

RESEARCH ARTICLE

On the role of transcription in positioning nucleosomes

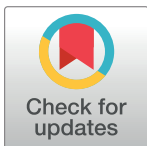
Zhongling Jiang , Bin Zhang *

Departments of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

* binz@mit.edu

Abstract

Nucleosome positioning is crucial for the genome's function. Though the role of DNA sequence in positioning nucleosomes is well understood, a detailed mechanistic understanding on the impact of transcription remains lacking. Using numerical simulations, we investigated the dependence of nucleosome density profiles on transcription level across multiple species. We found that the low nucleosome affinity of yeast, but not mouse, promoters contributes to the formation of phased nucleosomes arrays for inactive genes. For the active genes, a heterogeneous distribution of +1 nucleosomes, caused by a tug-of-war between two types of remodeling enzymes, is essential for reproducing their density profiles. In particular, while positioning enzymes are known to remodel the +1 nucleosome and align it toward the transcription start site (TSS), spacer enzymes that use a pair of nucleosomes as their substrate can shift the nucleosome array away from the TSS. Competition between these enzymes results in two types of nucleosome density profiles with well- and ill-positioned +1 nucleosome. Finally, we showed that Pol II assisted histone exchange, if occurring at a fast speed, can abolish the impact of remodeling enzymes. By elucidating the role of individual factors, our study reconciles the seemingly conflicting results on the overall impact of transcription in positioning nucleosomes across species.

 OPEN ACCESS

Citation: Jiang Z, Zhang B (2021) On the role of transcription in positioning nucleosomes. *PLoS Comput Biol* 17(1): e1008556. <https://doi.org/10.1371/journal.pcbi.1008556>

Editor: Anna R. Panchenko, Queen's University, CANADA

Received: August 26, 2020

Accepted: November 23, 2020

Published: January 8, 2021

Copyright: © 2021 Jiang, Zhang. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Funding: This work was supported by the National Institutes of Health (Grant 1R35GM133580-01) to BZ. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Nucleosome positioning plays a key role in the genome's function by regulating the accessibility of protein binding sites as well as higher-order chromatin organization. Though significant progress has been made towards studying the role of DNA sequence in positioning the nucleosomes, our understanding on the impact of transcription lags behind. Our study uses kinetic simulations to explore the role of DNA sequence specificity, transcription factor binding, enzyme remodeling, and Pol II elongation in positioning nucleosomes. It suggests that the differences in nucleosome density profiles observed at various transcription levels in yeast and mouse embryonic stem cells can be understood from a tug-of-war between two types of remodeling enzymes.

Introduction

Nucleosomes are the fundamental packaging unit of chromatin, comprising 147 base pairs (bp) of DNA wrapped around histone proteins [1]. Their formation helps to fit eukaryotic genomes inside the nucleus but also occludes the DNA binding of protein molecules, including regulatory factors and transcriptional machinery [2, 3]. The precise position of nucleosomes along the DNA sequence, therefore, can critically impact the function of the genome by regulating its accessibility [4–8]. Recent whole-genome sequencing-based studies have indeed revealed the depletion of nucleosomes at many promoter and enhancer regions to accommodate transcription [9, 10]. In addition, nucleosome positioning may affect gene expression indirectly by regulating higher-order chromatin organization [11–14]. For example, protein molecules such as Cohesin and CTCF have been shown to facilitate chromatin folding and the formation of so-called topologically associated domains [15, 16]. These domains promote enhancer-promoter contacts, and their formation relies on the accessibility of CTCF binding sites [17, 18]. Underpinning the molecular determinants of nucleosome positioning is, therefore, of fundamental interest and can provide insight into gene regulatory mechanisms.

Since the DNA molecule undergoes substantial distortion when wrapping around histone proteins, its intrinsic, sequence-specific property can impact the stability, and correspondingly position, of the formed nucleosomes [7, 19]. Numerous studies have found that nucleosomes preferentially occupy DNA segments that are more susceptible to bending and twisting. They have led to the discovery of periodic dinucleotides (AT and TA) along the nucleosome length [9, 20, 21] and intrinsically stiff poly(dA:dT) tracts at nucleosome-depleted regions [22]. Computational models based on such sequence features have been developed to predict *in vivo* nucleosome occupancy [23–25]. Accuracy of such predictions can be hampered, however, by the presence of a variety of processes and activities in the nucleus that may overwrite intrinsic positioning signals from the DNA [26–29].

Transcription is one of such processes that can alter the location of nucleosomes via chromatin remodeling and histone eviction [30]. Due to the consumption of ATP, the kinetics of these movements does not necessarily satisfy detailed balance, and the resulting nucleosome configurations may conflict with the thermodynamic distribution determined from the DNA sequence alone. The impact of transcription is evident from Fig 1 and S2 Fig, where average nucleosome density profiles for genes with varying levels of transcriptional activity are shown to exhibit striking differences. In particular, nucleosomes for more active genes (red) appear less ordered in yeast with less pronounced peaks and valleys when compared with inactive ones (blue). This disordering is particularly striking for the top 200 active genes (yellow). However, the opposite trend is observed for mouse embryonic stem cells (ESC), for which clear patterns emerge from a featureless profile as transcription level elevates. We note that the qualitative trends seen in mouse ESC are conserved across multi-cellular organisms [31–33].

In this paper, we carried out theoretical analysis and numerical simulations of a kinetic model that explicitly considers the impact of several key factors known to impact nucleosome positioning, including the DNA sequence, chromatin remodeling enzymes, and histone exchange (Fig 2). We found that a tug-of-war between two types of remodeling enzymes explains the observed difference between nucleosome density profiles at varying transcription levels and across species. In particular, remodeling enzymes that regulate and reduce inter-nucleosome spacing tend to drive the nucleosome array away from the transcription start site (TSS). On the other hand, positioning enzymes help to align nucleosomes towards the TSS. Competition between these enzymes results in two types of density profiles with well- and ill-positioned +1 nucleosome. Mixing the two profiles at different populations can give rise to results that qualitatively reproduce yeast or mouse ESC data. We further demonstrated that

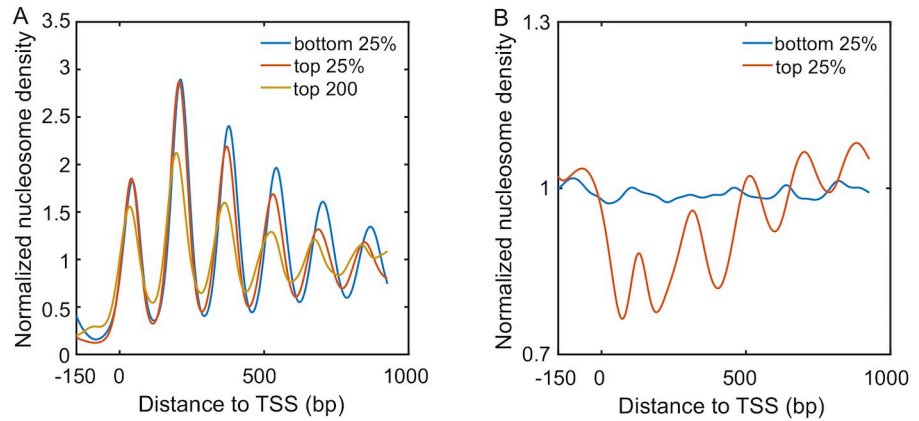


Fig 1. Normalized nucleosome density profiles obtained with a chemical mapping method for *S. cerevisiae* [34] (A) and mouse embryonic stem cells [31] (B) near TSS. After removing genes with more than one promoter [31, 35], 4151 and 18969 genes were considered here for *S. cerevisiae* and mouse, respectively. Genes were separated into quartiles depending on levels of transcription activities, with the bottom and top 25% corresponding to the most inactive and active genes, respectively. Result for the top 200 most active genes for yeast is also shown to highlight the decrease in amplitude with increased transcriptional activity. The density profiles were smoothed with the formula Eq. S1 in the S1 Supporting information. For the raw plots, see S1 Fig.

<https://doi.org/10.1371/journal.pcbi.1008556.g001>

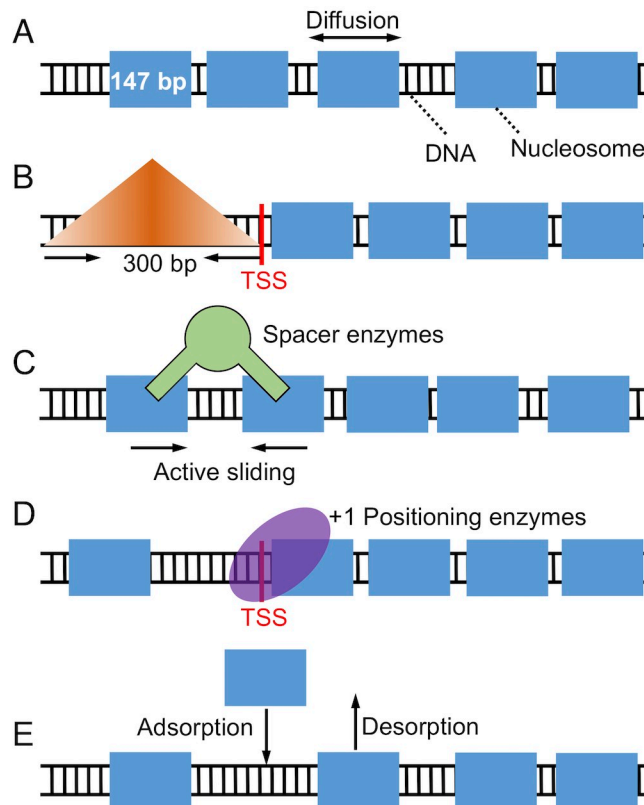


Fig 2. Illustration of the kinetic model used for studying nucleosome positioning that includes thermal diffusion. (A), a barrier in the promoter region that penalizes nucleosome binding (B), spacer enzyme remodeling (C), +1 nucleosome positioning, (D) and histone exchange (E). The DNA is drawn as a black ladder, and histone proteins are represented as blue rectangles. Remodeling enzymes are colored in green and use a pair of nucleosomes as substrate.

