Supplemental Material

RSC-dependent Constructive and Destructive Interference between Opposing Arrays of Phased Nucleosomes in Yeast

Dwaipayan Ganguli, Răzvan V. Chereji, James R. Iben, Hope A. Cole, and David J. Clark

Program in Genomics of Differentiation, Eunice Kennedy Shriver National Institute for Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

Contents

Supplementary Methods

Active Nucleosome Phasing Model

We propose the following active nucleosome phasing mechanism. An ATP-dependent remodeler starts from a potential barrier (nucleosome inaccessible region) located in the NDR and positions the first nucleosome (+1) against this barrier. After the +1 nucleosome is positioned, it acts as a new potential barrier (inaccessible region) for the next nucleosome. The remodeler then positions the +2 nucleosome next to the +1 nucleosome, and so on. We propose the following chromatin organization rule: the position of a nucleosome relative to a neighboring one is determined by a Gaussian distribution

$$g(d) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(d-D)^2}{2\sigma^2}}. \tag{1}$$

Thus, the distribution of the nucleosome repeat length peaks at $d = D$, and has a finite standard deviation σ. If we assume that the first potential barrier is fixed at location B, then the outer edge of the first nucleosome will be positioned at distance d from the barrier edge, according to Equation (1). Therefore, the dyad of the +1 nucleosome will be located at position x_1 relative to the barrier edge, with the probability

$$p_1(B + x_1) = g(x_1 + 73) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[(x_1+73)-D]^2}{2\sigma^2}}.$$

The distribution of the second nucleosome dyad position is given by the convolution of p_1 with g,

$$p_2(B + x_2) = \sum_{x_1=1}^{\infty} p_1(B + x_1) g(x_2 - x_1),$$

the distribution of the third nucleosome dyad position is given by the convolution of p_2 with g, and so forth. Note that the convolution of two Gaussian distributions is another Gaussian, with the mean and the variance equal to the sum of the two means and two variances, respectively. Thus, nucleosome $+n$ will have a dyad position at a distance x_n relative to the original barrier, with the probability distribution

$$p_n(B + x_n) = \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{[(x_n+73)-nD]^2}{2n\sigma^2}},$$

or equivalently

$$p_n(x) = \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{[(x-B+73)-nD]^2}{2n\sigma^2}}. \tag{2}$$

The probability of finding any nucleosome dyad at coordinate x is given by

$$p(x) = \sum_{n=1}^{\infty} p_n(x) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{[(x-B+73)-nD]^2}{2n\sigma^2}}. \tag{3}$$

This dyad distribution oscillates close to the potential barrier, but these oscillations are damped with the distance. Far from the original potential barrier the dyad density has an asymptote equal to $1/D$, because on average there is one new positioned nucleosome every D bp.

One can expect that arrays of nucleosomes are also positioned in the opposite direction, starting from a downstream potential barrier. In the case of bidirectional phasing, near a potential barrier two contributions must be averaged: the oscillatory pattern obtained by positioning the nucleosomes starting from that barrier and the average dyad density $1/D$, which is obtained when the nucleosomes are positioned starting from a distant barrier. In this case, we obtain the nucleosome dyad probability

$$\widetilde{p}(x) = \frac{p(x) + \frac{1}{D}}{2} = \frac{1}{2D} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{\sqrt{2\pi n\sigma^2}} e^{-\frac{[(x-B+73)-nD]^2}{2n\sigma^2}}. \tag{4}$$

