

Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*

Alexandra B Lantermann^{1,5}, Tobias Straub^{1,5}, Annelie Strålfors², Guo-Cheng Yuan^{3,4}, Karl Ekwall² & Philipp Korber¹

Positioned nucleosomes limit the access of proteins to DNA and implement regulatory features encoded in eukaryotic genomes. Here we have generated the first genome-wide nucleosome positioning map for *Schizosaccharomyces pombe* and annotated transcription start and termination sites genome wide. Using this resource, we found surprising differences from the previously published nucleosome organization of the distantly related yeast *Saccharomyces cerevisiae*. DNA sequence guides nucleosome positioning differently: for example, poly(dA-dT) elements are not enriched in *S. pombe* nucleosome-depleted regions. Regular nucleosomal arrays emanate more asymmetrically—mainly codirectionally with transcription—from promoter nucleosome-depleted regions, but promoters harboring the histone variant H2A.Z also show regular arrays upstream of these regions. Regular nucleosome phasing in *S. pombe* has a very short repeat length of 154 base pairs and requires a remodeler, Mit1, that is conserved in humans but is not found in *S. cerevisiae*. Nucleosome positioning mechanisms are evidently not universal but evolutionarily plastic.

The positioning of nucleosomes along eukaryotic chromosomes affects DNA accessibility and provides a key mechanism for the regulation of DNA-related processes such as transcription, replication, recombination and repair^{1–3}. Genome-wide nucleosome positioning maps have previously been analyzed in the unicellular budding yeast *S. cerevisiae* and in various metazoans, such as *Drosophila melanogaster*, *Homo sapiens* and *Caenorhabditis elegans*^{4–13}. These maps revealed strikingly similar general nucleosome positioning patterns at promoters in various species. A nucleosome-depleted region (NDR) close to the transcription start site is flanked up- and downstream by positioned nucleosomes (denoted the –1 and +1 nucleosomes, respectively) that are often the starting points for regular nucleosomal arrays. The observation of such stereotypical patterns and other well-defined nucleosome positions raises two questions: what determines nucleosome positioning? And how well conserved are nucleosome positioning mechanisms across species?

The contributions of DNA sequence features and protein factors to the determination of nucleosome positions are increasingly recognized^{5,6,8,12,14–23}. One prominent example of the role of DNA sequence elements is the existence of poly(dA-dT) tracts that intrinsically disfavor strong bending of DNA such as occurs upon nucleosome formation^{17,24} and that seem to be involved in excluding nucleosomes from promoter NDRs *in vivo*^{17,22,25}. More generally, however,

nucleosome positioning *in vivo* seems also to be determined by factors beyond intrinsic DNA sequence properties^{22,23}. In *S. cerevisiae*, ATP-dependent nucleosome remodelers such as Isw2 or RSC (refs. 6 and 23, respectively) can position nucleosomes over unfavorable DNA sequences *in vivo*. Furthermore, *in vitro* reconstitution of chromatin using DNA and histones under salt-gradient dialysis conditions only sometimes, as at the *S. cerevisiae* PHO84 promoter²⁶, leads to chromatin patterns that are close to those seen *in vivo*¹⁷, but usually, this approach does not recapitulate accurate *in vivo* nucleosome positioning²². Intriguingly, incubation of such salt-gradient dialysis chromatin with a yeast whole-cell extract in the presence of ATP shifts nucleosomes to their *in vivo* positions, as for the *S. cerevisiae* PHO5 and PHO8 promoters (refs. 27 and 28, respectively). This argues that protein factors, in addition to DNA and histones, have a necessary role in nucleosome positioning.

With regard to the second question, it is currently unclear how well conserved the interplay is between DNA sequence, histones, and other protein factors that determine nucleosome positioning. The extremely well-conserved structure of the nucleosome may suggest universally conserved nucleosome positioning mechanisms. This possibility is supported by recent reports on the universality of DNA-encoded nucleosome positioning signals⁸ and on a similar encoding of open promoter chromatin structures by DNA sequence in the relatively closely related yeasts *S. cerevisiae* and *Candida albicans*¹³. However,

¹Adolf-Butenandt-Institut, University of Munich, Munich, Germany. ²Karolinska Institutet, Department of Biosciences and Nutrition, Center for Biosciences NOVUM, Huddinge, Sweden. ³Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA. ⁴Department of Biostatistics and Computational Biology, Dana-Faber Cancer Institute, Boston, Massachusetts, USA. ⁵These authors contributed equally to this work. Correspondence should be addressed to K.E. (karl.ekwall@ki.se) or P.K. (pkorber@lmu.de).

Received 15 July 2009; accepted 19 November 2009; published online 31 January 2010; doi:10.1038/nsmb.1741

Figure 1 Alignment of genes at their TSS reveals a prominent NDR upstream and a regular nucleosomal array downstream of the TSS, with a shorter repeat length in *S. pombe* than in *S. cerevisiae*. **(a)** Overlay of nucleosome occupancy profiles of 4,013 *S. pombe* genes after TSS alignment. **(b)** Same as in **a**, along with the same type of alignment for *S. cerevisiae* genes. **(c)** MNase ladder analysis for *S. cerevisiae* and *S. pombe*. White asterisks mark the position of the tetranucleosomal fragments. The 1-kb ladder (NEB) was loaded as marker (M). **(d)** Spectral analyses of nucleosome occupancy profiles for *S. pombe* and *S. cerevisiae* reveal frequency peaks (marked by vertical lines) at 6.5 and 6 nucleosomes per 1,000 bp, respectively.

S. pombe chromosome segments inserted into mouse chromosomes adopt the nucleosome repeat length typical for mouse chromatin²⁹, and shuttle vectors are assembled into different chromatin structures in the distantly related *S. cerevisiae* and *S. pombe*³⁰. The latter observations argue for species-specific nucleosome positioning along the same DNA sequence.

To address these questions, we undertook the first genome-wide mapping of nucleosome positions in the fission yeast *S. pombe* together with a comprehensive comparison to published maps of nucleosome positioning in the budding yeast *S. cerevisiae*.

