3870–3885 *Nucleic Acids Research, 2012, Vol. 40, No. 9* *Published online 12 January 2012*
doi:10.1093/nar/gkr1311

Distal chromatin structure influences local nucleosome positions and gene expression

An Jansen[1,2], Elisa van der Zande[1,2], Wim Meert[1,2], Gerald R. Fink[3,*] and Kevin J. Verstrepen[1,2,*]

[1]Laboratory for Systems Biology, VIB, Bio-Incubator, Gaston Geenslaan 1, B-3001, Leuven, Belgium, [2]Laboratory for Genetics and Genomics, Centre of Microbial and Plant Genetics (CMPG), KU Leuven, Kasteelpark Arenberg 22, B-3001 Leuven, Belgium and [3]Whitehead Institute for Biomedical Research/M.I.T., Nine Cambridge Center, Cambridge, MA 02142, USA

Received August 18, 2011; Revised December 18, 2011; Accepted December 21, 2011

ABSTRACT

The positions of nucleosomes across the genome influence several cellular processes, including gene transcription. However, our understanding of the factors dictating where nucleosomes are located and how this affects gene regulation is still limited. Here, we perform an extensive *in vivo* study to investigate the influence of the neighboring chromatin structure on local nucleosome positioning and gene expression. Using truncated versions of the *Saccharomyces cerevisiae URA3* gene, we show that nucleosome positions in the *URA3* promoter are at least partly determined by the local DNA sequence, with so-called 'anti-nucleosomal elements' like poly(dA:dT) tracts being key determinants of nucleosome positions. In addition, we show that changes in the nucleosome positions in the *URA3* promoter strongly affect the promoter activity. Most interestingly, in addition to demonstrating the effect of the local DNA sequence, our study provides novel *in vivo* evidence that nucleosome positions are also affected by the position of neighboring nucleosomes. Nucleosome structure may therefore be an important selective force for conservation of gene order on a chromosome, because relocating a gene to another genomic position (where the positions of neighboring nucleosomes are different from the original locus) can have dramatic consequences for the gene's nucleosome structure and thus its expression.

INTRODUCTION

The DNA of eukaryotic cells is packaged into chromatin, a complex high-order structure consisting of DNA and its associated proteins. The basic repeating unit of chromatin is the nucleosome, formed when a stretch of DNA wraps around histone proteins (1). DNA in nucleosomes is less accessible than naked DNA, which influences a number of vital biological processes including replication, recombination and gene transcription (2–10). Genome functioning thus depends on the (exact and correct) positioning of nucleosomes on the DNA.

The recent availability of genome-wide nucleosome maps has greatly improved our understanding of *in vivo* nucleosome organization in the model eukaryote *Saccharomyces cerevisiae* (11–17). These studies revealed that the majority of nucleosomes are well-positioned, meaning that they occur at the same locations in most cells of a population. Positioned nucleosomes are separated by short fragments of linker DNA, the average linker length being ~18 bp in *S. cerevisiae*. Interestingly, specific patterns of nucleosome organization occur at the 5′-ends of genes. Here, a nucleosome-depleted region of ~150 bp (generally referred to as the 5′ nucleosome-free region or 5′ NFR) is surrounded by the highly positioned −1 and +1 nucleosomes. The 5′ NFR is located just upstream of the transcription start site (TSS), and many functional *cis*-regulatory sequences such as transcription factor binding sites reside in the 5′ NFR, where binding of transcription factors is not obstructed by the presence of nucleosomes (11–13,16–22).

Whereas the organization of nucleosomes in yeast is directed by a wide range of *trans* factors such as chromatin remodeling enzymes, the local DNA sequence is also a major determinant of nucleosome positioning (23).

*To whom correspondence should be addressed. Tel: +32 (0)16 751393; Email: kevin.verstrepen@biw.vib-kuleuven.be
Correspondence may also be addressed to Gerald R. Fink. Tel: +1 617 258 5215; Email: gfink@wi.mit.edu

AT content is a good predictor of nucleosome occupancy, and high AT content correlates with low nucleosome occupancy (24,25). Typically, one or more homopolymeric runs of polyA, referred to as poly(dA:dT) sequences, occur in the 5' NFR where they act as nucleosome-excluding sequences (11,12,15,19,26). Elegant experiments by Roy *et al.* (27) and Struhl and coworkers (28) showed that the poly(dA:dT) elements in promoters contribute to gene expression. In addition, ~20% of yeast promoters contain stretches of tandem repeats that are often extremely AT-rich and act as nucleosome-excluding sequences (29). It is unclear how AT-rich sequences function as antinucleosomal elements, but studies suggest that poly(dA:dT) tracts have unusual structural and mechanical properties that likely prevent incorporation into nucleosomes (30).

It has been proposed that antinucleosomal AT-rich sequences drive sequence-directed nucleosome positioning by forming boundaries against which nucleosomes are positioned (13,31). NFRs established by poly(dA:dT) tracts tend to be surrounded by well-positioned nucleosomes (11,13). In turn, these highly localized nucleosomes might dictate the position of neighboring nucleosomes because adjacent nucleosomes appear to impose packing constraints on each other, much like beads on a string (13,32–34). This model for sequence-directed organization of nucleosomes throughout the genome is called the 'statistical positioning' or 'barrier' model (13,31). Although the barrier model was proposed more than two decades ago (31,35), it has only recently been applied for nucleosome positioning around nucleosome-depleted promoter regions (11,13). Here, the +1 nucleosomes and to a lesser extent the −1 nucleosomes seem to form barriers that direct the positioning of other nucleosomes as far as 1 kb away (13). Several recent physical modeling studies have confirmed the barrier model for nucleosome organization surrounding 5′ NFRs (36–39). However, recent *in vivo* evidence suggests nucleosome organization at the 5′-ends of genes is aided by a non-statistical ATP-dependent nucleosome packing mechanism (40). Unraveling the contribution and importance of (long-range) statistical positioning on local nucleosome structure therefore requires further *in vivo* experiments.

This study provides novel *in vivo* evidence showing that nucleosome positions in promoters are strongly influenced by the surrounding chromatin context. We use the *URA3* gene, a well-studied sequence with a known nucleosome structure, to measure changes in nucleosome positioning and determine the consequences of an altered chromatin environment on *URA3* expression. Inserting the same *URA3* construct at its native locus as well as different genomic locations alters the chromatin structure of the promoter, as well as *URA3* activity. Our results confirm that poly(dA:dT) sequences shield stretches of DNA from each other's influence on nucleosome positioning. In the absence of a poly(dA:dT) barrier, local nucleosome positioning is influenced by the surrounding chromatin context, as well as local DNA sequence. These experiments show that adjacent chromatin structure is a contributing factor in conserving chromosomal gene order—transpositions to new sites risk loss of gene function despite the presence of intact coding sequences.

MATERIALS AND METHODS

Yeast strains and media

Strains, plasmids and primers used in this study are listed in Supplementary Tables SI, SII and SIII, respectively. Yeast strains are derived from S288C strain BY4741 (41). Strain construction strategies are described in the Supplementary Materials and Methods. Standard yeast media were prepared as described (42).

Growth conditions

Yeast cultures were grown overnight in 5 ml of YPD at 30°C in a rotating wheel unless otherwise noted. For the plate assays, YPD liquid cultures were grown overnight, diluted to OD_{600} 0.15, and grown to OD_{600} 0.8–1.2. Culture densities were adjusted to equivalence, serially diluted 4-fold, and spotted onto YPD, SC-ura and 5-FOA plates. Plates were incubated at 30°C for 3 days. Growth in liquid YPD and SC-ura medium was assayed using the Bioscreen C (Oy Growth Curves Ab Ltd.). YPD cultures were grown overnight, cells were collected, washed and resuspended in water. Bioscreen microtiter plates (Honeycomb II) were inoculated with 1×10^6 cells in YPD or SC-ura medium to a final volume of 300 µl per well. Plates were placed in the Bioscreen, and incubated at 30°C with constant shaking. Optical density was measured every 15 min for 24 h using a 600 nm filter. OD values were log transformed and plotted against measurement time-points. A regression line was fitted to the linear range of the plot, corresponding to exponential growth. The doubling time was defined as the inverse of the slope of the regression line. Each condition was assayed in triplicate.

Flow cytometry

Yeast cultures were grown overnight in a 96-well plate with each well containing 100 µl YPD. Overnight cultures were diluted into another 96-well plate containing 100 µl fresh YPD or SC-ura media and grown for 5 h. All plates were grown at 30°C while shaking. Cells were collected by centrifugation, and shortly before analysis resuspended in PBS buffer. The OD_{600} of each plate was measured, and for a number of representative samples the number of cells/ml was counted using a hemocytometer. All samples contained between 8×10^6 and 2×10^7 cells/ml. Samples were analyzed using a BD Influx cell sorter (BD Biosciences). Fluorescent intensities were examined using a 488 nm excitation and a 530 ± 40 nm emission wavelength filter to detect YFP-tagged cells. Results were analyzed using FlowJo software (Treestar Inc.). For all fluorescence measurements the cellular autofluorescence was measured using strain AJY248 without fluorescent reporters. For all experiments performed, a strain carrying an integrated copy of p*TDH3*-YFP was used as a positive control to account for any changes in illumination intensity. These fluctuations were

small, <10% of the expression level. All population-averaged fluorescence measurements were corrected for autofluorescence by substracting the population-averaged fluorescence measured for strain AJY248. After correcting for autofluorescence, all population-averaged fluorescence measurements were normalized by the population-averaged fluorescence measurements for strain AJY631 (Ura3-YFP with wild-type *URA3* promoter at the native *URA3* locus).

Quantitative PCR

YPD liquid cultures were grown overnight, diluted 1:200 in 50 ml YPD and grown for 5 h. Similarly, SC-ura cultures were diluted 1:100 in 50 ml SC-ura and grown for 6 h. Samples were collected and washed once with ice-cold water and then stored at −80°C. RNA was extracted from yeast cells by first spheroplasting cells for 1 h at 37°C using Solution A [Zymolyase (1 mg/ml; MP Biomedicals), 0.9 M sorbitol, 0.1 M EDTA pH 7.5, 14 mM β-mercaptoethanol)] and then using an ABI 6100 Nucleic Acid Prep Station and reagents (Applied Biosystems). Total RNA was reverse transcribed (Transcriptor First Strand cDNA Synthesis Kit, Roche), and qPCR was performed with Power SYBR Green PCR Master Mix (Applied Biosystems) on an Applied Biosystems Step One Plus system using a two-step program (15 s melting at 95°C and 60 s annealing-extension at 60°C). *URA3* levels were normalized to *ACT1*. *URA3* and *ACT1* were amplified using primers KV2478 × KV2479 and KV36 × KV37, respectively.

