Supplemental Material for

Accessibility of promoter DNA is not the primary determinant of chromatin-mediated gene regulation

Răzvan V. Chereji[1,†], Peter R. Eriksson[1,†], Josefina Ocampo[1,2,†], Hemant K. Prajapati[1], and David J. Clark[1,*]

[1]Division of Developmental Biology, *Eunice Kennedy Shriver* National Institute for Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

[2]Current address: Instituto de Investigaciones en Ingeniería Genética y Biología Molecular "Dr. Héctor N. Torres" (INGEBI-CONICET), Buenos Aires, Argentina.

[†]These authors contributed equally to this work.

[*]Corresponding author. E-mail: clarkda@mail.nih.gov

Contents

Supplemental Methods

Restriction enzyme digestion kinetics

The kinetics of enzyme processes can be investigated using the Michaelis–Menten framework, which we briefly describe in this section. Let's consider the following enzyme catalyzed reaction,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P,$$

where we denote the enzyme (AluI restriction enzyme in our case) with E, the DNA containing the AluI target site ("substrate") with S, the enzyme-substrate complex (enzyme bound to its site) by ES, and the cut DNA ("product") with P. The first step of association between the substrate and enzyme to form the complex ES is reversible, with forward and reverse rates k_1 and k_{-1}, respectively. The second step of DNA cleavage by AluI is assumed to be irreversible, happening with the rate k_2.

Let's denote the concentrations of each component as follows:

$$\begin{aligned}
[E] &- \text{the concentration of free enzyme } E;\\
[S] &- \text{the concentration of substrate } S;\\
[ES] &- \text{the concentration of enzyme-substrate complex } ES;\\
[P] &- \text{the concentration of product } P.
\end{aligned}$$

These concentrations satisfy the following reaction equations:

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = -k_1[S][E] + k_{-1}[ES], \tag{1}$$

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} = -k_1[S][E] + (k_{-1} + k_2)[ES], \tag{2}$$

$$\frac{\mathrm{d}[ES]}{\mathrm{d}t} = k_1[S][E] - (k_{-1} + k_2)[ES], \tag{3}$$

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_2[ES]. \tag{4}$$

Adding Eqs. (2) and (3) we obtain that

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} + \frac{\mathrm{d}[ES]}{\mathrm{d}t} = 0,$$

meaning that $[E](t) + [ES](t) = [E]_0$, so the total concentration of enzyme (either free or bound to DNA) remains constant during the experiment. Adding Eqs. (1), (3), and (4) we obtain that

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} + \frac{\mathrm{d}[ES]}{\mathrm{d}t} + \frac{\mathrm{d}[P]}{\mathrm{d}t} = 0,$$

so the total amount of DNA (either cut or uncut) is also constant, $[S](t) + [ES](t) + [P](t) = [S]_0$. Therefore, we can solve for $[E](t)$ and $[P](t)$,

$$\begin{aligned}
[E](t) &= [E]_0 - [ES](t),\\
[P](t) &= [S]_0 - [S](t) - [ES](t),
\end{aligned}$$

and substitute these into the initial system of four differential equations, which reduces to:

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} = -k_1[S]([E]_0 - [ES]) + k_{-1}[ES],$$
$$\frac{\mathrm{d}[ES]}{\mathrm{d}t} = k_1[S]([E]_0 - [ES]) - (k_{-1} + k_2)[ES].$$

This set of coupled nonlinear differential equations cannot be solved analytically, and numerical integration is required to obtain an accurate solution. However, the situation simplifies considerably if we assume that the cleavage reaction happens very fast compared to the binding and unbinding processes, i.e. if we assume that the process of searching the specific site of AluI is the rate limiting step. In this case, we can use the quasi-steady-state approximation (QSSA), which assumes that the reactive intermediate ES will be present at a low concentration that is changing very slowly throughout the course of the reaction, i.e. $\frac{\mathrm{d}[ES]}{\mathrm{d}t} \approx 0$. This type of equations and assumptions can be rigorously studied using singular perturbation theory, which is outside the scope of this article.

Using the QSSA we obtain

$$k_1[S]([E]_0 - [ES]) - (k_{-1} + k_2)[ES] = 0,$$

and

$$[ES] = \frac{[E]_0[S]}{K_m + [S]}$$

where $K_m = \frac{k_{-1}+k_2}{k_1}$ is known as the Michaelis constant. From Eq. (4) we obtain the production rate

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_2[ES] = \frac{V_{\max}[S]}{K_m + [S]},$$

where $V_{\max} = k_2[E]_0$. Substituting $[ES]$ into the equation for $[S]$ we obtain

$$\begin{aligned}\frac{\mathrm{d}[S]}{\mathrm{d}t} &= -k_1[S]\left([E]_0 - \frac{[E]_0[S]}{K_m + [S]}\right) + k_{-1}\frac{[E]_0[S]}{K_m + [S]} \\ &= \frac{[E]_0[S]}{K_m + [S]}\left(-k_1(K_m + \cancel{[S]}) + \cancel{k_1[S]} + k_{-1}\right) \\ &= -\frac{V_{\max}[S]}{K_m + [S]}.\end{aligned}$$

So, we obtained the Michaelis-Menten equation, which gives the rate of product creation and substrate decay

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = -\frac{\mathrm{d}[S]}{\mathrm{d}t} = \frac{V_{\max}[S]}{K_m + [S]}, \tag{5}$$

where the maximum rate is $V_{\max} = k_2[E]_0$, and the Michaelis constant is $K_m = \frac{k_{-1}+k_2}{k_1}$.

The Michaelis-Menten rate equation does not have a definite order with respect to the substrate, as in different regimes of the substrate concentration it can be simplified either to a first-order rate equation,

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} \approx -\frac{V_{\max}}{K_m}[S], \text{if } [S] \ll K_m,$$

or to a zero-order rate equation,

$$\frac{\mathrm{d}[S]}{\mathrm{d}t} \approx -V_{\max}, \text{if } [S] \gg K_m.$$

In both cases the rate equation can be easily integrated to obtain

$$[S](t) = [S]_0 \exp\left(-\frac{V_{\max}}{K_m}t\right) \text{ (first-order rate)},$$

or

$$[S](t) = [S]_0 - V_{\max}t \text{ (zero-order rate)},$$

where $[S](0) = [S]_0$.

In the general case, integrating Eq. (5) we obtain

$$\int \frac{K_m + [S]}{[S]}\mathrm{d}[S] = \int -V_{\max}\mathrm{d}t$$

or

$$K_m \log\left(\frac{[S](t)}{[S]_0}\right) + ([S](t) - [S]_0) + V_{\max}t = 0.$$

This is an exact solution of the Michaelis-Menten equation, but unfortunately it is an implicit nonlinear equation. The explicit closed-form solution of this equation Schnell and Mendoza (1997); Goličnik (2011) is rarely used, as it involves the Lambert W function, the special function that solves the transcendental equation

$$W(x)e^{W(x)} = x.$$

The explicit solution of the Michaelis-Menten rate equation Schnell and Mendoza (1997); Goličnik (2011) is

$$[S](t) = K_m W\left(\frac{[S]_0}{K_m}\exp\left(\frac{[S]_0 - V_{\max}t}{K_m}\right)\right), \tag{6}$$

which can be conveniently rewritten in a dimensionless form as

$$s(t) = W\left(s_0 \exp\left(s_0 - vt\right)\right), \tag{7}$$

where $s = [S]/K_m$ is the reduced concentration, and $v = V_{\max}/K_m$ is the first-order rate constant, which is directly proportional to the enzyme concentration,

$$v = \frac{V_{\max}}{K_m} = \frac{k_1 k_2}{k_{-1} + k_2}[E]_0.$$

In our experiments, instead of digesting the chromatin for different amounts of time, we digested the chromatin for a fixed time, $T = 20$ min, but using different concentrations of AluI, $[E]$. In this case, the reduced concentration of substrate at time $t = T$ can be written as a function of the enzyme concentration that was used in a particular experiment,

$$\begin{aligned} s([E]) &= W\left(s_0 \exp\left(s_0 - v([E])T\right)\right) \\ &= W\left(s_0 \exp\left(s_0 - k[E]\right)\right), \end{aligned} \tag{8}$$

which has a simple form, with only two free parameters:

$$k = \frac{v([E])T}{[E]} = \frac{k_1 k_2}{k_{-1} + k_2} T, \tag{9}$$
$$s_0 = \frac{[S]_0}{K_m}.$$

In a population of cells, a specific AluI site can be either accessible to the restriction enzyme, or inaccessible – blocked by nucleosomes or other DNA-binding proteins. Let A be the fraction of cells in which this AluI site is accessible and $1 - A$ the inaccessible fraction. Let's denote the concentration of accessible and inaccessible AluI sites by

$[D]$ – the concentration of accessible sites (free DNA);
$[B]$ – the concentration of inaccessible/blocked sites.