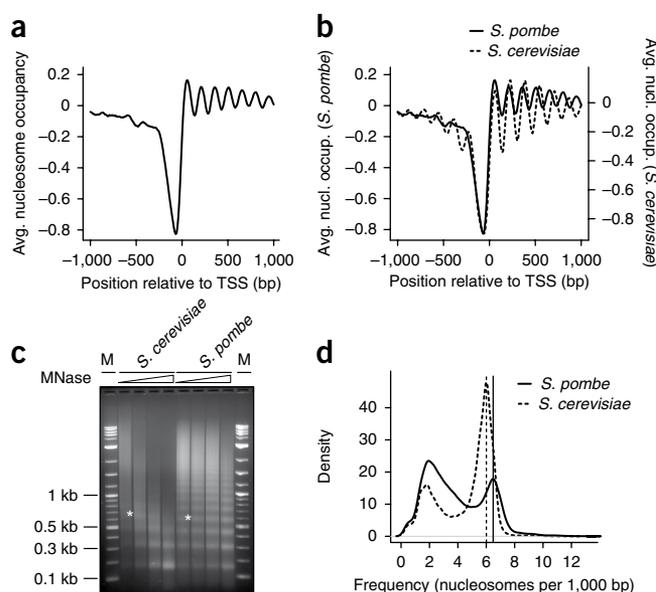
RESULTS

For genome-wide nucleosome mapping in *S. pombe*, we prepared mononucleosomal DNA by digesting chromatin with micrococcal nuclease (MNase) and hybridized it to a whole-genome 20-bp-resolution tiling microarray³¹. The accuracy of our method was extensively validated, as a comparison of nucleosome positions derived from classical MNase indirect end-labeling and from microarray data at 19 loci showed 94% of nucleosome borders coinciding in the two analyses (Supplementary Fig. 1a,b and Supplementary Table 1).

A short nucleosomal repeat length in *S. pombe*

As promoters in *S. cerevisiae*, as well as in fly and human cells, show the stereotypical nucleosome organization^{1,4–7,9,10,12}, we looked for a similar pattern at *S. pombe* promoters. We annotated the transcriptional start (TSS) and transcription termination sites (TTS) of 4,013 and 3,925 genes, respectively, using published transcriptome data³² (Supplementary Fig. 1c and Supplementary Table 2). An overlay of hybridization profiles after alignment at the TSS showed a pronounced NDR just upstream and a regular nucleosomal array downstream of the TSS (Fig. 1a). The positions of NDRs correlated very well with our annotated TSSs, even for unusually long 5' untranslated regions (Supplementary Fig. 2). The position of the NDR and of the +1 nucleosome relative to the TSS, and the presence of a regular downstream nucleosomal array, were in agreement with the corresponding overlay profile of *S. cerevisiae* promoters^{4–6} (Fig. 1b).

However, there were also important differences in nucleosome organization between *S. pombe* and *S. cerevisiae*. The average distance of nucleosome occupancy peaks—the nucleosome repeat length or spacing—was considerably shorter in *S. pombe* (Fig. 1b). We confirmed this difference in MNase ladders (Fig. 1c, Supplementary Fig. 3a) and used spectral analysis to scan the hybridization data for periodic nucleosome positioning patterns. This revealed prominent peaks at frequencies of 6.5 or 6 nucleosomes per 1,000 bp, translating to nucleosome repeat lengths of 154 bp and 167 bp, for *S. pombe* and *S. cerevisiae*, respectively (Fig. 1d). A broad peak at about 2 nucleosomes per 1,000 bp corresponds to low-frequency noise. Our measurement of the shorter repeat length for *S. pombe* resolves a past disagreement between one study reporting a *S. pombe* repeat length of 154 ± 2 bp³³ and another reporting the same spacing



for *S. pombe* as for *S. cerevisiae*³⁴; the latter was described as 165 ± 5 bp^{5,7,12,35}, in good agreement with our value.

No prominent array regularity upstream of NDRs

An even more striking difference was the lack of a positioned –1 nucleosome or a regular nucleosomal array upstream of the NDR in the *S. pombe* overlay pattern. As such regular features may be obscured in composite overlays if they belong to subgroups with offset array registers, we clustered promoter regions on the basis of their promoter nucleosome occupancy profiles (Fig. 2a,b). Indeed, some promoter organization subtypes (clusters 1, 3, 4 and 6) did show a –1 nucleosome but at different positions relative to the TSS. Nonetheless, a regular nucleosomal array upstream of the NDR was not visible in any cluster. In contrast, such arrays were very prominent when we performed the same clustering for nucleosome occupancy profiles from *S. cerevisiae* (clusters 2–5 in Supplementary Fig. 3c). Generally, the amplitude of the patterns was higher for the *S. cerevisiae* data owing to their higher resolution (4 bp). However, *S. cerevisiae* data⁴ with the same resolution (20 bp) as our *S. pombe* data still showed regular upstream arrays (Supplementary Fig. 3d). The lack of regular upstream patterns in our data is therefore unlikely to be due to the lower resolution. Further, the median intergenic distance in *S. pombe* is 442 bp, compared to 366 bp for *S. cerevisiae*, arguing against a more frequent disturbance of the upstream region in *S. pombe* by upstream genes. An alignment of genes without an upstream gene within 1 kb also showed no regular upstream arrays (Supplementary Fig. 3e).

High expression tends to correlate with open promoters

The clustering of *S. pombe* promoter nucleosome patterns also revealed differently pronounced NDRs (compare clusters 2 and 6 with the others in Fig. 2b; cluster 6 may even have two weak NDRs). NDRs are associated with gene activity in *S. cerevisiae*^{36,37}, and different promoter chromatin organizations were reported to be correlated with different expression levels of the encoded genes⁵. In *S. pombe*, promoters with a deeper NDR (genes in clusters 1, 3, 4 and 5) had significantly higher median expression levels (Fig. 2c; *P*-value $< 2.2 \times 10^{-16}$, two-sided Student's *t*-test). Accordingly, grouping genes on the basis of steady-state expression levels (Supplementary Fig. 4a) correlated higher average expression with deeper NDR troughs