Nucleosome positioning

Mononucleosomal DNA was prepared as described (11). Nucleosomal DNA was analyzed using primer pairs tiling the length of the *URA3* promoter region and surrounding sequences (about 1 kb in total). Primer pairs are listed in Supplementary Table SIV, and the primer combinations used for each strain are listed in Supplementary Table SV. Each primer pair generates 97–103 bp products, centered 15–30 bp away from the neighboring primer pair. The position at the center of the PCR products was used as the value for plotting points in nucleosome positioning maps. Quantification was performed using qPCR as described above. The nucleosomal DNA enrichment was calculated as the log ratio of nucleosomal DNA to that of total genomic DNA. Each 96-well plate contained a primer pair targeting the *PHO5* promoter as a normalization control for plate-to-plate variation. Biological replicates of each sample were analyzed.

The nucleosomal DNA enrichment was plotted against genomic coordinate. Plots were smoothed using a moving average of three adjacent values and normalized to the tallest peak within the promoter. Peaks of PCR signal represent enrichment of DNA segments protected by nucleosomes from micrococcal nuclease digestion, and valleys represent nucleosome-free segments. Peak positions correspond to nucleosome positions, while peak height is a measure for nucleosome occupancy. Peak positions and height were determined using the multi-peak fitting tool of the IGOR Pro 6 data analysis software (WaveMetrics Inc.). Highly reproducible nucleosome position maps were obtained for a control region (the *PHO5* promoter). However, as noted by other researchers using this method (8), we were unable to rigorously establish a linear relationship between nucleosome occupancy and protection from micrococcal nuclease. Therefore, peak heights should be viewed merely qualitatively, and the interpretation of this data is limited to the deduction of nucleosome positions. Each experiment was performed in duplicate. Measured peak positions in Table 1 and Supplementary Tables SVI and SVII are averages of the values obtained in both experiments.

Pearson correlation coefficients were calculated to examine the dependence between variables. We also calculated the Spearman correlation coefficients (data not shown), which in most cases were similar to the Pearson correlation coefficients, except for the correlation between nucleosome positions and *URA3* expression in the *lys2*::*URA3* mutants containing construct *URA3*-163. Here, the Pearson correlation coefficient is skewed because expression at position 721 is exceptionally high; therefore, the reported correlation coefficients do not take into account the values measured for this strain. *P*-values were calculated using the Student's *t*-test.

RESULTS

URA3 as a reporter of local chromatin structure

To investigate the influence of chromatin context on local nucleosome positioning, we inserted the *URA3* gene at various locations in the *S. cerevisiae* genome and explored the consequences for *URA3* nucleosome distribution. *URA3* was chosen because the nucleosome structure of this gene is well characterized and reflects the nucleosome pattern of a typical yeast gene, containing a 5′ NFR that is surrounded by two highly localized nucleosomes (Figure 1A, 2A) (27,32,43). Moreover, expression of *URA3* can be easily estimated by growth on two different substrates: SC-ura, a growth medium lacking uracil and allowing growth only if *URA3* is induced, and 5-FOA, allowing growth in the absence of basal *URA3* transcription (44,45).

The *URA3* promoter contains a poly(dA:dT) barrier sequence (Figure 1A). To determine whether this sequence shields the *URA3* gene from the influence of surrounding chromatin, we created a series of truncated *URA3* constructs in which the *URA3* promoter including the poly(dA:dT) tract, is progressively deleted (5′–3′). The intergenic region between the *URA3* START site and the STOP codon of the first gene upstream (*GEA2*) is 366 bp long (Figure 1A), and progressive deletions resulted in truncated *URA3* constructs with between 265 bp and 100 bp of promoter sequence proximal to the *URA3* START codon remaining. The corresponding mutants are referred to by the length of the remaining truncated *URA*3 promoter (e.g. 'mutant 163' or 'strain 163' only has the 163 bp most proximal to its START codon remaining, see Figure 1).

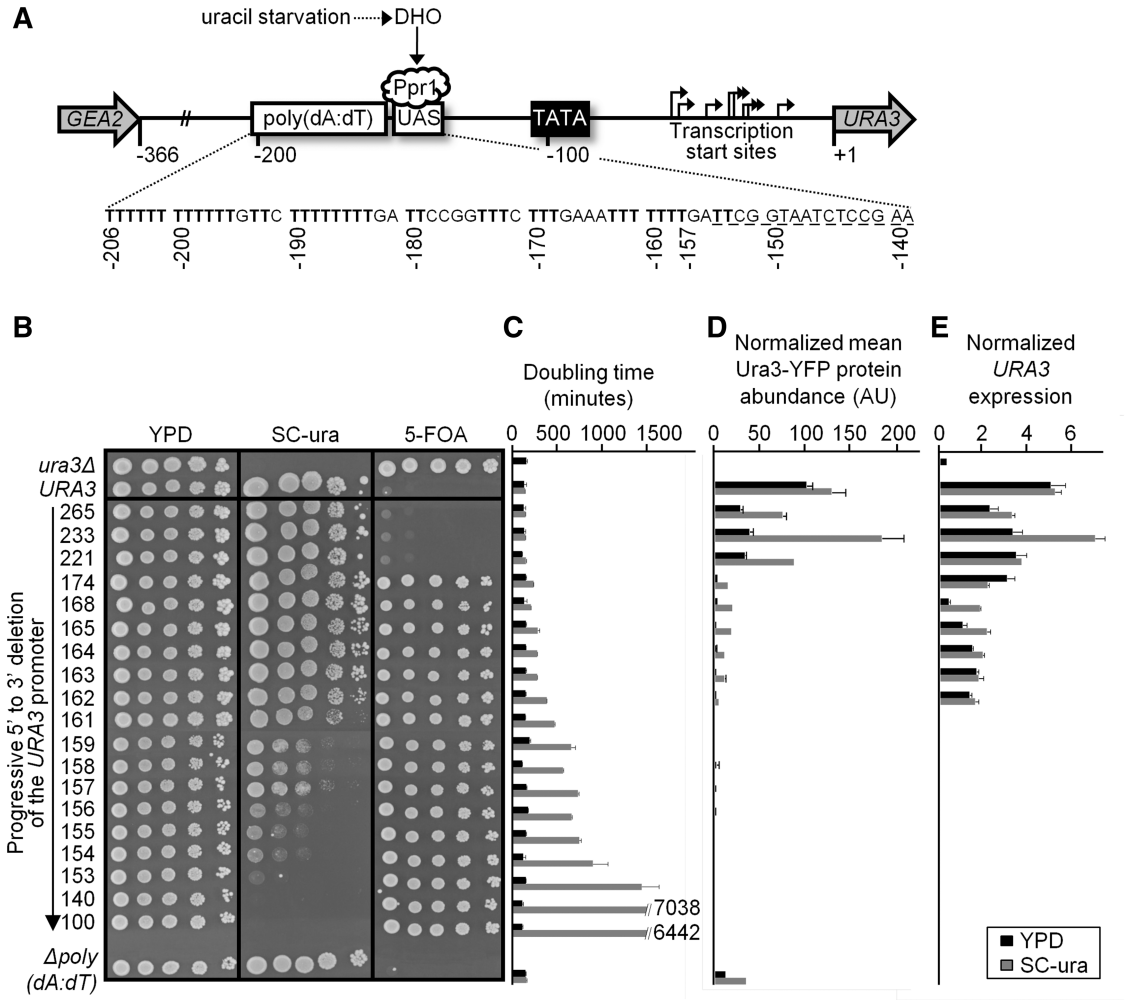

Figure 1. Progressive 5′–3′ deletion of the *URA3* promoter results in a gradual decrease in *URA3* expression. (**A**) The *URA3* promoter has a poly(dA:dT)-rich region required for basal expression between positions −206 and −157, a UAS bound by the transcriptional activator Ppr1 between positions −154 and −139 (dashed underline) and two functionally distinct (constitutive and regulatory) TATA elements between positions −109 and −80. The promoter has eight different transcription initiation sites, consisting of sites inducible by Ppr1 at positions −41, −38, −33, −28 and −19, and constitutive sites at positions −60, −55 and −47. DNA-bound Ppr1 is not transcriptionally active and Ppr1-induced expression requires increased levels of dihydroorotic acid (DHO). Adapted from Ref. (27). (**B**) Truncation of the *URA3* promoter leads to changes in growth on SC-ura and 5-FOA medium. Mutants are labeled by the number of remaining nucleotides in the *URA3* promoter. Δ*poly(dA:dT)* indicates a mutant strain only deleted for the *URA3* promoter poly(dA:dT) sequence. These mutants also show differences in doubling times (**C**), normalized mean Ura3-YFP protein abundance (**D**) and normalized *URA3* mRNA expression (**E**) for growth in liquid YPD (black bars) and liquid SC-ura (gray bars) media. In YPD medium, *URA3* protein and mRNA levels for strains 161–100 did not differ significantly from the negative control. In SC-ura medium, we were unable to obtain reproducible *URA3* protein and mRNA levels for strains 161–100 because of the very poor growth of these strains in this medium. Error bars denote standard deviation.

Progressive 5′–3′ deletion of the *URA3* promoter results in a gradual decrease in *URA3* expression and a perturbed nucleosome pattern

First, we examined a series of truncated *URA3* reporter strains with the *URA3* gene at its native genomic location (Figure 1). *URA3* expression was estimated by growth on different substrates (Figure 1B and C). The reporter strains show only minor growth differences in non-selective YPD medium (Figure 1B and C, black bars). On solid SC-ura medium, progressive promoter reduction results in decreased capability to grow and eventually total loss of growth (Figure 1B). This pattern is confirmed by doubling time measurements in liquid SC-ura medium (Figure 1C, gray bars). Here, the doubling time is initially not affected by the decrease in promoter size. However, upon removal of most of the poly(dA:dT) sequence, further promoter reduction results in a gradual increase of the doubling times, presumably reflecting the inability of these strains to induce *URA3* to sufficiently high levels to support fast growth (Figure 1C, gray bars). On 5-FOA medium, mutants lacking at least half of the poly(dA:dT) tract gain the ability to grow, indicating the absence of basal *URA3* activity (Figure 1B).