<https://doi.org/10.1371/journal.pcbi.1008556.g002>

fast kinetics of histone eviction/adsorption, if induced by RNA polymerase (Pol) II elongation, could reduce or abolish the impact of remodeling enzymes. Our study, therefore, provides insight into the role of transcription in positioning nucleosomes and reconciles the seemingly conflicting trend across species.

Results

DNA sequence contributes to the barrier for forming phased nucleosome array

A striking difference between yeast and mouse ESC is their distinct nucleosome density profiles for genes with minimal transcription activity (blue lines in Fig 1). While for yeast, these genes exhibit oscillatory patterns with well-positioned nucleosomes, the corresponding curve for mouse ESC does not show significant features. Given their low level of transcription, the impact of remodelers and transcription factors are potentially minimum. We wondered whether contributions from DNA sequences could explain nucleosome distributions in these genes.

We extracted the sequences surrounding TSS for 1000 genes with the lowest transcription level from yeast and mouse genome. Using a model introduced by van Noort and coworkers that quantifies nucleosome occupancy based on a periodic function of dinucleotides [24], we determined the nucleosome affinity profile for each DNA segment. As shown in Fig 3A and S3 Fig, the average affinity for yeast genes quantified in terms of binding energy peaks at promoters located on the left of TSS. Promoters of *S. cerevisiae* are, therefore, inherently nucleosome repelling. Mouse genes, on the other hand, exhibit the opposite trend, with the same region being most favorable for nucleosome formation. The difference in promoters' nucleosome affinity is particularly interesting in light of the statistical positioning model [10, 36], which argues that the presence of a repulsive potential could create nucleosome-free regions and align downstream nucleosomes.

To more directly evaluate the impact of DNA sequences, we carried out simulations for each inactive gene to determine their average density profiles using the predicted sequence-specific nucleosome affinity. Details for these simulations are provided in the *Materials and Methods*. As shown in Fig 3B, it is evident that for yeast but not mouse, there is a depletion of nucleosomes on the left side of TSS. This depletion gives rise to weak peaks at +1 and +2 nucleosomes. Notably, the height of the peaks do not differ significantly from the results for mouse genes and is much lower than the ones seen in Fig 1A.

The less prominent features seen in simulated density profiles are consistent with results from nucleosomes reconstituted in vitro using only genomic DNA and histones [37]. They can be attributed to the relatively small fluctuation in predicted nucleosome affinity. Additional transcription factors and remodeling enzymes could take advantage of the weakened affinity to occupy promoters [38], even for inactive genes, further driving the depletion of nucleosomes and effectively raising the barrier height. Without over complicating the model, we incorporated the effect of these proteins with a repulsive potential. As shown in Fig 3C, it similarly increases from TSS to the center of the promoter region (-150 bp) as in yeast affinity profile but with a larger slope. For simplicity, we further removed DNA-sequence specific interactions and applied uniform nucleosome binding energy across the lattice for all downstream analysis (see Table in S1 Table). Simulations carried out with the promoter potential and a uniform binding energy resulted in a density profile with clear oscillatory patterns and amplitudes comparable to those seen in experiments. If the promoter potential was removed, the simulations produced a uniform density profile as expected. Therefore, the intrinsic

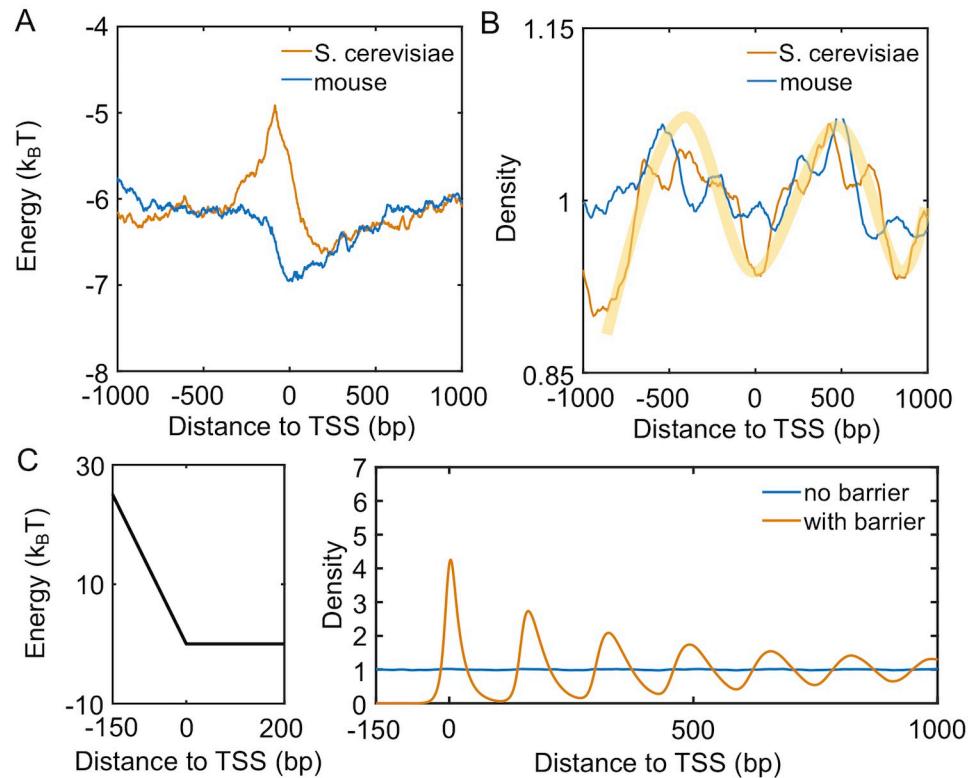


Fig 3. Intrinsic differences in the nucleosome affinity of promoter sequences contribute to the formation of phased nucleosome arrays in yeast, but not mouse, inactive genes. (A) Average nucleosome affinity computed for the 1000 genes with lowest transcription level for yeast (yellow) and mouse (blue). (B) Nucleosome density profiles for the corresponding sequence-specific affinity shown in part A. The shaded yellow curve is shown as a guide for the eye. (C) Nucleosome density profiles for two kinetics models that incorporate a barrier in the promoter region or not. An illustration of the promoter potential is shown on the left.

<https://doi.org/10.1371/journal.pcbi.1008556.g003>

property of yeast promoter sequences and the binding of additional protein molecules, or a lack thereof, help recreate the nucleosome density profiles of inactive genes.

Spacer enzymes induce nucleosome condensation and ill-positioned +1 nucleosome

We next focus on the impact of transcription on nucleosome occupancy in mouse. Fig 1B suggests that as the transcription level increases, nucleosomes become more aligned, as evidenced by the emergence of peaks and valleys. This change could arise from the establishment of nucleosome-free regions at gene promoters to accommodate the arrival of the transcription machinery. Two additional features of the density profile cannot be readily explained by the statistical positioning model, however. First, compared to the curves for yeast (Fig 1A) and from simulations (Fig 3C), the +1 nucleosome in mouse shows a much lower occupancy. Second, the spacing between nucleosomes decreases with the increase of transcriptional activity (S2(B) Fig). Here nucleosome spacing is measured as the distance between two neighboring peaks. Its decrease has also been confirmed in a recent single-cell study that directly measured the distance between nucleosomes from the same DNA molecule [39]. In the following, we explore mechanisms in addition to statistical positioning that can explain these two features.

We note that the decrease of inter-nucleosome spacing upon transcription is indeed a conserved phenomenon and can be readily seen from the yeast profiles as well. In addition, nanopore sequencing of long DNA segments that contain multiple nucleosomes has confirmed the same trend in *Drosophila* [32]. A possible explanation for the spacing change is the recruitment of spacer enzymes to actively transcribed genes. These enzymes use a pair of nucleosomes as their substrate and act as rulers to adjust the length of the linker DNA [27, 37, 40, 41]. Numerical simulations have confirmed their impact on inter-nucleosome distances via examining the so-called radial distribution profile [42–44]. The impact of these spacer enzymes on nucleosome density profiles near TSS remains unclear, however.

We carried out simulations to study the distribution of nucleosomes with the presence of spacer enzymes and a promoter potential. As detailed in the *Materials and Methods* section, these enzymes bind with a pair of neighboring nucleosomes and move them closer by one bp at every step. The rate of such remodeling steps is independent of the underlying energy landscape, and the enzymes break the detailed balance. Using a theory developed by us [44], we mapped the non-equilibrium model with enzymes onto an equivalent and renormalized equilibrium system with effective, attractive interactions between nucleosomes. To determine the distribution of nucleosomes for this effective equilibrium system, we used artificial dynamics that significantly reduces the computational time needed for statistical convergence. As shown in S4 Fig, the average distance between neighboring nucleosomes indeed decreases upon the introduction of enzymes. To our surprise, however, the density profile resembles that for active genes from mouse (see Fig 4A). In contrast to the typical decreasing trend seen in Fig 3C, the height of individual peaks gradually increases as they move away from TSS. We note that the presence of the barrier potential is crucial to reveal the effect of spacer enzymes. Simulations that include spacer enzymes but without the barrier potential produced a density profile resembling the one for silent genes, i.e., the no barrier or enzyme case (S5 Fig). For simulations of genes with intermediate expression level and weakened barrier potential and enzyme remodeling rate, see S6 Fig.