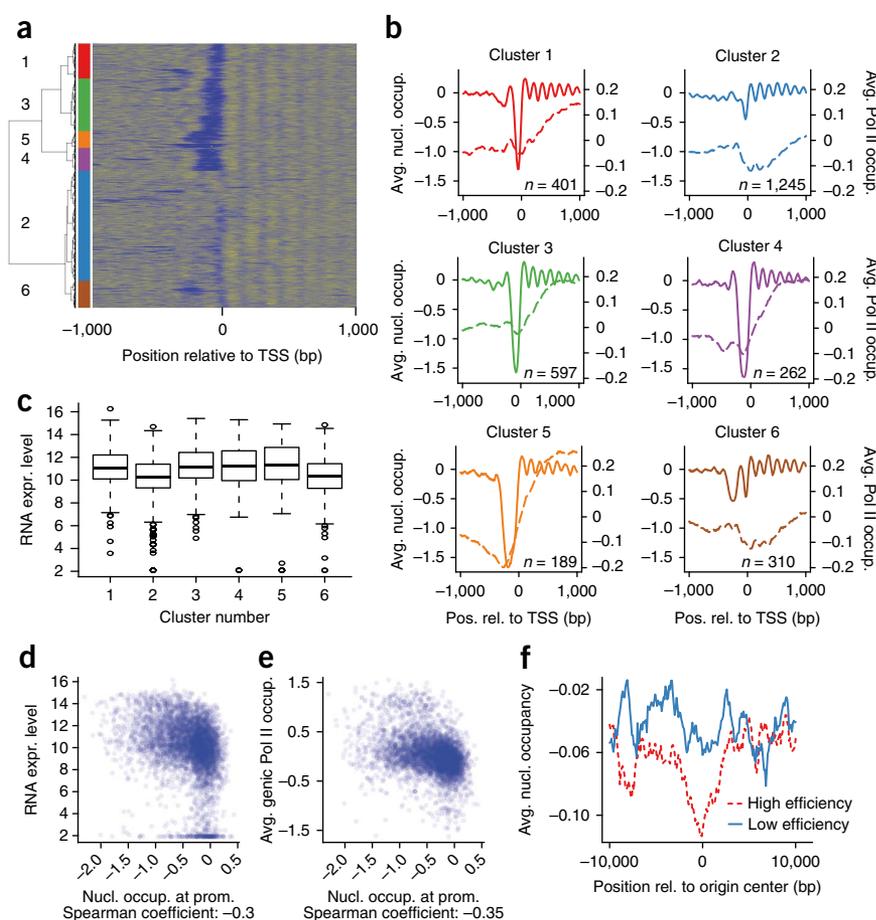


Figure 2 Subtypes of promoter chromatin organization are not generally predictive for levels of gene expression or RNA polymerase II occupancy. **(a)** Nucleosome occupancy profiles were clustered according to the pattern surrounding the TSS. **(b)** As in **Figure 1a**, but for gene clusters as derived in **a**. In addition, average RNA polymerase II occupancy levels are shown, and the number of genes in each cluster (*n*) is given. **(c)** Box plot analysis of expression data for gene clusters as in **a**. **(d)** Scatter-plot correlation of promoter nucleosome occupancy with gene expression. **(e)** Scatter-plot correlation of promoter nucleosome occupancy with RNA polymerase II occupancy over the transcript. **(f)** Overlay of nucleosome occupancy profiles after alignment at high- and low-efficiency replication origins.

Fig. 2f) correlated well with the efficiency of *S. pombe* replication origins³⁹. This nucleosome depletion over high efficiency origins (*P*-value < 2.2×10^{-16} , two-sided Wilcoxon test) experimentally confirmed an earlier prediction⁸, but it was of a different quality—that is, spread out over a larger region and less extensive (**Supplementary Fig. 5a**)—than the NDRs at promoters.

In *S. cerevisiae*, there is a positioned nucleosome followed by a NDR at the 3' end of transcription units^{6,7,12}. Such an NDR was also weakly discernible in *S. pombe* but without clear nucleosome positioning (**Supplementary Fig. 5b**). Alignment at stop codons or at TTSs

(**Supplementary Fig. 4b**) and lower promoter nucleosome occupancy (**Supplementary Fig. 4c**). Thus, we observed the same trend—in fact, somewhat more pronounced—of more open promoter chromatin at more highly expressed genes that was seen previously by others for *S. cerevisiae*⁵ (compare the same representation for *S. cerevisiae* in **Supplementary Fig. 4d**) and by ourselves, at lower resolution, for *S. pombe*³⁸.

Nonetheless, we note that a gene-by-gene correlation of promoter nucleosome occupancy and steady-state expression levels was rather poor—aside from the higher occupancy seen at silent genes—both in *S. pombe* and in *S. cerevisiae* (**Fig. 2d** and **Supplementary Fig. 4e**). The same was true for a correlation of promoter nucleosome occupancy with RNA polymerase II occupancy (**Fig. 2e**), which may provide a more direct readout of chromatin effects as it is less dependent on post-transcriptional processes. This reflects the fact that a trend of averages need not necessarily provide reliable predictions on a single-gene basis. Here this may be because the dynamic range of transcription levels is much greater than that of nucleosome occupancy. Furthermore, for moderately and weakly expressed genes, on which much of the correlation is based, it is difficult to accurately measure a possibly transient dissociation of nucleosomes during the relatively rare passage of polymerase.

Over coding regions, the average nucleosome occupancy was unaffected by expression levels in *S. pombe* (**Supplementary Fig. 4f**), which is in contrast to the situation in *S. cerevisiae*, where nucleosome occupancy was higher in the coding regions of highly expressed genes⁵.

In agreement with previous suggestions, nucleosome depletion over large regions (note the different scales of the *x* and *y* axes in

did not make much difference in this regard. Thus, a positioned nucleosome 3' of genes does not seem to be a strong universal feature.

As has been observed in budding yeast^{2,3,26,27}, nucleosome occupancy in intergenic regions was lower than in genic regions in *S. pombe* (**Supplementary Fig. 5c**), which may be a result of the prevalence of NDRs at promoters.

A particularly interesting case of NDR formation occurred at promoters bound by the co-repressors Tup11, Tup12 and Ssn6 (ref. 40), as the NDRs were very deep and broad (**Supplementary Fig. 5d**). In budding yeast, however, the homologs Tup1 and Ssn6 did not affect promoter NDRs but generated regular nucleosome positioning at the *FLO1*, *RNR2*, *RNR3*, *ANB1* and *SUC2* loci and at several genes specific to cells of the **a** mating type⁴¹. In a *S. pombe* *tup11 tup12* double mutant, the nucleosome occupancy at promoter NDRs was no different from that in wild-type *S. pombe* (**Supplementary Fig. 5e**). Thus, either other factors, intrinsic properties of Tup11 and Tup12 target promoters or both—but not the co-repressors themselves—seem to cause this special promoter nucleosome pattern.

Species-specific reading of DNA sequence features

Despite some commonalities in nucleosome organization between the two yeast species, we were struck by the differences and wondered whether they were the result of different nucleosome positioning mechanisms. The contribution of the DNA sequence to nucleosome positioning is strong enough that computational models, such as our previously developed N-score algorithm¹⁶, can be trained on experimental nucleosome positioning data to allow some prediction of nucleosome occupancy from the DNA sequence alone^{5,8,14–18,21}.



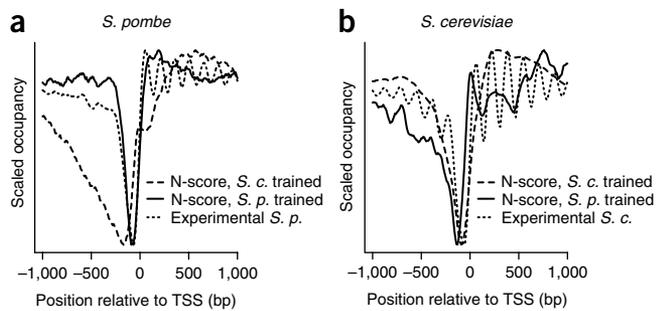


Figure 3 Nucleosome occupancy responds differently to DNA sequence in *S. pombe* and *S. cerevisiae*. (a) Scaled overlays of nucleosome occupancy, as in **Figure 1a**, and of N-score calculations after training with *S. pombe* or *S. cerevisiae* hybridization data and application to the *S. pombe* genome sequence. (b) As in **a**, but with experimental data for *S. cerevisiae* and N-score calculations applied to the *S. cerevisiae* genome sequence.

