Supplementary Methods

1. MNase digestion and sample preparation for high throughput sequencing
Cells were grown to OD_{600} of 0.8-1.0 at 30°C from overnight starter diluted to 100 ml of YPD medium at OD_{600} of 0.15. Formaldehyde solution (37%) was added to final concentration of 1%, followed by shaking at 30 °C for 15 min. The formaldehyde was quenched by adding glycine to 0.125 M and letting to stand at room temperature for 5 min. Cells were collected by centrifugation at 3000 g for 5 min and washed twice with same volume of ice-cold water. At this point pellet can be stored at -80°C. Cells were resuspended in 10 ml Buffer Z (1M sorbitol, 50 mM Tris-HCl pH = 7.5, freshly added 10 mM β- mercaptoethanol), 70 µl of zymolase (1000units/ml) was added and the cells were incubated at 30°C for 30 min. Then the spheroplasts were collected by centrifugation at 3000g for 10 min, washed with 10 ml Buffer Z and resuspended in 0.5 ml Buffer NP (1M sorbitol, 10 mM Tris- HCl pH = 7.5, 50 mM NaCl, 5 mM MgCl2, 1 mM CaCl2, freshly added 1 mM β- mercaptoethanol, NP-40 0.075%, 0.5 mM spermidine). Then the sample was split for 2 aliquots, digested for 20 min at 37°C with 0.5 and 2 units of MNase. The reactions were stopped by addition of 0.1 ml of 5% SDS with 50 mM EDTA to each reaction. Cross-linking was reversed at 65°C for 4 h, samples were diluted with 0.4 ml water and 5 µl of RNase A (10 mg/ml , Sigma) were added. The samples were incubated at 37°C for 1 h, then 5 µl of Proteinase K (20 mg/ml) were added to each sample and were incubated at 65°C for 1 h. The DNA was extracted by phenol-chloroform-isoamyl alcohol (PCI), washed with chloroform-isoamyl alcohol (CI) and precipitated with ethanol. The digestion products were resolved on a 2% agarose gel, and mononucleosomal size band (~150 bp) was isolated by Wizard SV gel clean-up kit (Promega). An additional ethanol precipitation was performed to improve purity of the sample and to concentrate the input DNA.

2. Control experiments with naked DNA
For control, cells grown to OD_{600} of 0.8-1.0 were collected by centrifugation at 3000 g for 5 min and washed twice with same volume of ice-cold water. Cells were resuspended in 10 ml Buffer Z, 70 µl of zymolase was added and the cells were incubated at 30°C for 30 min. Then the spheroplasts were collected by centrifugation at 3000g for 10 min and washed with 10 ml Buffer Z. The samples were resuspended in 1 ml 50 mM EDTA, 0.3% SDS and incubated at 65°C for 20 min. After chilling on ice, 0.2 ml of 3M potassium acetate, 11.5 % acetic acid solution was added and samples were allowed to stand for 2 h on ice. Precipitate was removed by 5 min centrifugation at 5000 rpm at 4°C. DNA was recovered by ethanol precipitation, dissolved in 0.4 ml water and extracted by PCI, followed by CI wash and ethanol precipitation. Naked genomic DNA was resuspended in 0.3 ml Buffer NP, the sample was split for 3 aliquots, each digested for 10 min at 37°C with 0.05 to 0.2 units of MNase. The reactions were stopped by addition of 20 µl of 5% SDS with 50 mM EDTA, re-united, and 5 µl of Proteinase K were added. After 1 h incubation at 37°C, DNA was extracted by PCI, washed with CI and precipitated with ethanol. The digestion products were resolved on a 2% agarose gel, and band of 150±20 bp size was isolated by Wizard SV gel clean-up kit (Promega). An additional ethanol precipitation was performed to improve purity of the sample and to concentrate the input DNA.

3. Library preparation and sequencing
The similar amounts of DNA from three independent biological repeats were mixed and diluted to 1 ng/ml concentration. The ChIP-Seq DNA Sample Prep Kit (IP-102-1001) and the standard protocol by Illumina were used to create libraries for sequencing, with the modification: the fragments with ligated adaptors were subjected directly to amplification, skipping the gel purification and size selection step. All other steps of cluster formation and sequencing procedure were performed following standard protocols for Illumina GA2 instrument.

4. Analysis of high-throughput sequencing
For each lane of Illumina sequencing, reads of 34-40bps were mapped to the genomic sequences of *S. cerevisiae* and *S. paradoxus* with two separate runs of Eland, allowing up to two mismatches within the first 32bps. ~50% of the reads were mapped to a single location in only one of the genomes, or were mapped to single locations in both genomes but with at least two more mismatches to one genome (e.g. a read with one mismatch to *S. cerevisiae* should have at least three mismatches to *S. paradoxus*). These reads could be confidently mapped to a specific location in one of the genomes. Even if there is a sequencing error in a certain read, that read would still have higher similarity to the correct genome than to the other genome. The only exception for this would be reads in which all base-pairs that differ between the species have sequencing errors and thus do not discriminate between the species. The probability for this to occur is equal to p^s, where p is the sequencing error-rate (per base), and s is the number of sites that differ between the species at the respective genomic position. Our estimated error-rate is ~0.15 for an entire read (i.e. each read has on average 0.15 sequencing errors) and ~0.0045 for a single base (i.e. p=0.0045). This means that at genomic positions with a single base difference between the species we would exclude less than one percent (~0.45%) of the reads due to sequencing-errors and at positions with two base differences we would exclude only ~0.002% of the reads (20 out of a million). Thus, in our analysis of reads that differ between the species by at least 2 mutations, the combined sequencing of the two species has a negligible effect (exclusion of less than ~0.002% of the reads).

Since reads of ~36bps corresponded to the ends of ~150bp fragments, the location of each mapped read was converted into the expected center position of the original DNA fragment. This was done by assuming a constant fragment length for each sample. This length was estimated as the median distance between peaks of reads in the forward strand and consecutive peaks of reads from the reverse strand.

The above analyses were performed for the combined samples of the two species, for the samples of the hybrid and for the control samples of naked DNA taken from the two species or from the hybrid. Thus, all biases resulting from the mapping procedure, sequencing errors or alignment errors should affect the species, hybrid and control datasets in the same way and thus would not lead to errors in our final analysis. For example, if a certain region could not be mapped to *S. paradoxus* due to sequencing error, then we would get zero signal for *S. paradoxus* both in the species and in the control sample, and thus would not call this position a nucleosome gain/loss.

We analyzed genomic positions that correspond to aligned promoters and coding regions of the two species (Kellis *et al*, 2003), which included 1kb upstream and 1kb

downstream of the aligned start codons. We excluded alignments with sequence gaps, with alignment gaps beyond 20bps, and genes without one-to-one orthology relationships. Similar analysis was performed for the 2kb regions surrounding the aligned stop codons and for the aligned sequences of the entire genes.

5. Filtering of nucleosomes does not seem to affect our main conclusions

The reported analysis is for ~50% of the yeast nucleosomes since we wanted to focus on the most reliable cases, and excluded the following genes and nucleosomes:

1. Genes that lack one-to-one-orthology relationships and/or complete sequence alignments (~15% of nucleosomes).
2. Differential occupancy similar to that seen in the genomic DNA controls (~30% of nucleosomes). MNase bias affects most genome-wide studies of nucleosome positioning and is particularly important in inter-species comparisons. We thus took special care to avoid it.
3. Nucleosome-reads that could not discriminate between the species due to high conservation of the respective genomic positions and our use of co-sequencing of the two species. Here, we demanded that mapped sequence reads will discriminate between the genomes by at least two-nucleotides, filtering out ~13% of nucleosomes.
4. We focused on the promoters and 2kb of each gene and thus ignored some portions of long genes, filtering out ~5% of nucleosomes.

To examine whether this filtering biases our results we relaxed our filtering criteria by (i) Adding reads that differ by only one base between the species (additional ~9% nucleosomes), (ii) including the complete sequences of genes (additional ~5% nucleosomes), and (iii) not correcting for MNase bias (similar to previous studies, additional ~30% nucleosomes). We repeated the analysis for ~50,000 nucleosomes, which consists of ~80% of all nucleosomes. The results were practically identical to those from our previous analysis. In particular, the fraction of nucleosomal differences remained ~10%, the fraction of cis effects remained ~70%, and there was still no correlation between nucleosomal and expression changes (Fig. S11).

This analysis still excluded ~4% of genomic positions with high sequence conservation (reads that cannot discriminate between the species) and genes without one-to-one-orthology relationships and/or complete sequence alignment. We verified that there is no considerable difference in the GC-content of addressable nucleosomes with non-addressable nucleosomes (Fig. S11).

