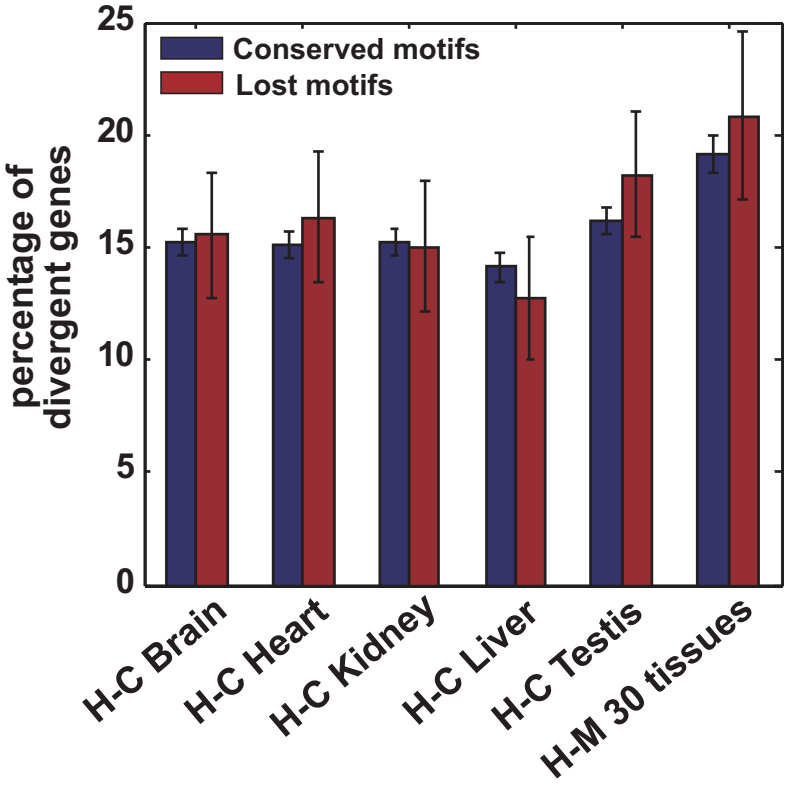

Figure S1. Expression divergence of mammalian genes with predicted loss of TF binding sites. (a) We predicted the loss of transcription factor binding sites by searching for exact matches to consensus motifs that are conserved in the aligned promoters of several species, but mutated in the promoter of another species (top). To identify the most reliable cases, we demanded that the lost motif have at least two disrupting mutations (that do not match the consensus motif) and no other binding site for the same TF is present in that promoter. The percentage of genes with divergent expression (expression divergence>mean+std) is shown among the genes with conserved (blue) and lost (red) TF binding sites**.** Errorbars were calculated by bootstrapping. The x-axis in corresponds to five different measures of divergence between human and chimpanzee tissues (H-C; taken from Khaitovich et al. (Khaitovich et al. 2005)), and another measure of divergence between 30 tissues of human and mouse (H-M; taken from Su et al. (Su et al. 2004)).