We trained the N-score algorithm¹⁶ on hybridization data from *S. pombe* and from *S. cerevisiae*⁵ and applied both model versions to the genome sequences of both yeasts. The N-score did a very good job of predicting the NDR and overall nucleosome occupancy in the species for which it was trained but performed considerably worse when applied cross species (**Fig. 3a,b**). This was even more apparent when comparing individual clusters of promoter nucleosome organizations (**Supplementary Fig. 6a,b**), where the NDR was often poorly met and sometimes peaks and troughs even coincided (clusters 1, 2 and 6 in **Supplementary Fig. 6a** and clusters 5 and 6 in **Supplementary Fig. 6b**). Accordingly, the DNA sequence rules as reflected in the N-score parameters for each species are different (**Supplementary Table 3**). Some of the most discriminative features in *S. cerevisiae* are the structural parameters^{5,42} “tip” (rotation about long base-pair axis), “minor_mobility” (mobility to bend toward minor groove) and “minor_size” (minor groove size), none of which are top predictors for *S. pombe*. By contrast, the most discriminative features for *S. pombe* are the sequence CGTTA, “nucleosome probability” (probability of contacting the nucleosome core) and “wedge” (helix deflection angle), none of which are top predictors for *S. cerevisiae*.

One of the most extensively studied sequence features in *S. cerevisiae* is poly(dA-dT), which is a strong nucleosome exclusion signal^{1-4,25}.

However, we did not observe a similar trend in *S. pombe*, where in fact poly(dA-dT) sequences occurred less frequently in NDRs than elsewhere. For example, 8.8% of NDR probes contained the pentamer AAAAA, compared to 12.4% elsewhere (**Supplementary Table 4**). Poly(dA-dT) enrichment in NDRs is less common in human, chicken and fly than in worm and budding yeast^{2,8,11}; thus, in this regard, fission yeast is more similar to the former three species than to budding yeast. Unexpectedly, however, *S. pombe* NDRs are enriched almost fourfold for the sequence CGTTA as compared to other genomic regions (**Supplementary Table 4**).

Likewise, the unrelated model for predicting nucleosome positioning developed by Kaplan *et al.*¹⁷ performed well for *S. cerevisiae* but poorly for *S. pombe*, as it predicted a peak of nucleosome occupancy within the promoter NDR (**Supplementary Fig. 7a,b**).

Collectively, DNA sequence properties are interpreted in markedly different ways by the nucleosome positioning machineries in the two yeasts, which is in keeping with the differential nucleosome positioning on shuttle vectors³⁰ and over the *S. cerevisiae* *HIS3* locus after integration into the *S. pombe* genome²⁵. Species-specific nucleosome positioning factors may override purely biophysical DNA sequence properties and thus limit the power of models based only on the interaction of histones and DNA for predicting *in vivo* nucleosome positions¹⁷. Accordingly, previous researchers have reported²² that intrinsic histone-DNA interactions are not sufficient to determine nucleosome positions in *S. cerevisiae in vivo*.

Array formation is co-directional with transcription

According to the “statistical nucleosome positioning model”⁴³, some genomic regions function as boundary elements—that is, as alignment sites for the formation of regular nucleosomal arrays through a passive and statistical queuing process. NDRs have been suggested to provide such boundary function^{1,4,12}, although it is unclear exactly what corresponds to the boundary²². We now argue that the alignment of nucleosomes to promoter NDRs, at least in *S. pombe*, is not entirely passive but rather is a directional and active process.

We found that, similarly to earlier conclusions drawn from data of other species^{4,5,8,10,22}, the directionality of nucleosome array formation in *S. pombe* seems to be linked to transcription. First, an alignment at all NDRs in *S. pombe* did not reveal a regular array pattern emanating in either direction (**Fig. 4a**). Instead, a nucleosomal array

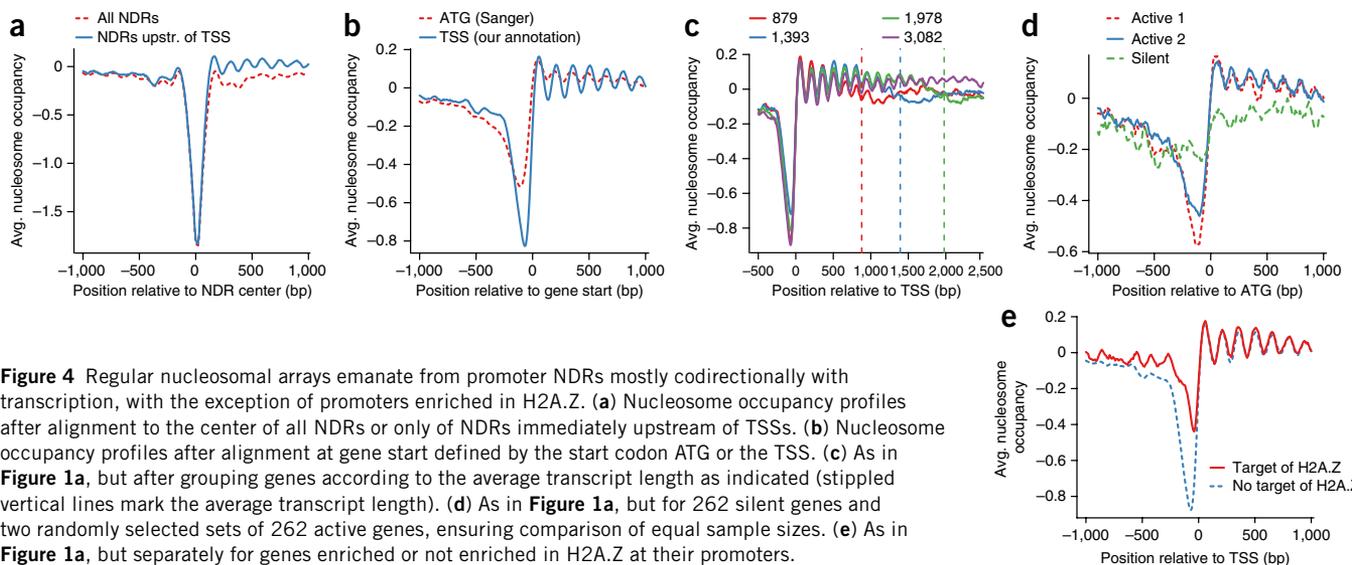


Figure 4 Regular nucleosomal arrays emanate from promoter NDRs mostly codirectionally with transcription, with the exception of promoters enriched in H2A.Z. (a) Nucleosome occupancy profiles after alignment to the center of all NDRs or only of NDRs immediately upstream of TSSs. (b) Nucleosome occupancy profiles after alignment at gene start defined by the start codon ATG or the TSS. (c) As in **Figure 1a**, but after grouping genes according to the average transcript length as indicated (stippled vertical lines mark the average transcript length). (d) As in **Figure 1a**, but for 262 silent genes and two randomly selected sets of 262 active genes, ensuring comparison of equal sample sizes. (e) As in **Figure 1a**, but separately for genes enriched or not enriched in H2A.Z at their promoters.