To ensure that growth on SC-ura and 5-FOA medium reflects Ura3 activity, we also measured Ura3 protein and *URA3* mRNA levels. The abundance of a Ura3-YFP fusion protein in single cells was determined using flow

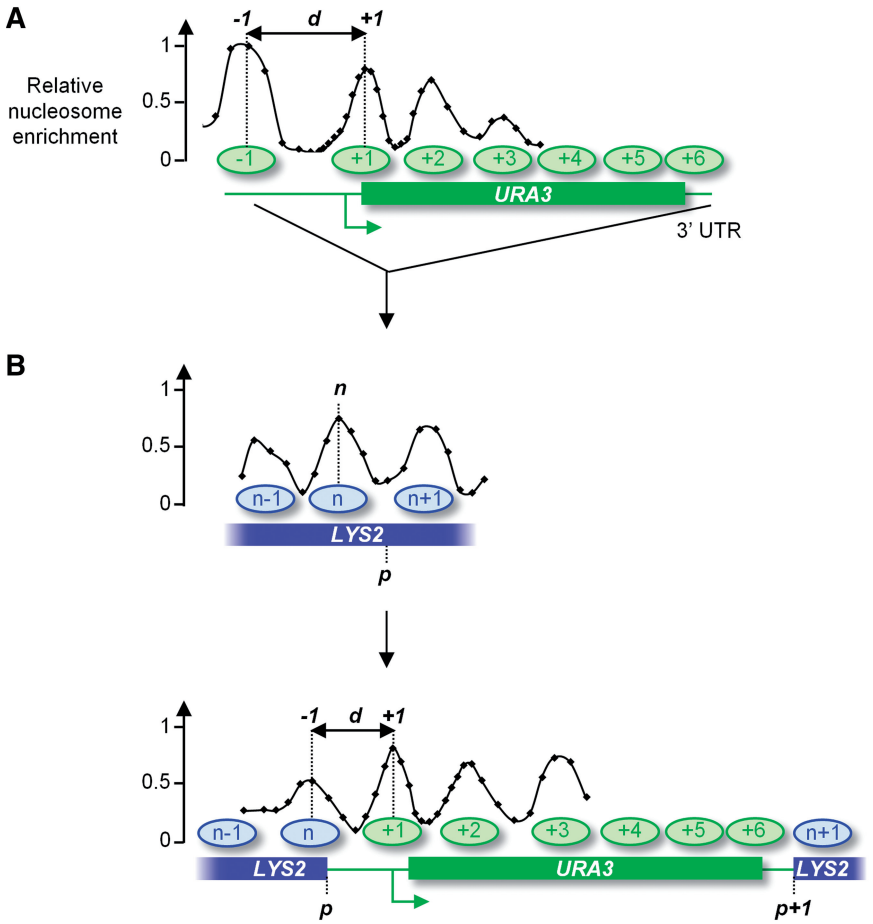

Figure 2. The nucleosome profile of the *URA3* promoter is determined by the positions of the −1 and +1 nucleosomes. (**A**) Using tiling qPCR, we determined the nucleosome profile of the full-length *URA3* promoter at the native *URA3* locus. Peaks of PCR signal represent enrichment of DNA segments protected by nucleosomes from micrococcal nuclease digestion, and valleys represent nucleosome-free segments. Peak positions corresponding to the centers of the −1 and +1 nucleosomes were calculated (see 'Materials and Methods' section), as well as the distance *d* between the −1 and +1 nucleosome centers (Table 1). (**B**) The *URA3* gene including the (truncated) promoter was inserted at various locations *p* in the *LYS2* gene. Before insertion, *n* is the position of the center of the first nucleosome upstream of the insertion site *p*. This upstream nucleosome might influence the positions of the −1 and +1 nucleosomes of the *URA3* promoter after insertion. Coordinates for *n* are relative to the insertion site *p*. After insertion, the positions of the centers of the −1 and +1 nucleosomes were determined, as well as the distance *d* between the −1 and +1 nucleosome centers (Table 1). Coordinates for −1 and +1 are relative to the translation start site (ATG) of *URA3*.

cytometry (Figure 1D). *URA3* mRNA levels were measured using quantitative PCR (qPCR) (Figure 1E). The data obtained with both methods are strongly correlated ($R = 0.86$, $P < 0.1$ in YPD and $R = 0.99$, $P < 0.05$ in SC-ura). In YPD medium, Ura3 levels are generally low, and no protein or mRNA is measured in the mutants that have lost basal *URA3* expression, which agrees with their ability to grow on 5-FOA medium (Figure 1B and D, black bars). Higher Ura3 levels are found when cells were grown in SC-ura medium. In keeping with the growth assays on SC-ura, *URA3* expression decreases gradually with decreasing promoter length (Figure 1C and D, gray bars). Taken together, progressive 5′–3′ deletion of the *URA3* promoter causes gradual changes in *URA3* expression.

Next, we examined nucleosome positioning in the wild-type *URA3* promoter and a subset of truncated promoter mutants using tiling qPCR (Figure 2A, Supplementary Figure S1, Table 1). As illustrated in Figure 2A, peaks of PCR signal represent enrichment of DNA segments, while valleys represent nucleosome-free segments. Peak positions corresponding to the centers of the −1 and +1 nucleosomes were calculated (Figure 2A, Supplementary Figure S1, Table 1) (see 'Materials and Methods' section). In addition, the distance *d* between the −1 and +1 nucleosome centers was determined (Table 1). The native *URA3* promoter contains a 5′ NFR of ~150 bp surrounded by well-positioned nucleosomes, thereby adopting the nucleosome distribution of a typical yeast promoter (Figure 2A, Supplementary Figure S1, Table 1) (11,13,43). Using two sequence-based computational models (15,21), we first predicted the nucleosome positions of the mutant *URA3* constructs (Supplementary Figure S2). Both models predict the NFR to gradually shorten and become more nucleosome-occupied when the *URA3* promoter undergoes progressive deletions. Our experimental data confirm these predictions, and also indicate that nucleosome positions in the *URA3* promoter become less well-defined and fuzzier in these constructs (Supplementary Figure S1, Table 1).

Table 1. Positions of the *URA3*-1 and +1 nucleosomes

Insertion site[a]	Promoter length[b]	Position −1 nucleosome[c]	Position +1 nucleosome[c]	Distance[d]
native	366	−303.12 ± 2.50	−5.59 ± /-9.75	297.53+/-12.25
native	163	Noisy	Noisy	n/a
native	162	Noisy	Noisy	n/a
native	161	Noisy	Noisy	n/a
native	Δ*poly(dA:dT)*	−244.88 ± 0.04	−59.81 ± 3.42	185.07 ± 3.46
LYS2-721	366	−472.83 ± 3.26	−1.95 ± 1.49	470.88 ± 4.74
LYS2-721	221	−272.97 ± 2.91	−33.01 ± 2.43	239.96 ± 5.33
LYS2-721	174	−290.38 ± 3.03	Noisy	n/a
LYS2-721	168	−277.97 ± 3.14	−44.13 ± 4.08	233.84 ± 7.22
LYS2-721	167	−270.31 ± 3.41	−58.70 ± 8.63	211.61 ± 12.04
LYS2-721	165	−276.00 ± 1.82	−59.65 ± 2.23	216.35 ± 4.05
LYS2-721	164	−282.89 ± 3.30	−48.15 ± 3.22	234.74 ± 6.52
LYS2-721	163	−259.29 ± 2.26	−34.51 ± 2.16	224.78 ± 4.42
LYS2-721	162	−266.41 ± 5.22	Noisy	n/a
LYS2-721	161	−270.91 ± 4.95	Noisy	n/a
LYS2-721	Δ*poly(dA:dT)*	−404.02 ± 16.43	−69.57 ± 1.23	334.45 ± 15.20
LYS2-450	163	−196.27 ± 3.14	−38.50 ± 10.12	157.76 ± 13.26
LYS2-500	163	−233.22 ± 3.15	−46.22 ± 1.53	187.00 ± 4.68
LYS2-540	163	−284.30 ± 4.23	−57.64 ± 1.11	226.66 ± 5.34
LYS2-575	163	Noisy	Noisy	n/a
LYS2-721	163	−259.29 ± 2.26	−34.51 ± 2.16	224.78 ± 4.42
LYS2-770	163	Noisy	−24.56 ± 9.38	n/a
LYS2-800	163	−229.08 ± 6.38	−30.11 ± 3.80	198.97 ± 10.18
LYS2-860	163	−284.38 ± 4.24	−24.98 ± 1.69	259.40 ± 5.93
LYS2-865	163	−270.06 ± 3.63	−46.28 ± 3.25	223.77 ± 6.88
LYS2-1950	163	−261.42 ± 3.01	−24.98 ± 5.86	236.44 ± 8.87
LYS2-2200	163	−209.71 ± 11.58	−45.93 ± 5.75	163.78 ± 17.33
LYS2-3050	163	−223.66 ± 2.97	−41.43 ± 1.32	182.24 ± 4.29

[a] Insertion site of the *URA3* reporter gene, at the native *URA3* locus ('native') or at the *LYS2* gene (e.g. '*LYS2*-721' indicates an insertion site is located 721 bp downstream of the *LYS2* START site).
[b] Length of the remaining truncated *URA3* promoter (e.g. '162' indicates the promoter has the 162 bp most proximal to its START codon remaining). Δ*poly(dA:dT)* indicates the deletion of the poly(dA:dT) sequence only.
[c] Position of the nucleosome center relative to the *URA3* START site.
[d] Distance between the −1 and +1 nucleosome centers (Figure 2). The NFR width can be calculated by subtracting 147 bp from the distance.

Poly(dA:dT) functions as a border element for nucleosome positioning and is required for high basal and induced *URA3* expression

The previous results show that in mutants with progressive truncations of the *URA3* promoter, *URA3* expression and promoter nucleosome positions were only affected after the poly(dA:dT) element was at least partly deleted. To investigate directly the role of the poly(dA:dT) tract in the *URA3* promoter, we constructed a mutant in which the 50 bp poly(dA:dT) tract was removed (without removing any other *URA3* promoter sequences). Deleting the poly(dA:dT) tract causes a large decrease in *URA3* expression. However, the cells still manage to induce *URA3* transcription to sufficient levels to allow growth on medium lacking uracil (SC-ura), and basal expression is still too high to allow growth on 5-FOA medium (Figure 1B, D and E). An identical poly(dA:dT) deletion was constructed in a strain containing the *URA3* gene at another genomic location, and the results obtained for this strain were similar to the results reported for the native *URA3* locus (Figure 3A and C, bottom row).

To investigate whether changes in chromatin structure might explain the observed change in *URA3* activity, we determined the nucleosome positions over the *URA3* promoter in the Δ*poly(dA:dT)* mutants (Supplementary Figure S1E, Figure 4L, Table 1). Deletion of the poly(dA:dT) tract greatly reduces the NFR. The −1 and +1 nucleosomes have moved towards each other, leaving only a short linker region in which the UAS is located, presumably leaving this site accessible. The two TATA boxes and the different TSSs are now covered by the +1 nucleosome, presumably making them less accessible. Thus, the poly(dA:dT) tract appears to help the formation of the NFR and for maintaining the accessibility of several *cis* regulatory elements of the *URA3* promoter. These results confirm that the poly(dA:dT) tract functions as an important border element for directing nucleosome positions in the *URA3* promoter, which in turn affects expression of the *URA3* gene.

Although removal of the poly(dA:dT) tract causes important shifts in nucleosome positioning, the *URA3* promoter still has highly localized nucleosomes, indicating that other elements also contribute to the formation of these nucleosomes. The presence of highly localized nucleosomes after deletion of the poly(dA:dT) element is also predicted by sequence-based computational models (Supplementary Figure S2), indicating that the signals to position these nucleosomes may be encoded in the local DNA sequence.

