Supplemental Figure Legends

Supplemental Table 1. Abundance of polyG elements of different length and strength varies significantly between different species in a philogenetically coherent way.

(A) Number of polyG elements of different lengths with 0 mismatches in all species. **(B)** Number of polyG elements of different in vitro strength for all species (Methods).

Supplemental Table 2. Presence of PolyA and PolyG sequences in NFRs.

(A) Percentage of PolyG and PolyA sequences of different length that reside in the 5' NFRs in all species.

Supplemental Table 3. New GRF motifs predicted by our motif finding algorithm across the 13 yeast species. Shown are the different GRF motifs (rows) predicted by our method. The first column lists the Position Specific Scoring Matrices (PSSMs), the second column shows the closest *S. cerevisiae* PSSM match found, a statement if the best matching PSSM had a prominent mismatch, and a question mark if no reasonable match was found. The third column displays a check mark for every species that had the same or very similar PSSM prediction. Most predicted GRF binding sites are not detected in *S. cerevisiae*. For example, the PSSM for *S. cerevisiae* transcription factor Pbf1 is associated with nucleosome depletion in *C. glabrata* in a species-specific manner, while cell cycle factors Swi4 and Mbp1 may act as GRFs in *S. castelli* and other species.

Supplemental Figure 1. Cbf1 exhibits weak anti-nucleosomal activity in *S. cerevisiae*. **(A)** *CBF1* deletion in *S. cerevisiae* results in only a mild increase in nucleosome occupancy at Cbf1 motifs. Left: genes with significant matches to the Cbf1 binding site (purple). Right: difference between *cbf1*Δ and wildtype strains in nucleosome abundance at each gene in *S. cerevisiae* (rows). Genes are aligned by the +1 nucleosome/NFR boundary (red arrow) and ranked from gain (top, yellow) to loss (bottom, blue) in nucleosome occupancy over their NFR. (**B, C**) Distributions of the difference in nucleosome occupancy at all intergenic Cbf1 CACGTGA sites between *cbf1*Δ and wildtype strains in (**B**) *S. cerevisiae*, and in (**C**) *C. albicans*. **(D)** Only slight increase in nucleosome occupancy in CACGTGA Cbf1 sites in *S. cerevisiae* in the *cbf1*Δ strain. Shown is mean log2 nucleosome occupancy (Y axis) at all *S. cerevisiae* genes with a CACGTGA Cbf1 motif match in their promoter in wild type (blue) and *cbf1*Δ (red) strains. Genes are aligned by the location of the CACGTGA Cbf1 motif (located at position 0 on the x-axis).

Supplemental Figure 2. GRF deletion effects on polyA or polyG occupancy do not result from motif co-occurrence. (**A**) Shown is a scatter plot of the enrichment (KS test, log10 of KS p value) of the Cbf1 motif (Y axis) or poly G elements (X axis) in the promoters of genes from different gene sets in *C. albicans*. Notably, genesets enriched for Cbf1 motifs, such as genes encoding ribosomal protein subunits (red circle) are depleted of polyG elements. Similar results hold for gene by gene analysis (not shown). (**B**) Shown are for every gene set the change in nucleosome occupancy in *cbf1*Δ (x axis,

log10 of the KS statistic, negative values indicate increased nucleosome occupancy in the *cbf1*Δ strain) versus the enrichment (KS test, log10 of KS p value) of the Cbf1 motif (blue) or poly G elements (black) in their genes' promoters. Genesets exhibiting significant increases in nucleosome occupancy in the *cbf1*Δ are enriched for Cbf1 binding sites, and depleted of polyG elements. (**C**) As in **A**, but for enrichment of polyA elements (x axis) vs. enrichment of Reb1 sites (y axis). Note that polyAs largely do not overlap with Reb1 at functional genesets. (**D**) Nucleosome occupancy changes over polyA elements between wild-type and *reb1* mutants (Y axis) is plotted for all polyA elements (red), or only for polyA elements without a nearby Reb1 binding motif (blue, see Methods). Clearly, nucleosome loss over PolyA elements in *reb1* mutants does not result from sliding of nucleosomes onto nearby Reb1 binding sites. Similar results hold for analyses of Abf1, Rsc3, Cbf1 and Sap1 (not shown).

Supplemental Figure 3. MNase digestion level and temperature do not affect the increased nucleosome occupancy at Sap1 sites in *S. pombe Sap1*^{ts} strain grown in a restrictive temperature. (**A**) We measured nucleosome occupancy genome-wide in *sap1*^{ts} strain using different amounts of MNase to rule out the possibility that the amount of MNase affects our conclusions. In a nomally digested sample, the ratio of mono to dinucleosomes on a gel is about 4:1, whereas an overdigested sample will have mono to dinucleosome ratio greater than 5:1 (Weiner et al. 2010). Shown is the mean nucleosome occupancy at all genes with a significant Sap1 motif match in their upstream promoter for an overdigested *sap1*^{ts} strain grown in restrictive temperature (35˚C) (pink), a normally digested *sap1*^{ts} strain grown in restrictive temperature (35˚C) (red), a wildtype strain

grown at restrictive temperature (35˚C) (blue), and a wild type strain grown at 30˚C (cyan). Genes are aligned by the location of the Sap1 motif. (**B**) Increased nucleosome occupancy over 5-mers reflecting the Sap1 half-sites in *sap1*^{ts} strain (at restrictive temperature, 35˚C) compared to wild-type (at 30˚C). Shown is the mean nucleosome occupancy (log2) for each 5-mer in the wild type (X axis) and the *sap1*^{ts} strain (Y axis). Sap1 half-sites are labeled. The only additional site with increased occupancy is the intrinsic sequence GGGGG. (**C**) Increased nucleosome occupancy in 7-mers reflecting the Sap1 half-sites and polyG in *sap1*^{ts} strain compared to wild-type (both at restrictive temperature, 35˚C). Shown is the mean nucleosome occupancy (log2) for each 7-mer in the wild type (X axis) and the *sap1*^{ts} strain (Y axis). Triangles: Sap1 sites; diamonds: polyG elements. As with 5-mers, 7-mers analysis shows that polyG and Sap1 binding variants are the only 7-mers that have an increased nucleosome occupancy in *sap1*^{ts} strain compared to wild-type (occupancy difference >.5).

Supplemental Figure 4. Intrinsic and trans-regulated nucleosome positioning sequences in *C. elegans*. (**A**) PolyG is an intrinsically anti-nucleosomal sequence in *C. elegans*. Shown are in vivo mean nucleosome occupancy levels (Y axis) for polyG sequences of different lengths (X axis) in *C. elegans* (Valouev et al. 2008). (**B**) A putative GRF binding site in *C. elegans* inferred by our method. We used *in vitro* estimates of each 7mers nucleosome disfavoring potential from *S. cerevisiae* and *C. albicans*. The likely motif in *C. elegans* includes the 8-mer CGGCAAAT, which is extremely abundant in non-coding DNA (Subirana and Messeguer 2010). This 8-mer frequently forms clusters of dimers and trimers that may serve as punctuation marks along the genome,

depleting DNA of nucleosomes and creating a favorable site for homologous DNA recognition (Subirana and Messeguer 2010).