6. Analysis of nucleosome positioning

After mapping the reads to the aligned positions in the two genomes, we obtained the number of reads that mapped to each base-pair of the alignments. This data was then transformed to "nucleosome occupancy", i.e. the number of reads that cover each base-pair, assuming that reads correspond to mono-nucleosome fragments of 150bps. Nucleosome occupancy data was used in several analysis for comparing overall occupancy at specific regions (e.g. Fig. 5), but this data is less suitable for prediction of exact nucleosome positions. Thus, we also defined "nucleosome scores", by Gaussian filtering of the number of reads at each base-pair, with a window of 50bps and standard deviation of 25bps (Albert *et al*, 2007). This transformation produces sharper peaks and allows a better estimation of nucleosome positions. We estimated the positions of nucleosomes as peaks of nucleosome scores, which were (i) not

among the 10% peaks with lowest scores, and (ii) not within 100bps of another peak of higher score. The number of nucleosomes defined by these criteria corresponded to ~80% of nucleosomal DNA and 20% of linker DNA, as estimated by previous studies (Lee *et al*, 2007). Each predicted nucleosome was also assigned an "occupancy level" defined as the number of reads mapped to at most 30bp from the estimated nucleosome center position.

Nucleosome scores from all samples were normalized to the same distribution using percentile normalization. These normalized scores were used in all figures to enable a visual comparison of the nucleosome patterns. The raw data of mapped reads and the normalized nucleosome scores will be available at the GEO and SRA databases.

7. Inter-species comparison of nucleosome positioning

At each aligned region we identified pairs of nucleosomes with the most similar positions in the two species. If two nucleosomes from one species paired with the same nucleosome from the other species, then the one which is more distant from the single nucleosome was regarded as a possible nucleosome gain/loss. Paired nucleosomes whose positions differed by at least 30bps between the species (and at most 80bp) were regarded as a possible nucleosome shift. Nucleosomes whose occupancy level differed by at least 2-fold (after correcting for the overall difference in occupancy levels between the corresponding samples of the two species) were regarded as a possible occupancy change.

Each potential nucleosome gain/loss was also required to have at least 2-fold higher occupancy at the species with nucleosome (compared with the species without nucleosome) and that this nucleosome will be supported by at least 8 reads (that map to at most 30bps of the predicted center position). Similarly, a potential nucleosome shift (from position $P_{s.cer}$ to $P_{s.par}$) was required (i) to have at least 2-fold more *S. cerevisiae* reads around $P_{s.cer}$ and 2-fold more *S. paradoxus* reads around $P_{s.par}$, (ii) that each of these nucleosomes will be supported by at least 8 reads, and (iii) that the inter-species difference in reads distribution (at the region containing $P_{s.cer}$, $P_{s.par}$ and 30bps to each direction) would be significant ($P<0.05$, two-sample t-test).

To further increase the confidence of the predicted nucleosomal changes we repeated the analysis above only for the reads that mapped to the forward strand and (separately) only for the reads that mapped to the reverse strand. We required that (i) potential changes would pass all of the above thresholds in either one of the strands, (ii) that nucleosomes at potential changes are mapped in the forward and reverse analyses to within 30bps of their positions in the combined analysis, and (iii) that nucleosomes with potential shifts were mapped in both the forward and reverse analyses to within 20bps of their positions in the combined analysis, but to more than 20bps away from the position of the nucleosomes in the other species.

8. Controlling for MNase bias

Following standard procedures, our protocol for isolating mono-nucleosomes involved DNA digestion by MNase. Since MNase has an inherent sequence-dependent bias (Dingwall *et al*, 1981; Horz and Altenburger, 1981), observed inter-species differences could also reflect sequence divergence unrelated to nucleosome positioning. To control for this possibility, we repeated the experiment using naked

DNA, with duplicates, and excluded all differences that might arise due to MNase bias (Fig. S1).

Naked DNA from the two species (or the hybrid) was digested with MNase, pooled (for the two species) and sequenced. Sequencing data was processed as the nucleosomal data and averaged over the four samples (duplicates for the two species combined and duplicates for the hybrid). The same procedure for comparison of nucleosome positioning between the two species was performed for naked DNA, and for each position with differential occupancy, gain/loss or shift observed in nucleosomal DNA we calculated the difference in number of mapped reads for the naked DNA experiment. As shown in Fig. S1 for occupancy changes, with a threshold of zero, ~40% of the changes remain. In these cases nucleosomal and naked DNA must have opposite direction of differences, indicating that at least 40% of the observed occupancy changes are not caused by MNase digestion bias. With a threshold of one (i.e. 2-fold difference in naked DNA) which is the same threshold used to define occupancy changes at the nucleosomal DNA, ~64% of the changes remain, suggesting that at most 64% of the observed occupancy changes are not caused by MNase bias. Thus, the percentage of changes that are not due to MNase bias is probably between 40% and 64%. We chose to use an intermediate threshold of 1.5-fold occupancy difference (~0.6 in $\log_2$) in the naked DNA which includes 54% of the observed occupancy changes in nucleosomal DNA and excludes the rest. Similar criteria was used to exclude gains/losses and shifts (if any one of the nucleosome centers had a difference larger than 1.5-fold in the naked DNA read density. Note that this control probably also eliminates many real differences that are correlated with MNase bias and thus we underestimate the amount of occupancy changes, but eliminate most of the artifacts.

This analysis excluded ~30% of all nucleosomes, and thus we identified ~2400 inter-species nucleosome differences (in each of the strains, i.e. WT or mutants) among ~24,000 nucleosomes in which we can identify such differences, which correspond to a 10% frequency of nucleosome differences. This frequency increases when we relax the criteria for defining differences but remains at 10%-20% for various criteria (not shown).

9. Comparison of biological repeats

Nucleosome positioning for the mutant strains of *Δhtz1*, *Δgcn5* and *Δisw1* in *S. cerevisiae* and in *S. paradoxus* were each measured with two biological repeats. For all other analyses in this work the biological repeats were combined into a single dataset. However, in order to estimate the number of changes that would be observed for the same strain in identical conditions we compared these biological repeats (Fig. S2). The number of inter-species differences is ~3-fold higher than the number of differences observed between biological repeats. Moreover, less than one percent of the differences between biological repeats are consistently observed in at least 3 comparisons (only 3 occupancy changes and zero losses or shifts), in contrast to approximately half of the inter-species changes, suggesting the reliability of the observed changes.

10. Cis versus trans differences in nucleosome positioning

At each position of an inter-species difference we compared the difference in occupancy between the two species (Δparents) to the difference in occupancy between

the two hybrid alleles (Δhybrid). For loss and occupancy changes these differences were defined as $\log_2(O_{s.cer}/O_{s.par})$, where $O_{s.cer}$,$O_{s.par}$ are the number of reads mapped to within 30bps of the predicted location of the difference, in *S. cerevisiae* and *S. paradoxus*, respectively, or in the corresponding hybrid alleles. For Shifts, these differences were defined as $\log_2(O^1_{s.cer}/O^1_{s.par})+ \log_2(O^2_{s.par}/O^2_{s.cer})$, where O^1,O^2 are the same as above for the positions of the *S. cerevisiae* and *S. paradoxus* nucleosomes, respectively. We expect to have Δhybrid approximately equal to Δparents in cis-differences and approximately zero in trans-differences, and thus we compared |Δhybrid-Δparents| to |Δhybrid|. If one of these terms is larger than the other by at least a factor of 1.2 then we classify the difference as a cis or trans effect, respectively. To avoid spurious assignments, trans-differences were also required to have |Δhybrid|<0.6 (or |Δhybrid|<1.5 if Δhybrid has an opposite sign to Δparents). Cis differences were required to have |Δhybrid|>0.6. In some cases $O_{s.cer}$ or $O_{s.par}$ were zero so that Δhybrid or Δparents (and therefore also |Δhybrid-Δparents|) were infinite. To avoid misclassification of cis-differences due to inflation of |Δhybrid-Δparents|, all cases in which Δhybrid has the same sign as Δparents and |Δhybrid|>1 were also defined as cis.

The above analysis is inevitably parameter dependent, and therefore we also used other analyses to estimate the overall frequency of cis and trans differences. First, we varied the parameters described above. Second, we used a different strategy to obtain a general estimate of the number of cis and trans differences. The amount of trans differences was estimated as twice the number of differences in which Δhybrid and Δparents have opposite signs, and the remaining differences were estimated as cis (Fig. 2c). The rationale behind this is that if every difference is either only cis or only trans, then trans-differences should have Δhybrid values with a symmetric distribution around zero, such that in half of the cases Δhybrid has the same sign as Δparents and in the other half it has the opposite sign. In contrast, cis-differences should always have the same sign of Δhybrid and Δparents. The different analyses gave similar results with 50%-90% cis-differences and in most cases around 70% (not shown). Notably, the predicted frequency of cis-differences was highest for shifts. We suspect that this is because trans-shifts are often small, as observed for *ISW2* (Whitehouse *et al*, 2007) and *ISW1* (unpublished data), and are thus not detected by our stringent criteria for identifying large shifts. Indeed, when we relax the criteria of shifts to larger than 15bps (instead of 30bps), the frequency of cis-differences drops from 83% to 72%, which we therefore use as the estimate in Fig. 2c.