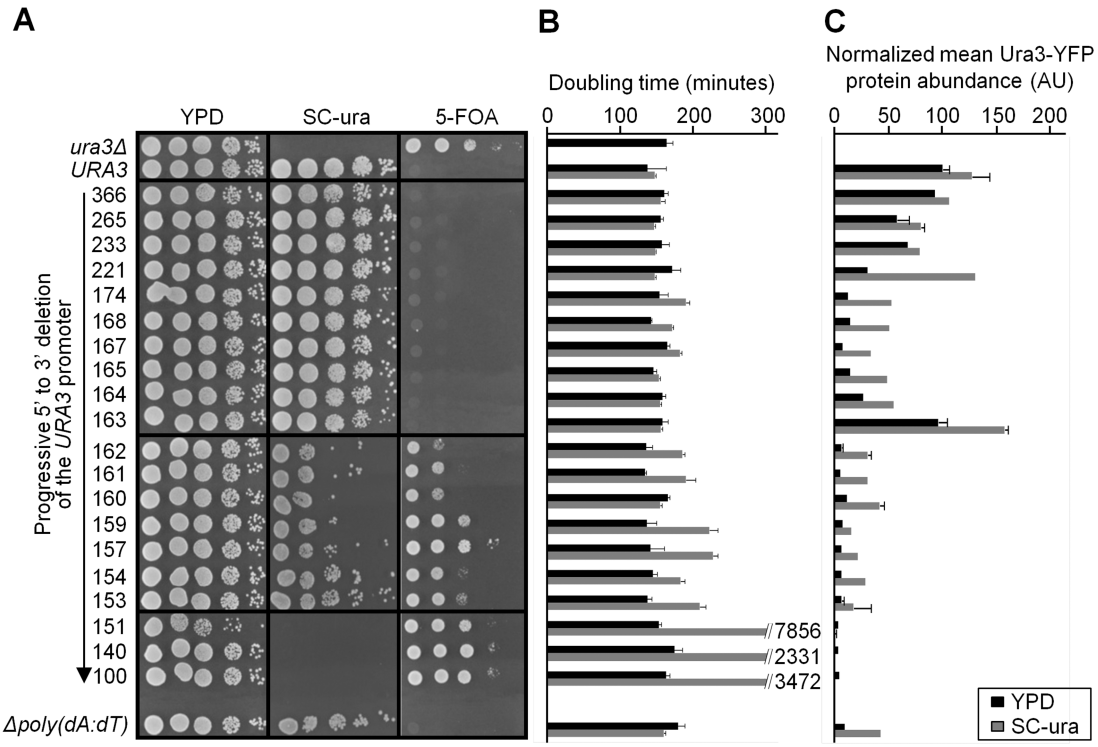

Figure 3. Growth and *URA3* expression in the *lys2::URA3* promoter mutants. The truncated *URA3* constructs were inserted into the *LYS2* gene, 721 bp downstream of the *LYS2* translational start site (ATG). 5′ to 3′ progressive deletion mutants are labeled by the number of remaining nucleotides in the *URA3* promoter. Δ*poly(dA:dT)* indicates a mutant strain only deleted for the *URA3* promoter poly(dA:dT) sequence. (**A**) Growth on solid YPD, SC-ura and 5-FOA media. (**B**) Doubling times, and (**C**) normalized mean Ura3-YFP protein abundance for growth in liquid YPD (black bars) and liquid SC-ura (gray bars) media. Error bars denote standard deviation.

Expression of truncated *URA3* constructs depends on genomic location

To investigate if the expression profile and nucleosome structure of the various truncated *URA3* reporters depend on the surrounding chromatin context, we relocated the reporters to a different genomic location. The *URA3* constructs were inserted into the *LYS2* gene, 721 bp downstream of the *LYS2* START codon. The *LYS2* ORF was chosen as an insertion site because it shows a gradual change from well-positioned nucleosomes near the START codon to more fuzzy, less-positioned nucleosomes further downstream, offering multiple 'chromatin environments' in which the *URA3* constructs can be inserted (see further). To exclude the possibility that *LYS2* transcription affects *URA3* expression and the nucleosome profile of the *URA3* promoter, we examined these phenotypes in a *lys2::URA3* mutant in which the *LYS2* promoter was deleted and *LYS2* transcription was abolished (Supplementary Figure S3, Supplementary Table SVI).

We examined the growth profiles of 20 different truncated *URA3* promoter mutants integrated at position 721 of the *LYS2* gene (Figure 3A and B). All strains display similar growth patterns in non-selective YPD medium (Figure 3A and B, black bars). Based on their growth on selective media, the mutants can be divided into three groups: (i) mutants 265-163 grow on SC-ura medium but not on 5-FOA medium, indicating *URA3* is being induced as well as basally expressed, (ii) mutants 162-153 grow on SC-ura medium and gain the ability to grow on 5-FOA medium, indicating a reduction in basal *URA3* expression and (iii) mutants 151-100 fail to grow on SC-ura medium while displaying growth on 5-FOA medium, indicating the absence of both basal and induced *URA3* expression (Figure 3A). Interestingly, growth on SC-ura for the constructs at this ectopic location differs from the growth observed at the native locus (Figure 1B and C and Figure 3A and B). Instead of a gradual increase, the doubling time remains fairly constant in the mutants of the first group, only to increase slightly in the mutants of the second group. Abrupt transition from growth to no growth distinguishes the second and the third group of mutants (Figure 3A and B).

We chose to investigate *URA3* expression by measuring Ura3 protein concentrations using flow cytometry rather than by determining mRNA levels. We already showed that Ura3 protein and *URA3* mRNA measurements strongly correlate (R between 0.86 and 0.99, see above), and we found that protein measurements are more robust in strains that express *URA3* at extremely low concentrations. Moreover, determining protein levels on a single cell basis allows us to examine differences between individual cells in a population (so-called expression noise, see below).

The growth patterns of the different mutants correlate with their intracellular Ura3 protein levels (Figure 3C). In

non-selective YPD medium and in inducing SC-medium, we observe the same general trend. First, Ura3-YFP levels drop more or less gradually (mutants 366–167), after which they increase again (mutants 165–163). Mutants with even shorter *URA3* promoters (mutants 162–100) show a dramatic drop in Ura3-YFP levels (Figure 3C). Interestingly, in a number of mutants (e.g. strains 167 and 162–160), *URA3* gene expression in YPD is almost the same, but these strains show markedly different growth on 5-FOA medium (Figure 3A and C). This indicates that a narrow difference in *URA3* expression (leading to different Ura3 levels) may distinguish strains growing on 5-FOA medium from strains not growing on 5-FOA medium. In addition, it is possible that the observed ability to grow on 5-FOA medium is a result of an increased variation in Ura3 levels between individual cells within a population [so-called expression noise, (46)]. Such an increase in expression noise could result in a subset of the cells no longer basally expressing *URA3* (allowing growth on 5-FOA medium), whereas others still show basal expression, explaining the significant levels of Ura3 present in YPD medium. However, the intracellular Ura3-YFP signals do not show a difference between noise levels in the strains with shorter promoters (162 and shorter) compared to those with longer promoters (163 and up) (see Supplementary Figure S4).

Nucleosome positions of truncated *URA3* constructs depend on genomic location

To investigate if the chromatin structure of the various mutants depends on the genomic location of the *URA3* marker, we determined the nucleosome positions of various truncated *URA3* promoters when inserted at nucleotide 721 of the *LYS2* locus (Figure 4, Table 1). For the full-length *URA3* promoter (mutant 366), we find a NFR surrounded by two well-positioned nucleosomes. Due to the altered genomic context upstream of the *URA3* promoter, the position of the −1 nucleosome changes, but the position of the +1 nucleosome remains the same. As a consequence, the NFR doubles in size (Figure 4B, Table 1). However, this does not affect *URA3* expression (Figure 3), presumably because the accessibility of the functional *cis* elements in the *URA3* promoter is not affected.

The experimentally determined nucleosome positions for the full-length *URA3* promoter are consistent with the predictions of the computational models, although the experimentally determined NFR is smaller than what is predicted (Supplementary Figure S5) (15,21). For the truncated *URA3* constructs inserted at the *LYS2* locus, the models predict a gradual shrinking of the 5′ NFR associated with the progressive deletion of the *URA3* promoter (Supplementary Figure S5). Although we observe a similar trend in our qPCR data, the actual placing of the predicted nucleosomes is not identical to the positions that were experimentally determined in this study. We believe that at least part of this discrepancy may be explained by differences in methodology and data analysis (experimental differences, e.g. in MNase digestion; quantitative PCR data versus next generation sequencing).

The predictions made by the models are similar (but not identical) to the predictions made for the mutants at the native *URA3* locus (Supplementary Figure S2 and S5). In the Kaplan *et al*. model, the −1 and +1 nucleosomes can be clearly identified and the predictions show that the −1 nucleosome gradually moves inward while the position of the +1 nucleosome does not alter (Supplementary Figure S5A–H). In the predictions made by the Field *et al*. model, individual nucleosomes are more difficult to distinguish, especially in the strains with short *URA3* promoters, which appear to be characterized by a high overall nucleosome occupancy (Supplementary Figure S5I–P).

Experimental examination of the nucleosome profile shows the presence of a highly positioned −1 nucleosome in all of the strains (Figure 4, Table 1). Interestingly, the −1 nucleosomes occupy similar (but not identical) positions relative to the *URA3* translation start site (ATG) in all of the mutants. The positions occupied by the −1 nucleosomes in the truncated promoter mutants are shifted downstream relative to the position occupied by the −1 nucleosome in the native, full-length *URA3* promoter (average downstream shift is 29 nt; Figure 4, Table 1). It should be noted that the sequences incorporated in the −1 nucleosome are different in each of the mutants; indeed, as the *URA3* promoter shortens, the −1 nucleosome gradually moves upstream further into the *LYS2* gene, and as a result, the position of this nucleosome relative to the *URA3* translation start site remains similar. Thus, it appears that the position of this nucleosome is not determined by the upstream *LYS2* sequence, but rather by the *URA3* sequence, downstream *URA3* nucleosomes, and/or upstream *LYS2* nucleosomes. In addition, the size of the truncated promoter or the presence of a (partial) poly(dA:dT) element do not affect the position of the −1 nucleosome.

In most strains, a clearly positioned +1 nucleosome can also be detected (Figure 4, Table 1). However, in the mutants with very short *URA3* promoters (strains 162 and 161), the region normally occupied by the +1 nucleosome shows relatively low nucleosome occupancy. Thus, it appears that the gradual shortening of the *URA3* promoter eventually results in the loss of a (positioned) +1 nucleosome. In addition, we could not detect the presence of a +1 nucleosome in mutant 174 (Figure 4, Table 1). Where a positioned +1 nucleosome is present, our data show that it has shifted upstream towards the −1 nucleosome (Figure 4, Table 1).

Thus, insertion of the various *URA3* mutants in a non-native genomic location causes shifts in the nucleosome pattern compared to the pattern observed when these mutants are located at the native *URA3* site. Gradual truncation of the *URA3* promoter at the non-native *LYS2* site causes relocation of the −1 and +1 nucleosomes, resulting in the shortening of the 5′ NFR (Figure 4, Table 1).