Then, at the beginning of the reaction we have that

$$[D](0) = [D]_0 = A[S]_0,$$
$$[B](0) = [B]_0 = (1 - A)[S]_0,$$

and the corresponding dimensionless variables

$$d_0 = \frac{[D]_0}{K_m} = As_0,$$
$$b_0 = \frac{[B]_0}{K_m} = (1 - A)s_0.$$

At the end of the digestion, according to Eq. (8) we have that

$$\begin{aligned} d([E]) &= W\left(d_0 \exp\left(d_0 - k[E]\right)\right) \\ &= W\left(As_0 \exp\left(As_0 - k[E]\right)\right), \end{aligned}$$

while the amount of inaccessible/blocked substrate remains constant during the cleavage reaction,

$$b([E]) = b_0 = (1 - A)s_0.$$

The fraction of AluI sites that remain uncut after time T is

$$\begin{aligned} f_{\text{uncut}}([E]) &= \frac{d([E]) + b([E])}{d_0 + b_0} \\ &= \frac{W\left(As_0 \exp\left(As_0 - k[E]\right)\right) + (1 - A)s_0}{s_0} \\ &= \frac{W\left(As_0 \exp\left(As_0 - k[E]\right)\right)}{s_0} + (1 - A). \end{aligned}$$

The fraction of AluI sites that were cut before time T is

$$\begin{aligned} f_{\text{cut}}([E]) &= 1 - f_{\text{uncut}}([E]) \\ &= A - \frac{W\left(As_0 \exp\left(As_0 - k[E]\right)\right)}{s_0}. \end{aligned}$$

So, the fractions of AluI sites that are cut and uncut at the end of the experiment, depend on four variables: $[E]$ – the concentration of restriction enzyme used in the experiment; A – the fraction of accessible sites; k – the first-order cleavage rate; and s_0 – the substrate concentration. We obtained that

$$f_{\text{cut}}([E], A, k, s_0) = A\left(1 - \frac{W\left(As_0 \exp\left(As_0 - k[E]\right)\right)}{As_0}\right) \tag{10}$$

$$f_{\text{uncut}}([E], A, k, s_0) = 1 - A + A\frac{W\left(As_0 \exp\left(As_0 - k[E]\right)\right)}{As_0} \tag{11}$$

In Fig. S9 we show the dependence of f_{cut} and f_{uncut} on the four parameters: $[E]$, A, k, and s_0. We see that the accessibility parameter, A, dictates the asymptotic (plateau) values for f_{cut} and f_{uncut} (Fig. S9A):

$$\lim_{[E]\to\infty} f_{\text{cut}} = A,$$
$$\lim_{[E]\to\infty} f_{\text{uncut}} = 1 - A.$$

The rate parameter k dictates the initial slopes of the f_{cut} and f_{uncut} curves (Fig. S9B), and the s_0 parameter determines the overall shape of the curves (Fig. S9C): a low s_0 results in a first-order reaction (exponential decay of the substrate), while a high s_0 results in a zero-order reaction (linear decay of the substrate).

Fig. S10 shows the fraction of AluI sites that were (Fig. S10A) or were not cut (Fig. S10B) when different concentrations of AluI were used in the cleavage reaction. Although there is a significant variability among the plateau levels for different sites (percentiles represented as different shades of red and blue in Figs. S10A and S10B, respectively), the overall shape of the medians suggest an exponential decay of the concentration of accessible AluI sites, consistent with a first-order decay reaction. By fitting the medians of f_{cut} and f_{uncut} using Eqs. (10) and (11), respectively, we obtained the following fitted parameters: $A = 0.209$, $k = 0.048$, and a very small s_0 parameter ($s_0 \ll 1$).

As we pointed out before, when $s_0 \ll 1$ (or equivalently $[S] \ll K_m$), the Michaelis-Menten rate equation becomes a first-order rate equation, with the solution

$$[S](t) = [S]_0 \exp\left(-\frac{V_{\max}}{K_m}t\right).$$

At the end of the reaction (at time T), the reduced concentration as a function of the enzyme concentration $[E]$ is

$$s([E]) = s_0 \exp\left(-k[E]\right),$$

where $s = [S]/K_m$ and k is given by Eq. (9). Similarly, in the case of two types of substrate (accessible and blocked sites), we obtain that

$$d([E]) = d_0 \exp\left(-k[E]\right) = As_0 \exp\left(-k[E]\right)$$

and

$$b([E]) = b_0 = (1 - A)s_0.$$

Therefore, the fraction of DNA that was not cut by AluI during the cleavage reaction becomes

$$f_{\text{uncut}}([E], A, k) = \frac{d([E]) + b([E])}{s_0} = Ae^{-k[E]} + (1 - A), \tag{12}$$

while the cut fraction is

$$f_{\text{cut}}([E], A, k) = 1 - f_{\text{uncut}}([E], A, k) = A\left(1 - e^{-k[E]}\right). \tag{13}$$

Clearly, the asymptotic values of these two fractions are

$$\lim_{[E]\to\infty} f_{\text{cut}} = A,$$
$$\lim_{[E]\to\infty} f_{\text{uncut}} = 1 - A,$$

and the accessibility parameter A can be easily obtained from these asymptotes. The other parameter, the first-order rate constant k, can be obtained by analyzing the initial slope of $\log(f_{\text{uncut}})$. We have that

$$\log\left(f_{\text{uncut}}([E], A, k)\right) = \log\left(Ae^{-k[E]} + (1 - A)\right).$$

When $[E] \to 0$, we have that

$$e^{-k[E]} \approx 1 - k[E],$$

and

$$\begin{aligned}\log\left(f_{\text{uncut}}([E], A, k)\right) &\approx \log\left(A(1 - k[E]) + (1 - A)\right)\\ &= \log\left(1 - kA[E]\right)\\ &\approx -kA[E]\end{aligned}$$

so for the initial stages of digestion, $\log(f_{\text{uncut}})$ decreases linearly with $[E]$, with a slope given by

$$\text{Slope} = \frac{\partial}{\partial[E]} \log(f_{\text{uncut}}) = -kA \tag{14}$$

Using the plateau level for f_{cut} or f_{uncut} we obtain a quantitative estimation for DNA accessibility A, which is not available through other methods, such as MNase-seq, ATAC-seq, or DNase-seq. Knowing the accessibility of each AluI site and measuring the initial slope of $\log(f_{\text{uncut}})$, using Eq. (14) we can compute the rate constant k for each site. If different AluI sites are located in genomic regions characterized by different chromatin openness, then we expect the overall rates k to vary from region to region: AluI sites located in open chromatin will be characterized by higher cleavage rates k, while the sites located in closed chromatin will be characterized by lower rates k. Surprisingly, when different mouse liver chromatin states were analyzed to determine their typical digestion rate constants, we observed only a relatively small difference among the respective k's (less than 5-fold difference) and not the expected orders-of-magnitude difference. This indicates that the accessibilities of euchromatin and heterochromatin are very similar. Therefore, DNA accessibility is *not* the primary determinant of chromatin-mediated gene regulation.