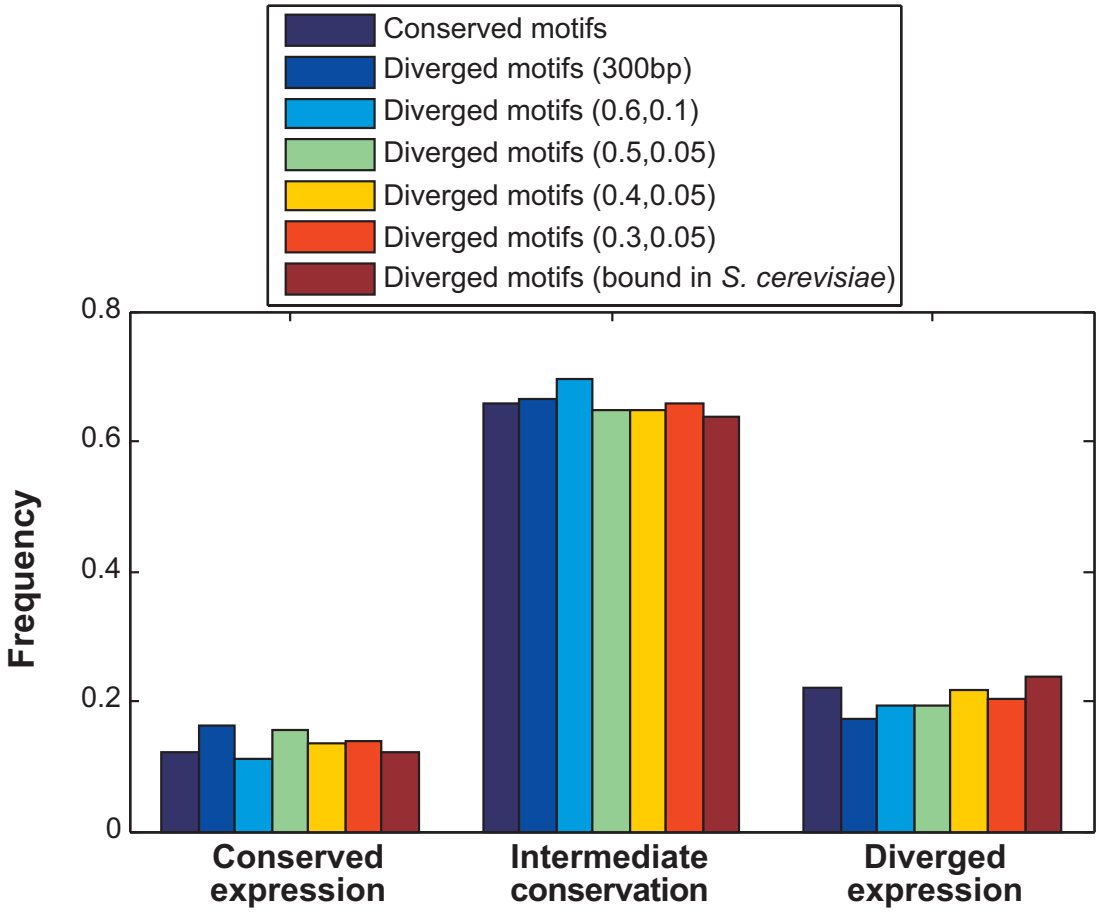

Figure S2. Diverged motifs have no detectable impact on gene expression using various parameters. We compared the fraction of genes with diverged expression among those with conserved and diverged motifs and found that there is no significant difference between the two gene sets (p>0.05). To verify the robustness of this result we used different parameters to define conserved and diverged motifs and found similar results. The results are shown for conserved motifs using our initial threshold (specificity of 0.5), and for diverged motifs using several sets of parameters (other sets were also tested and led to similar results). The initial parameters were specificity of 0.5 and 0.1 for conservation and divergence, respectively, promoter length of 600bp, and no binding criteria. Each dataset is indicated by the parameters that were altered relative to the initial analysis: the second dataset was based on shorter promoters (300bp) since most known binding sites are concentrated at these regions; the third through sixth datasets were based on different specificity thresholds, as indicated; the seventh dataset included only promoters which were bound by the respective TF in *S. cerevisiae* (Harbison et al. 2004) and diverged in other species. Notably, both the gene sets with conserved and diverged motifs had higher percentage of genes with diverged expression (~20%) than that of genes with conserved expression(~15%), suggesting that binding by TFs increases the likelihood of expression divergence whether it is conserved or not.