11. Promoter regulatory sites

Transcription start sites were defined as the median of those defined from multiple previous studies (Lee *et al*, 2007; Miura *et al*, 2006; Nagalakshmi *et al*, 2008). Transcription factor binding sites in *S. cerevisiae* were estimated from previous work (MacIsaac *et al*, 2006) using the dataset of P<0.005 and no conservation criteria, and their presence in *S. paradoxus* was estimated from the conservation of the binding site motifs in the aligned positions (the aligned *S. paradoxus* sequence does not have to be completely conserved but only to match the consensus motif defined by MacIsaac et al.). Binding sites that are present in S. paradoxus and absent in S. cerevisiae were identified by searching for promoters where*:* (i) *S. paradoxus* has a match to the consensus motif, (ii) the aligned position in *S. cerevisiae* is mutated and does not

match the consensus motif, and (iii) no binding was identified for the respective TF in *S. cerevisiae*. Note that these cases (cer-, par+) do not rely on experimental support for TF binding and are thus less confident than the opposite cases (cer+, par-).

Supplementary Figures

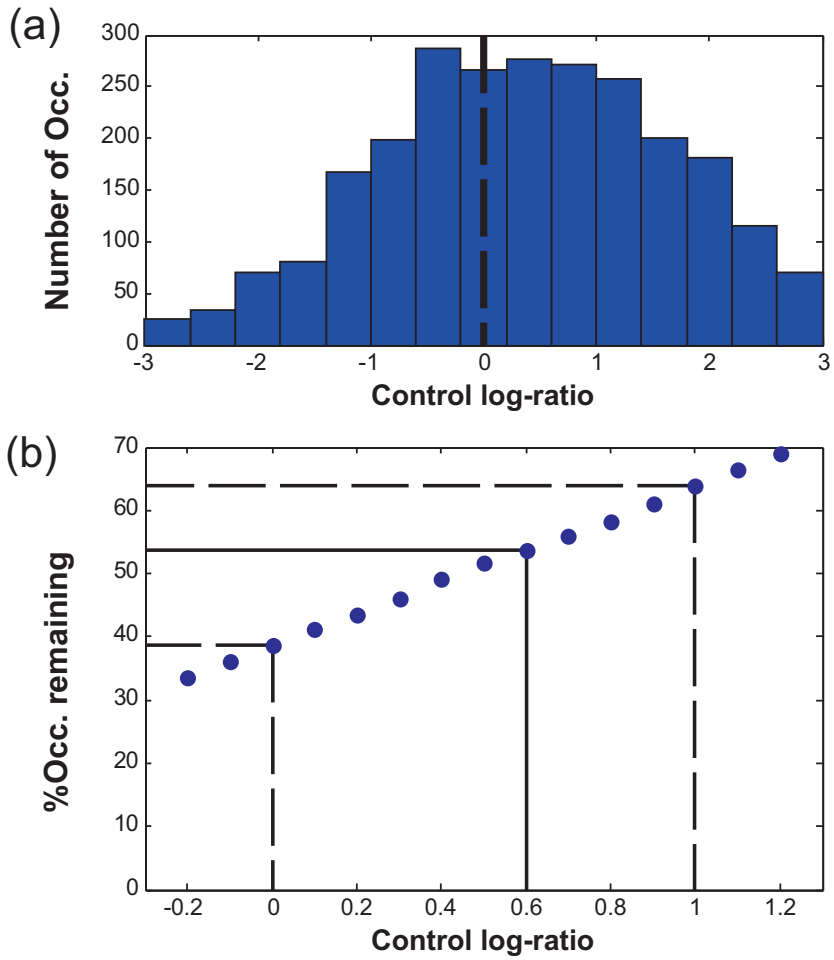


Figure S1. Controlling for MNase bias with naked DNA.
(a) Histogram of $\log_2$-ratios from the naked DNA experiment, where positive values indicate the same difference as in the nucleosomal DNA and negative values indicate an opposite difference compared with nucleosomal DNA (e.g. if *S. cerevisiae* had higher occupancy in nucleosomal DNA but lower occupancy in naked DNA). (b) The percentage of occupancy changes that remain after excluding those that show a difference in naked DNA beyond each threshold is shown on the *y*-axis, and the thresholds are shown on the *x*-axis. We chose to use an intermediate threshold of 1.5-fold occupancy difference (~0.6) in the naked DNA which includes 54% of the observed occupancy changes in nucleosomal DNA and excludes the rest.

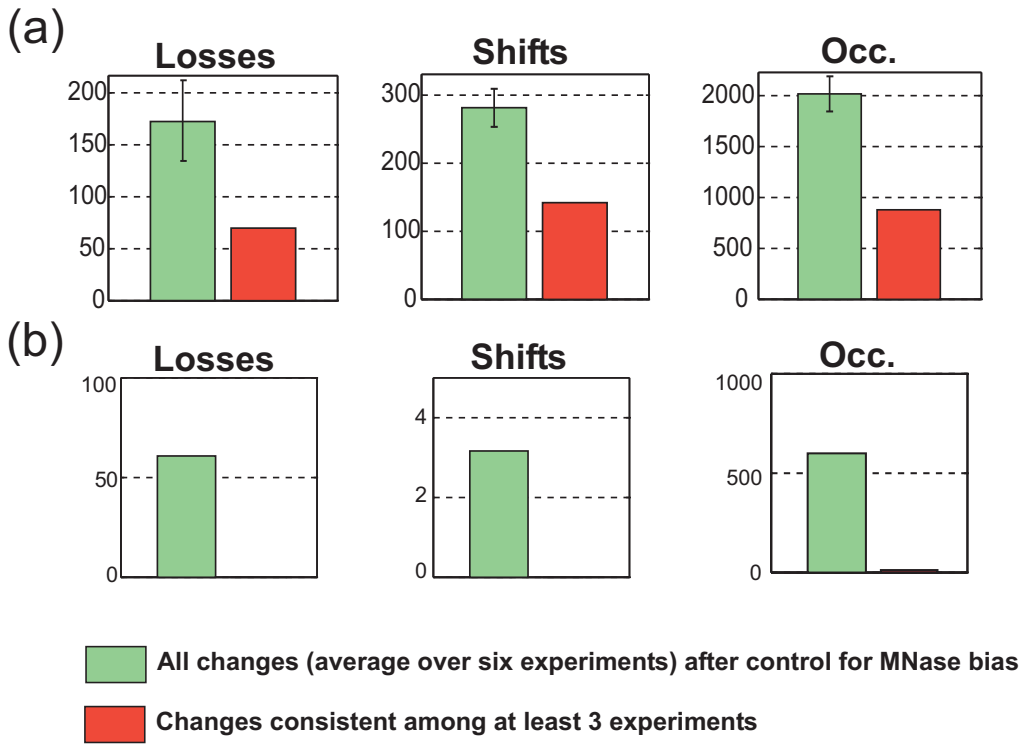


Figure S2. Number of changes observed between the species and in comparison of biological replicate experiments.
(a) Number of inter-species differences in nucleosome positioning. The average number across the six comparisons (WT and 5 mutants) is shown in green and the number of changes which are consistently observed in most comparisons (at least 3) is shown in red.
(b) Comparison of biological repeats. As in (a), green and red indicates the average number of changes and the number of consistent changes (at least 3 comparisons) among six comparisons of biological repeats (3 mutants for *S. cerevisiae* and for *S. paradoxus*).