Supplementary Figures

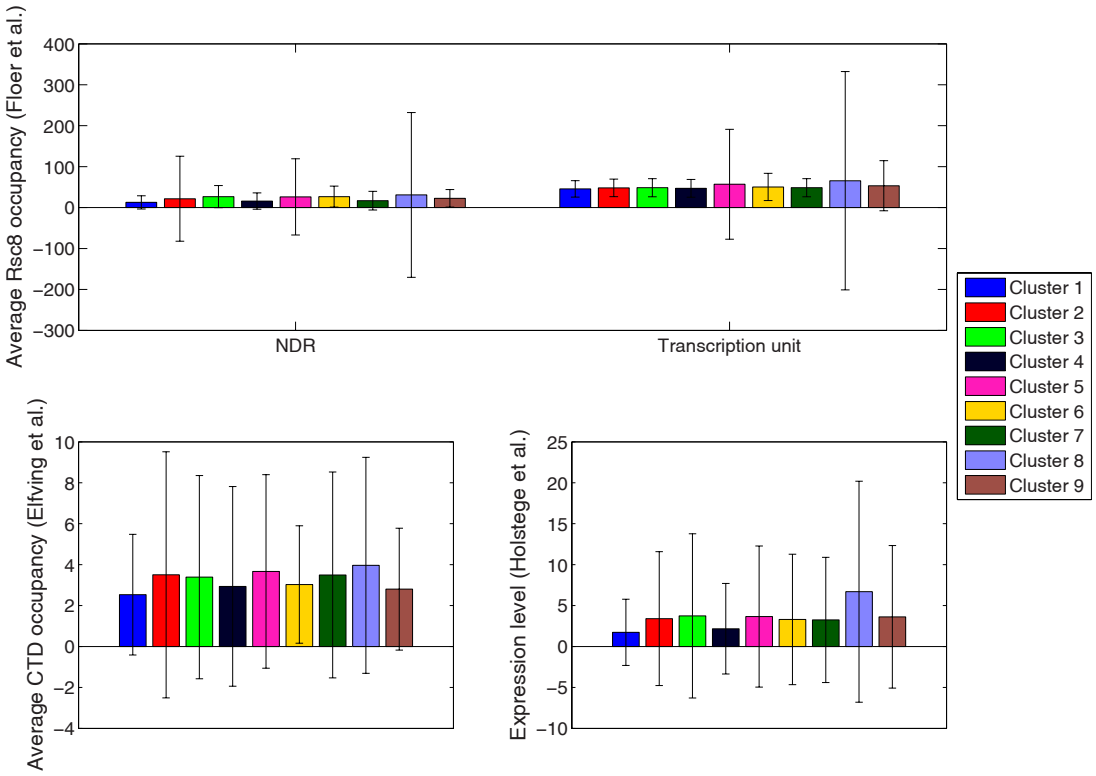

Figure S1. No significant correlation between the +1/-1 nucleosome shift clusters and Rsc8 occupancy, average Pol II occupancy over ORFs, or gene expression. The nine gene-clusters are defined by the directional shifts of the +1 and -1 nucleosomes relative to the TSS (Fig. 4). RSC occupancy data derived from Floer et al. (2010); Pol II CTD ChIP data are from Elfving et al. (2014) and expression (mRNA) data are from Holstege et al. (1998). Average values are shown with error bars indicating the standard deviation.