Nucleosome positions of truncated *URA3* constructs determine *URA3* expression

Do the nucleosome profiles explain *URA3* expression patterns and the corresponding Ura3 protein levels? In

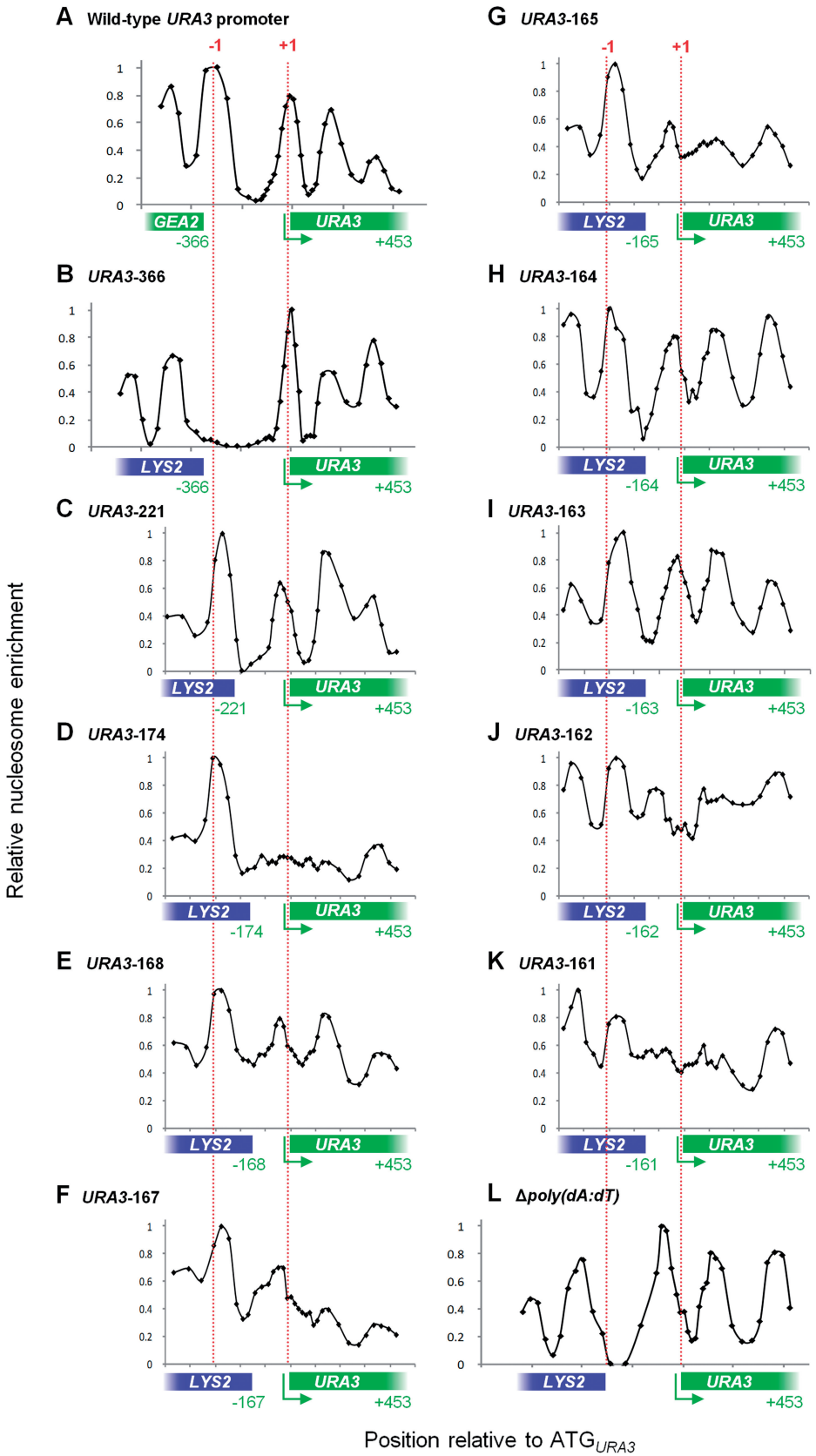

Figure 4. Nucleosome positions of the native *URA3* promoter and selected *lys2::URA3* promoter mutants. Using tiling qPCR, we determined the nucleosome positions of (**A**) the wild-type *URA3* promoter at its native locus (same as Figure 2A). In addition, we determined the nucleosome profile of selected *lys2::URA3* promoter mutants, i.e. mutants (**B**) 366, (**C**) 221, (**D**) 174, (**E**) 168, (**F**) 167, (**G**) 165, (**H**) 164, (**I**) 163, (**J**) 162, (**K**) 161 and (**L**) Δ*poly(dA:dT)*. Peak positions corresponding to nucleosome positions were measured (Table 1). Red dotted lines mark the positions of the −1 and +1 nucleosomes in the wild-type *URA3* promoter at its native locus.

the strains with a positioned +1 nucleosome, there is a strong correlation between the position of this nucleosome, as measured in non-selective YPD medium, and *URA3* expression in YPD medium. The more this nucleosome shifts upstream, the lower *URA3* expression. ($R = 0.80$, $P < 0.0005$) (Figure 3C, Table 1). This is most likely the result of an altered accessibility of the TSSs. In the native *URA3* promoter, the TSSs reside inside but close to the border of the +1 nucleosome. However, in the truncated *URA3* constructs, the +1 nucleosome shifts upstream and the TSSs become incorporated in the +1 nucleosome, presumably resulting in a decreased accessibility. In the strains that lack a (positioned) +1 nucleosome, and specifically in those strains with the shortest promoters (mutants 162 and 161), only very low *URA3* expression can be measured, both in non-selective (YPD) and selective (SC-ura) media (Figure 3C, Table 1). In addition, these strains have gained the ability to grow on 5-FOA medium, indicating the absence of basal *URA3* expression (Figure 3A). These data indicate that the presence of a positioned +1 nucleosome is required for efficient *URA3* expression.

Our data also show a moderate correlation between the position of the −1 nucleosome and *URA3* expression, both measured in YPD medium ($R = -0.57$, $P < 10^{-5}$) (Figure 3C, Table 1). The positions of the −1 nucleosome only vary modestly in the different truncated constructs, and overall, the distance to the *URA3* translation start remains similar (Table 1). Therefore, in contrast to the +1 nucleosome, minor changes in the position of the −1 nucleosome are not expected to exert large effects on *URA3* expression.

In strains containing a truncated *URA3* construct with a positioned +1 nucleosome, the distance between the −1 and +1 nucleosomes is similar, with a clear NFR in-between the −1 and +1 nucleosomes (Table 1). Although this NFR is only about half the size of the NFR of the full-length *URA3* promoter at the native locus, the Ppr1 binding site and the TATA boxes are accessible in all strains with a positioned +1 nucleosome. The size of the NFR is comparable in all mutant strains, and the width of the distance between the −1 and +1 nucleosomes also shows a moderate correlation with *URA3* expression ($R = 0.62$, $P < 10^{-4}$).

For a selection of mutants, we determined the *URA3* promoter nucleosome positions under different inducing (SC-ura) and non-inducing (YPD, 5-FOA) conditions (Supplementary Figure S6, Supplementary Table SVII). These data show that the −1 and +1 nucleosomes are organized independent of *URA3* promoter induction. This is not entirely unexpected because the transcriptional activator Ppr1 is always bound to the *URA3* promoter (induction occurs when DNA-bound Ppr1 is activated by dihydroorotic acid, see Figure 1A). These results confirm the results obtained in rich (YPD) medium and further indicate that a positioned +1 nucleosome as well as a clear NFR are required for efficient *URA3* expression.

Together, these results show that *URA3* expression correlates with the nucleosome organization of the promoter. The environment around position 721 of the *LYS2* ORF influences nucleosome positioning in the truncated *URA3* inserts, suggesting that long-range effects play a role in determining nucleosome positioning and gene expression.

Genomic locus determines growth and expression patterns of *lys2*::*URA3* reporter strains

The previous results show that progressive deletion of the *URA3* promoter alters *URA3* expression and chromatin structure, and that these changes are different when *URA3* is present in its native genomic location, or when it is inserted ectopically at position 721 of the *LYS2* gene.

To further investigate the influence of the genomic environment on local *URA3* activity and chromatin structure, we inserted a truncated *URA3* promoter into twelve different locations within the *LYS2* gene (Figure 5A). Construct *URA3*-163 (i.e. the 163 bp most proximal to the START codon remain, see Figure 1A) was used because our previous results indicate that this construct is sensitive to the surrounding chromatin in which it is embedded. In parallel, the same experiments were performed with construct *URA3*-162 (see Supplementary Figures S7, S9, S10 and S13 and Table SVI).

The insertion sites were chosen based on the published nucleosome positions in the *LYS2* gene (12) and confirmed in this study using tiling qPCR (Figure 5A). More specifically, we inserted the reporter constructs in sites occupied by positioned nucleosomes, sites at the edges of or in between positioned nucleosomes, and sites depleted for nucleosomes. Nine sites (450–865) are located within 1 kb from the well-positioned +1 nucleosome that borders the 5′ NFR in the *LYS2* promoter. The nucleosomes in this region are highly localized. At sites 450, 860 and 865 the *URA3* gene is inserted at a nucleosome 'peak', indicating that a nucleosome is positioned here. Sites 540, 721 and 770 on the other hand are located in a 'valley', a linker region where no nucleosome is present. The other sites (500, 575 and 800) are located at the transition between a peak and a valley, and indicate a site occupied by a nucleosome, but close to one of the nucleosome borders. A final group of insertion sites (1950, 2200 and 3051) are located well beyond 1 kb from the highly localized +1 nucleosome in a region with low nucleosome occupancy, likely containing delocalized or fuzzy nucleosomes, i.e. nucleosomes occupying different locations in the genomes of the cells in a population (Figure 5A).