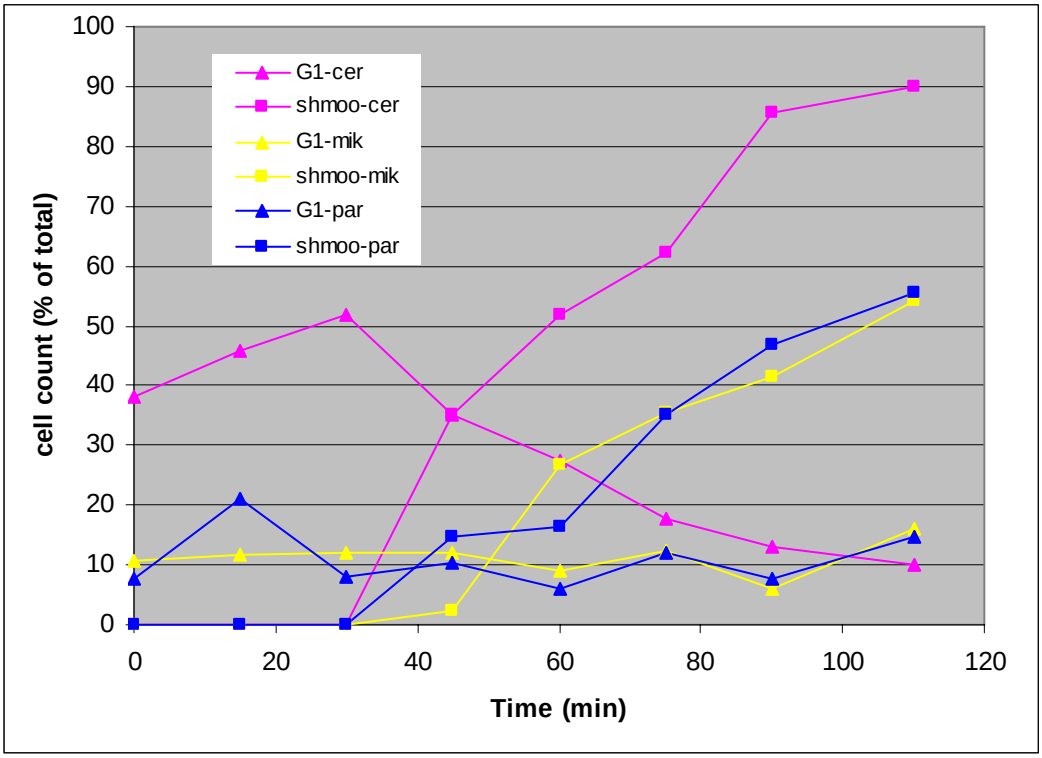

Figure S3. Synthetic α-factor induces mating in three yeast species. A-type haploid cells from *S. cerevisiae*, *S. paradoxus* and *S. mikatae* were each subjected to α-factor (5μM). The percentage of cells with a visible bud (i.e. at G1) or shmoo is shown at different time points.