Examining the simulated nucleosome arrays revealed a wide range of configurations with both well- and ill-positioned +1 nucleosome. We first ordered the nucleosome arrays along the y -axis based on the position of the first nucleosome and computed the corresponding local nucleosome density profiles. The results are shown in the left panel of Fig 4B, with representative, traditional one-dimensional profiles presented on the right. The top configurations exhibit a well-defined +1 nucleosome, and the corresponding density profile resembles that of a statistical positioning model shown in Fig 3C. For many of the configurations near the bottom of the plot, the +1 nucleosome shifts away from the TSS, giving rise to a wide nucleosome free region. A mixture of these configurations with varying +1 nucleosome positions results in the final profile shown in Fig 4A.

The inclusion of spacer enzymes can, therefore, impact both inter-nucleosome distances and the position of the +1 nucleosome. Without these enzymes, nucleosomes will occupy all accessible DNA regions while staying as far apart from each other as possible to maximize entropy. This tendency for an equal partition of the DNA is the essence of the statistical positioning model. It will ensure the confinement of the +1 nucleosome in a narrow region between the TSS and the +2 nucleosome. On the other hand, spacer enzymes introduce effective attraction between nucleosomes and cause them to aggregate rather than staying farther apart [44]. The entire array of nucleosomes now behaves as a single entity, and individual nucleosomes are no longer uniformly distributed across the genome. The free, collective movement of the entire nucleosome array with respect to the TSS, again driven by entropy, will result in ill-positioned +1 nucleosome.

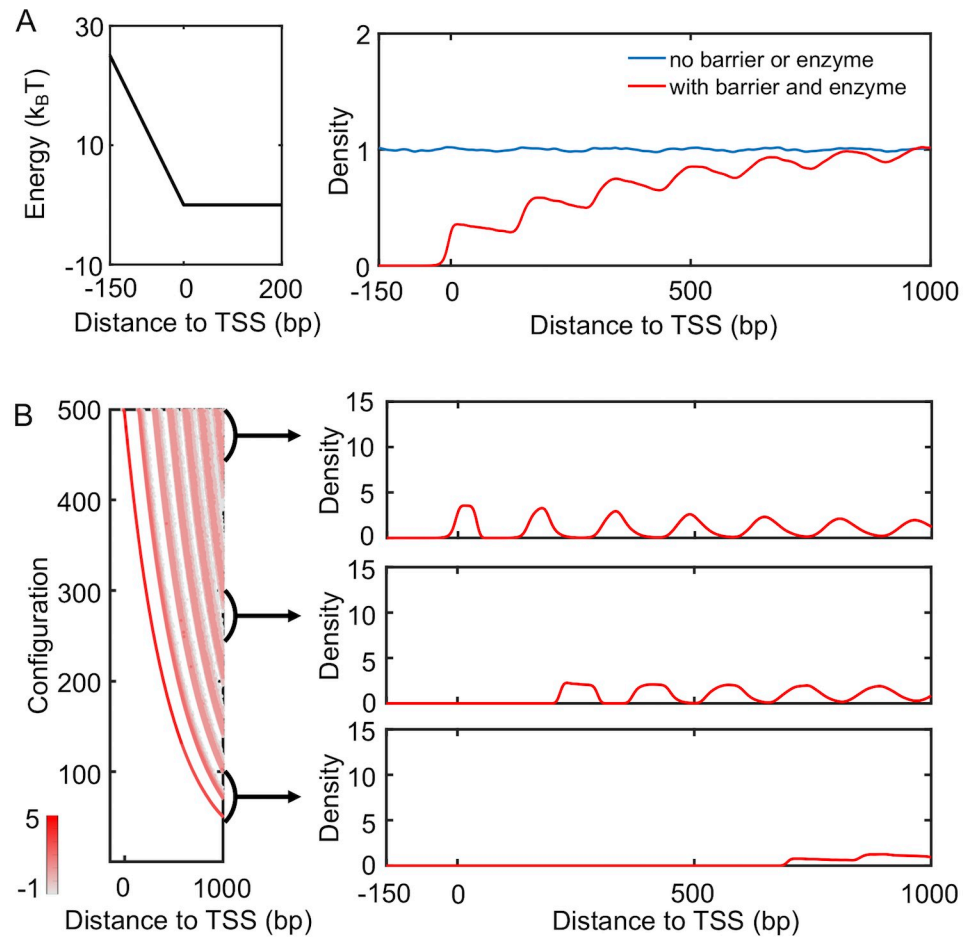


Fig 4. Spacer enzymes drive nucleosome condensation and the formation of ill-positioned +1 nucleosome. (A) Simulated nucleosome density profile using a model that includes a promoter potential (left) and spacer enzymes reproduce results for mouse active genes. The no barrier result is identical to that shown in Fig 3C. (B) Nucleosome configurations exhibit a wide distribution of positions for the +1 nucleosome. A scatter plot for the density distribution of all simulated configurations ordered by the position of the first nucleosome is shown on the left. Each row was computed via averaging over 1000 independent configurations. The right panel presents three example one-dimensional profiles in which the +1 nucleosome gradually shifts away from the TSS.

<https://doi.org/10.1371/journal.pcbi.1008556.g004>

A mixture of profiles with well- and ill-positioned +1 nucleosome reproduces yeast results

The presence of a promoter potential and spacer enzymes leads to the formation of two types of nucleosome density profiles with well- and ill-positioned +1 nucleosome. A mixture of the two types qualitatively reproduces the experimental results for active mouse genes. We next investigated whether the same mixture but with different levels of population can explain yeast nucleosome density profiles.

The more pronounced patterns seen in yeast profiles suggest that configurations with well-positioned +1 nucleosome should dominate. We note that many positioning enzymes are known to align nucleosomes toward the TSS [37]. To mimic the impact of these molecules, we introduced an additional positioning potential between the region of 0 and 30 bp from TSS. This approximate treatment avoids making explicit assumptions regarding the molecular mechanisms of remodeling enzymes that remain largely unknown.

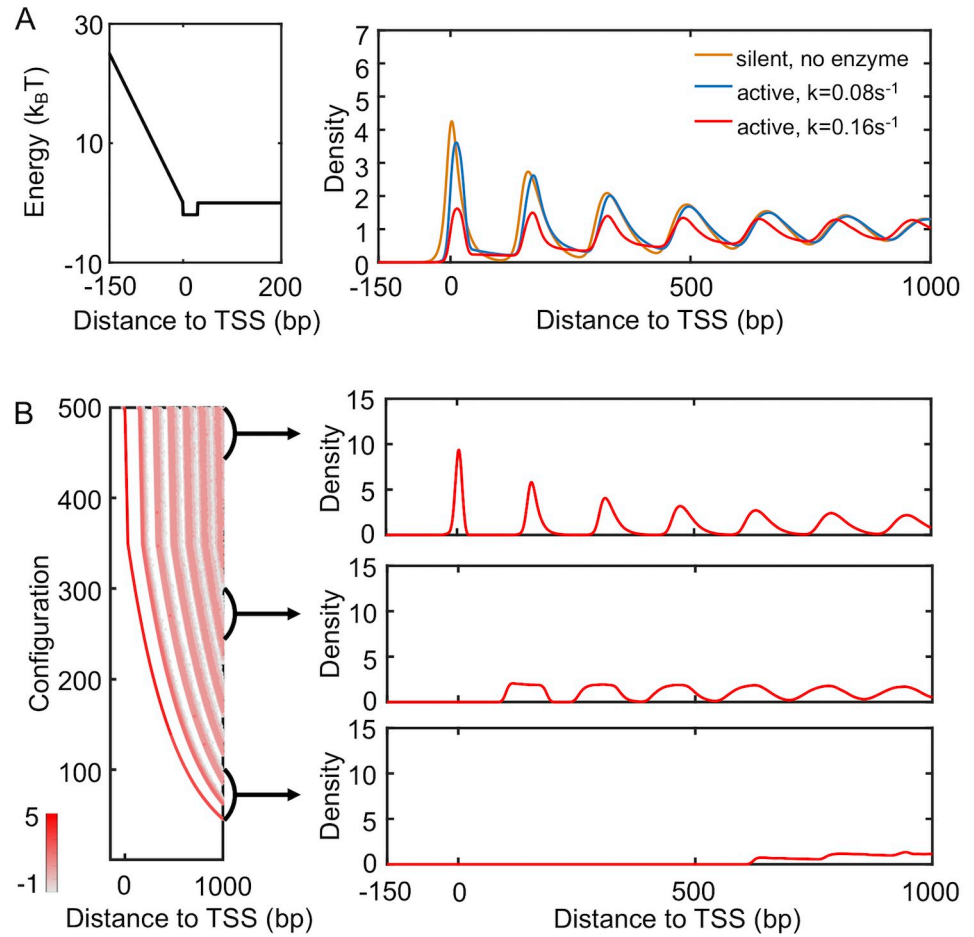


Fig 5. A mixture of configurations with well- and ill-positioned +1 nucleosome reproduce yeast density profiles. (A) Nucleosome density profiles determined with the presence of a positioning potential of $-2k_B T$ introduced to the region following the promoter (left). Comparison between the two density profiles with varying rates for spacer enzymes confirms that more active genes with higher enzyme rates exhibit lower peaks. The yellow curve from Fig 3C that corresponds to the profile of yeast silent genes is shown for reference. (B) Scatter plot for the density distribution of all simulated configurations ordered by the position of the +1 nucleosome (left), with example one-dimensional profiles shown on the right. Each row was computed via averaging over 1000 independent configurations.

<https://doi.org/10.1371/journal.pcbi.1008556.g005>

As shown in Fig 5, the new potential succeeds in attracting nucleosomes to the TSS. The resulting density profile (red) now resembles those from yeast and the height of the peaks decreases as they move away from the TSS. We note that as transcription activity decreases, enzymes will be recruited less to the genes, and their effective remodeling rate will be smaller. Slower spacer enzymes with a rate of $k = 0.08s^{-1}$ compete less effectively with the positioning enzymes, and the relative population of configurations with ill-positioned +1 nucleosome decreases. Correspondingly, the nucleosome profile exhibits higher peaks (blue) than that for more active genes with a faster enzyme rate, consistent with the dependence on transcription activity seen in experimental results (Fig 1A).