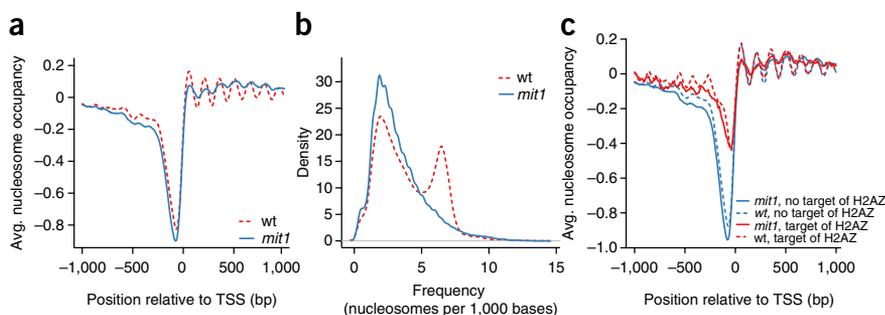


Figure 5 The Snf2-type remodeler ATPase Mit1 is critical for the regularity of nucleosomal arrays. (a) As in **Figure 1a** for wild type (wt) and *mit1*. (b) As in **Figure 1d** for wt and *mit1*. (c) Stippled profiles are the same as in **Figure 4e**, and solid profiles show the corresponding overlays for the *mit1* mutant.

became apparent only after alignment at promoter NDRs but not at any other type of NDR (data not shown), and only if the alignment was in the same transcriptional orientation (**Fig. 4a**). This argues strongly that the generation of regular nucleosomal arrays at promoter NDRs is not symmetrical but rather is codirectional with transcription. The alignment at the NDR center in **Figure 4a** emphasizes the NDR depth, whereas the alignment at the TSS in **Figure 1a** yields a more pronounced amplitude for the nucleosomal array. In general, the more distinct the point of alignment is for all individual patterns, the more distinct the composite alignment patterns become. Accordingly, and similarly to what occurs in *S. cerevisiae*⁵, an alignment at the start codon (ATG) yielded less pronounced array amplitudes than an alignment at the TSS (**Fig. 4b**). The greater distinctness of the array in **Figure 1a** therefore suggests that the transcription-related point of alignment (TSS) is more relevant for setting the array register than the chromatin-related (NDR) or translation-related (ATG) points of alignment. Second, the overlay of TSS-aligned RNA polymerase II occupancy profiles showed polymerase enrichment underlying the regular arrays (**Fig. 2b**). Third, transcript length correlated well with array extent (**Fig. 4c**), in contrast to the uniform dampening of the oscillatory pattern regardless of transcript length that would be expected in the case of a purely passive queuing mechanism. Fourth, little detectable regular array pattern was observed at silent as compared to active gene promoters (**Fig. 4d**). As we could not assign TSSs for silent genes, we aligned them at the ATG, which generates discernible arrays for active (**Fig. 4b,d**) but not for silent genes (**Fig. 4d**).

Intriguingly, in addition to the downstream array, at promoters harboring the histone variant H2A.Z, a weaker but appreciable upstream array was also visible (**Fig. 4e**). Similar to some observations in *S. cerevisiae*⁴⁴, these H2A.Z-containing promoters drive genes with lower expression levels on average⁴⁵, which correlated with a less pronounced NDR (**Fig. 4e**). It remains to be determined whether such upstream arrays are linked to the recently described role of H2A.Z in antisense RNA suppression in *S. pombe*⁴⁶.

Active phasing involves the remodeler Mit1

If the NDR sets the boundary and transcription determines the direction, there remains the question of what sets the regular spacing between the nucleosomes. Members of the ISWI class of nucleosome-remodeler ATPases function as nucleosome spacing factors⁴⁷ and generate directional nucleosome positioning over energetically unfavorable DNA sequences in *S. cerevisiae*⁶. However, the *S. pombe* genome encodes a different set of remodeler ATPases than does the *S. cerevisiae* genome⁴⁸. Most strikingly, there is no ISWI remodeler but, among others, a Mi-2 type of remodeler, Mit1, that

is conserved in humans but not present in *S. cerevisiae*. Mit1 was purified from *S. pombe* as a subunit of the SHREC complex⁴⁹, which is involved in nucleosome positioning at the heterochromatic mating type locus but is also associated with euchromatic regions. This makes it a prime candidate for a nucleosome positioning factor in euchromatic coding regions also. Indeed, the TSS-aligned nucleosome occupancy profile of a *mit1* deletion mutant showed a strikingly diminished amplitude—that is, a compromised regularity—of the downstream nucleosomal array as compared to the wild type (**Fig. 5a**). The prevalent frequency of 6.5 nucleosomes per 1,000 bp seen in spectral analysis of wild-type *S. pombe* (**Fig. 1d**) was not discernible in the *mit1* mutant (**Fig. 5b**). Notably, not only the downstream arrays but also the weaker upstream arrays at H2A.Z-containing promoters were compromised in the absence of Mit1 (**Fig. 5c**). This argues that Mit1 is critically important for regular nucleosome spacing downstream and upstream of the promoter NDRs, supporting our interpretation of an active instead of a passive nucleosome alignment process.

The Mit1 effect seems to be specific, as it was not observed in the absence of another remodeler, Fft3 (**Supplementary Fig. 7c**), which is the homolog of *S. cerevisiae* Fun30 (ref. 48).

Deletion of *mit1* affected the expression of numerous genes; 300 were up- and 367 downregulated (using a twofold change as threshold; **Supplementary Table 5**). The widespread effects on euchromatic genes suggest a genome-wide function of Mit1. How the decreased regularity of nucleosome phasing relates to changes in expression levels remains to be further studied.

We also checked whether there were heterochromatin-specific effects on nucleosome occupancy in the *mit1* mutant. We defined heterochromatin by the presence of the histone H3 Lys9 (H3K9) dimethyl mark⁵⁰. Only 23 genes in heterochromatin are active enough to allow TSS annotation. This low number precludes analysis of the type shown in **Figure 1a** for heterochromatic genes only, and the results of the analysis shown in **Figure 4a** were unchanged when these genes were omitted (data not shown). Heterochromatic regions contain many repetitive sequences, which were so far excluded from our analysis. Nonetheless, separate spectral analyses of euchromatin and heterochromatin regions, including repetitive sequences, revealed wider spacing for heterochromatin (**Supplementary Fig. 7d**; compare with **Fig. 1d**), which is consistent with a lower median nucleosome occupancy in heterochromatin compared to euchromatin (**Supplementary Fig. 7e**). This distribution was unchanged in the *mit1* mutant (**Supplementary Fig. 7f**). Owing to the inclusion of repetitive sequences, such results are preliminary, but they are potentially very interesting, as a lower nucleosome occupancy in heterochromatin is unexpected in the light of the repressive function. We speculate that a wider spacing—that is, a longer linker length—in heterochromatin may be more conducive to tighter forms of higher-order folding.