For each insertion site, we characterized growth of the resulting strains (Figure 5B and C). All mutants grow similarly on (non-selective) YPD medium, although small variations in doubling time can be observed (Figure 5B and C, black bars). In SC-ura medium (selecting for induction of *URA3*), several strains exhibit a doubling time comparable to the doubling time measured in YPD medium, whereas other strains show moderately increased doubling times (Figure 5B,C, grey bars). Whereas growth differences in SC-ura medium are subtle, two strains exhibit very different growth patterns on 5-FOA medium, a measure for basal *URA3* expression (Figure 5B). Strains that exhibit the longest doubling times on SC-ura medium are able to grow on 5-FOA medium (Figure 5B anC, grey bars). Ura3 protein levels (Figure 5D) agree with the observed growth patterns. As

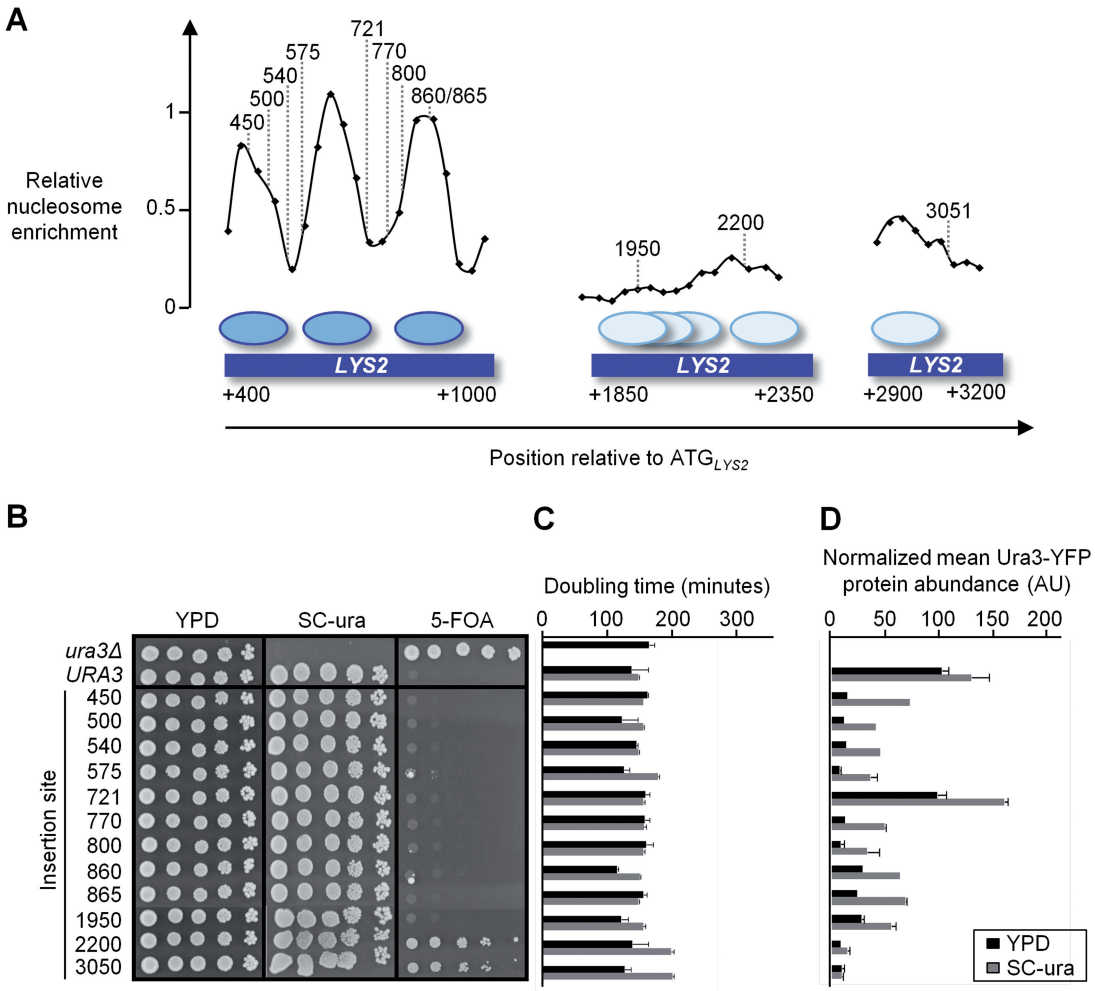

Figure 5. Expression of a truncated *URA3* gene depends on the genomic context. A truncated *URA3* construct in which 163 bp of the *URA3* promoter remain (*URA3*-163), was inserted at various locations in the *LYS2* gene. (**A**) Different insertion sites of the truncated *URA3*-163 construct within the *LYS2* ORF. The nucleosome profile for regions of the *LYS2* gene as it was experimentally determined by tiling qPCR is shown. Dotted lines indicate insertion sites which are labeled by their location inside the *LYS2* ORF (coordinates refer to the position relative to the *LYS2* translation start). (**B**) Different insertion sites lead to different growth patterns on SC-ura and 5-FOA medium. Mutants are labeled by their location inside the *LYS2* ORF. These mutants also show differences in doubling times (**C**), and normalized mean Ura3-YFP protein abundance (**D**) for growth in liquid YPD (black bars) and liquid SC-ura (gray bars) media. Error bars denote standard deviation.

expected, the strains that are able to grow on 5-FOA medium show very low Ura3-YFP protein abundance (Figure 5B and D). Taken together, for each reporter construct (*URA3*-163 and *URA3*-162, see Supplementary Figure S7), the growth and *URA3* expression phenotypes vary greatly between the different insertion sites.

Nucleosome organization in the *lys2*::*URA3* reporter strains correlates with expression

To determine whether the genomic location of the *URA3* reporter constructs influences the nucleosome positions in the *URA3* promoter, and if these differences correlate with the differences in *URA3* expression, we ascertained the nucleosome positions of all the *lys2*::*URA3* mutants described in the previous section (Supplementary Figure S8, Table 1). The majority of mutants have highly localized nucleosomes occupying the *URA3* promoter, although some have fuzzy −1 (loci 575, 770) and/or +1 (locus 575) nucleosomes. When we compare the positions of the localized −1 and +1 nucleosomes relative to their positions in the full-length *URA3* promoter at its native locus, we observe a shift of both nucleosomes towards the NFR.

Can *URA3* expression levels in the different *lys2*::*URA3* mutants be explained by the altered nucleosome positions? The position of the +1 nucleosome correlates poorly with *URA3* expression ($R = 0.40$, $P < 10^{-9}$), while the position of the −1 nucleosome has a moderate impact ($R = -0.67$, $P < 5 \times 10^{-8}$). However, our data shows that *URA3* expression is most strongly affected by the width of the NRF ($R = 0.80$, $P < 5 \times 10^{-8}$). To further explore the correlation between nucleosome positioning and *URA3* expression, we investigated another series of mutants, which contain the *URA3*-162 inserted at 12 different loci in the *LYS2* gene (Supplementary Figures S7, S9, Supplementary Table SVI). Here, we obtain highly

similar data, confirming that *URA3* expression correlates poorly with the position of the +1 nucleosome ($R = 0.33$, $P < 10^{-8}$), moderately with the position of the −1 nucleosome ($R = -0.59$, $P < 5 \times 10^{-8}$) and strongly with the size of the NFR ($R = 0.77$, $P < 5 \times 10^{-7}$). Hence, for the different insertion sites in *LYS2*, the width of the NFR is the major determinant of *URA3* expression, while the position of the +1 nucleosome has only a small impact.

These results are different from the data we obtained for the progressively truncated *URA3* constructs at locus 721 (see above). At locus 721, we found that *URA3* expression was most strongly influenced by the position of the +1 nucleosome. When we take a closer look at the data in Table 1, we notice that the extent of variation in the position of the +1 nucleosome is similar in both series of mutants (i.e. *LYS2*-721 and *URA3*-163). In contrast, the width of the NFR varies greatly depending on the insertion site, while at locus 721, the NFR size remains fairly stable (Table 1). This might explain why NFR width only moderately affects expression at locus 721, while having a large impact at other loci. Together, the data obtained for the mutants at locus 721 and the mutants at the different insertion sites indicate that the size of the NFR executes a larger effect on *URA3* expression than the position of the +1 nucleosome.

We also examined whether the nucleosome positions of the *URA3*-163 and *URA3*-162 *lys2::URA3* constructs could be predicted by the existing computational models (Supplementary Figure S10; 15,21). Here, it is hard to compare the *overall nucleosome occupancy* predicted by the models with the *exact nucleosome positions* determined in our analysis. However, by comparing NFR width, it becomes clear that the models often fail to accurately predict a nucleosome-depleted region that is comparable in size to the NFRs measured in this study. This disparity might indicate that the experimentally determined (*in vivo*) nucleosome positions may be influenced by *trans* factors that are not accounted for in the computational models. However, we did not find any evidence for changes in *trans*-factor binding sites that may be created or deleted in the in the *URA3*-162 and *URA3*-163 constructs (Supplementary text, Supplementary Figures S11, S12, Supplementary Tables SVI, SVIII).

Nucleosome organization in the *lys2::URA3* reporter strains depends on chromatin context

Finally, we examined whether the chromatin context of the insertion site determines the positions of the −1 and +1 nucleosomes in the *lys2::URA3* reporter strains. Two parameters were used to study the effects of chromatin context. The first parameter is the position of nucleosome *n*, the nucleosome immediately upstream of each insertion site before insertion of the *URA3* gene (Figure 2B). This nucleosome acts as the −1 nucleosome of the *URA3* promoter after insertion. Our data show that there is a good correlation between the position of the -1 nucleosome after insertion and the position of nucleosome *n* before insertion ($R = -0.74$, $P < 5 \times 10^{-5}$) (Figure 6A). This result may suggest that the position of nucleosome *n* does not change upon insertion of *URA3*. However, when *URA3* is inserted at or close to the center of nucleosome n, this nucleosome moves upstream, away from the *URA3* promoter. On the other hand, when *URA3* is inserted close to the nucleosome border or in a region devoid of nucleosomes, nucleosome n shifts downstream, towards the *URA3* promoter.

In addition, nucleosome *n* sometimes also affects the position of the *URA3* +1 nucleosome, even though the effect is relatively weak (but statistically significant, $R = 0.38$, $P < 5 \times 10^{-4}$) (Figure 6B). Finally, the distance between the −1 and +1 nucleosomes is strongly influenced by the position of nucleosome *n* ($R = 0.82$, $P < 0.005$) (Figure 6C). The same analysis was performed for *URA3*-162 and similar but (slightly) weaker correlation coefficients were obtained (Supplementary Figure S13A–C).

Second, we investigated whether the AT content at the junction between the upstream *LYS2* sequence and the *URA3* promoter after insertion of the *URA3* gene correlates with nucleosome positioning. For each of the insertion sites, we calculated the %AT content at the junction between the upstream *LYS2* sequence and the truncated *URA3* promoter (Table 2). We also checked whether a new, long poly(dA:dT) tract was formed at the junction, but this was not the case. We found large variations in %AT content, between 56% and 80%, and we plotted %AT content against the NFR size and the positions of the −1 and +1 nucleosomes after insertion of the truncated *URA3*-162 promoter at various locations of the *LYS2* gene (Figure 6D–F). The results show that AT content has a mild effect on the position of the −1 and +1 nucleosomes, and the distance between the −1 and +1 nucleosomes ($R = -0.33$, $P < 5 \times 10^{-10}$, $R = -0.28$, $P < 10^{-9}$ and $R = 0.25$, $P < 5 \times 10^{-7}$, respectively) (Figure 6D–F). Stronger correlations were obtained for construct *URA3*-162, with the exception of nucleosome +1 (Supplementary Figure S13D–F).

DISCUSSION

In this study, we used truncated versions of the *S. cerevisiae URA3* gene to investigate the influence of the neighboring chromatin structure on local nucleosome positioning and gene expression. Our results show that progressive 5′–3′ truncation of the *URA3* promoter increases the sensitivity of the promoter to the influence of surrounding chromatin. In these strains, the nucleosome positions change and become dependent on the genomic location. In turn, the altered nucleosome positions affect *URA3* activity.

