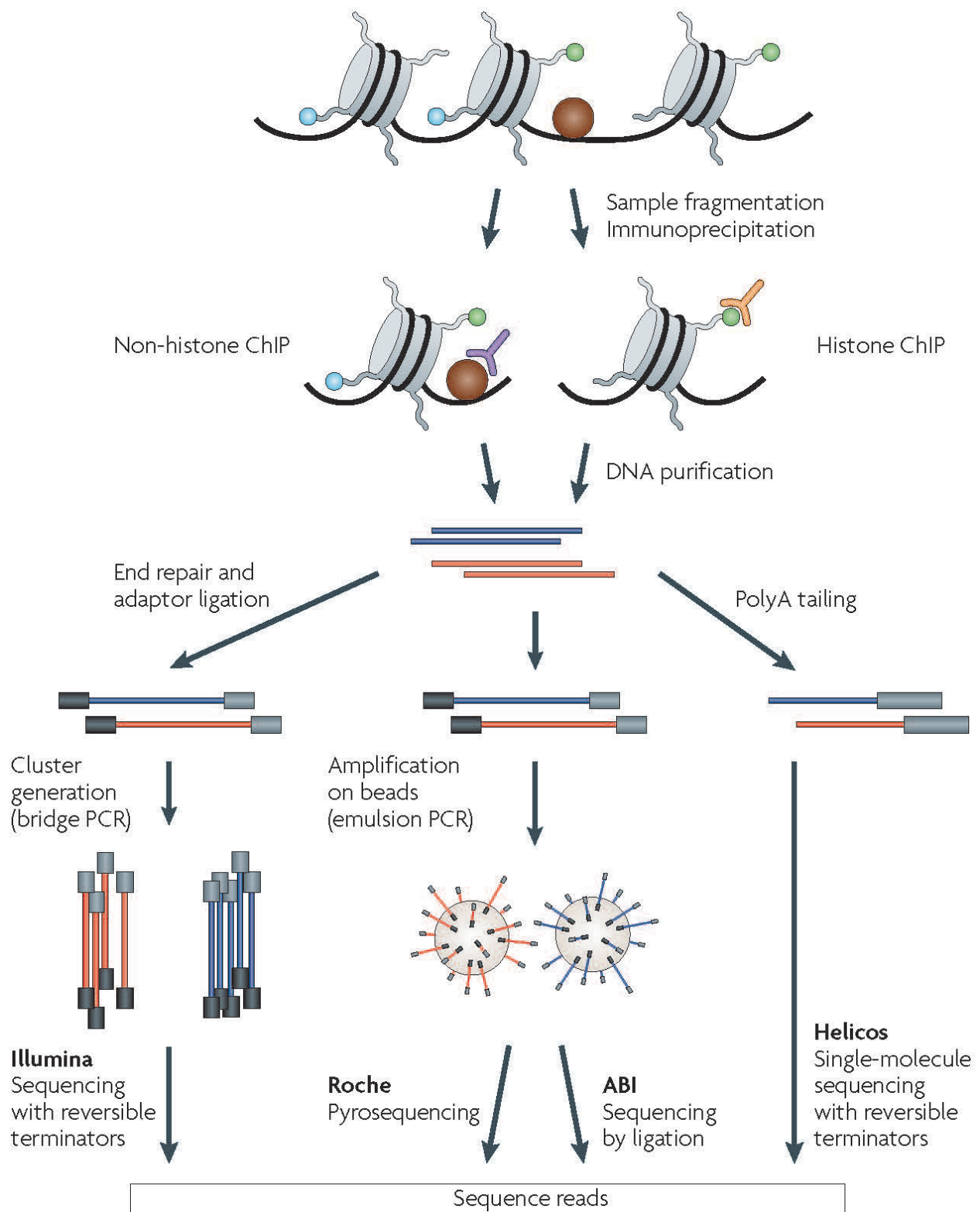


Figure 10. Overview of a ChIP-seq experiment [66] (<https://www.nature.com/nrg/journal/v10/n10/full/nrg2641.html>).

3.2.1. Data management and genome alignment

Sequencing quality control (QC) is to check the reads quality (fraction of mapped reads) and depth of coverage. Tools BWA and Bowtie are widely used in reads alignments. During the

alignment process, multiple-mapping reads and duplication reads are often filtered so as to remove overrepresented regions of the genome due to technical bias [60]. Reads filtering can be performed with SAMtools or Picard tools.

3.2.2. Data visualization

Data visualization helps to observe the reads distribution at specific locus. The Integrative Genomics Viewer (IGV) [69], which is developed by the University of California Santa Cruz (UCSC), is one of the most powerful tools to visualize. In IGV, the multiple types of annotation data are integrated, including gene information, epigenetic and expression data, single-nucleotide polymorphisms (SNPs), repeat elements and functional information from the ENCODE, and other research projects. IGV accepts many types of data formation including BED, BedGraph, GFF, WIG, and BAM files, which allow to compare with publicly data.

3.2.3. Identification of enriched regions

With respect to nucleosomes sequencing data, there are two basic tasks in analysis. One is to calculate the nucleosome profile (reads coverage) both along the genomic coordinate and near the regulatory sites (for instance the TSSs). This helps to directly check the quality of MNase digestion and DNA sequencing. The other task is to infer the precise nucleosomes positions (dyad position) using the nucleosome profile so as to identify the nucleosome alteration among different cell types.

3.2.3.1. MNase-seq data

For single-end MNase-seq data, one method to make nucleosome profile is as follows [70]. First, the length of each read was extended 73 bp in the 3' direction, and the Watson-strand reads and Crick-strand reads were oppositely shifted 73 bp. The absolute nucleosome occupancy value of each genomic site was expressed as the number of reads covering the genomic sites. Second, nucleosome occupancy was scaled by dividing the occupancy value by the average nucleosome occupancy of the whole genome; i.e., the nucleosome occupancy was expressed as the fold change of the absolute occupancy relative to the average occupancy. Reads can also be shifted 73 bp toward the 3' direction, which represented the midpoint [60].

With paired-end sequencing, it is assumed that the nucleosome midpoint is consistent with the midpoint of the forward and reverse reads. Unless the reads are from the on type cell (single cell), nucleosome positions actually represent the average positions in cell population. Therefore, the overlapping reads have to be clustered over genomic regions [60].

Calling nucleosomes actually is to find the peak positions along the nucleosome profile. DANPOS is one tool that can identify nucleosome positions [37]. Also, it allows us to detect three categories of nucleosome dynamics, such as position shift, fuzziness change, and occupancy change, using a uniform statistical framework using MNase-seq datasets. Other tools can be found in **Table 2**.

3.2.3.2. DNase-seq data, ATAC-seq, and ChIP-seq

From the DNase/ATAC/ChIP-seq datasets, nucleosome position cannot be directly inferred, but they provide information about the opening chromatin and protein-binding regions,

which associate nucleosome depletion. Therefore, for these datasets, peaking calling is one central task. MACS identifies genome-wide locations of transcription/chromatin factor binding or histone modification, including removing redundant reads, adjusting read position, calculating peak enrichment, and estimating the empirical false discovery rate (FDR) (<http://liulab.dfci.harvard.edu/MACS/index.html>) [71]. Based on the position-adjusted reads, MACS slides a window of size $2d$ across the genome to identify regions that are significantly enriched relative to the genome background. The P-value is derived from Poisson distribution. When a control sample is available, MACS can also estimate an empirical false discovery rate (FDR) for every peak by exchanging the ChIP-seq and control samples and identifying peaks in the control sample using the same set of parameters used for the ChIP-seq sample.

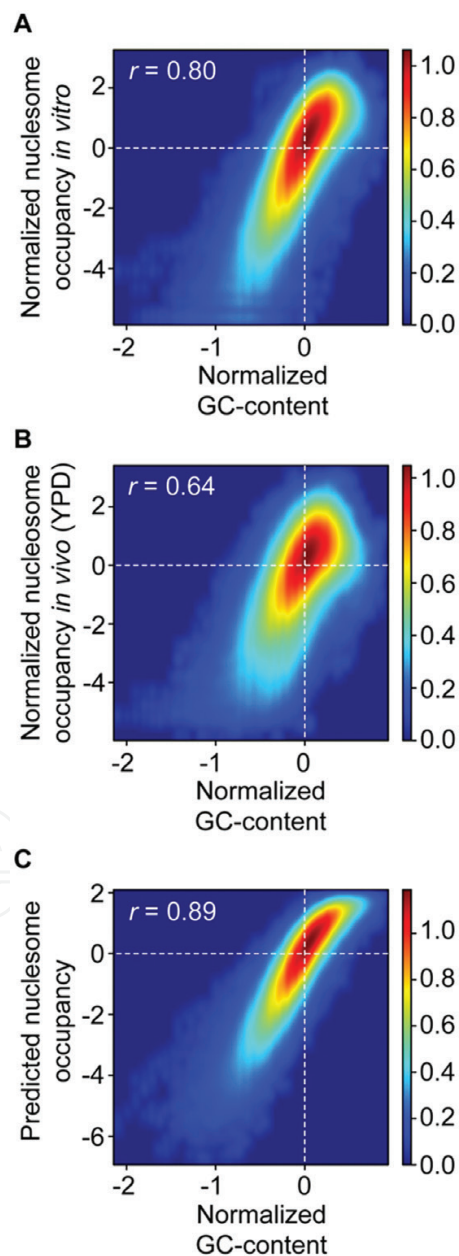


Figure 11. The GC content in 147 base pair windows is strongly correlated to nucleosome occupancy. Density plot comparison between the normalized centered GC content in 147 base pair windows (x axis) and (A) the *in vitro* reconstituted, (B) the *in vivo* (YPD). GC-content is normalized as the \log_2 GC-content within 147 base pairs [75].