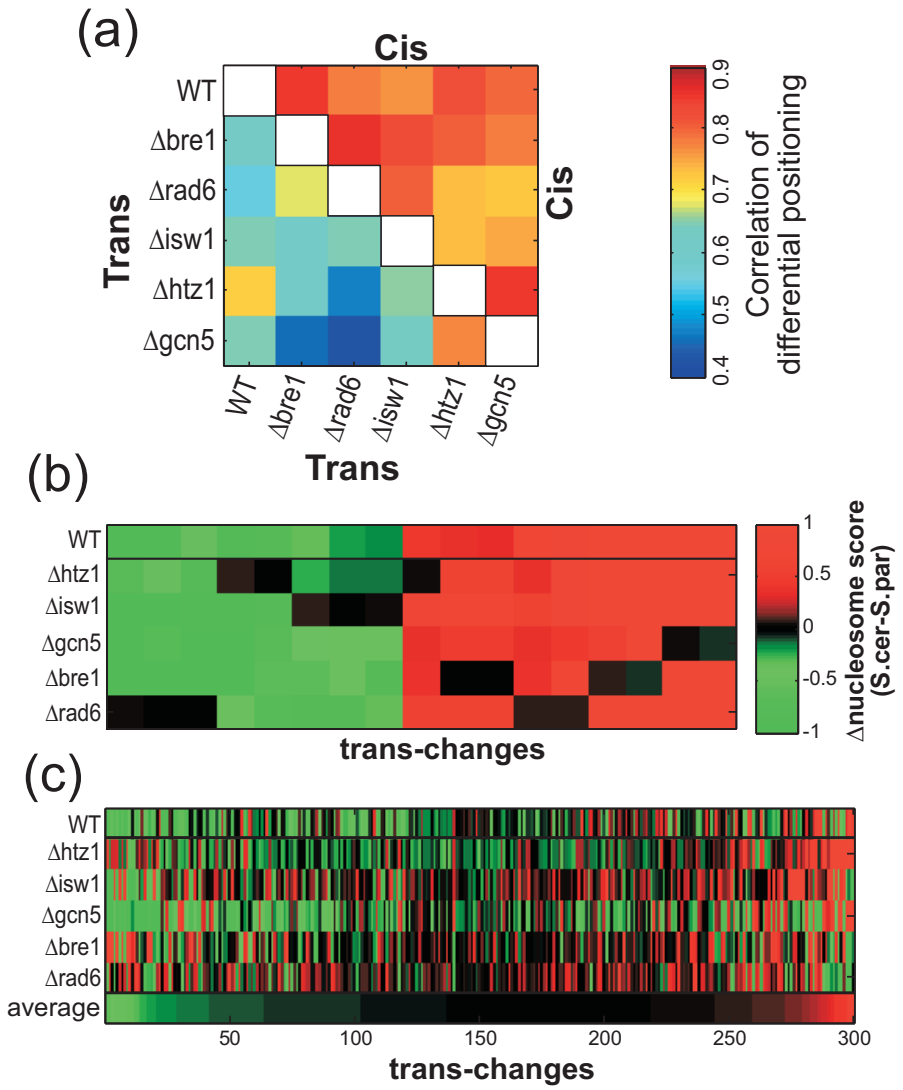


Figure S3. Properties of cis and trans changes in nucleosome positioning.
(a) Changes in cis are highly correlated among the different strains (WT and deletion strains) while changes in trans differ among strains. Shown are the correlations of differential occupancy between the different strains, for all positions with changes in cis (upper right) or all positions with changes in trans (lower left).
(b) Few trans-changes are only dependent on a particular chromatin regulator. Each column represents an inter-species trans-effect, and the colors indicate the difference in nucleosome score (Δscore) at the position of this effect for each of the strains (WT and mutants). Shown are 17 examples of trans-changes in which Δscores are similar among all but one of the strains. Overall, there were 53 trans-changes in which Δscores were different in one strain compared to all others ($P<0.05$) and |Δscore| was lowest in this strain, consistent with the possibility that these trans-changes are generated by divergence in the activity of the chromatin regulator deleted in the corresponding strain.
(c) Many trans-changes are variable among the different strains. Examples of the more frequent scenario in which Δscores are variable among the strains but that no single strain (or pair of strains) could account for this variability. Overall, there were 351 trans-changes in which (i) the STD of Δscores among the strains was at least 2-fold higher than the average of Δscores, and (ii) Δscores were not significantly different ($P>0.05$) in one or two of the trains with lowest |Δscores| compared with the other strains. These cases suggest complex regulation by multiple factors and may indicate that the trans-changes reflect overall differences in the state of cells from the two species, for example, as would be caused by differences in environmental sensing and signal transduction, rather than the effect of particular chromatin regulators. The average over all strains is also shown as the lowest row.

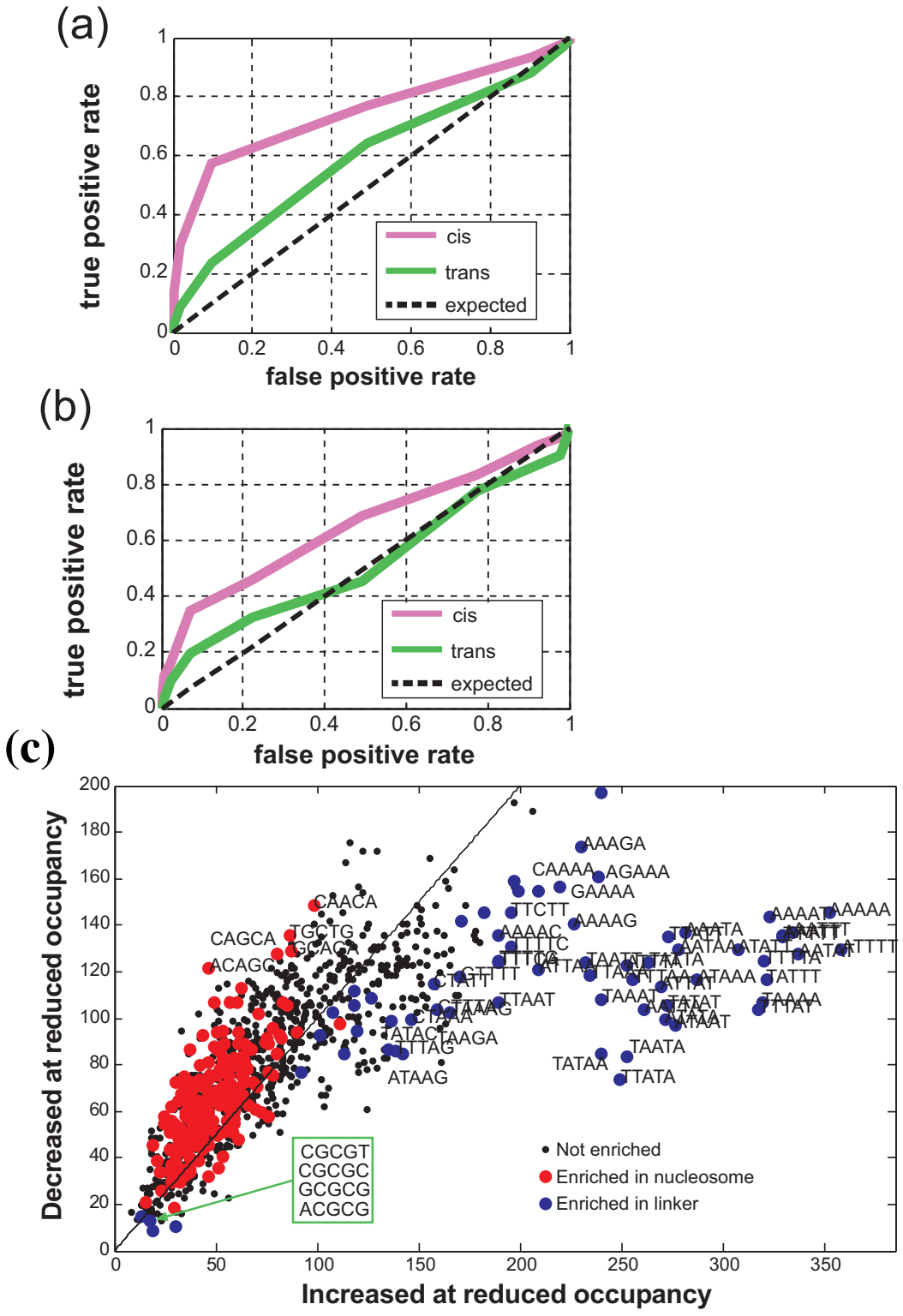


Figure S4. Sequence-based prediction of Occ. and Shift changes.
(a-b) For each predicted occupancy (a) or shift (b) change we calculated the difference in the predicted nucleosome scores (Field *et al*, 2008) of the two species at the location of change, and compared with the score differences at thousand randomly

selected locations in which nucleosome positioning is conserved (control). For each difference in scores we examined the frequency of control regions that have higher difference (false positive rate) and the frequency of nucleosome changes with higher score difference (true positive rate). The predictive power of cis occupancy changes is almost as high as that of cis nucleosome losses (Fig. 2d), but the predictive power of cis shifts is much lower. In contrast to our expectation, we do find some predictive power for trans occupancy or shift changes. One possible explanation for this is that some trans changes are also enhanced in cis (i.e. cis-trans interaction) while we predict them to be only trans. Another possibility is that trans differences are more easily identified if there is also MNase bias that increases their observed inter-species differences. This bias should also appear in the hybrid, but by itself could be too low to identify as cis difference. This MNase bias could then be captured by the sequence model allowing it to partially predict those trans differences. Finally, we could have made errors in classification of cis effects as trans due to noise in our data.