When the poly(dA:dT) tract is removed without removing any up- or downstream sequence, the NFR in the *URA3* promoter is no longer present and *URA3* expression decreases dramatically. This indicates that the poly(dA:dT) tract helps to establish the NFR and optimal *URA3* expression. Although this result agrees with previous findings (5,28), we find that the *URA3* promoter in the *Δpoly(dA:dT)* strain still retains highly localized nucleosomes. Thus, although the poly(dA:dT) tract directs the formation of the NFR, other elements

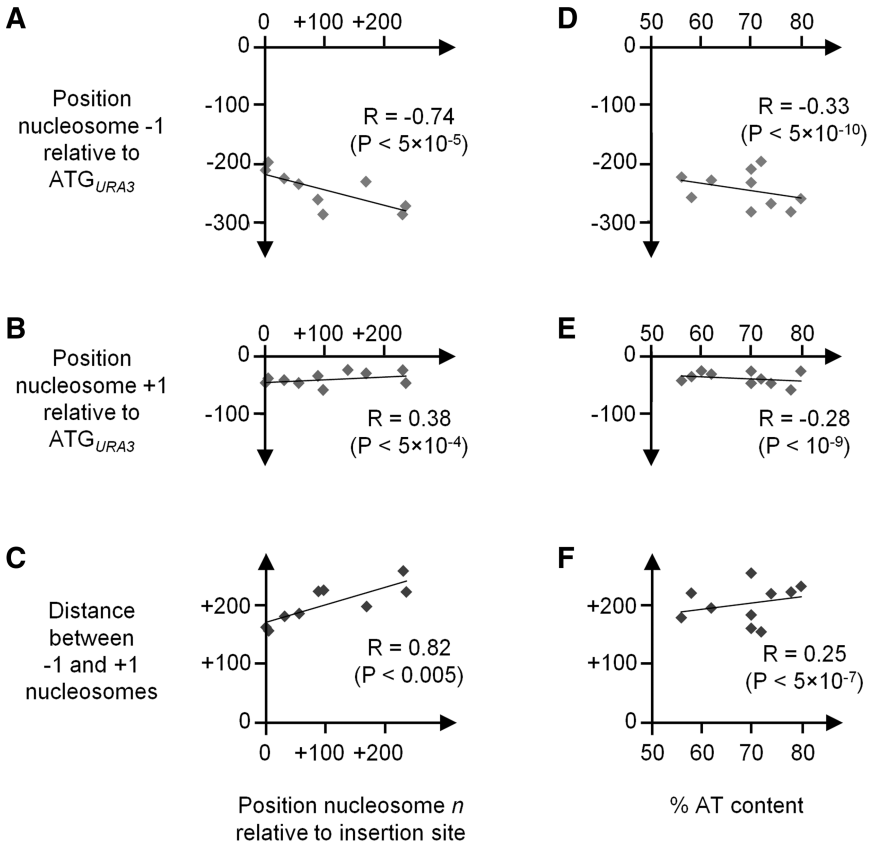

Figure 6. Genomic context influences the positions of the −1 and +1 nucleosomes of a truncated *URA3* promoter. The genomic context of each insertion site was quantified by measuring the position of the first nucleosome *n* upstream of each insertion site before insertion of the *URA3* gene (Figure 2B) and %AT content at the junction between the upstream *LYS2* sequence and the *URA3* promoter. The positions of the −1 nucleosome (**A**), the positions of +1 nucleosome (**B**) and the distance between the −1 and +1 nucleosomes (**C**) of the *URA3*-163 promoter after insertion into various locations of the *LYS2* gene are plotted against the position of nucleosome *n* before insertion. The positions of the −1 nucleosome (**D**), the positions of +1 nucleosome (**E**) and the distance between the −1 and +1 nucleosomes (**F**) of the *URA3*-163 promoter after insertion into various locations of the *LYS2* gene are plotted against %AT content at the junction between the upstream *LYS2* sequence and the *URA3* promoter. For each plot, the Pearson correlation coefficient was calculated.

Table 2. AT content of the junction sequence

Mutant	Junction sequence (40 nt upstream and 10 nt downstream)	% AT	Length of poly(dA:dT) > 75% purity (nt)
native *URA3* poly(dA:dT)	TTTTTTTTTTTTGTTCTTTTTTTTGATTCCGGTTTCTTTGAAATTTTTTT	82	50
LYS2-450	GTTCCGTTTGGCCTTTTTGGAAAACCAAGATTTCAAATTA*TTTTTTTGATT*	72	14
LYS2-500	AGCATCATTTAGTGGACTTTGCTTTGAATTTGGATACCAG*TTTTTTTGATT*	70	11
LYS2-540	TAATAATGCGCATGTTTTGAACTTAATTTATAACAGCTTA*TTTTTTTGATT*	78	14
LYS2-575	GCTTACTGTATTCGAATGAAAGAGTAACCATTGTTGCGGA*TTTTTTTGATT*	68	11
LYS2-721	TAAGAACTTGGGCTGGTGCGATTTCGTGGGGTGTATTCAC*TTTTTTTGATT*	58	11
LYS2-770	AGGACAATGCTGAAGCCTTCCCAGAGAGAACCTGTGTTGT*TTTTTTTGATT*	60	17
LYS2-800	CCTGTGTTGTGGAGACTCCAACACTAAATTCCGACAAGTC*TTTTTTTGATT*	62	13
LYS2-860	ACATCAACCGCACTTCTAACATAGTTGCCCATTATTTGAT*TTTTTTTGATT*	70	17
LYS2-865	AACCGCACTTCTAACATAGTTGCCCATTATTTGATTAAAA*TTTTTTTGATT*	74	11
LYS2-1950	AGAATTGAATAAAGAAAAATTTGTGAACAACTGGTTTGTT*TTTTTTTGATT*	80	17
LYS2-2200	CATCCATTGGTAAGAGAAAACATTACTTTAGTTCGCAAAA*TTTTTTTGATT*	70	11
LYS2-3050	CCCACGTCAGGGCCAAGGATGAAGAAGCTGCATTTGCAAG*TTTTTTTGATT*	56	11

in the *URA3* promoter must also help in establishing nucleosome organization.

Truncations of the *URA3* promoter that remove both the poly(dA:dT) tract and the upstream region of the *URA3* promoter show more dramatic differences in nucleosome organization than mutants in which only the poly(dA:dT) tract was removed. Moreover, when the promoter is truncated beyond a certain point, expression

and nucleosome positions become sensitive to the genomic context. More specifically, our data indicate that the positions of the upstream nucleosomes influence the position of the −1 nucleosome and the width of the NFR (i.e. the distance between the −1 and +1 nucleosomes). Interestingly, the position of the +1 nucleosome is in some cases also influenced by the (distal) chromatin context of the insertion site, indicating that nucleosome positions can influence each other over longer distances. However, the effect is relatively weak ($R = 0.38$; $P < 5 \times 10^{-4}$). This could be due to (a combination of) several factors. The local DNA sequence at the position of this nucleosome may play a primary role in positioning this nucleosome, and in addition, it is possible that *trans* factors such as ATP-dependent nucleosome remodeling proteins also contribute to the positioning of the key +1 nucleosome, as recently shown by Pugh and coworkers (40).

Changes in the *URA3* nucleosome pattern are correlated with *URA3* expression and *URA3*-dependent growth phenotypes. Our data show that *URA3* activity is largely determined by the extent of the NFR and the presence and exact position of the +1 and −1 nucleosomes. A narrow NFR results in decreased *URA3* expression, most likely due to decreased accessibility of the Ppr1 binding site and/or the TATA boxes that are located in the NFR. This effect appears to influence both basal and induced expression. In addition, the induction of different *URA3* promoters results in little change in nucleosome positioning but large changes in expression (Supplementary Figure S6). Thus, in this case, it appears that chromatin architecture solely sets accessibility, and the TF increases expression without gross modifications to the nucleosomal positions. Finally, we also observe that an upstream shift of the +1 nucleosome results in a reduction of *URA3* expression. The TSSs, normally located at the border of this nucleosome, gradually become covered by the +1 nucleosome as this nucleosome moves upstream.

In summary, our data indicate that nucleosome barriers such as poly(dA:dT) sequences are crucial for obtaining an accurate promoter nucleosome organization and correct gene expression. In addition, we provide novel *in vivo* evidence that the positions of neighboring nucleosomes in the genomic surroundings can also influence the local promoter nucleosome positioning, even though the effect is relatively weak. Thus, our data agrees with a thermodynamic model for nucleosome positioning, as proposed by the barrier model. However, other studies have convincingly shown that *trans* factors, such as chromatin remodeling factors, also play a key role. In fact, the data in Figure 5 and Supplementary Figure S7, when compared with each other, argue for a second layer of nucleosome organization in addition to statistical positioning. Here, two constructs that differ by only one nucleotide are inserted at various locations in the *LYS2* gene. When inserted at the same locus, these constructs sometimes show large differences in growth and *URA3* activity. This difference is most apparent for locus 721, and is caused by strikingly different nucleosome patterns in both constructs (Figure 4, Table 1). Because both constructs differ by only one nucleotide, and a different nucleosome pattern is not predicted by thermodynamic models (Supplementary Figure S10), this might indicate that *trans* factors are at play (although we did not find any evidence for the involvement of a *trans* factor).

One of the most intriguing properties of (eukaryotic) genomes is the relative high degree of conservation of gene synteny. Multiple selective forces act on synteny, and our study shows that nucleosome organization may play an important role. Relocating a gene to another genomic position may have dramatic consequences for its nucleosome structure, which in turn influences its expression. Nucleosome positioning may therefore be a limiting factor in natural recombination processes where genes may be deregulated after they are relocated to different parts of the genome. Moreover, in genetic engineering applications, the insertion site of a construct can be a crucial factor to obtain optimal expression and regulation of an ectopic sequence. This necessitates investigating multiple independent transformants to account for phenotypic differences caused by the different positions of inserts, rather than the function of the insert, or to find mutants that have the most optimal expression pattern. Our results indicate that the use of longer constructs that contain strong nucleosome boundary elements may help to reduce the influence of the chromosomal location on the expression of ectopic constructs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Materials and Methods, Supplementary Figures 1–13, Supplementary Tables I–VIII and Supplementary references (47–58).

ACKNOWLEDGEMENTS

The authors wish to thank T. Snoek and K. Voordeckers for critique of the manuscript, and Verstrepen and Fink lab members for fruitful discussions about this research.

FUNDING

European Molecular Biology Organization (EMBO) Young Investigator Program, Human Frontier Science Program (HFSP; grant number RGY79/2007); European Research Council (ERC) Young Investigator Grant (grant number 241426); Flemisch Institute for Biotechnology (VIB); University of Leuven (KU Leuven); Research Foundation – Flanders (FWO-Vlaanderen); FWO-Odysseus program; Agency for Innovation by Science and Technology (IWT); AB InBev Baillet-Latour Foundation (to K.J.V.); by the National Institutes of Health (NIH; grant numbers GM040266 and GM035010 to G.F.R.); University of Leuven (KU Leuven) BOF-BDB Fellowship (to A.J.). Funding for open access charge: VIB.