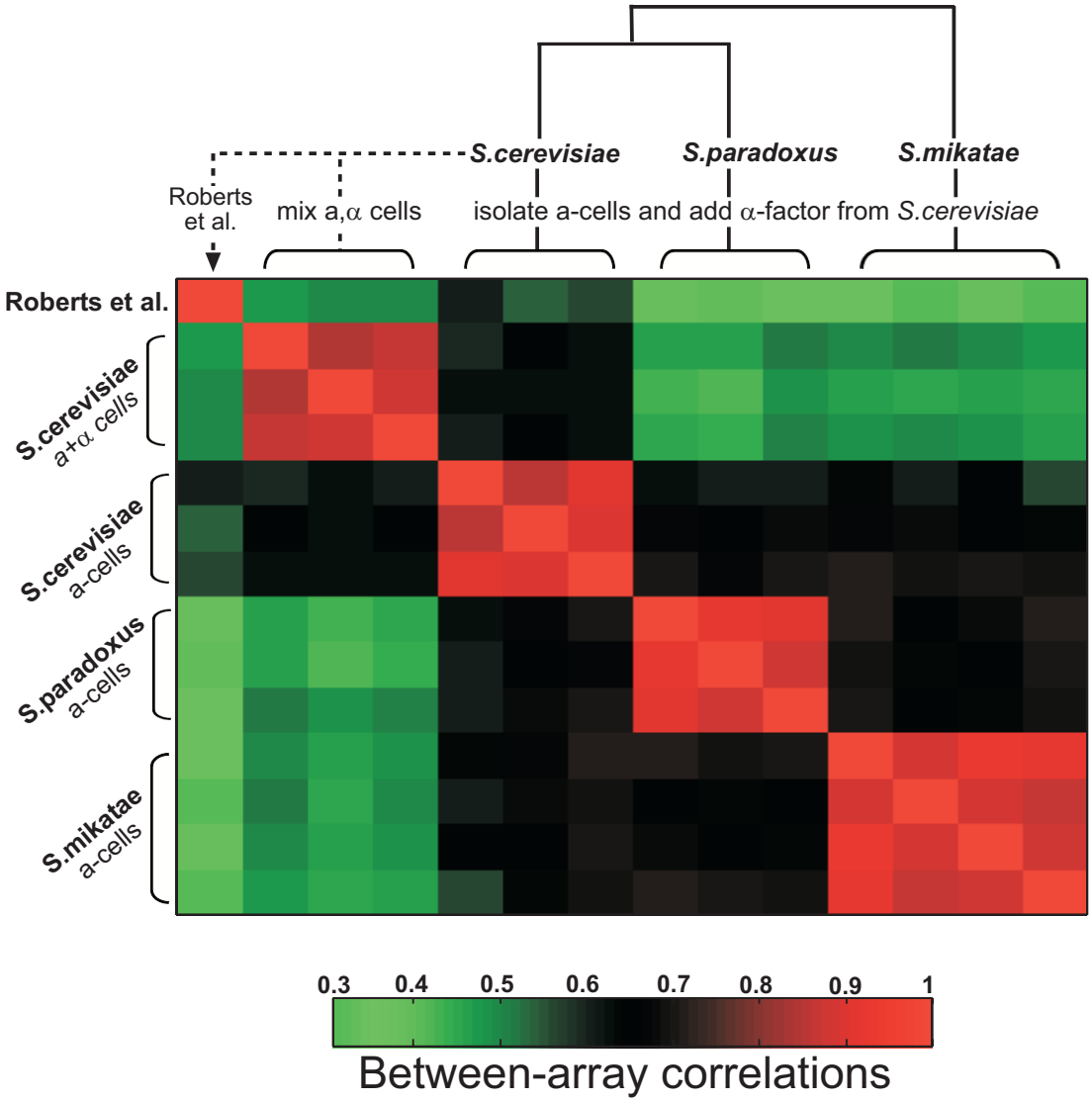

Figure S4. Correlations between the mating expression program in different experiments and species. In addition to the α-factor experiments, we measured the combined response of a and α cells from *S.cerevisiae* after mixing them, and downloaded the mating response as previously measured by Roberts et al. (Roberts et al. 2000). The correlations among all these genomic responses were calculated over 3248 genes with a significant response in at least one experiment. The relatively low correlation between the α-factor and mating experiments (r~0.63) is probably due to the fact that only a-cells are examined in the α-factor experiment while both cell types are measured (and averaged) in the mating experiment. Indeed, several α-specific genes were only upregulated in the mating experiment (e.g. STE3, MATA1).

```
SGD_Scer_STE12/YHR084W      MKVQITNSRTEEILKVQANNENDEVSKATPGEVEESLRLIGDLKFFLATA
MIT_Spar_c35_10073          MKVQITNSRTEEILKVQTNNESDEVSKATPSEVEESLRLIDDLKFFLATA
MIT_Smik_c313_9948          MKVQITNSRTEEILKIQTGNESDEASKATPSEVEESLRLIDDLKFFLATA
                            ***************:*:.**.**.*****.*********.*********

SGD_Scer_STE12/YHR084W      PVNWQENQIIRRYYLNSGQGFVSCVFWNNLYYITGTDIVKCCLYRMQKFG
MIT_Spar_c35_10073          PVNWQENQIIRRYYLNSGQGFVSCVFWNNLYYITGTDIVKCCLYRMQKFG
MIT_Smik_c313_9948          PVNWQENQIIRRYYLNSGQGFVSCVFWNNLYYITGTDIVKCCLYRMQKFG
                            **************************************************

SGD_Scer_STE12/YHR084W      REVVQKKKFEEGIFSDLRNLKCGIDATLEQPKSEFLSFLFRNMCLKTQKK
MIT_Spar_c35_10073          REVVQKKKFEEGIFSDLRNLKCGIDATLEQPKSEFLSFLFRNMCLKTQKK
MIT_Smik_c313_9948          REVVQKKKFEEGIFSDLRNLKCGIDATLEQPKSEFLSFLFRNMCLKTQKK