3.3. GC-content and cutting bias

GC content bias means the variability between the GC content in a region and the count of fragments/reads mapped to it. The bias can dominate the signal of interest for analyses and leads to false positive. More seriously, the bias tends to be different among the samples; thus, there is no general method to remove it [72]. Two facts associate the variability. One is GC content which is heterogeneous among the genome. In yeast, the open reading frames (ORFs) with similar GC contents at silent codon positions are significantly clustered on chromosomes [73]. Moreover, GC content varies along the genome and is often correlated with functionality. The other is MNase that has a cutting bias. Kinetic analysis indicates that the rate of cleavage is 30 times greater at the 5' side of A or T than at G or C [74].

Most current correction methods follow a common path. Both fragment counts and GC counts are binned to a bin-size of choice [72]. Then, the conditional mean fragment count per GC value is modeled by assuming smoothness. At last, a predicted count is estimated for each bin based on the bin's GC. The predictions represent one normalization for the original signal [72].

Another aspect is that GC-content is predictive for nucleosome position both in vivo and in vitro (**Figure 11**) [75]. That is, nucleosomal DNA sequences tend to be enriched in GC base pairs.

4. The transcription regulation and nucleosome positioning

4.1. The +1 and -1 nucleosomes and nucleosome-free regions (NFRs) near transcription start sites (TSSs)

Nucleosome positioning is in gene regulation since the DNA packing on the surface of the histone octamer can occlude the binding sites of transcription factors (TFs) on genomic DNA. That is to say, the nucleosome positioning at promoters negatively regulates gene transcription by preventing TFs binding. Typically, nucleosomes are depleted around transcription start sites (TSSs), resulting a nucleosome-free region (NFR) that is flanked by two well-positioned nucleosomes (+1 and -1 nucleosomes). In downstream of the TSS, nucleosomes are equally spaced as a nucleosome array. At 3' direction gene (transcription termination sites (TTS)), there is also a NFR, called 3' NFR. Additionally, the poly (dA:dT) sequences are found in the 5' and 3' NFRs, where they act as nucleosome-excluding sequence. The characteristics of nucleosome organization are found in multiple species, including yeast, worms, flies and humans. In such an organization, the NFRs are often the TFs binding regions. Transcriptional activation involves several steps in yeast [56]. Firstly, special chemical modifications (acetylation and methylation (H3K4me3)) occur on histones of the -1 and +1 nucleosomes (**Figure 12**). The acetylation marks can be recognized by bromodomain modules of the SAGA histone acetyltransferase complex and Bdf1. SAGA and TFIID then deliver TBP to promoters. Then, the pre-initiation complex (PIC) is mounted.

It is suggested that NFRs at promoters result from a competition between TF and nucleosome binding because that incorporating competition with TFs improves the prediction performance for nucleosome positioning, particularly in promoter regions [76]. Moreover, the mechanism is not restricted to a few promoters, but is the typical configuration along the

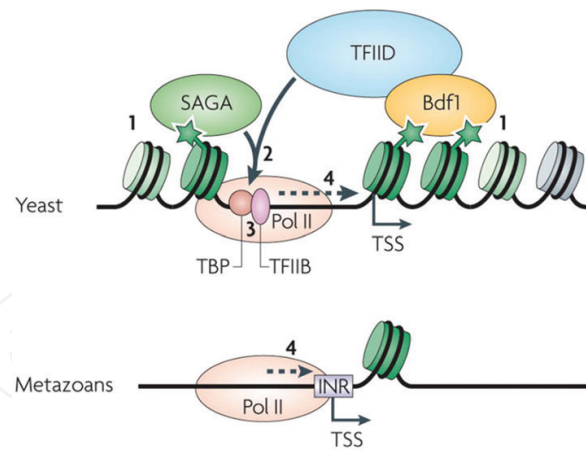


Figure 12. Transcription initiation in budding yeast, including four steps [56].

genomes. Interestingly, it was reported that of the 158 yeast TFs, only 10–20 significantly contribute to inducing NFRs, and these TFs are highly enriched for having direct interactions with chromatin remodelers [76].

Therefore, theoretically, nucleosome level at promoters should negatively associate gene expression level. As expected, for the acid phosphatase inducible PHO5 gene, a significant cell-to-cell variation was found in nucleosome positions and the nucleosome shift correlates with changes of gene expression (**Figure 13**) [77]. However, nucleosome positioning is not absolute, and even with major shifts in gene expression, some cells fail to change nucleosome configuration. We found in human CD4⁺ T cells, a wider NFR at promoters of housekeeping genes and highly expressed genes [78].

4.2. The difference of nucleosome organization among species

The current studies suggest that almost all eukaryotic organisms hold the nucleosome organization characteristics at the 5' end of gene, namely a NFR flanked by two (+1 and -1) well-positioned nucleosomes and followed by an array of nucleosomes downstream of TSSs [37, 70, 79]. But compared with multicellular organisms fly, worm, and human, yeast is very simple. Nucleosome organization exhibits some differences. First, averagely, yeast has a short linker DNA. The linker DNA is 18 bp in *S. cerevisiae*, ~28 bp in *Drosophila melanogaster* and *Caenorhabditis elegans*, and ~38 bp in human [56]. Second, the dyad position of the +1 nucleosome relative to TSS appears to vary in different organisms. In yeast, the dyad of the nucleosome is at ~50–60 bp downstream of the TSS [80]. However, in *Drosophila*, the dyad is found at 135 bp downstream of the TSS, reflecting the differences in transcriptional regulatory mechanisms. Third, the 10–11 bp periodicities of the specific dinucleotides (such as AA/TA/TT-GC, WW-SS) pronounce stronger in yeast than in other multicellular organism. In other words, from single-cell organism to multicellular organism, genomic DNA needs to bear more of “encoding information” to meet a more complex regulation requirement. In genomic DNA of multicellular organism, more of TF binding sites are embedded, which will disturb the coding for other information, such as the coding for nucleosome positioning. Forth, in multicellular organism, there exists exons and introns, thus having a splicing process in transcription. It was