Duplex qPCR

Primers and Taqman hydrolysis probes were designed such that fluorescence will be detected only if the sequence is not digested with AluI (uncut): the target AluI site is located in between the probe and the primer bound to the same strand, such that the probe cannot be displaced if AluI has cut the DNA. The reference sequence does not contain an AluI site and so should be a measure of the total DNA independent of AluI digestion. Four different target primer sets and their hydrolysis probes (located in the yeast *ARG1*, *HIS4* and *GAL1* genes) and two different reference primer sets and probes (located in *FIN1* and *BIM1*) were designed using the OligoAnalyzer Tool (IDT, USA) and the GenScript Real-time PCR (TaqMan) Primer Design tool (Genscript, USA). Probes for target sequences were labelled at the 5'-end with 6FAM (emission maximum 517 nm); reference probes were labelled at the 5'-end with VIC (emission maximum 551 nm). All probes were labelled with the QSY quencher at the 3'-end. The sequences of the primers and probes are listed below.

Five of the libraries generated in each qDA-seq experiment with arrested yeast cells were used as template in 20 μl duplex qPCR reactions (0, 0.8, 3.2, 29.6, and 54 nM AluI) containing 2.5 μl template at $\sim$10 ng/μl and TaqMan Fast Advanced Master Mix (Applied Biosystems 4444557). qPCR reactions were set up in 96-well plates (Applied Biosystems 4346906) using MicroAmp optical adhesive sealing (Applied Biosystems 4311971). qPCR was performed in a 7500 Fast Real-Time PCR System (Applied Biosystems, USA) in standard mode. After optimization, the final concentration used for primers and probes was 0.25 μM. The PCR cycling parameters used were: 2 min initial hold at 55°C, 5 min initial denaturation at 95°C, followed by 40 cycles of 30 s at 95°C, 30 s at 52°C and 30 s at 60°C. A No Template Control was run with every set of primers and probes, which gave C_t values > 40. All samples were measured in triplicate. PCR efficiencies were estimated using two different methods (dilution series and curve-fitting using LinRegPCR software) with similar results. For the AluI-digested samples, the uncut fraction was obtained as the ratio of the target probe to the reference probe,

$$\frac{N_0^{\text{target}}}{N_0^{\text{ref}}} = \frac{(1+E^{\text{ref}})^{C_t^{\text{ref}}}}{(1+E^{\text{target}})^{C_t^{\text{target}}}},$$

where N_0 is the initial template concentration, E is the PCR efficiency for the target or reference primer pairs, and C_t is the threshold cycle.

The values obtained from this equation were normalized to the undigested control to obtain the uncut fraction. The uncut fraction was subtracted from 1 to obtain the AluI-cut fraction. Duplex PCR reactions were performed using one of two references: *FIN1* was used as a reference with *ARG1* targets; *BIM1* was used as a reference with the *GAL1* and *HIS4* targets.

List of primers and probes used in qPCR:
ARG1 A5 forward primer: CCTGCAAGTTCGTTTGTG
ARG1 A5 probe: 6FAM-ACAGGTCAACGCTGTGTACG-QSY
ARG1 A5 reverse primer: AGGTCTTGCCAAAGAGG
ARG1 A8 forward primer: CATCATACCAATTTTGAATAGAAG
ARG1 A8 probe: 6FAM-CGCTCGGAATAAGTTGACGTAG-QSY
ARG1 A8 reverse primer: AATGTTCTGTTCTCATGAAG
GAL1 A3 forward primer: GTGCGTCCTCGTCTTCA
GAL1 A3 probe: 6FAM-CTTTATTGTTCGGAGCAGTGCGG-QSY
GAL1 A3 reverse primer: GCCAGGTTACTGCCAATT
HIS4 A4 forward primer: ATTAGCGCCATGAAGAACG
HIS4 A4 probe: 6FAM-TGGGTAGTGCCTTGTGATCCG-QSY

HIS4 A4 reverse primer: ACAGTCAGACTCGTTGGAAT
BIM1 forward primer: ACGCGACATAGAAATCCTG
BIM1 probe: VIC-TGAGACCATCACCGGCCATG-QSY
BIM1 reverse primer: CCACCTTCTTGACGAACC
FIN1 forward primer: ACAGATGGAATGAAGCATAG
FIN1 probe: VIC-CCTTCAGACATTCCGGCGAC-QSY
FIN1 reverse primer: GGCCTATCCAGGCTTTG
The probes were obtained from ThermoFisher.

Supplemental Figures

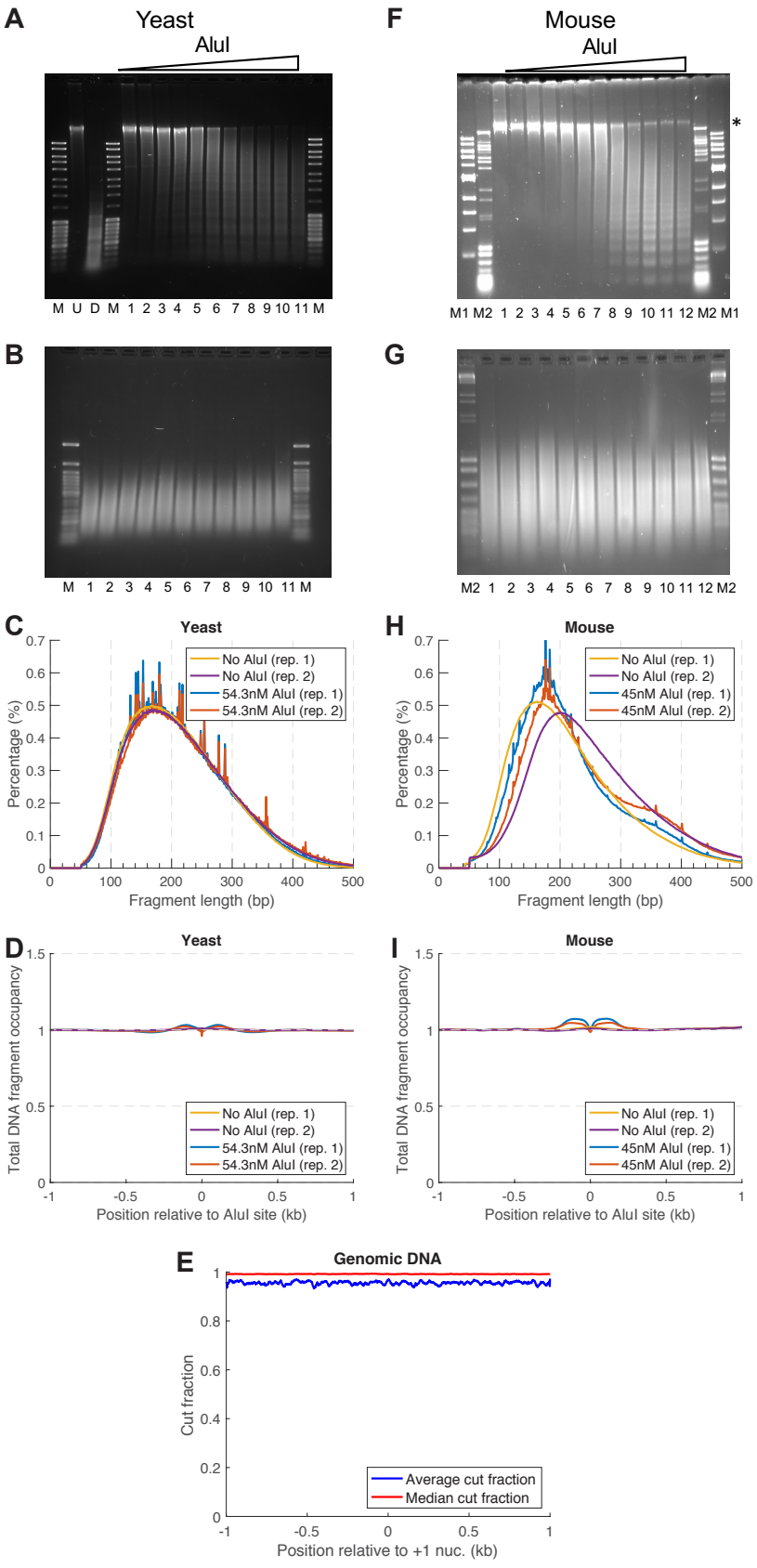

Figure S1. Caption on the next page.