Introducing the positioning potential also breaks the symmetry across the TSS (S7 Fig). The nucleosome density profile on its left hand side is less prominent as the one on the right and shown in Fig 5A. This asymmetry is qualitatively consistent with experimental observations. It is worth mentioning that the yeast results shown in Fig 1A deviate from the simulated profiles

in that the second peak is higher than the first. However, this same trend was not observed in recent single-molecule studies that simultaneously measure the position of multiple nucleosomes from the same DNA [45, 46]. Additional experiments are needed to evaluate the robustness and mechanism of this particular feature further.

Fast histone exchange abolishes the impact of spacer enzymes

For highly transcribed genes, in addition to remodeling enzymes, Pol II could impact the positioning of nucleosomes as well. As it elongates along the DNA, Pol II could cause partial or complete loss of histone proteins [28, 47–49]. In the following, we investigate the impact of Pol II induced histone eviction on nucleosome density profiles.

Specifically, we carried out stochastic simulations that explicitly model nucleosome diffusion, enzyme remodeling, and histone eviction and absorption. We assumed that the eviction and absorption rates depend only on inter-nucleosome and nucleosome-DNA interactions and are independent of remodeling enzymes (see [Materials and methods](#)). The ratio of the two was tuned to ensure a density of approximately 0.88. The basal rate constant $r_{\text{const}} = 0.1\text{s}^{-1}$ was estimated from the transcription rate 1min^{-1} for the most active genes [50] with the assumption of full eviction for all nucleosomes.

The resulting nucleosome density profile is shown in [Fig 6](#) (purple). It differs significantly from the one obtained from a kinetic model that only included remodeling enzymes and a promoter potential (red), which was also shown in [Fig 4A](#). The impact of remodeling enzymes in these simulations is significantly reduced. In particular, a pronounced peak emerges near TSS, and the density profile now traces well the result from a model with only the promoter potential (blue). The decrease of inter-nucleosome spacing cannot be observed in the radial distribution profile either ([S8 Fig](#)). We found that the impact of histone eviction on the density profile depends on its rate and gradually diminishes as the rate slows down ([S9 Fig](#)). It is worth noting that a significantly smaller rate (10^{-8}s^{-1}) is needed to reveal the impact of remodeling enzymes on nucleosome spacing. Since a decrease of inter-nucleosome spacing is readily seen in experimental nucleosome density profiles, we anticipate that complete eviction of histone octamers to be rare [28, 51–53].

The competition between Pol II and spacer enzymes on positioning nucleosomes can be understood as following. The diffusive dynamics driven by thermal motions defines an equilibrium distribution of nucleosomes along the DNA. This distribution depends both on inter-nucleosome and nucleosome-DNA interactions. Spacer enzymes modify this distribution by introducing an effective attractive potential between nucleosomes. The two-dimensional dynamics of histone exchange can give rise to, yet, another steady-state distribution. Unless slowed down substantially, histone exchange can lead to faster relaxation kinetics when compared with nucleosome movements restricted to one dimension. It will essentially overwrite any impact caused by spacer enzymes or diffusion on nucleosome distribution. If the rates for histone eviction and adsorption satisfy detailed balance with regard to the normal potential for nucleosome-DNA and inter-nucleosome interactions, the steady-state distribution determined from histone exchange kinetics should be consistent to the equilibrium distribution obtained from pure diffusion. This consistency explains the agreement between purple and blue lines seen in [Fig 6A](#). On the other hand, if the two rates were modified to account for the effective interaction potential induced by spacer enzymes, the steady-state distribution will reproduce the one dictated by spacer enzymes. Such kinetics, though less meaningful biologically, can prove beneficial for reducing the computational cost of stochastic simulations (see [Materials and methods](#) for more discussions).

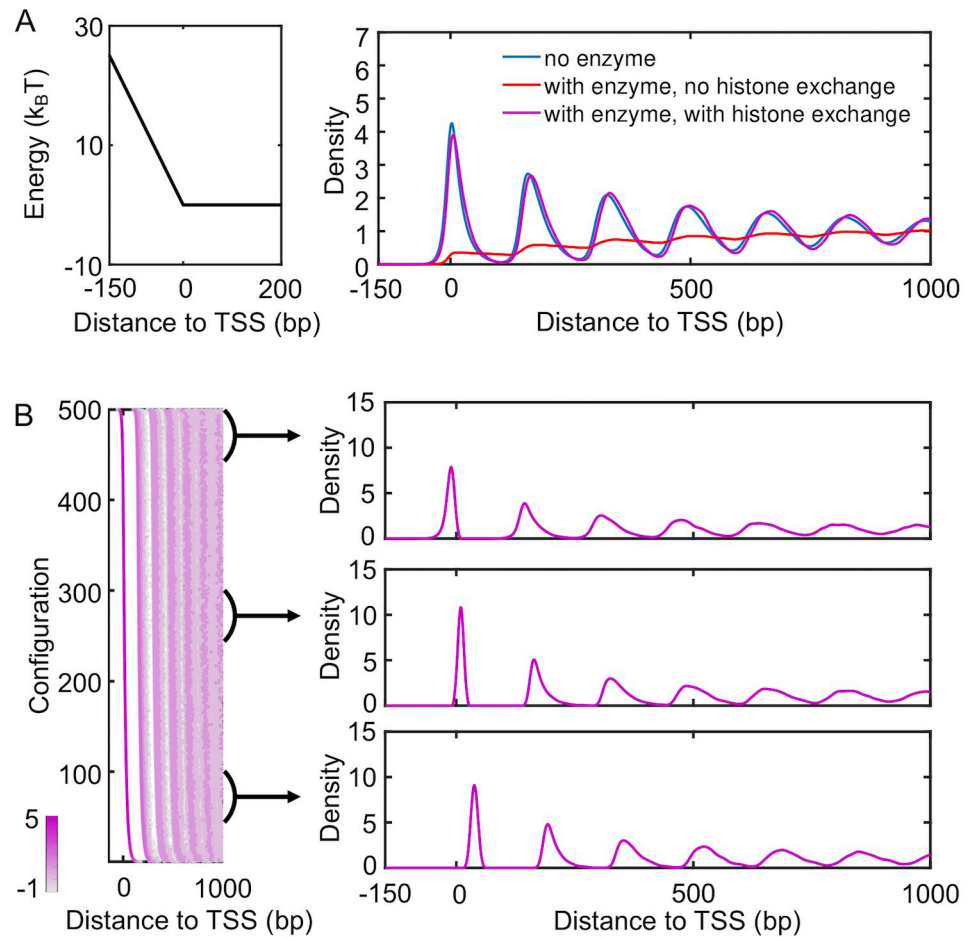


Fig 6. Impact of histone exchange kinetics on nucleosome density profiles. (A) Nucleosome density profiles determined from kinetic models that only includes a promoter potential (blue), that includes both a promoter potential and spacer enzymes (red), and that includes a promoter potential, spacer enzymes and histone exchange (purple). The red curve is identical to the one shown in Fig 4A. An illustration of the promoter potential is shown on the left. (B) Scatter plot for the density distribution of all simulated configurations ordered by the position of the +1 nucleosome (left), with example one-dimensional profiles shown on the right. Each row was computed via averaging over 1000 independent configurations.

<https://doi.org/10.1371/journal.pcbi.1008556.g006>

Conclusions and discussion

In this paper, we investigated the impact of transcription on nucleosome positioning. By partitioning genes based on their transcriptional activity, we determined the corresponding nucleosome density profiles for both yeast and mouse. A striking difference for inactive genes was observed between the two species. Similar featureless profiles as that from mouse have been observed for inactive genes from *Drosophila* [32] and human [33] as well. Analyzing the nucleosome binding affinity of DNA sequences suggests that while yeast promoters are nucleosome repelling, the opposite holds true for mouse promoters. This difference could contribute to the formation of phased nucleosome arrays in yeast, but not mouse ESC, via the statistical positioning mechanism. The nucleosome attracting promoters appear to be a rule of multi-cellular organisms rather than an exception in mouse, as shown in prior studies [23, 54, 55]. They might function to suppress the expression of certain genes crucial for cell differentiation [55].

We further carried out stochastic simulations to study the variation of nucleosome density profiles as the transcriptional activity elevates. Focusing on qualitative trends rather than quantitative agreements allowed us to extract a minimalist model of nucleosome positioning. We discovered that a tug-of-war between two types of enzymes is the key to rationalize the observed trends. In particular, enzymes that use a pair of nucleosomes as substrate can induce nucleosome condensation and tend to shift the nucleosome array away from TSS, giving rise to density profiles with ill-positioned +1 nucleosome. Positioning enzymes, on the other hand, can counteract this effect and align the +1 nucleosome back to TSS. A combination of density profiles with well- and ill-positioned +1 nucleosome can qualitatively reproduce *in vivo* results from both yeast and mouse ESC.

Possible candidates for spacer enzymes include Chd1 and ISW1 and INO80 may function as positioning enzymes. Density profiles determined from nucleosomes reconstituted *in vitro* with and without these enzymes indeed appear to consistent with our model predictions [37]. However, validating the impact of these enzymes *in vivo* can be complicated by the presence of additional enzymes with redundant roles. In addition, some of the enzymes have been reported to play significant roles both in positioning the +1 nucleosome and regulating nucleosome spacing.