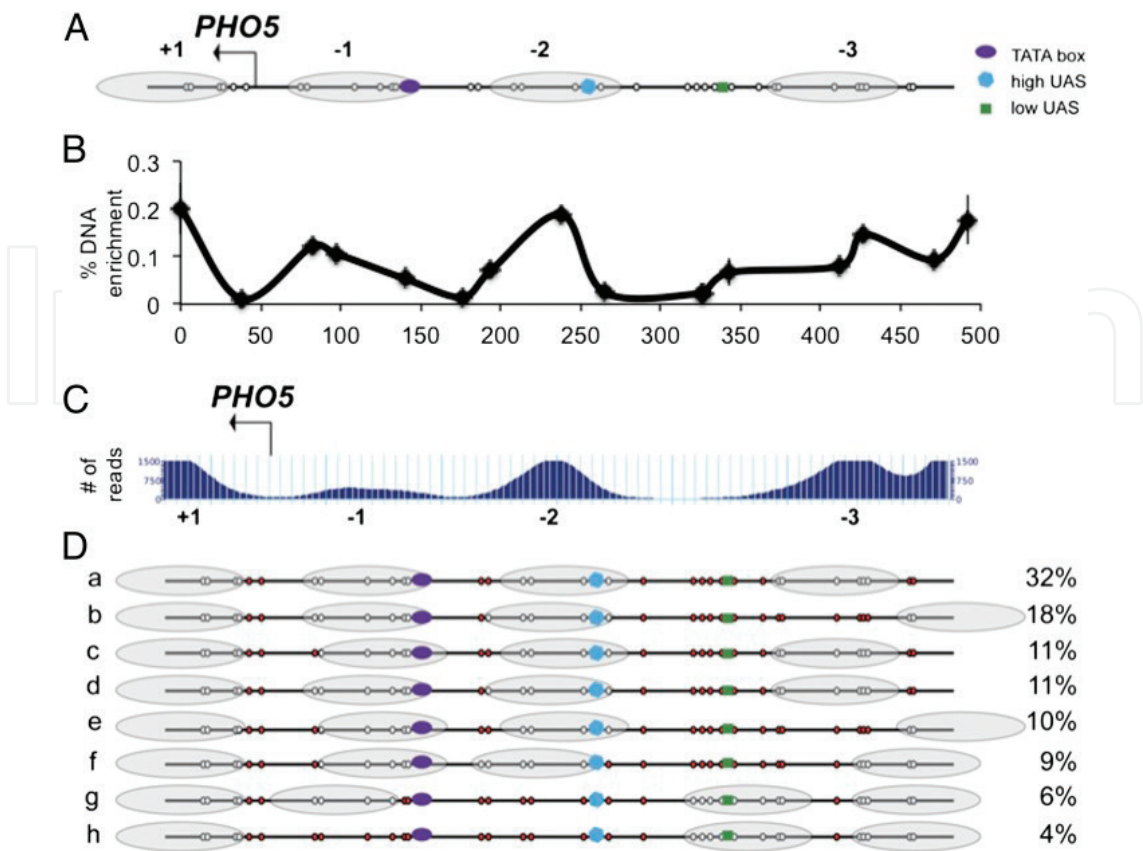


Figure 13. The heterogeneity of nucleosome architecture at the *PHO5* promoter. (A) Canonical position of nucleosomes (gray ovals) in the *PHO5* promoter. White circles indicate the locations of cytosines of GC dinucleotides. (B) Mapping of mononucleosomal DNA of cells grown in rich media using a nucleosome-scanning assay. Enrichment of mononucleosomal DNA (y axis) is indicated by the midpoints of each amplicon (x axis), and error bars represent 1 SD from two independent biological replicates. (C) MNase-seq track of the *PHO5* promoter from cells grown in rich media. (D) Nucleosome architecture of 806 cells from three bulk populations revealed eight conformations (a–h). Nucleosomes are depicted as gray ovals. Red circles indicate methylated cytosines, and white circles indicate unmethylated cytosines that are part of GC dinucleotides. The fraction of total cells that demonstrated each protection pattern is indicated on the right. The SDs for the three experiments were all less than 1.5% [77] (the figure legend is re-written according to the literature).

found that nucleosomes are also well-positioned at both ends of the exon in multicellular [81]. But yeast lacks the feature since its genomic DNA does not include introns. Moreover, upon stress or mutation, nucleosome dynamics frequently occurs at promoters in yeast cells [70]. But in human cells, the nucleosomes alter mainly at enhancers [82, 83].

5. Nucleosome alteration (dynamics) during stress and histone mutation

5.1. Nucleosome alteration upon mutating at modifiable histone residues

Histones are the fundamental element of nucleosomes, and histone mutation do have direct influence on the genome-wide nucleosome organization.

Mutations in histone H3 N-terminal can affect the binding of Chd1, RSC, and SWI/SNF on chromatin, thus having a role in repositioning nucleosomes [84]. Using a native gel electrophoresis

experiment, we can quantitatively track the loss of nucleosome in different histone mutations. As for influence on RSC repositioning, mutations of H3 R42A and R49A rank the first, both raise the original rate in wild-type nucleosomes up to 2.1-fold. H3 I51A mutations has the least effect on products of RSC directed remodeling (**Figure 14**), indicating that H3 I51A is capable of suppressing the nucleosome-unraveling function of RSC.

SWI/SNF-independent (Sin) mutants have various effects on nucleosome alteration. Class I Sin mutants like H4 R45 has the greatest effect; they may completely vanish certain protein-DNA interactions. Influence of class II mutants is relatively mild, which just do little modification on solvent structure as well as the histone octamer main chain conformation,

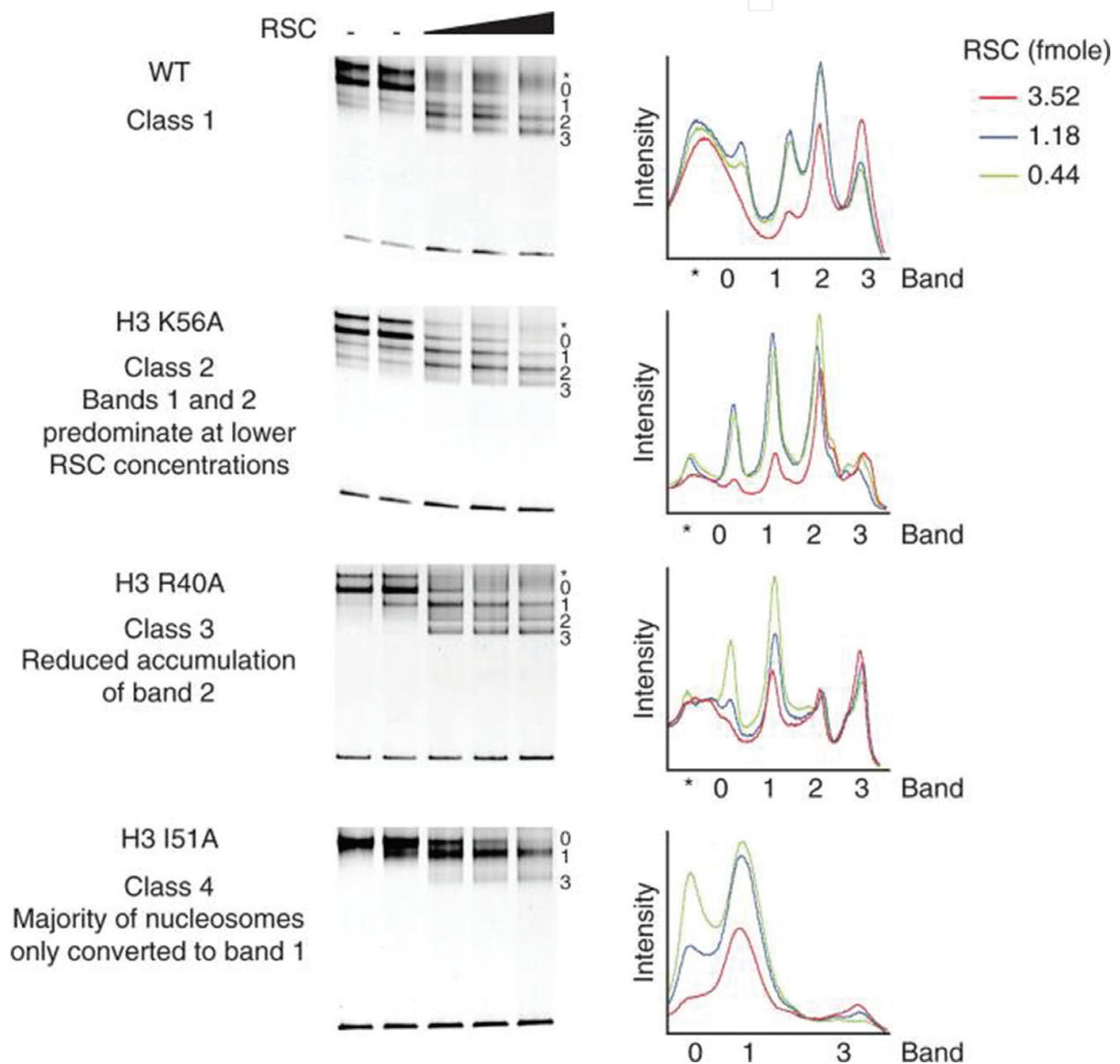


Figure 14. RSC remodeling products in nucleosome with different H3 mutations. H3 mutations could influence the repositioned products, and these mutations are separated into four classes according to the distribution pattern of the products (i.e., bands 1,2 and 3, and band 0 is the origin position). We use four line chart describing band intensities of the RSC remodeled nucleosomes one to one [84] (the legend is rewritten according to the literature).

and class III mutants merely weaken the interactions between octamer and DNA. The changes of protein-NDA interactions lead to an increment of nucleosome-sliding rates.

Histone depletion also has influences on nucleosome. For instance, +1 nucleosomes will notably shift away from the TSS (transcription start sites) when conducting histone H4 depletion in nucleosomes, and +2, +3, and +4 nucleosomes also showed different levels of movement away from the TSS [85]. This was first founded in the study of Harm van Bakel et al., with an excellent idea of researching nucleosome reposition under promoter-closing condition (Figure 15).

H3 depletion also causes changes of nucleosome occupancy in genome-wide manner [86]. Depleting HHT1 and using GAL1 promoter to control HHT2 (HHT1 and HHT2 are H3 coding genes) as HHT2 gene nearly does not express when strains grow in dextrose but not in galactose, histone H3 completely disappeared in *S. cerevisiae* (Figure 16). In this way, Andrea J. got four strains with different types and carbon sources. In strains with the histone H3 deleted (3 hours), severe changes in nucleosome organization were observed from normal histone levels strain (3 hours) while two types of strains are quite similar in the start (Figure 16B). Upon H3 depletion, weakness appears over the whole nucleosomes. An overall view of whole-genome correlation between nucleosome occupancy profiles of normal wild type and H3 depletion strains exhibit an expected decrease compared with ones of H3 shut-off strains grown in galactose (e.g., H3 not deleted) and normal wild type (Figure 16). More clearly, there is an evident nucleosome positioning decrement along with the movement from +1 and +2 nucleosomes to the gene's transcription termination site (TTS).