Figure S1. Digestion of yeast and mouse nuclei with AluI. (A-D) Yeast nuclei. **(A)** Gel electrophoretic analysis of AluI-digested DNA from nuclei (AluI titration; lanes 1-11). U: Undigested genomic DNA; D: fully digested with AluI. M: Mass ruler (Thermo-Fisher). **(B)** The same AluI titration samples after sonication. M: 50-bp ladder (NEB) **(C)** Length histograms of all sequenced DNA fragments after sonication with no AluI and at the highest concentration of AluI (54 nM). Note that the very sharp peaks protruding from the curve in the length histograms for the AluI-digested DNA correspond to the AluI fragments expected from a digest of the rDNA locus on chromosome XII (∼200 copies of a 9.1 kb repeat comprising ∼15% of the yeast genome): predicted AluI fragments in this length range are 452, 441, 421, 417, 356, 278, 254, 222, 218, 213, 180, 153, 145, 141, 132, 66, 61 and 55 bp. **(D)** Coverage plot for all sequenced AluI fragments aligned on the ∼40,000 AluI sites. These plots show only small changes in coverage over AluI sites after AluI digestion, indicating that the recoveries of DNA fragments with and without AluI ends are very similar. **(E)** Complete AluI digestion of yeast genomic DNA. Fraction cut as a function of distance from the center of the +1 nucleosome on all ∼5,000 yeast genes. The median cut fraction is >99% as expected. The mean cut fraction is >96%. The latter is slightly skewed by a few uncut sites, which are mostly explained by genomic sequence differences between this strain and the S288C reference genome. **(F-I)** Equivalent plots for mouse liver nuclei. M1: 1 kb marker (NEB); M2: mixture of λ-DNA digested with BstEII and pBR322 digested with MspI (NEB). In **(F)**, the AluI-resistant long DNA fragment (>14 kb; marked by *) was identified by paired-end sequencing (after gel-purification and sonication) as major mouse satellite DNA (multiple tandem repeats lacking an AluI site).

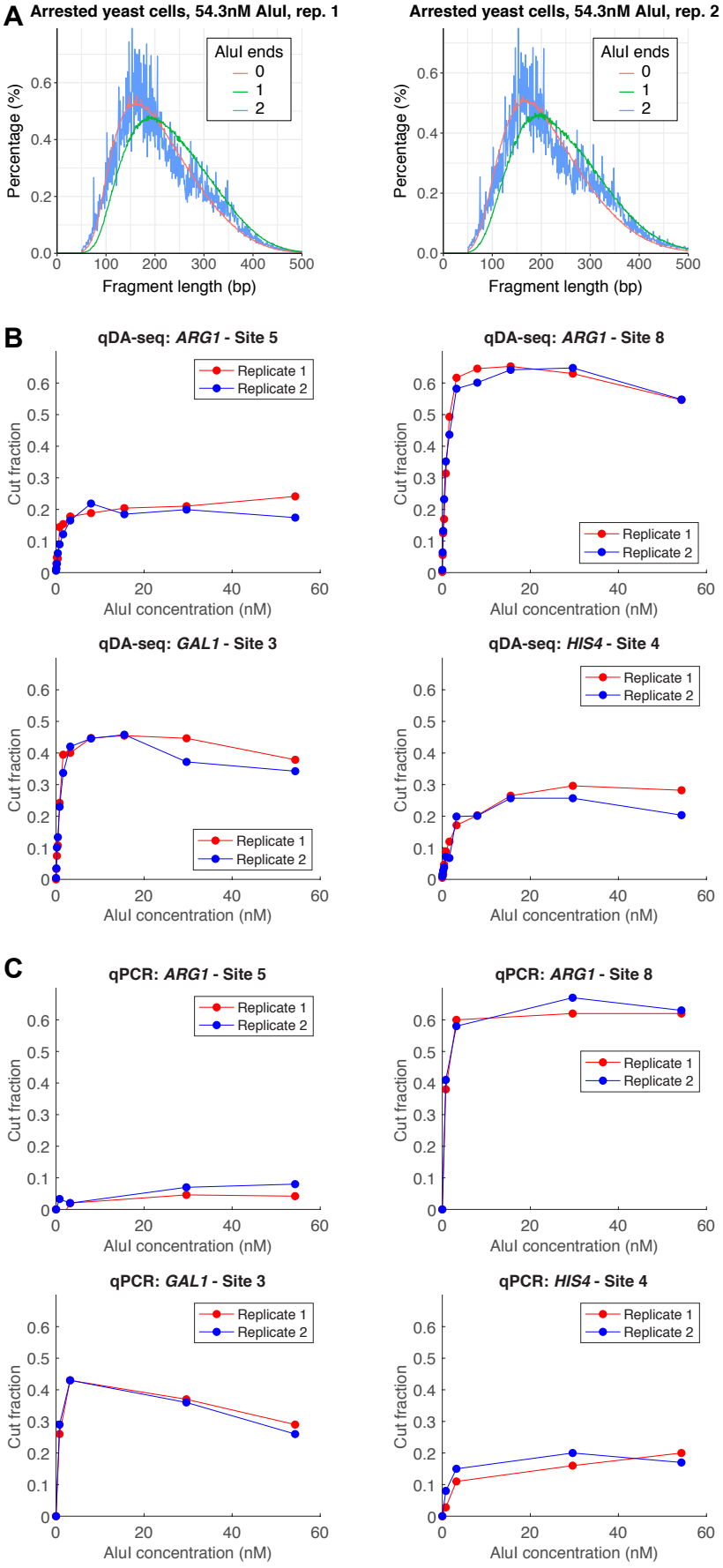

Figure S2. Caption on the next page.