We note that many details on nucleosome positioning remain unknown. When designing the kinetic model, we opted for simplicity and interpretability over quantitative accuracy. In particular, we did not explicitly consider DNA sequence effect for results shown in Figs 4 and 5. To examine the robustness of our conclusions on remodeling enzymes, we performed additional simulations that explicitly incorporated sequence specific nucleosome binding energy profiles. As shown in S10 Fig, we were able to qualitatively reproduce the same trends using models that incorporate nucleosome binding affinity of *S. cerevisiae* and mouse ESC genes (see Table in S2 Table). Spacer enzymes were able to create ill-positioned +1 nucleosomes even with the presence of DNA sequence specific binding, while positioning enzymes remain successful at aligning the +1 nucleosomes towards TSS.

To evaluate the robustness of our conclusions on model parameters, we computed a phase diagram with respect to the rate of spacer enzymes and the strength of the positioning potential. As shown in Fig 7, the competition between two opposing forces that favor yeast and mouse-like density profiles is evident and can be seen across a wide range of parameters.

A direct consequence of the competition between the two types of remodeling enzymes is the heterogeneity in the position of +1 nucleosomes, as is evident in Figs 4 and 5. Also, compared to yeast, the +1 nucleosomes from mouse ESC should exhibit a larger variance. Similarly, active genes in yeast are predicted to possess less well-defined +1 nucleosome positions than the inactive ones. We note that the position of individual +1 nucleosomes cannot be determined from the bulk data analyzed in Fig 1. To validate these predictions, we analyzed single-molecule data produced by Zhao [39] and Au group [45]. As shown in S11 and S12 Figs, consistent with model predictions, there is indeed a significant variance in the +1 nucleosome position for individual genes from both species. The relative trend between mouse and yeast genes and between yeast active and inactive genes is consistent with model predictions. The single-molecule results, therefore, provide strong support for the mechanism uncovered from our simulations.

It is worth emphasizing that in addition to the competition between remodeling enzymes, other factors could contribute to the experimental results. For example, we focused our analysis on a single gene, but the results shown in Fig 1 are from averaging over many genes. In particular, individual genes could differ in the range of the promoter potential shown in Fig 4A due to the binding of a diverse set of transcription factors [38]. This variation in promoter potential could result in further heterogeneity of the +1 nucleosome and the lowering of the

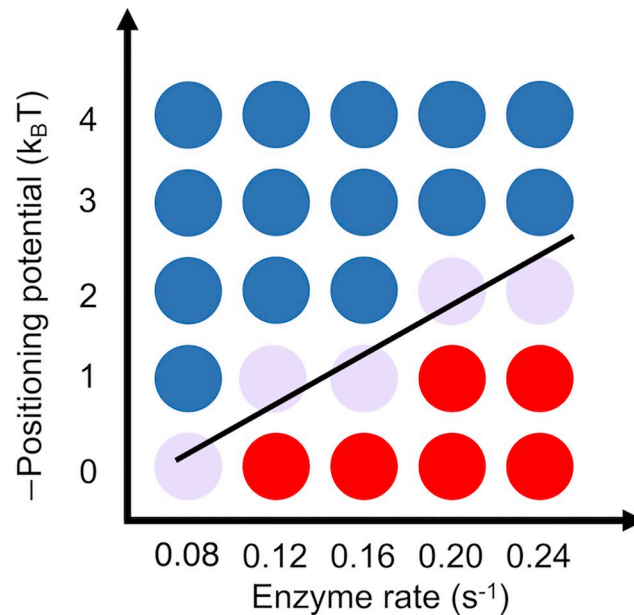


Fig 7. Phase diagram of nucleosome density profile with respect to the strength of the positioning potential and the rate of spacer enzymes. Blue and red dots correspond to yeast and mouse-like patterns, respectively. Each dot represents an independent stochastic simulation using the parameters specified in the two axes. We defined mouse-like nucleosome density profiles as the ones with a gradual increase of peak height, i.e., the 2nd peak is higher than the 1st one. Yeast-like nucleosome density profiles as the ones with a gradual decrease of peak height, i.e., the 2nd peak is lower than the 1st one. The black line connecting the purple dots represents the phase transition line.

<https://doi.org/10.1371/journal.pcbi.1008556.g007>

corresponding peak in the density profile. Accounting for such gene-specific features, though will not impact the conclusions here based on common features shared by all genes, will be crucial for predicting nucleosome positioning *in silico*.

Materials and methods

Kinetic model of nucleosome positioning

We consider a one-dimensional lattice model to study the positioning of nucleosomes along the DNA sequence (Fig 2). Each lattice site s represents a single bp and is assigned with a nucleosome binding energy V_s . In most cases, V_s was set to a constant value, such that the nucleosome binding energy of a 147 bp long DNA segment is $V_i = -42k_B T$ [56], to focus on the impact of remodeling enzymes. When the DNA sequence effect was explicitly considered, we determined V_s using the periodic function of dinucleotides introduced by van Noort and coworkers [24]. The length of the lattice is 14700 bp, and the periodic boundary condition was enforced to eliminate any end effect. Nucleosome density was set as 0.88, a typical value found near the gene coding regions in yeast [53].

To account for the excluded volume effect, a pair potential was introduced between neighboring nucleosomes i and $i + 1$ as

$$u(\Delta x_i) = \begin{cases} \text{infinity} & \Delta x < 147 \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where $\Delta x_i = x_{i+1} - x_i$, x_i corresponds to the dyad position of the i -th nucleosome, and by definition, each nucleosome occupies a region of 147 bp in length [1].

Nucleosomes can move along the DNA via diffusive motion with a rate of

$$d = \frac{D}{(\Delta l)^2} e^{-\beta \Delta U/2}, \quad (2)$$

where $D = 1\text{bp}^2/\text{s}$ is the diffusion coefficient and $\Delta l = 1\text{bp}$ is the step size [57]. ΔU denotes the change of the total energy before and after nucleosome movement. The rate expression was designed such that detailed balance is satisfied [56].

In addition to thermal motions, positions of nucleosomes can be altered by transcription related activities as well [58]. In the following, we consider the impact of three major factors related to transcription.

First, transcription factors, preinitiation complex, and Pol II are known to compete with histone proteins to bind gene promoters [38, 59]. We incorporated these proteins' effect as an energetic barrier centered at 150 bp upstream of TSS to penalize nucleosome formation. Similar treatment has been used by Padinhateeri and coworkers to create a nucleosome-free region near TSS [60]. As illustrated in Fig 2B, the barrier is symmetric with respect to the center. Its triangular shape allows nucleosomes to occupy the promoter region with a finite probability. The mathematical expression for this promoter potential is provided in the [S1 Supporting information](#).

Second, active transcription can recruit remodeling enzymes to alter the position of nucleosomes at the expense of ATP. While several types of remodeling enzymes have been discovered, here we focus on ISW1-like enzymes that modulate inter-nucleosome spacing and INO80-like enzymes that adjust the position of the +1 nucleosome [27, 37, 40, 61–64]. Following Möbius *et al.* [43], we assumed that spacer enzymes bind to neighboring nucleosomes that are within 332 bp at a rate of 0.16 s^{-1} [62, 65], and randomly move one of them toward the other by one bp (Fig 2C). For the positioning enzymes, we modeled their effect with a positioning potential located near TSS.

Finally, transcription of the gene body by Pol II could displace nucleosomes completely off the DNA [28, 47]. To account for such disrupt events, we explicitly modeled absorption and desorption of histone proteins with rate expressions

$$r_{\text{on}} = r_{\text{const}} e^{-\beta \Delta V/2}, \quad r_{\text{off}} = r_{\text{const}} e^{\beta \Delta V/2}. \quad (3)$$

r_{const} is the rate constant, and $\Delta V = \Delta U + \mu$ includes both the change in the system's total energy ΔU and a chemical potential μ . The impact of Pol II was incorporated into the rate constant and the chemical potential, whose value was tuned to ensure an average system density of 0.88 (see Table in [S1 Table](#)).

Details of stochastic simulations

We carried out stochastic simulations using the Gillespie algorithm [66] to determine steady-state nucleosome density profiles.

Simulations without Pol II facilitated histone exchange. In several of the kinetic models explored in the *Results* Section, the effect of histone eviction from Pol II was not explicitly considered. Without remodeling enzymes, these models describe systems with equilibrium statistics since the diffusive dynamics follows detailed balance (Eq 2). When remodeling enzymes are present, as shown in our previous study [44], the kinetic model can be rigorously mapped onto an effective equilibrium system with renormalized temperature and potential, detailed expressions for which are provided in the SI. For such *one-dimensional* equilibrium or quasi-equilibrium systems, there is a well defined, unique distribution for each model that depends only on inter-nucleosome potentials and DNA sequence. These distributions are independent

of the kinetic schemes used in stochastic simulations as long as they satisfy detailed balance. Therefore, for their determination, we simulated only “artificial” absorption and desorption kinetics with rates defined in Eq 3 and $r_{\text{const}} = 12s^{-1}$. Renormalized potentials were used to determine the change in the system’s total energy ΔU if remodeling enzymes were introduced in the kinetic model. The two-dimensional dynamics for histone exchange helps to alleviate the topological constraint and jamming dynamics experienced if the system is restricted to one dimension with all nucleosomes bound to the DNA. It can significantly reduce the computational time needed for convergence. In the large number limit, the statistics of the grand canonical ensemble with histone exchange should be equivalent to that of a system restricted to one dimension with fixed nucleosome number. In S13 Fig, we showed that, for the system size considered here, the fluctuation in nucleosome number is small and has minimal impact on the resulting density profile.

We carried out 200 independent simulations for kinetic models lacking DNA sequence specificity. To investigate the impact of DNA sequences, we also separately carried out 1000 simulations for both yeast and mouse. Each one of these simulations incorporates a nucleosome binding affinity profile predicted from the sequence of an inactive gene. All simulations lasted for 5000 seconds and were initialized with over 80 nucleosomes randomly distributed over the lattice. 2500 configurations were recorded every two seconds in each simulation to determine the density profiles.