DISCUSSION

Our comparative genomics analysis of nucleosome positioning in the two distantly related yeasts *S. pombe* and *S. cerevisiae* showed surprisingly different mechanisms of nucleosome positioning with respect to the roles of DNA sequence features, NDR boundary elements and remodelers and also in regard to the roles of the histone variant H2A.

Z and the co-repressor complex Tup11–Tup12–Ssn6. This argues for the evolutionary plasticity of nucleosome positioning mechanisms and against the existence of a universal nucleosome positioning code. We suggest that this plasticity provides an important degree of flexibility for the evolution of genomes and their regulatory networks.

We agree with earlier conclusions^{12,22} that the statistical nucleosome positioning model⁴³ is probably correct in the sense that DNA sequence features co-define nucleosome positioning or nucleosome depletion only in relatively few cases at boundary elements, but not for the majority of individual nucleosomes. However, in light of asymmetrical array formation, which is especially pronounced in *S. pombe*, the definition of the array register start at the boundary by itself is not sufficient. There must be a mechanism to define the direction of array formation starting from the boundary and a mechanism to set the spacing. We propose that in *S. pombe* the transcriptional machinery is involved in setting the direction and the ATPase subunit Mit1 of the SHREC complex is critical for actively generating the regular spacing.

Recently, histone acetylation in coding regions by the histone acetyltransferase Gcn5 was linked to transcriptional elongation in *S. pombe*⁵¹. Interestingly, both the transcription elongation defect and the reduced acetylation at H3K14 in a *gcn5* mutant could be completely suppressed by deletion of *clr3*, the histone deacetylase component of SHREC. This further suggests a role in euchromatic transcribed regions for the SHREC complex, and we speculate that this histone acetyltransferase–histone deacetylase interplay in coding regions may be coupled to the active nucleosome phasing by the Mit1 remodeler.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/nsmb/>.

Accession codes. The raw data for the microarray hybridizations reported here are deposited at the Gene Expression Omnibus (GSE16040).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank H. Bhuiyan and J. Walfridsson for generating the *S. pombe* expression data during their work in the group of K. Ekwall, R.R. Barrales (group of J.J. Ibeas, Universidad Pablo de Olavide, Sevilla, Spain) for bringing the first *S. pombe* strains into the Korber group, F. Thoma (ETH Zürich, Switzerland) for advice on chromatin analysis in *S. pombe*, F. Fagerström-Billai at the BEA microarray facility at Novum, Karolinska Institutet, for assistance, and F. Müller-Planitz (Adolf-Butenandt-Institut, Univ. Munich) for help with MATLAB. We are grateful for the communication of replication origin coordinates by C. Heichinger (Univ. Zürich) and of TSS coordinates by W. Lee (Stanford Univ.) and N. Dutrow (Univ. Utah). We thank H. Madhani and co-workers (Univ. California San Francisco) for sharing data before publication and for comments on the manuscript. This work was funded by the German Research Community (Transregio 5; P.K. and co-workers), the 6th Framework Programme of the European Union (NET programme; P.K. and K.E. and co-workers), the Swedish Cancer Society and Swedish Research Council (K.E. laboratory) and the Claudia Adams Barr Program (G.-C.Y.).

AUTHOR CONTRIBUTIONS

A.B.L. carried out all preparation and experimental analysis of biological material besides the actual microarray hybridizations, which were done at the BEA Affymetrix core facility at Novum with the help of A.S. A.B.L. did the *S. pombe* TSS and TTS annotation. T.S. did the bioinformatics analyses. G.-C.Y. provided the N-score codes, applied the model of Kaplan et al.¹⁷, analyzed DNA sequence features and gave advice on bioinformatics. A.S. and K.E. introduced A.B.L. to the work with *S. pombe* and microarrays. K.E. provided strains and reference data. P.K. and K.E. initiated, designed and supervised the study. A.B.L. and T.S. generated the

figures. P.K. and A.B.L. wrote the paper. A.B.L. and T.S. contributed equally to the study. All authors discussed results and commented on the manuscript

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at <http://www.nature.com/nsmb/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

- Jiang, C. & Pugh, B.F. Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172 (2009).
- Radman-Livaja, M. & Rando, O.J. Nucleosome positioning: how is it established, and why does it matter? *Dev. Biol.* published online (13 June 2009).
- Segal, E. & Widom, J. What controls nucleosome positions? *Trends Genet.* **25**, 335–343 (2009).
- Yuan, G.C. *et al.* Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**, 626–630 (2005).
- Lee, W. *et al.* A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **39**, 1235–1244 (2007).
- Whitehouse, I., Rando, O.J., Delrow, J. & Tsukiyama, T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031–1035 (2007).
- Shivaswamy, S. *et al.* Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* **6**, e65 (2008).
- Field, Y. *et al.* Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* **4**, e1000216 (2008).
- Mavrich, T.N. *et al.* Nucleosome organization in the *Drosophila* genome. *Nature* **453**, 358–362 (2008).
- Schones, D.E. *et al.* Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898 (2008).
- Valouev, A. *et al.* A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* **18**, 1051–1063 (2008).
- Mavrich, T.N. *et al.* A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18**, 1073–1083 (2008).
- Field, Y. *et al.* Gene expression divergence in yeast is coupled to evolution of DNA-encoded nucleosome organization. *Nat. Genet.* **41**, 438–445 (2009).
- Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442**, 772–778 (2006).
- Ioshikhes, I.P., Albert, I., Zanton, S.J. & Pugh, B.F. Nucleosome positions predicted through comparative genomics. *Nat. Genet.* **38**, 1210–1215 (2006).
- Yuan, G.C. & Liu, J.S. Genomic sequence is highly predictive of local nucleosome depletion. *PLoS Comput. Biol.* **4**, e13 (2008).
- Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366 (2009).
- Peckham, H.E. *et al.* Nucleosome positioning signals in genomic DNA. *Genome Res.* **17**, 1170–1177 (2007).
- Badis, G. *et al.* A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol. Cell* **32**, 878–887 (2008).
- Parnell, T.J., Huff, J.T. & Cairns, B.R. RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. *EMBO J.* **27**, 100–110 (2008).
- Gupta, S. *et al.* Predicting human nucleosome occupancy from primary sequence. *PLoS Comput. Biol.* **4**, e1000134 (2008).
- Zhang, Y. *et al.* Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 847–852 (2009).
- Hartley, P.D. & Madhani, H.D. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**, 445–458 (2009).
- Drew, H.R. & Travers, A.A. DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **186**, 773–790 (1985).
- Sekinger, E.A., Moqtaderi, Z. & Struhl, K. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* **18**, 735–748 (2005).
- Wippo, C.J. *et al.* Differential cofactor requirements for histone eviction from two nucleosomes at the yeast PHO84 promoter are determined by intrinsic nucleosome stability. *Mol. Cell. Biol.* **29**, 2960–2981 (2009).
- Korber, P. & Hörz, W. In vitro assembly of the characteristic chromatin organization at the yeast PHO5 promoter by a replication-independent extract system. *J. Biol. Chem.* **279**, 35113–35120 (2004).
- Hertel, C.B., Langst, G., Hörz, W. & Korber, P. Nucleosome stability at the yeast PHO5 and PHO8 promoters correlates with differential cofactor requirements for chromatin opening. *Mol. Cell. Biol.* **25**, 10755–10767 (2005).
- McManus, J. *et al.* Unusual chromosome structure of fission yeast DNA in mouse cells. *J. Cell Sci.* **107**, 469–486 (1994).
- Bernardi, F., Zatchej, M. & Thoma, F. Species specific protein–DNA interactions may determine the chromatin units of genes in *S. cerevisiae* and in *S. pombe*. *EMBO J.* **11**, 1177–1185 (1992).