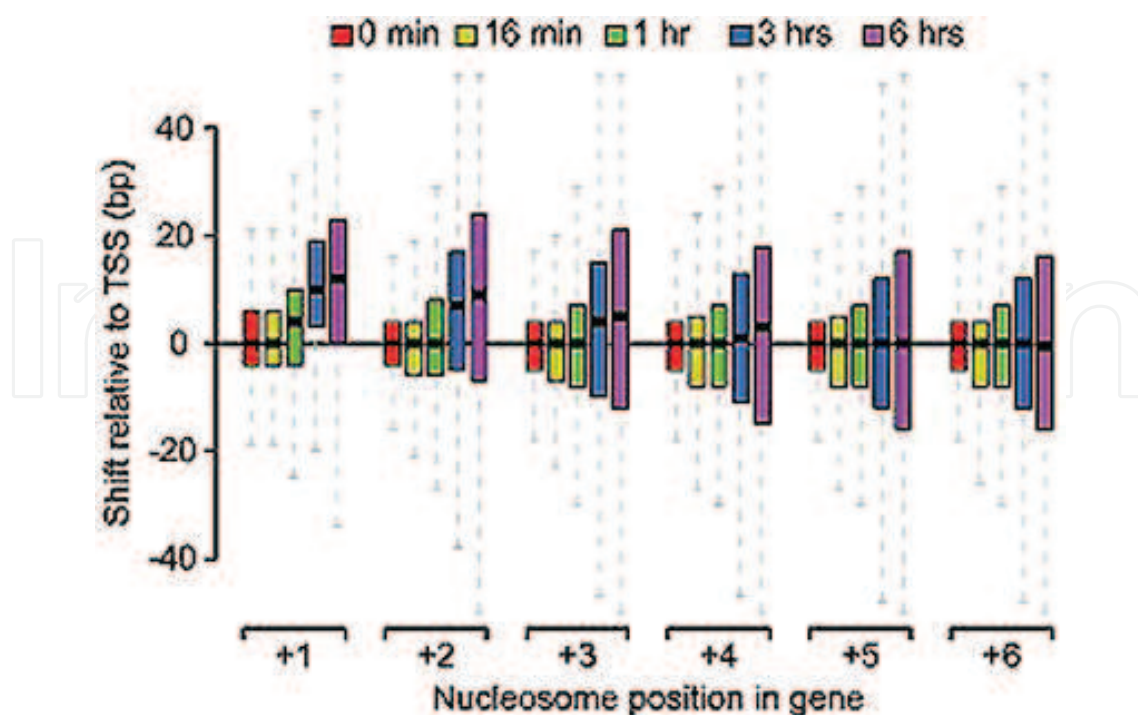


Figure 15. A box plot of nucleosome shifts relative to TSS caused by histone H4 depletion [85].

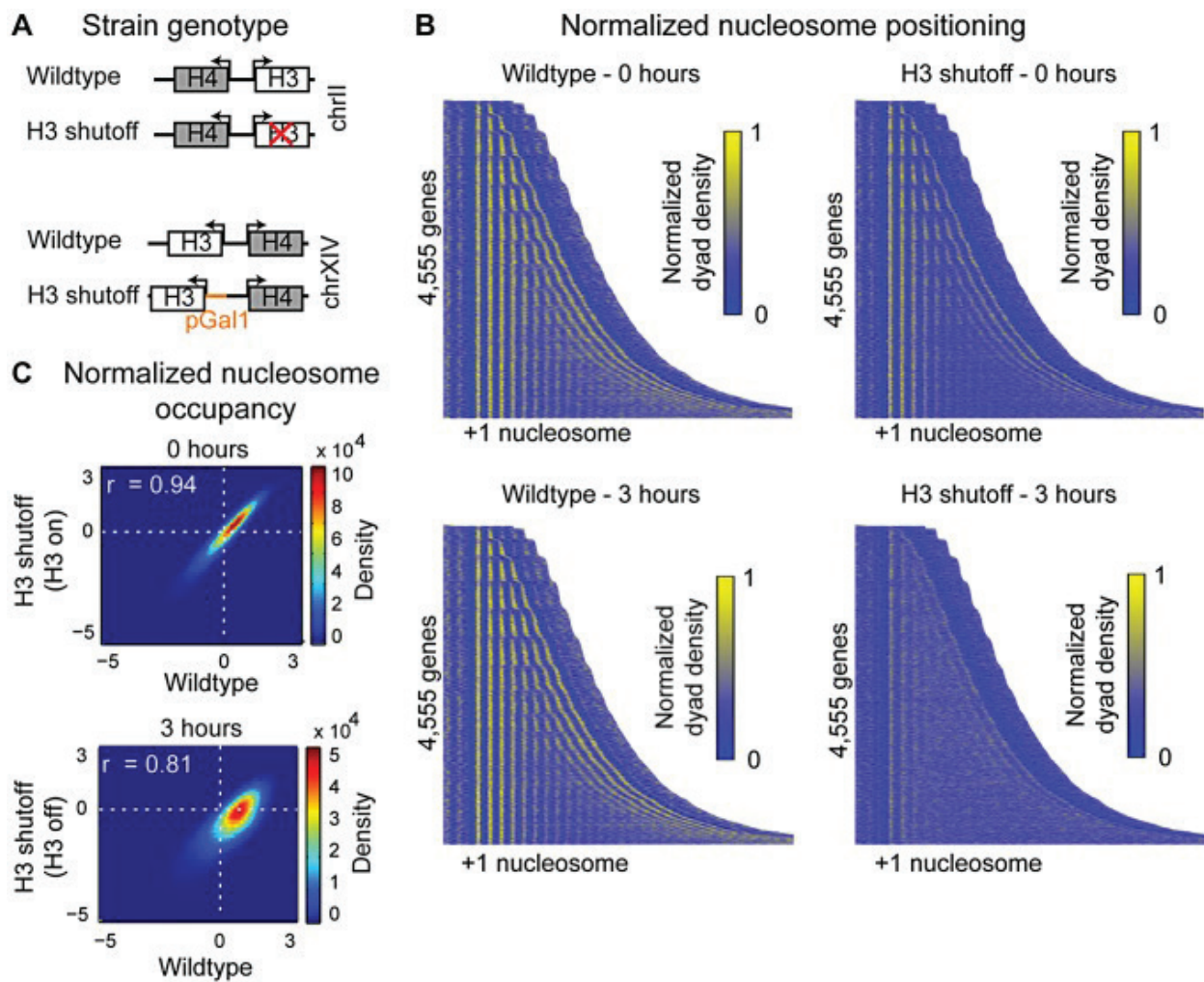


Figure 16. Histone H3 depletion alters nucleosome occupancy genome-wide. (A) Mechanism for H3 depletion in strains. (B) The average, Gaussian-smoothed dyad density for the wild-type and H3 shutoff strains at 0 and 3 hours were aligned on the +1 nucleosomes for 4555 genes. H3 depletion results in a loss of regular positioning of nucleosomes internal to the gene. (C) Log2 normalized average nucleosome occupancy in the wild-type strain versus the H3 shutoff strain. Genome-wide nucleosome occupancy in the two strains is similar prior to H3 depletion ($r = 0.94$) but decreases following H3 depletion ($r = 0.81$) [86] (the figure legend is rewritten according to the literature).

5.2. Nucleosome alteration upon heat shock for yeast

Several kinds of changes on carbon source for yeast can alter nucleosome positioning directly or indirectly. After a heat shock, nucleosome occupancy usually becomes higher at promoters that are repressed and the condition is on the contrary at activated ones [87]. A negative correlation is suggested between nucleosome occupancy and transcription levels caused by heat shock. PAPAS is a long non-coding RNA (lncRNA) and was tested carrying out help in the repression of Pol I transcription as it is upregulated by heat shock [88]. CHD4/NuRD is the remodeling complex that could prevent transcription in a way of accessing nucleosomes which should have bound around promoters onto the transcriptional off position. An examination for nucleosome positioning in normal and heat-shocked cells indicated that heat shock led to a promoter-bound nucleosome movement

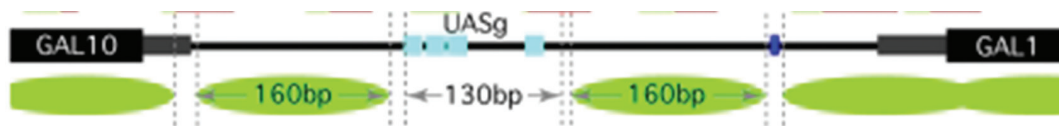


Figure 17. GAL 1, GAL 10, and Gal 4 binding sites (cyan rectangles) locus. Nucleosomes are presented by green circles, and a TATA box is located between Gal4 and GAL1, with blue rectangles [89] (the figure legend is rewritten according to the literature).

award downstream position via promoting PAPAS expression which could induce recruitment of CHD4/NuRD to rDNA [88].