Simulations with Pol II facilitated histone eviction. Pol II and remodeling enzymes can evict and assemble nucleosomes during transcription [47, 49]. This two-dimensional kinetics for histone exchange defines a steady-state distribution consistent with the reaction rates defined in Eq 3. The one-dimensional spacer enzymes, by themselves, can give rise to another steady-state distribution that depends on enzyme kinetics. If the rate expressions for histone exchange were modified to account for the effective interaction induced by enzymes, the two steady-state distributions are consistent with each other. This consistency inspired our use of artificial exchange kinetics to accelerate computer simulations mentioned above. Biologically, however, histone exchange rates most likely do not depend on spacer enzymes, and the two distributions will be in conflict.

To rigorously account for the impact of both kinetics, we performed stochastic simulations that explicitly include diffusion, enzyme remodeling, and histone eviction and absorption as well. A total of 2500 independent 5×10^5 -second-long simulations were performed. Only 200 configurations in the last 400 seconds of each simulation were collected with an equal time interval to compute the density profiles. These simulations were again initialized with randomly placed nucleosomes over the lattice.

Data processing

Genome-wide mappings of nucleosome positions obtained with a chemical mapping method are available for *S. cerevisiae* and mouse embryonic stem cells in the NCBI database with accession number GSE36063 and GSE82127. Compared to the micrococcal nuclease digestion, followed by high-throughput sequencing (MNase-seq) [29], the chemical mapping approach is affected less by sequence preference or nucleosome unwrapping and can provide base pair resolution of nucleosome center positions [31, 34]. We point out, however, that the qualitative trend shown in Fig 1A has been observed for data obtained with MNase-seq as well [30].

To determine the transcription level of individual genes, we downloaded RNA-seq data using accession number GSE52086 for yeast [67] and GSE82127 for mouse [31]. Genes with more than one promoter [31, 35] were removed from analysis. DNA sequences surrounding

TSS were extracted from the Eukaryotic Promoter Database [68, 69] based on the ID provided in the RNA-seq data.

Supporting information

S1 Supporting information. Supporting information for “On the role of transcription in positioning nucleosomes”.

(PDF)

S1 Table. Values of the chemical potential used to ensure an average density of 0.88. $V_i = -42k_B T$ is the equilibrium nucleosome binding free energy [56], and $V_i + \mu$ provides an estimation of the effective nucleosome affinity in each model.

(PDF)

S2 Table. Values of the chemical potential used to ensure an average density of 0.88 or 0.78 on simulations with DNA sequence specific nucleosome binding. \bar{V}_i is the mean nucleosome binding free energy of each gene calculated with the model introduced by van Noort and coworkers [24], and $\bar{V}_i + \mu$ provides an estimation of the effective nucleosome affinity in each model.

(PDF)

S1 Fig. The raw nucleosome density profiles shown in Fig 1 of the main text before smoothing. See text *Nucleosome density profile smoothing* in [S1 Supporting information](#) for details on the smoothing function.

(PDF)

S2 Fig. Normalized nucleosome density profiles for *S. cerevisiae* and mouse near TSS. (A) *S. cerevisiae* [34]. (B) mouse [31]. After removing genes with more than one promoter [31, 35], 4151 and 18969 genes were considered here for *S. cerevisiae* and mouse, respectively. These genes were separated into quartiles depending on their transcription activities, with the bottom and top 25% corresponding to the most inactive and active genes, respectively. For *S. cerevisiae*, using the first five peaks for each curve, we estimated the inter-nucleosome spacing as 164.75 bp for bottom 25%, 165.25 bp for bottom 25%-50%, 163.75 bp for top 25%-50%, 161.5 bp for top 25%, and 159.75 bp for top 200. For mouse, the spacing was similarly estimated as 189.75 bp for top 25%-50% and 188.5 bp for top 25%.

(PDF)

S3 Fig. Average nucleosome affinity computed for *S. cerevisiae* and mouse genes with different expression levels. (A) *S. cerevisiae*. (B) mouse. The same procedure as in [Fig 3A](#) of the main text was used in computing these profiles.

(PDF)

S4 Fig. Comparison between radial distribution profiles for kinetic models with or without spacer enzymes and average nucleosome spacing as a function of the remodeling rate of spacer enzymes. (A) Comparison between radial distribution profiles for kinetic models with (red) or without (blue) spacer enzymes. The plots were computed from the same data used in [Fig 4A](#) of the main text. It is evident that, upon the introduction of spacer enzymes, inter-nucleosome distances decrease. See text: *Definition of the density profile and radial distribution function* in [S1 Supporting information](#) for more details. (B) Average nucleosome spacing as a function of the remodeling rate of spacer enzymes. The spacing was determined using the average distance between neighboring nucleosomes. The nucleosome pair flanking the TSS

was excluded from distance estimation.
(PDF)

S5 Fig. Nucleosome density profile obtained from simulations that include spacer enzymes but not the barrier potential. Nucleosome density profile obtained from simulations that include spacer enzymes but not the barrier potential (blue), which was obtained by averaging 500 independent 5000-second-long simulations to achieve a better ensemble average under the strong clustering effect of spacer enzymes. For reference, the profile determined with the presence of both the barrier potential and spacer enzymes (red), i.e., the red curve in Fig 4A of the main text, is also shown.
(PDF)

S6 Fig. Simulated nucleosome density profile for mouse genes with intermediate expression level (purple). For reference, the corresponding profile of silent (blue) and active (red) genes, i.e., the two curves in Fig 4A of the main text, are also shown. We note that unlike yeast, mouse inactive genes differ from the active genes in two ways. First, mouse inactive genes do not have a well defined nucleosome free region nor barrier potential. Second, mouse inactive genes lack the remodeling enzymes that are known to be associated with transcription. This is the key reason that we used “no barrier or enzyme” in Fig 4A. Yeast inactive genes, on the other hand, do have a well defined nucleosome free region and barrier potential as evidenced in Fig 1A. Because of the presence of differences in both barrier potential and remodeling enzymes between mouse active and inactive genes, studying “middle” genes in mouse requires tuning two parameters rather than one single parameter as in yeast. In the simulations used for intermediate genes, we decreased the height of the barrier potential from 25 to $2k_B T$ and reduced the rate of spacer enzymes from 0.16 to $0.12 s^{-1}$. Other simulation setups were kept the same as for active genes.
(PDF)

S7 Fig. The nucleosome density profile across both sides of the TSS exhibits asymmetric patterns. The same data for the blue curve in Fig 5A of the main text was used here to determine the density profile.
(PDF)

S8 Fig. Fast histone change kinetics abolishes the impact of remodeling enzymes on inter-nucleosome spacing. To measure the inter-nucleosome spacing, we determined the radial distribution profiles from the same data used to compute the nucleosome density profiles shown in Fig 6A of the main text. The result from a kinetic model with spacer enzymes and histone exchange (purple) is almost identical to that from a model with no enzymes (blue) and does not exhibit any decrease in inter-nucleosome spacing.
(PDF)

S9 Fig. Dependence of nucleosome density profile and inter-nucleosome spacing on the rate of histone exchange kinetics that was varied from $10^{-2} s^{-1}$, to $10^{-6} s^{-1}$ and to $10^{-8} s^{-1}$. (A) $10^{-2} s^{-1}$. (B) $10^{-6} s^{-1}$. (C) $10^{-8} s^{-1}$. The radial distribution profiles for the kinetic models without enzyme and with enzyme but no histone exchange are identical to those shown in S8 Fig.
(PDF)

S10 Fig. Nucleosome density profiles determined using DNA sequence-specific binding energy support the robustness of conclusions presented in the main text. (A) The corresponding plot to Fig 3C of the main text. The blue line was computed without the barrier potential and is identical to the mouse result shown in Fig 3B of the main text. We used the

nucleosome binding profiles of mouse genes in the corresponding simulations. The red line was obtained from simulations with the barrier potential and the nucleosome binding profiles of *S. cerevisiae* genes. (B,C) The corresponding plots to Fig 4A of the main text. Simulations performed for part B used the same set up as those in the main text, except with the addition of nucleosome binding profiles computed using mouse genes. Simulations performed for part C uses a shifted promoter potential and a nucleosome density of 0.78. Both of these two changes are supported by the experimental data shown in Fig 1 of the main text. (D) The corresponding plot to Fig 5A of the main text. Simulations were performed using the same set up as those in the main text, except with the addition of nucleosome binding profiles computed using yeast genes. See text *Simulations with DNA sequence specific nucleosome binding* in S1 Supporting information for simulation details.

(PDF)

S11 Fig. Single-cell data supports the heterogeneous distribution of the +1 nucleosome position in mouse ESC. The single-cell nucleosome mapping data for mouse ESC was obtained from the NCBI database with accession number GSE96688 [39]. The +1 nucleosome was defined as the first nucleosome in the downstream 1000 bp region of TSS. Only mapped nucleosomal fragments with a length of 140-180 bp were retained, and the centers of fragments were considered as the nucleosome dyad position. We further discarded segments for which the first nucleosome is more than 1000 bp away from the TSS. This treatment left out 121 genes (18,969 in total) for which no nucleosomes were found in all samples within 1000 bp of TSS. (A) Genes exhibit large variance in the position of the +1 nucleosomes. Each row represents an individual gene, and they were sorted based on the mean distance between +1 nucleosome and TSS. The solid line correspond to the mean position and the light blue errorbars represent standard deviations. (B) Probability distribution of the variance shown in part A. (C) Probability distribution of distance between +1 nucleosome position of a given gene $x[n]$ and the smallest position of that gene found across all cells $x[1]$. We note that a potential caveat of single cell experiments is that they might suffer from sparse data coverage. Failing to detect some nucleosomes could misguide data analysis, causing the assignment of +2 or +3 as +1 nucleosome and an exaggeration of the variance seen in part A. Though it would be difficult to rule out artifacts from missing data, we can gauge their impact by considering extreme cases. In particular, we consider the case that all nucleosome arrays in different cells have identical +1 nucleosome positions, and the variance arises purely from missing data and misassignment. In this case, the probability distribution of $x[n] - x[1]$ should exhibit clear peaks separated by nucleosome repeat length. As shown in part C, no significant peaks can be seen in the distribution. Therefore, the +1 nucleosome position variance should not arise purely as a result of data artifact, but reflects the true heterogeneity across cells.