Conflict of interest statement. None declared.

REFERENCES

1. Kornberg,R.D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science*, **184**, 868–871.
2. Wu,T.C. and Lichten,M. (1994) Meiosis-induced double-strand break sites determined by yeast chromatin structure. *Science*, **263**, 515–518.
3. Zhang,Z., Shibahara,K. and Stillman,B. (2000) PCNA connects DNA replication to epigenetic inheritance in yeast. *Nature*, **408**, 221–225.
4. Lipford,J.R. and Bell,S.P. (2001) Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell*, **7**, 21–30.
5. Sekinger,E.A., Moqtaderi,Z. and Struhl,K. (2005) Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell*, **18**, 735–748.
6. Guffanti,E., Percudani,R., Harismendy,O., Soutourina,J., Werner,M., Iacovella,M.G., Negri,R. and Dieci,G. (2006) Nucleosome depletion activates poised RNA polymerase III at unconventional transcription sites in Saccharomyces cerevisiae. *J. Biol. Chem.*, **281**, 29155–29164.
7. Field,Y., Fondufe-Mittendorf,Y., Moore,I.K., Mieczkowski,P., Kaplan,N., Lubling,Y., Lieb,J.D., Widom,J. and Segal,E. (2009) Gene expression divergence in yeast is coupled to evolution of DNA-encoded nucleosome organization. *Nat. Genet.*, **41**, 438–445.
8. Lam,F.H., Steger,D.J. and O'Shea,E.K. (2008) Chromatin decouples promoter threshold from dynamic range. *Nature*, **453**, 246–250.
9. Han,M. and Grunstein,M. (1988) Nucleosome loss activates yeast downstream promoters in vivo. *Cell*, **55**, 1137–1145.
10. Han,M., Kim,U.J., Kayne,P. and Grunstein,M. (1988) Depletion of histone H4 and nucleosomes activates the PHO5 gene in Saccharomyces cerevisiae. *EMBO J.*, **7**, 2221–2228.
11. Yuan,G.C., Liu,Y.J., Dion,M.F., Slack,M.D., Wu,L.F., Altschuler,S.J. and Rando,O.J. (2005) Genome-scale identification of nucleosome positions in S. cerevisiae. *Science*, **309**, 626–630.
12. Lee,W., Tillo,D., Bray,N., Morse,R.H., Davis,R.W., Hughes,T.R. and Nislow,C. (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.*, **39**, 1235–1244.
13. Mavrich,T.N., Ioshikhes,I.P., Venters,B.J., Jiang,C., Tomsho,L.P., Qi,J., Schuster,S.C., Albert,I. and Pugh,B.F. (2008) A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.*, **18**, 1073–1083.
14. Shivaswamy,S., Bhinge,A., Zhao,Y., Jones,S., Hirst,M. and Iyer,V.R. (2008) Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.*, **6**, e65.
15. Field,Y., Kaplan,N., Fondufe-Mittendorf,Y., Moore,I.K., Sharon,E., Lubling,Y., Widom,J. and Segal,E. (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.*, **4**, e1000216.
16. Jiang,C. and Pugh,B.F. (2009) A compiled and systematic reference map of nucleosome positions across the Saccharomyces cerevisiae genome. *Genome Biol.*, **10**, R109.
17. Whitehouse,I., Rando,O.J., Delrow,J. and Tsukiyama,T. (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature*, **450**, 1031–1035.
18. Lee,C.K., Shibata,Y., Rao,B., Strahl,B.D. and Lieb,J.D. (2004) Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat. Genet.*, **36**, 900–905.
19. Struhl,K. (1985) Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl Acad. Sci. USA*, **82**, 8419–8423.
20. Albert,I., Mavrich,T.N., Tomsho,L.P., Qi,J., Zanton,S.J., Schuster,S.C. and Pugh,B.F. (2007) Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature*, **446**, 572–576.
21. Kaplan,N., Moore,I.K., Fondufe-Mittendorf,Y., Gossett,A.J., Tillo,D., Field,Y., LeProust,E.M., Hughes,T.R., Lieb,J.D., Widom,J. *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, **458**, 362–366.
22. Bai,L., Charvin,G., Siggia,E.D. and Cross,F.R. (2010) Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle. *Dev. Cell*, **18**, 544–555.
23. Jansen,A. and Verstrepen,K.J. (2011) Nucleosome positioning in. *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.*, **75**, 301–320.
24. Tillo,D. and Hughes,T.R. (2009) G+C content dominates intrinsic nucleosome occupancy. *BMC Bioinformatics*, **10**, 442.
25. Hughes,A. and Rando,O. (2009) Chromatin 'programming' by sequence - is there more to the nucleosome code than %GC? *J. Biol.*, **8**, 96.
26. Anderson,J.D. and Widom,J. (2001) Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. *Mol. Cell Biol.*, **21**, 3830–3839.
27. Roy,A., Exinger,F. and Losson,R. (1990) cis- and trans-acting regulatory elements of the yeast URA3 promoter. *Mol. Cell Biol.*, **10**, 5257–5270.
28. Iyer,V. and Struhl,K. (1995) Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J.*, **14**, 2570–2579.
29. Vinces,M.D., Legendre,M., Caldara,M., Hagihara,M. and Verstrepen,K.J. (2009) Unstable tandem repeats in promoters confer transcriptional evolvability. *Science*, **324**, 1213–1216.
30. Segal,E. and Widom,J. (2009) Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.*, **19**, 65–71.
31. Kornberg,R.D. and Stryer,L. (1988) Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Res.*, **16**, 6677–6690.
32. Bernardi,F., Zatchej,M. and Thoma,F. (1992) Species specific protein–DNA interactions may determine the chromatin units of genes in S.cerevisiae and in S.pombe. *EMBO J.*, **11**, 1177–1185.
33. Thoma,F. and Simpson,R.T. (1985) Local protein-DNA interactions may determine nucleosome positions on yeast plasmids. *Nature*, **315**, 250–252.
34. Thoma,F. and Zatchej,M. (1988) Chromatin folding modulates nucleosome positioning in yeast minichromosomes. *Cell*, **55**, 945–953.
35. Fedor,M.J., Lue,N.F. and Kornberg,R.D. (1988) Statistical positioning of nucleosomes by specific protein-binding to an upstream activating sequence in yeast. *J. Mol. Biol.*, **204**, 109–127.
36. Milani,P., Chevereau,G., Vaillant,C., Audit,B., Haftek-Terreau,Z., Marilley,M., Bouvet,P., Argoul,F. and Arneodo,A. (2009) Nucleosome positioning by genomic excluding-energy barriers. *Proc. Natl Acad. Sci. USA*, **106**, 22257–22262.
37. Vaillant,C., Palmeira,L., Chevereau,G., Audit,B., d'Aubenton-Carafa,Y., Thermes,C. and Arneodo,A. (2010) A novel strategy of transcription regulation by intragenic nucleosome ordering. *Genome Res.*, **20**, 59–67.
38. Mobius,W. and Gerland,U. (2010) Quantitative test of the barrier nucleosome model for statistical positioning of nucleosomes up- and downstream of transcription start sites. *PLoS Comput. Biol.*, **6**
39. Chevereau,G., Palmeira,L., Thermes,C., Arneodo,A. and Vaillant,C. (2009) Thermodynamics of intragenic nucleosome ordering. *Phys. Rev. Lett.*, **103**, 188103.
40. Zhang,Z., Wippo,C.J., Wal,M., Ward,E., Korber,P. and Pugh,B.F. (2011) A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science*, **332**, 977–980.
41. Brachmann,C.B., Davies,A., Cost,G.J., Caputo,E., Li,J., Hieter,P. and Boeke,J.D. (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, **14**, 115–132.
42. Guthrie,C. and Fink,G.R. (1991) *Guide to yeast genetics and molecular biology*. Academic Press, San Diego, California.
43. Tanaka,S., Livingstone-Zatchej,M. and Thoma,F. (1996) Chromatin structure of the yeast URA3 gene at high resolution provides insight into structure and positioning of nucleosomes in the chromosomal context. *J. Mol. Biol.*, **257**, 919–934.
44. Boeke,J.D., LaCroute,F. and Fink,G.R. (1984) A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.*, **197**, 345–346.

45. Gottschling,D.E., Aparicio,O.M., Billington,B.L. and Zakian,V.A. (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell*, **63**, 751–762.
46. Raser,J.M. and O'Shea,E.K. (2005) Noise in gene expression: origins, consequences, and control. *Science*, **309**, 2010–2013.
47. Ausubel,F.M. (1987) *Current Protocols in Molecular Biology*. Greene Pub. Associates and Wiley-Interscience: J. Wiley, New York.
48. Gietz,R.D., Schiestl,R.H., Willems,A.R. and Woods,R.A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, **11**, 355–360.
49. Sherman,F., Fink,G.R. and Hicks,J. (1991) *Methods in yeast genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
50. Storici,F., Lewis,L.K. and Resnick,M.A. (2001) In vivo site-directed mutagenesis using oligonucleotides. *Nat. Biotechnol.*, **19**, 773–776.
51. Chattoo,B.B., Sherman,F., Azubalis,D.A., Fjellstedt,T.A., Mehnert,D. and Ogur,M. (1979) Selection of lys2 mutants of the yeast Saccharomyces cerevisiae by the utilization of alpha-aminoadipate. *Genetics*, **93**, 51–65.
52. Zaret,K.S. and Sherman,F. (1985) alpha-Aminoadipate as a primary nitrogen source for Saccharomyces cerevisiae mutants. *J. Bacteriol.*, **162**, 579–583.
53. Sheff,M.A. and Thorn,K.S. (2004) Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. *Yeast*, **21**, 661–670.
54. Pernambuco,M.B., Winderickx,J., Crauwels,M., Griffioen,G., Mager,W.H. and Thevelein,J.M. (1996) Glucose-triggered signalling in Saccharomyces cerevisiae: different requirements for sugar phosphorylation between cells grown on glucose and those grown on non-fermentable carbon sources. *Microbiology*, **142**, 1775–1782.
55. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual.*, 2 edn. Cold Spring Harbor Laboratory Press, Plainview, NY.
56. Newburger,D.E. and Bulyk,M.L. (2009) UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res*, **37**, D77–D82.
57. Berger,M.F., Philippakis,A.A., Qureshi,A.M., He,F.S., Estep,P.W. III and Bulyk,M.L. (2006) Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol.*, **24**, 1429–1435.
58. Zhu,C., Byers,K.J., McCord,R.P., Shi,Z., Berger,M.F., Newburger,D.E., Saulrieta,K., Smith,Z., Shah,M.V., Radhakrishnan,M. *et al.* (2009) High-resolution DNA-binding specificity analysis of yeast transcription factors. *Genome Res.*, **19**, 556–566.