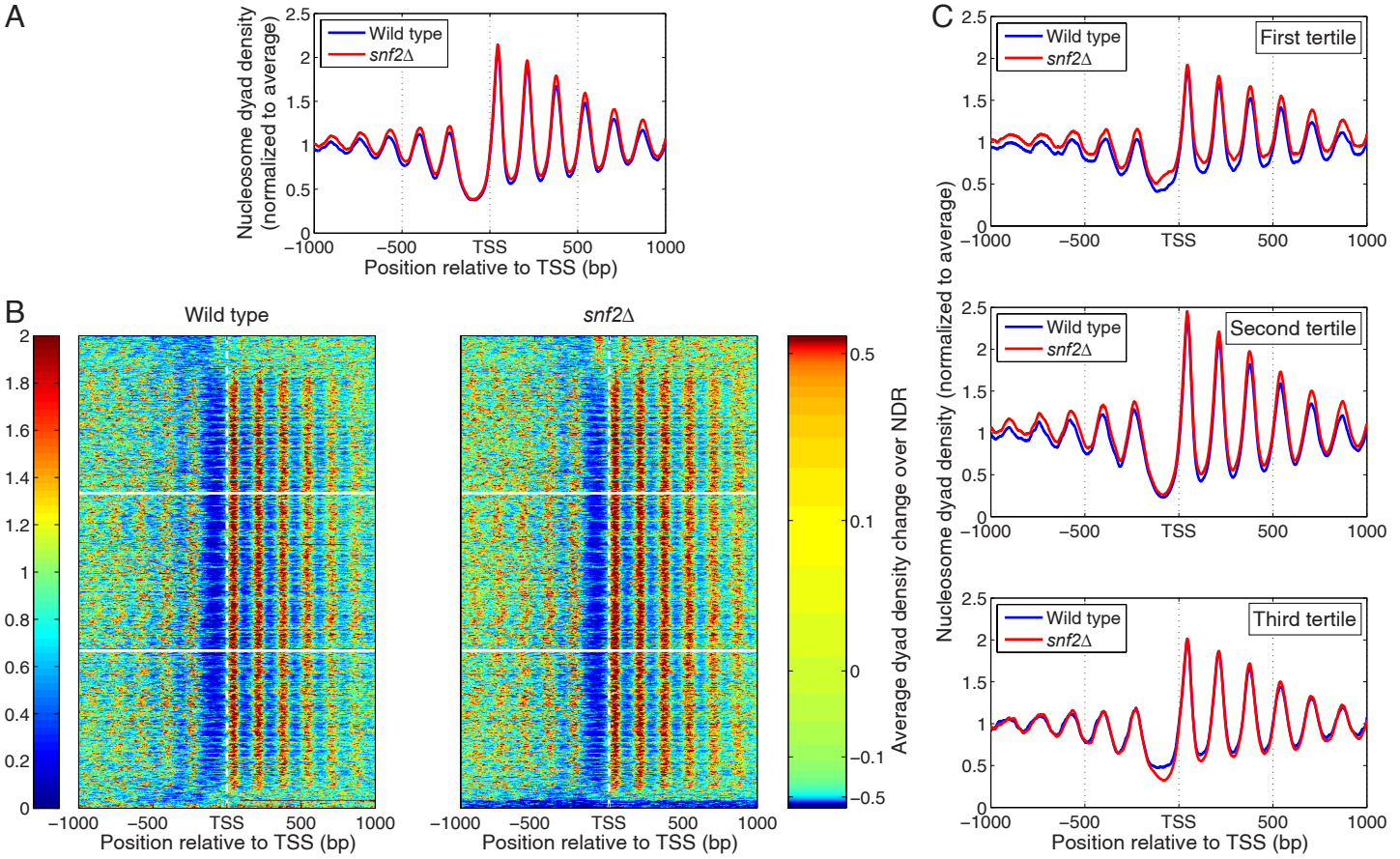

Figure S2. SWI/SNF does not affect nucleosome phasing at a global level. (A) Phasing relative to the TSS for all genes in wild type and *snf2*Δ cells. (B) Heat map analysis of all genes, sorted for change in NDR occupancy. Genes displaying the largest increase in NDR occupancy in *snf2*Δ cells relative to wild type are at the top. Note: Most of the genes at the very bottom of the *snf2*Δ heat map derive from the rDNA repeats; they show a large decrease in NDR occupancy in *snf2*Δ cells, which suggests that the *snf2*Δ strain has fewer repeats than the wild type. (C) Average nucleosome dyad density profiles for the top, central and bottom gene-tertiles delimited by the white horizontal lines in B. Heat maps and plots were smoothed as described in the legend to Fig. 3.