(c) 5-mers enriched at positions of occupancy changes.

The same analysis as in Fig. 2e was performed for occupancy changes. As in Fig. 2e, reduced nucleosome occupancy is associated with AT-rich 5-mers: 51 5-mers were enriched ($P<0.01$) at positions of reduced occupancy, including 32 with only A or T nucleotides and 19 with only one G or C nucleotide. 5-mers which are bound by the RSC complex were not enriched at positions of reduced occupancy, although they are generally associated with linker regions.

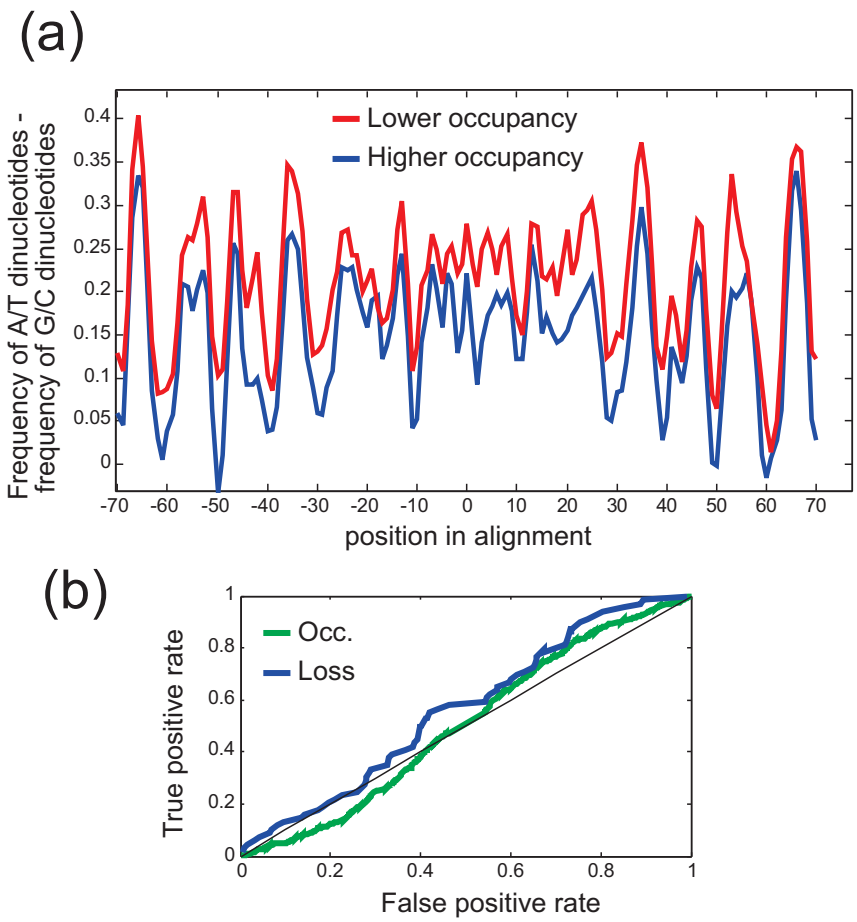


Figure S5. Dinucleotide patterns are not predictive of inter-species differences in nucleosome positioning. Previous studies have shown a specific pattern of A/T dinucleotides and an approximately opposite pattern of G/C dinucleotides at aligned nucleosomal sequences(Field *et al*, 2008; Segal *et al*, 2006). These patterns were taken from Segal et al.(Segal *et al*, 2006) and combined by subtracting the frequency of G/C dinucleotides from the frequency of A/T dinucleotides. For each position of nucleosome loss or occupancy change, we aligned the surrounding sequence of 150bp with the combined pattern of dinucleotides.

(a) The average pattern (frequency of A/T dinucleotides minus frequency of G/C dinucleotides) is shown for aligned positions of occupancy changes, with the sequences of higher occupancy in blue and sequences of lower occupancy in red. Sequences of the lower and higher occupancy nucleosomes have similar periodic patterns but with a difference in A/T content: nucleosomes with low occupancy have higher A/T content (higher values at the *y*-axis), consistent with the nucleosome-disfavoring activity of AT-rich sequences.

(b) For each nucleosome loss or occupancy change we compared the correlation of the respective sequences from the two species with the pattern of dinucleotides (after alignment with the pattern). We then examined whether nucleosome loss or reduced occupancy could be predicted by a reduced correlation with the pattern of dinucleotides (true positives) and compared this with the difference in correlation with the pattern at randomly selected positions (false positives). The results show that the pattern of dinucleotides is not predictive of changes in occupancy (green) or nucleosome losses (blue).

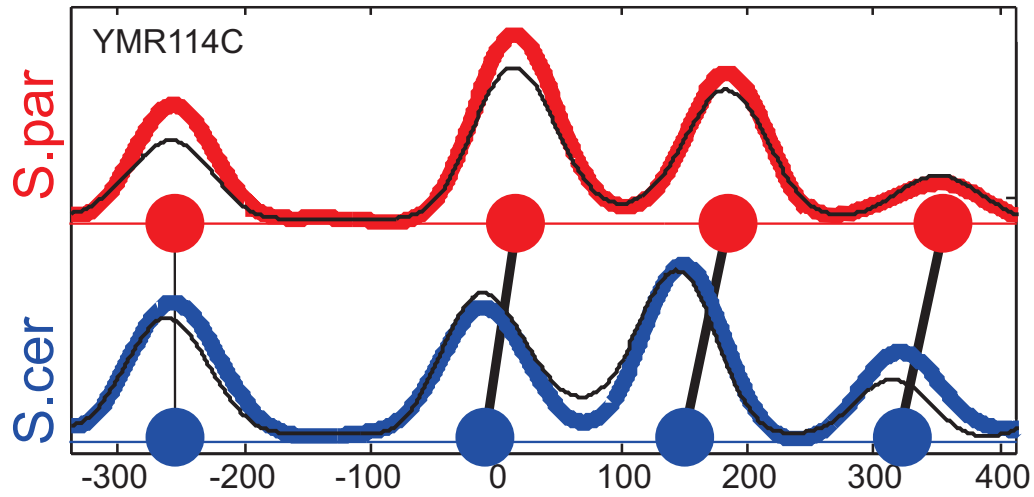


Figure S6. Shift propagation at the YMR114C gene. Red and blue curves display the nucleosome scores of the two species and the black lines display the nucleosome scores of the corresponding hybrid alleles. Nucleosomes +1, +2 and +3 are shifted downstream in *S. paradoxus*, compared to *S. cerevisiae*, but the -1 nucleosome has no shift.