5.3. Nucleosome alteration upon changing carbon source for yeast

Several kinds of changes on carbon source for yeast can alter nucleosome positioning directly or indirectly. Gal4, a transcriptional activator discovered in *S. cerevisiae* has been intensively studied. Two genes, GAL1 and GAL10 are both regulated by Gal4 (**Figure 17**) [89]. It was found that GAL1 promoter nucleosomes became absent from cells grown for many generations in galactose. But by ChIP experiments, Gal4 is found always present both before and after the nutrition shift. In fact, the follow-up Gal80-absence comparison revealed that galactose could remove Gal80 from nucleosomes, an inhibitor of Gal4. Then the recruiting function of freed Gal4 is quickly motivated, leading SWI/SWF binding to the genes. And this always goes with promoter nucleosomes removal as another two ChIP experiments shows.

Besides, the influence of glucose on nucleosome reassembly was affected by the presence of galactose [89]. The transcription factor Msn2, which is recognized with stress-response feature, not only participates in quite a number of environmental stress response as a mediator but also proactively functions in the restructure activities of nucleosome-depleted region (NDR) during transcriptional reprogramming [90]. Msn2 usually binds to small parts of stress response elements (STREs) and a glucose-to-glycerol downshift could apparently promote Msn2 occupancy near STREs (**Figure 18**). Moreover, the nutrition downshift-stress also enables Msn2 to promote the nucleosome repositioning over promoters of genes. It is concluded that Msn2 has a main function of removing the nucleosomes-binding to promoter regions during gene activation and acts negative role in these regions when genes expression is in low level.

5.4. Nucleosome alterations caused by mutations at modifiable histone residues in *S. cerevisiae*

Histone proteins can be modified by chemical modifications on particular residues. We examined the effect of substituting modifiable residues of four core histones with the non-modifiable residue alanine on nucleosome dynamics [70]. We mapped the genome-wide nucleosomes in 22 histone mutants of *S. cerevisiae* and compared the nucleosome alterations relative to the wild-type strain. The results indicated that different types of histone mutation resulted in different phenotypes and a distinct reorganization of nucleosomes. Nucleosome occupancy was altered at telomeres, but not at centromeres. The first nucleosomes upstream (−1) and downstream (+1) of the TSS were more dynamic than other nucleosomes (**Figure 19**). Mutations in histones affected the nucleosome array downstream of the TSS. Highly expressed genes, such as ribosome genes and genes involved in glycolysis, showed increased nucleosome

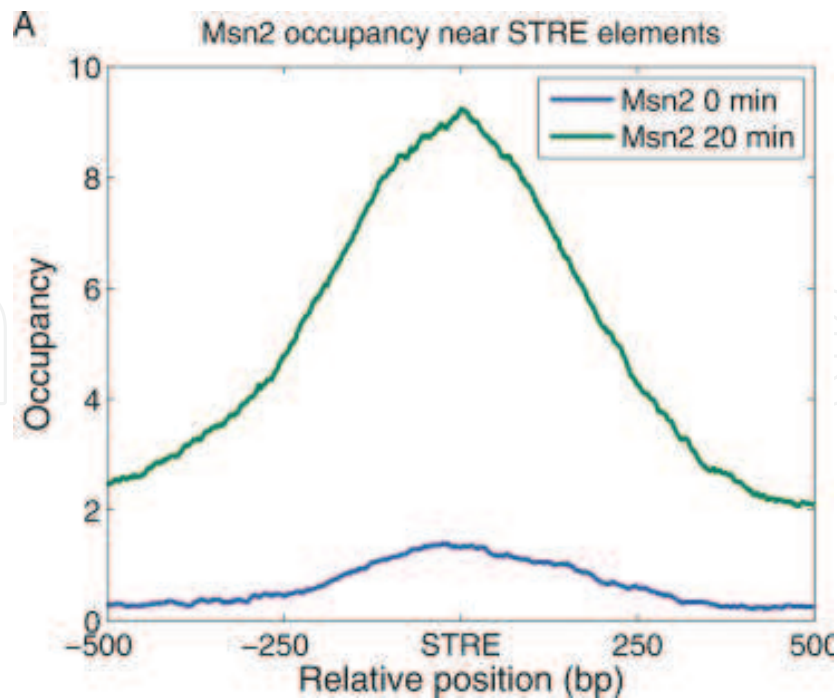


Figure 18. The average Msn2 occupancy near a certain STREs before a carbon source (glucose) replacement with glycerol and 20 min later [89] (the legend is re-written according to the literature).

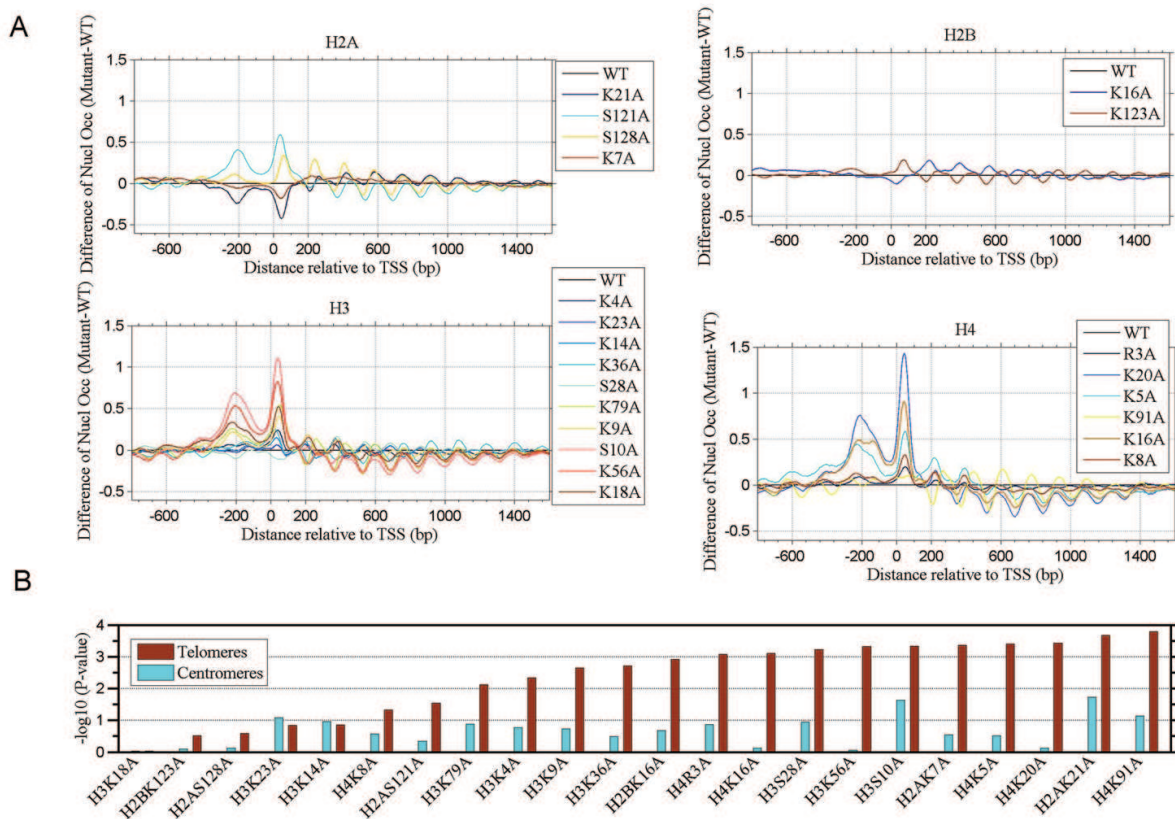


Figure 19. Nucleosome occupancy was altered at telomeres and promoters upon mutating the modifiable residues of the histones. (A) Average differences in nucleosome occupancy around the TSS between mutant and wild-type strains. (B) Difference significance of nucleosome occupancy at telomeres and centromeres between each mutant and the wild-type strain. All P-values ($-\log_{10}$) were calculated with a two-sample t-test [70].

occupancy in many types of histone mutant. In particular, the H3K56A mutant exhibited a high percentage of dynamic genomic regions, decreased nucleosome occupancy at telomeres, increased occupancy at the +1 and -1 nucleosomes, and a slow growth phenotype under stress conditions. Our findings provide insight into the influence of histone mutations on nucleosome dynamics.