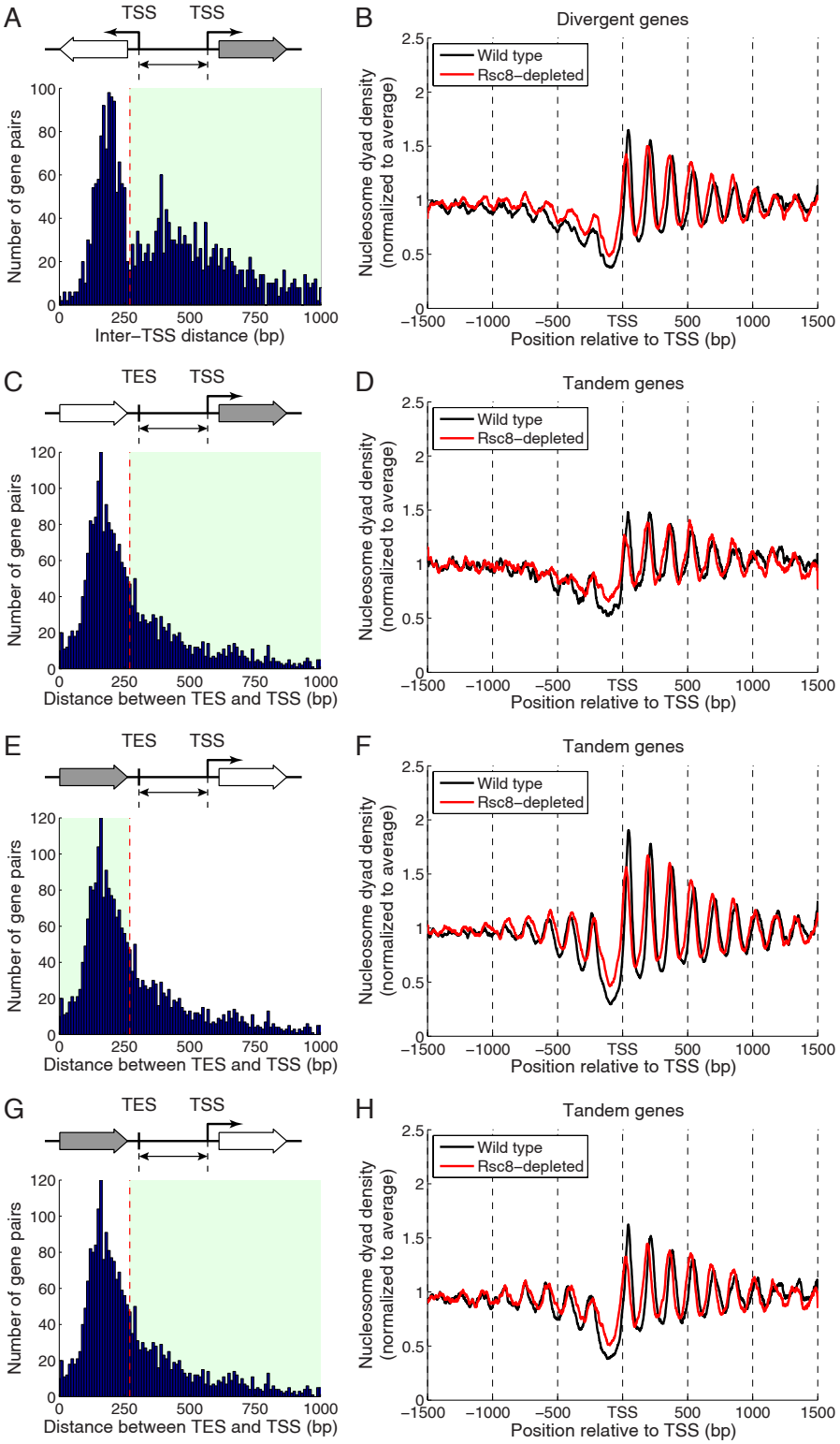

Figure S3. Nucleosome phasing on divergent and tandem genes (additional analysis for Fig. 7). (A) Distribution of inter-TSS distances for divergent genes. Green box indicates genes with inter-TSS distances > 270 bp. (B) TSS phasing plot for divergent genes with inter-TSS distances > 270 bp (compare with Fig. 7A,B). (C) Distribution of distances between the TES and TSS for tandem genes. Green box indicates genes with TES-TSS distances > 270 bp. (D) TSS phasing plot for the downstream gene in tandem gene pairs with TES-TSS distances > 270 bp (compare with Fig. 7C,D). (E,G) Distribution of distances between the TES and TSS for tandem genes. (F,H) TSS phasing plots for the upstream gene in tandem gene pairs with TES-TSS distances < 270 bp or > 270 bp, respectively.