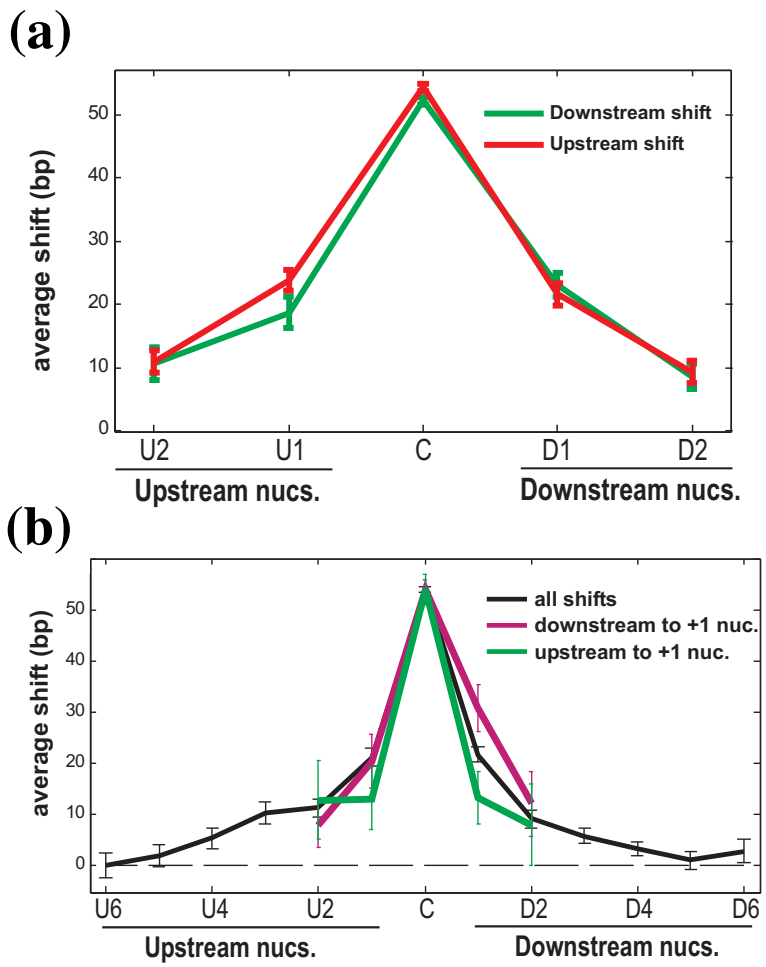


Figure S7. Shifts propagate to both directions and through +1 nucleosomes with strong positioning signals.
(a) Shifts propagate to both directions. The analysis in Fig. 3c was done separately for shifts in which the *S. cerevisiae* nucleosome is shifted downstream (green) and upstream (red). A downstream shift (green) could physically affect the downstream nucleosomes (D1 and D2) and thus the downstream nucleosomes could be "pushed". However, the upstream nucleosomes (U1 and U2) are in the opposite direction of the shift so their shifts are expected due to statistical positioning. The downstream shifts of the D1 nucleosomes are only slightly larger than the downstream shifts of the U1 nucleosomes and this difference disappears when comparing the D2 and U2 nucleosomes. Similarly, upstream shifts should only "push" the U1,U2 nucleosomes but their upstream shifts are only slightly larger than the upstream shifts of the D1,D2 nucleosomes. These results suggest that the propagation of shifts is mostly due to statistical positioning and not because of steric effects.
(b) Shifts propagate through +1 nucleosomes with strong positioning signals. The analysis in Fig. 3d was repeated for a third of the +1 nucleosomes with strongest nucleosome-favoring signals (Ioshikhes *et al*, 2006).

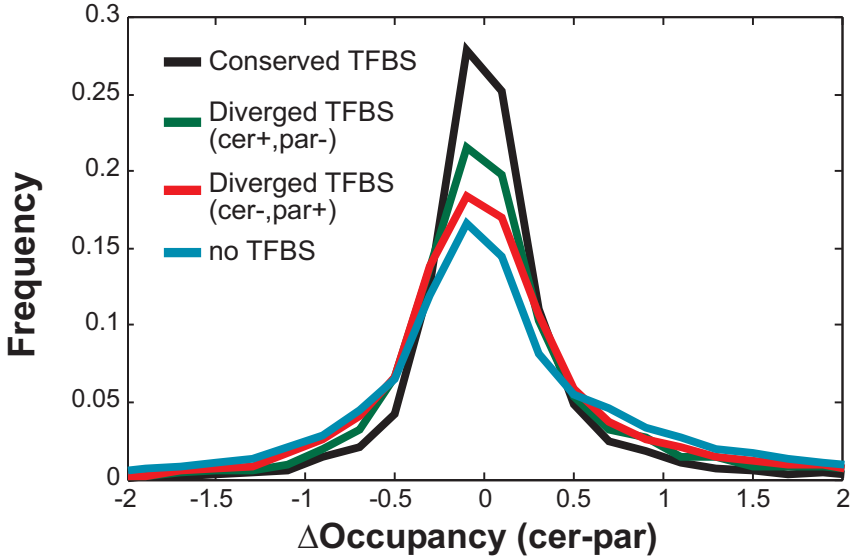
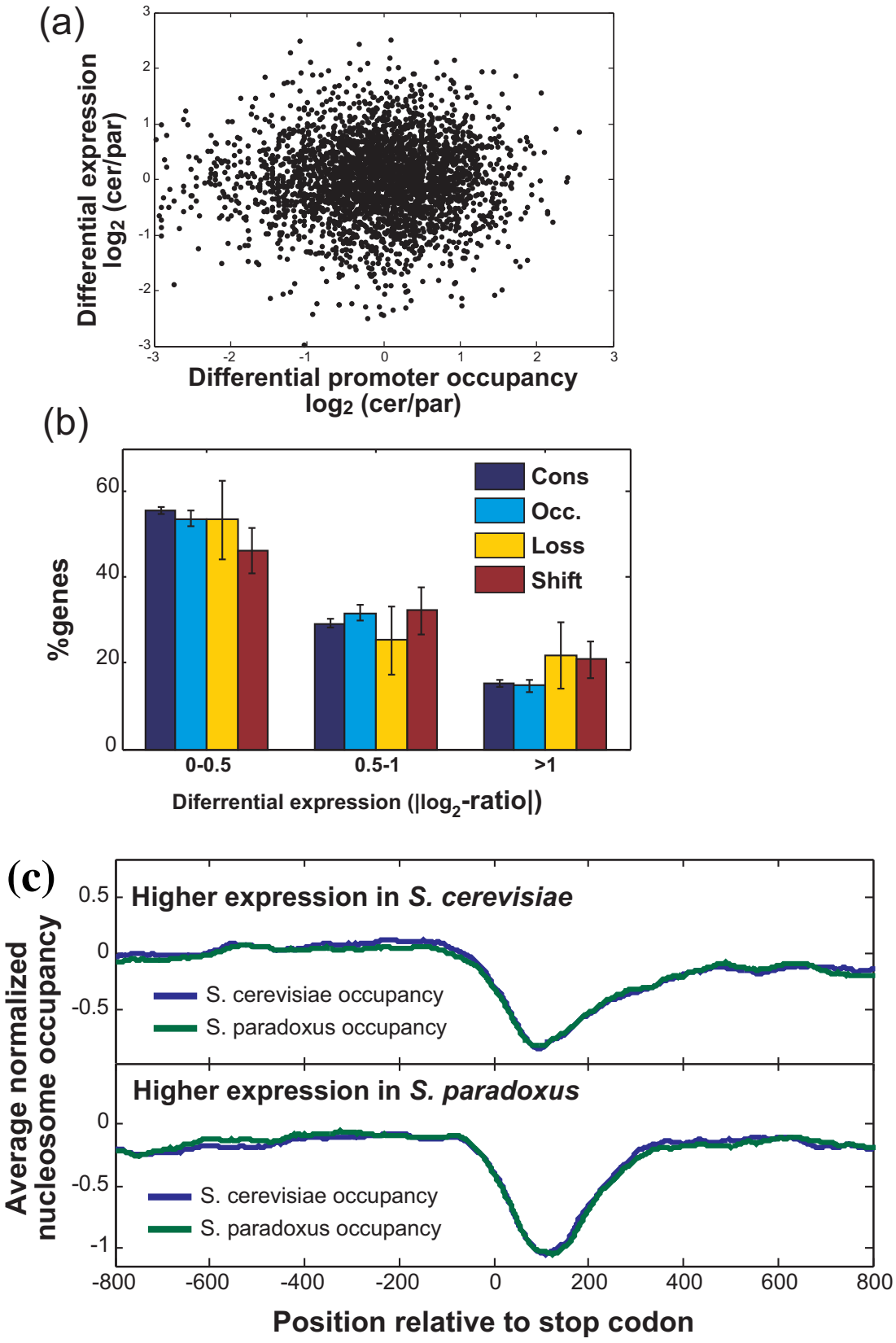


Figure S8. Divergence of nucleosome occupancy at positions of diverged transcription factor binding sites (TFBSs). TFBSs were classified as conserved or two classes of diverged (present in s. cerevisiae but not in S. partadoxus and vice versa) as described in the Supplementary methods (section 11). As a control, we also examined promoter positions without TFBSs, but at the same distance from the start codons (by shuffling the genes but not the positions of conserved TFBS). Shown are the distributions of inter-species differences in nucleosome occupancy at each of the four types of positions. Diverged TFBSs have an intermediate conservation of nucleosome occupancy, higher than no TFBSs but lower than conserved TFBSs. Notably, these distributions are not skewed for diverged TFs, indicating that divergence of TFBSs is not preferentially associated with either increase of decrease of nucleosome occupancy.