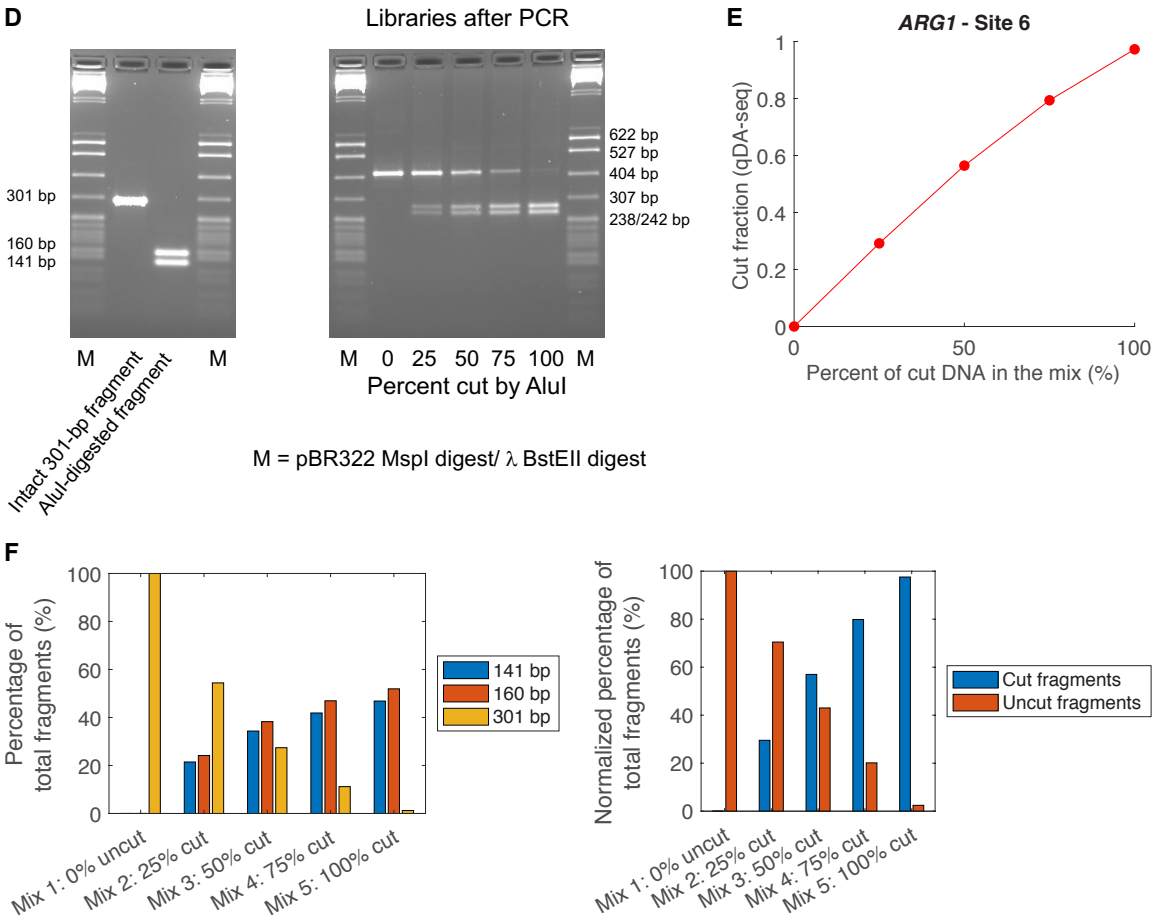

Figure S2. Validation of qDA-seq sequencing data. **(A)** Length histograms for DNA fragments with AluI and/or sonicated ends. To check for a potential bias towards shorter AluI-cut fragments as compared to their uncut counterparts, the DNA fragments were separated into the three possible categories: (i) both ends result from AluI cleavage; (ii) one end results from AluI cleavage and the other from breakage by sonication; (iii) both ends result from breakage by sonication. Samples: arrested yeast cell nuclei digested with 54 nM AluI (biological replicates 1 and 2). These three groups of fragments have similar length distributions, indicating that AluI-cut DNA fragments are not shorter than uncut DNA fragments. We suggest that sonication preferentially breaks longer fragments, resulting in similar distributions for uncut and AluI-cut fragments. **(B)** qDA-seq plots for four selected AluI sites in arrested yeast cell nuclei (see Fig. S4 for their locations). **(C)** qPCR plots for the same sites shown in **(B)**. AluI-cut fractions were measured by qPCR in five of the same libraries used for qDA-seq (data in **(B)**). See Supplemental Methods for details. There is close agreement between replicates and qPCR gives similar plateau values to qDA-seq for the same samples. **(D-F)** The qDA-seq method accurately reports the AluI-cut fraction. **(D)** A 301-bp DNA fragment corresponding to chromosome XV nucleotides 219,937 to 220,237 was obtained by PCR. This fragment contains AluI site 6 at nt 220,096 in the ARG1 ORF (see Fig. 2A for designation of AluI sites). A portion of the purified DNA fragment was cut to completion with AluI, yielding fragments of 160 bp and 141 bp. The DNA fragments were checked in a 3% agarose gel stained with ethidium bromide (left panel). The DNA concentrations of the uncut and cut fragment were measured. Five mixtures were prepared: 0% cut, 25% cut, 50% cut, 75% cut and 100% cut. Indexed paired-end libraries were made from each mixture using the same protocol described for yeast libraries and checked in a 3% agarose gel (right panel). Note that the addition of adaptor to both ends of the DNA followed by PCR increases the lengths of all three fragments by 120 bp. **(E)** Cut fractions obtained by qDA-seq for the five samples containing a mix of intact 301-bp fragments and AluI-digested fragments. The percent of AluI-cut DNA in each sample is indicated on the horizontal axis: 0%, 25%, 50%, 75%, 100%. **(F)** Bar plot of the percentage of fragments having the lengths of 141, 160, and 301 bp. Left panel contains the raw percentages, where cut fragments of 141 bp and 160 bp were counted separately. After computing the average count for the short fragments (141 bp and 160 bp) as an estimation of the cut fragments, we obtained the normalized percentage of fragments that were cut by AluI (right panel).

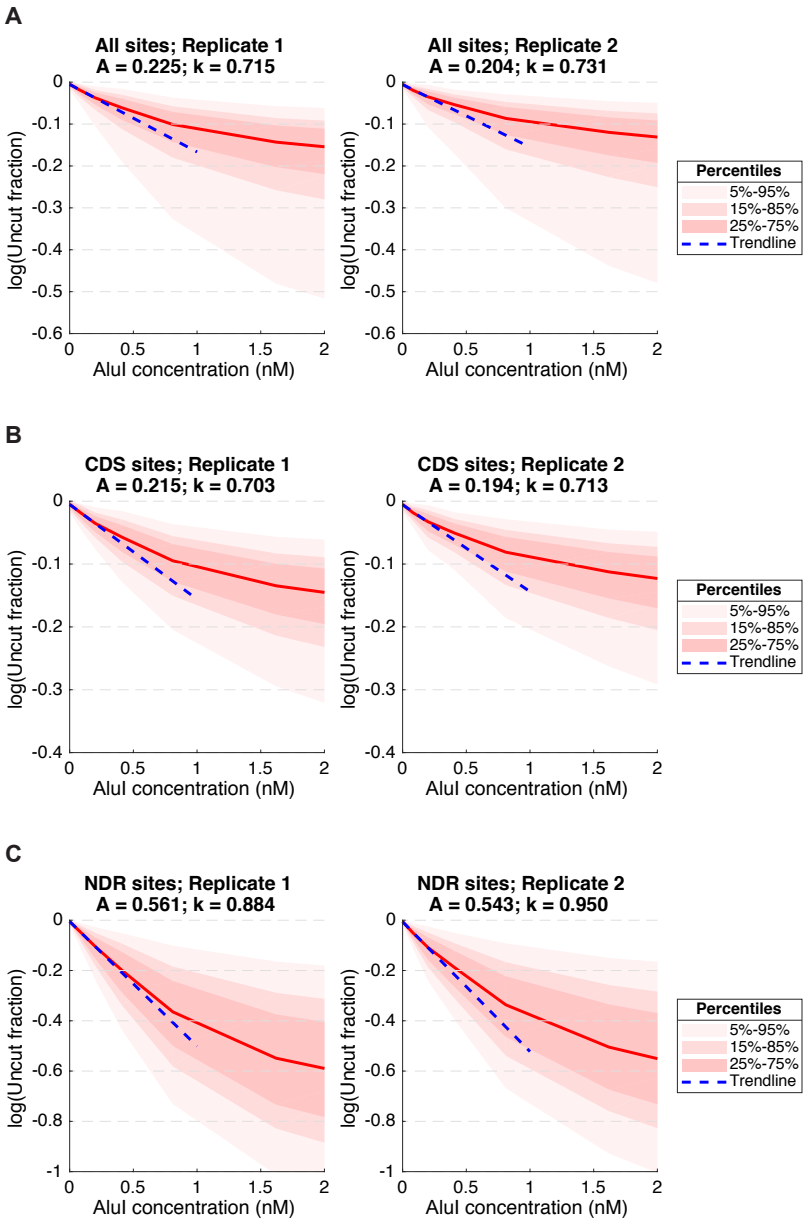

Figure S3. Mean initial digestion rates of AluI sites in yeast cells. The natural logarithm of the uncut fraction vs. AluI concentration for two biological replicate experiments. **(A)** All AluI sites. **(B)** AluI sites in gene bodies. **(C)** AluI sites in promoter NDRs (defined by the locations of the +1 and -1 nucleosomes obtained from Chereji et al. (2018)). See Supplemental Methods for a full mathematical treatment.