(PDF)

S12 Fig. Single molecule data supports the heterogeneous distribution of the +1 nucleosome position in *S. cerevisiae*. Data produced from the Au group [45] were used for the analysis. The +1 nucleosome was defined as the first nucleosome found in the [-50, 1000] bp region around TSS. We included 50 bp upstream of the TSS since the +1 nucleosomes peak in yeast is closer to the TSS than mouse ESC, as is evident in Fig 1 of the main text. (A) Variance of +1 nucleosome position across different genes. Each row represents an individual gene, and they were sorted based on the mean distance between +1 nucleosome and TSS. The solid line correspond to the mean position and the light blue errorbars represent standard deviations. (B,C) Probability distribution of the +1 nucleosome position for various group of genes.

(PDF)

S13 Fig. Accuracy of the simulation algorithm with “artificial” kinetics for histone absorption/desorption. (A) Probability distribution for the number of nucleosomes bound to the DNA computed from the simulation data used for creating Fig 4A (red) of the main text. Over 90% of the simulated configurations falls into the ± 10 interval of the peak nucleosome number 86. We note that the length of free DNA region is only 98 nucleosomes due to the presence of the barrier potential in the promoter region. The peak of the distribution with 86 nucleosomes, therefore, provides a density of approximately 0.88. (B) Comparison between nucleosome density profiles obtained from artificial kinetics (red) and from rigorous stochastic simulations of diffusion and enzyme remodeling (blue). In these additional simulations, no explicit histone exchange was considered, and the 86 nucleosomes remain bound to the one-dimensional lattice at all time. We performed 500 simulations that lasted for 7.5×10^6 seconds and were initialized from randomly distributed nucleosomes. 1000 configurations were extracted from the last 5×10^5 s of each trajectory at every 500 seconds to compute the density profile. The good agreement between the two density profiles provide strong support for the accuracy of the algorithm with artificial kinetics used in the main text. (PDF)

Acknowledgments

The authors thank Dr. Kin Fai Au and Yunhao Wang for sharing single-molecule long read sequencing data of nucleosomes, and Dr. van Noort for sharing the software on computing nucleosome affinity from DNA sequences.

Author Contributions

Conceptualization: Zhongling Jiang, Bin Zhang.

Data curation: Zhongling Jiang, Bin Zhang.

Formal analysis: Zhongling Jiang, Bin Zhang.

Investigation: Zhongling Jiang, Bin Zhang.

Methodology: Zhongling Jiang, Bin Zhang.

Project administration: Bin Zhang.

Resources: Zhongling Jiang, Bin Zhang.

Software: Zhongling Jiang, Bin Zhang.

Supervision: Bin Zhang.

Validation: Zhongling Jiang, Bin Zhang.

Visualization: Zhongling Jiang, Bin Zhang.

Writing – original draft: Zhongling Jiang, Bin Zhang.

Writing – review & editing: Zhongling Jiang, Bin Zhang.

References

1. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997; 389(6648):251–260. <https://doi.org/10.1038/38444> PMID: 9305837
2. Bai L, Morozov AV. Gene regulation by nucleosome positioning. *Trends Genet*. 2010; 26(11):476–483. <https://doi.org/10.1016/j.tig.2010.08.003> PMID: 20832136

3. Voong LN, Xi L, Wang JP, Wang X. Genome-wide mapping of the nucleosome landscape by micrococcal nuclease and chemical mapping. *Trends Genet.* 2017; 33(8):495–507. <https://doi.org/10.1016/j.tig.2017.05.007> PMID: 28693826
4. Schiessel H, Widom J, Bruinsma RF, Gelbart WM. Polymer reptation and nucleosome repositioning. *Phys Rev Lett.* 2001; 86(19):4414–4417. <https://doi.org/10.1103/PhysRevLett.86.4414> PMID: 11328188
5. Möbius W, Neher RA, Gerland U. Kinetic accessibility of buried DNA sites in nucleosomes. *Phys Rev Lett.* 2006; 97(20):208102. <https://doi.org/10.1103/PhysRevLett.97.208102> PMID: 17155720
6. Jiang C, Pugh BF. Nucleosome positioning and gene regulation: Advances through genomics. *Nat Rev Genet.* 2009; 10(3):161–172. <https://doi.org/10.1038/nrg2522> PMID: 19204718
7. Freeman GS, Lequieu JP, Hinckley DM, Whitmer JK, de Pablo JJ. DNA shape dominates sequence affinity in nucleosome formation. *Phys Rev Lett.* 2014; 113(16):168101. <https://doi.org/10.1103/PhysRevLett.113.168101> PMID: 25361282
8. Parsons T, Zhang B. Critical role of histone tail entropy in nucleosome unwinding. *J Chem Phys.* 2019; 150(18):185103. <https://doi.org/10.1063/1.5085663> PMID: 31091895
9. Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, et al. A genomic code for nucleosome positioning. *Nature.* 2006; 442(7104):772–778. <https://doi.org/10.1038/nature04979> PMID: 16862119
10. Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, et al. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* 2008; 18(7):1073–1083. <https://doi.org/10.1101/gr.078261.108> PMID: 18550805
11. Grigoryev SA, Woodcock CL. Chromatin organization—The 30nm fiber. *Exp Cell Res.* 2012; 318(12):1448–1455. <https://doi.org/10.1016/j.yexcr.2012.02.014> PMID: 22394510
12. Grigoryev SA. Nucleosome spacing and chromatin higher-order folding. *Nucleus.* 2012; 3(6):493–499. <https://doi.org/10.4161/nucl.22168> PMID: 22990522
13. Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. *Cell Res.* 2012; 22(3):490–503. <https://doi.org/10.1038/cr.2012.15> PMID: 22270183
14. Gilbert N. Biophysical regulation of local chromatin structure. *Curr Opin Genet Dev.* 2019; 55:66–75. <https://doi.org/10.1016/j.gde.2019.06.001> PMID: 31306886
15. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012; 485(7398):376–380. <https://doi.org/10.1038/nature11082> PMID: 22495300
16. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature.* 2012; 485(7398):381–385. <https://doi.org/10.1038/nature11049> PMID: 22495304
17. Schoenfelder S, Fraser P. Long-range enhancer-promoter contacts in gene expression control. *Nat Rev Genet.* 2019; 20(8):437–455. <https://doi.org/10.1038/s41576-019-0128-0> PMID: 31086298
18. Hnisz D, Day DS, Young RA. Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell.* 2016; 167(5):1188–1200. <https://doi.org/10.1016/j.cell.2016.10.024> PMID: 27863240
19. Widom J. Role of DNA sequence in nucleosome stability and dynamics. *Q Rev Biophys.* 2001; 34(3):269–324. <https://doi.org/10.1017/S0033583501003699> PMID: 11838235
20. Satchwell SC, Drew HR, Travers AA. Sequence periodicities in chicken nucleosome core DNA. *J Mol Biol.* 1986; 191(4):659–675. [https://doi.org/10.1016/0022-2836\(86\)90452-3](https://doi.org/10.1016/0022-2836(86)90452-3) PMID: 3806678
21. Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, et al. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature.* 2007; 446(7135):572–576. <https://doi.org/10.1038/nature05632> PMID: 17392789
22. Segal E, Widom J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr Opin Struct Biol.* 2009; 19(1):65–71. <https://doi.org/10.1016/j.sbi.2009.01.004> PMID: 19208466
23. Tompitak M, Vaillant C, Schiessel H. Genomes of multicellular organisms have evolved to attract nucleosomes to promoter regions. *Biophys J.* 2017; 112(3):505–511. <https://doi.org/10.1016/j.bpj.2016.12.041> PMID: 28131316
24. van der Heijden T, van Vugt JJFA, Logie C, van Noort J. Sequence-based prediction of single nucleosome positioning and genome-wide nucleosome occupancy. *Proc Natl Acad Sci.* 2012; 109(38):E2514–E2522. <https://doi.org/10.1073/pnas.1205659109> PMID: 22908247
25. Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, et al. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature.* 2009; 458(7236):362–366. <https://doi.org/10.1038/nature07667> PMID: 19092803