(a)
Differential expression
log2 (cer/par)
Differential promoter occupancy
log2 (cer/par)
(b)
%genes
Cons
Occ.
Loss
Shift
0-0.5
0.5-1
>1
Diferrential expression (|log2-ratio|)
(c)
Average normalized
nucleosome occupancy
Higher expression in S. cerevisiae
S. cerevisiae occupancy
S. paradoxus occupancy
Higher expression in S. paradoxus
S. cerevisiae occupancy
S. paradoxus occupancy
Position relative to stop codon

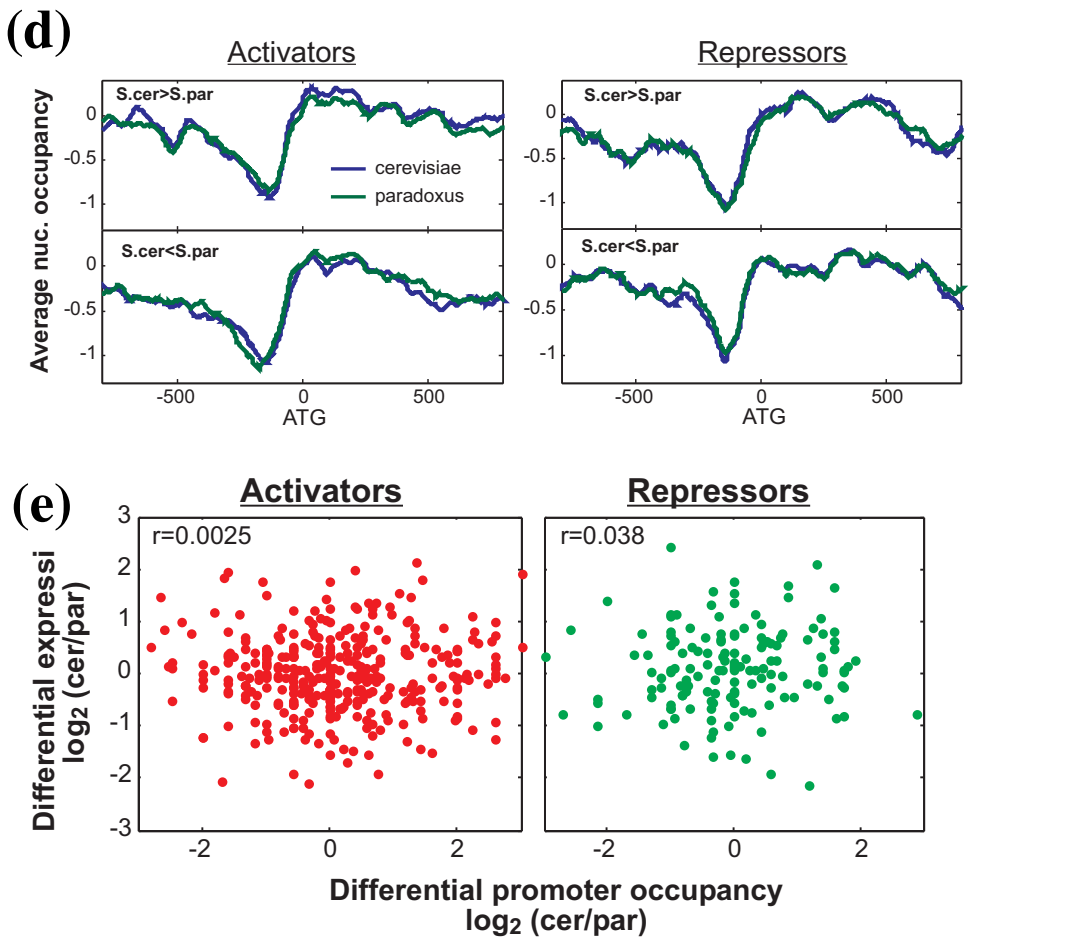
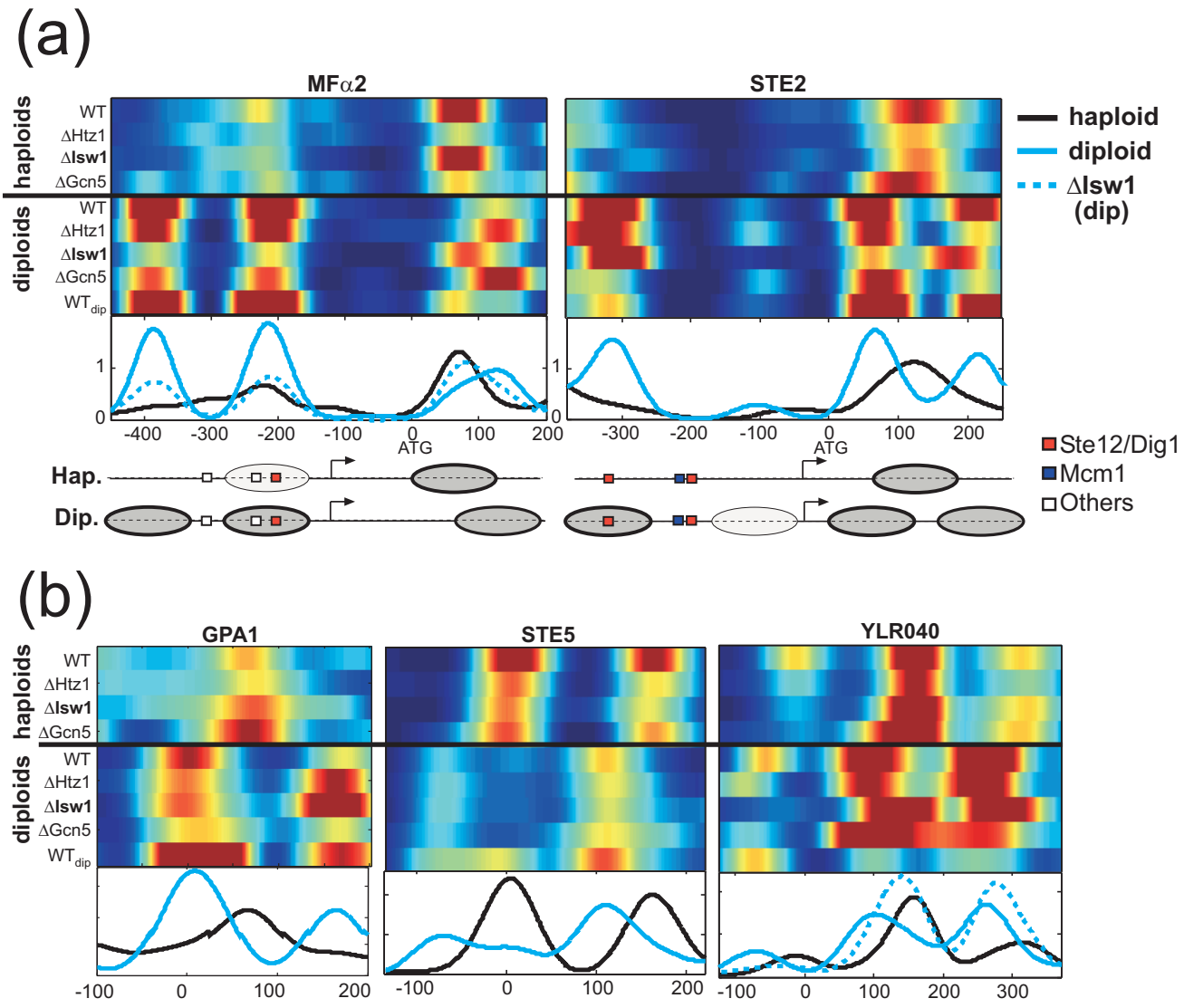


Figure S9. No correlation between divergence of nucleosome positioning and gene expression.
(a) No correlation between differential promoter nucleosome occupancy (log-ratio of the number of reads that map to [-300..-50] in the two species) and differential expression (log-ration of the expression levels in the two species). Similar results were obtained when we used differential occupancy at the coding-regions, at 3'-UTRs or at TFBSs (not shown).
(b) Genes with conserved nucleosome positioning and those with differences in nucleosome positioning (Occ., Loss and Shift) were divided into three bins according to their differential expression among the two species. Genes with differential nucleosome positioning are not significantly enriched with differential expression ($p>0.05$ for each of the three classes compared with the conserved class). Error-bars were calculated by bootstrapping.
(c) No correlation between inter-species changes in nucleosome occupancy at the region of transcription termination and inter-species changes in gene expression. Shown is the average nucleosome occupancy of the two species for all genes with higher *S. cerevisiae* (top) or *S. paradoxus* (bottom) expression levels.
(d-e) No correlation between differential occupancy and differential expression when separately analyzing genes with known *S. cerevisiae* binding sites only for activators (left) and genes with known binding sites for repressors (right). Activator and repressor transcription factors were defined based on Gene Ontology (GO) annotations and *S. cerevisiae* binding sites were taken from MacIsaac et al. (e) For each of the two gene-sets, the average normalized occupancy of the two species is shown for genes with higher (top) and lower (bottom) *S. cerevisiae* expression level, compared to *S. paradoxus*, by at least 1.5-fold. (e) Scatterplot of differential occupancy at promoter positions of TFBSs versus differential expression of the corresponding genes, for activating TFBSs (left) and repressing TFBSs (right). Correlation coefficients are indicated above and both are not statistically significant ($P>0.5$).