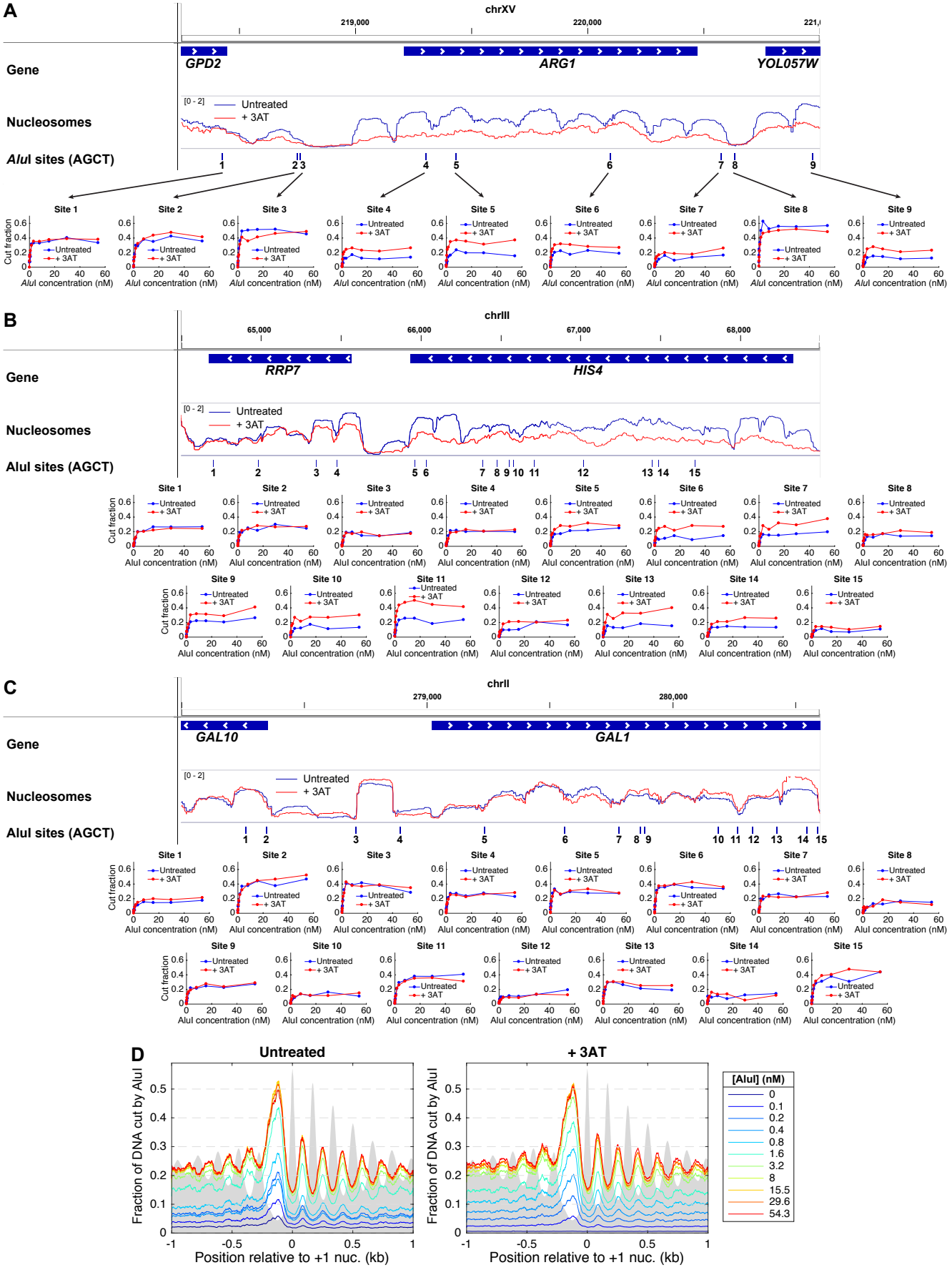

Figure S4. Caption on the next page.

Figure S4. Heavy transcription after 3AT induction results in increased AluI accessibility in yeast. After induction with 3AT, a small number of Gcn4-dependent genes, including **(A)** *ARG1* and **(B)** *HIS4*, show significant nucleosome loss over the coding region, sometimes extending into the flanking regions, as well as disrupted nucleosome spacing and a wider NDR, all of which correlate with heavy transcription Cole et al. (2011, 2014). AluI accessibility data and MNase-seq data are shown (blue line: control; red line: +3AT). **(C)** *GAL1* (not induced by 3AT). **(D)** Mean AluI accessibility as a function of distance from the center of the +1 nucleosome (defined in Chereji et al. (2018)) on all ∼5,000 yeast genes for exponentially growing cells before and after treatment with 3AT for 20 min (cf. arrested cells; Fig. 3A). Grey area: MNase-seq data showing nucleosome dyad distributions (fragment centers) on an arbitrary scale.

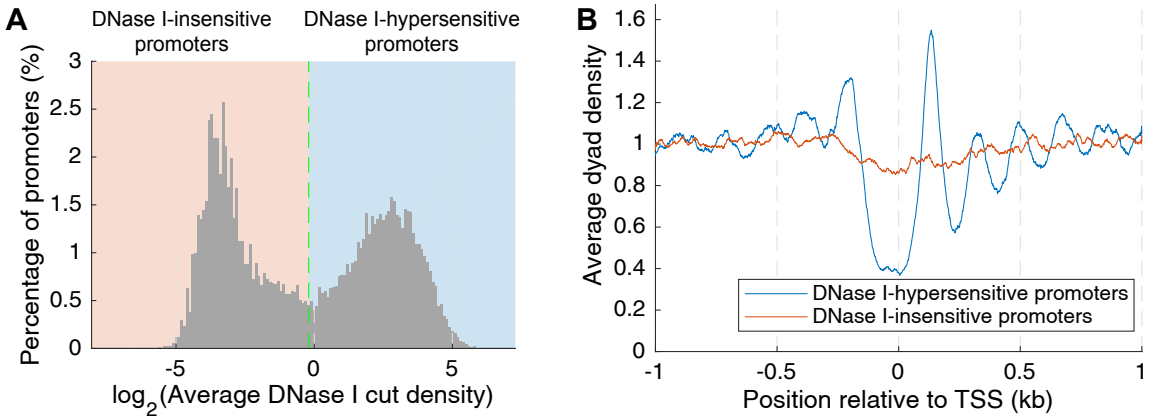

Figure S5. Mouse genes in liver can be divided into two distinct chromatin states using promoter DNase I hypersensitivity data. (A) Histogram of the number of promoters with a given DNase I cut density. The dashed line indicates the dividing line used in the DNase I heatmap in Fig. 4C (white line). **(B)** Nucleosome phasing on genes with DNase I-hypersensitive promoters and DNase I-insensitive promoters (data above and below the white line in the MNase-Exo-seq heatmap in Fig. 4C).