                            **************************************************

SGD_Scer_STE12/YHR084W      QKVFFWFSVAHDKLFADALERDLKRESLNQPSTTKPVNEPALSFSYDSSS
MIT_Spar_c35_10073          QKVFFWFSVAHDKLFADALERDLKRESLNQPSTTRSVNEPALSFSYDSSS
MIT_Smik_c313_9948          QKVFFWFSVAHDKLFADALERDLKRESLNQPSTTKPINEPALSFSYDSSS
                            **********************************:.:**************

SGD_Scer_STE12/YHR084W      DKPLYDQLLQHLDSRRPSSTTKSDNSPPKLESENFKDNELVTVTNQPLLG
MIT_Spar_c35_10073          DKPLYDQLLQHLDSRRPSSATKSDNSPSKLESENFKDTELVTVTNQPLLG
MIT_Smik_c313_9948          DKPLYDQLLQHLESRRPSSAIKSDDSPSKLESENFKDAELVTVTNQPLLG
                            ************:******: ***:**.********* ************

SGD_Scer_STE12/YHR084W      VGLMDDDAPESPSQINDFIPQKLIIEPNTLELNGLTEETPHDLPKNTAKG
MIT_Spar_c35_10073          VGLIDDDTPESPSQINDFIPQKLIVEPNTLELNSLAEETSHALPKNAAKG
MIT_Smik_c313_9948          VDLMDR-TPESPSQVDDFAPQKLIIEPNTLELNSLADETSHVLPKNAAKS
                            *.*:*  :******::** *****:********.*::**.* ****:**.

SGD_Scer_STE12/YHR084W      RDEEDFPLDYFPVSVEYPTEENAFDPFPPQAFTPAAPSMPISYDNVNERD
MIT_Spar_c35_10073          RDEEDFPLDYFPVSVEYPTEENAFDPFPQQAFTPAAPSMPISYDNVSERD
MIT_Smik_c313_9948          RDEEDFPLDYFPVSVEYPTEENAFDPFPQQAFTPAAPSIPISYDNVSERE
                            **************************** *********:*******.**:

SGD_Scer_STE12/YHR084W      SMPVNSLLNRYPYQLSVAPTFPVPPSSSRQHFMTNRDFYSSNNNKEKLVS
MIT_Spar_c35_10073          SMPVNSLLNRYPYQLSVAPTFPVPPSSSRQHFMTNRDFYSSNNNKEKLAS
MIT_Smik_c313_9948          SMPVNSLLNRYPYQLSVAPTFPVPPSSSRQHFMTNRDFYSSNNNKEKLVS
                            **********************************************.*

SGD_Scer_STE12/YHR084W      PSDPTSYMKYDEPVMDFDESRPNENCTNAKSHNSGQQTKQHQLYSNNFQQ