(a)
MFα2
STE2
haploids
diploids
WT
ΔHtz1
Δlsw1
ΔGcn5
WT dip
haploid
diploid
Δlsw1 (dip)
-400 -300 -200 -100 0 100 200
-300 -200 -100 0 100 200
ATG
Hap.
Dip.
Ste12/Dig1
Mcm1
Others
(b)
GPA1
STE5
YLR040
-100 0 100 200
-100 0 100 200
-100 0 100 200 300

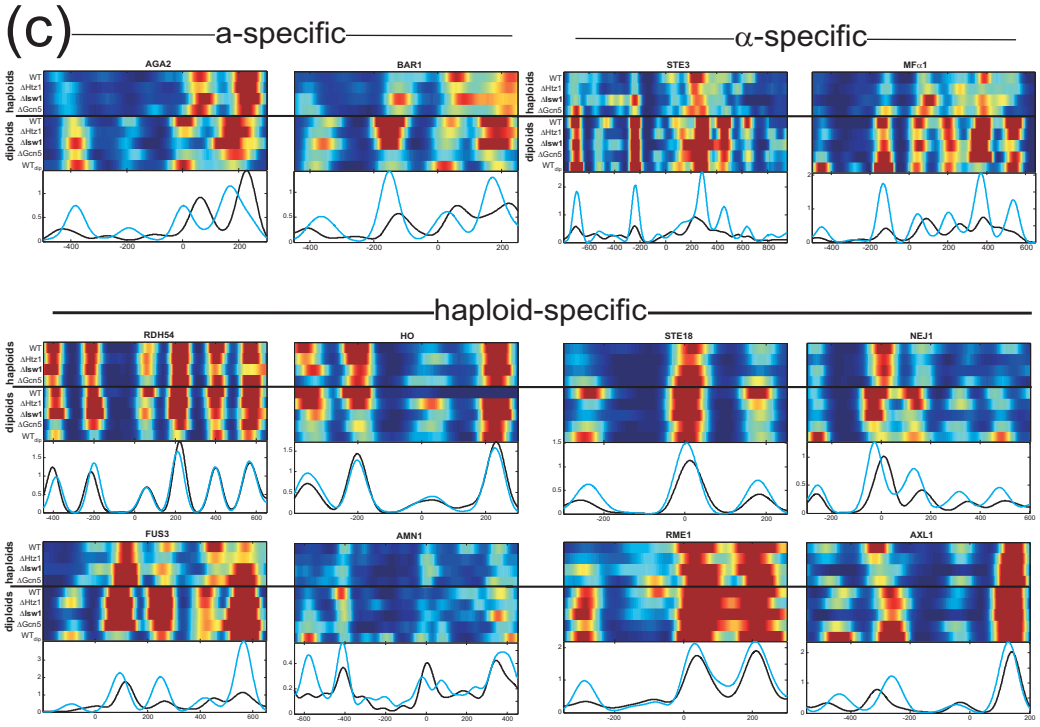


Figure S10. Differential nucleosome positioning between haploids and diploids among haploid-specific genes.

(a) Examples of two genes in which nucleosome occupancy is higher in the diploids and results in lower accessibility of known transcription factor binding sites, which might explain the differences in expression. Shown are patterns of nucleosome positioning in haploids and diploids, for the a-specific gene *STE2* (data shown for *S. cerevisiae* a and a/α strains) and for the α-specific gene *MFα2* (data shown for *S. paradoxus* α and a/α strains). Curves represent the average nucleosome scores of haploids (black), diploids (cyan), and *Δisw1* diploids at MFα2 (dashed cyan). Predicted nucleosome positions, TSS and binding sites (red for Ste12, blue for Mcm1 and white for others) are indicated below.

(b) In these examples, the differences in occupancy are concentrated at the coding region and so their relationship with gene expression is unclear. In contrast to most other genes, *STE5* has higher occupancy at haploids than in diploids.

(c) All haploid-specific genes analyzed here which are not shown in (a-b), divided to a-specific genes, α-specific genes and all other haploid-specific genes.

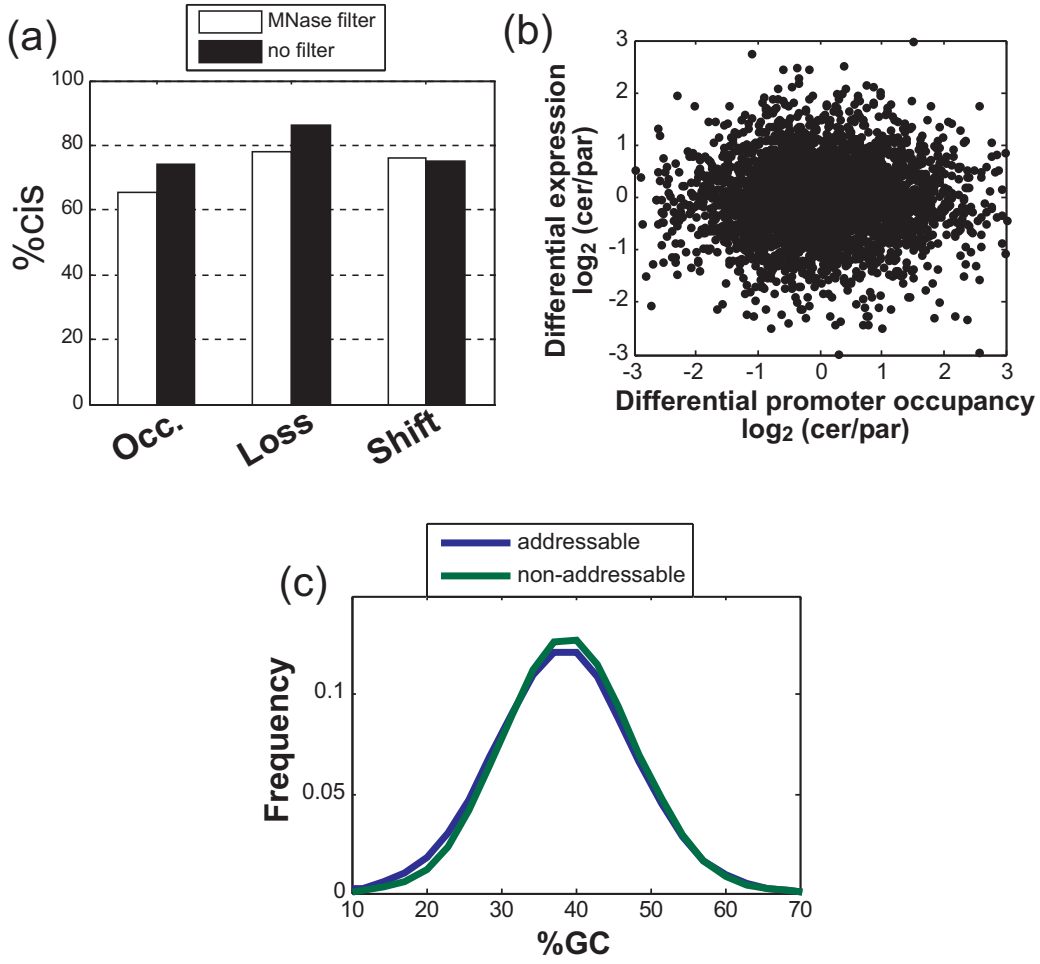


Figure S11. Filtering of nucleosomes does not seem to affect our main conclusions. We relaxed the criteria for excluding genomic regions (see Supplementary Methods section #5) and repeated our analysis. **(a)** The percentage of cis effects remained ~70% in this analysis, with a slight increase when we did not exclude differences that are consistent with MNase bias. **(b)** Differential promoter occupancy remained uncorrelated to differential expression (as in Fig. S9). Similarly, there was no correlation with differential expression for differential occupancy at coding regions, 3'-UTRs and for genes with nucleosome-specific differences (gain/loss or shift) (not shown). **(c)** Genomic regions that cannot be addressed in our analysis due to high sequence conservation (i.e. reads cannot be mapped to a particular genome) or to lack of one-to-one orthology relationships or sequence alignments have a similar GC-content to genomic regions that are addressed.

References

Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF (2007) Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature* **446:** 572-576.

Dingwall C, Lomonossoff GP, Laskey RA (1981) High sequence specificity of micrococcal nuclease. *Nucleic Acids Res* **9:** 2659-2673.

Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, Lubling Y, Widom J, Segal E (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput Biol* **4:** e1000216.

Horz W, Altenburger W (1981) Sequence specific cleavage of DNA by micrococcal nuclease. *Nucleic Acids Res* **9:** 2643-2658.

Ioshikhes IP, Albert I, Zanton SJ, Pugh BF (2006) Nucleosome positions predicted through comparative genomics. *Nat Genet* **38:** 1210-1215.

Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423:** 241-254.

Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39:** 1235-1244.

MacIsaac KD, Wang T, Gordon DB, Gifford DK, Stormo GD, Fraenkel E (2006) An improved map of conserved regulatory sites for Saccharomyces cerevisiae. *BMC Bioinformatics* **7:** 113.

Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, Ito T (2006) A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. *Proc Natl Acad Sci U S A* **103:** 17846-17851.

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320:** 1344-1349.

Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, Wang JP, Widom J (2006) A genomic code for nucleosome positioning. *Nature* **442:** 772-778.

Whitehouse I, Rando OJ, Delrow J, Tsukiyama T (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450:** 1031-1035.