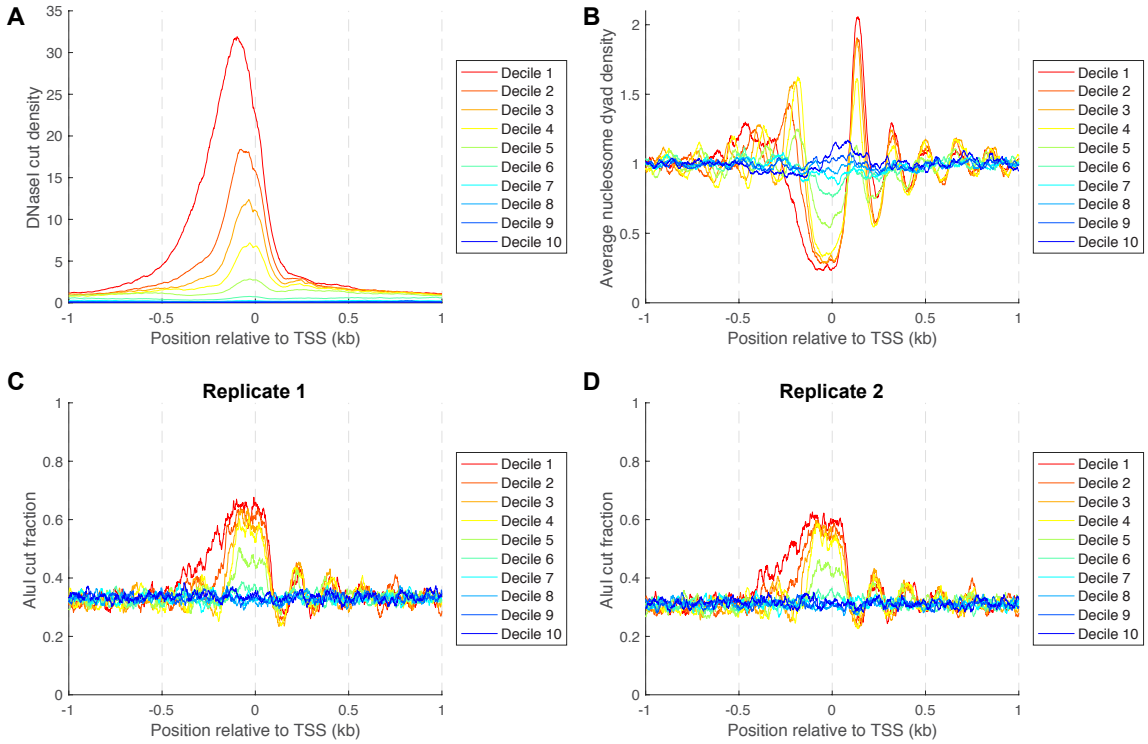

Figure S6. Fine analysis of the data in Fig. 4. (A) All mouse promoters were ranked according to their DNase I hypersensitivity (Fig. 4C) and divided into deciles, such that decile 1 contains the most hypersensitive promoters. The plot shows the DNase I cut density for each decile. **(B)** MNase-Exo-seq data for the DNase I deciles. **(C,D)** AluI data for each decile (replicates 1 and 2). The DNase I data show progressively weaker promoter peaks for deciles 1 - 10, as expected. The MNase and AluI data show that the accessibility of the NDR does not change much in the top four deciles; the main difference is that the NDR extends farther upstream in decile 1, which is linked with high transcriptional activity, at least in yeast. Deciles 6 -10 are very similar, with average accessibility close to the genomic average.

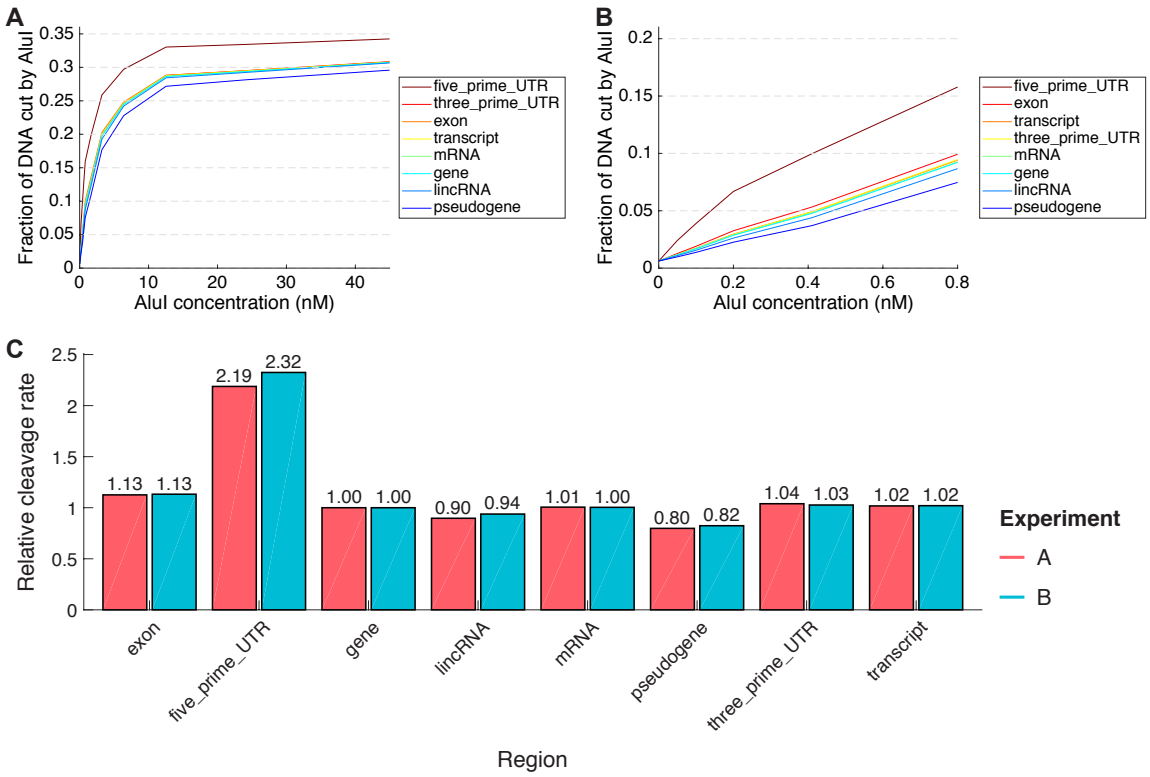

Figure S7. Relative AluI cleavage rates for sites in various annotated regions of the mouse genome (mm10). (A) AluI digestion kinetics for 8 types of annotated regions of the mouse genome. The asymptotic value of f_{cut} represents the accessible fraction of each region, A. **(B)** Close up view on the domain of low AluI concentrations. The initial slope of each of these curves is equal with $k \cdot A$, so it can be used as a measure of the cleavage rate constant corresponding to each region, k. **(C)** Comparison of the cleavage rate constants corresponding to 8 annotated regions (two biological replicate experiments). See Supplemental Methods for details of the analysis.

A

		AluI (nM)	Million reads	Coverage
Arrested yeast cells	Replicate 1	0	13.9	250.8
		0.1	12.4	224.5
		0.2	15.0	273.8
		0.4	8.8	158.7
		0.8	14.3	261.0
		1.6	20.3	360.0
		3.2	14.9	265.1
		8	17.5	299.2
		15.5	13.2	235.9
		29.6	16.6	305.4
		54.3	19.3	353.7
Arrested yeast cells	Replicate 2	0	14.3	266.1
		0.1	17.0	332.0
		0.2	9.8	193.8
		0.4	17.8	317.7
		0.8	11.7	223.5
		1.6	17.8	328.9
		3.2	13.1	238.4
		8	9.0	173.5
		15.5	10.1	201.9
		29.6	23.9	456.8
		54.3	15.0	284.8

		AluI (nM)	Million reads	Coverage
Asynchronous yeast cells	Replicate 1	0	6.3	126.7
		0.1	17.4	345.2
		0.2	7.8	147.5
		0.4	8.5	160.8
		0.8	16.3	313.4
		1.6	10.9	207.2
		3.2	15.6	294.1
		8	15.3	277.8
		15.5	17.3	335.6
		29.6	7.8	155.4
		54.3	15.2	281.7
Asynchronous yeast cells	+3AT; Replicate 1	0	23.2	451.1
		0.1	14.3	264.0
		0.2	8.5	166.0
		0.4	11.6	232.4
		0.8	9.5	183.6
		1.6	38.6	712.4
		3.2	14.3	281.0
		8	29.1	555.8
		15.5	14.9	301.6
		29.6	14.4	288.9
		54.3	24.7	493.1

		AluI (nM)	Million reads	Coverage
Asynchronous yeast cells	Replicate 2	0	6.4	147.2
		0.1	6.5	149.9
		0.2	6.8	151.0
		0.4	7.7	171.7
		0.8	6.9	150.3
		1.6	8.9	202.8
		3.2	25.3	458.7
		8	19.7	362.6
		15.5	22.6	407.0
		29.6	24.6	439.4
		54.3	22.2	371.6
Asynchronous yeast cells	+3AT; Replicate 2	0	10.8	197.8
		0.1	15.5	302.1
		0.2	16.5	302.5
		0.4	14.7	271.1
		0.8	12.6	234.5
		1.6	20.1	378.2
		3.2	18.0	328.1
		8	13.9	248.3
		15.5	15.7	303.9
		29.6	18.4	361.7
		54.3	22.2	455.8

		AluI (nM)	Million reads	Coverage
Mouse liver cells	Replicate 1	0	126.2	15.6
		0.05	221.6	32.6
		0.1	161.5	20.6
		0.2	244.0	32.9
		0.4	122.6	14.7
		0.8	116.2	15.7
		1.6	124.8	14.2
		3.2	115.0	14.8
		6.4	128.0	18.7
		12.5	109.1	14.0
		24.1	130.6	15.7
		45	130.3	16.0
Mouse liver cells	Replicate 2	0	115.3	15.5
		0.05	207.9	28.4
		0.1	227.5	30.0
		0.2	190.0	26.1
		0.4	124.4	16.8
		0.8	122.6	16.2
		1.6	131.0	16.8
		3.2	123.7	16.0
		6.4	129.7	17.2
		12.5	129.3	15.6
		24.1	122.4	16.5
		45	118.9	15.7