MIT_Spar_c35_10073          PSDPTSYMKYDEPVIDFEESRVNENCTNAKSHNSGQQTKQHQLYPNNFQQ
MIT_Smik_c313_9948          PSDPTSYIKYDEPVIDLEESRLNENGTNSKSYNSGQQTKQHQLYSNNFQQ
                            *******:******:*::*** *** **:**:************.*****

SGD_Scer_STE12/YHR084W      SYPNGMVPGYYPKMPYNPMGGDPLLDQAFYGADDFFFPPEGCDNNMLYPQ
MIT_Spar_c35_10073          PYPNGMVPGYYPKMPYNPMVGDPLLDQAFYGADDLFFPPEGCDNNMLYPQ
MIT_Smik_c313_9948          SYPNGMVPGYYPKMPFNPMVGDSLLDQAFYGADDLFFPPEGCDSSVIYPQ
                            .**************:*** **.***********:********..::***

SGD_Scer_STE12/YHR084W      TATSWNVLPPQAMQPAPTYVGRPYTPNYRSTPGSAMFPYMQSSNSMQWNT
MIT_Spar_c35_10073          TATSWNLLPPQAMQPAPTYVPRPYTPNYRSTPGSAMFPYMQSSNSMQWNT
MIT_Smik_c313_9948          TATSWNVLPPQAMQPAPTYVPRPYTPNYRSTPGSAMFPYMQNSNSVQWNT
                            ******:************* ********************.***:****

SGD_Scer_STE12/YHR084W      AVSPYSSRAPSTTAKNYPPSTFYSQNINQYPRRRTVGMKSSQGNVPTGNK
MIT_Spar_c35_10073          AVSPYSSRAPSTTAKNYPPSTFYSQNMNQYPRRRTVGMKSSQGNVPTGNK
MIT_Smik_c313_9948          AVSPYRSRAPSTTAKNYPPSTFYPQNMNQYSRRRTVGMKASQANVPTNNK
                            ***** *****************.**:***.********:**.****.**

SGD_Scer_STE12/YHR084W      QSVGKSAKISKPLHIKTSAYQKQYKINLETKARPSAGDEDSAHPDKNKEI
MIT_Spar_c35_10073          QSVGKSAKISKPLHIKTSAYQKQYKINLETKTRSSAGDEDSAHPDKSKEI
MIT_Smik_c313_9948          LSVGKSAKVSKPLHIKTSAYQKQYKINLEAKTKPITGDEDSVHSDKSREV
                             *******:**********************:*::. :******.*.**.:*:

SGD_Scer_STE12/YHR084W      SMPTPDSNTLVVQSEEGGAHSLEVDTNRRSDKNLPDAT
MIT_Spar_c35_10073          SMPTPDSNTLVIQSEEDGVDSPEVETNGRSNKNLTDAT
MIT_Smik_c313_9948          SMPTPDSNTLVIQSEEDGVNSPEVQTSGVSKKKLPDTT
                            ***********:****.*..* **:*.  *.*:*.*:*
```