6. Htz1 dynamics on chromatin and its effect on nucleosome stability

6.1. Htz1 and nucleosome, Htz1 and transcription

Yeast histone H2A variant Htz1, which is called H2A.Z in mammalian, plays important roles in DNA transactions. Zhang et al. gave a detailed study for the genome-wide dynamics of Htz1 [91]. Firstly, Htz1 occupancy is highly reproducible ($r \geq 0.94$). Secondly, Bdf1 (a component of Swr1 complex), Gcn5 (a histone acetyltransferase) and histone acetylation all play a part in Htz1 occupancy, as well as Swr1. At several specific locations, Swr1 complex is indispensable to meet the requirements for Htz1 deposition. There are obvious correlations between Htz1 and some histone acetylation, implying Htz1 occupies genes in their repressed/basal states, and Htz1 occupancy was reduced in strains with little Gcn5 or Bdf1. Thirdly, Htz1 shows much greater preference than the poor performance of H2A in occupying promoters. Htz1 occupancy is negatively correlated to the presence of a TATA box, suggesting that the occupancy prefers TATA-less promoters. Fourthly, gene activation associates Htz1 loss from promoters.

Zhang et al. presented a model to explain the mechanism that how Htz1 works to regulate transcription (**Figure 20**) [91]. In the particular repressed/basal genes, a nucleosome with Htz1 occupies the promoters and tends to TATA-less regions. Bdf1, a component of SWR1 complex, could promote the process and helps targeting. Loss of Bdf1 could confer a decrement of Htz1 occupancy. SWR1 complex is necessary for deposition as its recruitment involves physical interactions between SWR1 and DNA sequence-specific transcriptional regulators, physical interactions between SWR1 and promoter binding initiation factors and finding histone modification via Bdf1 or other SWR1 components. Gcn5 does not just help target deposition, but also acetylates H3K14 and other residues which may well be the primary reason for the association between Htz1 occupancy and histone acetylation. But Htz1 takes no particular role in making favor of repressing genes, even though it is observed in a high frequency during repression. In fact, Htz1 keeps a balance presence between the repressed and basal states for full activation. When genes states transit from basal to active, chromatin remodeling factors take in action and activators bind to the enhancer. All of the above likely contribute to the Htz1 nucleosome replacement, which promotes activation via giving way to occupancy of certain transcription factors (**Figure 20**).

Martins-Taylor et al. studied Htz1 in a new aspect and revealed that there were some relationship between Htz1 and the cell-cycle progression requirement of establishing transcriptional silencing [92]. Htz1 appeared to work in a direct way to restrict the spread of silent chromatin

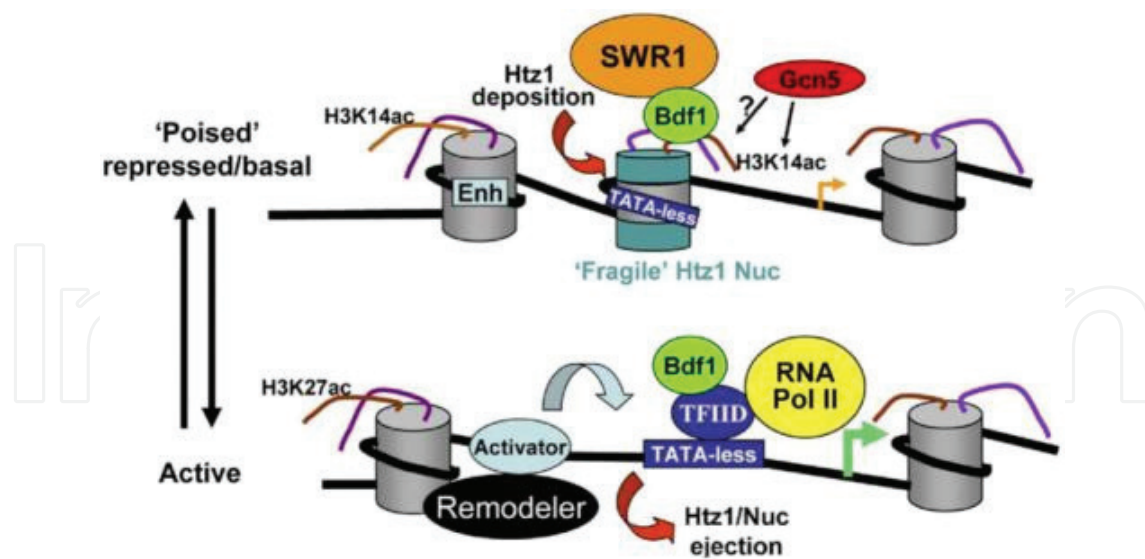


Figure 20. Transcriptional regulation model of Htz1 [91].

from the telomere, and the deletion of genes coding Htz1 could make the establishment of silent chromatin independent from cell-cycle progression.

6.2. Nucleosome, Pol II, Chz1, Htz1, and Spt16

Pol II (RNA polymerase II) promotes the transcription of DNA and is positively associated with the transcription rate. At the beginning of transcription initiation, Pol II are assembled with general transcription factors (GTFs) to make up the pre-initiation complex, binding onto the promoter to initiate transcription [93]. The Htz1 generally occupies the Pol II promoters and affects the combination of GTFs with Pol II, thus inhibiting transcription [91, 94]. Chz1 is an H2B-specific chaperone that delivers Htz1 for H2A substitution [95]. The transcriptional elongation factor FACT is an indispensable component in achieving the process of eliminating the nucleosome block in transcriptional elongation [96]. In yeast cells, Spt16 and Pob3 are the counterparts of FACT. Spt16 destroys the nucleosomes before the running of Pol II complex and reconstructs them after the running. Also, Spt16 has a role of chaperone.

We revealed that Spt16 and Pol II interact with each other and together affect or be affected by gene transcription as they both bind at exposed gene regions, and are positively correlated with the transcription rate (**Figure 21**) [97]. Importantly, Spt16 prefers genes without Htz1 only when Chz1 exists. This discrimination may not be caused for that there are direct interaction mechanism, but is probably to meet the need of transcription initiation. It is found that Chz1 deletion prevents Htz1 occupancy at promoters and telomeres in previous study. Also, in the *chz1*-deletion mutant, Spt16 binding at ribosomal genes was lost, suggesting that Chz1 is prior in Htz1-bound genes and thus Spt16 has no more binding chances.

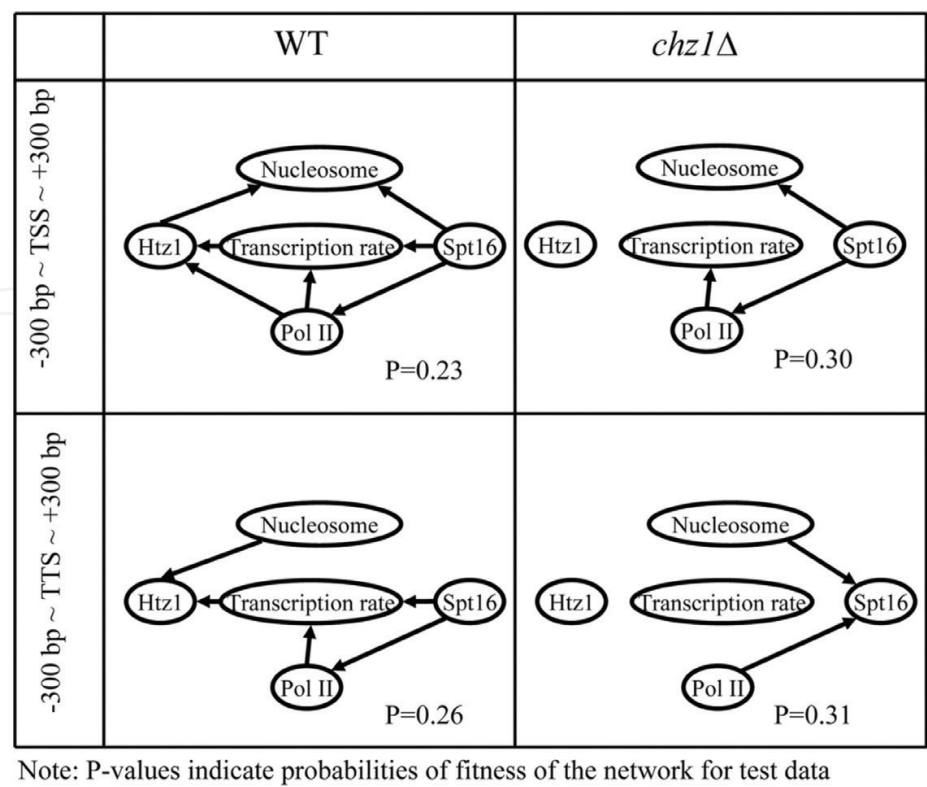


Figure 21. Inferring causal relationships between Spt16, Htz1, Pol II, nucleosome occupancy and transcription rate using Bayesian networks [97]. P-Value indicate probabilities of fitness of the network for test data.

Funding

The writing work was supported by the National Natural Science Foundation of China (No. 31371339 and No. 81660471) and Natural Science Foundation of Xinjiang Province of China (No. 2015211C057).