B

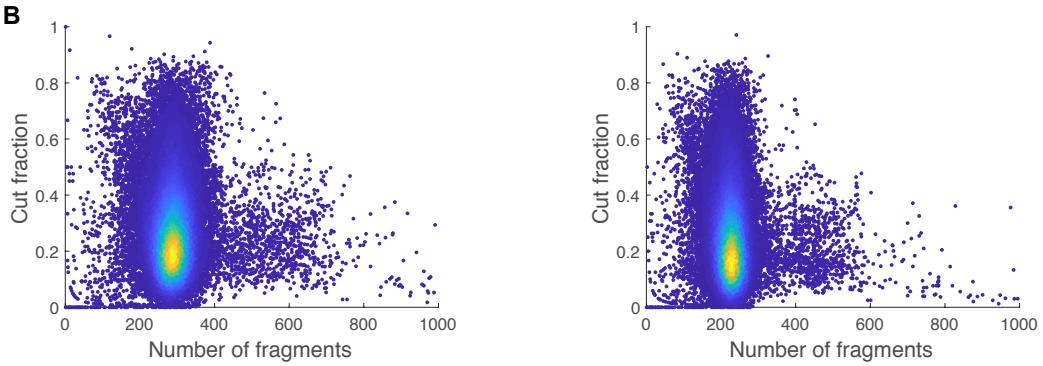

Figure S8. Sequencing depth statistics. (A) Table of the total number of read pairs (in millions) and the corresponding genomic coverage for all data sets used in this study. **(B)** Density plots showing the distribution of the cut fraction as a function of the number of reads for ∼40,000 AluI sites, in arrested yeast cells digested with 54.3nM AluI (two biological replicates).

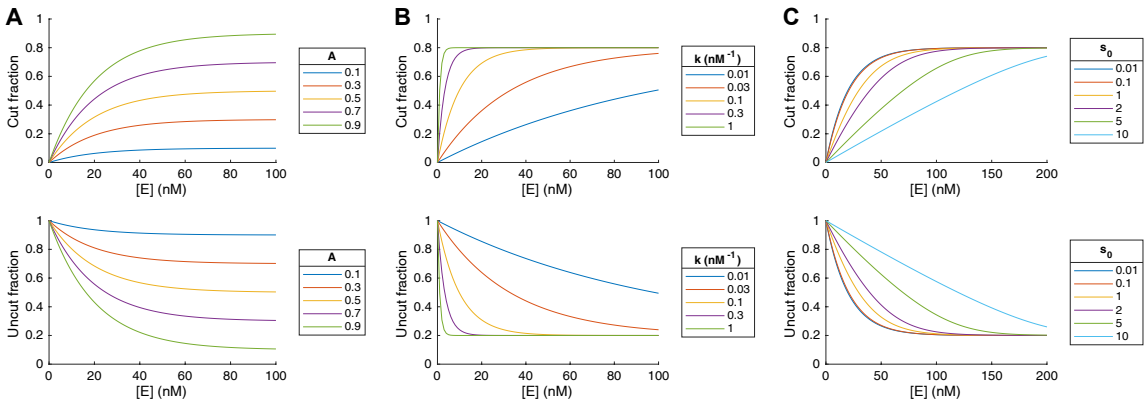

Figure S9. Predictions of the exact solution of the Michaelis-Menten equation (Eqs. (10) **and** (11)**).** The predicted dependence of the fraction of AluI sites that are cut (upper panels) or remain uncut (lower panels) on different parameters: **(A)** DNA accessibility A, **(B)** reaction rate k, and **(C)** the initial substrate concentration s_0 ($[S]_0/K_m$).

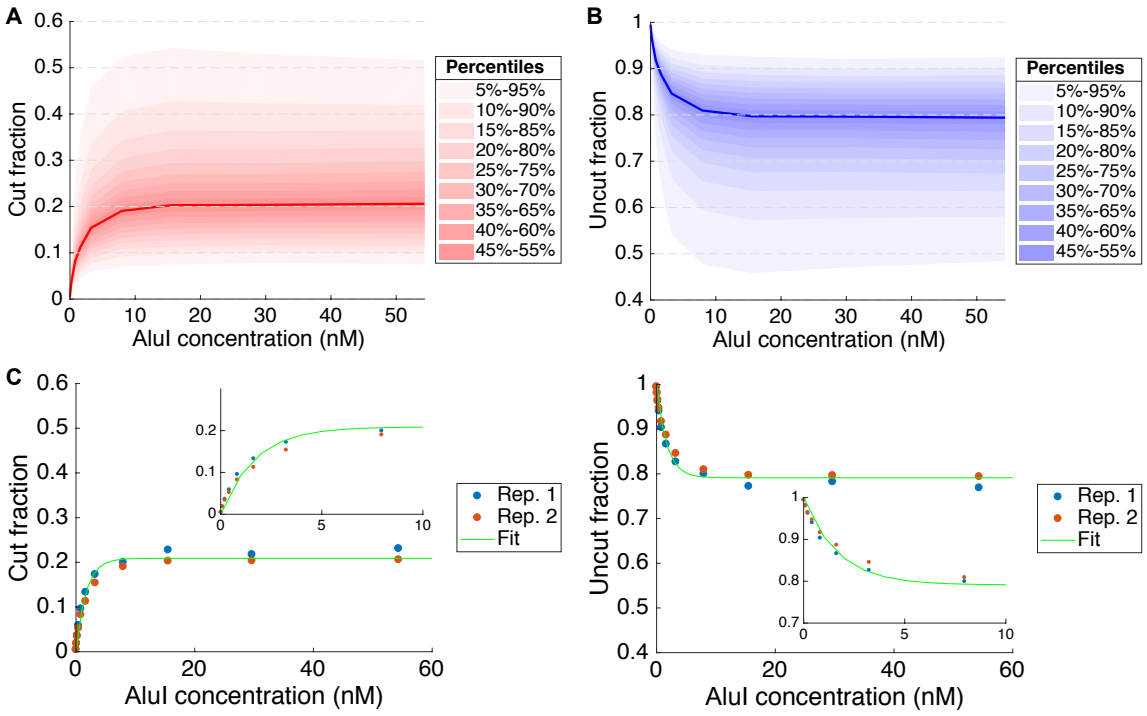

Figure S10. Restriction enzyme cleavage in yeast. (A) The fraction of sites cut by different concentrations of AluI in 20 min. The median cut fraction for each AluI concentration is shown with a red line, and different percentiles of the cut fraction distribution are shown with different shades of red. **(B)** The fraction of sites that were not cut by AluI after 20 min. The median value is shown with a blue line, and different percentiles are shown with different shades of blue. **(C)** Fits for the cut and uncut fractions, using the exact solution of the Michaelis-Menten model – Eqs. (10) and (11). Fitted parameters: $A = 0.209$, $k = 0.594$, $s_0 \ll 1$ (too small to get a precise estimate).

Supplemental References

Chereji RV, Ramachandran S, Bryson TD, and Henikoff S. 2018. Precise genome-wide mapping of single nucleosomes and linkers in vivo. *Genome Biol.* **19**: 19.

Cole HA, Howard BH, and Clark DJ. 2011. Activation-induced disruption of nucleosome position clusters on the coding regions of Gcn4-dependent genes extends into neighbouring genes. *Nucleic Acids Res.* **39**: 9521–9535.

Cole HA, Ocampo J, Iben JR, Chereji RV, and Clark DJ. 2014. Heavy transcription of yeast genes correlates with differential loss of histone H2B relative to H4 and queued RNA polymerases. *Nucleic Acids Res.* **42**: 12512–12522.

Goličnik M. 2011. Exact and approximate solutions for the decades-old Michaelis-Menten equation: Progress-curve analysis through integrated rate equations. *Biochem. Mol. Biol. Educ.* **39**: 117–125.

Schnell S and Mendoza C. 1997. Closed form solution for time-dependent enzyme kinetics. *J. Theor. Biol.* **187**: 207–212.