31. Lantermann, A., Stralfors, A., Fagerstrom-Billai, F., Korber, P. & Ekwall, K. Genome-wide mapping of nucleosome positions in *Schizosaccharomyces pombe*. *Methods* **48**, 218–225 (2009).
32. Dutrow, N. *et al.* Dynamic transcriptome of *Schizosaccharomyces pombe* shown by RNA-DNA hybrid mapping. *Nat. Genet.* **40**, 977–986 (2008).
33. Godde, J.S. & Widom, J. Chromatin structure of *Schizosaccharomyces pombe*. A nucleosome repeat length that is shorter than the chromatosomal DNA length. *J. Mol. Biol.* **226**, 1009–1025 (1992).
34. Bernardi, F., Koller, T. & Thoma, F. The *ade6* gene of the fission yeast *Schizosaccharomyces pombe* has the same chromatin structure in the chromosome and in plasmids. *Yeast* **7**, 547–558 (1991).
35. Thomas, J.O. & Furber, V. Yeast chromatin structure. *FEBS Lett.* **66**, 274–280 (1976).
36. Lee, C.K., Shibata, Y., Rao, B., Strahl, B.D. & Lieb, J.D. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat. Genet.* **36**, 900–905 (2004).
37. Bernstein, B.E., Liu, C.L., Humphrey, E.L., Perlstein, E.O. & Schreiber, S.L. Global nucleosome occupancy in yeast. *Genome Biol.* **5**, R62 (2004).
38. Wirén, M. *et al.* Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. *EMBO J.* **24**, 2906–2918 (2005).
39. Heichinger, C., Penkett, C.J., Bahler, J. & Nurse, P. Genome-wide characterization of fission yeast DNA replication origins. *EMBO J.* **25**, 5171–5179 (2006).
40. Fagerström-Billai, F., Durand-Dubief, M., Ekwall, K. & Wright, A.P. Individual subunits of the Ssn6-Tup11/12 corepressor are selectively required for repression of different target genes. *Mol. Cell. Biol.* **27**, 1069–1082 (2007).
41. Malavé, T.M. & Dent, S.Y. Transcriptional repression by Tup1-Ssn6. *Biochem. Cell Biol.* **84**, 437–443 (2006).
42. Ponomarenko, J.V. *et al.* Conformational and physicochemical DNA features specific for transcription factor binding sites. *Bioinformatics* **15**, 654–668 (1999).
43. Kornberg, R.D. & Stryer, L. Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Res.* **16**, 6677–6690 (1988).
44. Guillemette, B. *et al.* Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol.* **3**, e384 (2005).
45. Buchanan, L. *et al.* The *Schizosaccharomyces pombe* Jmjc-protein, Msc1, prevents H2A.Z localization in centromeric and subtelomeric chromatin domains. *PLoS Genet.* **5**, e1000726 (2009).
46. Zofall, M. *et al.* Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. *Nature* **461**, 419–422 (2009).
47. Varga-Weisz, P.D. *et al.* Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**, 598–602 (1997).
48. Flaus, A., Martin, D.M., Barton, G.J. & Owen-Hughes, T. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* **34**, 2887–2905 (2006).
49. Sugiyama, T. *et al.* SHREC, an effector complex for heterochromatic transcriptional silencing. *Cell* **128**, 491–504 (2007).
50. Cam, H.P. *et al.* Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* **37**, 809–819 (2005).
51. Johnsson, A. *et al.* HAT-HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress. *EMBO Rep.* **10**, 1009–1014 (2009).

ONLINE METHODS

Genome-wide nucleosome mapping. *S. pombe* strains were obtained from K. Ekwall: wild type (HU303, *h*⁻), *mit1* (HU1295, *h*⁻, *mit1::kanMX6*, *leu1-32*, *ade6-210*, *ura4-DS/E*), *fft3* (HU1939, *h*⁻, *fft3::hph*, *leu1-32*, *ade6-210*, *ura4-DS/E* or *D18*), *tup11 tup12* (Hu0946, *h*⁺, *tup11::ura4*, *tup12::ura4*, *ade6-M210*, *leu1-32*, *ura4-D18*). Genome-wide nucleosome mapping for *S. pombe* as well as MNase ladders and MNase indirect end-labeling for *S. pombe* and *S. cerevisiae* were done as described in detail^{31,52}. Briefly, we cross-linked logarithmically growing cells with formaldehyde and lysed them with zymolyase, and then isolated nuclei and immediately digested them with MNase to yield DNA fragments of mononucleosomal length. Before hybridization, we further fragmented the mononucleosomal DNA and biotin-labeled it with terminal deoxynucleotidyl transferase. We used the Affymetrix *S. pombe* Tiling 1.0FR array, which comprises 1.2 million probes representing the complete *S. pombe* genome at 20-bp resolution mapped to the *S. pombe* genome version of 15 September 2004. We used DNase I-fragmented genomic DNA from *S. pombe* as hybridization control. Restriction enzyme sites and probes used for MNase indirect end-labeling were as in **Supplementary Table 1**.

Annotation of TSS and TTS. We annotated transcriptional start and termination sites in *S. pombe* by visual inspection of the data from Dutrow *et al.*³² (**Supplementary Table 2**). We used these annotations to demarcate transcript length in **Figures 2e** and **4c** and in **Supplementary Figures 4f** and **5c**. For transcriptome analysis, we used the <http://bioserver.hci.utah.edu/BioInfo/index.php/Software:IGB> website and navigated from there to the Das2 server, where transcriptome data are available.