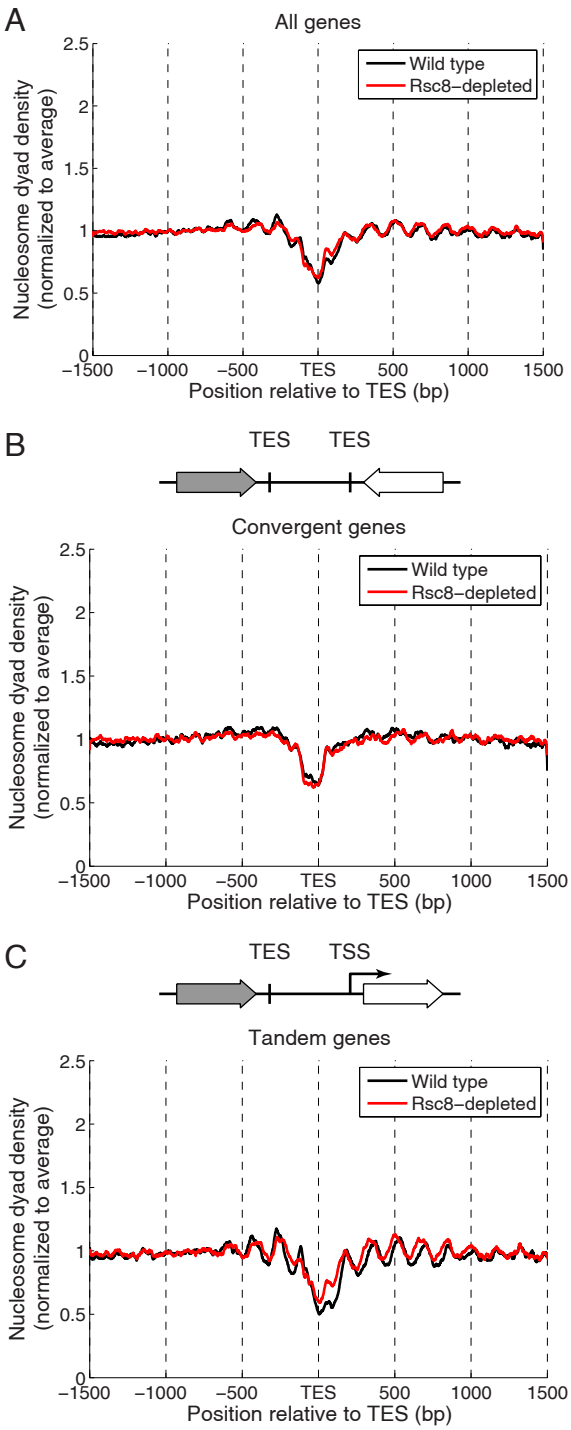

Figure S4. Little or no nucleosome phasing with respect to the transcript end site (TES). (A) Average nucleosome dyad density in a window of 3 kb centered on the TES. There is an NDR at the TES, but there is very little phasing (compare to Fig. 2A). (B) Convergent gene pairs show a weak 3'-NDR but no phasing relative to the TES. (C) The weak oscillations near the TES observed in A are due to tandem gene-pairs; the TSS immediately downstream is presumably responsible for the phasing. All profiles were smoothed using a moving average filter with the span of 21 bp.