Author details

Hongde Liu^{1*}, Weiheng Ma¹, Jiahao Xie¹, Huamei Li¹, Kun Luo², Donghui Luo², Lei Liu³ and Xiao Sun¹

*Address all correspondence to: liuhongde@seu.edu.cn

1 State Key Laboratory of Bioelectronics, School of Biological Science & Medical Engineering, Southeast University, Nanjing, China

2 Department of Neurosurgery, Xinjiang Evidence-Based Medicine Research Institute, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, China

3 College of Foreign Language & Literature, Northwest Normal University, Lanzhou, China

References

- [1] Volpe TA, Kidner C, Hall IM, Teng G, Grewal SIS, Martienssen RA. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*. 2002; **297**(5588):1833-1837
- [2] Collins FS, Lander ES, Rogers J, Waterston RH, Conso IHGS. Finishing the euchromatic sequence of the human genome. *Nature*. 2004; **431**(7011):931-945
- [3] Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. *Nature*. 1997; **389**(6648):251-260
- [4] Li GH, Zhu P. Structure and organization of chromatin fiber in the nucleus. *FEBS Letters*. 2015; **589**(20):2893-2904
- [5] Arnaudo AM, Molden RC, Garcia BA. Revealing histone variant induced changes via quantitative proteomics. *Critical Reviews in Biochemistry and Molecular Biology*. 2011; **46**(4):284-294
- [6] Reynolds SM, Bilmes JA, Noble WS. Learning a weighted sequence model of the nucleosome core and linker yields more accurate predictions in *Saccharomyces cerevisiae* and *Homo sapiens*. *PLoS Computational Biology*. 2010; **6**(7):e1000834
- [7] Benevolenskaya EV. Histone H3K4 demethylases are essential in development and differentiation. *Biochemistry and Cell Biology*. 2007; **85**(4):435-443
- [8] Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007; **129**(4):823-837
- [9] Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, Vakoc AL, Kim JE, Chen JJ, Lazar MA, et al. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Molecular and Cellular Biology*. 2008; **28**(8):2825-2839
- [10] Rosenfeld JA, Wang ZB, Schones DE, Zhao K, DeSalle R, Zhang MQ. Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics*. 2009; **10**(1):143
- [11] Richmond TJ. Hot papers—Crystal structure—Crystal structure of the nucleosome core particle at 2.8 angstrom resolution by K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond—Comments. *Scientist*. 1999; **13**(23):15-15
- [12] White CL, Suto RK, Luger K. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *The EMBO Journal*. 2001; **20**(18):5207-5218
- [13] Yuan GC, Liu YJ, Dion MF, Slack MD, LF W, Altschuler SJ, Rando OJ. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science*. 2005; **309**(5734):626-630

- [14] Satchwell SC, Drew HR, Travers AA. Sequence periodicities in chicken nucleosome core DNA. *Journal of Molecular Biology*. 1986;**191**(4):659-675
- [15] Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, Wang JP, Widom J. A genomic code for nucleosome positioning. *Nature*. 2006;**442**(7104):772-778
- [16] Struhl K, Segal E. Determinants of nucleosome positioning. *Nature Structural & Molecular Biology*. 2013;**20**(3):267-273
- [17] Herzel H, Weiss O, Trifonov EN. 10-11 bp periodicities in complete genomes reflect protein structure and DNA folding. *Bioinformatics*. 1999;**15**(3):187-193
- [18] Zhong J, Luo K, Winter PS, Crawford GE, Iversen ES, Hartemink AJ. Mapping nucleosome positions using DNase-seq. *Genome Research*. 2016;**26**(3):351-364
- [19] Liu H, Wu J, Xie J, Yang X, Lu Z, Sun X. Characteristics of nucleosome core DNA and their applications in predicting nucleosome positions. *Biophysical Journal*. 2008;**94**(12):4597-4604
- [20] Liu H, Lin S, Cai Z, Sun X. Role of 10-11bp periodicities of eukaryotic DNA sequence in nucleosome positioning. *Bio Systems*. 2011;**105**(3):295-299
- [21] Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, Lubling Y, Widom J, Segal E. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Computational Biology*. 2008;**4**(11):e1000216
- [22] Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, et al. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*. 2009;**458**(7236):362-366
- [23] Guo SH, Deng EZ, LQ X, Ding H, Lin H, Chen W, Chou KC. iNuc-PseKNC: A sequence-based predictor for predicting nucleosome positioning in genomes with pseudo k-tuple nucleotide composition. *Bioinformatics*. 2014;**30**(11):1522-1529
- [24] Brogaard K, Xi LQ, Wang JP, Widom J. A map of nucleosome positions in yeast at base-pair resolution. *Nature*. 2012;**486**(7404):496-501
- [25] Xi LQ, Fondufe-Mittendorf Y, Xia L, Flatow J, Widom J, Wang JP. Predicting nucleosome positioning using a duration Hidden Markov Model. *BMC Bioinformatics*. 2010;**11**(1):346
- [26] Wang LG, Huang HJ, Dougherty G, Zhao Y, Hossain A, Kocher JPA. Epidaurus: Aggregation and integration analysis of prostate cancer epigenome. *Nucleic Acids Research*. 2015;**43**(2):e7
- [27] Di Gesu V, Lo Bosco G, Pinello L, Yuan GC, Corona DFV. A multi-layer method to study genome-scale positions of nucleosomes. *Genomics*. 2009;**93**(2):140-145
- [28] Locke G, Tolkunov D, Moqtaderi Z, Struhl K, Morozov AV. High-throughput sequencing reveals a simple model of nucleosome energetics. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;**107**(49):20998-21003
- [29] Cui F, Zhurkin VB. Rotational positioning of nucleosomes facilitates selective binding of p53 to response elements associated with cell cycle arrest. *Nucleic Acids Research*. 2014;**42**(2):836-847

- [30] Cui F, Zhurkin VB. Structure-based analysis of DNA sequence patterns guiding nucleosome positioning in vitro. *Journal of Biomolecular Structure & Dynamics*. 2010;**27**(6): 821-841
- [31] Levitsky VG, Katokhin AV, Podkolodnaya OA, Furman DP, Kolchanov NA. NPRD: Nucleosome positioning region database. *Nucleic Acids Research*. 2005;**33**:D67-D70
- [32] Nguyen N, Vo A, Won KJ. A wavelet-based method to exploit epigenomic language in the regulatory region. *Bioinformatics*. 2014;**30**(7):908-914
- [33] Stolz RC, Bishop TC. ICM Web: The interactive chromatin modeling web server. *Nucleic Acids Research*. 2010;**38**:W254-W261
- [34] Gabdank I, Barash D, Trifonov EN. FineStr: A web server for single-base-resolution nucleosome positioning. *Bioinformatics*. 2010;**26**(6):845-846
- [35] Chen W, Lin H, Feng PM, Ding C, Zuo YC, Chou KC. iNuc-PhysChem: A sequence-based predictor for identifying nucleosomes via physicochemical properties. *PLoS One*. 2012;**7**(10):e47843
- [36] Weiner A, Hughes A, Yassour M, Rando OJ, Friedman N. High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. *Genome Research*. 2010;**20**(1):90-100
- [37] Chen KF, Xi YX, Pan XW, Li ZY, Kaestner K, Tyler J, Dent S, He XW, Li W. DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Research*. 2013;**23**(2):341-351
- [38] Meyer CA, He HSH, Brown M, Liu XS. BINOCh: Binding inference from nucleosome occupancy changes. *Bioinformatics*. 2011;**27**(13):1867-1868
- [39] Woo SS, Zhang XK, Sauteraud R, Robert F, Gottardo R. PING 2.0: An R/bioconductor package for nucleosome positioning using next-generation sequencing data. *Bioinformatics*. 2013;**29**(16):2049-2050
- [40] Humburg P, Helliwell CA, Bulger D, Stone G. ChIPseqR: Analysis of ChIP-seq experiments. *BMC Bioinformatics*. 2011;**12**(1):39
- [41] Quintales L, Vazquez E, Antequera F. Comparative analysis of methods for genome-wide nucleosome cartography. *Briefings in Bioinformatics*. 2015;**16**(4):576-587
- [42] Schopflin R, Teif VB, Muller O, Weinberg C, Rippe K, Wedemann G. Modeling nucleosome position distributions from experimental nucleosome positioning maps. *Bioinformatics*. 2013;**29**(19):2380-2386
- [43] Mammana A, Vingron M, Chung HR. Inferring nucleosome positions with their histone mark annotation from ChIP data. *Bioinformatics*. 2013;**29**(20):2547-2554
- [44] Fu K, Tang QZ, Feng JX, Liu XS, Zhang Y. DiNuP: A systematic approach to identify regions of differential nucleosome positioning. *Bioinformatics*. 2012;**28**(15):1965-1971