Generation of *S. pombe* expression data. We analyzed total RNA preparations by the hot phenol method³⁸ from cells grown to mid-logarithmic phase (5×10^6 cells per ml) in rich medium on the Affymetrix Yeast genome 2.0 array. We used Gene Spring (Agilent) to analyze the microarray data. We performed three independent experiments with RNA preparations from wild type and two with preparations from *mit1* mutant cells. Genes that were identified as being reproducibly up- or downregulated by a factor of 2 or greater in *mit1* compared to wild-type cells are listed in **Supplementary Table 5**.

Processing of microarray data. We carried out signal processing and downstream analyses using R/Bioconductor (<http://www.r-project.org>, <http://www.bioconductor.org>). All functions were called using default parameters if not indicated otherwise. *S. pombe* genome sequences and annotations (version of 16 July 2008) were obtained from the Sanger Genome Project (http://www.sanger.ac.uk/Projects/S_pombe). We mapped Affymetrix GeneChip *S. pombe* Tiling 1.0FR array probes to the genome dated 16 July 2008 using NCBI MegaBlast (<http://www.ncbi.nlm.nih.gov/blast/megablast.shtml>) and removed probes matching more than one genomic location. We normalized the raw signals using the “vsn” algorithm⁵³ and calculated the nucleosome occupancy as \log_2 of the ratio of mononucleosome to genomic DNA signals. This study comprises four biological replicates of wild type, three biological replicates of *mit1* and two biological replicates of *fft3* and *tup11 tup12* mononucleosome hybridizations. All of them were normalized to four genomic DNA preparations, three of which were obtained from the wild-type strain and one from the *mit1* mutant.

We calculated cumulative profiles by averaging sliding window values along genomic features as denoted in the figure legends. If not indicated otherwise, we used a step size of 10 bp and a window size of 50 bp for nucleosome signals and a step size of 10 and window size of 200 bp for RNA polymerase II data. We clustered nucleosome occupancy patterns based on the region from -370 to +500 bp relative to the TSS using “hclust” (R package “Stats”) on scaled profiles using Ward’s minimum variance method. For scatter-plot correlation analysis, we used the average nucleosome occupancy in the region from -300 to 0 bp relative to the TSS. We calculated spectral densities using “spec.pgram” (R package “Stats”) on equally spaced (50-bp window, 10-bp step size) nucleosome occupancy signals including de-meaning, a padding proportion of 1 and Daniell smoother widths of 5. We sampled spectral densities for 1-kb windows with a 500-bp overlap all along the chromosomes. We processed Affymetrix Yeast Genome 2.0 Array data

with “gcrma” to calculate expression values. For **Figure 4a**, we defined NDRs with a hidden Markov model calculated with TileMap (<http://biogibbs.stanford.edu/~jihk/TileMap/index.htm>). We defined all signals that revealed a strong depletion of the nucleosome density over ten or more following probes as NDRs. In total, 2,839 NDRs were identified this way, two-thirds of them localized in promoter regions (-500 bp to +100 bp relative to TSS). Only those NDRs that could be unambiguously assigned as being closest to the TSS were used for the alignment in **Figure 4a**. Of the other identified NDRs, 73 were localized within transcripts (from +100 bp to -100 bp relative to TTS), 43 in the 3' regions of genes (-100 bp to +200 bp relative to TTS) and 586 elsewhere.

External data sources. *S. pombe* RNA polymerase II binding data⁵⁴ were derived from ArrayExpress (E-MTAB-18). We remapped the origin annotations from Heichinger *et al.*³⁹ to the *S. pombe* genome version dated 16 July 2008. The 50% of origins above the median efficiency were defined as “high efficiency” and the 50% below as “low efficiency” origins, respectively. *S. pombe* H2A.Z occupancy data were from Buchanan *et al.*⁴⁵ and Tup11 and Tup12 occupancy data from Fagerström-Billai *et al.*⁴⁰. *S. cerevisiae* nucleosome mapping data⁵ were obtained from ArrayExpress (E-MEXP-1172) and processed as for the *S. pombe* data. The *S. cerevisiae* TSS annotations according to Lee *et al.*⁵ were kindly communicated by W. Lee. *S. cerevisiae* RNA expression data were from David *et al.*⁵⁵.

Analysis of DNA sequence contributions to nucleosome positioning. The predicted nucleosome occupancy was calculated by using the N-score model we developed previously¹⁶ with minor modifications⁵⁶. Briefly, for each species, we selected 8,000 loci each corresponding to the highest or lowest log ratio in the microarray data as a training set. The microarray data for *S. cerevisiae* were obtained from Lee *et al.*⁵. The 129-bp genomic sequence centered at each locus was extracted, converted to 16 dinucleotide frequencies and wavelet-transformed with the Haar basis. We then built a stepwise logistic regression classification model by combining wavelet energies coefficients, word counts¹⁸ and structural parameters^{5,42} as predicting variables. Each model was applied to calculate the genome-wide scores for both species. We used the 2008 genome version for *S. pombe* and the 2003 genome version for *S. cerevisiae*.

The analysis of the frequency of DNA ‘sequence words’ in NDRs (**Supplementary Table 4**) was done as described⁴ with probes of all NDRs as defined by TileMap for **Figure 4a** and all other probes as reference set.

Preparation of figures. Bioinformatic data analysis plots were generated using the statistical package of R/Bioconductor. File size of large plots was reduced by conversion to TIFF files in Adobe Photoshop CS2. Hybridized Southern blots were exposed to X-ray films (Fuji Super RX) at -80 °C using intensifier screens (DuPont, Lightening Plus). Films were scanned in CMYK modus (MikroTek ScanMaker i900). Agarose gel images were taken with a gel documentation apparatus (Peglab) using the software Vision Capt (Peglab). Digital images of agarose gels or scanned Southern blots were imported in Adobe Photoshop CS2 and further manipulated by conversion into grayscale format and linear level adjustment. The resolution was reduced from 16-bit to 8-bit channel. All plots and images were imported in Adobe Illustrator CS2 for final figure layout.

URL. The MATLAB code for the authors’ N-score model can be downloaded from G.-C.Y.’s website (<http://bcb.dfc.harvard.edu/~gcyuan>).

52. Svaren, J., Venter, U. & Hörz, W. *In vivo* analysis of nucleosome structure and transcription factor binding in *Saccharomyces cerevisiae*. *Methods Mol. Genet.* **6**, 153–167 (1995).
53. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. & Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18** (Suppl. 1), S96–S104 (2002).
54. Wilhelm, B.T. *et al.* Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239–1243 (2008).
55. David, L. *et al.* A high-resolution map of transcription in the yeast genome. *Proc. Natl. Acad. Sci. USA* **103**, 5320–5325 (2006).
56. Yuan, G.C. Targeted recruitment of histone modifications in humans predicted by genomic sequences. *J. Comput. Biol.* **16**, 341–355 (2009).