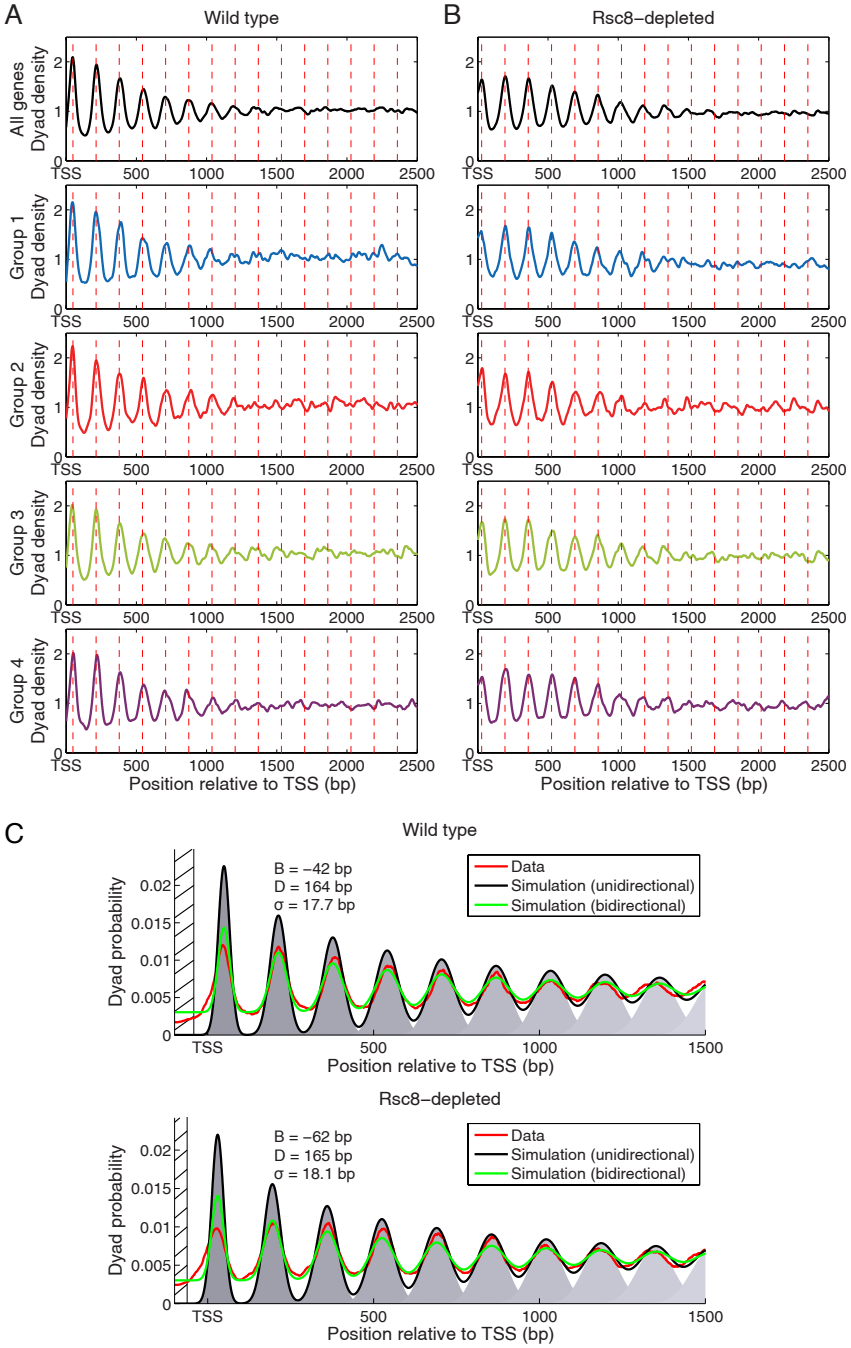

Figure S5. Phasing interference analysis of convergent gene pairs with long inter-TSS distances. (A) Wild type cells. (B) Rsc8-depleted cells. All convergent gene-pairs with long inter-TSS distances (between 3.5 and 4.5 kb; pink boxes in Fig. 8A) were analysed as in Figure 8B. Top panels show the average for all of these convergent gene-pairs. The lower four panels correspond to four phase groups 1 - 4, representing gene-pairs with phase differences of 0 (in phase), 40, 80 (almost exactly out of phase) and 120 bp, respectively. (C) Model for active nucleosome positioning and decay in phasing power. Average nucleosome dyad density profiles were calculated for genes ¿ 2 kb and aligned by their TSS. The data were converted into a probability of finding a dyad at each position relative to the TSS by normalizing the average dyad density such that the maximum nucleosome occupancy, obtained by symmetrically extending each dyad read to 147 bp, equals one (red line). We fit the three parameters of the active nucleosome positioning model (see Methods): average value, D, and standard deviation, , of the distance of the +1 nucleosome from the barrier in the NDR or for the +2 nucleosome from the +1 nucleosome, etc. The third parameter, B, is the position of the initial barrier with respect to the TSS. The predicted dyad probabilities of successive nucleosomes are shown by different shades of grey. The predicted probabilities of finding any dyad at a given position are shown with black and green curves, corresponding to the unidirectional and bidirectional phasing models, respectively. The Pearson's linear correlation coefficients between the real distributions and the predicted ones are 0.95 for wild type and 0.92 for Rsc8-depleted cells.