- [45] Vainshtein Y, Rippe K, Teif VB. NucTools: Analysis of chromatin feature occupancy profiles from high-throughput sequencing data. *BMC Genomics*. 2017;**18**(1):158
- [46] Liu LJ, Xie JM, Sun X, Luo K, Qin ZS, Liu HD. An approach of identifying differential nucleosome regions in multiple samples. *BMC Genomics*. 2017;**18**(1):135
- [47] Lai WKM, Buck MJ. ArchAlign: Coordinate-free chromatin alignment reveals novel architectures. *Genome Biology*. 2010;**11**:12
- [48] Straver R, Oudejans CBM, Sistermans EA, Reinders MJT. Calculating the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles. *Prenatal diagnosis*. 2016;**36**(7):614-621
- [49] Nellore A, Bobkov K, Howe E, Pankov A, Diaz A, Song JS. NSeq: A multithreaded java application for finding positioned nucleosomes from sequencing data. *Frontiers in Genetics*. 2012;**3**:320
- [50] Lai WK, Bard JE, Buck MJ. ArchTEx: Accurate extraction and visualization of next-generation sequence data. *Bioinformatics*. 2012;**28**(7):1021-1023
- [51] Polishko A, Bunnik EM, Le Roch KG, Lonardi S. PuFFIN—A parameter-free method to build nucleosome maps from paired-end reads. *BMC Bioinformatics*. 2014;**15**(Suppl 9): S11
- [52] Zhang Y, Shin H, Song JS, Lei Y, Liu XS. Identifying positioned nucleosomes with epigenetic marks in human from ChIP-Seq. *BMC Genomics*. 2008;**9**:537
- [53] Vignali M, Hassan AH, Neely KE, Workman JL. ATP-dependent chromatin-remodeling complexes. *Molecular and Cellular Biology*. 2000;**20**(6):1899-1910
- [54] Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, part II: ATP-dependent chromatin remodeling. *Trends in Molecular Medicine*. 2007;**13**(9):373-380
- [55] Mayes K, Qiu ZJ, Alhazmi A, Landry JW. ATP-dependent chromatin remodeling complexes as novel targets for cancer therapy. *Advances in Cancer Research*. 2014;**121**:183-233
- [56] Jiang CZ, Pugh BF. Nucleosome positioning and gene regulation: Advances through genomics. *Nature Reviews Genetics*. 2009;**10**(3):161-172
- [57] Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research*. 2008;**18**(7):1073-1083
- [58] Zentner GE, Henikoff S. Surveying the epigenomic landscape, one base at a time. *Genome Biology*. 2012;**13**(10):250
- [59] Meyer CA, Liu XS. Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nature Reviews Genetics*. 2014;**15**(11):709-721
- [60] Tsompana M, Buck MJ. Chromatin accessibility: A window into the genome. *Epigenetics & Chromatin*. 2014;**7**(1):33

- [61] Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. High-resolution mapping and characterization of open chromatin across the genome. *Cell*. 2008;**132**(2):311-322
- [62] Song L, Crawford GE. DNase-seq: A high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor Protocols*. 2010;**2010**(2) pdb. prot5384
- [63] Weintraub H, Groudine M. Chromosomal subunits in active genes have an altered conformation. *Science*. 1976;**193**(4256):848-856
- [64] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*. 2013;**10**(12):1213-1218
- [65] Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols in Molecular Biology*; **109** 21.29.1-9
- [66] Park PJ. ChIP-seq: Advantages and challenges of a maturing technology. *Nature Reviews Genetics*. 2009;**10**(10):669-680
- [67] Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD. FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Research*. 2007;**17**(6):877-885
- [68] Rhee HS, Pugh BF. ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. *Current Protocols in Molecular Biology*. 2012 Chapter 21:Unit 21.24
- [69] Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*. 2013;**14**(2):178-192
- [70] Liu H, Wang P, Liu L, Min Z, Luo K, Wan Y. Nucleosome alterations caused by mutations at modifiable histone residues in *Saccharomyces cerevisiae*. *Scientific Reports*. 2015; **5**:15583
- [71] Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nature Protocols*. 2012;**7**(9):1728-1740
- [72] Benjamini Y, Speed TP. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Research*. 2012;**40**(10):e72-e72
- [73] Bradnam KR, Seoighe C, Sharp PM, Wolfe KH. G+C content variation along and among *Saccharomyces cerevisiae* chromosomes. *Molecular Biology and Evolution*. 1999;**16**(5):666-675
- [74] Dingwall C, Lomonosoff GP, Laskey RA. High sequence specificity of micrococcal nuclease. *Nucleic Acids Research*. 1981;**9**(12):2659-2673
- [75] Chung HR, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, Sperling SR, Vingron M. The effect of micrococcal nuclease digestion on nucleosome positioning data. *PLoS One*. 2010;**5**(12):e15754

- [76] Ozonov EA, van Nimwegen E. Nucleosome free regions in yeast promoters result from competitive binding of transcription factors that interact with chromatin modifiers. *PLoS Computational Biology*. 2013;**9**(8):e1003181
- [77] Small EC, Xi L, Wang JP, Widom J, Licht JD. Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;**111**(24):E2462-E2471
- [78] Liu H, Luo K, Wen H, Ma X, Xie J, Sun X. Quantitative analysis reveals increased histone modifications and a broad nucleosome-free region bound by histone acetylases in highly expressed genes in human CD4+ T cells. *Genomics*. 2013;**101**(2):113-119
- [79] Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, et al. Nucleosome organization in the drosophila genome. *Nature*. 2008;**453**(7193):358-362
- [80] Radman-Livaja M, Rando OJ. Nucleosome positioning: How is it established, and why does it matter? *Developmental Biology*. 2010;**339**(2):258-266
- [81] Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcarcel J, Guigo R. Nucleosome positioning as a determinant of exon recognition. *Nature Structural & Molecular Biology*. 2009;**16**(9):996-1001
- [82] Skalska L, Stojnic R, Li J, Fischer B, Cerda-Moya G, Sakai H, Tajbakhsh S, Russell S, Adryan B, Bray SJ. Chromatin signatures at Notch-regulated enhancers reveal large-scale changes in H3K56ac upon activation. *The EMBO Journal*. 2015;**34**(14):1889-1904
- [83] Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;**459**(7243):108-112
- [84] Somers J, Owen-Hughes T. Mutations to the histone H3 alpha N region selectively alter the outcome of ATP-dependent nucleosome-remodelling reactions. *Nucleic Acids Research*. 2009;**37**(8):2504-2513
- [85] van Bakel H, Tsui K, Gebbia M, Mnaimneh S, Hughes TR, Nislow C. A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. *PLoS Genetics*. 2013;**9**(5):e1003479
- [86] Gossett AJ, Lieb JD. In vivo effects of histone H3 depletion on nucleosome occupancy and position in *Saccharomyces cerevisiae*. *PLoS Genetics*. 2012;**8**(6):e1002771
- [87] Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nature Genetics*. 2004;**36**(8):900-905
- [88] Zhao Z, Dammert MA, Hoppe S, Bierhoff H, Grummt I. Heat shock represses rRNA synthesis by inactivation of TIF-IA and lncRNA-dependent changes in nucleosome positioning. *Nucleic Acids Research*. 2016;**44**(17):8144-8152
- [89] Bryant GO, Prabhu V, Floer M, Wang X, Spagna D, Schreiber D, Ptashne M. Activator control of nucleosome occupancy in activation and repression of transcription. *PLoS Biology*. 2008;**6**(12):2928-2939

- [90] Elfving N, Chereji RV, Bharatula V, Bjorklund S, Morozov AV, Broach JR. A dynamic interplay of nucleosome and Msn2 binding regulates kinetics of gene activation and repression following stress. *Nucleic Acids Research*. 2014;**42**(9):5468-5482
- [91] Zhang HY, 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**(2):219-231
- [92] Martins-Taylor K, Sharma U, Rozario T, Holmes SG. H2A.Z (Htz1) controls the cell-cycle-dependent establishment of transcriptional silencing at *Saccharomyces cerevisiae* telomeres. *Genetics*. 2011;**187**(1):89-104
- [93] Sims RJ 3rd, Mandal SS, Reinberg D. Recent highlights of RNA-polymerase-II-mediated transcription. *Current Opinion in Cell Biology*. 2004;**16**(3):263-271
- [94] Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature*. 2007;**446**(7135):572-576
- [95] Luk E, ND V, Patteson K, Mizuguchi G, WH W, Ranjan A, Backus J, Sen S, Lewis M, Bai Y, et al. Chz1, a nuclear chaperone for histone H2AZ. *Molecular Cell*. 2007;**25**(3):357-368
- [96] Formosa T. FACT and the reorganized nucleosome. *Molecular BioSystems*. 2008;**4**(11): 1085-1093
- [97] Liu H, Luo K, Zhou Z, Mu Y, Wan Y. Histone chaperone Chz1 facilitates the disfavouring property of Spt16 to H2A.Z-containing genes in *Saccharomyces cerevisiae*. *The Biochemical Journal*. 2014;**460**(3):387-397