26. Segal E, Widom J. What controls nucleosome positions? *Trends Genet.* 2009; 25(8):335–343. <https://doi.org/10.1016/j.tig.2009.06.002> PMID: 19596482
27. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol.* 2017; 18(7):407–422. <https://doi.org/10.1038/nrm.2017.26> PMID: 28512350
28. Venkatesh S, Workman JL. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol.* 2015; 16(3):178–189. <https://doi.org/10.1038/nrm3941> PMID: 25650798
29. Chereji RV, Clark DJ. Major determinants of nucleosome positioning. *Biophys J.* 2018; 114(10):2279–2289. <https://doi.org/10.1016/j.bpj.2018.03.015> PMID: 29628211
30. Chereji RV, Morozov AV. Functional roles of nucleosome stability and dynamics. *Brief Funct Genomics.* 2014; 14(1):50–60. <https://doi.org/10.1093/bfgp/elu038> PMID: 25275099
31. Voong LN, Xi L, Sebeson AC, Xiong B, Wang JP, Wang X. Insights into nucleosome organization in mouse embryonic stem cells through chemical mapping. *Cell.* 2016; 167(6):1555–1570. <https://doi.org/10.1016/j.cell.2016.10.049> PMID: 27889238
32. Baldi S, Krebs S, Blum H, Becker PB. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. *Nat Struct Mol Biol.* 2018; 25(9):894–901. <https://doi.org/10.1038/s41594-018-0110-0> PMID: 30127356
33. Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A. Determinants of nucleosome organization in primary human cells. *Nature.* 2011; 474(7352):516–522. <https://doi.org/10.1038/nature10002> PMID: 21602827
34. Brogaard K, Xi L, Wang JP, Widom J. A map of nucleosome positions in yeast at base-pair resolution. *Nature.* 2012; 486(7404):496–501. <https://doi.org/10.1038/nature11142> PMID: 22722846
35. Engel SR, Dietrich FS, Fisk DG, Binkley G, Balakrishnan R, Costanzo MC, et al. The Reference Genome Sequence of *Saccharomyces cerevisiae*: Then and Now. *G3: Genes, Genomes, Genetics.* 2014; 4(3):389–398. <https://doi.org/10.1534/g3.113.008995>
36. Kornberg RD, Stryer L. Statistical distributions of nucleosomes: Nonrandom locations by a stochastic mechanism. *Nucleic Acids Res.* 1988; 16(14):6677–6690. <https://doi.org/10.1093/nar/16.14.6677> PMID: 3399412
37. Krietenstein N, Wal M, Watanabe S, Park B, Peterson CL, Pugh BF, et al. Genomic nucleosome organization reconstituted with pure proteins. *Cell.* 2016; 167(3):709–721. <https://doi.org/10.1016/j.cell.2016.09.045> PMID: 27768892
38. Rhee HS, Pugh BF. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature.* 2012; 483(7389):295–301. <https://doi.org/10.1038/nature10799> PMID: 22258509
39. Lai B, Gao W, Cui K, Xie W, Tang Q, Jin W, et al. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. *Nature.* 2018; 562(7726):281–285. <https://doi.org/10.1038/s41586-018-0567-3> PMID: 30258225
40. Ocampo J, Chereji RV, Eriksson PR, Clark DJ. The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. *Nucleic Acids Research.* 2016; 44(10):4625–4635. <https://doi.org/10.1093/nar/gkw068> PMID: 26861626
41. Ocampo J, Chereji RV, Eriksson PR, Clark DJ. Contrasting roles of the RSC and ISW1/CHD1 chromatin remodelers in RNA polymerase II elongation and termination. *Genome research.* 2019; 29(3):407–417. <https://doi.org/10.1101/gr.242032.118> PMID: 30683752
42. Florescu AM, Schiessel H, Blossey R. Kinetic control of nucleosome displacement by ISW1/ACF chromatin remodelers. *Phys Rev Lett.* 2012; 109(11):118103. <https://doi.org/10.1103/PhysRevLett.109.118103> PMID: 23005680
43. Mobius W, Osberg B, Tsankov AM, Rando OJ, Gerland U. Toward a unified physical model of nucleosome patterns flanking transcription start sites. *Proc Natl Acad Sci.* 2013; 110(14):5719–5724. <https://doi.org/10.1073/pnas.1214048110> PMID: 23509245
44. Jiang Z, Zhang B. Theory of active chromatin remodeling. *Phys Rev Lett.* 2019; 123(20):208102. <https://doi.org/10.1103/PhysRevLett.123.208102> PMID: 31809105
45. Wang Y, Wang A, Liu Z, Thurman AL, Powers LS, Zou M, et al. Single-molecule long-read sequencing reveals the chromatin basis of gene expression. *Genome Res.* 2019; 29(8):1329–1342. <https://doi.org/10.1101/gr.251116.119> PMID: 31201211
46. Shipony Z, Marinov GK, Swaffer MP, Sinnott-Armstrong NA, Skotheim JM, Kundaje A, et al. Long-range single-molecule mapping of chromatin accessibility in eukaryotes. *Nat Methods.* 2020; 17(3):319–327. <https://doi.org/10.1038/s41592-019-0730-2> PMID: 32042188
47. Kulaeva OI, Hsieh FK, Chang HW, Luse DS, Studitsky VM. Mechanism of transcription through a nucleosome by RNA polymerase II. *BBA—Gene Regul Mech.* 2013; 1829(1):76–83.

48. Teves SS, Weber CM, Henikoff S. Transcribing through the nucleosome. *Trends Biochem Sci.* 2014; 39(12):577–586. <https://doi.org/10.1016/j.tibs.2014.10.004> PMID: 25455758
49. Kulaeva OI, Hsieh FK, Studitsky VM. RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. *Proc Natl Acad Sci.* 2010; 107(25):11325–11330. <https://doi.org/10.1073/pnas.1001148107> PMID: 20534568
50. Pelechano V, Chávez S, Pérez-Ortín JE. A Complete Set of Nascent Transcription Rates for Yeast Genes. *PLoS ONE.* 2010; 5(11):e15442. <https://doi.org/10.1371/journal.pone.0015442> PMID: 21103382
51. Zhao J, Herrera-Diaz J, Gross DS. Domain-Wide displacement of histones by activated heat shock factor occurs independently of Swi/Snf and is not correlated with RNA polymerase II density. *Mol Cell Biol.* 2005; 25(20):8985–8999. <https://doi.org/10.1128/MCB.25.20.8985-8999.2005> PMID: 16199876
52. Kuryan BG, Kim J, Tran NNH, Lombardo SR, Venkatesh S, Workman JL, et al. Histone density is maintained during transcription mediated by the chromatin remodeler RSC and histone chaperone NAP1 in vitro. *Proc Natl Acad Sci.* 2012; 109(6):1931–1936. <https://doi.org/10.1073/pnas.1109994109> PMID: 22308335
53. Oberbeckmann E, Wolff M, Krietenstein N, Heron M, Ellins JL, Schmid A, et al. Absolute nucleosome occupancy map for the *Saccharomyces cerevisiae* genome. *Genome Res.* 2019; 29(12):1996–2009. <https://doi.org/10.1101/gr.253419.119> PMID: 31694866
54. Vavouri T, Lehner B. Chromatin organization in sperm may be the major functional consequence of base composition variation in the human genome. *PLoS Genet.* 2011; 7(4):e1002036. <https://doi.org/10.1371/journal.pgen.1002036> PMID: 21490963
55. Tillo D, Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Field Y, et al. High nucleosome occupancy is encoded at human regulatory sequences. *PLoS ONE.* 2010; 5(2):e9129. <https://doi.org/10.1371/journal.pone.0009129> PMID: 20161746
56. Padinhateeri R, Marko JF. Nucleosome positioning in a model of active chromatin remodeling enzymes. *Proc Natl Acad Sci.* 2011; 108(19):7799–7803. <https://doi.org/10.1073/pnas.1015206108> PMID: 21518900
57. Ranjith P, Yan J, Marko JF. Nucleosome hopping and sliding kinetics determined from dynamics of single chromatin fibers in *Xenopus* egg extracts. *Proc Natl Acad Sci.* 2007; 104(34):13649–13654. <https://doi.org/10.1073/pnas.0701459104> PMID: 17698962
58. Struhl K, Segal E. Determinants of nucleosome positioning. *Nat Struct Mol Biol.* 2013; 20(3):267–273. <https://doi.org/10.1038/nsmb.2506> PMID: 23463311
59. Pugh BF, Venters BJ. Genomic organization of human transcription initiation complexes. *PLoS ONE.* 2016; 11(2):e0149339. <https://doi.org/10.1371/journal.pone.0149339> PMID: 26866362
60. Parmar JJ, Marko JF, Padinhateeri R. Nucleosome positioning and kinetics near transcription-start-site barriers are controlled by interplay between active remodeling and DNA sequence. *Nucleic Acids Res.* 2014; 42(1):128–136. <https://doi.org/10.1093/nar/gkt854> PMID: 24068556
61. Blosser TR, Yang JG, Stone MD, Narlikar GJ, Zhuang X. Dynamics of nucleosome remodelling by individual ACF complexes. *Nature.* 2009; 462(7276):1022–1027. <https://doi.org/10.1038/nature08627> PMID: 20033040
62. Deindl S, Hwang WL, Hota SK, Blosser TR, Prasad P, Bartholomew B, et al. ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. *Cell.* 2013; 152(3):442–452. <https://doi.org/10.1016/j.cell.2012.12.040> PMID: 23374341
63. Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell.* 2013; 154(3):490–503. <https://doi.org/10.1016/j.cell.2013.07.011> PMID: 23911317
64. Yen K, Vinayachandran V, Batta K, Koerber RT, Pugh BF. Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell.* 2012; 149(7):1461–1473. <https://doi.org/10.1016/j.cell.2012.04.036> PMID: 22726434
65. Qiu Y, Levandosky RF, Chakravarthy S, Patel A, Bowman GD, Myong S. The Chd1 chromatin remodeler shifts nucleosomal DNA bidirectionally as a monomer. *Mol Cell.* 2017; 68(1):76–88. <https://doi.org/10.1016/j.molcel.2017.08.018> PMID: 28943314
66. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *J Phys Chem.* 1977; 81(25):2340–2361. <https://doi.org/10.1021/j100540a008>
67. Martín GM, King DA, Green EM, Garcia-Nieto PE, Alexander R, Collins SR, et al. Set5 and Set1 cooperate to repress gene expression at telomeres and retrotransposons. *Epigenetics.* 2014; 9(4):513–522. <https://doi.org/10.4161/epi.27645> PMID: 24442241

68. Dreos R, Ambrosini G, Groux R, Périer RC, Bucher P. The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. *Nucleic Acids Res.* 2017; 45(D1):D51–D55. <https://doi.org/10.1093/nar/gkw1069> PMID: 27899657
69. Dreos R, Ambrosini G, Périer RC, Bucher P. The Eukaryotic Promoter Database: expansion of EPDnew and new promoter analysis tools. *Nucleic Acids Res.* 2015; 43(D1):D92–D96. <https://doi.org/10.1093/nar/gku1111> PMID: 25378343