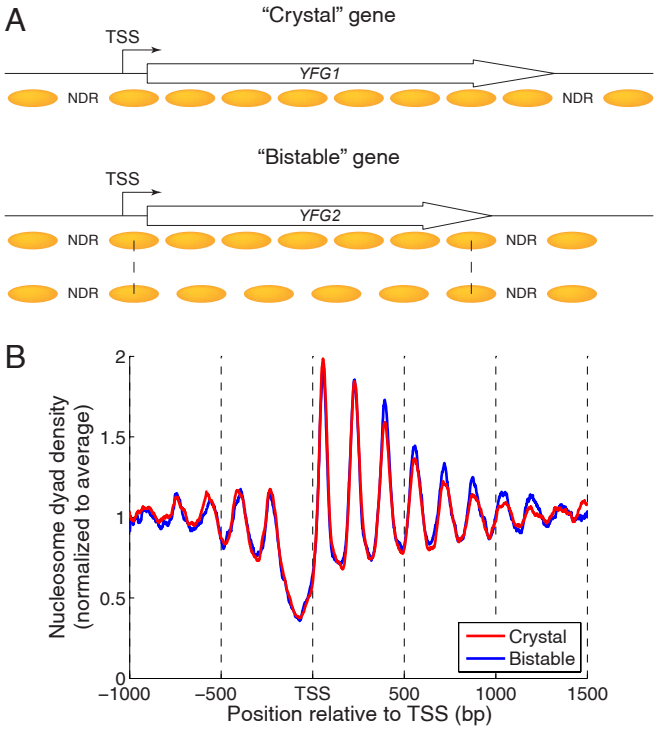

Figure S6. No evidence for nucleosomal arrays with different spacing (bistable genes). (A) Altered spacing model of Vaillant et al. (2010). A "crystal" gene has a single nucleosomal array with ∼165 bp spacing. A "bistable" gene has two alternative nucleosomal arrays with n or $n+1$ nucleosomes, with altered spacing such that the positions of the first and last nucleosomes are the same in both arrays. They identified 1940 crystal genes, 946 bistable genes and 1668 genes with more complex chromatin structures. (B) Average nucleosome dyad density for crystal genes (red line) and bistable genes (blue line) in wild type cells. The two phasing profiles are almost identical, with typical spacing of ∼165 bp.

Supplementary References

Elfving N, Chereji RV, Bharatula V, Björklund S, Morozov AV, Broach JR. 2014. A dynamic interplay of nucleosome and Msn2 binding regulates kinetics of gene activation and repression following stress. *Nucl. Acids Res.* **42**: 5468–5482.

Floer M, Wang X, Prabhu V, Berrozpe G, Narayan S, Spagna D, Alvarez D, Kendall J, Krasnitz A, Stepansky A, Hicks J, Bryant GO, Ptashne M. 2010. A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell* **141**: 407–418.

Holstege, FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717–728.

Vaillant C, Palmeira L, Chevereau G, Audit B, d'Aubenton-Carafa Y, Thermes C, Arneodo A. 2010. A novel strategy of transcription regulation by intragenic nucleosome ordering. *Genome Res.* **20**: 59–67.