Figure S5. Multiple alignment of STE12 in S. cerevisiae, S.paradoxus and S. mikatae. The binding domain is between residues 41 and 204 (Yuan and Fields 1991).

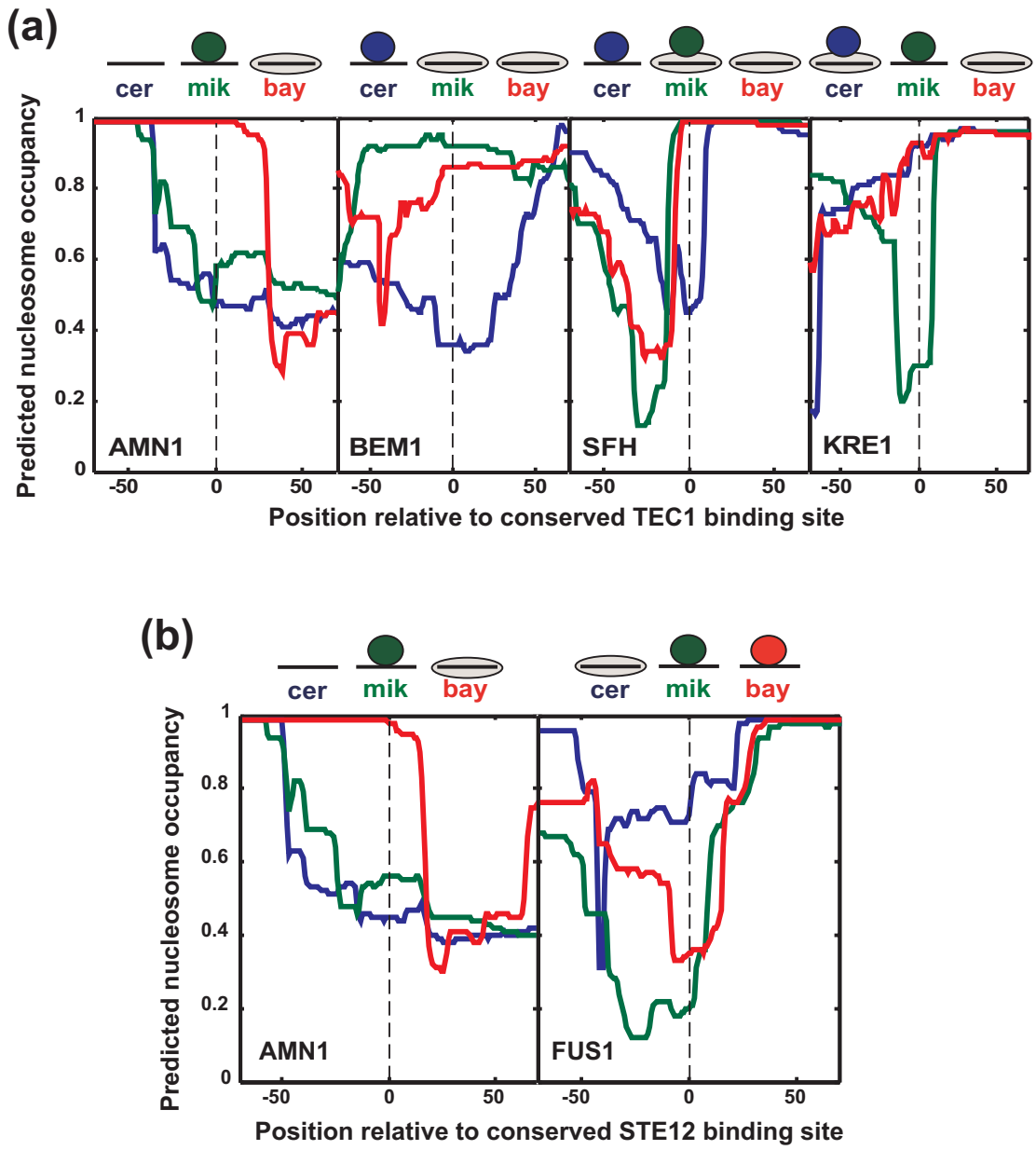

Figure S6. Predicted nucleosome occupancy among genes with differential binding of TEC1 and STE12 in pseudohyphal conditions. We examined the predicted nucleosome occupancy (Segal et al. 2006) at genes with conserved motifs but divergent binding of TEC1 (**a**) and STE12 (**b**) in pseudohyphal conditions (Borneman et al. 2007). Genes with interspecies differences in nucleosome occupancy of at least 0.5 were selected and analyzed. For each of these genes, the predicted nucleosome occupancy in *S. cerevisiae*, *S. mikatae* and *S. bayanus* is shown for the region surrounding the conserved binding site. The correspondence between changes in TF binding and changes in nucleosome positions is indicated above each plot; binding of the respective TFs in the three species (Borneman et al. 2007) is indicated by circles, and high nucleosome occupancy is indicated by gray ovals. Only 2 of the 9 promoters which are bound have high nucleosome occupancy, while 7 of the 9 promoters which are not bound have high nucleosome occupancy (p=0.03, by shuffling the binding of TFs).

Borneman, A.R., T.A. Gianoulis, Z.D. Zhang, H. Yu, J. Rozowsky, M.R. Seringhaus, L.Y. Wang, M. Gerstein, and M. Snyder. 2007. Divergence of transcription factor binding sites across related yeast species. *Science* **317:** 815-819.

Harbison, C.T., D.B. Gordon, T.I. Lee, N.J. Rinaldi, K.D. Macisaac, T.W. Danford, N.M. Hannett, J.B. Tagne, D.B. Reynolds, J. Yoo, E.G. Jennings, J. Zeitlinger, D.K. Pokholok, M. Kellis, P.A. Rolfe, K.T. Takusagawa, E.S. Lander, D.K. Gifford, E. Fraenkel, and R.A. Young. 2004. Transcriptional regulatory code of a eukaryotic genome. *Nature* **431:** 99-104.

Khaitovich, P., I. Hellmann, W. Enard, K. Nowick, M. Leinweber, H. Franz, G. Weiss, M. Lachmann, and S. Paabo. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309:** 1850-1854.

Roberts, C.J., B. Nelson, M.J. Marton, R. Stoughton, M.R. Meyer, H.A. Bennett, Y.D. He, H. Dai, W.L. Walker, T.R. Hughes, M. Tyers, C. Boone, and S.H. Friend. 2000. Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* **287:** 873-880.

Segal, E., Y. Fondufe-Mittendorf, L. Chen, A. Thastrom, Y. Field, I.K. Moore, J.P. Wang, and J. Widom. 2006. A genomic code for nucleosome positioning. *Nature* **442:** 772-778.

Su, A.I., T. Wiltshire, S. Batalov, H. Lapp, K.A. Ching, D. Block, J. Zhang, R. Soden, M. Hayakawa, G. Kreiman, M.P. Cooke, J.R. Walker, and J.B. Hogenesch. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* **101:** 6062-6067.

Yuan, Y.L. and S. Fields. 1991. Properties of the DNA-binding domain of the Saccharomyces cerevisiae STE12 protein. *Mol Cell Biol* **11:** 5910-5918.